Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Fisica

Diffusive approach for non-linear beam dynamics in a circular accelerator

Relatore:
Prof. Armando Bazzani

Correlatore:
Dott. Massimo Giovannozzi

Presentata da:
Carlo Emilio Montanari

Anno Accademico 2018/2019

La distinzione tra arte, filosofia, scienza non la conoscevano Empedocle, Dante, Leonardo, Galileo, Cartesio, Goethe, Einstein, né gli anonimi costruttori delle cattedrali gotiche, né Michelangelo; né la conoscono i buoni artigiani di oggi, né i fisici esitanti sull'orlo del conoscibile

---

Primo Levi

Sommario

Nella fase di design e simulazione di un acceleratore circolare, risulta fondamentale avere una buona comprensione del moto betatronico non-lineare delle singole particelle, e di come questo interferisce con la qualità del fascio. In questo lavoro vengono studiati sistemi Hamiltoniani sottoposti a perturbazioni stocastiche tramite un framework diffusivo, basato sull'equazione di Fokker-Planck. Tale studio viene poi applicato all'analisi del moto betatronico non-lineare e al problema dell'Apertura Dinamica. In particolare, vengono impostate le basi per formulare un metodo di interpolazione di processi diffusivi simil-Nekhoroshev e viene proposta una procedura sperimentale per misurare gli effetti di diffusione locale all'interno di un acceleratore. L'Apertura Dinamica è una quantità chiave per il comportamento a lungo termine di un acceleratore, tuttavia, la misura di questa quantità nelle simulazioni presenta serie difficoltà dal punto di vista computazionale. È dunque nel nostro interesse riuscire a formulare una legge che descriva la dipendenza dal tempo dell'Apertura Dinamica.

Abstract

In the design phase and simulation of a circular accelerator, it is fundamental to have a proper understanding of the single-particle non-linear betatronic motion, and of how such dynamics interfere with beam quality. In this work we study stochastically-perturbed Hamiltonian systems using a diffusive framework, based on the Fokker-Plank equation. This study is then applied to the analysis of the non-linear betatronic motion and to the Dynamic Aperture problem. In particular, we lay down the basis for formulating an interpolation procedure for Nekhoroshev-like diffusive processes and we propose an experimental procedure for measuring the local diffusive behaviour inside an accelerator. Dynamic Aperture is a key quantity for the long-term behaviour of an accelerator. However, the measure of this quantity via simulation presents serious computational difficulties. Because of that it is in our interest to develop a functional law which models the DA time-dependence.

Contents

Introduction

In the design stage of high-energy particle accelerators, especially those based on superconducting magnets, such as LHC, HL-LHC or the proposed FCC-hh, it is a fundamental issue to take into account the effects of unavoidable non-linear errors in the magnetic field.

One key quantity to consider is the Dynamic Aperture (DA) of an accelerator [1]. The DA is the amplitude of the phase-space region where stable motion occurs. Understanding the features of this stable region is fundamental for controlling the long-term beam dynamics and the phase space region where one can safely operate with the beam.

Multiple studies [2, 3] highlighted how the evolution of DA can be described as a coexistence of weakly chaotic regions, whose escape rate can be described in terms of a Nekhoroshev-like estimate, and invariant KAM tori, affected only by Arnold diffusion, whose relevance is still debated.

As for experimental data, DA measurements at the LHC have been already carried out at injection energy using different approaches, like the standard kick method [4], but also using an innovative method based on the Nekhoroshev-like scaling law of the DA [5]. These experimental studies are also particularly relevant in view of the future High Luminosity LHC (HL-LHC) project [6], for which the operational strategy to set the non-linear correctors in the high-luminosity IRs is still under study.

One of the most critical aspects in DA studies is the computational time required for tracking the non-linear dynamics for long period of times, i.e. for time scales comparable to the operational times of LHC. This makes the investigation of better designs and configurations extremely long and difficult.

A way to tackle adequately this problem would be to develop a robust model for the time dependence of DA. Such robust model would then be interpolated with the result of a short simulation in order to, ideally, extrapolate the long-term behaviour.

We are going to focus on a recent interpolation model based on a diffusive framework [7]. This diffusive framework is based on important results on the dynamics of stochastically-perturbed Hamiltonians.

Any real macroscopic physical system cannot realise the symplectic deterministic character of the Hamiltonian dynamics at arbitrarily small scales due to the unavoidable presence of external random perturbations that destroy the time-coherence in the orbits' evolution. Nevertheless some results of perturbation theory turn out to be robust enough with respect to perturbations of the system under consideration and they can provide effective laws in the study of the stability problem. Nekhoroshev's theorem [8] is among these results and his corresponding estimate for the orbit stability time has been applied in may fields, from celestial mechanics to accelerator physics.

These considerations and results about physical systems, combined with the theory of dynamics of stochastic Hamiltonians, give us the theoretical foundations for the diffusive framework we are interested in utilising for talking the Dynamic Aperture problem and the formulation of valid time-dependent evolution laws.

The final result of this framework is a Fokker-Planck description of the beam shape evolution, with a Nekhoroshev-like diffusion coefficient. Such process can be directly compared with the experimental measures of beam losses which happens at the various collimator locations of the machine. These measures can in fact compared with the outgoing current at an absorbing boundary condition of the Fokker-Planck process.

In this Master Thesis, are going to review some characteristics of this diffusive framework and lay down some first 1D analysis tools for testing the validity of this Fokker-Planck approximation and investigate possible interpolation procedures for the probability current at absorbing boundary conditions.

Moreover, since this framework implies the existence of local diffusive behaviours of the beam distribution *inside* the stable region of phase space, we are going to propose a possible experimental procedure for providing an experimental measurement of local diffusion behaviours in a circular accelerator. This is a critical step if we want to improve the validity of this model beyond the theoretical domain. This procedure will be also simulated in first instance in a simple 1D scenario.

In Chapter 1 we review the fundamental theory on betatron motion and its description via symplectic maps and normal form theory. We also present the concept of one-turn maps and some fundamental tools for approximating

the non-linear terms in the betatron oscillations. Moreover, we also introduce some elements of Birkhoff normal forms and we review the most important particle tracking codes used at CERN for accelerator physics.

In Chapter 2 we introduce the Hénon map and its main characteristics. The simplicity of this map, combined with its physical features, proves to be a very important tool for DA studies and stability analysis.

In Chapter 3 we introduce the concept of Dynamic Aperture, as well as its operative definition when performing a tracking simulation. We will also review its most important time-dependent models in literature, as well as the most consolidated experimental procedures to measure it in an accelerator.

In Chapter 4 we introduce the theory of stochastically-perturbed Hamiltonian systems and review a method to derive a Fokker-Planck equation for describing the dynamics of such systems. We also introduce the fundamental concepts of the Averaging principle, which is fundamental when dealing with the averaging of the angle variable in action-angle dynamical systems.

In Chapter 5 we illustrate the fundamental concepts behind the diffusive framework for describing the time dependence of DA. Next, we illustrate some tools and methods to work with such framework.

Finally, in Chapter 6 we present some preliminary result we obtained via numerical simulations and we illustrate some possible direction of analysis that these results suggest.

Chapter 1

Symplectic maps and non-linear betatron motion

In an accelerator, the transverse particle motion can be described by means of a closed orbit[1] and a small-amplitude betatron oscillation around the closed orbit. This terminology was first derived from researches on transverse particle motion in a betatron in the seminal work of D. Kerst and R. Serber [9]. Now it is used for transverse motion in every type of accelerator.

In this chapter we shall introduce the betatronic motion for a magnetic lattice and its modelling by means of non-linear symplectic maps, which is a standard procedure in accelerator physics [10]. The modelling of the betatronic motion by means of symplectic maps [11] is the essential tool to study the dynamic aperture of a circular collider that is the main topic of this work.

In Section 1.1, we present the equation for the linear betatronic motion from which the theoretical approach is constructed, specifying the fundamental approximations and presenting the fundamental objects for describing the transverse motion in an accelerator. The concept of symplectic transfer map for magnetic elements is introduced in Section 1.2. The composition of such maps to build a one-turn map is presented in Section 1.3. The transfer map for linear magnetic elements are presented in Section 1.4. Some basic concepts of non-linear transfer map theory and methodology are presented in Section 1.5, in order to properly understand the concepts presented in Chap-

[1] In a synchrotron, a closed orbit is defined as a particle trajectory that closes on itself after a complete revolution.

ter 2. In Section 1.6 we briefly present the basic results of Birkhoff's normal forms. Finally, in Section 1.7 we present the 3 main single-particle tracking codes used at CERN.

1.1 Betatron motion

We shall now review the fundamental theory that describes betatron motion. A complete review on the topic is given in [12].

1.1.1 Particle motion and Frenet-Serret Co-ordinate System

The Hamiltonian of a charged particle in an electromagnetic field is given by:

$$H = c\sqrt{m^2c^2 + (\mathbf{P} - e\mathbf{A})^2} + e\mathbf{\Phi} \tag{1.1}$$

where we are expressing the electromagnetic field in terms of the vector potential $\mathbf{A}$ and the scalar potential $\mathbf{\Phi}$

$$\mathbf{E} = -\nabla\Phi - \partial\mathbf{A}/\partial t, \qquad \mathbf{B} = \nabla \times \mathbf{A} \tag{1.2}$$

using the canonical momentum $\mathbf{P} = \mathbf{p} + e\mathbf{A}$, where $\mathbf{p}$ is the mechanical momentum of the particle. The Hamilton's equations of motion are

$$\dot{x} = \frac{\partial H}{\partial p_x}, \quad \dot{P}_x = -\frac{\partial H}{\partial x}, \quad \text{etc.} \tag{1.3}$$

where the overdot stands for the derivative with respect to time t and $(x, P_x), \ldots$ pairs are conjugate phase-space co-ordinates.

It is necessary to adapt this Hamiltonian in a convenient choice of coordinates. Considering the toroidal symmetry of the problem, we can distinguish between two types of motion:

1. **longitudinal**, i.e. along the accelerator circumference (also referred to as *reference trajectory*);

2. **transverse**, i.e. in the plane normal to the longitudinal motion.

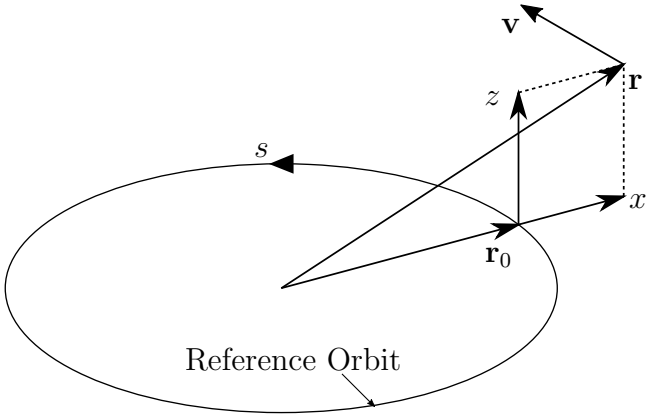


Figure 1.1: Curvilinear coordinate system for particle motion in a circular accelerator. $\mathbf{r}_0(s)$ is the reference orbit, $\hat{x}, \hat{s}$, and $\hat{z}$ form the basis of the curvilinear coordinate system. Any point in phase space can be expressed as $\mathbf{r} = \mathbf{r}_0 + x\hat{x} + z\hat{z}$. In this figure, x and z are betatron coordinates.

Let $\mathbf{r}_0(s)$ be the reference trajectory (see Fig. 1.1), where s is the arc length measured along the closed orbit from a reference initial point. Let us then consider the tangent unit vector to the closed orbit

$$\hat{s}(s) = \frac{\mathrm{d}\mathbf{r}_0(s)}{\mathrm{d}s} \tag{1.4}$$

The unit vector perpendicular to the tangent vector and on the tangential plane is

$$\hat{x}(s) = -\rho(s)\frac{\mathrm{d}\hat{s}(s)}{\mathrm{d}s} \tag{1.5}$$

where $\rho(s)$ defines the radius of curvature. The unit vector orthogonal to the tangential plane is then given by

$$\hat{z}(s) = \hat{x}(s) \times \hat{s}(s) \tag{1.6}$$

The vectors $\hat{x}, \hat{s}$ and $\hat{z}$ form the orthonormal basis for the right-handed Frenet-Serret curvilinear coordinate system with

$$\begin{aligned} \hat{x}'(s) &= \frac{1}{\rho(s)}\hat{s}(s) + \tau(s)\hat{z}(s) \\ \hat{z}'(s) &= -\tau(s)\hat{x}(s) \end{aligned} \tag{1.7}$$

where the prime denotes a differentiation with respect to s, and $\tau(s)$ is the torsion of the curve. For simplicity, we can neglect the torsion and work in a plane-only geometry, where $\tau(s) = 0$. The particle trajectory around the reference orbit can then be expressed as:

$$\mathbf{r}(s) = \mathbf{r}_0(s) + x\hat{x}(s) + z\hat{z}(s) \tag{1.8}$$

We now shall express the equation of motion in terms of the reference orbit co-ordinate system (x, s, z). In order to do so, we perform a canonical transformation by using the generating function

$$F_3(\mathbf{P}; x, s, z) = -\mathbf{P} \cdot [\mathbf{r}_0(s) + x\hat{x}(s) + z\hat{z}(s)] \tag{1.9}$$

where $\mathbf{P}$ is the momentum in the original Cartesian coordinate system. The conjugate momenta for the coordinates (x, s, z) is given by

$$\begin{aligned} p_s &= -\frac{\partial F_3}{\partial s} = (1 + x/\rho)\mathbf{P} \cdot \hat{s}, \\ p_x &= -\frac{\partial F_3}{\partial x} = \mathbf{P} \cdot \hat{x}, \\ p_z &= -\frac{\partial F_3}{\partial z} = \mathbf{P} \cdot \hat{z} \end{aligned} \tag{1.10}$$

This leads us to the new Hamiltonian

$$H = e\Phi + c\left\{m^2c^2 + \frac{(p_s - eA_s)^2}{(1 + x/\rho)^2} + (p_x - eA_x)^2 + (p_z - eA_z)^2\right\}^{1/2} \tag{1.11}$$

where A_s, A_x and A_z are obtained by applying the transformation in Eq. (1.10) to the vector $\mathbf{A}$, i.e.

$$A_s = (1 + x/\rho)\mathbf{A} \cdot \hat{s}, \quad A_x = \mathbf{A} \cdot \hat{x}, \quad A_z = \mathbf{A} \cdot \hat{z} \tag{1.12}$$

At this point we want to use s as the independent variable instead of time. To do so, we can use the relation $dH = (\partial H/\partial p_x)dp_x + (\partial H/\partial p_s)dp_s = 0$ or

$$x' = \frac{\mathrm{d}x}{\mathrm{d}s} = \frac{\dot{x}}{\dot{s}} = \left(\frac{\partial H}{\partial p_x}\right)\left(\frac{\partial H}{\partial p_s}\right)^{-1}, \qquad \text{etc.}, \tag{1.13}$$

this leads us to

$$\begin{aligned} t' = \frac{\partial p_s}{\partial H},\ H' &= -\frac{\partial p_s}{\partial t} \\ x' = -\frac{\partial p_s}{\partial p_x},\ p_x' &= \frac{\partial p_s}{\partial x} \\ z' = -\frac{\partial p_s}{\partial p_z},\ p_z' &= \frac{\partial p_s}{\partial z} \end{aligned} \tag{1.14}$$

where the prime always denotes differentiation with respect to s. This is Hamilton's equation of motion with s as the independent variable, $-p_s$ as the new Hamiltonian, and the conjugate phase-space coordinates given by

$$x, p_x; \qquad z, p_z; \qquad t, -H. \tag{1.15}$$

When the scalar and vector potentials ϕ and $\mathbf{A}$ are time independent, the new Hamiltonian $-p_s$ is also time independent. The transformation reduces the degrees of freedom from three to two at the price of having the new Hamiltonian depending on the new variable s. Because of the circular nature of the circular accelerator, such dependence on s is periodic. As we will see, this periodicity is exploited in the analysis of linear and non-linear betatron motion. The new Hamiltonian $\tilde{H} = -p_s$ is then expressed by

$$\tilde{H} = -\left(1 + \frac{x}{\rho}\right)\left[\frac{(H - e\phi)^2}{c^2} - m^2c^2 - (p_x - eA_x)^2 - (p_z - eA_z)^2\right]^{1/2} - eA_s \tag{1.16}$$

The total energy and momentum of the particles are $E = H - e\phi$ and $p = \sqrt{E^2/c^2 - m^2c^2}$, respectively. Since the transverse momenta p_x and p_z are much smaller than the total one, we can expand the Hamiltonian up to second order in p_x and p_z

$$\tilde{H} \approx -p\left(1 + \frac{x}{\rho}\right) + \frac{1 + x/\rho}{2p}\left[(p_x - eA_x)^2 + (p_z - eA_z)^2\right] - eA_s \tag{1.17}$$

1.1.2 EM fields in Frenet-Serret Co-ordinate System

In the Frenet-Serret co-ordinate system $(\hat{x}, \hat{s}, \hat{z})$, the scale factor becomes

$$h_x = 1, \qquad h_s = 1 + \frac{x}{\rho}, \qquad h_z = 1 \tag{1.18}$$

We have then the following equations for the electromagnetic fields:

$$\begin{aligned}
\nabla\Phi &= \frac{\partial\Phi}{\partial x}\hat{x} + \frac{1}{h_s}\frac{\partial\Phi}{\partial s}\hat{s} + \frac{\partial\Phi}{\partial z}\hat{z} \\
\nabla\cdot\mathbf{A} &= \frac{1}{h_s}\left[\frac{\partial\left(h_s A_1\right)}{\partial x} + \frac{\partial A_2}{\partial s} + \frac{\partial\left(h_s A_3\right)}{\partial z}\right] \\
\nabla\times\mathbf{A} &= \frac{1}{h_s}\left[\frac{\partial A_3}{\partial s} - \frac{\partial\left(h_s A_2\right)}{\partial z}\right]\hat{x} + \left[\frac{\partial A_1}{\partial z} - \frac{\partial A_3}{\partial x}\right]\hat{s} + \\
&\quad + \frac{1}{h_s}\left[\frac{\partial\left(h_s A_2\right)}{\partial x} - \frac{\partial A_1}{\partial s}\right]\hat{z} \\
\nabla^2\Phi &= \frac{1}{h_s}\left[\frac{\partial}{\partial x}h_s\frac{\partial\Phi}{\partial x} + \frac{\partial}{\partial s}\frac{1}{h_s}\frac{\partial\Phi}{\partial s} + \frac{\partial}{\partial z}h_s\frac{\partial\Phi}{\partial z}\right]
\end{aligned} \tag{1.19}$$

where $A_1 = \mathbf{A}\cdot\hat{x}$, $A_2 = \mathbf{A}\cdot\hat{s}$, and $A_3 = \mathbf{A}\cdot\hat{z}$. In accelerator physics, we are interested in the case with zero electric potential $\Phi = 0$, furthermore, for an accelerator with transverse magnetic fields, we can assume $A_x = A_z = 0$. All considered, we can directly express the two-dimensional magnetic field as

$$\mathbf{B} = B_x(x,z)\hat{x} + B_z(x,z)\hat{z} \tag{1.20}$$

where

$$B_x = -\frac{1}{h_s}\frac{\partial\left(h_s A_2\right)}{\partial z} = -\frac{1}{h_s}\frac{\partial A_s}{\partial z}, \quad B_z = \frac{1}{h_s}\frac{\partial\left(h_s A_2\right)}{\partial x} = \frac{1}{h_s}\frac{\partial A_s}{\partial x} \tag{1.21}$$

with $A_s = h_s A_2$. Using Maxwell's equation $\nabla\times\mathbf{B} = 0$, we have

$$\frac{\partial}{\partial z}\frac{1}{h_s}\frac{\partial A_s}{\partial z} + \frac{\partial}{\partial x}\frac{1}{h_s}\frac{\partial A_s}{\partial x} = 0 \tag{1.22}$$

For a straight geometry, i.e. $h_s = 1$, we have that $\nabla_\perp^2 A_s = 0$ and that A_s can be expanded in power series as

$$A_s = B_0\,\mathrm{Re}\left[\sum_{n=0}^{\infty}\frac{b_n + ia_n}{n+1}(x+iz)^{n+1}\right] \tag{1.23}$$

and obtain then $B_z = \frac{\partial A_s}{\partial x}$ and $B_x = -\frac{\partial A_s}{\partial z}$. The normalization constant B_0 is chosen equal to the main dipole field strength, so that $b_0 = 1$. This leads us to obtain $B_0 b_0 = -[B\rho]/\rho$, where $B\rho$ is defined as **momentum rigidity**

of the beam, ρ is the **bending radius**, and $b_0 = 1$. The resulting magnetic flux density is in the end given by

$$B_z + iB_x = B_0 \sum_{n=0}^{\infty} (b_n + ia_n)(x + iz)^n \tag{1.24}$$

with

$$b_n = \frac{1}{B_0 n!} \frac{\partial^n B_z}{\partial x^n}\bigg|_{x=z=0} \quad a_n = \frac{1}{B_0 n!} \frac{\partial^n B_x}{\partial x^n}\bigg|_{x=z=0} \tag{1.25}$$

where b_n, a_n are called $2(n+1)$th multipole coefficients. The effective multipole field on the beams then becomes

$$\frac{1}{B\rho}(B_z + iB_x) = \mp\frac{1}{\rho} \sum_{n=0}^{\infty} (b_n + ia_n)(x + iz)^n \tag{1.26}$$

where the $-$ and $+$ signs are used for particles with positive and negative charges respectively.

1.1.3 Equation of Betatron Motion

If we disregard the effects of synchrotron motion (whose frequency is usually very slow when compared to the typical betatron motion frequencies), we have a completely time-independent Hamiltonian equation for betatron motion, whose equations are given by:

$$x' = \frac{\partial \tilde{H}}{\partial p_x}, \quad p_x' = -\frac{\partial \tilde{H}}{\partial x}, \quad z' = \frac{\partial \tilde{H}}{\partial p_z}, \quad p_z' = -\frac{\partial \tilde{H}}{\partial z} \tag{1.27}$$

where $\tilde{H}$ is given by Eq. (1.17). Combining them with the transverse magnetic fields given by Eq. (1.21), the betatron equations of motion are given by

$$\begin{cases} x'' - \frac{\rho + x}{\rho^2} = \pm\frac{B_z}{B\rho}\frac{p_0}{p}\left(1 + \frac{x}{\rho}\right)^2 \\ z'' = \mp\frac{B_x}{B\rho}\frac{p_0}{p}\left(1 + \frac{x}{\rho}\right)^2 \end{cases} \tag{1.28}$$

the upper and lower signs correspond again to the positive and negative charged particle, respectively, p is the particle's momentum, p_0 is the momentum of a reference particle, $B\rho = p_0/e$ is the **magnetic rigidity** and e is the charge of the particle.

At this point, we can consider a magnetic field linear in both x and z. We set $B_z = B\rho K_1 x$ and $B_x = B\rho K_2 z$. If we define $K_x = \frac{1}{\rho^2} - K_1$ and $K_z = K_2$, Eqs. (1.28) become

$$x'' + K_x(s)x = 0, \qquad z'' + K_z(s)z = 0 \tag{1.29}$$

which are known in literature as **Hill's equations**.

1.1.4 Analysis of Hill's equations

If we consider the dependence on s in the equations (1.29), it is possible to write down a solution starting from the ansatz [10]

$$\begin{cases} x(s) = \sqrt{\varepsilon_x \beta_x(s)}\cos(\psi_x(s) + \delta_x) \\ z(s) = \sqrt{\varepsilon_z \beta_y(z)}\cos(\psi_z(s) + \delta_z) \end{cases} \tag{1.30}$$

these equations represent two harmonic oscillators where the amplitudes $\beta_x(s), \beta_z(s)$ and phase advances $\psi_x(s), \psi_z(s)$ are all dependent on s. The constant values (invariant of motion) $\varepsilon_x, \varepsilon_z, \delta_x, \delta_z$ are related to the initial conditions.

Applying the ansatz (1.30) to Eq. (1.29), we obtain the following

$$\begin{aligned} &\dot{\psi}_x\dot{\beta}_x + \ddot{\psi}_x\beta_x = 0 \\ &\dot{\psi}_z\dot{\beta}_z + \ddot{\psi}_z\beta_z = 0 \\ &\ddot{\beta}_x\frac{1}{2} - \dot{\beta}_x^2\frac{1}{4\beta_x} - \beta_x\dot{\psi}_x^2 + \beta_x K_x(s) = 0 \\ &\ddot{\beta}_z\frac{1}{2} - \dot{\beta}_z^2\frac{1}{4\beta_z} - \beta_z\dot{\psi}_z^2 + \beta_z K_z(s) = 0 \end{aligned} \tag{1.31}$$

where, for notation convenience, we used the dot to represent a derivative with respect to s instead of the prime. From the first two equations we obtain the relation between amplitude and phase advance

$$\psi_x(s) = c_x \int_0^s \frac{d\sigma}{\beta_x(\sigma)} \qquad \psi_z(s) = c_z \int_0^s \frac{d\sigma}{\beta_z(\sigma)} \tag{1.32}$$

with c_x, c_z constant factors and with everything set so that $\psi_x(0) = \psi_z(0) = 0$. Setting $c_x = c_z = 1$ determines β_x, β_y uniquely. Of course, since gradients and the bending radius are periodic of period L, i.e. the total length of the

circular accelerator, β_x, β_y may be taken to be periodic as well with the same period.

In accelerator physics, β_x and β_y are called the **betatron functions** of the accelerator, and they determine the dimension and the shape of the beam itself; the constants $\varepsilon_x, \varepsilon_y$ are the **emittances** of a single particle in the transversal phase space, which we shall present in detail in Subsec. 1.1.5; and the phase advances per turn divided by 2π

$$\nu_x = \frac{\psi_x(L)}{2\pi}, \qquad \nu_y = \frac{\psi_y(L)}{2\pi} \tag{1.33}$$

are called the **linear tunes** of the circular accelerator.

1.1.5 Courant-Snyder invariant and emittance

We will now analyse the physical meaning of the emittance ε. Let y be one of the transverse co-ordinates. The linear betatron motion is described by Hill's equation

$$y'' + K(s)y = 0 \tag{1.34}$$

and it is possible to write the general solution in the form

$$y(s) = a\sqrt{\beta_y(s)}\cos\left(\psi_y(s) + \xi_s\right) \tag{1.35}$$

with a and ξ fixed by the initial condition and

$$\psi_y(s) = \int_0^s \frac{\mathrm{d}s'}{\beta_y\left(s'\right)} \tag{1.36}$$

We now perform the following transformation into action-angle variables

$$y = \sqrt{2\beta_y J_y}\cos\psi_y \quad p_y = -\sqrt{2\beta_y J_y}\sin\psi_y \tag{1.37}$$

we can observe that the following relation holds true

$$p_y = \beta_y\frac{\mathrm{d}y}{\mathrm{d}s} - \frac{1}{2}\frac{\mathrm{d}\beta_y}{\mathrm{d}s}y \tag{1.38}$$

in fact, substituting the solution of $y(s)$, we have

$$p_y = \beta_y\frac{\mathrm{d}y}{\mathrm{d}s} - \frac{1}{2}\frac{\mathrm{d}\beta_y}{\mathrm{d}s}y = -a\sqrt{\beta(s)}\sin\left(\psi_s\right) \tag{1.39}$$

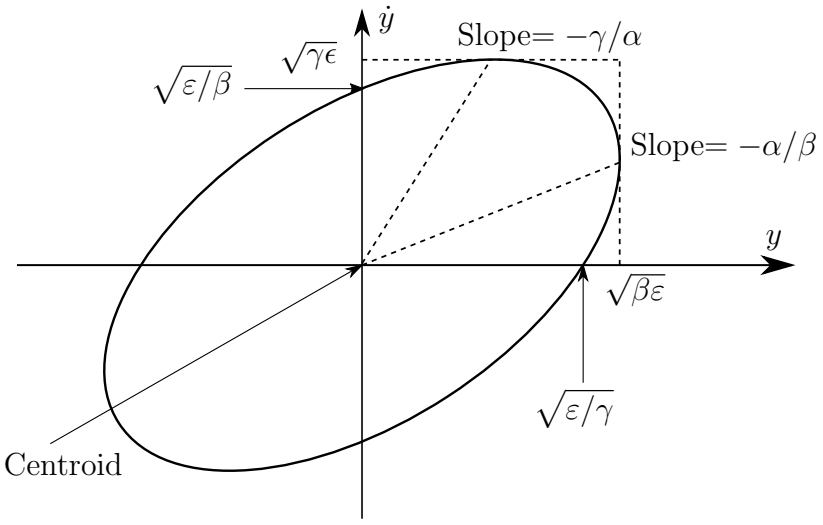


Figure 1.2: The Courant-Snyder invariant ellipse. The area enclosed by the ellipse is equal to $\pi\varepsilon$, where ε is twice the betatron action; α, β and γ are the Twiss parameters.

and we can define the **Courant-Snyder ellipse** on the (y, y') plane

$$C\left(y, y'\right) = \frac{y^2 + p_y^2}{\beta_y(s)} = \gamma y^2 + 2\alpha y y' + \beta y'^2 \tag{1.40}$$

Substituting the solutions for y and p_y, we obtain

$$C\left(y, y'\right) = \frac{1}{\beta}\left(a^2\beta\cos^2\psi + a^2\beta\sin^2\psi\right) = a^2 \equiv \varepsilon \tag{1.41}$$

this implies that the area enclosed by such an ellipse (see Fig. 1.2) will be equal to $\pi\varepsilon$.

Being also $a = \sqrt{\varepsilon}$, we can write (and thus satisfy the ansatz of Eq. (1.30))

$$y(s) = \sqrt{\varepsilon_y\beta_y}\cos\psi_y \tag{1.42}$$

It is clear that $\varepsilon_y = 2J_y$, i.e. twice the value of the action variable of the particle.

This ε is called **Courant-Snyder invariant** and is a property relative to a single particle. It should not be confused with the **emittance**, which is relative to a beam, i.e. a distribution of particles in phase space. However, these two quantities are deeply related, as we will show.

Let $\rho(y, y')$ be a normalised beam distribution in phase space. The average, the standard deviations and the correlation coefficients are given by

$$\langle y \rangle = \int \mathrm{d}y\, \mathrm{d}y'\, y\rho\left(y, y'\right) \quad \langle y' \rangle = \int \mathrm{d}y\, \mathrm{d}y'\, y'\rho\left(y, y'\right) \tag{1.43}$$

$$\sigma_y^2 = \int \mathrm{d}y\, \mathrm{d}y'\, (y - \langle y \rangle)^2 \rho\left(y, y'\right) \quad \sigma_{y'}^2 = \int \mathrm{d}y\, \mathrm{d}y'\, \left(y' - \left\langle y' \right\rangle\right)^2 \rho\left(y, y'\right) \tag{1.44}$$

$$\sigma_{yy'} = \int \mathrm{d}y\, \mathrm{d}y'\, (y - \langle y \rangle)\left(y' - \left\langle y' \right\rangle\right) \rho\left(y, y'\right) \tag{1.45}$$

Thus, we can define the **rms emittance** as

$$\varepsilon_{\mathrm{rms}} = \sqrt{\sigma_y^2 \sigma_{y'}^2 - \sigma_{yy'}^2} = \sigma_y \sigma_{y'} \sqrt{1 - r^2} \tag{1.46}$$

where r is the correlation coefficient. We now want to relate the single-particle, betatron emittance with this statistical inferred value.

Assuming that particles are uniformly distributed in an ellipse

$$y^2/a^2 + y'^2/b^2 = 1 \tag{1.47}$$

we have that the total phase-space area under the ellipse is $\mathcal{A} = \pi ab = 4\pi\varepsilon_{\mathrm{rms}}$. The factor 4 is often used in literature for the definition of the full emittance, i.e. $\varepsilon = 4\varepsilon_{\mathrm{rms}}$, to ensure that the phase-space area of such an ellipse is $\pi\varepsilon$. Such rms emittance is invariant under a coordinate rotation.

In accelerators, described by Hill's equation, particles are distributed in the Courant-Snyder ellipse

$$C(y, y') = \gamma y'^2 + 2\alpha y y' + \beta y^2 \tag{1.48}$$

Using the previous result and the coordinate rotation, it is possible to show that

$$\epsilon_{\mathrm{rms}} = \frac{\sigma_y^2}{\beta} = \frac{\sigma_{y'}^2}{\gamma}, \quad r = -\frac{\alpha}{\sqrt{\beta\gamma}} \tag{1.49}$$

or

$$\begin{pmatrix} \sigma_y^2 & \sigma_{yy'} \\ \sigma_{yy'} & \sigma_{y'}^2 \end{pmatrix} = \epsilon_{\mathrm{rms}} \begin{pmatrix} \beta & -\alpha \\ -\alpha & \gamma \end{pmatrix} \tag{1.50}$$

1.2 Element maps

We want to consider the dynamics of a single particle in a circular magnetic lattice composed of L magnetic elements denoted $\mathcal{M}^{(1)}, \ldots, \mathcal{M}^{(L)}$. We focus our analysis on the horizontal and vertical axes x, z, with the formalism

specified in the previous Section. By neglecting the coupling with the longitudinal motion, we can analyse only the transverse plane (x, z) dynamics. We analyse the motion in a 4D phase space, whose points are given by:

$$\mathbf{x} = \begin{pmatrix} x \\ p_x \\ z \\ p_z \end{pmatrix} \quad \text{or alternately} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}; \tag{1.51}$$

where

$$p_x \equiv \frac{\mathrm{d}x}{\mathrm{d}s} \quad p_z \equiv \frac{\mathrm{d}z}{\mathrm{d}s} \tag{1.52}$$

this second notation will be used whenever we need to express an implicit sum over the index of x_k.

The gradients can be well approximated by using s-independent functions inside the single magnets $\mathcal{M}^{(l)}$ that constitute the magnetic lattice of length s_l (like shown in Fig. 1.3). As a consequence, it is convenient to use the transfer map $\mathbf{M}^{(l)}$ of the magnetic element $\mathcal{M}^{(l)}$, which is located between the positions s_{l-1} and s_l. With $\mathbf{M}^{(l)}$ we denote a function that transforms the phase space coordinates $\mathbf{x}(s_{l-1})$ into $\mathbf{x}(s_l)$ by using Eq. (1.28):

$$\mathbf{x}(s_l) = \mathbf{M}^{(l)}(\mathbf{x}(s_{l-1})) \qquad \mathbf{M}^{(l)} : \mathbf{R}^4 \to \mathbf{R}^4 \tag{1.53}$$

Since the equations of transverse motion (1.28) are obtained from an $s-$dependent Hamiltonian, $\mathbf{M}^{(l)}$ can be interpreted as an Hamiltonian flow that propagates an initial condition $\mathbf{x}(s_{l-1})$ to $\mathbf{x}(s_l)$.

A generic map $\mathbf{M} : \mathbf{R}^4 \to \mathbf{R}^4$ is **symplectic** if its Jacobian $\mathbf{M}_J$, given by the equation

$$\mathbf{M}_{J\,k,l}(\mathbf{x}) \equiv \frac{\partial M_k}{\partial x_l}(\mathbf{x}), \tag{1.54}$$

is a symplectic matrix for every $\mathbf{x}$, meaning

$$\mathbf{M}_J(\mathbf{x})\mathbf{J}\mathbf{M}_J^T(\mathbf{x}) = \mathbf{J} \tag{1.55}$$

where $\mathbf{M}_J^T$ denotes the transposed matrix and $\mathbf{J}$ is defined as

$$\mathbf{J} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \tag{1.56}$$

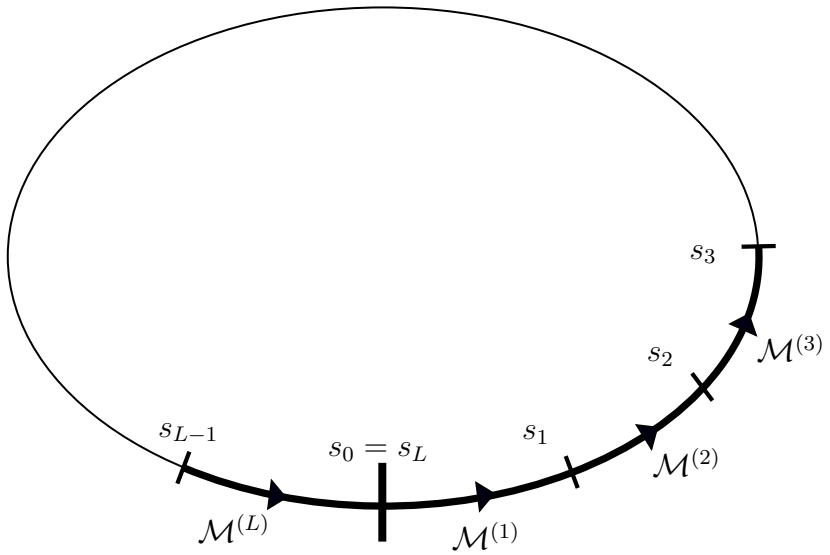


Figure 1.3: Simple sketch of a circular accelerator represented as a composition of magnetic elements. (If we take $s_0 = 0$, the length of the accelerator is s_L)

Another equivalent definition of symplectic map is given in terms of Poisson brackets

$$\{f, g\} = \sum_{i=x,y} \frac{\partial f}{\partial i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial i} \tag{1.57}$$

where $f, g : \mathbf{R}^4 \to \mathbf{R}$ and are defined on the phase space. In fact the definition of symplectic map (1.55) is equivalent to the following Poisson bracket notation on the components of the map $\mathbf{M}$:

$$\{M_i, M_j\} = \mathbf{J}_{i,j} \qquad i, j = 1, \ldots, 4. \tag{1.58}$$

Finally, a map $\mathbf{M} : \mathbf{R}^2 \to \mathbf{R}^2$ is defined area preserving if its Jacobian $\mathbf{M}_j$ has determinant equal to one. More explicitly

$$\det(\mathbf{M_j}) \equiv \frac{\partial M_1}{\partial x}\frac{\partial M_2}{\partial p} - \frac{\partial M_1}{\partial p}\frac{\partial M_2}{\partial x} = 1 \tag{1.59}$$

this condition is equivalent to the symplectic condition in $\mathbf{R}^2$. In fact, the symplectic condition can be considered as the natural generalisation of the area-preserving condition to higher dimensions.

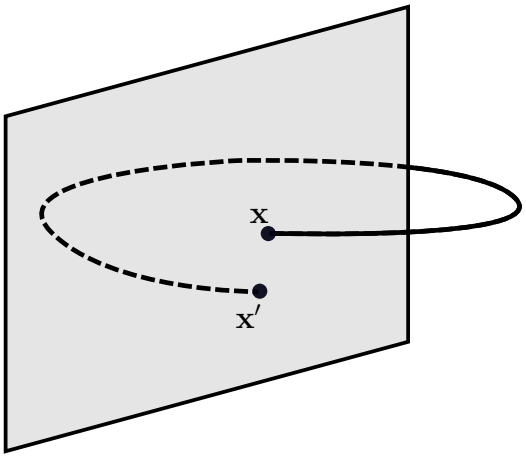


Figure 1.4: Simple sketch of the concept behind a Poincaré section of a circular accelerator.

1.3 One-turn maps

A one-turn map $\mathbf{M}$ is the Poincaré map of the circular accelerator at section $s = s_0$, and it's given by composing single-element maps:

$$\mathbf{M} = \mathbf{M}^{(L)} \circ \mathbf{M}^{(L-1)} \circ \ldots \mathbf{M}^{(2)} \circ \mathbf{M}^{(1)} \tag{1.60}$$

This map M transforms the phase space coordinates $\mathbf{x}(s_0)$ of a particle into new coordinates $\mathbf{x}(s_L)$ which correspond to the particle after one full turn (See Fig. 1.4). The equation of motion thus is

$$\mathbf{x}'(s_L) = \mathbf{M}(\mathbf{x}(s_0)). \tag{1.61}$$

It is possible of course to expand this map in a Taylor series around the origin (which in our definitions is a fixed point) in the form

$$x_j(s_L) = \sum_{i_1,i_2,i_3,i_4 \geq 0} m_{j;i_1,i_2,i_3,i_4} (x_1(s_0))^{i_1} (x_2(s_0))^{i_2} (x_3(s_0))^{i_3} (x_4(s_0))^{i_4} \tag{1.62}$$

with $j = 1, 2, 3, 4$ and $m_{j;i_1,i_2,i_3,i_4}$ real coefficients. These coefficients are dependent on the beam energy, which is considered constant in our stability studies (since we are considering the beam after the acceleration phase is ended).

1.4 Linear one-turn maps.

We can separate the linear part of the map from the non-linear contribution and obtain the formulation

$$\mathbf{M}(\mathbf{x}) = \mathbf{L}(\mathbf{x}) + O(|\mathbf{x}|^2) \tag{1.63}$$

where $\mathbf{L}$ is the linearised map as a 4×4 matrix. This formulation allows us to write directly its action on the coordinates as:

$$x'_j = \sum_{k=1}^{4} \mathbf{L}_{j,k} x_k \tag{1.64}$$

Moreover, with linearised maps, the composition operation simply becomes a multiplication of matrices.

In this framework, given a particle with initial condition $\mathbf{x}(s_0)$, we can computing its **orbit** accordingly to this basic tracking algorithm

$$\mathbf{x}(s_{jL+l}) = \mathbf{M}^{(l)}(\mathbf{x}(s_{jL+l-1})) \tag{1.65}$$

where $\mathbf{x}(s_{jL+l})$ is the point of the orbit at $s = s_l$ after j full revolutions. The orbits at a fixed section s_j are given by the sequence $\{\mathbf{x}(s_{jL+l})\}_{j=0}^{\infty}$ for fixed l. This can be obtained with the iteration of the one-turn map $\mathbf{M}(s_l)$

$$\mathbf{x}(s_{(j+1)L+l}) = \mathbf{M}(s_l; \mathbf{x}(s_{jL+l})) \tag{1.66}$$

1.4.1 Magnetic linear transfer maps

The three fundamental ingredients of a circular accelerator are drifts, dipoles and quadrupoles (neglecting all the non-linearities caused by errors and correctors).

In order to describe them in terms of linear transfer maps, we can consider the linearisation of Eq. (1.29). If we consider the simplest cases in which we have constant gradients and bending radius, it is possible to describe these three fundamental elements.

Drift

The drift is a straight empty section, i.e. $K_x = 0, K_z = 0$. Its corresponding linear map $\mathbf{L}_{\text{drift}}$ can be written as

$$\mathbf{L}_{\text{drift}} = \begin{pmatrix} 1 & l & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & l \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.67}$$

where l is the drift element length.

Dipole

The dipole is a rectangular straight section of length l with constant dipolar field, i.e. $K_x = B_0, K_z = 0$ or the other way around. Its corresponding linear map $\mathbf{L}_{\text{dipole}}$ can be written as

$$\mathbf{L}_{\text{dipole}} = \begin{pmatrix} 1 & K_x^{-1} \sin K_x & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos K_x & \rho \sin K_x \\ 0 & 0 & -K_x \sin K_x & \cos K_x \end{pmatrix} \tag{1.68}$$

This is valid for rectangular magnets, and also takes into account the edge focusing effect [13].

Quadrupole

The quadrupole is a straight section of length l with a constant quadrupolar field and it can be either a defocusing quadrupole or a focusing one, i.e. we can have either $K_x = K, K_z = -K$ or $K_z = K, K_x = -K$. The corresponding linear maps $\mathbf{L}_{\text{dquad}}$ and $\mathbf{L}_{\text{fquad}}$ are given by

$$\mathbf{L}_{\text{dquad}} = \begin{pmatrix} \cosh\left(\sqrt{K}l\right) & \frac{1}{\sqrt{K}} \sinh\left(\sqrt{K}l\right) & 0 & 0 \\ \sqrt{K} \cosh\left(\sqrt{K}l\right) & \cosh\left(\sqrt{K}l\right) & 0 & 0 \\ 0 & 0 & \cos\left(\sqrt{K}l\right) & \frac{1}{\sqrt{K}} \sin\left(\sqrt{K}l\right) \\ 0 & 0 & -\sqrt{K} \sin\left(\sqrt{K}l\right) & \cos\left(\sqrt{K}l\right) \end{pmatrix} \tag{1.69}$$

$$\mathbf{L}_{\text{fquad}} = \begin{pmatrix} \cos\left(\sqrt{K}l\right) & \frac{1}{\sqrt{K}} \sin\left(\sqrt{K}l\right) & 0 & 0 \\ -\sqrt{K} \sin\left(\sqrt{K}l\right) & \cos\left(\sqrt{K}l\right) & 0 & 0 \\ 0 & 0 & \cosh\left(\sqrt{K}l\right) & \frac{1}{\sqrt{K}} \sinh\left(\sqrt{K}l\right) \\ 0 & 0 & \sqrt{K} \cosh\left(\sqrt{K}l\right) & \cosh\left(\sqrt{K}l\right) \end{pmatrix} \tag{1.70}$$

One-turn map for the Hill equation

Let us now try to build the one-turn map for a generic Hill equation in the form of Eq. (1.29), for the computation of the one-turn linear map, we start by letting

$$\mathbf{w} = \begin{pmatrix} \sqrt{\varepsilon_x}\sin\delta_x \\ \sqrt{\varepsilon_x}\cos\delta_x \\ \sqrt{\varepsilon_z}\sin\delta_z \\ \sqrt{\varepsilon_z}\cos\delta_z \end{pmatrix} \tag{1.71}$$

then, we can write Eq. (1.30) in the compact form

$$\mathbf{x}(s) = \mathbf{T}(s)\mathbf{R}(\psi(s))\mathbf{w} \tag{1.72}$$

where $\mathbf{T}(\mathbf{s})$ is

$$\mathbf{T}(\mathbf{s}) = \begin{pmatrix} \sqrt{\beta_x(s)} & 0 & 0 & 0 \\ \frac{-\alpha_x(s)}{\sqrt{\beta_x(s)}} & \frac{1}{\sqrt{\beta_x(s)}} & 0 & 0 \\ 0 & 0 & \sqrt{\beta_z(s)} & 0 \\ 0 & 0 & \frac{-\alpha_z(s)}{\sqrt{\beta_z(s)}} & \frac{1}{\sqrt{\beta_z(s)}} \end{pmatrix} \tag{1.73}$$

with $\alpha_x(s) = -\dot{\beta}_x(s)/2$ and $\alpha_z(s) = -\dot{\beta}_z(s)/2$, and $\mathbf{R}(\psi(s))$ direct sum of two rotation matrices

$$\mathbf{R}(\psi(s)) = \begin{pmatrix} \cos\psi_x(s) & \sin\psi_x(s) & 0 & 0 \\ -\sin\psi_x(s) & \cos\psi_x(s) & 0 & 0 \\ 0 & 0 & \cos\psi_z(s) & \sin\psi_z(s) \\ 0 & 0 & -\sin\psi_z(s) & \cos\psi_z(s) \end{pmatrix} \tag{1.74}$$

Starting to evaluate Eq. (1.29) at $s = s_0 = 0$, we have $\mathbf{x}(0) = \mathbf{T}\mathbf{w}$, Therefore

$$\mathbf{x}(s) = \mathbf{T}(s)\mathbf{R}(\psi(s))\mathbf{T}(0)^{-1}\mathbf{x}(0) \tag{1.75}$$

It is possible of course to build a one-turn map from Eq. (1.75) by setting $s = s_L$ and taking into account the periodicity of $\mathbf{T}(s_l) = \mathbf{T}(0)$. We can then write directly

$$\mathbf{x}' = \mathbf{L}\mathbf{x}, \qquad \mathbf{L} = \mathbf{T}\mathbf{R}\mathbf{T}^{-1} \tag{1.76}$$

where $\mathbf{R} \equiv \mathbf{R}(\psi(\mathbf{s_L})), \mathbf{T} \equiv \mathbf{T}(0), \mathbf{x} \equiv \mathbf{x}(0)\text{and}\mathbf{x}' \equiv \mathbf{x}(s_L)$. If we perform the matrix products in this last equation we obtain the explicit expression of $\mathbf{L}$,

which is known in literature as the Twiss matrix

$$\mathbf{L} = \begin{pmatrix} \mathbf{L}_x & 0 \\ 0 & \mathbf{L}_z \end{pmatrix} \tag{1.77}$$

where

$$\mathbf{L}_x = \begin{pmatrix} \cos(2\pi\nu_x) + \alpha_x \sin(2\pi\nu_x) & \beta_x \sin(2\pi\nu_x) \\ -\frac{1+\alpha_x^2}{\beta_x} \sin(2\pi\nu_x) & \cos(2\pi\nu_x) - \alpha_x \sin(2\pi\nu_x) \end{pmatrix} \tag{1.78}$$

and $\mathbf{L}_z$ has a very similar expression. $\beta_x, \beta_z, \alpha_x, \alpha_z$ are evaluated at $s = 0$.

1.5 Non-linear transfer maps

In general, it is not possible to compute exactly the transfer map of a non linear element $\mathcal{M}^{(l)}$. Therefore it becomes necessary to develop different strategies in order to deal with sextupole and higher order multipoles.

For dealing with the second order differential equation (1.28) (and higher approximations of it) we start by writing it as a first order differential equation for $\mathbf{x} \equiv (x, p_x, z, p_z)$:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = \mathbf{A}(s)\mathbf{x} + \mathbf{f}(x, z; s) \tag{1.79}$$

with $\mathbf{A}$ linear part of the equation and $\mathbf{f}$ the non-linear contributions.

Let us now define $\mathbf{L}(s, s_{l-1})$ as the fundamental matrix for the linear system so that

$$\begin{aligned} \frac{\mathrm{d}\mathbf{L}(s, s_{l-1})}{\mathrm{d}s} &= \mathbf{A}(s)\mathbf{L}(s, s_{l-1}) \\ \mathbf{L}(s_{l-1}, s_{l-1}) &= \mathbf{I} \\ \mathbf{L}(s'', s')\mathbf{L}(s', s) &= \mathbf{L}(s'', s) \qquad \text{(semigroup property)} \end{aligned} \tag{1.80}$$

then, the non-linear Eq. (1.79) has the following solution in integral form

$$\begin{aligned} \mathbf{x}(s) = \mathbf{L}(s, s_{l-1})\mathbf{x}(s_{l-1}) + \int_{s_{l-1}}^{s} \mathbf{L}(s_{l-1}, s')\mathbf{f}(x, y; s')\, ds' = \\ = \mathbf{L}(s, s_{l-1}) \left(\mathbf{x}(s_{l-1}) + \int_{s_{l-1}}^{s} \mathbf{L}(s_{l-1}, s')\mathbf{f}(x, y; s')\, ds' \right) \end{aligned} \tag{1.81}$$

with $\mathbf{x}(s_{l-1})$ initial condition.

Usually, on a single magnetic element, $\mathbf{A}(s)$ has constant value and, therefore, it is possible to write a simpler form $\mathbf{A}(s) = \mathbf{A}$ and, consequently

$$\mathbf{L}(s, s_{l-1}) = e^{(s-s_{l-1})\mathbf{A}} \quad \text{for} \quad s_{l-1} \leq s \leq s_l \tag{1.82}$$

where we are using the standard exponential matrix definition through the Taylor series

$$e^{\mathbf{sA}} \equiv \sum_{n=0}^{+\infty} s^n \frac{\mathbf{A}^n}{n!} \tag{1.83}$$

With the problem set in this form, it is now possible to apply different approaches for adequately taking into consideration the non linear term in Eq. (1.79) and obtain approximated solutions for the equations of motion.

1.5.1 Thick lens approximation.

This approximation is based on a recursive method: as first step we compute an approximate solution $\mathbf{x}^{(0)}(s)$ that solves the linear equation

$$\mathbf{x}^{(0)}(s) = \mathbf{L}(s, s_{l-1})\mathbf{x}(s_{l-1}) \tag{1.84}$$

next, the successive approximations $\mathbf{x}^{(n)}(s)$ are given by substituting $\mathbf{x}^{(n-1)}(s)$ into the r.h.s. of Eq. (1.81):

$$\mathbf{x}^{(n)}(s) \equiv \mathbf{L}(s, s_{l-1}) \left(\mathbf{x}^{(n-1)}(s_{l-1}) + \int_{s_{l-1}}^{s} \mathbf{L}(s_{l-1}, s')\mathbf{f}(x^{(n-1)}, y^{(n-1)}; s')\, ds' \right) \tag{1.85}$$

At each step we get a better approximation to the solution, but it's an approximation that is not exactly symplectic, regardless of the polynomial order n used in the initial conditions. We will have in fact inevitable violations by terms of order $|\mathbf{x}|^n$.

1.5.2 Thin lens approximation.

In this approximation, instead, we build the polynomial approximation of the element map $\mathbf{M}^{(l)}$ by concentrating the non-linearity in one or more places of the interval $[s_{l-1}, s_l]$. With this approach it is possible to obtain maps that are exactly symplectic (this reason and the easiness to implement these maps

numerically, makes this approach the standard in many tracking codes). This approach can be considered only when the magnetic length is small.

The **one-kick approximation** (Fig. 1.5) approaches Eq. (1.81) by concentrating the non-linear content $\mathbf{f}$ only at the beginning of the magnetic element by substituting

$$\mathbf{f}(x, z; s) \to \mathbf{f}(x, z; s)\ell\delta(s - (s_{l-1} + 0)) \tag{1.86}$$

where the plus zero indicates that the non-linearity acts 'immediately after' the beginning of the element. This makes the computation of the rest of the transfer map completely straightforward: from $s_{l-1} + \varepsilon$ to s_l we have

$$\mathbf{x}(s_l) = \mathbf{L}(s_l, s_{l-1} + \varepsilon)\mathbf{x}(s_{l-1} + \varepsilon) \tag{1.87}$$

The non-linear kick has no effect on the coordinate x:

$$\int_{s_{l-1}}^{s_{l-1}+\varepsilon} \frac{\mathrm{d}x}{\mathrm{d}s}\, ds = x(s_{l-1} + \varepsilon) - x(s_{l-1}) = \int_{s_{l-1}}^{s_{l-1}+\varepsilon} p_x(s)\, ds = \mathcal{O}(\varepsilon) \tag{1.88}$$

instead it inserts a discontinuity in the momentum p_x:

$$\int_{s_{l-1}}^{s_{l-1}+\varepsilon} \frac{\mathrm{d}p_x}{\mathrm{d}s}\, ds = p_x(s_{l-1} + \varepsilon) - p_x(s_{l-1}) = \ell f_x(x, z; s_{l-1}) + \mathcal{O}(\varepsilon) \tag{1.89}$$

the same is valid for z, p_z. By applying the limit $\varepsilon \to 0$ we obtain what it's called the kick map:

$$\begin{pmatrix} x \\ p_x \\ z \\ p_z \end{pmatrix}_{s=s_{l-1}+0} = \begin{pmatrix} x \\ p_x + \ell f_x(x, z; s) \\ z \\ p_z + \ell f_z(x, z; s) \end{pmatrix}_{s=s_{l-1}} \tag{1.90}$$

Finally, the one-kick approximation is given by the composition of L with the kick map:

$$\mathbf{x}_{\text{one-kick}}(s_l) = \mathbf{L}(s_l, s_{l-1})(\mathbf{x}(s_{l-1}) + \ell\mathbf{f}(x, z; s_{l-1})) \tag{1.91}$$

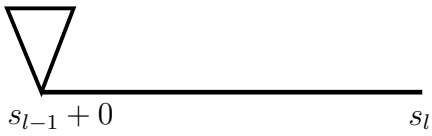


Figure 1.5: Simple sketch of the one-kick approximation.

1.5.3 Repeated kicks approximation.

It is possible to improve the accuracy arbitrarily by splitting the magnetic element into m equal portions and applying to each of them the single-kick approximation. In other words it is possible to develop a numerical scheme of arbitrary precision.

This leads to symplectic maps whose error can be estimated to be $\mathcal{O}(l^2/m)$. This analysis comes from the fact that these approximations are related to standard results in numerical analysis on the rectangular and trapezoidal rules of integration [14].

1.6 An introduction to Birkhoff normal form

We have seen so far how it is possible to analyse the betatronic motion of a particle by using the transfer map description of the magnetic lattice. In the linear case, the transfer matrix approach is equivalent to the Courant-Snyder theory: in the normal co-ordinates representation the transfer matrix is a pure rotation. However, when the non-linear effects due to the multipolar components of the magnetic field are taken into account, a different approach is needed to transform the system into a standard form, i.e. the so-called normal form.

If we perform a specific non-linear change of co-ordinates on a non-linear map, it is possible to reach a form that exhibits explicit symmetries depending on the absence or presence of resonance relations among the linear tunes.

The mathematical theory which gives us the tools for such analysis is given by Birkhoff normal forms. A complete review of its utilisation for non-linear betatronic motion is given in [11], in this Section we shall introduce the main theorems and results that define the Birkhoff normal form theory [15].

Let us start by considering in a phase space $\mathbb{R}^{2n}$ a smooth Hamiltonian system H having an elliptic equilibrium point at the origin. In the neighbourhood of the origin, we can have a canonical system of coordinates (p, q)

(possibly defined only in a neighbourhood of the origin) in which the Hamiltonian takes the form

$$H(p,q) = H_0(p,q) + H_1(p,q) \tag{1.92}$$

where

$$H_0(p,q) = \sum_{l=1}^{n} \omega_l \frac{p_l^2 + q_l^2}{2}, \qquad \omega_l \in \mathbb{R} \tag{1.93}$$

and H_1 is a smooth function having a zero of order 3 at the origin. The equations of motion for (1.92) take the form

$$\begin{aligned} \dot{p}_l &= -\omega_l q_l - \frac{\partial H_1}{\partial q_l} \\ \dot{q}_l &= -\omega_l p_l + \frac{\partial H_1}{\partial p_l} \end{aligned} \tag{1.94}$$

Since H_1 has a zero of order three, its gradient starts with quadratic terms. Thus, in the linear approximation, we obtain a system of n independent harmonic oscillators.

We now present the Birkhoff theorem for H and in the following we will denote by $x = (p,q)$ the whole set of variables.

Theorem 1. (Birkhoff). For any integer $N \geq 0$, there exist a neighbourhood $\mathcal{U}_N$ of the origin and a canonical transformation $\mathcal{T}_N : \mathbb{R}^{2n} \supset \mathcal{U}_N \to \mathbb{R}^{2n}$ which transforms the Hamiltonian system (1.92) in a Birkhoff Normal Form up to order N, i.e.

$$H^{(N)} \equiv H \circ \mathcal{T}_N = H_0 + Z^{(N)} + \mathcal{R}^{(N)} \tag{1.95}$$

where the $\circ$ denotes the combination of operators and $Z^{(N)}$ is a polynomial of degree $N+2$ which Poisson commutes with H_0, i.e. $\{H_0, Z^{(N)}\} = 0$, and $\mathcal{R}^{(N)}$ is a small remainder such that

$$\left|\mathcal{R}^{(N)}(x)\right| \leq C_N \|x\|^{N+3}, \qquad \forall x \in \mathcal{U}_N \tag{1.96}$$

moreover, one has

$$\|x - \mathcal{T}_N(x)\| \leq C_N \|x\|^2, \qquad \forall x \in \mathcal{U}_N \tag{1.97}$$

such inequality is fulfilled by the inverse transformation $\mathcal{T}_N^{-1}$.

If the frequencies are non-resonant, it can be shown that the function $Z^{(N)}$ depends only on the action variables

$$I_j \equiv \frac{p_j^2 + q_j^2}{2} \tag{1.98}$$

therefore, in the normalised system the actions becomes integrals of motion. In the resonant case, the normal form becomes more complicated and new specific phenomena, like exchange of energy among the oscillators that constitutes the normalised system (beatings).
Remark: The remainder $\mathcal{R}^{(N)}$ is very small in a small neighbourhood of the origin. In particular, it is of order ε^{N+3} in a ball of radius ε.

1.7 Particle tracking codes for beam dynamics studies

Particle-tracking codes are used to evaluate the long-term stability of accelerators and to simulate particle losses. Simulations can identify potential dangers for the accelerator's components, such as superconducting magnet quenching caused by an excessive loss of particles in the coils, and suggest machine settings to avoid them.

More specifically, single-particle tracking codes simulate the motion of a large number of non-interacting particles for many different initial conditions and seeds, for pseudo-random realizations of stochastic perturbation terms. Collective effects in a full beam of particles are in general considered using other types of codes.

We shall now present the most important particle-tracking codes used at CERN both for analysing existing rings and developing new ones:

- **SixTrack** [16] is a single-particle, six-dimensional symplectic tracking code used for studies of particle dynamics in circular accelerators. It has been used extensively for the LHC dynamic aperture studies and collimation studies. The core program is a single executable file made of about 70000 lines of Fortran code. SixTrack is based on the four-dimensional tracking code RACETRACK [17], and expands it by including the longitudinal degrees of freedom and a highly-refined post-processing analysis capabilities.

- **The Methodical Accelerator Design code (MAD-X)** [18] is mainly used for the design of accelerators, with particular emphasis on optimisation of lattice optics, but includes capabilities for particle tracking. This code is de facto the official software utilised at CERN, and even elsewhere, for designing new accelerator machines.

- **The Polymorphic Tracking Code (PTC)** [19], is an independent code that has been also embedded into MAD-X, thus extending greatly its capabilities. Note that PTC is a library of tracking routines and not a pre-built simulation tool like SixTrack or MAD-X. The high accuracy of the tracking models, based on high-order symplectic integrators, is unfortunately achieved at the cost of sizeable reduction of computing speed.

An adequate simulation of an accelerator ring like LHC, performed with SixTrack, can require entire days of computing on large clusters, as we will see in Chapter 3, this limitation in the computing times leads to some potential issues we want to address when discussing the Dynamic Aperture.

Chapter 2

The Hénon map

In this chapter we introduce the Hénon map [20], which is the simplest non-trivial symplectic polynomial map. Despite its simplicity, this map presents dynamical behaviours which are typical of most polynomial maps. Moreover, it is possible to show that it can be derived from basic, but important Hamiltonian models of non-linear betatronic motion. Because of that, this model is ideal for analysing the key features of non-linear dynamics and applications to non-linear one-turn maps.

In Section 2.1 we introduce the 2D Hénon map and show how it can be deduced [21] from the Hamiltonian of a 2D betatronic motion for a lattice made of periodic cells with a single sextupole non-linearity in the single-kick approximation. In Section 2.2 we present the main theorems that give us extremely important insights on the behaviour of this map (and give us the most important theoretical pillars for this work). In Section 2.3 we present briefly the 4D version of the Hénon map, among with some variations that allow the simulation of more complex accelerator behaviours and some insights about the dynamical theorems that are still valid from the 2D version of the map.

2.1 The 2D model

The Hénon map is the composition of a rotation of angle ω with an addition operation that adds the quantity $\hat{x}^2$ to $\hat{p}$, leaving $\hat{x}$ unchanged. In real

coordinates $(\hat{x}, \hat{p})$ it reads

$$\begin{pmatrix} \hat{x}' \\ \hat{p}' \end{pmatrix} = \mathbf{R}(\omega) \begin{pmatrix} \hat{x} \\ \hat{p} + \hat{x}^2 \end{pmatrix}; \qquad \mathbf{R}(\omega) = \begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix} \tag{2.1}$$

This is one of the simplest examples of non-integrable Hamiltonian map, and, with its basic 2D polynomial symplectic map, it can be used as a model of betatronic motion with a sextupole effect. In fact, the Hénon map can be obtained as a Poincaré map of a periodic Hamiltonian.

Let us consider the following Hamiltonian

$$H = \frac{p^2}{2} + \frac{x^2}{2\beta^2} - \frac{K_2 x^3}{6} \sum_{n=-\infty}^{+\infty} \delta(s - n\ell) \tag{2.2}$$

where s is the longitudinal coordinate $\in [0, s_L]$ and $\ell = s_L/L$. This Hamiltonian can describe the horizontal betatronic oscillations inside of a ring made with L identical FODO cells of length ℓ with a thin sextupole. β denotes the average of $\beta(s)$ on the cell and K_2 denotes the integrated sextupole gradient

$$\beta = \frac{1}{\ell} \int_0^{\ell} \beta(s)\, ds \qquad K_2 = \int_0^{l_S} k_2(s)\, ds \tag{2.3}$$

with l_S sextupole length. The equations of motion of the Hamiltonian (2.2) are

$$\begin{cases} \dot{x} = p \\ \dot{p} = -\frac{x}{\beta^2} + \frac{K_2}{2} x^2 \sum_{n=-\infty}^{+\infty} \delta(s - n\ell) \end{cases} \tag{2.4}$$

In order to integrate these and derive the Hénon map, we firstly observe that $x(s)$ is continuous while $p(s)$ is discontinuous at $s = n\ell$. By denoting the left and right limits of the discontinuity as $x_n^{\pm} = x(n\ell \pm 0)$ and $p_n^{\pm} = p(n\ell \pm 0)$, we integrate (2.4) in $n\ell - \varepsilon \le s \le n\ell + \varepsilon$ and, by letting $\varepsilon \to 0$, we obtain

$$p_n^+ - p_n^- = +\frac{K_2}{2} x_n^2, \qquad x_n^+ = x_n^- \equiv x_n \tag{2.5}$$

the one-turn map from (x_n, p_n^-) to (x_{n+1}, p_{n+1}^-) is then obtained by computing firstly x_n, p_n^+ using Eq. (2.4) and adding next the discontinuous term (2.5). This leads to

$$\begin{aligned} \begin{pmatrix} x_{n+1} \\ p_{n+1}^- \end{pmatrix} &= \begin{pmatrix} \cos\frac{\ell}{\beta} & \beta \sin\frac{\ell}{\beta} \\ -\frac{1}{\beta}\sin\frac{\ell}{\beta} & \cos\frac{\ell}{\beta} \end{pmatrix} \begin{pmatrix} x_n \\ p_n^+ \end{pmatrix} = \\ &= \begin{pmatrix} \sqrt{\beta} & 0 \\ 0 & \frac{1}{\sqrt{\beta}} \end{pmatrix} \mathbf{R}\left(\frac{\ell}{\beta}\right) \begin{pmatrix} \frac{1}{\sqrt{\beta}} & 0 \\ 0 & \sqrt{\beta} \end{pmatrix} \begin{pmatrix} x_n \\ p_n^- + \frac{K_2}{2} x_n^2 \end{pmatrix} \end{aligned} \tag{2.6}$$

Then, in order to reach the simplified form showed at the beginning, it's necessary to perform firstly the following change of coordinates

$$\hat{x} = \frac{x}{\sqrt{\beta}}, \qquad \hat{p} = p\sqrt{\beta} \tag{2.7}$$

next, since we are left with an Hénon map with a coefficient of the quadratic term equal to $K_2\beta^{3/2}/2$, we need to perform a not area-preserving scaling. We use then the scaled dimensionless coordinates

$$\begin{aligned} \hat{X} &= \frac{K_2\beta^{3/2}}{2}\hat{x} = \frac{K_2\beta}{2}x \\ \hat{P} &= \frac{K_2\beta^{3/2}}{2}\hat{p} = \frac{K_2\beta^2}{2}p \end{aligned} \tag{2.8}$$

which let us end up with the Hénon map

$$\begin{pmatrix} \hat{X}_{n+1} \\ \hat{P}_{n+1} \end{pmatrix} = \mathbf{R}\begin{pmatrix} \ell/\beta \end{pmatrix}\begin{pmatrix} \hat{X}_n \\ \hat{P}_n + \hat{X}_n^2 \end{pmatrix} \tag{2.9}$$

2.2 Important theorems

In this section we shall present the most important results on the 2D Hénon map in the framework of stability properties in the neighbourhood of the origin, i.e. the reference trajectory and a fixed point of the map.

The qualitative behaviour of the Hénon map is given, like any other discrete map $\mathbf{F}$, by its fixed points and cycles, i.e. all the $(\hat{x}, \hat{p})$ such that $\mathbf{F}^{\circ m}(\hat{x}, \hat{p}) = (\hat{x}, \hat{p})$, where m is the number of iterations.

Let us recall briefly the standard classification of fixed points of order 1, i.e. $m = 1$ for an area preserving map $\mathbf{F}$: once we found a fixed point $(\hat{x}_0, \hat{p}_0)$, we linearise $\mathbf{F}$ in the neighbourhood of that point and obtain

$$\begin{pmatrix} \hat{x}' \\ \hat{p}' \end{pmatrix} = \begin{pmatrix} \hat{x}_0 \\ \hat{p}_0 \end{pmatrix} + \mathbf{A}\begin{pmatrix} \hat{x} - \hat{x}_0 \\ \hat{p} - \hat{p}_0 \end{pmatrix} + \mathcal{O}((\hat{x} - \hat{x}_0)^2 + (\hat{p} - \hat{p}_0)^2) \tag{2.10}$$

where $\mathbf{A}$ is the Jacobian matrix of $\mathbf{F}$ at the fixed point. Since $\mathbf{F}$ is area-preserving by definition, we have that $\det(\mathbf{A}) = 1$ and we classify the fixed point with the following principle

$$\begin{cases} \textit{hyperbolic} & \text{if } |\mathrm{Tr}(\mathbf{A})| > 2 & \text{real eigenvalues } \lambda^{\pm 1} \\ \textit{elliptic} & \text{if } |\mathrm{Tr}(\mathbf{A})| < 2 & \text{complex eigenvalues } e^{\pm i\alpha},\ \alpha \in \mathbb{R} \\ \textit{parabolic} & \text{if } |\mathrm{Tr}(\mathbf{A})| = 2 & \text{equal eigenvalues } \lambda = 1 \end{cases} \tag{2.11}$$

The 2D Hénon map has 2 fixed points: an elliptical one at the origin $(0,0)$ and one hyperbolic (for $\omega \neq 0$) at $(2\tan\frac{\omega}{2}, -2\tan^2\frac{\omega}{2})$. For this work, we are interested in investigating the stable nature of the elliptic point of the map and, in particular, understand how the presence of the non-linear terms (which are neglected in the linearisation around the elliptic point) preserves or disrupts the structure of the orbits.

It is known in literature that, in the case of elliptic fixed points, the geometry of the linear orbits is not preserved in general and it is possible to have no more formal stability, i.e. it is possible to find unstable orbits for arbitrarily small amplitudes. However, it is possible to have scenarios in which only the resonant frequency orbits topology is broken.

A theorem that provides a deep insight on such a breaking mechanism is the Poincaré-Birkhoff theorem, which states that having a map whose phase space structure is an amplitude-dependent rotation, i.e. a twist map, like the Hénon map, all the invariant curves that have a resonant tune are broken into islands by arbitrarily small perturbations.

Theorem 2. (Poincaré-Birkhoff) [22]. Let F be the following twist map

$$\mathbf{F} : \begin{cases} \theta' = \theta + \rho \mod 2\pi \\ \rho' = \rho \end{cases} \tag{2.12}$$

where $\rho \in [0, R]$, $R > 0$ and $\theta \in [0, 2\pi]$. Let $\rho^* = 2\pi p/q$ be a circle of periodic orbits and let $\mathbf{F}_\varepsilon$ be a perturbed map

$$\mathbf{F}_\varepsilon : \begin{cases} \theta' = \theta + \rho + \varepsilon f(\theta, \rho) \mod 2\pi \\ \rho' = \rho + \varepsilon g(\theta, \rho) \end{cases} \tag{2.13}$$

with f, g regular on their domains. Then, there are two circles on which $\mathbf{F}_\varepsilon^{\circ q}$, i.e. the map $\mathbf{F}_\varepsilon$ iterated q times, induces rotations in the opposite directions and a curve Γ between them on which $\mathbf{F}_\varepsilon^{\circ q}$ induces radial displacements. The intersections of Γ and $\mathbf{F}_\varepsilon^{\circ q}\Gamma$ are fixed points: half of them elliptic and half hyperbolic. Therefore the map $\mathbf{F}_\varepsilon$ has kq elliptic fixed points and kq hyperbolic fixed points, with $k \in \mathbb{N}$ and $k \geq 1$.

In Fig. 2.1 the appearance of a chain of 5 islands for the Hénon map close to resonance 5 is shown. From this figure it seems clear that, at least from a numerical point of view, most of the orbits of the map are only deformed circles, i.e. the topology of such orbits is preserved under perturbation.

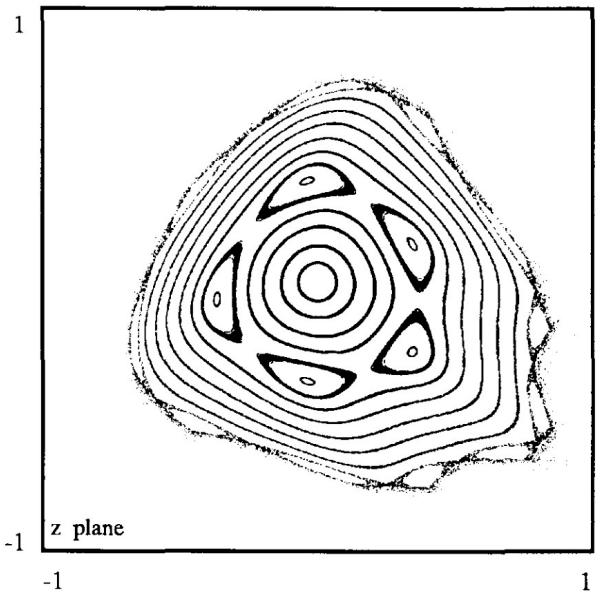


Figure 2.1: Chain of 5 islands in the Hénon map for $\omega/2\pi = 0.205$. (Source: [21])

This then does not exclude the possibility that orbits with strongly irrational frequency could just be deformed by the perturbation, without having their topology completely broken. A precise statement in such matter is provided by Kolmogorov–Arnold–Mooser (KAM) theory, which is the theory of perturbations for conditionally periodic motions of Hamiltonians, when considering infinite times. In particular, KAM theory gives rigorous results about the topological stability of perturbed orbits and extends the results one could obtain in an heuristic approach using the averaging principle (which will be presented in Section 4.4). Under some assumptions, KAM theory ensures the existence of invariant curves whose frequency is diophantine.

An angular frequency $\omega = (\omega_1, \omega_2, \ldots, \omega_d) \in \mathbb{R}^d$ is called **diophantine** if it satisfies the estimate:

$$\left|e^{i\mathbf{q}\cdot\omega} - 1\right|^{-1} \leq \gamma|\mathbf{q}|^{\eta} \tag{2.14}$$

where $\mathbf{q}$ is a vector of integers and γ and η are positive constants. Such diophantine frequencies are dense in $\mathbb{R}^d$ and, for $d = 1$, have a Lebesgue measure of 1 in the interval $[0, 1]$, i.e. the probability of finding a non-diophantine

number is zero, as for rational numbers.

Theorem 3. (Kolmogorov–Arnold–Moser) [23, 24, 25]. Let $\mathbf{F}_\varepsilon$ be the following perturbed area-preserving twist map:

$$\mathbf{F}_\varepsilon : \begin{cases} \theta' = \theta + \rho + \varepsilon f(\theta, \rho) \\ \rho' = \rho + \varepsilon g(\theta, \rho) \end{cases} \tag{2.15}$$

with f, g analytic in $D : |\rho| \leq \rho_0, |\mathrm{Im}(\theta)| \leq \theta_0$, let also $\omega/2\pi$ be a diophantine frequency. Then, the circle $\rho = \omega$ of the unperturbed twist map, i.e. $\varepsilon = 0$ is mapped into a closed orbit of the perturbed system (2.15) for ε sufficiently small. This theorem is also valid for symplectic maps in $\mathbb{R}^{2d}$ for $d \geq 2$.

Since in a 2D phase space any region whose boundary is an invariant curve is invariant and the curve itself is a topological barrier to the motion, i.e. a curve makes a perfect confinement, the existence of invariant curves given by this theorem is sufficient to prove the stability of a fixed point, provided that the KAM theory can be applied.

In higher dimensional systems, the stability problem is not solved by KAM theory since the invariant tori do not separate the phase space. However if one is interested in orbit stability for large, but finite times (and as we will see in Chapter 3, this is our case), there is an extremely important theorem for symplectic maps [26, 27] which generalises the well-known theorem formulated by Nekhoroshev [8] for Hamiltonian flows.

Theorem 4. (Nekhoroshev). Let $F(x)$ be a symplectic map in a phase space of dimension $2d, (d \geq 2)$, analytic in a polydisc of unit radius, having the origin as an elliptic fixed point. Assuming that the frequency vector $\omega \in \mathbb{R}^d$ is diophantine as defined in (2.14), then any orbit with initial point in a polydisc of radius $\rho/2$ will remain in a polydisc of radius ρ for a time $t \leq T$, where

$$T = T_* \exp\left[\left(\frac{\rho_*}{\rho}\right)^A\right] \tag{2.16}$$

provided that $\rho \leq \rho_*$, with the constants A, T_*, ρ_* dependent on the diophantine constants η, γ.

2.3 The 4D Hénon map

The 4D Hénon map can be obtained as a Poincaré section of a linear magnetic lattice with a sextupole non-linearity included in a single-kick approximation.

It is usually presented in the following form:

$$\begin{pmatrix} x' \\ p'_x \\ z' \\ p'_z \end{pmatrix} = \mathbf{L} \begin{pmatrix} x \\ p_x + \frac{K_2}{2}(x^2 - z^2) \\ z \\ p_z - K_2 x z \end{pmatrix} \tag{2.17}$$

where K_2 is a numerical value which represents the sextupolar strength in the system and $\mathbf{L}$ is the linear part of the transfer map of the lattice, which, under Courant-Snyder coordinates and an adequate re-scaling of coordinates, can be written as a direct product of two two-dimensional rotations

$$\mathbf{L} = \begin{pmatrix} R(\omega_x) & 0 \\ 0 & R(\omega_z) \end{pmatrix} \tag{2.18}$$

Despite being so simple, this map contains most of the physical features of a non-linear magnetic lattice [11]. Because of that, it is extremely useful to perform fast and straightforward computer tracking simulation for testing new models for beam dynamics, even for long times.

Moreover, the 4D Hénon map, as a quadratic map, has the important characteristic of having an attracting point at infinite amplitude. That makes it a proper tool for defining correctly the Dynamic Aperture problem (which will be presented in Chapter 3).

It is also possible to formulate different variants of the 4D Hénon map to include more complex and even time-dependent effects, like tune modulation effects [28], which have a strong impact on the beam and drastically enhance long-term particle losses. These tune modulations are caused in a real accelerator by unavoidable external effects such as power supply ripple, or synchro-betatron coupling, via the residual uncompensated chromaticity, and can be modelled by a set of non-linear oscillations whose linear frequencies are modulated. The resulting model is called Modulated Hénon map:

$$\begin{pmatrix} x^{(n+1)} \\ p_x^{(n+1)} \\ z^{(n+1)} \\ p_z^{(n+1)} \end{pmatrix} = \mathbf{R} \begin{pmatrix} x^{(n)} \\ p_x^{(n)} + ([x^{(n)}]^2 - [z^{(n)}]^2) \\ z^{(n)} \\ p_z^{(n)} - 2x^{(n)} z^{(n)} \end{pmatrix} \tag{2.19}$$

$$\mathbf{R} = \begin{pmatrix} R(\omega_x^{(n)}) & 0 \\ 0 & R(\omega_y^{(n)}) \end{pmatrix} \tag{2.20}$$

where the linear frequencies $\omega_x^{(n)}, \omega_z^{(n)}$ are slowly varying with the discrete time n following the equations

$$\omega_x^{(n)} = \omega_{x0}\left(1 + \varepsilon \sum_{k=1}^{m} \varepsilon_k \cos(\Omega_k n)\right) \tag{2.21}$$

$$\omega_z^{(n)} = \omega_{z0}\left(1 + \varepsilon \sum_{k=1}^{m} \varepsilon_k \cos(\Omega_k n)\right) \tag{2.22}$$

Such modulations can be implemented so that the values ε_k, Ω_k can mimic very well the known characteristics observed in existing accelerators (in [28] are presented values that mimic the tune ripple observed in the quadrupole of SPS).

Obtaining knowledge on the true topology of a 4D map and its orbits is in general extremely difficult. Making 2D projections of the orbits gives only poor information and it constitutes the main difficulty in the full comprehension of the dynamics of the map.

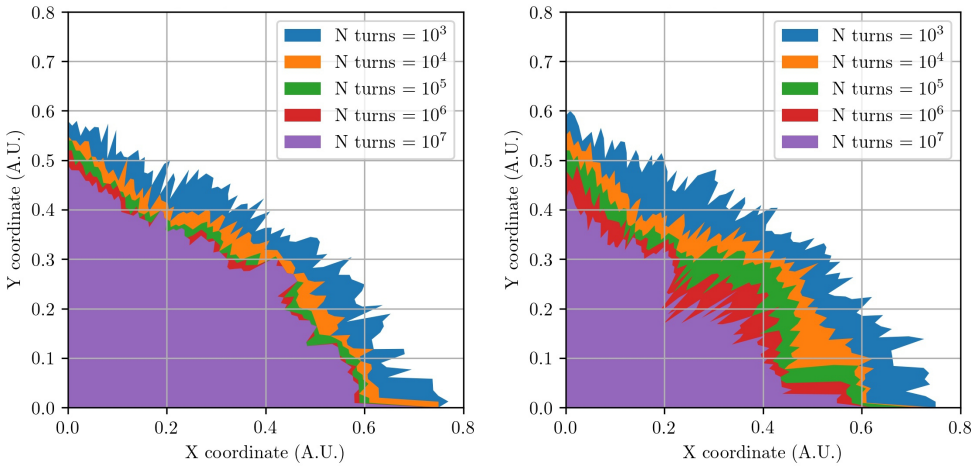
The only important established results are the theorems valid for generic $2n$ dimensional systems: the KAM Theorem 3 and the Nekhoroshev estimate Th. 4. It must be said, however, that the KAM theorem for 4D mappings implies the existence of invariant 2D tori which are **not** topological barriers like the KAM tori in the 2D mappings. Arnold in fact shows how in a region 'filled' with KAM tori there are still some initial conditions for which an amplitude variation is present, no matter how small we choose the neighbourhood of the origin. This phenomenon is called **Arnold diffusion** and, since the measure of these initial conditions is extremely small and the divergence times extremely long, its actual importance in physical scenarios of accelerator physics is still debated.

However, in phase space regions in which KAM theory is not dominant, i.e. the measure of KAM tori is small, it is possible to observe the creation of weakly chaotic regions in which diffusive-like phenomena can be observed. In these regions, the estimates provided by the Nekhoroshev theorem can still be used.

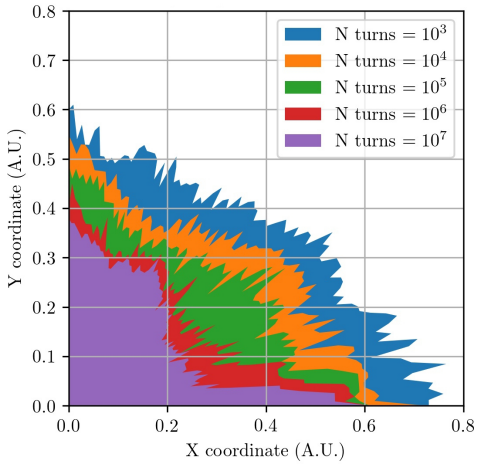
In general, an exhaustive analysis of the dynamics of such 4D maps is far from being reached, but it constitutes the starting point for formulating new valid models for inspecting topics such as the Dynamic Aperture model.

For example, in Ref. [29], two fit models for Dynamic Aperture are tested on the modulated Hénon map, on which Nekhoroshev-like behaviours are

observed. In Fig. 2.2 it is possible to observe the time evolution of the stable region for the modulated Hénon map for different values of ε (a proper review of the Dynamic Aperture concept is given in the next chapter).



(a) $\varepsilon = 0$ (b) $\varepsilon = 32$



(c) $\varepsilon = 64$

Figure 2.2: Time evolution of the stability region for the modulated Hénon map for different ε values ((X, Y) corresponds to (x, z) in our notation). (Source: [29])

Chapter 3

Dynamic Aperture

The Dynamic Aperture (DA) is defined as the amplitude of the phase-space region of a circular accelerator where stable motion occurs for a time scale equal to the operational cycle time of the accelerator under consideration (in LHC the typical time scale for experiments data-taking is ~ 10 hours, after that the luminosity loss requires the execution of a new fill). The DA is one of the key quantities that are considered for the design of modern colliders based on superconducting magnets, such as Tevatron [30], HERA [31], RHIC [32], SSC [33], and LHC [34]; and it is also a key figure of merit for the design of the HL-LHC project [6] as well for future collider studies.

In this chapter we shall present the main aspects of this quantity and review the most important models of its functional dependence from the relevant parameters, which constitutes the starting point of this work. In Section 3.1 we present and comment the DA definition. In Section 3.2 we briefly present the DA interpolation problem and the most important studies about it. Finally, in Section 3.3, we illustrate the experimental method utilised for performing experimental measurements of DA in a circular accelerator such as LHC.

3.1 Dynamic Aperture definition

A complete discussion on the DA definition, its computation, and its accuracy can be found in Refs. [1, 28].

In the application to particle accelerators it is convenient to introduce the concept of DA in these terms:

Definition 1. (Dynamic Aperture). Let A be the physical aperture of the accelerator, i.e. the subset of the phase space that can be confined in the beam pipe, and let $\mathbf{M}$ the one-turn map of the magnetic lattice. We define the formal DA as

$$\mathcal{D}(N) = \bigcap_{n=0}^{N} \mathbf{M}^{(n)}(A) \tag{3.1}$$

We recall that $\mathbf{M}$ has an elliptic fixed point at the origin.
Let $z \in \mathcal{D}(N)$, we define $\Pi_X z = x$ as the projection of z on the configuration space; the set

$$\Pi_X \mathcal{D}(N) \tag{3.2}$$

may have a very complex topology (it is a measurable set) so that it is convenient to compute the convex envelope of the connected component containing the origin, i.e. we neglect the islands of stability.
To define the measure of such a component, we proceed as follow: for each direction $\hat{\theta} \in S^d$ (where S^d is the unit sphere or a sector of the unit sphere) in the configuration space, we define

$$R(\hat{\theta}; N) = \lambda_* \quad s.t. \quad \lambda\hat{\theta} \in \Pi_X \mathcal{D}(N) \quad \forall \quad \lambda \in [0, \lambda_*] \tag{3.3}$$

so that we can finally define the Dynamic Aperture as

$$D(N) = \frac{1}{\mu(S^d)} \int_{S^d} R(\hat{\theta}; N) d\hat{\theta} \tag{3.4}$$

where μ is the volume measure in the configuration space.

The value of N needs to be adapted for a proper time frame. In a mathematical sense, stable motion implies bounded motion for $N \to \infty$. In our accelerator context, stable motion and particle stability can be linked to a maximum number of turns $N_{\max}$, where the value of $N_{\max}$ is set on the basis of the specific device or application under consideration, e.g. in LHC the revolution frequency is 11.245 kHz [35], which implies $\sim 10^9$ turns for a standard 10-hour luminosity fill.

A valid numerical declination of Def. 1 for four-dimensional symplectic mappings which model betatron motion is provided in Ref. [1]. More specifically, it presents a valid method for fast numerical estimation of DA, as well as estimation of the associated errors and optimisations of the integration steps.

Let us operate on a 4D phase space on which we have defined a one-turn map $\mathbf{M}$ for the betatron motion with the formalism presented in Chapter 1. If we consider an ensemble of initial conditions defined on a polar grid ($x = r\cos\theta, p_x = 0, y = r\sin\theta, p_y = 0$), $0 \leq \theta \leq \pi/2$, where x, y are expressed in units σ_x, σ_y of beam dimension, and we track them for up to $N_{\max}$ turns to assess their stability, then we can define the DA as:

$$D(N) = \frac{2}{\pi}\int_0^{\pi/2} r(\theta; N)\, d\theta \equiv \langle r(\theta; N)\rangle \tag{3.5}$$

where $r(\theta; N)$ is the last stable amplitude, i.e. $x^2 + y^2 < r_{\max}$ for every iteration of $\mathbf{M}$, not disconnected from the origin for up to N turns in the direction θ. We can say that r is the computable version of the 'ideal' $R(\theta; N)$ given in Eq. (3.3).

In order to evaluate the error on the DA numerical estimation (3.5), we just need to consider that its implementation into a computer code implies a discretization over the radial variable r and one over the angular variable θ [28]. Assuming we execute a complete tracking over N_θ different angles and N_r different amplitudes for initial condition, we have $\Delta r = r_{\max}/N_r$ and $\Delta\theta = \pi/(2N_\theta)$. Then, we can obtain an error estimation via Gaussian sum in quadrature

$$\Delta D = \sqrt{\left(\frac{\partial D}{\partial r}\frac{\Delta r}{2}\right)^2 + \left(\frac{\partial D}{\partial \theta}\frac{\Delta \theta}{2}\right)^2} \tag{3.6}$$

which leads us to

$$\Delta D = \sqrt{\frac{(\Delta r)^2}{4} + \left\langle\left|\frac{\partial r}{\partial \theta}\right|\right\rangle^2 \frac{(\Delta\theta)^2}{4}} \tag{3.7}$$

where $\langle|\partial r/\partial\theta|\rangle$ is the average of the finite differential measures obtained in the simulation process. From this last equation we can deduce that the step in r must be equal to the step in θ times $\langle|\partial r/\partial\theta|\rangle$ to optimize the numerical integration steps.

3.2 Functional dependence

Simulating entire sets of initial conditions on different one-turn map setups is a CPU-intense task that becomes unsustainable when considering extremely

high $N_{\max}$ values or complex symplectic tracking models[1]. Moreover, the multipolar components of the various superconducting magnets are known only with a limited precision, so that one has to perform parametric studies in order to consider different realisations of the magnetic lattice. While it might not be a problem to execute scalable parallel batches of different realisations and initial conditions [36], realistic time scales in tracking simulation are still out of reach for proper accelerator physics reaserch.

Because of that, a robust model for the time dependence of DA would be essential for speeding up the reaserch and development of better machines. Therefore, it is in our interest to explore and build models to fit and, ideally, extrapolate the dependency of the DA on the number of turns and allow us the execution of shorter simulation for the same amount of information.

The main idea behind this framework is that long-term behaviour of DA, which is a heavy task to attack directly, can be extrapolated from the knowledge gained from numerical simulations performed over a smaller number of turns, but with a detailed scan of the phase space. Moreover, since it is possible to execute parallel computing over different initial conditions, it is possible to work with high fineness numerical estimation of DA and hope for 'high quality' coefficients for such interpolations.

Studies have explored possible models for this fit [37, 28] and an answer was provided by a combination of the fundamental results of KAM theory and Nekhoroshev theorem. More specifically, a model was based on the hypothesis that the phase space can be partitioned into two regions:

1. a central core, with $r < D_\infty$, where the phase space is full of KAM tori so that the Arnold diffusion phenomenon takes place for a set of orbits of extremely small measure (to the point that the physical value of the phenomenon itself is still debated);

2. an outer part, with $r > D_\infty$, where the stability and escape rate can be estimated with a Nekhoroshev-like estimate

$$N(r) = N_0 \exp\left[\left(\frac{r_*}{r}\right)^{1/\kappa}\right] \tag{3.8}$$

where $N(r)$ is the number of turns that are estimated to be stable for a particle with initial amplitude smaller than r.

[1]Researches like [28] presents simulations at $N_{\max} \sim 10^6 - 10^7$, while instead it would be necessary to reach values $\sim 10^9$.

From this separation, the following scaling law is formulated:

$$D(N) = D_\infty + \frac{b}{(\log N)^\kappa} \tag{3.9}$$

where D_∞ represents the asymptotic value of the stability domain (region 1) amplitude and b, κ are additional parameters connected to the Nekhoroshev-like process (region 2).

This model has provided good results in literature, however, experiences with the data analysis of numerical simulation of various configuration of LHC [28] and of experimental data obtained from Tevatron [3] showed that the fit parameters b, κ, D_∞ can assume signs that go beyond the boundaries predicted by the strict application of the model based on Nekhoroshev theorem.

Another thing we are interested in is the establishment of a direct link between the DA and the expected beam lifetime in a synchrotron, in order to transform a DA interpolating law into an expected beam-quality model. The approach proposed in [3] considers an initial 2D Gaussian distribution for a beam

$$\rho_G(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)} \tag{3.10}$$

then, transforming (3.10) to polar coordinates and applying the DA definition (3.5), i.e. assuming that all particles with starting amplitude beyond $D(N)$ are lost after N turns, we can directly compute the evolution of beam intensity N_b with the following equation

$$\frac{N_b(N)}{N_b(1)} = 1 - \int_{D(N)}^{+\infty} e^{-\frac{r^2}{2}} r\, dr = 1 - e^{-\frac{D^2(N)}{2}} \tag{3.11}$$

with $D(N) \xrightarrow[N\to 0]{} +\infty$. This last equation represents a starting point for building the direct connection we are looking for, and also allows us to establish some experimental procedures for evaluating DA from beam losses in a circular accelerator.

3.3 Experimental measures of Dynamic Aperture

Measuring DA in a synchrotron, as well as measuring any kind of non-linear observable, is a true challenge and an important goal that can lead to the

proper examination and understanding of non-linear single-particle motion.

Most of the last results in terms of DA measurements are obtained on the LHC, where many experimental observations were performed with a vast plethora of different configurations and combined with many intensive numerical simulations, performed since the design phase [4, 38, 39, 5].

The first reference makes use of a 'standard-kick' method of beam analysis, based on inflicting a single strong kick to the beam, in order to measure at what precise amplitude fast losses are observed, i.e. by deflecting the beam we measure beam losses as a function of the displacement itself. This method proved to be very useful once for the first measures of DA when the LHC became operational, however, it has two major drawbacks:

1. This method requires deflecting the beam by an amount that is a sizeable fraction of the DA, this prevents its application at top energy in the LHC due to the enormous magnetic strength that would be required. In fact, measurements based on this technique were all performed at injection energy.

2. Non-negligible, fast beam losses are intrinsic to this method, as the risk of beam-induced quenches makes this feature not particularly suitable for a superconducting machine, especially for the extremely high energies we are interested for LHC.

The other three references, instead, makes use of a more advanced kicked-beam method where, instead of a single high amplitude kick, many small kick are used to deflect the beam. Such method and its main achievements are reviewed in detail in Ref. [5], here we present briefly the main idea behind it.

Starting from Eq. (3.11), one possible method for measuring DA is to measure beam loss following a multitude sequence of short amplitude kicks (see Fig. 3.1). Such kicks results necessary in the analysis, since, in the case of LHC, there are no manifested lifetime issues at injection time, suggesting that the DA is comparable to the mechanical aperture of the machine [38]. Therefore, these amplitude kicks acts as artificial reducer of beam lifetime and are necessary to obtain proper measures in adequate operational times.

A kick, which can induced with some of the non-linear correctors inside the circular collider, shifts the beam to larger amplitudes in the phase space, resulting in more particles passing beyond the DA and thus becoming lost. The distance of the excited beam from the dynamic aperture can then be determined from the measured beam losses, following the kick, while the

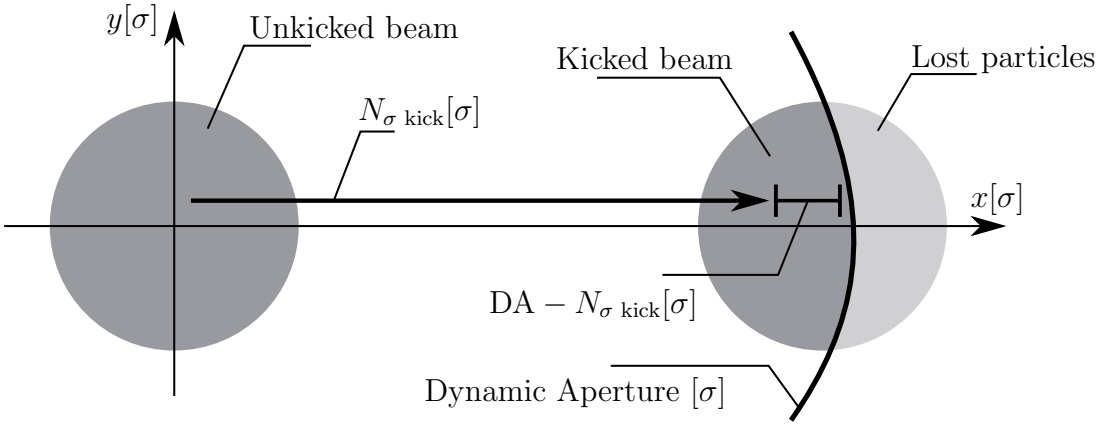


Figure 3.1: Illustration of the advanced kicked-beam method used to measure the Dynamic Aperture. The value $(\mathrm{DA} - N_{\sigma\ \mathrm{kick}})$, which we want to measure, can be determined from the beam loss over the many turns and kicks.

beam amplitude itself can be determined from the turn-by-turn beam position monitors (BPM) of the machine.

Adapting a little Eq. (3.11), in order to take into account the amplitude kicks, it is possible to compute the DA starting from the beam loss with [4]:

$$\frac{\Delta I}{I} = \frac{I}{\sqrt{2\pi}} \int_{\mathrm{DA}-N_{\sigma\ \mathrm{kick}}}^{+\infty} e^{-\frac{1}{2}N'^2}\, dN' \tag{3.12}$$

$$\frac{\mathrm{DA} - N_{\sigma\ \mathrm{kick}}}{\sqrt{2}} = \mathrm{erf}^{-1}\left[1 - 2\frac{\Delta I}{I}\right] \tag{3.13}$$

The advanced-kick method, performed in the last measurement sessions at LHC, can be summarised in this way:

- Prepare the machine in an operational-like configuration, with either the sextupole and octupole correctors turned off or setup to correct the known non-linearities of the machine. Also, with all the collimators set in parking position, with only the horizontal and vertical primary devices set to $12\,\sigma$ (this is meant to shadow the effects of the known machine aperture, measured to be larger than about $13\,\sigma$).

- Inject a pilot bunch with operational normalised emittance, i.e. $\approx 2\,\mu\mathrm{m}$ and $1.4 - 1.7 \times 10^{10}\mathrm{p}$.

- Apply repeated small kicks to blow-up the beam size until losses are observed at the predisposed primary collimators. This could be obtained with the transverse damper (ADT) of LHC and had the advantage of exciting the beam with white-noise, creating a well-reproducible Gaussian shape of the transverse beam distribution after the blow-up, thus obtaining the shape required for a proper utilisation of Eq. (3.13).

- Record beam losses and bunch intensities and analyse them off-line for studying the time-dependencies therein.

All the obtained measures showed good agreement with the utilised model on both the vertical and horizontal plane and the measured dynamic aperture agreed well within the safety margin specified in LHC design phase.

Moreover, it was possible to compare this non-linear observable with simulated values [4] obtained by particles tracked with advanced simulation tools, which aims to take into consideration all the known imperfections and non-linearities of the machinery. In particular, simulations with the PTC module within MAD-X and the symplectic integrator SixTrack were performed.

The comparison between the various setups of the machine (with or without non-linear correction) were in good agreement with the simulated values, and showed how the non-linear corrections significantly improved the value of DA. This demonstration of Dynamic Aperture correction in the LHC represents an important result that encourages deeper analysis on this topic.

Chapter 4

Stochastic Hamiltonians and Fokker-Planck equation

In this chapter we shall review the mathematical tools and physical methods [14, 40, 41] we will use in order to deal with stochastic Hamiltonian systems and, more in general, perturbed Hamiltonian systems.

In Section 4.1 we review the fundamental theoretical setup necessary when utilizing stochastically-perturbed Hamiltonian system, and we introduce the stochastic Liouville equation for expressing the evolution of the average dynamics. In Section 4.2 we derive the Fokker-Planck equation for describing such average dynamics, starting from a stochastic Liouville equation. In Section 4.3 we discuss the fast angle relaxation process on a toroidal surface, which is an important issue in order to properly execute averaging procedures for perturbed dynamical systems. Finally, in Section 4.4 we review the basic concepts of the Averaging Principle, and we show how it can be used for stochastically-perturbed dynamical systems, in order to reach a Fokker-Planck form for describing the average dynamics.

4.1 Stochastic Hamiltonian systems

The whole theory of stochastic Hamiltonian systems aims to solve the main problems that come up with the study of stochastic differential equations. Such equations cannot be directly tackled by taking into account the symplectic nature of the solutions, due to the non covariant characters of the Wiener process, which simulates the stochastic fluctuations.

While stochastic differential equations cannot be used as fundamental laws of physics, they can be used as fundamental tools for 'mesoscopic' approaches for physical problems in Stochastic Thermodynamics, i.e. problems in non-equilibrium Statistical Mechanics.

The irreversible character of the stochastic dynamics allows to introduce the time arrow and the entropy concept in a natural way, while instead makes the principles of classical mechanics for isolated systems incompatible.

In this chapter, we are interested in methods to reconcile the Hamiltonian nature of the Classical Mechanics, which is strictly related to the canonical character of the evolution equations, with the non-differentiable nature of the Wiener process, which describes the continuous innovative character of the stochastic effects.

Our goal is to obtain proper evolution laws for average macroscopic observables of systems that are continuously affected by unpredictable small perturbations from the external environment. It is important to remark that these macroscopic observables concerns average quantities (both in space and time) of extensive or intensive nature, whose evolution does not necessarily preserve the properties of Hamiltonian Mechanics. In order to construct such formulation, it is essential to require further hypothesis on the microscopic dynamics of the elementary elements of the statistical system, i.e. particles, in particular, we recall that:

- the microscopic dynamics should have a strong chaotic character, i.e. strong dependence from the initial conditions and/or the environmental conditions;
- the system can be decomposed into microscopic almost-independent subsystems (elementary components).

When both hypothesis are verified, the stochastic equations can be used as a mathematical model for Statistical Mechanics of physical systems and it is possible to include some principles of Classical Mechanics.

In a simplified approach we can assume that the microscopic dynamics of the single particles can be expressed as the superposition of a deterministic Hamiltonian, dependent only on the particle dynamical state, and a stochastic term, which simulates the interaction with the other particles and the environment, whose amplitude depends only on the particle dynamical state. This leads us to the following stochastic Hamiltonian system:

$$H(x,t) = H_0(x) + H_1(x)\xi(t) \tag{4.1}$$

where $\xi(t)$ is a regular stationary stochastic Gaussian process, i.e. we assume that the realisations $\xi(t)$ are continuous, with a correlation function

$$\langle \xi(t)\xi(t+\tau)\rangle = \sigma^2\Phi(\gamma\tau) \qquad \Phi(\gamma\tau) \simeq e^{-\gamma|\tau|} \tag{4.2}$$

where γ^{-1} is the correlation time scale, which defines the noise evolution with respect to the characteristic time scale of the unperturbed dynamics H_0. In this formulation, the white-noise limit is given by $\gamma, \sigma \to +\infty, \sigma^2/\gamma \to 1/2$, so that

$$\sigma^2\Phi(\gamma\tau) \to \delta(\tau) \tag{4.3}$$

When we consider Eq. (4.1) for describing an ensemble of independent particles, any particle feels a different realization of the noise that mimics the effect of the interaction with the environment, including possible interactions with other particles. For a given realization of the stochastic process $\xi(t)$, the evolution follows a symplectic dynamics for any initial condition.

For any continuous realization of $\xi(t)$, we can define the stochastic phase flow $x(t, x_0|\xi) = \mathbf{\Phi}^t_\xi(x_0)$ based on the Hamiltonian (4.1) and on the symplectic maps

$$\left(\frac{\partial \mathbf{\Phi}^t_\xi}{\partial x}\right)^T J \left(\frac{\partial \mathbf{\Phi}^t_\xi}{\partial x}\right) = J \tag{4.4}$$

where J is the usual symplectic matrix related to the canonical form of the Hamilton equations

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{4.5}$$

When we consider the limit of δ-correlated noise, i.e. white-noise, the stochastic phase flow can be associated to a stochastic differential equation that has to be derived from Eq. (4.1) and allows the symplectic character of the dynamics to be preserved in a probabilistic sense. We have the following symplectic stochastic phase flow

$$x(t+\Delta t) = \exp\left(\Delta t D_{H_0} + \Delta\xi D_{H_1}\right) x(t) \tag{4.6}$$

where the operators D_H are the Lie-derivative operators defined by the Poisson brackets

$$D_H = \frac{\partial H}{\partial x} J \frac{\partial}{\partial x} \tag{4.7}$$

Remark: $\Delta\xi$ is the increment of the stochastic process for a time Δt; for a differentiable noise we have $\Delta\xi \propto \Delta t$, but this is not true in the white-noise

limit $\Delta\xi \to \Delta w_t$, where w_t is a Wiener process, this limit is not a pointwise limit and it has to be considered as a weak-convergence in a probabilistic sense.

The explicit expansion of Eq. (4.6) reads

$$x(t+\Delta t) = x(t) + \Delta t D_{H_0} x(t) + \Delta\xi D_{H_1} x(t) + \frac{\Delta\xi^2}{2} D^2_{H_1} x(t) + O(\Delta t^2, \Delta\xi^3) \tag{4.8}$$

By construction, this scheme satisfies a symplectic condition with an error $\mathcal{O}(\Delta t^{3/2})$ (we recall the fact that $\Delta w_t \equiv \mathcal{O}(\Delta t^{1/2})$). In the limit $\Delta t \to 0$ we obtain the stochastic differential equation

$$dx = D_{H_0} x\, dt \,+\, D_{H_1} x\, dw_t \,+\, \frac{1}{2} D^2_{H_1} x\, dt \tag{4.9}$$

since Δw_t^2 is stochastically equivalent to Δt.

In order to relate Eq. (4.9) to a canonical form like Eq. (4.1), we can define the Stratonovich stochastic integral of a function $\sigma_k(x)$

$$\int \sigma_k(x) \circ dw_t = \int \sigma_k(x)\, dw_t + \frac{1}{2} \int \sigma_l(x) \frac{\partial \sigma_k}{\partial x_l}(x)\, dt \tag{4.10}$$

so that if we set

$$\sigma_k(x) = \frac{\partial H_1}{\partial x_l} J_{lk} \tag{4.11}$$

we can write (4.9) in the canonical form

$$dx = D_{H_0} x\, dt + D_{H_1} x \circ dw_t \tag{4.12}$$

In order to properly understand the white-noise limit in the context of stochastically perturbed Hamiltonian mechanics, we proceed to compute the evolution of the distribution function according to the **stochastic Liouville equation**.

We require a stationary condition on the system, meaning that for any realization $\xi(t)$ and any time t_0, there exists a realization $\xi'(t)$ such that

$$\mathbf{\Phi}^{t_0+t}_{\xi(t_0)} = \mathbf{\Phi}^{t}_{\xi'(0)} \tag{4.13}$$

in other words, the evolution is homogeneous in time. This stationary condition implies that the statistical properties of the evolution are invariant with

respect to the choice of the initial time, and that the decaying of the correlation can be justified if the future evolution turns out to be independent from the past as $t \to \infty$.

Let us now define the distribution function $\rho(x,t)$ as the probability to detect a particle at x after a time t given its initial condition. For a fixed regular realization $\xi(t)$, the evolution of the distribution function $\rho_\xi(x,t)$ is the solution of the stochastic Liouville equation

$$\frac{\partial \rho_\xi}{\partial t} = -\left(\frac{\partial H_0}{\partial x} + \xi(t)\frac{\partial H_1}{\partial x}\right) J \frac{\partial \rho_\xi}{\partial x} \tag{4.14}$$

The validity of the stochastic Liouville equation is strictly related to the regularity of the stochastic signal $\xi(t)$. More precisely, if one considers the system

$$\dot{x} = \sigma(x)\xi(x) \tag{4.15}$$

the stochastic phase flow reads

$$x(\Delta t) = x + \sigma(x)\int^{\Delta t} \xi(s)\, ds + \frac{\partial \sigma}{\partial x}\sigma(x)\int^{\Delta t}\int^{s} \xi(s')\xi(u)\, du\, ds' + \dots \tag{4.16}$$

and we therefore require the following limit

$$\lim_{\Delta t \to 0}\left(\frac{1}{\Delta t}\int^{\Delta t}\int^{s} \xi(s')\xi(u)\, du\, ds'\right) = 0 \tag{4.17}$$

If we assume the following correlation function

$$\langle \xi(s)\xi(u)\rangle = \gamma \exp(-\gamma|s-u|) \tag{4.18}$$

we obtain

$$\begin{aligned}\left\langle \int^{\Delta t}\int^{s} \xi(s)\xi(u)\, du\, ds \right\rangle &= \int^{\Delta t} [\exp(-\gamma s) - 1]\, ds \\ &= \frac{1}{\gamma}[1 - \exp(-\gamma \Delta t)] - \Delta t\end{aligned} \tag{4.19}$$

and the required limit holds in average if

$$\lim_{\Delta t \to 0} \frac{\gamma \Delta t}{2} = 0 \tag{4.20}$$

This means that we cannot perform the white-noise limit $\gamma \to \infty$ in the derivation of the stochastic Liouville equation. From a physical point of view, if the correlation time scale γ^{-1} is of the same order of the evolution time scale Δt, then the stochastic Liouville equation is not justified. On the contrary, if $\gamma \Delta t \ll 1$ then the fluctuation $\xi(t)$ can be considered a regular function and Eq. (4.14) is valid.

The distribution function $\rho_\xi(x,t)$ provides all the information for a statistical approach to the dynamics of the system. For a given observable $f(x)$, i.e. a regular function defined in the phase space of a single particle, we define the average value of f on the whole system as

$$\hat{f}_\xi(t) = \int f(x)\rho_\xi(x,t)\, dx \tag{4.21}$$

this is an extensive macroscopic observable and it represents a typical result of an experimental measure. However, in many cases, the realization of the process $\xi(t)$ is unknown. In a Statistical Physics approach one assumes that in an ensemble of particles, each particle is subjected to a different noise realization. This assumption is not always justified and in a linear case can lead to wrong results, but in a non-linear case, the non-linear character of the dynamics allows to consider independent the different particles[1]. In this scenario, the properties of the system can be well described by the expectation value of any observable with respect to all possible realizations of the noise $\xi(t)$. Such methodology is justified if the experimental measurements are not instantaneous, but are obtained with a time-averaging procedure

$$\bar{f}_\xi(t) = \frac{1}{T}\int_t^{t+T} \hat{f}_\xi(t)\, dt \tag{4.22}$$

where the time T is sufficiently long with respect to the correlation time scale of the stochastic process $\xi(t)$. Then $\bar{f}_\xi(t)$ is almost independent from the realization $\xi(t)$ in a probabilistic sense, and we can approximate

$$\bar{f}(t) = \left\langle \hat{f}_\xi(t) \right\rangle = \int f(x) \left\langle \rho_\xi(x,t) \right\rangle\, dx \tag{4.23}$$

where $\langle \rho_\xi(x,t) \rangle$ represents the expectation value of the distribution function with respect to all possible noise realizations.

[1] However one should prove that the distribution of the noise realizations is the same if one considers a single particle or an ensemble of particles.

The knowledge of $\langle \rho_\xi(x,t) \rangle$ gives all the information needed for a statistical mechanics approach to the study of the system and, from a mathematical point of view, this assertion means the existence of a generalized law of large numbers for a statistical system composed by many elementary non-interacting particles.

The classical result for statistical systems made of N 'almost' independent particles is that the particle distribution function is well approximated by $\langle \rho_\xi(x,t) \rangle$ with a statistical error of order $O(\sqrt{N(\Delta V)})$, where $N(\Delta V)$ is the number of particles contained in the volume ΔV, that can be considered very small. From the point of view of a test particle, $\langle \rho_\xi(x,t) \rangle$ is the probability density to find the particle in the volume element dV.

All these average values are computed by considering independent realizations of $\xi(t)$ distributed according to $d\mu(\xi)$. There are of course situations in which the fluctuations effects are not neglegible: it can be because the system contains a limited number of particles or because the independence assumption for the dynamics of different particles is not valid. In such a case, $\hat{f}_\xi$ is an expectation value with respect to the probability distribution $\langle \rho_\xi(x,t) \rangle$, and it just represents the average value for a sequence of repeated independent measures.

Deriving a valid evolution equation for the average distribution $\langle \rho_\xi(x,t) \rangle$ is a key issue for non-equilibrium Statistical Mechanics. For doing so, it is possible to proceed in a formal way from the stochastic Liouville equation (4.14), which guarantees the symplectic character of the dynamics.

4.2 Fokker-Planck derivation

We now want to obtain an equation for $\langle \rho_\xi(x,t) \rangle$, starting from an equation in the form of (4.14). Let us start from a stochastic Liouville equation in the form

$$\frac{\partial \rho_\xi}{\partial t} = -\xi(t) \frac{\partial H}{\partial x}(x,t) J \frac{\partial \rho_\xi}{\partial x} \tag{4.24}$$

the solution can be formally written as

$$\rho_\xi(x,t) = \mathbb{T} \exp\left(- \int_0^t ds\, \xi(s) \frac{\partial H}{\partial x}(x,s) J \frac{\partial}{\partial x} \right) \rho_0(x) \tag{4.25}$$

where $\rho_0(x)$ is the initial distribution and $\mathbb{T}$ is the time-ordering operator, which is necessary due to the non-commutativity of the Lie operators

$$D_H = \frac{\partial H}{\partial x} J \frac{\partial}{\partial x} \tag{4.26}$$

at different times. Of course this problem does not sussist if we are working with a time independent H.

In order to compute the expectation value of Eq. (4.25) with respect of all the noise realizations, we can use the commutation property of the $\mathbb{T}$ operator and the averaging computation. By expanding the exponential operator we get the terms

$$\sum_n \frac{1}{n!} \left\langle \mathbb{T} \left(\int_0^t -\xi(s) \frac{\partial H}{\partial x}(x,s) J \frac{\partial}{\partial x}\, ds \right)^n \right\rangle \tag{4.27}$$

and it is trival to prove that

$$\left\langle \mathbb{T} \left(\int_0^t -\xi(s) \frac{\partial H}{\partial x}(x,s) J \frac{\partial}{\partial x}\, ds \right)^n \right\rangle = \\ = \mathbb{T} \left\langle \left(\int_0^t -\xi(s) \frac{\partial H}{\partial x}(x,s) J \frac{\partial}{\partial x}\, ds \right)^n \right\rangle \tag{4.28}$$

We now want to perform an explicit computation of (4.27) in the case of the white noise limit. In the white noise limit for $\xi(t)$, the stochastic operator is a Gaussian operator

$$D(t,\xi) = \int_0^t ds\, \xi(s) \frac{\partial H}{\partial x}(x,s) J \frac{\partial}{\partial x} \tag{4.29}$$

is a Gaussian operator, i.e. for each time t the probability density of the linear operators $D(t,\xi)$ is a Gaussian function. We can use at this point the properties of the moments of a Gaussian distribution and obtain directly

$$\frac{1}{n!} \mathbb{T} \langle D^n(t;\xi) \rangle = \frac{1}{2^{n/2}(n/2)!!} \mathbb{T} \left[\left\langle D^2(t;\xi) \right\rangle \right]^{n/2} \tag{4.30}$$

where $\mathbb{T}$ acts on the operator that defines the variance.

We can use a more convenient notation for the variance operator, and we can introduce the operator

$$B(x,t) = \frac{1}{2}\left\langle D^n(t,\xi)\right\rangle = \\ = \frac{1}{2}\int_0^t ds \int_0^t du\ \left\langle \xi(s)\xi(u)\right\rangle \frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x}\frac{\partial H}{\partial x}(x,u)J\frac{\partial}{\partial x} \quad (4.31)$$

and obtain the following equality (for n even, otherwise the contribution is zero)

$$\frac{1}{n!}\mathbb{T}\langle D^n(t;\xi)\rangle = \frac{1}{(n/2)!}\mathbb{T}[B(t)]^{n/2} \quad (4.32)$$

so that in the end, in the case of a Gaussian process, we can write the average value of Eq. (4.25) as

$$\left\langle \rho_\xi(x,t)\right\rangle = \mathbb{T}\exp(B(t))\rho_0(x) \quad (4.33)$$

At this point we consider the following

$$\left\langle \xi(s)\xi(u)\right\rangle = \delta(s-u) \quad (4.34)$$

and we apply it to the operator $B(x,t)$, which then reads:

$$B(x,t) = \frac{1}{2}\int_0^t ds \int_0^t du\,\delta(s-u)\frac{\partial H}{\partial x}(x,s)J\frac{\partial H}{\partial x}(x,u)J\frac{\partial}{\partial x} \quad (4.35)$$

$$= \frac{1}{2}\int_0^t ds\,\frac{\partial H}{\partial x}(x,s)J\frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x} \quad (4.36)$$

giving us the formal identity:

$$\left\langle \xi(s)\xi(u)\right\rangle = \mathbb{T}\exp\left(\frac{1}{2}\int_0^t \frac{\partial H}{\partial x}(x,s)J\frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x}\,ds\right)\rho_0(x) \quad (4.37)$$

In the end, we obtain that, for the stochastic Liouville equation (4.24), when $\xi(t)$ is a Gaussian noise with correlation function (4.39), and in the white-noise limit for the process $\xi(t)$, the average distribution $\langle\rho_\xi(x,t)\rangle$ satisfies the Fokker-Planck equation

$$\frac{\partial\left\langle \rho_\xi(x,t)\right\rangle}{\partial t} = \frac{1}{2}\frac{\partial H}{\partial t}(x,t)J\frac{\partial}{\partial x}\frac{\partial H}{\partial x}(x,t)J\frac{\partial\left\langle \rho_\xi(x,t)\right\rangle}{\partial x} \quad (4.38)$$

If we have instead the following correlation function for $\xi(t)$

$$\langle \xi(s)\xi(u)\rangle = \sigma^2\phi(\gamma(s-u)) \qquad s > u \tag{4.39}$$

we have that this procedure is not formally correct, since this Gaussian noise is in general not commutative. However, the Fokker-Planck expression can still be a good approximation in the limit $\sigma^2\phi(\gamma\tau) \to \delta\tau$.

Remark: the white-noise limit for the process $\xi(t)$ is a singular limit. To keep the noise effect finite, the amplitude of the noise should diverge. The Liouville equation (4.24), as seen in the previous section, is justified only when the noise correlation time γ^{-1} is greater than the evolution time scale Δt considered for the system itself. Meaning, the physical evolution we are considering is described with a discrete map with an evolution time scale Δt and a noise correlation time scale $\gamma^{-1} \gg \Delta t$. This means that Eq. (4.38) approximates the physical dynamics only when correlations $\gamma^{-1} \ll 1$ and $\gamma\Delta t \ll 1$ hold.

This setup is consistent with the Stratonovich limit presented in (4.10), that is equivalent to the stochastic differential equation

$$dx = D_H x\, dw_t + \frac{1}{2} D_H^2 x\, dt \tag{4.40}$$

that is associated to the Fokker-Planck equation (4.38). In another possible point of view, one can assume the stochastic Liouville equation as the true physical model to describe the evolution and Eq. (4.38) to describe the average evolution in the limit of a short correlation time for the noise fluctuations.

The white-noise limit is the fastest way to approximate the operator $B(x,t)$ when one has a fast decaying in the noise correlation. It is however possible to compute the effect of a finite noise correlation in the case of a slow exponentially decaying Gaussian noise. This is possible only by assuming a Gaussian character for $\xi(t)$, if this does not apply, it is necessary to use the white-noise limit for recovering the Gaussian character by means of the Central Limit Theorem.

In order to generalize these previous results to a stochastic Hamiltonian in the form of Eq. (4.1), we can observe that we can reach the form of the stochastic Liouville equation (4.24) by performing the symplectic change of variables

$$x(t,\xi) = \mathbf{\Phi}_0^t(y(t,\xi)) \tag{4.41}$$

where $\mathbf{\Phi}_0^t(\cdot)$ is the phase flux associated to the deterministic Hamiltonian $H_0(x)$. This change of variables leads us to the new Hamiltonian

$$H(y,t) = H_1(\mathbf{\Phi}_0^t(y))\xi(t) \tag{4.42}$$

so that we can apply the obtained result to the corresponding stochastic Liouville equation and obtain the following Fokker-Planck equation

$$\frac{\partial\langle\rho_\xi\rangle}{\partial t}(y,t) = \frac{1}{2}\frac{\partial H_1}{\partial y}(\mathbf{\Phi}_0^t(y))J\frac{\partial}{\partial y}\frac{\partial H_1}{\partial y}(\mathbf{\Phi}_0^t(y))J\frac{\partial}{\partial y}\langle\rho_\xi\rangle(y,t) \tag{4.43}$$

since the Hamiltonian Liouville operator is covariant with respect to the canonical change of variables

$$\frac{\partial H_1}{\partial y}(\mathbf{\Phi}_0^t(y),t)J\frac{\partial}{\partial y}\Big|_{y=\mathbf{\Phi}_0^{-t}(x)} = \frac{\partial H_1}{\partial x}(x)J\frac{\partial}{\partial x} \tag{4.44}$$

Thanks to this, we can rewrite the Fokker-Planck equation (4.43) in the original variables

$$\frac{\partial\langle\rho_\xi\rangle}{\partial t}(\mathbf{\Phi}_0^{-t}(x),t) = \frac{1}{2}\frac{\partial H_1}{\partial x}(x)J\frac{\partial}{\partial x}\frac{\partial H_1}{\partial y}(x)J\frac{\partial}{\partial x}\langle\rho_\xi\rangle(\mathbf{\Phi}_0^{-t}(x),t) \tag{4.45}$$

where $\mathbf{\Phi}_0^{-t}$ denotes the inverse phase flow. Finally, we let $\langle\rho_\xi\rangle(\mathbf{\Phi}_0^{-t},t) = \rho(x,t)$ and perform the following explicit partial derivative calculation

$$\frac{\partial}{\partial t}\rho(x,t) = \frac{\partial H_0}{\partial x}(x)J\frac{\partial}{\partial x}\rho(x,t) + \frac{\partial\langle\rho_\xi\rangle}{\partial t}(\mathbf{\Phi}_0^{-t}(x),t) \tag{4.46}$$

in which, when we replace the last r.h.s. term with Eq. (4.45), we obtain the final Fokker-Planck equation

$$\frac{\partial\rho}{\partial t}(x,t) = -\frac{\partial H_0}{\partial x}J\frac{\partial\rho}{\partial x}(x,t) + \frac{1}{2}\frac{\partial H_1}{\partial x}(x)J\frac{\partial}{\partial x}\frac{\partial H_1}{\partial x}(x)J\frac{\partial\rho}{\partial x}(x,t) \tag{4.47}$$

In the end, we have that the average distribution function of the stochastic Hamiltonian system (4.1), in the limit of a white Gaussian noise, satisfies the Fokker-Planck equation (4.47).

4.3 Diffusion over a toroidal surface

We shall now investigate the process of fast diffusion of a probability distribution over a toroidal surface. We are interested in this process as such situations happens in action-angle Hamiltonian systems for the angle variable.

We start by defining the following notation for the noise of a Wiener process and its integral:

$$w(t) = \int_0^t \xi(s)\, ds \qquad w_1(t) = \int_0^t w(t)\, ds \tag{4.48}$$

we can write $w_1(t)$ in the form

$$w_1(t) = \int_0^t \int_0^s \xi(u)\, du\, ds' = \int_0^t \xi(u) \int_u^s ds'\, du = \int_0^t (s'-u)\xi(u)\, du \tag{4.49}$$

This can be convenient if we want to display the processes $w(t)$ and $w_1(s)$ in the standard form

$$\int_0^t K(t,s)\xi(s)\, ds \tag{4.50}$$

where in this context $K(t,s)$ is called the **kernel** of the noise. We have that $K = 1$ for $w(t)$ and $K = t - s$ for $w_1(t)$. This implies $\sigma^2 = t$ for Wiener noise $w(t)$ while instead, for its integral $w_1(t)$, we have

$$\sigma^2(t) = \int_0^t (t-u)^2\, du = \frac{t^3}{3} \tag{4.51}$$

Let us consider now a basic example of angular diffusion. Let x be an angular variable defined over the interval $[0,1]$. Let us define then

$$\begin{aligned} \dot{x} &= \omega + y \mod 1 \\ \dot{y} &= \varepsilon\xi(t) \end{aligned} \tag{4.52}$$

A diffusive solution for an $\hat{x}$ defined on all $\mathbb{R}$ in the form $\rho(\hat{x},t)$ can be converted as a solution valid for x defined on the torus $\rho_T(x,t)$ by using the following periodicization

$$\rho_T(x,t) = \sum_{k=-\infty}^{+\infty} \rho(x+k,t) \quad x \in [0,1],\, k = 0, \pm 1, \pm 2, \ldots \tag{4.53}$$

The solution of (4.52) on $\mathbb{R}$ is given by

$$y = y_0 + \varepsilon w(t) \qquad x = x_0 + (\omega + \varepsilon y_0)t + \varepsilon w_1(t) \tag{4.54}$$

by writing the deterministic part of $x(t)$ in the form $\langle x(t)\rangle = x_0 + (\omega + \varepsilon y_0)t$, we can write the probability density in $\mathbb{R}$ in the form

$$\rho(x,t) = \frac{\exp\left\{\frac{-(x-\langle x(t)\rangle)^2}{2\sigma^2}\right\}}{\sqrt{2\pi\sigma^2}} \qquad \sigma^2 = \varepsilon^2\frac{t^3}{3} \tag{4.55}$$

we can apply the periodization and then obtain

$$\rho_T(x,t) = \sum_{n=-\infty}^{+\infty} \frac{\exp\left\{\frac{-(x-\langle x(t)\rangle+n)^2}{2\sigma^2}\right\}}{\sqrt{2\pi\sigma^2}} \qquad n = 0, \pm 1, \pm 2, \ldots \tag{4.56}$$

If we then represent the solution in the form of Fourier series

$$\rho_T(x,t) = \sum_{k=-\infty}^{+\infty} f_k(t) e^{2\pi i k x} \tag{4.57}$$

the f_k coefficients are given by

$$f_k = \int_0^1 e^{-2\pi i k x}\rho_T(x,t)\,dx \tag{4.58}$$

$$= \sum_{n=-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^1 \exp\left\{\frac{-(x-\langle x(t)\rangle+n)^2}{2\sigma^2} - 2\pi i k x\right\} dx \tag{4.59}$$

by assuming real f_k coefficients, we obtain

$$\rho_T(x,t) = 1 + 2\sum_{k=1}^{\infty} e^{-2\pi^2\sigma^2 k^2} f_k \cos(2\pi k(x - \langle x(t)\rangle)) \tag{4.60}$$

The most important consequence of this example, is noticing how ρ_T relaxes to the uniform distribution as $e^{-\sigma^2}$, where, in this scenario, we have

$$\sigma^2 \propto t^3 \tag{4.61}$$

while, instead, for a Wiener process, i.e. the one that affects y, we expect to have

$$\sigma^2 \propto t \tag{4.62}$$

4.3.1 Considerations for simple Hamiltonian systems

Let us now consider two simple, but significative examples of stochastically-perturbed Hamiltonian systems in action-angle coordinates (J, θ) in the form

$$H = H_0(J) + \varepsilon H_1(\theta, J)\xi(t). \tag{4.63}$$

Example 1

Considering the following Hamiltonian

$$H = H_0(J) + \varepsilon J\xi(t) \tag{4.64}$$

we have that

$$\begin{aligned} \dot{J} &= 0 \\ \dot{\theta} &= \omega(J) + \varepsilon\xi \end{aligned} \tag{4.65}$$

where $\omega(J) = \partial H_0/\partial J$. This implies that θ describes a simple Wiener process whose solution can be written in the form

$$\theta = \theta_0 + \omega t + \varepsilon w(t) \tag{4.66}$$

which implies a variance

$$\sigma_\theta^2 \propto t \tag{4.67}$$

Example 2

Considering the following Hamiltonian

$$H = H_0(J) - \varepsilon\theta\xi(t) \tag{4.68}$$

we have that

$$\begin{aligned} \dot{J} &= \varepsilon\xi \\ \dot{\theta} &= \omega(J) \end{aligned} \tag{4.69}$$

from which follows immediately

$$J = J_0 + \varepsilon w(t) \tag{4.70}$$

executing than a Taylor expansion for $\dot{\theta}$, we obtain

$$\dot{\theta} = \omega(J_0) + \omega'(J_0)\varepsilon w(t) + o(\varepsilon^2) \tag{4.71}$$

which, if integrated, brings

$$\theta = \theta_0 + \omega(J_0)t + \varepsilon\omega'(J_0)w_1(t) \tag{4.72}$$

this solution, as seen at the beginning of the Section, implies a variance

$$\sigma_\theta^2 \propto t^3 \tag{4.73}$$

while instead, for the action variable J, we have a variance

$$\sigma_J^2 \propto t \tag{4.74}$$

4.4 Averaging Principle

If we apply a small perturbation (e.g. a stochastic process with small ε) to an integrable system, the integrals of motion of the original system are expected to slowly change, and the order of magnitude of the time scale over which this change becomes considerable is generally dependent on the inverse of the perturbation parameter.

Let us assume a 1D system in action-angle coordinates, in which the angular coordinate is on a toroidal surface. Let us also assume it is possible to write the unperturbed equations of motion in this form

$$\begin{aligned} \dot{I} &= 0, \\ \dot{\theta} &= \omega(I) \end{aligned} \tag{4.75}$$

A small perturbation in the system, parametrized by a small quantity ε, creates new terms in the equations, these terms can be written in the following form

$$\begin{aligned} \dot{I} &= \varepsilon f(I, \theta, \varepsilon), \\ \dot{\theta} &= \omega(I) + \varepsilon g(I, \theta, \varepsilon) \end{aligned} \tag{4.76}$$

where f and g are periodic in θ. The action variable I, whose variation is $\mathcal{O}(\varepsilon)$ is called *slow variable*, instead, the angle variable θ, which varies as $\mathcal{O}(1)$ is called *fast variable*.

The quantities which originally were the integrals of motion start to slowly evolve. On time scales of order 1, these evolutions are small, on time scales of order $1/\varepsilon$ these evolutions can start to be relevant.

Usually, we are interested in knowing the behaviour in time of the slow variable. The idea behind the *averaging principle* is to replace the perturbed system (4.76) with the averaged system:

$$\dot{J} = \varepsilon F(J), \qquad F(J) = \frac{1}{2\pi} \oint_{\mathbb{T}} f(J, \theta, 0)\, d\theta \tag{4.77}$$

where, because θ evolves faster than I, its contribution to the variation of J over a period can be averaged. This allows us to describe the approximated evolution of the slow variables for time scales of order smaller or equal than $1/\varepsilon$.

As Vladimir I. Arnol'd specifies in [41], this principle is not a theorem, but a physical proposition, that is a vaguely stated, and, strictly speaking, false assertion. Such assertions often happen to be fruitful sources for mathematical theorems. Moreover, the problem of the influence of small Hamiltonian perturbations on an integrable Hamiltonian system was called by Poincaré the *fundamental problem of dynamics*.

We will show how it is possible to set up the averaging principle firstly for a 1D stochastic Hamiltonian in action-angle coordinates, and then for a generic multidimensional stochastic Hamiltonian in action-angle coordinates. We are mostly interested in reaching a valid expression for the probability distribution of the system in the form of a Fokker-Planck equation by averagin over the angular variable, expecting a faster relaxation to an uniform distribution over the toroidal space as shown in the previous Section.

4.4.1 Averaging a 1D stochastic Hamiltonian

Let us consider a 1D stochastically perturbed Hamiltonian system, described in action-angle coordinates, with the origin in a critical stable point of H_0, in the form

$$H = H_0(J) + \varepsilon H_1(\theta, J)\xi(t). \tag{4.78}$$

We have that the probability density $\rho(\theta, J, t)$ satisfies the equation

$$\frac{\partial \rho}{\partial t} + \Omega \frac{\partial \rho}{\partial \theta} = \frac{\varepsilon^2 \sigma^2}{2} \{H_1, \{H_1, \rho\}\} \tag{4.79}$$

where $\{,\}$ are the Poisson brackets and $\Omega(J) = \partial H_0 / \partial J$.

From the analysis in the previous Section, we know that a toroidal angle variable relaxes in much faster times compared to the action relaxation scales

($t_\theta \propto \varepsilon^{-2/3}$ for the angle variable while for the action we have instead $t_J \propto \varepsilon^{-2}$). Because of that it is possible to simplify this last equation into a Fokker-Planck equation for $t \gg t_\theta$: the distribution for the action variable $\rho(J,t)$, averaged over the angle variable, satisfies the equation

$$\frac{\partial}{\partial t}\rho(J,t) = \frac{\partial}{\partial J}D(J)\frac{\partial}{\partial J}\rho(J,t) \tag{4.80}$$

where

$$D(J) = \frac{\varepsilon^2\sigma^2}{2}\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{\partial H_1}{\partial\theta}\right)^2 d\theta \equiv \frac{\varepsilon^2\sigma^2}{2}\left\langle\left(\frac{\partial H_1}{\partial\theta}\right)^2\right\rangle_\theta \tag{4.81}$$

Proof: Let us start by considering the stochastic Liouville equation

$$\frac{\partial\rho}{\partial t} + \Omega(J)\frac{\partial\rho}{\partial\theta} + \varepsilon\xi(t)\{\rho, H_1\} = 0 \tag{4.82}$$

let us then write ρ in the form $\rho = \rho_0 + \varepsilon\rho_1$, where ρ_0 is the average component and ρ_1 the fluctuating component with zero average. Considering that $\langle\xi\rangle = 0$, we have that the average value of Eq. (4.82) is

$$\frac{\partial\rho_0}{\partial t} + \Omega(J)\frac{\partial\rho_0}{\partial\theta} + \varepsilon^2\{\langle\xi(t)\rho_1\rangle, H_1\} = 0 \tag{4.83}$$

Subtracting now (4.83) from (4.82), we obtain the equation

$$\frac{\partial\rho_1}{\partial t} + \Omega(J)\frac{\partial\rho_1}{\partial\theta} = -\xi(t)\{\rho_0, H_1\} + O(\varepsilon) \tag{4.84}$$

which we want to solve for ρ_1 so that we can substitute it into Eq. (4.83).

To do so, let us first execute the following change of variables

$$\begin{aligned}\theta &\to \theta - \Omega\tau\\ t &\to t - \tau\end{aligned} \tag{4.85}$$

which allows us to write

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\rho_1(\theta - \Omega\tau, J, t-\tau) = \xi(t-\tau)\{\rho_0, H_1\}(\theta - \Omega\tau, J, t-\tau) \tag{4.86}$$

integrating this last equation from $\tau = 0$ to $\tau = t$, we obtain

$$\rho_1(\theta, J, t) = -\int_0^t \{\rho_0, H_1\}(\theta - \Omega\tau, J, t-\tau)\xi(t-\tau)\,d\tau \tag{4.87}$$

where we took advantage over the fact that $\rho_1(\theta, J, 0) = 0$. Multiplying then both members for $\xi(t)$ and computing the average over all the noise realizations, we have, in the case of a Wiener noise:

$$\langle \xi(t)\rho_1(\theta, J, t)\rangle = -\frac{1}{2}\sigma^2\{\rho_0, H_1\}(\theta, J, t) \tag{4.88}$$

if we then replace this result into Eq. (4.83), we obtain, excluding terms of order ε^3 and higher, the equation for the average density

$$\frac{\partial \rho_0}{\partial t} + \Omega(J)\frac{\partial \rho_0}{\partial \theta} = \frac{\varepsilon^2\sigma^2}{2}\{\{\rho_0, H_0\}, H_0\}. \tag{4.89}$$

And, if we have that the relaxation time scale of the angle is faster than the action diffusion time scale, we can say in good approximation that $\rho = \rho(J, t)$ does not depend on the angle variable θ. This leads us to the double Poisson bracket expansion

$$\{\{\rho_0, H_0\}, H_0\} = \frac{\partial}{\partial J}\left[\left(\frac{\partial H_1}{\partial \theta}\right)^2 \frac{\partial \rho_0}{\partial J}\right] - \frac{\partial}{\partial \theta}\left[\frac{\partial \rho_0}{\partial J}\frac{\partial H_1}{\partial \theta}\frac{\partial H_1}{\partial J}\right] \tag{4.90}$$

by taking then the angular average of this last equation, the second term in the r.h.s becomes zero and we found the expression of $D(J)$ to be integrated over the whole torus.

4.4.2 Averaging principle for generic stochastic Hamiltonians

We still consider an perturbed Hamiltonian system in action-angle coordinates in presence of a weak chaotic dynamics in phase space, but this time we will consider a more generic approach valid for higher dimensions. We will work with the following stochastically-perturbed Hamiltonian system

$$H = H_0(I) + \xi(t)H_1(\theta, I) \tag{4.91}$$

where (θ, I) are the multidimensional action-angle variables and the noise realization $\xi(t)$ depends on the initial conditions of the orbit.

Let us start by introducing the slow variable

$$\phi = \theta - \Omega(I)t \tag{4.92}$$

where $\Omega(I) = \frac{\partial H_0}{\partial I}(I)$. This leads us to the following new Hamiltonian

$$H = \xi(t) H_1(\phi + \Omega(I)t, I) \tag{4.93}$$

via the generating function

$$F(\theta, J) = \theta J - H_0(J)t \tag{4.94}$$

In order to find an approximate solution of the previous dynamics, we consider the evolution of the angle-action variables for a time $T \gg \lambda$. This leads us to the map

$$\begin{aligned} \Delta\phi_j(T) &= \int_0^T \frac{\partial H_1}{\partial I_j}(\phi + \Omega(I)t, I)\xi(t)\, dt - \\ &\quad - \int_0^T t \frac{\partial H_1}{\partial \theta_k}(\phi + \Omega(I)t, I)\frac{\partial \Omega_k}{\partial I_j}\xi(t)\, dt \\ \Delta I_j(T) &= -\int_0^T \frac{\partial H_1}{\partial \theta_j}(\phi + \Omega(I)t, I)\xi(t)\, dt \end{aligned} \tag{4.95}$$

where $\Delta\phi_j(T) = \phi(T) - \phi(0)$ and $\Delta I(T) = I(T) - I(0)$. We want to inspect the second integral in the formula of $\Delta\phi_j(T)$ by performing the following integration by parts (we are of course truncating the expansion while doing so)

$$\begin{aligned} \int_0^T t \frac{\partial H_1}{\partial \theta_k}(\phi + \Omega(I)t, I)\frac{\partial \Omega_k}{\partial I_j}\xi(t)\, dt &\simeq \\ &\simeq \frac{\partial \Omega_k}{\partial I_j}\left[T \int_0^T \frac{\partial H_1}{\partial \theta_k}(\phi + \Omega(I)t, I)\xi(t)\, dt - \right. \\ &\quad \left. - \int_0^T \int_0^t \frac{\partial H_1}{\partial \theta_j}(\phi + \Omega(I)s, I)\xi(s)\, ds\, dt \right] \end{aligned} \tag{4.96}$$

where we can combine these two integrals and replace their arguments using the second equation in (4.95) and obtain

$$\int_0^T t \frac{\partial H_1}{\partial \theta_k}(\phi + \Omega(I)t, I)\frac{\partial \Omega_k}{\partial I_j}\xi(t)\, dt \simeq \frac{\partial \Omega_k}{\partial I_j}\int_0^T [\Delta I_k(T) - \Delta I_k(t)]\, dt \tag{4.97}$$

Therefore, we can say that if the action dynamics can be considered a stationary process, the main contribution to the angular dynamics is given by

$$\Delta\phi_j \simeq -\frac{\partial \Omega_k}{\partial I_j}\int_0^T \Delta I_k(t)\, dt \tag{4.98}$$

Therefore if the system is non-degenerate, i.e. the matrix $\partial\Omega_k/\partial I_j$ is not singular, the increment of the angle variables are the integral of the increments of the action and, as we saw in the previous sections, we expect a much faster relaxation to a uniform distribution. From this result we have that in the evolution of the action variables we can approximate the distribution of the angle variables with a uniform distribution. To get an approximation of the action dynamics, we consider the change of the action up to terms of order $\mathcal{O}(\|H_1\|^2)$

$$\Delta I_j = \int_0^T \frac{\partial H_1}{\partial \theta_j}\xi(s)\,ds + \int_0^T\int_0^t \frac{\partial^2 H_1}{\partial I_k \partial\theta_j}\frac{\partial H_1}{\partial\theta_k}\xi(t)\xi(s)\,ds\,dt - \\ - \int_0^T\int_0^t \frac{\partial^2 H_1}{\partial\theta_k\partial\theta_j}\left[\frac{\partial H_1}{\partial I_k} - \frac{\partial\Omega}{\partial I_k}\frac{\partial H_1}{\partial\theta_k}t\right]\xi(t)\xi(s)\,ds\,dt \quad (4.99)$$

Let us then assume that the angles ϕ are uniformly distributed so that we can average on the noise realizations and on the angles in a single step. The actions I perform a stochastic dynamics, whose average value is

$$\begin{aligned}\langle \Delta I_j\rangle_\phi &= \frac{\partial}{\partial I_k}\int_0^T\int_0^t \left\langle \frac{\partial H_1}{\partial\theta_j}\frac{\partial H_1}{\partial\theta_k}\right\rangle e^{-(t-s)/\lambda}\,ds\,dt \\ &= \frac{1}{2}\frac{\partial}{\partial I_k}\left\langle \frac{\partial H_1}{\partial\theta_j}\frac{\partial H_1}{\partial\theta_k}\right\rangle \int_0^T\int_0^T e^{-|t-s|/\lambda}\,ds\,dt\end{aligned} \quad (4.100)$$

where we have neglected the terms that are derivatves with respect to an angle variable and used the decorrelation law of $\xi(t)$. The corresponding variance is estimated by

$$\langle(\Delta I_j - \langle\Delta I_j\rangle_\phi)(\Delta I_k - \langle\Delta I_k\rangle_\phi)\rangle_\phi = \\ = \left\langle \frac{\partial H_1}{\partial\theta_j}\frac{\partial H_1}{\partial\theta_k}\right\rangle \int_0^T\int_0^T e^{-|t-s|/\lambda}\,ds\,dt \quad (4.101)$$

In the limit of slow diffusion time and in the approximation of a fast angle relaxation, we can describe the action dynamics by a stochastic process of the form

$$\Delta I_j = -\sqrt{T\lambda\|H_1\|^2}\sqrt{\left\langle \frac{\partial H_1}{\partial\theta_j}\frac{\partial H_1}{\partial\theta_l}\right\rangle}\hat{\xi}_l + \frac{T\lambda\|H_1\|^2}{2}\frac{\partial}{\partial I_k}\left\langle \frac{\partial H_1}{\partial\theta_j}\frac{\partial H_1}{\partial\theta_k}\right\rangle \quad (4.102)$$

where $\hat{H}_1 = H_1/\|H_1\|$ and $\hat{\xi}_l$ are identical independent distributed random variables with zero mean value and unitary variance. We can interpret the quantity $T\lambda\|H_1\|^2$ as the time step $\Delta\tau$ of the diffusion time (even though it has a dimension of an action): the continuous limit is recovered when $T \to \infty$ (so that $T \gg \lambda$) and $\|H_1\|^2 T \to 0$. In other words, T should be sufficiently long in order to consider the angles relaxing to a uniform distribution and the noise decorrelated, but $\|H_1\|^2$ has to be so small that the actions do not evolve in a time T. We also recall the limit requested for the validity of the stochastic Langevine equation

$$\lim_{T\to\infty} \frac{1}{T\lambda} \int_0^T \int_0^T e^{-|t-s|/\lambda}\, ds\, dt = 1 \tag{4.103}$$

Finally, introducing the diffusion time notation $\tau = \lambda\|H_1\|^2 t$, we have that Eq. (4.102) is the approximation of the solution of the stochastic differential equation

$$dI_j = -\sqrt{\left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_l} \right\rangle}\, d\omega_l(\tau) + \frac{1}{2}\frac{\partial}{\partial I_k}\left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k} \right\rangle d\tau \tag{4.104}$$

In order to completely check the consistency of our claims, let us consider again the angle dynamics in the diffusion time (we omit for convenience the indices)

$$\Delta\phi = \left\langle \frac{\partial \hat{H}_1}{\partial I} \right\rangle_\phi \sqrt{T\lambda\|H_1\|^2}\, \xi - \frac{\partial \Omega}{\partial I} \int_0^T (I(T) - I(t))\, dt \tag{4.105}$$

by applying the approximation (4.104) we just obtained, we can compute directly the fluctuating part of the action term $(I(T) - I(t))$ and obtain

$$\mathrm{Var}[I(T) - I(t)] = \left\langle \left(\frac{\partial \hat{H}_1}{\partial \theta} \right)^2 \right\rangle (\lambda\|H_1\|^2) \frac{T^3}{2} \tag{4.106}$$

At this point, if $\partial\Omega/\partial I \propto \mathcal{O}(1)$ the stochastic approximation of the action variables we performed implies that the relaxation process for the angles ϕ must be archived after a time t_ϕ, where $\lambda\|H_1\|^2 t_\phi^3 \simeq \mathcal{O}(1)$. Then, we can estimate

$$t_\phi \simeq \lambda^{-1/3}\|H_1\|^{-2/3} \tag{4.107}$$

so that the choice $T \simeq t_\phi$ provides the result

$$\Delta\tau \simeq \lambda^{2/3} \|H_1\|^{4/3} \tag{4.108}$$

which vastly proves the assumption of fast angle relaxation, when compared to the diffusion time scale.

The assumption on the fast relaxation of the angles ϕ, necessary to derive the equation (4.104), can be satisfied if the estimate $\|H_1\|^2 \ll \lambda^{-1}$ holds for the decorrelation time scale of the random fluctuations. This condition is necessary to describe the stochastic Hamiltonian dynamics as a diffusion process and it implies that the approach is consistent even if the Ljapounov exponent, characterizing the chaotic region, is small.

The evolution of the distribution function $\rho(I, \tau)$ at the diffusion time scale is well approximated by the solution of the Fokker-Planck equation

$$\frac{\partial \rho}{\partial \tau} = \frac{1}{2} \frac{\partial}{\partial I_j} \left\langle \frac{\partial H_1}{\partial \theta_j} \frac{\partial H_1}{\partial \theta_k} \right\rangle \frac{\partial}{\partial I_k} \rho(I, \tau) \tag{4.109}$$

where in this equation the slow diffusion time coefficient τ has the dimension of the square of an action, so that the diffusion coefficient is adimensional.

Chapter 5

A diffusive framework for non-linear beam dynamics

5.1 A simple model

As we discussed in the previous chapters, quasi-integrable Hamiltonian systems are provided with a set of invariant KAM tori of large measure. Moreover, for multidimensional systems, transport in phase space occurs only over a set of initial conditions of extremely small measure, through the topological mechanism of Arnold's diffusion. This phenomenon is characterized in general by an extremely long timescale for finite transport phenomena[42].

In a system described by a symplectic polynomial map with an elliptic fixed point at the origin, KAM invariant tori exist in a neighbourhood of the fixed point, but their measure gradually decreases as the distance from the fixed point increases, and a weakly-chaotic region takes the place of the broken tori. In this scenario, the Nekhoroshev Theorem 4 makes possible to elaborate some estimates of the timescales.

As explained in Sec. 3.2, the concept of Dynamic Aperture defines the region where a 'pure' KAM theory applies, i.e. the measure of the set of chaotic orbits is neglegible so that Arnold's diffusion is the only effective transport mechanism. Beyond the Dynamic Aperture, one expects the appearance of a large, weakly-chaotic region, generated by the overlap of serveral nonlinear resonances, where the orbits can be trapped for a long enough time, i.e. longer than the operational time of the machine, before a fast escape to infinity occurs.

The amplitude of the chaotic region depends on the non-linear terms in the map and, according to perturbation theory, the action variables are almost conserved in this region, whereas the angle variables follow a dynamics governed by chaotic fluctuations over a characteristic timescale. In the chaotic region, the diffusive behaviour can be extremely difficult to describe, due to the underlying geometrical structures associated with the Hamiltonian character of the phase space[43].

Moreover, in real scenarios, like applications to beam dynamic problems, it is not possible to avoid the presence of small external random perturbations [44], that can destroy the microscopic structures of the phase space responsable of the trapping of orbits for long time (stiking phenomenon). In this way we may have an homogeneous region of chaoticity in the phse space. These considerations justify the description of the orbit diffusion by means of a stochastically perturbed Hamiltonian, which has been presented and studied in Section 4.1. Such proposed Hamiltonian model has the following form

$$H(\theta, I, t) = H_0(I) + \varepsilon\xi(t)H_1(\theta, I) \tag{5.1}$$

with (θ, I) action-angle variables and $\xi(t)$ a regular, stationary stochastic noise with zero mean value and unit variance, which mimics the effects of the chaotic dynamics. Moreover, H_1 is normalized so that the parameter ε is what defines the diffusion timescale $\varepsilon^2 t$ of the process.

In this model, the term $H_0(I)$ represents the regular, deterministic part of the magnetic lattice in analysis, while the perturbation term $H_1(\theta, I)$ measures the effects linked with the presence of small random fluctuations, taking into account also the phase-space inhomogeneities.

For utilizing the stochastic model (5.1) for the DA problem, some considerations must be made:

- This model alone cannot describe the fast escape of a particle orbit *to infinity*, therefore it needs to be equipped with an absorbing barrier at a given distance to the origin (this is typical for every physical model).

- The evaluation based on this model must be made with $\varepsilon \ll 1$ in terms of the action threshold $I_{\mathrm{da}}(t_*)$ for which the probability that a particle with $I_0 \leq I_{\mathrm{da}}$ is lost at the absorbing boundary for $t \leq t_*$ is less than an *a priori* given small, albeit non zero value[1].

[1]Taking a zero value would mean not taking into considerations effects like the Arnold

- In the limit $t_* \to +\infty$, all the particles are lost with probability 1, unless the perturbation vanishes at a finite distance from the origin (which is not our case).

From perturbation theory, a possible approach for estimating the norm of $H_1(\theta, I)$ is based on the analysis of the asymptotic character of the perturbation series. For the case of a symplectic polynomial map, provided there are no dominating low-order resonances in the phase space, there is a generic estimate of the reminder of the Birkhoff's Normal Form which gives [26]:

$$\|R_n(I)\| \propto (n!)^\eta \left(\frac{I}{I_*}\right)^{n/2} \tag{5.2}$$

where the factorial term takes into account the number of contributions due to the structure of the functional equations, the exponent η is related to the number of degrees of freedom, and I_* is related to the strength of the non-linear terms.

For each I there exist an optimal order for the Normal Form reminder defined by

$$n^\eta = \left(\frac{I}{I_*}\right)^{\frac{1}{2}} \qquad \Rightarrow \qquad n = \left(\frac{I_*}{I}\right)^{-1/2\eta} \tag{5.3}$$

which, via substitution into Eq. (5.2), leads us to a Nekhoroshev-like estimate:

$$\|R_n(I)\| \propto \exp\left[-\eta \left(\frac{I_*}{I}\right)^{1/2\eta}\right] \tag{5.4}$$

Our assumption is that the estimate for the optimal remainder, which scales with the action I, is a fingerprint of the non-integrability of the dynamics and can be applied in the region of weak chaoticity to measure the fluctuations in the dynamics that destroys the long-term stability of the orbits.

We can operate with the estimate (5.4) and assume that it gives also a measure of the orbits diffusion in phase space. We can therefore work with the stochastic model (5.1) assuming that:

$$\|H_1(\theta, I)\| \simeq \exp\left[-\left(\frac{I_*}{I}\right)^{\alpha}\right] \tag{5.5}$$

diffusion which, as said before, breaks the possibility to have stable regions for infinite times.

These steps are the result of the fundamental conjecture that the non-existence of KAM tori over a certain region is intrinsically connected to the appearance of chaotic regions. This conjecture is fundamental in this preliminary stage in order to justify this approach.

To obtain an equation for $\rho(\theta, I, t)$, it is possible to apply to model (5.1) the averaging procedure over the angle variable discussed in Sec. 4.4. Of course, this approach can be justified when we are able to distinguish four timescales[45]:

1. the noise decorrelation timescale γ^{-1};
2. the averaging timescale $T \gg \gamma^{-1}$ at which $\xi(t)$ can be well approximated as white-noise;
3. the angle relaxation timescale $\propto \varepsilon^{-4/3} \gg T$;
4. the action diffusion timescale $\propto \varepsilon^{-2}$;

then, in the limit $\varepsilon \to 0, T \to +\infty$ with γ finite and $\varepsilon^2 T \ll 1$, we can well approximate ρ as the solution of the Fokker-Planck equation

$$\frac{\partial \rho}{\partial \tau} = \frac{1}{2}\frac{\partial}{\partial I}\left(\left\langle\left(\frac{\partial H_1}{\partial \theta}\right)^2\right\rangle_\theta \frac{\partial \rho}{\partial I}\right) \equiv \frac{1}{2}\frac{\partial}{\partial I}\left(D(I)\frac{\partial \rho}{\partial I}\right) \tag{5.6}$$

where τ is a slow effective time $\propto \varepsilon^2\sqrt{\gamma^{-1}T}t$. Since $\lim_{I\to 0} h(I) = 0$, we have a natural boundary at $I = 0$, we just need to add an absorbing boundary condition at $I = I_\mathrm{a}$ which, as explained at the beginning, represents the starting position of a known fast escape to infinity or, also, the presence of a collimator.

5.2 Approaches for diffusion measurements

5.2.1 Stochastic Symplectic Maps

Let us consider the following stochastic symplectic map

$$\begin{pmatrix} x_{n+1} \\ p_{n+1} \end{pmatrix} = \begin{pmatrix} \cos\Omega(I) & \sin\Omega(I) \\ -\sin\Omega(I) & \cos\Omega(I) \end{pmatrix} \begin{pmatrix} x_n \\ p_n + \epsilon\xi_{n+1} f(x_n) \end{pmatrix} \tag{5.7}$$

where ξ_n are independent random variables with zero mean value and unit variance and ϵ is a perturbation parameter (the independence condition for the variables ξ_n can be weakened in the limit of $\epsilon \ll 1$). The function $f(x) = \mathcal{O}(x^2)$ simulates the presence of non-linear terms (or the remainder of a perturbation theory) so that the origin is a stable fixed point.

Without loss of generality we assume that the integrable part is defined by the frequency

$$\Omega(I) = \omega_0 + \omega_1 I + \omega_2 \frac{I^2}{2} \quad I = \frac{p^2 + x^2}{2} \tag{5.8}$$

we assume also that the linear frequency ω_0 is not resonant.

As a stochastic Hamiltonian system, the stationary solution is the uniform one in phase space. This implies that the dynamics is certainly unstable, i.e. for any given initial condition, the probability to leave a given disk around the origin tends to 1 as $n \to \infty$. However, we are interested in

- the expectation value of the first passage time at an absorbing barrier as a function of the initial neighbourhood of the origin;
- the average time to lose a given percentage of the initial population;
- the evolution of an initial distribution function of particles.

The action dynamics of the map can be approximated in the following way:

$$I_{n+1} = I_n + \epsilon \xi_{n+1} \sqrt{2I_n} \cos\left(\theta_n + \Omega\left(I_n\right)\right) f\left(\sqrt{2I_n} \sin\left(\theta_n + \Omega\left(I_n\right)\right)\right) + \\ + \frac{\epsilon^2}{2} \xi_{n+1}^2 f^2\left(\sqrt{2I_n} \sin\left(\theta_n + \Omega\left(I_n\right)\right)\right) \tag{5.9}$$

Due to the symplectic nature of the map, the term $\mathcal{O}(\epsilon)$ has the form

$$\sqrt{2I} \cos(\theta) f(\sqrt{2I} \sin(\theta)) = \frac{\partial F}{\partial \theta}(\sqrt{2I} \sin(\theta)) \tag{5.10}$$

where $F(x)$ is the primitive of $f(x)$, this implies that its angular mean value is zero.

In order to apply properly the averaging theorem, we need to consider the fast relaxation of the angle variable in a diffusion time scale ϵ^{-2}, this

requires the condition $\omega_2 = \mathcal{O}(1)$. We end up then with the following diffusion coefficient for the action variable

$$D(I) = \left\langle \left(\frac{\partial F}{\partial \theta}(\theta, I) \right)^2 \right\rangle_\theta \tag{5.11}$$

At this point, we are interested in formulating a Fokker-Planck process to describe the action dynamic of this stochastic map. To do so, we can interpret $\epsilon \xi_n(t) F(x)$ as the interpolating Hamiltonian for the following map

$$\begin{pmatrix} x_{n+1} \\ p_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ p_n + \epsilon \xi_{n+1} f(x_n) \end{pmatrix} \tag{5.12}$$

where $\xi_n(t)$ is still a stepwise stochastic process. Then, we can perform the following time-dependent change of variables to the map

$$\begin{aligned} \phi_n &= \theta_n - \Omega(I_n)\, n \\ J_n &= I_n \end{aligned} \tag{5.13}$$

and obtain the following interpolating Hamiltonian

$$H = \epsilon \xi_n(t) F(J, \phi + \Omega(J) t) \tag{5.14}$$

to which we can apply the procedure presented in Chapter 4 and describe the action dynamics with a Fokker-Planck equation in the self-adjoint form with the diffusion coefficient given by Eq. (5.11).

In order to study the features of the a Nekhoroshev-like diffusion, we can set the following

$$F(x) = \int_0^x \exp\left(-\frac{x_*^2}{u^2} \right) du = \int_0^{\sqrt{2I}\sin\theta} \exp\left(-\frac{2I_*}{u^2} \right) du \tag{5.15}$$

where we have defined $I_* = x_*/2$. The corresponding diffusion coefficient is then given by

$$D(I) = 2I \left\langle \cos^2\theta \exp\left(-2 \frac{I_*}{I \sin^2\theta} \right) \right\rangle \tag{5.16}$$

To compute the integral, since the only angular terms are $\sin^2\theta$ and $\cos^2\theta$, we can consider only the positive quadrant of phase space, therefore we obtain

$$D(I) = \frac{4I}{\pi} \int_0^{\pi/2} \cos^2\theta \exp\left(-2 \frac{I_*}{I \sin^2\theta} \right) d\theta \tag{5.17}$$

If we perform then the change of variables $u = \sin^{-1}\theta$ we end up with

$$D(I) = \frac{4I}{\pi}\int_1^\infty \frac{1}{u^2}\sqrt{1-\frac{1}{u^2}}\exp\left(-2\frac{I_*}{I}u^2\right)\,du \leq \\ \leq \frac{8I}{3^{3/2}\pi}\int_0^\infty \exp\left(-2\frac{I_*}{I}u^2\right)\,du \quad (5.18)$$

so that we can write

$$D(I) \simeq cI^{3/2}\exp\left(-2\frac{I_*}{I}\right) \quad (5.19)$$

for a suitable constant c.

Stochastic noise and deterministic noise

Such approximation is valid for an ensemble of particles with independent noise realizations ξ_n with zero mean and unit variance, i.e. each particle has its own realization of the noise and the noise itself has zero autocorrelation.

However, such noise characteristics are not useful for emulating the effects of non-linear magnetic fields in the accelerator lattice. In fact, these effects can be better emulated with a 'deterministic noise' which has a single realization common for every particle and has a non-zero autocorrelation.

The results in Ref. [43], and most of the references therein, conjecture that it is still possible to observe most of the diffusive behaviours in phase space, even with an ensemble of particles all provided with the same noise realization.

Moreover, the authors also conjecture that such diffusive processes can be seen also if we replace the ξ_n random variables with a chaotic (but still deterministic) map, like a Standard Map[46]. This would imply that our approximation to a Fokker-Planck equation can be justified for couplings with maps that are 'chaotic enough'.

In order to verify these conjectures, we will utilize three different kinds of noise when simulating the symplectic map (5.9), and then compare the evolution of the ensemble with the corresponding Fokker-Planck process, computed with a Crank-Nicolson integrator (see Appendix A for the implementation detail).

The three noises we use are:

1. Random noise with different realizations for each particle. This will give us a baseline result for our approximation.

2. Random noise with the same realization for all the particles and different values of autocorrelation, in order to observe at which point the approximation 'breaks'.

3. Deterministic noise generated with a Standard Map, with different level of chaoticity.

A correlated random noise can be simulated by using the auxiliary dynamics

$$\xi_{n+1} = \gamma \xi_n + w_n \tag{5.20}$$

where $0 < \gamma < 1$ and w_n are independent normalized processes. We can analyse this process by letting $\eta_n = \gamma^n \xi_n$. We get

$$\eta_{n+1} = \eta_n + \gamma^{-(n+1)} w_n \tag{5.21}$$

then, letting $\eta_0 = 0$, we get

$$\eta_n = \sum_{k=1}^{n} \gamma^{-(k+1)} w_k \quad \xi_n = \sum_{k=1}^{n} \gamma^{n-(k+1)} w_k \tag{5.22}$$

and this leads us to $E(\xi_n) = 0$ when the initial condition is $\xi_0 = 0$ and, as for the correlation:

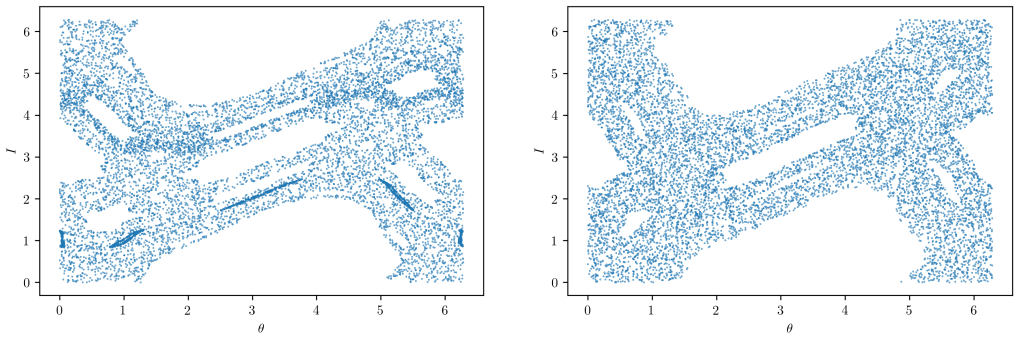
$$\begin{aligned} E\left(\xi_n \xi_m\right) &= \sum_{k=1}^{n} \sum_{h=1}^{m} \gamma^{n-(k+1)+m-(h+1)} E\left(w_k w_h\right) \\ &= \gamma^{n-m} \sum_{h=1}^{m} \gamma^{2m-2(h+1)} \end{aligned} \tag{5.23}$$

where $n > m$.

As for the Standard Map noise, we consider the following implementation

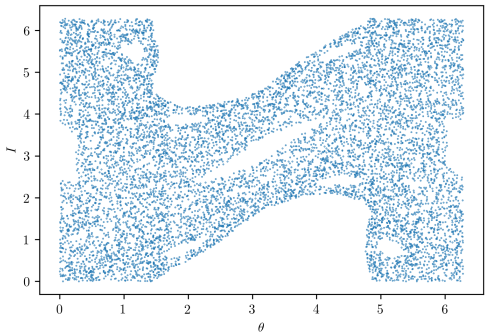
$$\begin{aligned} J_{n+1} &= J_n + K \sin\left(\psi_n\right) \qquad \mod 2\pi \\ \psi_{n+1} &= \psi_n + J_{n+1} \qquad\qquad\quad \mod 2\pi \end{aligned} \tag{5.24}$$

and use the array of J_n as our deterministic noise, with different values for K in order to simulate different levels of chaoticity. For a graphical visualization of some standard maps with different values of K, refer to Fig. 5.1.

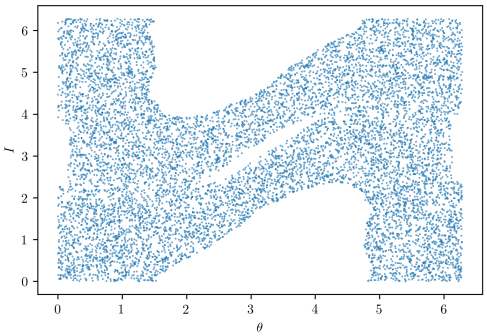


(a) ($K = 1.25$)

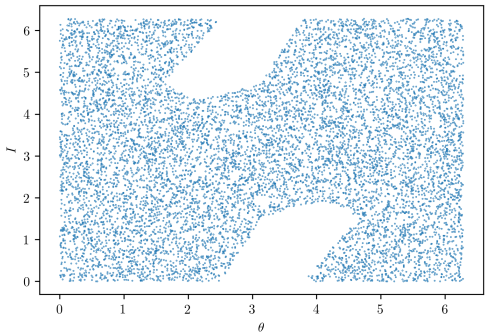
(b) ($K = 1.50$)



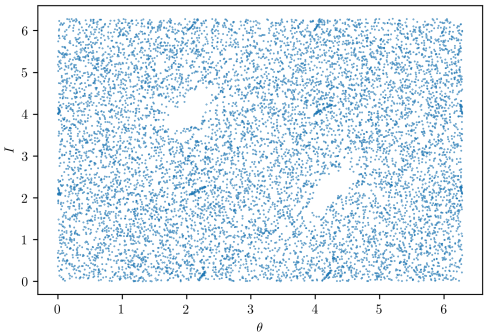
(c) ($K = 1.75$)



(d) ($K = 2.00$)



(e) ($K = 3.00$)



(f) ($K = 5.00$)

Figure 5.1: Standard maps with different K coefficients iterated 10^5 times. Starting condition: $(I_0 = 0.5, \theta_0 = 0.5)$. We observe as the strength of chaoticity increases with K.

5.2.2 First passage time

Let us start by considering a generic diffusion problem

$$\frac{\partial \rho}{\partial t}(x,t) = L_{FP}\rho(x,t) \tag{5.25}$$

where L_{FP} is a generic Fokker-Planck operator, with $x \in \Omega$ and with an absorbing boundary condition at $\partial\Omega$.

By definition the probability to be in Ω at time t, starting from an initial condition $\delta(x-a)$ is

$$P_\Omega(t|a) = \int_\Omega \rho(x,t|a)\, dx \tag{5.26}$$

Let t_a be the first passage time at $\partial\Omega$ for a realization of an initial condition $\partial(x-a)$, and let $p\,(t|a)$ the probability distribution of such t_a, we have by definition

$$\int_t^\infty p(s|a)ds = P_\Omega(t|a) \tag{5.27}$$

since $\partial\Omega$ is an absorbing boundary condition, i.e. the first passage time correspond to the loss time.

We can write the following

$$p(t|a) = \int_\Omega \frac{\partial}{\partial t}\rho(x,t|a)dx = \int_\Omega L_{FP}\rho(x,t|a)dx \tag{5.28}$$

this leads us to an equation for the average first passage time $\langle t_a(a)\rangle$

$$\begin{aligned}\langle t_a(a)\rangle &= \int_0^\infty tp(t|a)\, dt \\ &= \int_\Omega \int_0^\infty t\frac{\partial}{\partial t}\rho(x,t|a)\, dt\, dx \\ &= -\int_0^\infty \int_\Omega \rho(x,t|a)\, dx\, dt\end{aligned} \tag{5.29}$$

where we have performed an integration by parts. We can use then the adjoint operator $L_{FP}^\dagger$ and take into account the boundary conditions

$$L_{FP}^\dagger\langle t_a(a)\rangle = -\int_0^\infty \int_\Omega L_{FP}^\dagger\rho(x,t|a)\, dx\, dt \tag{5.30}$$

Remark: in this context, $L_{FP}^{\dagger}$ acts on the initial state a.

From the Kolmogorov relation for Markov processes, we know that

$$\frac{\partial}{\partial s}\int_{\Omega} dt\rho\left(x,t|x_0,s\right)\rho\left(x_0,s|a,0\right)dx_0 = 0 \tag{5.31}$$

so that we can write

$$\begin{aligned}\int_{\Omega} dx_0 \left[\rho\left(x_0,s|a,0\right)\frac{\partial}{\partial s}\rho\left(x,t|x_0,s\right)+\rho\left(x,t|x_0,s\right)\frac{\partial}{\partial s}\rho\left(x_0,s|a,0\right)\right] = \\ = \int_{\Omega} dx_0 \left[\rho\left(x_0,s|a,0\right)\frac{\partial}{\partial s}\rho\left(x,t|x_0,s\right)+\rho\left(x,t|x_0,s\right)L_{FP}\rho\left(x_0,s|a,0\right)\right]\end{aligned} \tag{5.32}$$

and obtain

$$\int_{\Omega} dx_0 \left[\frac{\partial}{\partial s}\rho\left(x,t|x_0,s\right)+L_{FP}^{\dagger}\rho\left(x,t|x_0,s\right)\right]\rho\left(x_0,s|a,0\right) = 0 \tag{5.33}$$

if we are considering a stationary process, we can also write

$$\frac{\partial}{\partial s}\rho\left(x,t|x_0,s\right) = \frac{\partial}{\partial s}\rho\left(x,t-s|x_0,0\right) = -\frac{\partial}{\partial t}\rho\left(x,t-s|x_0,0\right) \tag{5.34}$$

Letting $s \to 0$, we get the adjoint equation

$$\frac{\partial}{\partial t}\rho(x,t|a,0) = -L_{FP}^{\dagger}\rho(x,t|a,0) \tag{5.35}$$

Combining this last result with Eq. (5.30), we have the following equation for the average first passage time

$$\begin{aligned}L_{FP}^{\dagger}\langle t_a(a)\rangle &= -\int_0^{\infty}\int_{\Omega} L_{FP}^{\dagger}\rho(x,t|a)\,dx\,dt \\ &= \int_{\Omega}\int_0^{\infty}\frac{\partial}{\partial t}\rho(x,t|a,0)\,dt\,dx \\ &= -\int_{\Omega}\delta(x-a)\,dx \\ &= -1\end{aligned} \tag{5.36}$$

where the last passage is given by the fact that $\lim_{t\to\infty}\rho(x,t|a) = 0$, due to the presence of an absorbing boundary condition. In an explicit form, we end up with the equation

$$L_{FP}^{\dagger}\langle t_a(a)\rangle = -1 \tag{5.37}$$

As for the variance of the distribution $p(t|a)$, we can write the following

$$\begin{aligned}\sigma^2(a) &= \int_0^\infty t^2 p(t|a)\,dt - \langle t_a(a)\rangle^2 \\ &= -2\int_0^\infty \int_\Omega t\rho(x,t|a)\,dx\,dt - \langle t_a(a)\rangle^2\end{aligned} \tag{5.38}$$

where we executed an integration by parts. Then, we have

$$L_{FP}^\dagger \sigma^2(a) = -2\int_0^\infty \int_\Omega t L_{FP}^\dagger \rho(x,t|a)\,dx\,dt\; - L_{FP}^\dagger \langle t_a(a)\rangle^2 \tag{5.39}$$

using the relation

$$\begin{aligned}2\int_\Omega \int_0^\infty t\frac{\partial}{\partial t}\rho(x,t|a,0)\,dt\,dx &= -2\int_\Omega \int_0^\infty \rho(x,t|a,0)\,dt\,dx \\ &= -2\int_0^\infty tp(t|a)dt = -2\langle t_a(a)\rangle\end{aligned} \tag{5.40}$$

we get the following equation

$$L_{FP}^\dagger \sigma^2(a) = -2\langle t_a(a)\rangle - L_{FP}^\dagger \langle t_a(a)\rangle \tag{5.41}$$

Now we want to specify Eq. (5.37) and Eq. (5.41) to a self adjoint-case, like the one we are interested in

$$L_{FP} = \frac{1}{2}\frac{\partial}{\partial x}D(x)\frac{\partial}{\partial x} \tag{5.42}$$

By applying directly Eq. (5.37), we can evaluate the average first passage time $\langle t_a(a)\rangle$ for the first passage time starting from a point a

$$\frac{1}{2}\frac{\partial}{\partial a}D(a)\frac{\partial}{\partial a} = -1 \tag{5.43}$$

so that we can obtain immediately

$$\langle t_a(a)\rangle = 2\int_a^{x_a} \frac{x}{D(x)}\,dx \tag{5.44}$$

where of course we have set $\langle t_a(x_a)\rangle = 0$ for the absorbing boundary condition located at x_a. From the relation in Eq. (5.41) and from the following relation

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial x}D(x)\frac{\partial}{\partial x}\langle t_a(x)\rangle^2 &= \frac{\partial}{\partial x}D(x)\langle t_a(x)\rangle\frac{\partial\langle t_a\rangle}{\partial x} \\ = -2\langle t_a(x)\rangle + D(x)&\left(\frac{\partial\langle t_a\rangle}{\partial x}\right)^2\end{aligned} \tag{5.45}$$

we can write

$$\frac{1}{2}\frac{\partial}{\partial a}D(a)\frac{\partial\sigma^2}{\partial a} = -D(a)\left(\frac{\partial\langle t_a\rangle}{\partial a}\right)^2 = -\frac{4a^2}{D(a)} \tag{5.46}$$

and obtain the following result

$$\sigma^2(a) = 8\int_a^{x_a}\frac{x^2}{D(x)}\int_x^{x_a}\frac{1}{D(y)}\,dy\,dx \tag{5.47}$$

5.2.3 Current interpolation

Let us explicitly compute the current at an absorbing barrier I_a for the Fokker-Planck equation

$$\frac{\partial\rho}{\partial t} = \frac{1}{2}\frac{\partial}{\partial I}D(I)\frac{\partial}{\partial I}\rho \tag{5.48}$$

with a Nekhoroshev-like diffusion coefficient $D(I) = h^2(I)$

$$h(I) = c\exp\left[\left(-\frac{I_*}{I}\right)^\alpha\right] \tag{5.49}$$

In order to simplify the form of $h(I)$ it is possible to rescale the time in the form of $\tau = c^2t$, this allows us to *ignore* the c constant for all the following calculations.

If we perform the change of variables described in Section B.2

$$J = -\int_I^{I_a}\frac{dI}{h(I)} = -\int_I^{I_a}\exp\left[\left(\frac{I_*}{I}\right)^\alpha\right]dI \tag{5.50}$$

where I_a is the absorbing boundary, we change the form of the diffusion equation into a Smoluchowski equation

$$\frac{\partial\hat\rho}{\partial\tau} = \frac{1}{2}\frac{\partial}{\partial J}\frac{\mathrm{d}V(J)}{\mathrm{d}J}\hat\rho + \frac{1}{2}\frac{\partial^2\hat\rho}{\partial J^2} \tag{5.51}$$

where

$$V(J) = -\ln(h(I)) = \left(\frac{I_*}{I(J)}\right)^\alpha \tag{5.52}$$

and

$$\hat\rho(J,\tau) = \rho(I(J),\tau)h(I(J)) \tag{5.53}$$

An approach to study the probability current at an absorbing barrier condition is presented in Appendix B, in particular, we will make use of Eq. (B.23), by considering a linear approximation for $V(J)$ so that we can use the results valid for a linear potential $V(J) = -\nu J$.

We linearise the potential at the position J_0

$$V(J) = \left(\frac{I_*}{I_0}\right)^{\alpha} - \frac{\alpha}{I_*}\left(\frac{I_*}{I_0}\right)^{\alpha+1} \exp\left[\left(-\frac{I_*}{I_0}\right)^{\alpha}\right](J - J_0) \tag{5.54}$$

this allows us to define the value $\nu(I_0)$ as

$$\nu\left(I_0\right) = \frac{\alpha}{I_*}\left(\frac{I_*}{I_0}\right)^{\alpha+1} \exp\left[\left(-\frac{I_*}{I_0}\right)^{\alpha}\right] \tag{5.55}$$

this approximation is effective if I_0 is near the absorbing barrier and $I_* \gg I_a$.

For an initial condition $\delta(I - I_0)$, we get the current of the Smoluchowski Equation (5.51)

$$\mathcal{J}(t) = -\exp\left[-\frac{\left(J_0 + \frac{\nu(I_0)}{2}t\right)^2}{2t}\right]\frac{J_0}{\sqrt{2\pi}t^{3/2}} \tag{5.56}$$

where the main contribution to the current behaviour is due to the dependence on J_0 which can be approximated with

$$J_0 = -\int_{I_0}^{I_a} \exp\left[\left(\frac{I_*}{I}\right)^{\alpha}\right] dI \simeq -\exp\left[\left(\frac{I_*}{I_0}\right)^{\alpha}\right] \tag{5.57}$$

Alternative form for the diffusion coefficient

As we saw in Subsection 5.2.1, we might be interested in interpolating Nekhoroshev-like diffusions combined with a polynomial term I^{β}, resulting in the following (we neglect already the c term)

$$D(I) = I^{\beta}\exp\left[-2\left(\frac{I_*}{I}\right)^{\alpha}\right] \tag{5.58}$$

which leads us to

$$h(I) = I^{\beta/2}\exp\left[-\left(\frac{I_*}{I}\right)^{\alpha}\right] \tag{5.59}$$

Executing the previous calculations on this new $h(I)$, we end up with the change of variable

$$J_0 = -\int_{I_0}^{I_a} I^{-\beta/2} \exp[(I_*/I)^\alpha] \tag{5.60}$$

and the corresponding potential

$$V(J) = -\ln(h(I)) = -\frac{\beta}{2}\ln(I(J)) + \left(\frac{I_*}{I(J)}\right)^\alpha \tag{5.61}$$

and, linearising it at J_0

$$\begin{aligned} V(J) = &-\frac{\beta}{2}\ln(I_0) + \left(\frac{I_*}{I_0}\right)^\alpha - \\ &- \left[\frac{\beta}{2I_0} + \frac{\alpha}{I_*}\left(\frac{I_*}{I_0}\right)^{\alpha+1}\right] I_0^{\beta/2} \exp\left[\left(-\frac{I_*}{I_0}\right)^\alpha\right](J - J_0) \end{aligned} \tag{5.62}$$

we reach the value for $\nu(I_0)$

$$\nu(I_0) = \left[\frac{\beta}{2I_0} + \frac{\alpha}{I_*}\left(\frac{I_*}{I_0}\right)^{\alpha+1}\right] I_0^{\beta/2} \exp\left[\left(-\frac{I_*}{I_0}\right)^\alpha\right] \tag{5.63}$$

which we can plug with everything else in Eq. (5.56) to obtain a valid approximation for the current at the absorbing barrier.

5.3 A possible experimental procedure

As we reviewed in Sec. 3.3, the modern approach for measuring the Dynamic Aperture of an accelerator consists of applying repeated small kicks to blow-up the beam size until losses, which starts to happen when the beam crosses the DA perimeter, are measured.

This approach proved to be very practical for measuring the amplitude depedence of the stability time. However, it does not provide informations on what happens *inside* the Dynamic Aperture region.

In the diffusive framework we propose for the Dynamic Aperture problem, we are interested in defining a way to measure the local behaviour of the

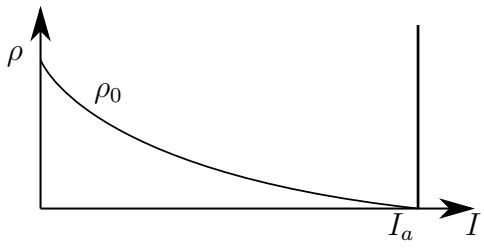
beam at the various region of the phase space. In particular, we are interested in comparing the analytical results with an experimental procedure for measuring the local diffusion coefficient $D(I)$ for a given value of the action variable I.

A possible way to do so is to obtain a valid measure of the first passage time for different values of I, and then connect this measurements with the expected average first passage time given by Eq. (5.44).

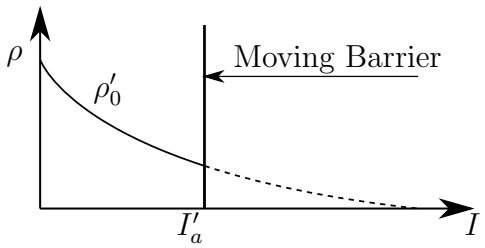
Such measurement can be obtained by applying a fast double-movement procedure to the beam collimators (see Fig. 5.2 for the detailed procedure). This procedure allows us to throw away part of the beam distribution, leaving a consistent gap between the remaining part and the absorbing barrier. The 'cutting point' of the beam can be approximated as a fast logistic damping to zero.

The time spent between this double-movement procedure and a non-zero current measurement at the absorbing barrier gives us a first passage time $t(I)$ for a precise region in the phase space.

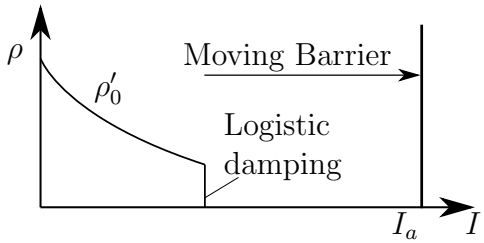
Iterating such procedure for different positions in the phase space can allow us to obtain an interpolating function for $t(I)$, which can give us the informations we are looking for about the local diffusion coefficients.



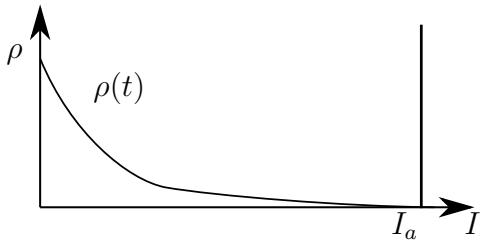
(a) The initial setup of the beam is configured as an exponential distribution ρ_0 and the collimator is at the standard position I_a.



(b) In an initial phase, the collimator is kept very close to the beam, to the point that we obtain is a loss of the tail of the beam and end up with a new initial distribution ρ'_0, cutted at the position I'_a. This phase is fast enough so that we do not have a relaxation of the distribution.



(c) Immediatly after, the collimator is brought back to the original position I_a. If the procedure is fast enough, we can consider the starting condition ρ'_0 with a logistic damping in I'_a.



(d) After a certain period with no beam losses, dependent on the chosen positions for the collimator, we measure a current at the absorbing barrier I_a at the time t. This gives us a first passage time for the position I'_a.

Figure 5.2: Scheme of the experimental procedure.

Chapter 6

Numerical results

6.1 Comparison between discrete diffusion and FP process

In order to inspect the validity of the Fokker-Planck approximation for a stochastic symplectic map, presented in Subsection 5.2.1, we consider an ensemble of 10^5 particles in a map with the numerical parameters given in Table 6.1, and we perform on them 10^4 iterations. The particles are initially in a gaussian initial condition with $I_{\text{mean}} = 3.0$ and $I_\sigma = 0.2$ (see Fig. 6.1).

These iterations are performed with the 3 different kind of noise presented, with various coefficient values, in order to test different regimes of chaoticity. The results are then compared with the results of the Crank-Nicolson integration of the corresponding Fokker-Planck approximation of the map.

ω_0	0.7
ω_1	1.3
ω_2	2.1
ϵ	0.5
x_*	4.0
x_{a}	3.0

Table 6.1: Numerical parameters utilized for the numerical simulation of a stochastic symplectic map like in Subsection 5.2.1.

Baseline result. In Fig. 6.2 we can observe comparison between the ensemble evolved with a different realization of independent random noise for each individual particle. This is exactly the type of noise we utilized as basis for our theory and, in fact, we can see how the Crank-Nicolson integration follows almost exactly the evolution of the ensemble.

Correlated random noise. Differently to the baseline case, here we are assigning to every particle the same noise realization and, moreover, we also consider incresing levels of correlation γ for the random noise itself.

In Fig. 6.3 (a, b), we see the $\gamma = 0$ case, which is basically the same noise of the baseline scenario, and we can see how the FP approximation still holds up even when every particle has the same noise realization (we just observe some fluctuations in the distribution).

In the other plots of Fig. 6.3, we can see how, for increasing values of γ, these fluctuations starts to become more and more significant. This starts to severely break the validity of our approximation for high values of γ, as we can see in Fig. 6.4, where for the highest values of γ we have no resemblance with the Fokker-Planck approximation.

This implies that for high noise correlation values it is necessary to reformulate adequately the approximation itself by executing a proper time rescaling and by including the correlation terms in other parts of the approximation.

Standard map noise. Since a random noise is not a 'proper tool' for making physical models, we are interested in the validity of our approximation when we replace the random noise with the deterministic chaotic motion of a standard map.

In Figs. 6.5 and 6.6, we see the effect of such 'deterministic noise' on the ensemble for different values of the K coefficient of the standard map.

In particular, we can see in Fig. 6.5 how high values of K, corresponding to very chaotic standard maps, returns an ensemble evolution which can be properly approximated with a Fokker-Planck process. In Fig. 6.6, instead, we see how the low chaoticiy of the map ends up completely breaking our approximation.

6.2 First passage times

In this Section we shall inspect the first preliminary results obtained by simulating the proposed experimental procedure presented in Sec. 5.3. We simulated a Fokker-Planck process characterized by a Nekhoroshev diffusion

$$D(I) = c \exp\left[-2\left(\frac{I_*}{I}\right)^{\alpha}\right] \tag{6.1}$$

the process we simulate is characterized by the numerical parameters presented in Table 6.2, and we integrate it with the Crank-Nicolson scheme for 10^5 steps with the time delta set at $k = 10^3$.

I interval	$[0, 5]$
c	1.0
I_*	30.0
α	1

Table 6.2: Numerical parameters for a Fokker-Planck process with Nekhoroshev diffusion.

The initial distribution we consider is a uniform distribution $\rho_0(I) = 1$ to which we applied a logistic damping at 50 different points over the interval $[3.5, 4]$ (see Fig. 6.7(a) for reference).

For each different damped initial condition, we considered the current evolution step after step (see Fig. 6.7(b)). It is possible to observe in Fig. 6.7(c) how each different damping position causes different times of zero measured current before the ramp up that characterize the currents themselfs.

Finally, in Fig. 6.7(d), we display a log scale in which we compare 3 time measures for each damping point analyzed I_0:

1. The analytical formula for the expected average first passage time, presented in Eq. (5.37), computed for the starting point I_0.

2. The time at which we measured the maximum peak of the current in our Crank-Nicolson integration.

3. The time at which we measured the first non-zero current in our Crank-Nicolson integration. This would be the actual observable we could be able to obtain from our proposed experimental procedure.

From what we can see from this plot, we can observe how these three measures appear to differ only for a multiplicative constant. This could imply an interesting way to extrapolate information about the local diffusion coefficient (for example by multiplying the measured time with a constant and then apply the inverse formula of Eq. (5.37)).

This preliminary result will be inspected in future reaserch, as well as a proper background theory in order to quantify properly this behaviour.

6.3 Current interpolation

In this Section we analyze the accuracy of the equations presented in Subsection 5.2.3 for estimating the current at the absorbing barrier for a Fokker-Planck process with Nekhoroshev and Nekhoroshev-like diffusion coefficients. The current we compare the equations with is computed with the Crank-Nicolson scheme (Appendix A).

More specifically, we are interested in evaluating the utilization of these equations for an interpolating procedure, in order to retrieve α, I_* and β, that characterize the diffusive behaviour of the process (we expect to exclude the parameter c after a normalization procedure).

We distinguish two different scenarios, one in which we have a pure Nekhoroshev diffusion, i.e. $\beta = 0$, and one in which we have a Nekhoroshev-like diffusion which also presents a polynomial term, i.e. $\beta \neq 0$.

The two Fokker-Planck processes we integrate are characterized by the numerical parameters persented in Tab. 6.3 and, as initial condition, we consider an uniform distribution $\rho_0(I) = 1$ combined with a logistic damping at $I = 4.40$. The logistic damping is chosen steep enough so that we can consider just an uniform distribution for $[0, I]$, i.e. the initial distribution goes from 1 to zero in an interval smaller than $\Delta I = 0.01$. It is possible to observe the evolution of these two separated processes in Fig. 6.8 and 6.10.

In Fig. 6.8(b) and 6.10(b) we present a direct comparison between the measured current and the corresponding estimated current obtained by utilizing our analytical formula stated in Eq. (5.56), with the exact same parametes utilized for the Fokker-Planck process simulated with the Crank-Nicolson scheme.

In Fig. 6.9, we slightly modify either the value I_* or α and keep the other one correct in order to evaluate the sensitivity of the current obtained from the equation, and perform the same comparison with the integrated Fokker-

I interval	$[0, 5]$
I_*	30.0
α	0.9
β	0.0

(a)

I interval	$[0, 5]$
I_*	35.0
α	0.9
β	1.5

(b)

Table 6.3: Numerical parameters for the 2 Fokker-Planck processes utilized for the current interpolation studies.

Planck process with the parameters given by Tab. 6.3(a). The resulting analytical curves are then compared with the original curve by calculating a L^2 norm of the difference. This norm is then normalized over the L^2 norm of the simulated current itself. This gives us a relative error of the interpolation.

In Fig. 6.11, we execute the same procedure but for the integrated Fokker-Planck process with the parameters given by Tab. 6.3(b), also analyzing how modifying the β exponent affects our formula.

From these results, we can see how our formula appears to be adequately accurate to replicate the behaviour of a numerical simulated current. Moreover, we can also see how it is strongly affected by small changes in the parameters. This implies that, if we know the 'correct' value of every parameter minus one, it is possible to extrapolate the value of the remaining one.

More specifically, we observe that our formula describes well a current with $\beta = 0$, despite a little underestimation, given by the linearization we performed on the potential. This underestimation becomes more consistent for the case with $\beta = 1.5$ and further analysis is reuqired before we can decide to use it for actual interpolation procedures.

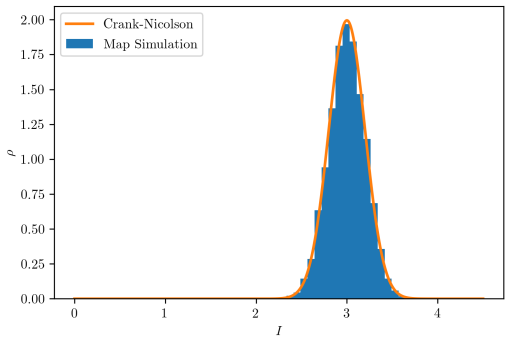
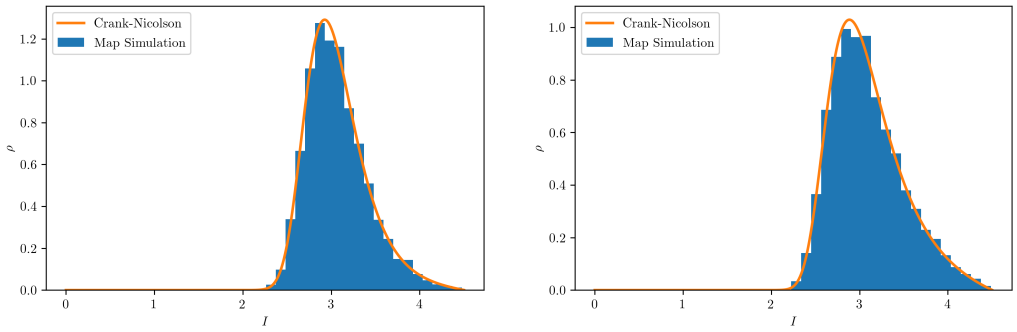
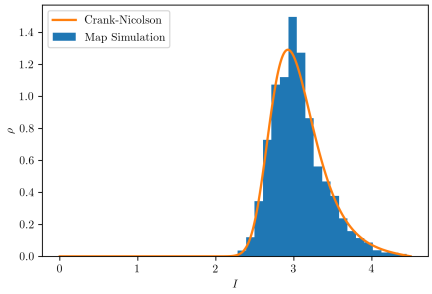


Figure 6.1: Initial distribution ($t = 0$) for the ensemble of 10^5 particles.

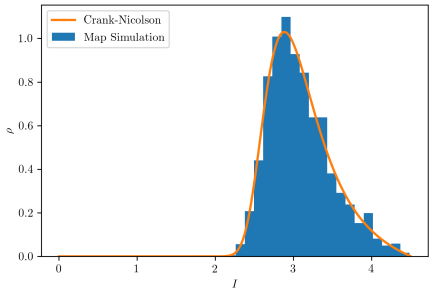


(a) ($t = 500$) (b) ($t = 1000$)

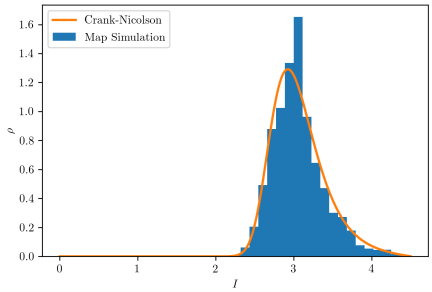
Figure 6.2: Ensemble of 10^5 particles iterated with independent random noise with $\gamma = 0$. Map parameters are given in Tab. 6.1. Each particle has its own noise realization. The Fokker-Planck approximation appears to be valid.



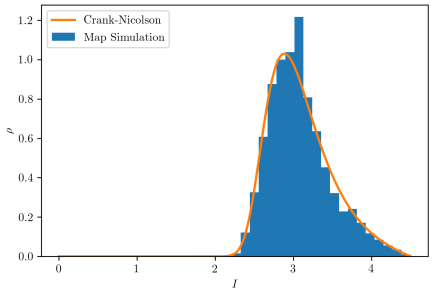
(a) $(t = 500, \gamma = 0.0)$



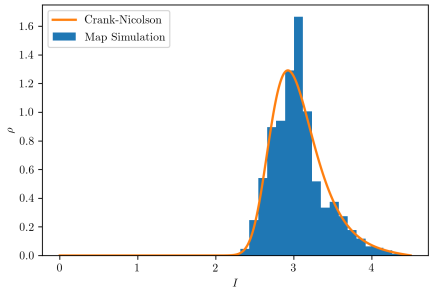
(b) $(t = 1000, \gamma = 0.0)$



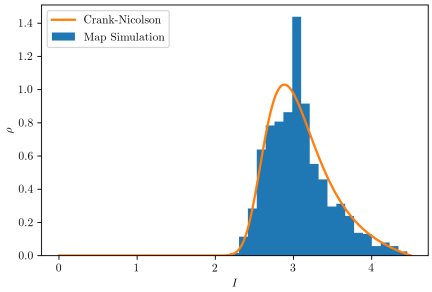
(c) $(t = 500, \gamma = 0.1)$



(d) $(t = 1000, \gamma = 0.1)$

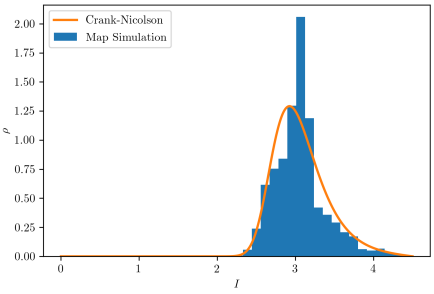


(e) $(t = 500, \gamma = 0.25)$

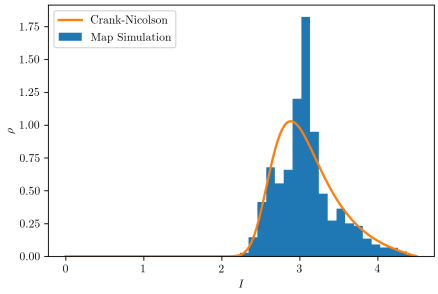


(f) $(t = 1000, \gamma = 0.25)$

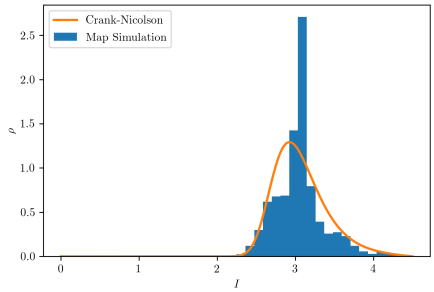
Figure 6.3: Ensemble of 10^5 particles iterated with random noise with various values of γ. Map parameters are given in Tab. 6.1. All particles are provided with the same noise realization. For higher values of γ we start to see the Fokker-Planck approximation breaking.



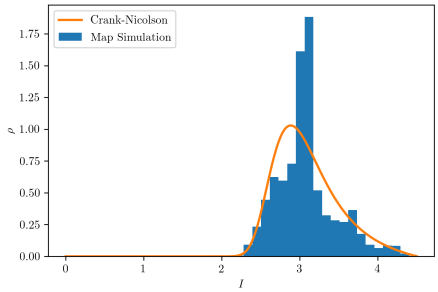
(a) $(t = 500, \gamma = 0.4)$



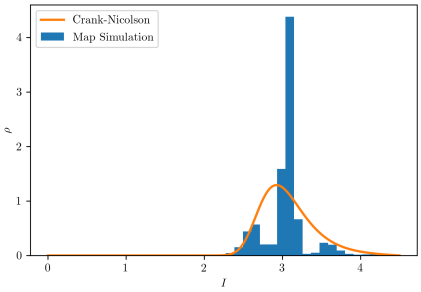
(b) $(t = 1000, \gamma = 0.4)$



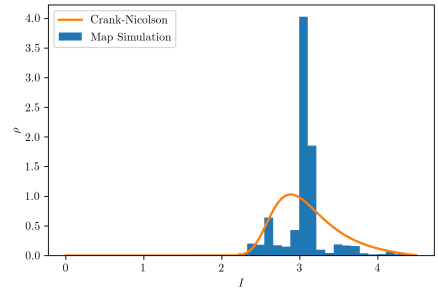
(c) $(t = 500, \gamma = 0.5)$



(d) $(t = 1000, \gamma = 0.5)$

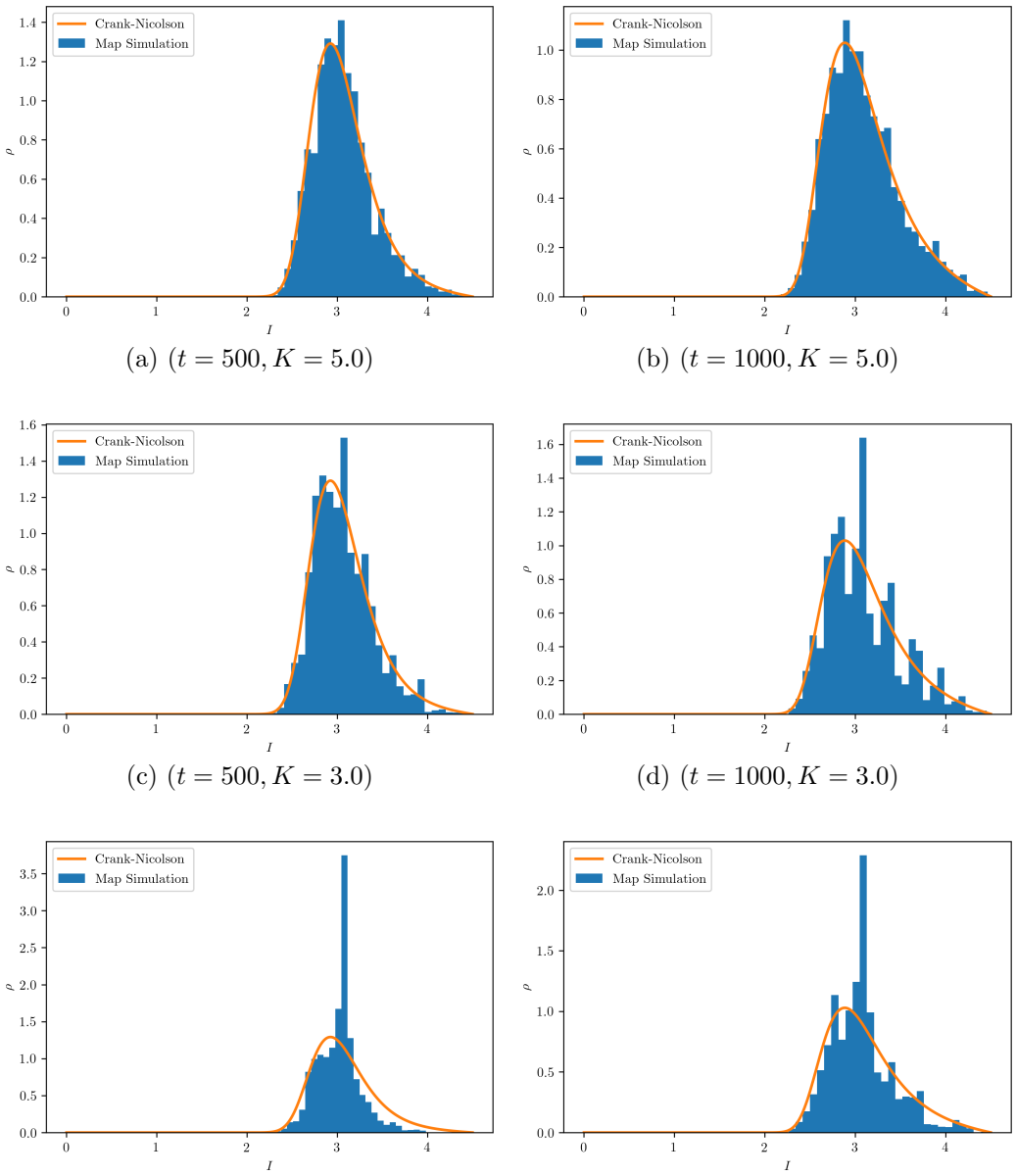


(e) $(t = 500, \gamma = 0.75)$



(f) $(t = 1000, \gamma = 0.75)$

Figure 6.4: Ensemble of 10^5 particles iterated with random noise with various values of γ. Map parameters are given in Tab. 6.1. All particles are provided with the same noise realization. For these high values of γ we see how the Fokker-Planck approximation is completely broken.

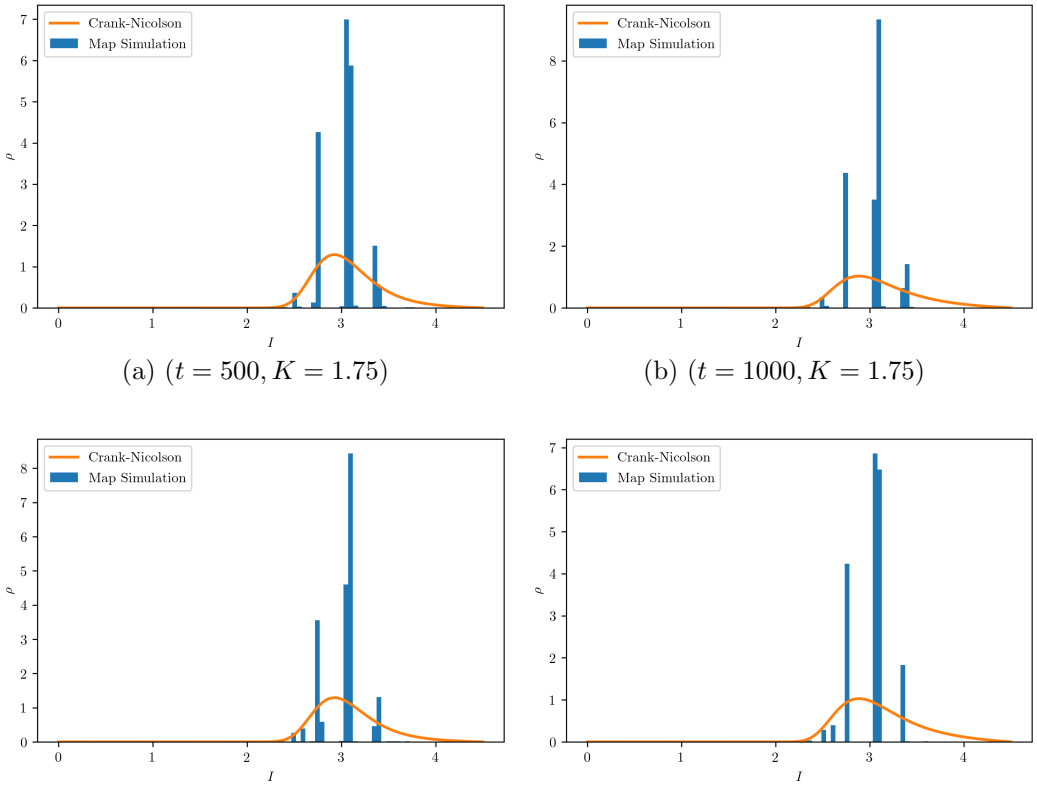


(a) $(t = 500, K = 5.0)$ (b) $(t = 1000, K = 5.0)$

(c) $(t = 500, K = 3.0)$ (d) $(t = 1000, K = 3.0)$

(e) $(t = 500, K = 2.0)$ (f) $(t = 1000, K = 2.0)$

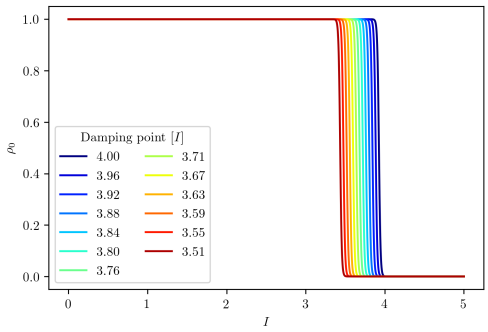
Figure 6.5: Ensemble of 10^5 particles iterated with deterministic noise obtained from a standard map with various values of K. Map parameters are given in Tab. 6.1. We can see how the Fokker-Planck approximation is valid for high enough values of K, i.e. highly chaotic standard maps.



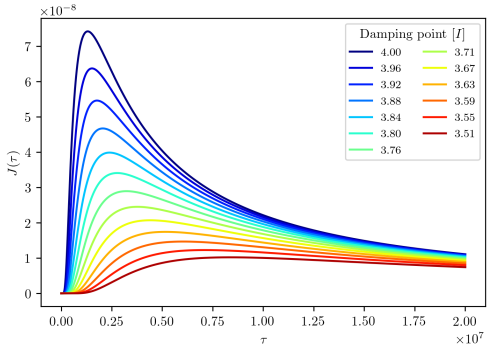
(a) $(t = 500, K = 1.75)$ (b) $(t = 1000, K = 1.75)$

(c) $(t = 500, K = 1.5)$ (d) $(t = 1000, K = 1.5)$

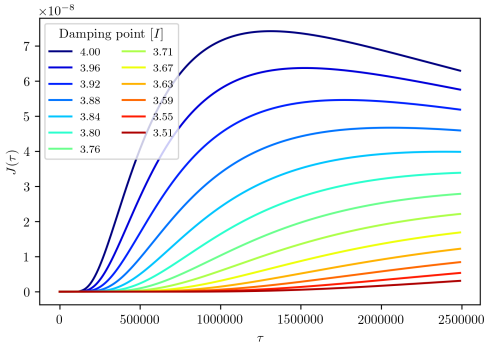
Figure 6.6: Ensemble of 10^5 particles iterated with deterministic noise obtained from a standard map with various values of K. Map parameters are given in Tab. 6.1. We can see how the Fokker-Planck approximation is completely broken for low enough values of K, i.e. weakly chaotic standard maps. More specifically, we observe a severe grouping of particles into narrow regions that will require further investigations.



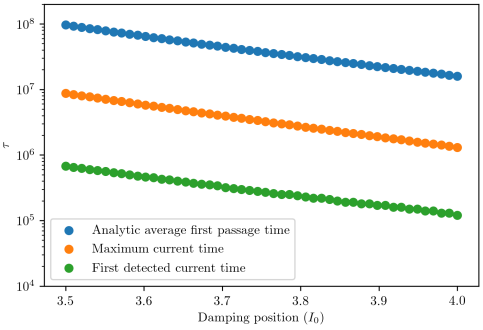
(a) Display of some damped initial conditions considered.



(b) Measured current at absorbing barrier for different damping points.

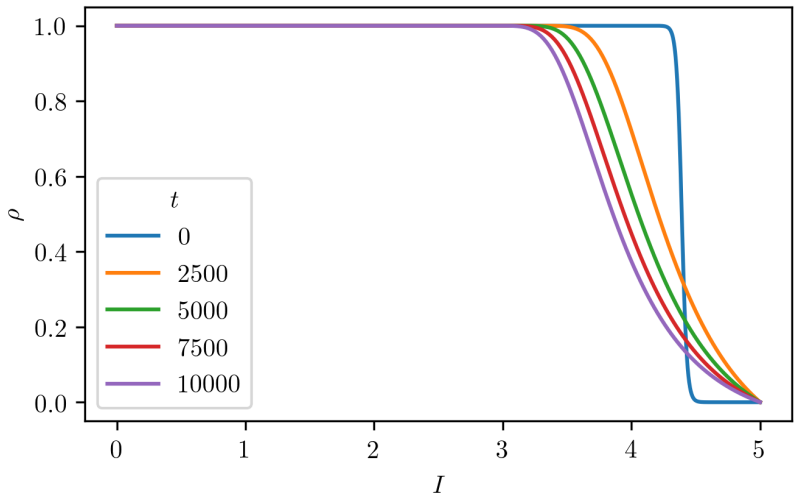


(c) Zoom-in of (b) to better observe the ramp-up of the various currents.

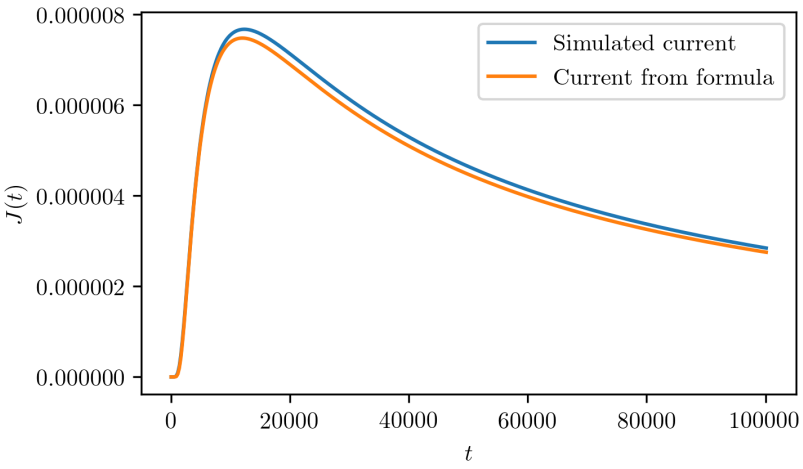


(d) Comparison of times. The blue dots are the analytical estimation of the average first passage time, the orange dots are the times at which maximum current was measured in the FP process, green dots are the times at which the first current was detected.

Figure 6.7: Analysis of multiple damped initial conditions for a Fokker-Planck process with Nekhoroshev diffusion and parameters given in Tab. 6.2. Figures (a,b,c) only show 13 elements out of 50 to improve readability.

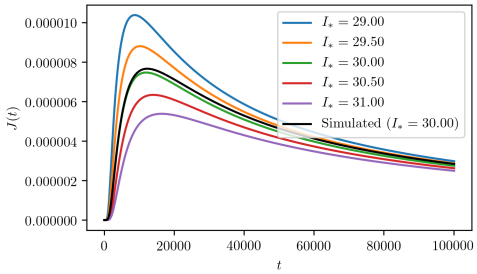


(a) Evolution of $\rho(I, t)$ for different times.

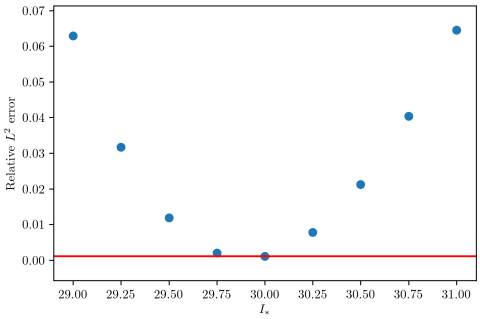


(b) Comparison between the simulated current and the estimated current obtained from Eq. (5.56).

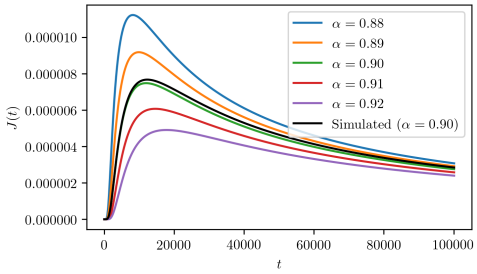
Figure 6.8: Time evolution of the Fokker-Planck process with parameters given by Tab. 6.3(a) and basic comparison with the corresponding current formula.



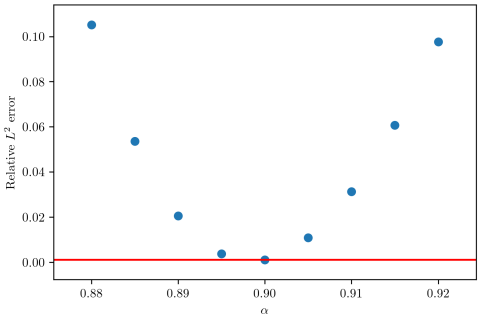
(a) Comparison between the simulated current (black curve) and different current estimations obtained from Eq. (5.56) for different I_* parameters.



(b) Relative L^2 norm of the difference between the various curves in Subfig.(a) and the actual current.

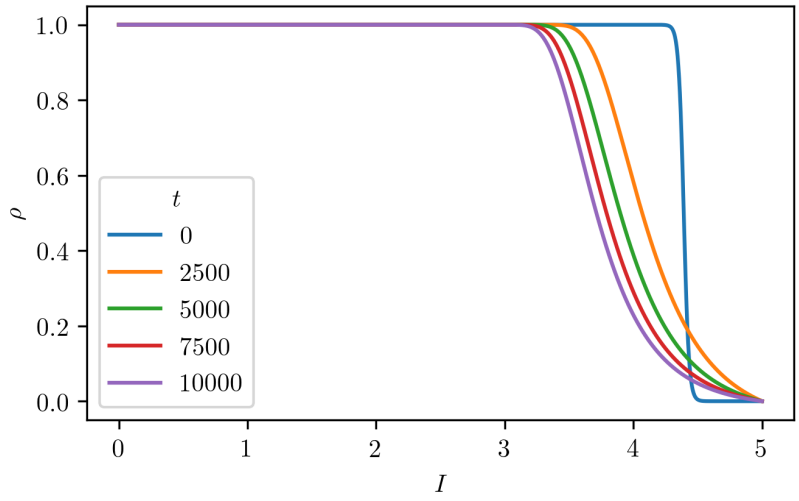


(c) Comparison between the simulated current (black curve) and different current estimations obtained from Eq. (5.56) for different α parameters.

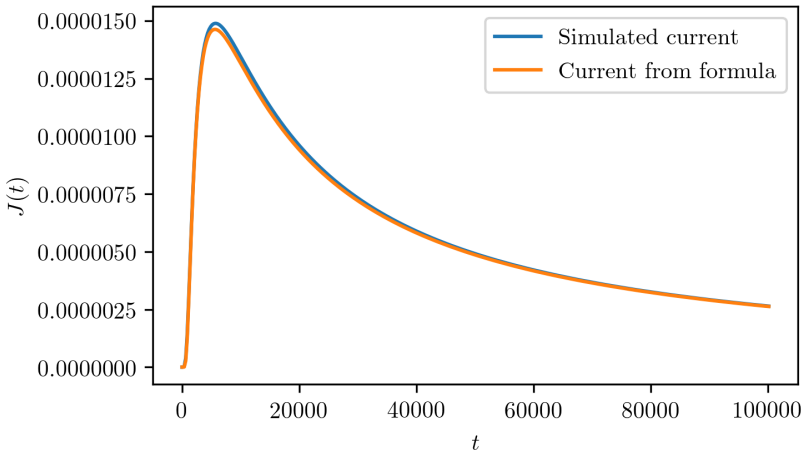


(d) Relative L^2 norm of the difference between the various curves in Subfig.(c) and the actual current.

Figure 6.9: Current interpolation performances and comparisons for the Fokker-Planck process with parameter given by Tab. 6.3(c) and the actual current.

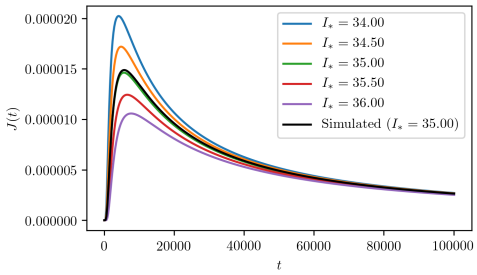


(a) Evolution of $\rho(I, t)$ for different times.

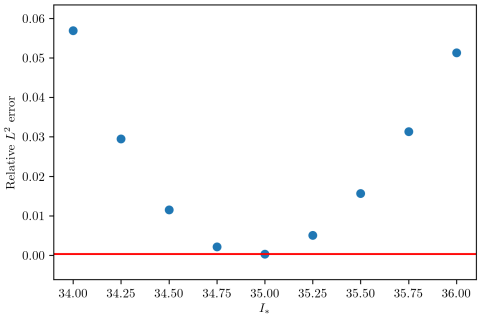


(b) Comparison between the simulated current and the current estimation obtained from Eq. (5.56).

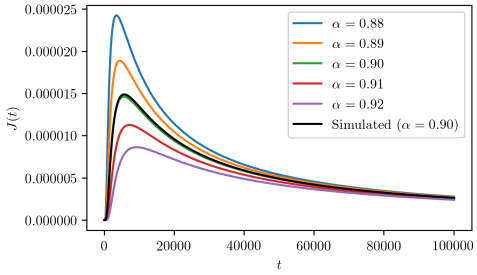
Figure 6.10: Time evolution of the Fokker-Planck process with parameters given by Tab. 6.3(b) and basic comparison with the corresponding current formula.



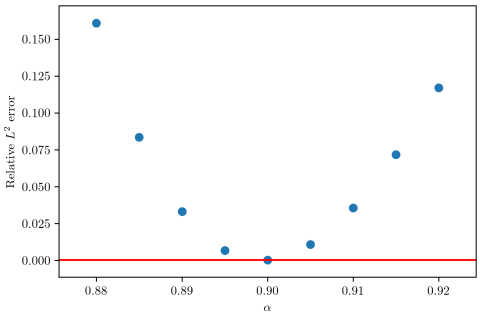
(a) Comparison between the simulated current (black curve) and different current estimations obtained from Eq. (5.56) for different I_* parameters.



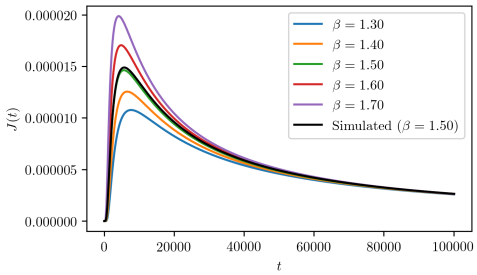
(b) Relative L^2 norm of the difference between the various curves in Subfig.(a) and the actual current.



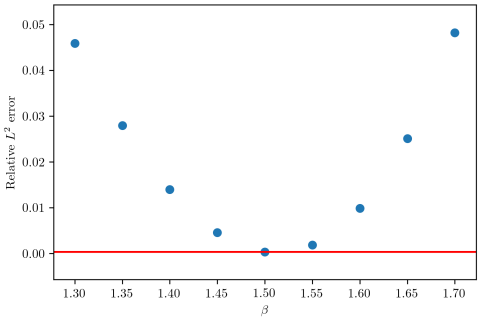
(c) Comparison between the simulated current (black curve) and different current estimations obtained from Eq. (5.56) for different α parameters.



(d) Relative L^2 norm of the difference between the various curves in Subfig.(c) and the actual current.



(e) Comparison between the simulated current (black curve) and different current estimations from Eq. (5.56) for different β parameters.



(f) Relative L^2 norm of the difference between the various curves in Subfig.(e) and the actual current.

Figure 6.11: Current interpolation performances and comparisons for the Fokker-Planck process with parameter given by Tab. 6.3(b).

Chapter 7

Conclusions

In this Master Thesis, we inspected the main characteristic of a diffusion framework for the non-linear betatronic motion in circular accelerators. More specifically, we began to develop a number of 1D tools for analysing the characteristic of the framework and elaborate procedures for future comparisons with experimental data.

We can summarize our results in three main points:

- We have performed some initial comparisons between stochastic symplectic maps of various nature and the corresponding Fokker-Planck approximation, defining a baseline result and laying down some initial considerations on the quality of the approximation.

- We have presented a first idea of an interpolation procedure for the probability current generated at an absorbing barrier by a Fokker-Planck process with a Nekhoroshev-like diffusion coefficient. This procedure will be refined in future works.

- We have proposed a possible experimental procedure to inspect the local diffusive behaviours of a particle beam, based on the analysis of first passage times, and we have performed a first simple simulation to inspect the possibilities of such procedure.

More specifically, we have observed how the stochastic symplectic maps can be adequately approximated with a Fokker-Planck process when coupled with a 'chaotic enough' noise. However, we are interested in understanding how we can modify this approximation in order to be able to take into account

deterministic noises with strong correlation values. This will be necessary when coupling symplectic maps with chaotic maps that, for example, mimic tune ripple effects.

Moreover, we have showed that our analytic estimation of a Fokker-Planck current at an absorbing barrier condition is good at estimating the actual current of the process, computed with a Crank-Nicolson integration of the process. However, before we can say that it can be used for an interpolation procedure, further analysis are required, especially in order to verify if the small underestimation we observed can interfere consistently with the fitting procedure of the parameters.

Finally, we have performed a first numerical simulation of our proposed experimental procedure and found first promising result that suggest that this might be a good path to follow for developing a valid experimental procedure for testing our diffusive framework. This first result must be inspected in order to certify its robustness at least on a theoretical point of view. Moreover, it will be necessary to test whether or not this result can be replicated in the 2D domain.

In general, these first results are promising enough to justify further investigations on this topic, as most of the expected theoretical results of this diffusive framework are observed in our numerical simulations. However, we have also encountered some difficulties that will need to be addressed in future analysis, in order to make this diffusive framework reliable enough for more advanced implementations.

7.1 Future work

This is the conclusion of the thesis, but not of the research work related to this topic. The next steps aim to move this framework from the theoretical domain to a fully-fledged proposal that can be implemented adequately in particle-tracking simulation codes (as for the Dynamic Aperture interpolation part) and in particle accelerators (as for experimental protocols for measuring local diffusive behaviours).

To do so, our models will be compared against more and more realistic one-turn maps and simulation data (obtainable with SixTrack) until we will be able to compare the actual experimental data taken at the LHC during several experimental sessions.

Moreover, the final form of this diffusive framework will be required to be

implemented and validated adequately for 2D scenarios. This will imply a lot of further analysis in order to inspect the additional complexity induced by the increase in the phase space dimensionality (we are already developing tools for this topic, such as a valid 2D implementation of a Crank-Nicolson scheme, presented in Appendix A.2).

Appendix A

The Crank-Nicolson method

For executing the numerical integration of the Fokker-Planck equation, we utilized the Crank-Nicolson method, which guarantees an $\mathcal{O}(k^2, h^2)$ accuracy, i.e. second order accuracy on both space and time, and unconditional stability. In this appendix we shall present the method and the implementation details. For more insights about this topic we invite the reader to consult Ref. [47].

Given a partial differential equation

$$\frac{\partial u}{\partial t} = F\left(u, x, t, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}\right) \tag{A.1}$$

the Crank-Nicolson (CN) method can be presented as a combination of the Forward Euler method with the Backward Euler method. Of course, we are not talking of an 'average' combination between these two methods, since we are dealing with implicit dependences between the various solutions:

$$\begin{aligned}
&\frac{u_m^{n+1} - u_m^n}{k} = F_m^n\left(u, x, t, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}\right) && \text{(Forward)} \\
&\frac{u_m^{n+1} - u_m^n}{k} = F_m^{n+1}\left(u, x, t, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}\right) && \text{(Backward)} \\
&\frac{u_m^{n+1} - u_m^n}{k} = \frac{1}{2}\left[F_m^n\left(u, x, t, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}\right) + F_m^{n+1}\left(u, x, t, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}\right)\right] && \text{(A.2)}
\end{aligned}$$

where u_m^n indicates the value of u in the position m at the time n on the discretised space-time lattice.

Since the CN Eq. (A.2) is an implicit method, in order to obtain the 'next' value of u it is necessary to solve for each step a system of linear equations.

There are two main advantages in using this method for integrating a PDE, instead of using an explicit method. Firstly, the CN method is a second-order method both in space and time, meaning that the truncation error in the scheme is in the order of $\mathcal{O}(k^2, h^2)$, where k is the discrete time interval and h the discrete space interval. Secondly, the Crank-Nicolson method is unconditionally stable, meaning that h and k have no bounds in the values they can take without having the numerical error growing significantly step after step.

The simplest example of application of the CN method is to the heat equation with absorbing boundary conditions. Let us define

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \qquad 0 < x < 1, t > 0 \tag{A.3}$$

with initial condition and boundaries

$$u(x, t = 0) = u_0(x) \qquad 0 \leq x \leq 1 \tag{A.4}$$

$$u(0, t) = g_0(t) = 0, u(1, t) = g_1(t) = 0 \qquad t \geq 0 \tag{A.5}$$

Firstly, we consider a discretization over the x interval into $M + 1$ elements with fineness $1/M$, these elements will be labelled from 0 to M. Next, we consider the basic finite difference form of the components of the PDE

$$\frac{\partial u_m}{\partial t} \Rightarrow \frac{u_m^{n+1} - u_m^n}{k} \tag{A.6}$$

$$\frac{\partial^2 u^n}{\partial x^2} \Rightarrow \frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} \tag{A.7}$$

replacing then the terms in scheme Eq. (A.2), we obtain the complete expression (see Fig. A.1 for a stencil representation)

$$\frac{u_m^{n+1} - u_m^n}{k} = \frac{1}{2}\left[\frac{u_{m+1}^n - 2u_m^n + u_{m-1}^n}{h^2} + \frac{u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}}{h^2}\right] \tag{A.8}$$

This particular case can then be simplified in this vectorial form

$$\left(\mathbf{I} - \frac{k}{2h^2}\mathbf{A}\right)\mathbf{u}^{n+1} = \left(\mathbf{I} + \frac{k}{2h^2}\mathbf{A}\right)\mathbf{u}^n + \mathbf{b} \tag{A.9}$$

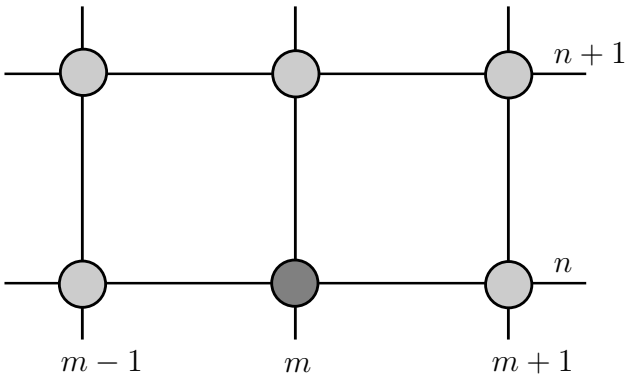


Figure A.1: Stencil representation of the 1D Crank-Nicolson scheme.

where we have

$$\mathbf{A} = \begin{pmatrix} -2 & 1 & 0 & \cdot & \cdot & 0 \\ 1 & -2 & 1 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & 0 & 1 & -2 & 1 \\ 0 & \cdot & \cdot & 0 & 1 & -2 \end{pmatrix} \tag{A.10}$$

$$\mathbf{b} = \frac{r}{2} \begin{pmatrix} u_0^n + u_0^{n+1} \\ 0 \\ \cdot \\ 0 \\ u_M^n + u_M^{n+1} \end{pmatrix} \equiv \frac{r}{2} \begin{pmatrix} g_0(t_n) + g_0(t_{n+1}) \\ 0 \\ \cdot \\ 0 \\ g_M(t_n) + g_M(t_{n+1}) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ 0 \\ 0 \end{pmatrix} \tag{A.11}$$

In general, with the CN scheme there is to solve a linear tridiagonal system for each iteration. Fortunately, the **Thomas algorithm** permits a quick solution of such systems in a complexity time of $\mathcal{O}(M)$, where instead a Gaussian elimination would require a complexity time of $\mathcal{O}(M^3)$.

The Thomas algorithm is very simple: given a tridiagonal system in the form

$$\begin{bmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & a_3 & b_3 & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ 0 & & & a_n & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ \vdots \\ d_n \end{bmatrix} \tag{A.12}$$

we compute two new sets of coefficients c'_i and d'_i with the following equations

$$c'_i = \begin{cases} \dfrac{c_i}{b_i} & ; \quad i = 1 \\ \dfrac{c_i}{b_i - a_i c'_{i-1}} & ; \quad i = 2, 3, \ldots, n-1 \end{cases} \tag{A.13}$$

$$d'_i = \begin{cases} \dfrac{d_i}{b_i} & ; \quad i = 1 \\ \dfrac{d_i - a_i d'_{i-1}}{b_i - a_i c'_{i-1}} & ; \quad i = 2, 3, \ldots, n. \end{cases} \tag{A.14}$$

the final solution can be then obtained via back substitution:

$$\begin{aligned} x_n &= d'_n \\ x_i &= d'_i - c'_i x_{i+1} \qquad ; \; i = n-1, n-2, \ldots, 1 \end{aligned} \tag{A.15}$$

A.1 CN scheme for a Fokker-Planck equation

In this work, the CN scheme is applied a particular kind of Fokker-Planck equation, we will see here the details of the implementation.

Let us consider the following PDE

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\alpha(x)\frac{\partial u}{\partial x}\right) + \beta(x)\frac{\partial u}{\partial x} + \gamma(x)u \tag{A.16}$$

for the construction of the CN integration scheme, it is possible to directly apply the following discretization

$$\frac{\partial u}{\partial t} \Rightarrow \frac{u_m^{n+1} - u_m^n}{k} \tag{A.17}$$

$$\begin{aligned} \frac{\partial}{\partial x}\left(\alpha(x)\frac{\partial u}{\partial x}\right) \Rightarrow & \left(\alpha_{m+\frac{1}{2}}\frac{u_{m+1}^n - u_m^n}{2h^2} - \alpha_{m-\frac{1}{2}}\frac{u_m^n - u_{m-1}^n}{2h^2}\right) + \\ & + \left(\alpha_{m+\frac{1}{2}}\frac{u_{m+1}^{n+1} - u_m^{n+1}}{2h^2} - \alpha_{m-\frac{1}{2}}\frac{u_m^{n+1} - u_{m-1}^{n+1}}{2h^2}\right) \end{aligned} \tag{A.18}$$

$$\beta(x)\frac{\partial u}{\partial x} \Rightarrow \beta_m\frac{u_{m+1}^n - u_{m-1}^n}{4h} + \beta_m\frac{u_{m+1}^{n+1} - u_{m-1}^{n+1}}{4h} \tag{A.19}$$

$$\gamma(x)u \Rightarrow \gamma_m\frac{u_m^n + u_m^{n+1}}{2} \tag{A.20}$$

which leads us to the complete scheme:

$$\begin{aligned}\left(-\frac{\alpha_{m+\frac{1}{2}}}{2h^2}-\frac{\beta_m}{4h}\right)u_{m+1}^{n+1}+\left(\frac{1}{k}+\frac{\alpha_{m+\frac{1}{2}}+\alpha_{m+\frac{1}{2}}}{2h^2}-\frac{\gamma}{2}\right)u_m^{n+1}\\ +\left(-\frac{\alpha_{m-\frac{1}{2}}}{2h^2}+\frac{\beta_m}{4h}\right)u_{m-1}^{n+1}=\\ =\left(\frac{\alpha_{m+\frac{1}{2}}}{2h^2}+\frac{\beta_m}{4h}\right)u_{m+1}^{n}+\left(\frac{1}{k}-\frac{\alpha_{m+\frac{1}{2}}+\alpha_{m+\frac{1}{2}}}{2h^2}+\frac{\gamma}{2}\right)u_m^{n}\\ +\left(\frac{\alpha_{m-\frac{1}{2}}}{2h^2}-\frac{\beta_m}{4h}\right)u_{m-1}^{n}\end{aligned}\tag{A.21}$$

A.2 CN scheme for a 2D Fokker-Plank equation

Since it will be necessary to properly simulate a 2D Fokker-Planck equation for testing adequately 2D DA models, we shall present briefly a time-efficient generalization of the Crank-Nicolson scheme to 2 dimensions, known under the name of Alternating Direction method (ADI), which is part of an even more generalized class of operator-splitting methods.

More specifically, a generalized form of the ADI methods, and a second-order ADI method for the parabolic equation with mixed derivatives (we will always consider absorbing boundary conditions, i.e. Dirichlet boundary conditions always equal to zero).

Firstly, we need to address how we can reshape the 2 dimensional arrays, which represents our 2D space, into one-dimensional arrays, so that we can operate with matrix-vector operations like in the 1D CN schemes seen before. The most simple reshape method to do so, is to follow the *lexicographic ordering*. In this method, given an $M \times M$ 2D vector $\mathbf{U}_{2D}$, we build an M^2

1D vector $\mathbf{U}_{1\mathrm{D}}$ where the following structure holds:

$$\mathbf{U}_{2\mathrm{D}} = \begin{pmatrix} U_{1,1} & U_{1,2} & . & U_{1,M} \\ U_{2,1} & U_{2,2} & . & U_{2,M} \\ . & . & . & . \\ . & . & . & . \\ U_{M,1} & U_{M,2} & . & U_{M,M} \end{pmatrix} \Rightarrow \mathbf{U}_{1\mathrm{D}} = \begin{pmatrix} U_{1,1} \\ U_{2,1} \\ . \\ . \\ U_{M,1} \\ U_{1,2} \\ U_{2,2} \\ . \\ . \\ U_{M,2} \\ . \\ . \\ . \\ U_{1,M} \\ U_{2,M} \\ . \\ . \\ U_{M,M} \end{pmatrix} \tag{A.22}$$

Let us begin by considering the following general form for a partial differential equation:

$$u_t = F \equiv F^{(0)} + F^{(1)} + F^{(2)} \tag{A.23}$$

where $F^{(1)}$ and $F^{(2)}$ are terms which group the derivates of u with respect to x and y only, respectively, and $F^{(0)}$ contains all other non-linear mixed derivatives terms.

In order to compute efficiently the various iterations of the solution vector $\mathbf{U}^n$ of Eq. (A.23), we can make use of the Douglas method for the heat equation in 2D:

$$\begin{aligned} \left(1 - \frac{r}{2}\delta_x^2\right) U'_{ml} &= \left(1 + \frac{r}{2}\delta_x^2 + r\delta_y^2\right) U^n_{ml}, \\ \left(1 - \frac{r}{2}\delta_x^2\right) U^{n+1}_{ml} &= U'_{ml} - \frac{r}{2}\delta_y^2 U^n_{ml} \end{aligned} \tag{A.24}$$

where $r = k/h^2$ and δ_x, δ_y represents the corresponding second order discrete derivation on the x, y coordinate. This scheme can be written in the

equivalent form (we omit for brevity the subscripts m, l):

$$\begin{aligned} W^{(0)} &= U^n + r\delta_x^2 U^n + r\delta_y^2 U^n, \\ W^{(1)} &= W^{(0)} + \frac{1}{2}(r\delta_x^2 W^{(1)} - r\delta_x^2 U^n), \\ W^{(2)} &= W^{(1)} + \frac{1}{2}(r\delta_y^2 W^{(2)} - r\delta_y^2 U^n), \\ U^{(n+1)} &= W^{(2)} \end{aligned} \tag{A.25}$$

The correspondence between notation (A.24) and notation (A.25) is given by the fact that $U' = W^{(1)}$.

Using the notation introduced in (A.23), we can rewrite (A.25) in the following form

$$\begin{aligned} W^{(0)} &= U^n + kF(U^n), \\ W^{(i)} &= W^{(i-1)} + \frac{1}{2}k\left(F^{(i)}(W^{(i)}) - F^{(i)}(U^n)\right), \qquad i = 1, 2; \\ U^{(n+1)} &= W^{(2)} \end{aligned} \tag{A.26}$$

where F comes from the definition in (A.23).

For $F^{(0)} \neq 0$, this method has accuracy $\mathcal{O}(k + h^2)$, for $F^{(0)} = 0$, we have accuracy $\mathcal{O}(k^2 + h^2)$. In order to archive accuracy $\mathcal{O}(k^2 + h^2)$ even for the more generic case, we can make use of the following modified scheme, presented in [48]:

$$\begin{aligned} W^{(0)} &= U^n + kF(U^n), \\ W^{(i)} &= W^{(i-1)} + \frac{1}{2}k\left(F^{(i)}(W^{(i)}) - F^{(i)}(U^n)\right), \qquad i = 1, 2; \\ V^0 &= W^0 + \frac{1}{2}k\left(F^0(W^2) - F^0(U^n)\right) \\ V^{(i)} &= V^{(i-1)} + \frac{1}{2}k\left(F^{(i)}(V^{(i)}) - F^{(i)}(U^n)\right), \qquad i = 1, 2; \\ U^{(n+1)} &= V^{(2)} \end{aligned} \tag{A.27}$$

the unconditional stability of this last method for a generic convection-diffusion equation has been investigated by the authors appropriately for a convection-diffusion equation.

Appendix B

Smoluchowski equation and probability current

In this appendix we shall estimate the probability current at an absorbing boundary for the Smoluchowski equation, by studying its spectral properties. Let us start from the Smoluchowski equation

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\frac{\mathrm{d}V}{\mathrm{d}x}\rho + \frac{D}{2}\frac{\partial^2 \rho}{\partial x^2} \tag{B.1}$$

it is possible to cast the diffusion operator in a self-adjoint form by introducing the distribution $p(x, y)$

$$\rho(x,t) = \exp\left[-\frac{V(x)}{D}\right] p(x,t) \tag{B.2}$$

which gives us the following

$$\frac{\partial p}{\partial t} = \frac{1}{2}\left[\frac{\mathrm{d}^2 V}{\mathrm{d}x^2} - \frac{1}{D}\left(\frac{\mathrm{d}V}{\mathrm{d}x}\right)^2\right] p + \frac{D}{2}\frac{\partial^2 p}{\partial x^2} \tag{B.3}$$

At this point, we want to study the solution for Eq. (B.3). It is possible to expand the solution in the form

$$p(x,t) = \sum_{\lambda} c_\lambda(t)\phi_\lambda(x) \tag{B.4}$$

where

$$c_\lambda(t) = c_\lambda(0)e^{-\lambda t} \tag{B.5}$$

and $\phi_\lambda(x)$ are the eigenfunctions of Eq. (B.3), meaning, functions that satisfy the following:

$$\frac{\mathrm{d}^2\phi_\lambda}{\mathrm{d}x^2} = \frac{1}{D}\left[a(x) - 2\lambda\right]\phi_\lambda(x) \tag{B.6}$$

where

$$a(x) = \left[\frac{1}{D}\left(\frac{\mathrm{d}V}{\mathrm{d}x}\right)^2 - \frac{\mathrm{d}^2V}{\mathrm{d}x^2}\right] \tag{B.7}$$

We can count on the orthogonality and completeness properties, which can be written as

$$\begin{aligned} \int \phi_\lambda(x)\phi_{\lambda'}(x)\,dx &= \delta(\lambda - \lambda') \\ \sum_\lambda \phi_\lambda(x)\phi_\lambda(x') &= \delta(x - x') \end{aligned} \tag{B.8}$$

and then, knowing that the initial condition $\rho(x,0)$ can be written in the new notations as

$$\begin{aligned} \rho(x,0) &= \exp\left[-\frac{V(x)}{D}\right] p(x,0) \\ p(x,0) &= \sum_\lambda c_\lambda(0)\phi_\lambda(x) \end{aligned} \tag{B.9}$$

we can directly write that

$$c_\lambda(0) = \int \exp\left[\frac{V(x)}{D}\right]\rho(x,0)\phi_\lambda(x)\,dx \tag{B.10}$$

if we apply that to the basic case $\rho(x,0) = \delta(x - x_0)$, we can write the expanded solution in the following form

$$\rho(x,t) = \exp\left[-\frac{V(x) - V(x_0)}{D}\right]\sum_\lambda e^{-\lambda t}\phi_\lambda(x_0)\phi_\lambda(x) \tag{B.11}$$

With that said, if we consider an absorbing boundary condition at $x = 0$, that's equivalent to impose $\phi_\lambda(0) = 0$ and gives us a probability current at the boundary which reads

$$\begin{aligned} J(t) &= D\frac{\partial\rho}{\partial x}(0,t) \\ &= \frac{D}{2}\exp\left[-\frac{(V(0) - V(x_0))}{D}\right]\sum_\lambda e^{-\lambda t}\phi_\lambda(x_0)\frac{\partial\phi_\lambda}{\partial x}(0) \end{aligned} \tag{B.12}$$

B.1 Potential in the form of $V(x) = -\nu x$

One specific case we are interested with is the $V(x) = -\nu x$ case, i.e. a constant drift towards the absorbing barrier at $x = 0$, this implies of course that $a(x) = \nu^2/(2D)$ and that the self-adjoint operator

$$-\frac{\nu^2}{2D} + \frac{D}{2}\frac{\partial^2}{\partial x^2} \tag{B.13}$$

is defined negative, i.e. all eigenvalues satisfies $\lambda \leq 0$. Let $-\lambda$ be an eigenvalue, the eigenvector equation reads

$$\frac{\partial^2 \phi_\lambda}{\partial x^2} = -\frac{1}{D}\left[2\lambda - \frac{\nu^2}{D}\right]\phi_\lambda(x) \tag{B.14}$$

where we recall the boundary condition $\phi(0) = 0$. By setting for convenience the value

$$k_\lambda = \sqrt{2\lambda - \frac{\nu^2}{D}} \tag{B.15}$$

we have for Eq. (B.13) the non-trivial solution

$$\phi_\lambda(x) = \frac{1}{\sqrt{\pi}} \sin\left(\frac{k_\lambda}{\sqrt{D}}x\right) \tag{B.16}$$

We have a zero-eigenvalue which corresponds to a trivial zero solution and an upper limit for the negative eigenvalues from the bound that comes from the value $-\nu^2/D$.

We now consider some direct examples to which apply what we have obtained.

If $\nu = 0$. Let us now consider as an example the probability current at the absorbing barrier $x = 0$ for the case $\nu = 0$, which implies $\phi_\lambda(x) \propto \sin\left(\sqrt{(2\lambda/D)}x\right)$, if we apply that to Eq. (B.12) with $\delta(x - x_0)$ as initial condition, all considering a continuous eigenvalue spectrum, we obtain the following

$$\begin{aligned} J(t, x_0) &= \frac{D}{2\pi}\int_0^\infty e^{-\lambda t} \sin\left(\sqrt{\frac{2\lambda}{D}}x_0\right)\sqrt{\frac{2\lambda}{D}}\; d\left(\sqrt{\frac{2\lambda}{D}}\right) \\ &= \frac{D}{\pi}\int_0^\infty e^{-\mu^2 D t/2}\sin(\mu x_0)\mu\, d\mu \end{aligned} \tag{B.17}$$

where we executed the change of variables $2\lambda = \mu^2 D$. At this point we execute the calculation and obtain:

$$\begin{aligned}
J(t,x_0) &= \frac{D}{\pi}\int_0^\infty e^{-\mu^2 Dt}\sin(\mu x_0)\mu\, d\mu \\
&= \frac{D}{2i\pi}\int_0^\infty e^{-\mu^2 Dt+i\mu x_0}\mu\, d\mu - \frac{D}{2i\pi}\int_0^\infty e^{-\mu^2 Dt-i\mu x_0}\mu\, d\mu \\
&= \frac{D}{2i\pi}\int_0^\infty e^{-\mu^2 Dt+i\mu x_0}\mu\, d\mu + \\
&\quad + \frac{D}{2i\pi}\int_{-\infty}^0 e^{-(-\mu)^2 Dt+i(-\mu)x_0}(-\mu)\, d(-\mu) \\
&= \frac{D}{2i\pi}\int_{-\infty}^\infty e^{-\mu^2 Dt+i\mu x_0}\mu\, d\mu \\
&= \frac{D}{2i\pi}\exp\left(-\frac{x_0^2}{2Dt}\right)\int_{-\infty}^\infty \exp\left[-\frac{Dt}{2}\left(\mu - i\frac{x_0}{Dt}\right)^2\right]\mu\, d\mu \\
&= \exp\left(-\frac{x_0^2}{2Dt}\right)\frac{x_0}{(t)^{3/2}\sqrt{2\pi D}}
\end{aligned} \tag{B.18}$$

Where in the last passages we reached a Gaussian integral form exploiting the fact that

$$\begin{aligned}
&\exp\left(\frac{-\mu^2 Dt\alpha \pm 2i\mu x_0}{2}\right) = \\
&\qquad = \exp\left[-\frac{Dt}{2}\alpha\left(\mu \pm i\frac{2x_0}{Dt\alpha}\right)^2\right]\exp\left(-\frac{x_0^2}{2Dt\alpha}\right)
\end{aligned} \tag{B.19}$$

If $\nu \neq 0$. If we instead have a drift field, we can just modify a bit the previous calculations and start with the following equation

$$\begin{aligned}
J(t,x_0) &= \frac{D}{\pi}\exp\left(-\frac{\nu x_0}{D}\right) \\
&\int_0^\infty e^{-\lambda t}\sin\left(\sqrt{\frac{2\lambda}{D}-\frac{\nu^2}{D^2}}x_0\right)\sqrt{\frac{2\lambda}{D}-\frac{\nu^2}{D^2}}\, d\left(\sqrt{\frac{2\lambda}{D}-\frac{\nu^2}{D^2}}\right)
\end{aligned} \tag{B.20}$$

By defining directly $2\lambda - \nu^2/D = k_\lambda^2 = \mu^2 D$, we end up with the relation

$$\begin{aligned}\lambda t \pm i\frac{k_\lambda}{\sqrt{D}}x_0 &= \frac{t}{2}\left(\mu^2 D + \frac{\nu^2}{D}\right) \pm i\mu x_0 \\ &= \frac{\nu^2}{2D}t + \frac{Dt}{2}\left(\mu \pm i\frac{x_0}{Dt}\right)^2 + \frac{x_0^2}{2Dt}\end{aligned} \tag{B.21}$$

which allows us to reach the following Gaussian expression

$$\exp\left[-\frac{\nu^2}{2D}t - \frac{x_0^2}{2Dt}\right]\exp\left[-\frac{Dt}{2}\left(\mu \pm i\frac{x_0}{Dt}\right)^2\right] \tag{B.22}$$

and we end up with the current

$$\begin{aligned}J(t,x_0) &= \exp\left(-\frac{\nu x_0}{D}\right)\exp\left(-\frac{\nu^2}{2D}t - \frac{x_0^2}{2Dt}\right)\frac{x_0}{t^{3/2}\sqrt{2\pi D}} \\ &= \exp\left(-\frac{(x_0+\nu t)^2}{2Dt}\right)\frac{x_0}{t^{3/2}\sqrt{2\pi D}}\end{aligned} \tag{B.23}$$

B.2 Change of variables for Smoluchowski form

We start with the following Fokker-Planck equation

$$\frac{\partial\rho(I,t)}{\partial t} = \frac{1}{2}\frac{\partial}{\partial I}\left[h^2(I)\frac{\partial\rho(I,t)}{\partial I}\right] \tag{B.24}$$

with $I \in [0, I_a], I_a \in \mathbb{R}^+$. We can introduce a new variable $x(I)$ so that

$$\begin{aligned}x &= -\int_I^{I_a} dI' h^{-1}(I') \\ \frac{dI(x)}{dx} &= h(I(x))\end{aligned} \tag{B.25}$$

and, to preserve the measure

$$\rho(I,t)dI = \rho(I(x),t)h(I(x))dx = \rho'(x,t)dx \tag{B.26}$$

The new Fokker-Planck equation reads then

$$\begin{aligned}\frac{\partial\rho'(x,t)}{\partial t} &= \frac{h(I(x))}{2}\frac{dx}{dI}\frac{\partial}{\partial x}\left[h(I(x))^2\frac{dx}{dI}\frac{\partial}{\partial x}\left(\frac{\rho'(x,t)}{h(I(x))}\right)\right] \\ &= -\frac{1}{2}\frac{\partial}{\partial x}\left[\frac{1}{h(I(x))}\frac{dh(I(x))}{dx}\rho'(x,t)\right] + \frac{1}{2}\frac{\partial^2\rho'(x,t)}{\partial x^2}\end{aligned} \tag{B.27}$$

defining then the effective potential

$$V(x) \equiv -\ln(h(I(x))) \tag{B.28}$$

Eq. (B.24) assumes the Smoluchowski form

$$\frac{\partial \rho'}{\partial t} = \frac{1}{2}\frac{\partial}{\partial x}\left[\left(\frac{dV(x)}{dx}\right)\rho'\right] + \frac{1}{2}\frac{\partial^2 \rho'}{\partial x^2} \tag{B.29}$$

where now the domain is $x \in (-\infty, 0]$.

Bibliography

[1] E. Todesco and M. Giovannozzi, "Dynamic aperture estimates and phase-space distortions in nonlinear betatron motion", *Physical Review E*, vol. 53, no. 4, p. 4067, 1996.

[2] M. Giovannozzi, W. Scandale, and E. Todesco, "Inverse logarithm decay of long-term dynamic aperture in hadron colliders", in *Proceedings of the 1997 Particle Accelerator Conference (Cat. No. 97CH36167)*, IEEE, vol. 2, 1997, pp. 1445–1447.

[3] M. Giovannozzi, "Proposed scaling law for intensity evolution in hadron storage rings based on dynamic aperture variation with time", *Physical Review Special Topics-Accelerators and Beams*, vol. 15, no. 2, p. 024 001, 2012.

[4] E. Maclean, R. Tomás, F. Schmidt, and T. Persson, "Measurement of nonlinear observables in the large hadron collider using kicked beams", *Physical Review Special Topics-Accelerators and Beams*, vol. 17, no. 8, p. 081 002, 2014.

[5] E. Maclean, M. Giovannozzi, and R. Appleby, "Innovative method to measure the extent of the stable phase-space region of proton synchrotrons", *Physical Review Accelerators and Beams*, vol. 22, no. 3, p. 034 002, 2019.

[6] A. G., B. A. I., B. O., F. P., L. M., R. L., and T. L., *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*, ser. CERN Yellow Reports: Monographs. Geneva: CERN, 2017. DOI: 10.23731/CYRM-2017-004. [Online]. Available: https://cds.cern.ch/record/2284929.

[7] A. Bazzani, O. Mazzarisi, M. Giovannozzi, and E. Maclean, "Diffusion in stochastically perturbed hamiltonian systems with applications to the recent lhc dynamic aperture experiments", 2019.

[8] N. N. Nekhoroshev, "An exponential estimate of the time of stability of nearly-integrable hamiltonian systems", *Uspekhi Matematicheskikh Nauk*, vol. 32, no. 6, pp. 5–66, 1977.

[9] D. W. Kerst and R. Serber, "Electronic orbits in the induction accelerator", *Phys. Rev.*, vol. 60, pp. 53–58, 1 Jul. 1941. DOI: 10.1103/PhysRev.60.53. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRev.60.53.

[10] E. D. Courant and H. S. Snyder, "Theory of the alternating-gradient synchrotron", *Annals of physics*, vol. 281, no. 1-2, pp. 360–408, 2000.

[11] A. Bazzani, G. Servizi, G. Turchetti, and E. Todesco, *A normal form approach to the theory of nonlinear betatronic motion*, CERN-94-02. CERN, 1994.

[12] S.-Y. Lee, *Accelerator physics*. World scientific publishing, 2018.

[13] K. Steffen, "Basic course on accelerator optics", 1985.

[14] G. Turchetti, *Modelli e Metodi Numerici della Fisica*. Dipartimento di Fisica Astronmia Università di Bologna, 2017.

[15] D. Bambusi, *An introduction to birkhoff normal form*, 2014.

[16] R. D. Maria *et al.*, "SixTrack Version 5: Status and New Developments", in *Proc. 10th International Particle Accelerator Conference (IPAC'19)*, (Melbourne, Australia), ser. International Particle Accelerator Conference, Geneva, Switzerland: JACoW Publishing, Jun. 2019, pp. 3200–3203. DOI: doi:10.18429/JACoW-IPAC2019-WEPTS043. [Online]. Available: http://jacow.org/ipac2019/papers/wepts043.pdf.

[17] A. Wrulich, "Racetrack-a computer code for the simulation of nonlinear particle motion in accelerators", 1984.

[18] H. Grote and F. Schmidt, "Mad-x-an upgrade from mad8", in *Proceedings of the 2003 Particle Accelerator Conference*, IEEE, vol. 5, 2003, pp. 3497–3499.

[19] F. Schmidt, E. Forest, and E. McIntosh, "Introduction to the polymorphic tracking code: Fibre bundles, polymorphic taylor types and" exact tracking"", Tech. Rep., 2002.

[20] M. Hénon, "Numerical study of quadratic area-preserving mappings", *Quarterly of applied mathematics*, pp. 291–312, 1969.

[21] A. Bazzani, P. Mazzanti, G. Servizi, and G. Turchetti, "Normal forms for hamiltonian maps and nonlinear effects in a particle accelerator", *Il Nuovo Cimento B (1971-1996)*, vol. 102, no. 1, pp. 51–80, 1988.

[22] V. I. Arnol'd and A. Avez, "Ergodic problems of classical mechanics", 1968.

[23] A. N. Kolmogorov, "On conservation of conditionally periodic motions for a small change in hamilton's function", in *Dokl. Akad. Nauk SSSR*, vol. 98, 1954, pp. 527–530.

[24] V. I. Arnol'd, "Small denominators and problems of stability of motion in classical and celestial mechanics", *Russian Mathematical Surveys*, vol. 18, no. 6, p. 85, 1963.

[25] J. Möser, "On invariant curves of area-preserving mappings of an annulus", *Nachr. Akad. Wiss. Göttingen, II*, pp. 1–20, 1962.

[26] A. Bazzani, S. Marmi, and G. Turchetti, "Nekhoroshev estimate for isochronous non resonant symplectic maps", *Celestial Mechanics and Dynamical Astronomy*, vol. 47, no. 4, pp. 333–359, 1989.

[27] G. Turchetti, *Number theory and physics berlin–heidelberg*, 1990.

[28] M. Giovannozzi, W. Scandale, and E. Todesco, "Dynamic aperture extrapolation in the presence of tune modulation", *Physical Review E*, vol. 57, no. 3, p. 3432, 1998.

[29] C. E. Montanari, "Evaluation of different scaling laws for dynamic aperture variation over time and intensity evolution in hadron storage rings", Sep. 2018. [Online]. Available: `https://cds.cern.ch/record/2638426`.

[30] N. Gelfand, "Calculations of the dynamic aperture at the tevatron", Tech. Rep., 1984.

[31] O. S. Brüning, W. Fischer, F. Willeke, and F. Schmidt, "Comparison of measured and computed dynamic aperture for the sps and the hera proton ring", *Part. Accel.*, vol. 54, pp. 223–235, 1996.

[32] Y. Luo, M. Bai, J. Beebe-Wang, W. Fischer, A. Jain, C. Montag, T. Roser, S. Tepikian, and D. Trobjevic, "Dynamic aperture evaluation at the current working point for rhic polarized proton operation", in *2007 IEEE Particle Accelerator Conference (PAC)*, IEEE, 2007, pp. 4363–4365.

[33] L. Schachinger and Y.-t. Yan, "Recent ssc dynamic aperture measurements from simulations", Tech. Rep., 1989.

[34] O. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, "Lhc design report, volume 1: The lhc main ring. cern, geneva, 2004", CERN-2004-003-V-1, Tech. Rep.

[35] CERN, Ed. (). Facts and figures about lhc, [Online]. Available: `https://home.cern/resources/faqs/facts-and-figures-about-lhc`.

[36] M. Giovannozzi and E. McIntosh, "Development of parallel codes for the study of nonlinear beam dynamics", *International Journal of Modern Physics C*, vol. 8, no. 02, pp. 155–170, 1997.

[37] M. Giovannozzi, W. Scandale, and E. Todesco, "Prediction of long term stability in large hadron colliders", *Part. Accel.*, vol. 56, no. CERN-LHC-PROJECT-REPORT-045, pp. 195–225, 1996.

[38] M. Albert, G. Crockford, S. Fartoukh, M. Giovannozzi, E. Maclean, A. MacPherson, R. Miyamoto, L. Ponce, S. Redaelli, H. Renshall, F. Roncarolo, R. Steinhagen, E. Todesco, R. Tomás, and W. Venturini Delsolaro, "First Experimental observations from the LHC Dynamic Aperture Experiment.", *Conf. Proc.*, vol. C1205201, no. CERN-ATS-2012-161, TUPPC081. 3 p, May 2012. [Online]. Available: `https://cds.cern.ch/record/1463302`.

[39] S. Cave, R. Maria, M. Giovannozzi, M. Ludwig, A. Macpherson, S. Redaelli, F. Roncarolo, M. Camillocci, and W. Delsolaro, "Experimental Observations from the LHC Dynamic Aperture Machine Development Study in 2012", no. CERN-ACC-2013-0165, 3 p, May 2013. [Online]. Available: `https://cds.cern.ch/record/1591297`.

[40] G. Turchetti, *Dinamica classica dei sistemi fisici*. Zanichelli, 1998.

[41] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt, *Mathematical aspects of classical and celestial mechanics*. Springer Science & Business Media, 2007, vol. 3.

[42] V. I. Arnol'd, "Instability of dynamical systems with many degrees of freedom", in *Doklady Akademii Nauk*, Russian Academy of Sciences, vol. 156, 1964, pp. 9–12.

[43] M. F. Mestre, A. Bazzani, P. M. Cincotta, and C. M. Giordano, "Stochastic approach to diffusion inside the chaotic layer of a resonance", *Physical Review E*, vol. 89, no. 1, p. 012 911, 2014.

[44] J. A. Ellison, "Noise effects in accelerators", in *Beam Measurement*, pp. 428–450. DOI: 10.1142/9789812818003_0019. eprint: https://www.worldscientific.com/doi/pdf/10.1142/9789812818003_0019. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/9789812818003_0019.

[45] R. Cogburn and J. A. Ellison, "A four-thirds law for phase randomization of stochastically perturbed oscillators and related phenomena", *Communications in Mathematical Physics*, vol. 166, no. 2, pp. 317–336, 1994.

[46] B. V. Chirikov, "Research concerning the theory of non-linear resonance and stochasticity", CM-P00100691, Tech. Rep., 1971.

[47] T. Lakoba, *Math 337, numerical diff equations*, 2016.

[48] K. In't Hout and B. Welfert, "Stability of adi schemes applied to convection–diffusion equations with mixed derivative terms", *Applied numerical mathematics*, vol. 57, no. 1, pp. 19–35, 2007.