Hamiltonian systems and Stochastic processes

Armando Bazzani
Department of Physics and Astronomy - Bologna University

March 19, 2026

Chapter 1

Stochastic Hamiltonian Systems

The microscopic deterministic nature of Hamiltonian systems poses a formidable problem when applied to statistical systems where the number of degree of freedoms prevents the solution of the motion equations. A possible approach is to consider the dynamics of a single particle and to introduce the interactions with other particles in an effective way. This approach has to consider some physical constraints: e.g. the existence of a preserved energy of the system during the interactions. We assume that the test particle dynamics is described by a time dependent Hamiltonian system

$$H(x,t) = H_0(x) + H_{int}(x, y(t)) \tag{1.1}$$

where $H_0(x)$ is the average energy of the system associated to a particle in the dynamical state x (it considers the interaction with the other particles as environment) and $H_1(x, y_\alpha)$ is a local interaction potential with a particle in the state $y(t)$. The interaction potential acts for a very small amount of time τ (local collisions) and the states $y(t)$ can be considered a stochastic process (molecular chaos).
If collision time scale τ is much less then the evolution time scale of H_0 we can separate the collision dynamics from the evolution and applying the CLT to get an effective description of the collision effects.
The concept of thermal bath is related to the definition of temperature and the Maximum Entropy Principle: the physical equilibrium is characterized by an exponential distribution of the energy states $H_0(x)$ with a characteristic energy given by the temperature T. Let x any microstate we have the MB equilibrium distribution

$$\rho_s(x) \propto \exp\left(-\frac{H_0(x)}{T}\right)$$

where the proportional constant is the partition function

$$A(T) = \int \exp\left(-\frac{H_0(x)}{T}\right) dx$$

assuming it is convergent. The MB distribution is related to the fluctuation dissipation relation discovered by Einstein: in a physical system we simulate the collision dynamics by

$$\begin{aligned} \dot{q} &= \frac{p}{m} \\ \dot{p} &= -\frac{\partial V}{\partial q} - \gamma p + \sqrt{2m\gamma T}\xi(t) \end{aligned} \tag{1.2}$$

where γ is a friction coefficient for the energy

$$H_0 = \frac{p^2}{2m} + V(q)$$

Using the stochastic dynamics we get

$$\Delta H_0 = -\frac{\gamma}{m}p^2\Delta t + \frac{p}{m}\sqrt{2m\gamma T}\int_t^{t+\Delta t} \xi(s)\,ds + \gamma T \left(\int_t^{t+\Delta t} \xi(s)\,ds\right)^2 + o(\Delta t)$$

The fluctuations $\xi(t)$ are characterized the correlation function:

$$\langle \xi(t+\Delta t)\xi(t)\rangle = \frac{1}{2\tau}e^{-\Delta t/\tau} \to \delta(\Delta t) \qquad \tau \to 0$$

τ is a correlation time scale that $\to 0$ in the white noise limit. An explicit integration gives

$$\int_t^{t+\Delta t} ds \int_t^s du \langle \xi(s)\xi(u)\rangle = \frac{\Delta t}{2} + \frac{\tau}{2}\left(e^{-\Delta t/\tau} - 1\right) \simeq \frac{\Delta t}{2}$$

assuming $\Delta t \gg \tau$. We denote by Δw_t the fluctuation term integrated over a time interval Δt in the white noise limit.
The emergence of dissipation γ follows from the physical conservation laws of collision dynamics. If we consider the average energy evolution

$$\langle \Delta H_0 \rangle = -\frac{2\gamma}{m}\left\langle \frac{p^2}{2m} \right\rangle \Delta t + \gamma T \Delta t$$

in the white noise limit. We observe that the dissipation and the fluctuation term coincides if we define the temperature $\langle p^2/2m \rangle = T/2$ in the equilibrium state. The Einstein relation has deep consequences: if we consider the Fokker-Planck equation for the evolution of the single particle distribution function

$$\frac{\partial \rho}{\partial t} = -D_{H_0}\rho + \gamma\frac{\partial}{\partial p}p\rho + m\gamma T\frac{\partial^2 \rho}{\partial x^2}$$

where D_{H_0} is the Lie derivative for the Hamiltonian H_0 ($D_{H_0}F = \{F, H_0\}$, the stationary solution for the distribution is

$$\rho_s(x) = A^{-1}(T)\exp\left(-\frac{H_0(q,p)}{T}\right) \qquad A(T) = \int \exp\left(-\frac{H_0(q,p)}{T}\right) dq\,dp \tag{1.3}$$

The momentum distribution is Gaussian and it follows

$$\int \frac{p^2}{2m}\rho_s(q,p)\,dq\,dp = T$$

The stochastic dynamics describes an ensemble of colliding particles in a thermal bath of temperature T, so that the FP equation describes the relaxation process. The equilibrium state is consistent with a thermodynamic description: let $\beta = T^{-1}$ by definition we get

$$\frac{\partial}{\partial \beta} \ln A(\beta) = -A^{-1}(\beta) \int H_0 \exp\left(-\frac{H_0(q,p)}{T}\right) dq\, dp = \langle H_0 \rangle = E(T)$$

where $E(T)$ is the internal energy. The we can identify the Helmholtz Free Energy

$$F = E - TS \qquad -T \ln A(T) = F$$

where S is the Entropy. We recall that F defines the equilibrium states of the system in a thermal bath, so that if $F(\eta)$ where η is a parameter that classifies the microstates, one can study the phase transitions using the bifurcation theory on $F(\eta)$. The definition of Entropy follows

$$S(T) = -\frac{\partial F}{\partial T} = \ln A(T) + \frac{E}{T} = -\int \rho_s \ln \rho_s \, dq\, dp \tag{1.4}$$

since

$$\ln \rho_s = -\ln A(T) - \frac{H_0}{T}$$

Remark: in the entropy definition it is understood that all the microstates are equivalent so that the invariant measure for the Hamiltonian dynamics is the Lebesgues measure. The distribution $\rho_s(T)$ is the maximum entropy distribution with the constraint that $\langle H_0 \rangle = \bar{E}$. Using the Lagrangian multiplier methods we get the condition

$$\delta \left[-\int (\ln \rho_s + \lambda H_0)\rho_s \, dq\, dp \right] = -\int (\ln \rho_s + \lambda H_0)\delta\rho_s \, dq\, dp$$

and it follows

$$\rho_s \propto \exp(-\lambda H_0)$$

where $\lambda = 1/T$ has to be computed inverting the relation $E(\lambda) = \bar{E}$.
To generalize this framework we write the stochastic dynamics of a thermal bath in a more general form (we set $m = 1$)

$$\dot{x} = D_{H_0} x + \sqrt{2T\gamma} D_{H_1} x \xi(t) - \gamma\{H_0, H_1\} D_{H_1} x \tag{1.5}$$

where $H_1 = q$. We compute the energy dissipation

$$\langle \Delta H_0 \rangle = -\gamma \langle \{H_0, H_1\}^2 \rangle \Delta t$$

and the energy fluctuations

$$\left\langle (H_0 - \langle H_0 \rangle)^2 \right\rangle = 2T\gamma \langle \{x, H_1\}^2 \rangle \Delta t$$

We have the generalized condition

$$\left\langle \frac{\{H_0, H_1\}^2}{2} \right\rangle = \langle \{x, H_1\}^2 \rangle T$$

which defines the temperature T according to the Einstein relation. The FP equation associated to the Hamiltonian (**??**) in the Stratonovich interpretation is written in the form

$$\frac{\partial \rho}{\partial t} = -D_{H_0}\rho + \gamma D_{H_1}\{H_0, H_1\}\rho + \gamma T D^2_{H_1}\rho \tag{1.6}$$

using the relations

$$\frac{\partial}{\partial x} D_{H_0} x\rho = \frac{\partial}{\partial x_j} J_{jk} \frac{\partial H_0}{\partial x_k}\rho = \frac{\partial \rho}{\partial x_j} J_{jk} \frac{\partial H_0}{\partial x_k} = D_{H_0}\rho$$

and

$$\frac{\partial}{\partial x}\{H_0, H_1\} D_{H_1} x\rho = \frac{\partial}{\partial x_j}\{H_0, H_1\}\rho J_{jk} \frac{\partial H_1}{\partial x_k} = D_{H_1}\{H_0, H_1\}\rho$$

Then the stationary solution is computed by choosing $\rho_s(x) = \rho_s(H_0(x))$ so that we get the condition

$$\gamma D_{H_1}\{H_0, H_1\}\rho_s(H_0) + \gamma T D^2_{H_1}\rho_s(H_0) = 0$$

We consider the sufficient condition

$$\{H_0, H_1\}\rho_s(H_0) + T\{H_0, H_1\}\frac{d\rho_s}{dH_0} = 0 \quad \Rightarrow \quad \frac{d}{dH_0}\ln \rho_s(H_0) = -T$$

and we recover the MB distribution. We remark that the stochastic equation has a covariant form in the Stratonovich interpretation: a canonical change of variable $x = T(y)$ in the white noise limit is covariant with respect the Stratonovich interpretation (see later).

1.0.1 Physical meaning of the stochastic equation

The stochastic differential equations cannot be directly studied with the methods of the dynamical systems theory due to the not physical characters of the Wiener process. From one hand they cannot be used as fundamental laws of Physics, but from the other hand they are a fundamental tool for a "mesoscopic" approach the physical problems and for a formulation of non-equilibrium Statistical Mechanics (Stochastic Thermodynamics) in Complex Systems Physics. The irreversible character of the stochastic dynamics allows to introduce the time arrow and the entropy concept in a natural way, but this is incompatible with the principles of classical mechanics for isolated systems. From a different point of view, one can say that the stochastic equations destroy the idea itself of isolated system: more precisely the "particles" are continuously affected by unpredictable small perturbations from the external environment (i.e. the concept of environment means the hidden degrees of freedom). The time reversing operation would imply to reverse the dynamical state of the whole environment (i.e. the whole universe) not only of the system particles. As a consequence even if the reverse dynamics is possible, but it is practically impossible to observe. It is also important to remark that any experimental measure concerns average quantities (both in space and time) of extensive or intensive observables, whose evolution does not necessary preserve the properties of Hamiltonian Mechanics. One assumes that the environment degrees of freedom satisfy the following conditions

- the microscopic dynamics should have a strong chaotic character (sensitive dependence from the initial conditions and/or from the environmental conditions);
- the macroscopic dynamics is affected by the environment through an additive term.

When both the hypotheses 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. One of the problems is 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 stochastic effects (statistical independence of the future from the past and irreversibility of the time arrow). We assume that the microscopic dynamics can be interpreted as the superposition of a deterministic Hamiltonian dependent only on the particle dynamical state and an additive stochastic term, whose amplitude depends only on the particle dynamical state. In a generic case one considers the a stochastic Hamiltonian in the form

$$H(x,t) = H_0(x) + H^p(x,\xi(t)) \tag{1.7}$$

where $\xi(t)$ is a regular stationary stochastic Gaussian process (i.e. we assume that the realizations $\xi(t)$ are continuous and we assume $<\xi>=0$) with a correlation function

$$\langle \xi(t)\xi(t+\tau)\rangle = \sigma^2\phi(\gamma\tau) \qquad \phi(\gamma\tau) \simeq \frac{\gamma}{2}e^{-\gamma|\tau|} \tag{1.8}$$

γ^{-1} is the correlation scale time, that defines the noise evolution with respect to the characteristic time scale of the unperturbed dynamics H_0 (the average values are taken over all the possible realizations of noise that can be associated to a probability measure $d\mu(\xi)$ in a functional space according to the path integral computation). We assume $H^p(x,0) = 0$ so that the perturbation can be written in the form

$$H^p(x,\xi(t)) = \xi(t)H_1(x) + \frac{\xi^2(t)}{2}H_2(x) + O(\xi^3) \qquad H_2(x) = \frac{\partial H_1}{\partial \xi}(x)$$

where we are interested in a small noise limit. We approximate the initial system by the Hamiltonian

$$H(x,t) = H_0(x) + H_1(x)\xi(t) + \frac{\xi^2(t)}{2}H_2(x) \tag{1.9}$$

We recall the white noise limit $\gamma \to \infty$ where the noise variance diverges

$$\lim_{\gamma\to\infty} \phi(\gamma\tau) \to \delta(\tau)$$

In the white noise limit we have

$$\lim_{\gamma\to\infty} \left\langle \int_0^\tau \xi^2(t)dt \right\rangle^2 = \sigma^2\tau$$

An explicit calculation gives

$$\left\langle \int_0^\tau \xi(t)\,dt \right\rangle^2 = \sigma^2 \frac{\gamma}{2} \int_0^\tau dt \int_0^\tau ds e^{-\gamma|t-s|} = \sigma^2 \int_0^\tau (1 - e^{-\gamma t}) dt$$

and

$$\left\langle \int_0^\tau \xi^2(t) dt \right\rangle^2 = \sigma^2 \left(\tau + \frac{e^{-\gamma\tau} - 1}{\gamma} \right)$$

Then letting in the limit $\gamma \to \infty$ when τ is fixed, we get the result $\sigma^2\tau$. We observe that the limit changes if we correlate the value of γ and τ (i.e. as we can fixed the product $\gamma\tau$ is the limit). For example if we fix γ the limit $\tau \to 0$ gives

$$\lim_{\tau\to 0} \left\langle \int_0^\tau \xi^2(t) dt \right\rangle^2 \simeq \sigma^2 \frac{\gamma\tau^2}{2}$$

that has different limit. This result is related to different physical models: if $\tau \gg \gamma^{-1}$ means that the evolution time scale is much greater than the fluctuation time scale and we can average over independent fluctuations. This is the diffusion limit where the mean square value of the fluctuations increases $\propto \tau$. In the white noise limit σ^2 plays the role of temperature. On the contrary if the evolution time scale is $\tau \ll \gamma^{-1}$ we have the ballistic case where we cannot average on the fluctuations that are an external forcing for the solution. In this case the solution is strongly dependent on the noise realization. Another case is when $\gamma\tau = 1$ and we have the limit

$$\lim_{\tau\to 0} \left\langle \int_0^\tau \xi^2(t) dt \right\rangle^2 \simeq \sigma^2 \frac{\tau}{2}$$

This is a diffusion limit, but the fluctuation time scale is of the same order of the evolution time scale: i.e. the environmental fluctuations could be induced by the particle dynamics itself. This is the expected case for Hamiltonian systems. As a consequence of the previous remark, the white noise limit has not a covariant character so it does not commute with changes of variable.

Without loss of generality, we neglect $H_2(x)$ in the following since its contribution is equivalent to an average contribution

$$\xi^2(t) H_2(x) \equiv \sigma^2 H_2(x)$$

in the white noise limit.

The Hamiltonian $H_0(x)$ is the particle energy whereas $H_1(x)$ is the perturbation Hamiltonian (in Statistical Mechanics one assumes that $H_1(x) \ll H_0(x)$. Moreover we consider the case $H_0(x) = E$ are compact invariant surfaces.

In applications to Statistical Mechanics each one considers an ensemble of independent 'particles' (i.e. copies of the system) whose evolution is defined by the Hamiltonian (**??**): i.e. any particles feel a different realization of the noise that mimics the effect of the interaction with the environment. This assumption can justified when the dynamics is non linear since the effect of fluctuations become independent at different points of the phase space since they depend on the non linear character of the unperturbed dynamics. For a given realization of the stochastic process $\xi(t)$ the evolution follows a symplectic dynamics for any initial condition. Then the equations of motion read

$$\dot{x} = J\left(\frac{\partial H_0}{\partial x} + \xi(t)\frac{\partial H_1}{\partial x}\right) = D_{H_0+\xi H_1}x \tag{1.10}$$

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

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

and D_H is the Poisson bracket operator $D_H x = \{x, H\}$. For any continuous realization $\xi(t)$, the stochastic phase flow $x(t, x_0|\xi) = \Phi^t_\xi(x_0)$ is defined by symplectic maps

$$\left(\frac{\partial \Phi^t_\xi}{\partial x}\right)^T J \frac{\partial \Phi^t_\xi}{\partial x} = J \tag{1.11}$$

The symplectic character of the phase flow (**??**) has to be justified from a physical point of view (i.e. one has to prove that for physical reasons the evolution must have a symplectic character since it represents a mechanical system). This is the case when we simulate the diffusion process in the phase space due to the local chaotic character of the dynamics. The evolution equation for the unperturbed energy along the trajectories reads

$$\frac{\partial H_0}{\partial t} = \xi(t)\frac{\partial H_0}{\partial x}^T J \frac{\partial H_1}{\partial x}$$

This equation is the backward equation for the evolution of the observable $H_0(x,t) = H_0(\Phi^t_\xi(x))$ along a stochastic trajectory for a given realization of the noise (the name backward is due to the fact that x is the initial condition). The solution can be formally written using the operator

$$H_0(x,t) = \exp\left(-\int_0^t \xi(s)\, ds \frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x}\right) H_0(x) \tag{1.12}$$

where we use $J^T = J$ and the commutative nature of the Gaussian operators

$$\xi(t)\frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x}$$

Remark: the fluctuation $\xi(t)$ is a scalar function in the Hamiltonian, so that any phase space variable is subjected to the same fluctuation and $H_1(x)$ is not time dependent. Then we have the relation

$$\begin{aligned} D_{H(t)}D_{H(s)}f &= \{\{f, H(s)\}, H(t)\} = \{\{H(t), H(s)\}, f\} + \{\{f, H(t)\}, H_(s)\} \\ &= \{\{H(t), H(s)\}, f\} + D_{H(s)}D_{H(t)}f \end{aligned}$$

so that the operators commutes if $\{H(t), H(s)\} = 0$. This is the case for the Hamiltonian $\xi(t)H_1(x)$ so that the formula (**??**) holds.
However if we consider the evolution of a generic observable $A(x)$ which is not an integral of motion for $H_0(x)$ the complete backward equation is

$$\frac{\partial A}{\partial t} = \frac{\partial A}{\partial x}^T J \frac{\partial H_0}{\partial x} + \xi(t)\frac{\partial A}{\partial x}^T J \frac{\partial H_1}{\partial x}$$

and the corresponding operators do not commute at different time so that it is not possible to write the solution in a exponential form.
The solution (**??**) depends on the Gaussian operator

$$\int_0^t \xi(s)\, ds \frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x}$$

(we assume $\xi(t)$ Gaussian), which has zero mean value and one can use the equality

$$\langle \exp(X) \rangle = \exp\left(\frac{\langle X \rangle^2}{2}\right)$$

which is valid for any Gaussian variable with $\langle X \rangle = 0$. Then the average observable evolution reads

$$\langle H_0(x,t) \rangle = \left\langle \exp\left(-\int_0^t \xi(s)\, ds \frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x}\right)\right\rangle H_0(x) \simeq \exp\left(\left[\frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x}\right]^2 t\right) H_0(x)$$

where we use the equality ($\xi(t)$ is assumed stationary)

$$\int_0^t \int_0^t \langle \xi(s_1)\xi(s_2) \rangle ds_1\, ds_2 \frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x} \frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x} = 2 \int_0^t \int_0^{s_1} c(s_1 - s_2) ds_1\, ds_2 \simeq t$$

and we perform the white noise limit

$$c(\tau) \simeq \frac{\gamma}{2} e^{-\gamma\tau} \qquad \gamma \gg 1$$

$H_0(x) = \langle H_0(x,0) \rangle$ is the initial condition and we recognize the solution of the partial differential equation

$$\frac{\partial \bar{H}_0}{\partial t} = \frac{1}{2} \frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x} \frac{\partial H_1}{\partial x}^T J \frac{\partial \bar{H}_0}{\partial x} \tag{1.13}$$

Remark: in the case of parametric noise the solution is more complicated since the perturbation Hamiltonian $H_1(x, \xi(t)$ produces non-commutative operators and the evolution operator is not exponential.
The symplectic character allows to introduce the Poisson bracket or the Lie derivative D_{H_1}

$$D_{H_1} F = \{F, H_1\} = -\frac{\partial H_1}{\partial x}^T J \frac{\partial F}{\partial x}$$

In the canonical variable $x = (q, p)$ and using the definition of J we get

$$D_{H_1} F = \sum_j \frac{\partial H_1}{\partial p_j} \frac{\partial F}{\partial q_j} - \frac{\partial H_1}{\partial q_j} \frac{\partial F}{\partial p_j}$$

and eq. (**??**) reads

$$\frac{\partial \bar{H}_0}{\partial t} = \frac{1}{2} D_{H_1}^2 \bar{H}_0 = \frac{1}{2} \frac{\partial H_1}{\partial x}^T J \frac{\partial}{\partial x} \left(\frac{\partial H_1}{\partial x}^T J \frac{\partial H_0}{\partial x}\right)$$

This equation holds for any first integral of motion of the unperturbed Hamiltonian. A particular case is the average dynamics when the observable is the single state of the system

$$\langle \Phi_\xi^t \rangle (x_0) = \int_\xi \Phi_\xi^t (x_0) d\mu(\xi) \tag{1.14}$$

We remark that the transformation $\Phi_\xi^t(x_0)$ is a symplectic map but this is not the case for the average value: indeed the symplectic condition reads (**??**)

$$\frac{\partial \Phi_\xi^t}{\partial x_0}^T J \frac{\partial \Phi_\xi^t}{\partial x_0} = J$$

so that the average value cannot be applied to each factor independently and the map (**??**) is not symplectic. From the Hamilton equation (**??**) the average procedure gives

$$\frac{d}{dt} \langle \Phi_\xi^t \rangle = J \left\langle H_0(\Phi_\xi^t) \right\rangle + \left\langle \xi(t) H_1(\Phi_\xi^t) \right\rangle$$

and in the white noise limit one gets

$$\begin{aligned} \langle \Phi_\xi^{\Delta t} \rangle (x_0) &\simeq \left\langle \exp \left(\Delta t D_{H_0} + \int_0^{\Delta t} \xi(s) D_{H_1} \, ds \right) \right\rangle x_0 + o(\Delta t) \\ &= \exp \left(\left[D_{H_0} + \frac{1}{2} D_{H_1}^2 \right] \Delta t \right) x_0 + o(\Delta t) \end{aligned}$$

By iterating this operator we can construct the solution $x(t)$ in the limit $\Delta t \to 0$

$$\bar{x}(t) = \exp \left(\left[D_{H_0} + \frac{1}{2} D_{H_1}^2 \right] t \right) x_0 + o(1)$$

This corresponds to the solution of the differential equation

$$\dot{\bar{x}} = D_{H_0(\bar{x})} \bar{x} + \frac{1}{2} D_{H_1(\bar{x})}^2 \bar{x}$$

whose character is not Hamiltonian. This is the average field approximation for the white noise limit.

1.1 Stochastic phase flows and diffusion equation

When one considers the white-noise limit the stochastic phase flow Φ_ξ^t can be associated to a stochastic differential equation that can to be derived from the stochastic Hamiltonian (**??**) to preserve the symplectic character in a probabilistic sense. The phase flow of the canonical equations (**??**) defines the stochastic phase flow

$$x(t + \Delta t) = \exp \left(\Delta t D_{H_0} + \Delta \xi D_{H_1} \right) x(t) \tag{1.15}$$

which is symplectic.
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 since $\Delta\xi \propto \sqrt{\gamma}\Delta t$ with $\gamma \to \infty$. In the white noise limit we consider $\Delta t \to 0$ but $\gamma\Delta t$ finite (i.e. the discretization time scale is of the same order of the correlation time scale), so that $\Delta\xi \simeq \sqrt{\Delta t}$ and it is necessary to consider the second order terms in the development. Using the operators D_H the explicit expansion reads

$$\begin{aligned} x(t+\Delta t) &= x(t) + \left[\Delta t D_{H_0} + \Delta\xi D_{H_1} + \frac{\Delta\xi^2}{2} D_{H_1}^2\right] x(t) + O(\Delta t^2, \Delta\xi^3) \\ &\simeq \exp\left(\Delta t D_{H_0} + \Delta\xi D_{H_1}\right) x(t) + O(\Delta t^{3/2}) \end{aligned}$$

where

$$\Delta\xi = \int_t^{t+\Delta t} \xi(s)\, ds$$

The relation holds in a statistical sense and in the white noise limit $\Delta\xi \to \Delta w_t$ and $\Delta\xi^2 \to \Delta t$ where w_t is a Wiener process (the limit is not a pointwise and it has to be considered a weak-convergence in a probabilistic sense) and the previous relation defines the Euler scheme for a stochastic differential equation. Then in the limit $\Delta t \to 0$ one gets the stochastic differential equation

$$dx = D_{H_0} x dt + D_{H_1} x dw_t + \frac{1}{2} D_{H_1}^2 x dt = D_{H_0} x dt + D_{H_1} x \circ dw_t \tag{1.16}$$

Remark: dw_t is a scalar stochastic process.
The equation (**??**) has not the canonical form (**??**), but its solution is stochastically equivalent to a symplectic phase flow in the sense of a L^2 mean square norm. The proof consider the possibility of representing a symplectic transformation using a Lie transformation

$$\Phi_{w_t}^{\Delta t} = \exp\left(D_{H_0} x \Delta t + D_{H_1} x \Delta w_t\right)$$

where Δt and Δw_t are given (in a simulation one has to choose the increments Δw_t independent each time step). Then we approximate

$$\exp\left(D_{H_0} x \Delta t + D_{H_1} x \Delta w_t\right) = D_{H_0} x \Delta t + D_{H_1} x \Delta w_t + \frac{1}{2} D_{H_1}^2 x \Delta t + O(\Delta t^{3/2})$$

in a L^2-average norm. Then in the limit $\Delta t \to 0$ one iterates the previous scheme at a finite time $t = N\Delta t$ with an error $O(\sqrt{\Delta t}) \to 0$.
It is possible to relate the equation (**??**) to a canonical form, by introducing the Stratonovich stochastic integral

$$\int \sigma_k(x) \circ dw_t = \int \sigma_k(x) dw_t + \frac{1}{2} \int \sigma_l(x) \frac{\partial \sigma_k}{\partial x_l}(x) dt \tag{1.17}$$

so that if we set

$$\sigma_k(x) = J_{kl} \frac{\partial H_1}{\partial x_l}$$

the equation (**??**) can be written in a canonical form

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

using the Stratonovich integral instead of the Ito stochastic integral. The Stratonovich equation recover a covariant form with respect to a symplectic change of variable $y = T(x)$: using the Ito formula we have

$$dy = \frac{\partial T}{\partial x}\left(D_{H_0(x)}x\,dt + D_{H_1(x)}x\,dw_t + \frac{1}{2}D^2_{H_1(x)}x\,dt\right) + \frac{1}{2}\frac{\partial^2 T}{\partial x^2}\left(D_{H_1(x)}x\right)^2\,dt$$

By definition of symplectic transformation we have

$$\frac{\partial T_i}{\partial x_k}\frac{\partial H}{\partial x_l}^T J_{lk} = \frac{\partial H}{\partial y_n}^T\frac{\partial T_n}{\partial x_l}^T J_{lk}\frac{\partial T_i}{\partial x_k} = \frac{\partial H}{\partial y_n}^T J_{ni}$$

so that

$$\frac{\partial T}{\partial x}D_{H_1(x)}x = D_{H_1(y)}y$$

Using the same algebraic properties it is possible to show

$$\frac{\partial T}{\partial x}D^2_{H_1(x)}x + \frac{\partial^2 T}{\partial x^2}\left(D_{H_1(x)}x\right)^2 = D^2_{H_1(y)}y$$

so we get the equation

$$dy = D_{H_0(y)}ydt + D_{H_1(y)}ydw_t + \frac{1}{2}D^2_{H_1(y)}ydt$$

which is the Stratonovich equation associated to the stochastic Hamiltonian $H_0(y) + H_1(y)\xi(t)$ is the white noise limit.
The statistical properties of any observable can be computed by the distribution function

$$\rho(x,t) = \left\langle \rho_0(\Phi_\xi^{-t}(x))\left|\frac{\partial\Phi_\xi^{-t}}{\partial x}\right|\right\rangle = \left\langle \rho_0(\Phi_\xi^{-t}(x))\right\rangle$$

since $\Phi^t\xi(x)$ is a symplectic map, where $\rho_0(x_0)$ is the initial probability distribution of the system and $\Phi_\xi^{-t}(x)$ the inverse of the stochastic phase flow with $\xi(t)$ given. The evolution of observables gives

$$\bar{O}(t) = \int \langle O(\Phi_\xi^t(x_0))\rangle \rho_0(x_0)\,dx_0 = \int O(x)\left\langle \rho_0(\Phi_\xi^{-t}(x)\right\rangle\,dx$$

To compute the evolution equation of the distribution function one considers the stochastic Liouville equation. We require a stationarity condition on the system: for any realization $\xi(t)$ and any time t_0, there exists a realization $\xi'(t)$ such that

$$\Phi^{t_0+t}_{\xi(t_0)} = \Phi^t_{\xi'(t_0)}$$

(i.e. the evolution is homogeneous in time). The stationary condition implies that the statistical properties of the evolution are invariant with respect to the choice of the initial time. The knowledge of $\Phi_\xi^t(x)$ gives a complete information on the system, but in many cases it is enough to know the distribution function $\rho(x,t)$ that gives the probability to detect a particle at x after a time t given its initial condition $\rho_0(x) = \delta(x - x_0)$. 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

$$\frac{\partial\rho_\xi}{\partial t} = -\frac{\partial}{\partial x}J\left(\frac{\partial H_0}{\partial x} + \xi(t)\frac{\partial H_1}{\partial x}\right)\rho_\xi = \left(\frac{\partial H_0}{\partial x} + \xi(t)\frac{\partial H_1}{\partial x}\right)J\frac{\partial\rho_\xi}{\partial x} \qquad (1.18)$$

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

$$\dot{x} = \sigma(x)\xi(t)$$

the stochastic phase flow reads

$$x(\Delta t) = x + \sigma(x)\int^{\Delta t} \xi(s)ds + \frac{\partial \sigma}{\partial x}\sigma(x)\int^{\Delta t}\int^{s} \xi(s)\xi(u)du\,ds + ...$$

We require that the following limit (at least in a weak form)

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t}\int^{\Delta t}\int^{s} \xi(s)\xi(u)du\,ds \to 0$$

Assuming a correlation function

$$\langle \xi(s)\xi(u)\rangle = \gamma \exp\left(-\gamma|s-u|\right)$$

we get

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

and the required limit holds in average if

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

This means that one 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 is not justified. Conversely if $\gamma \Delta t \ll 1$ then the fluctuation $\xi(t)$ can be considered a regular function and eq. (**??**) holds.
Remark: the time reversibility can be defined only in a statistical sense. For a given realization of the noise $\xi(t)$ one can consider the reverse realization $\xi(-t)$ at a fixed time $t = 0$; if there exists a realization $\xi'(t)$ such that $\xi'(t) = \xi(-t)$ the system is stochastically time-reversible. The time reversibility property may be too restrictive and in some cases one considers the reversibility condition only when the system is relaxed into a stationary equilibrium state (Onsager Theory). The equation (**??**) has the form of a continuity equation since the r.h.s. can be written as the divergence of the stochastic current

$$J_\xi(x,t) = -\left(\frac{\partial H_0}{\partial x} + \xi(t)\frac{\partial H_1}{\partial x}\right) J\rho(x,t)$$

Let $\Phi_\xi^{-t}(x)$ is the inverse of the phase flow Φ_ξ^t for a fixed realization $\xi(t)$ (the inverse exists as a consequence of the existence and uniqueness theorem for the solution of differential equations), we can write the solution of eq. (**??**) according to

$$\rho_\xi(x,t) = \rho_0\left(\Phi_\xi^{-t}(x)\right) \tag{1.19}$$

where $\rho_0(x)$ defines the initial condition. The distribution function $\rho_\xi(x,t)$ provides all the information for a statistical approach to the system. But in

many cases the realization of the process $\xi(t)$ is unknown so that one considers the expectation value of all the possible distribution function. In principle there is the problem to study $\rho_\xi(x,t)$ as a random function whose variance gives information on the reliability of the expectation value.
In Statistical Physics one assumes that considering an ensemble of particles, each particle is subjected by a different noise realization and the average on the possible realization corresponds to the average on the ensemble when the probability distribution is weakly dependent on the initial conditions. The knowledge of $\langle \rho_\xi(x,t) \rangle$ gives all the information for a statistical mechanics approach to the study of the considered systems and, from a mathematical point of view, this means the existence of a generalized law of large numbers for a system composed by many identical non-interacting particles. The classical result for statistical systems containing 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 $O(1/\sqrt{N(\Delta V)})$ where $N(\Delta V)$ is the number of particles contained in the volume ΔV in the limit $N \to \infty$ and $\Delta V \to 0$. The average value is computed by considering independent realizations $\xi(t)$ distributed according to $d\mu(\xi)$. Of course there are situations in which the fluctuations effects are important both because the system contains a limited number of particles and because the independence assumption for the dynamics of different particles fails.

1.2 Diffusion equation for the distribution function

Deriving an evolution equation for the average distribution $\langle \rho_\xi(x,t) \rangle$ is a key issue for non-equilibrium Statistical Mechanics . We proceed in a formal way from the stochastic Liouville equation (**??**). We takes advantage by the following *Lemma*:

Given a stochastic Liouville of the form

$$\frac{\partial \rho_\xi}{\partial t} = -\xi(t)\frac{\partial H}{\partial x}(x,t)J\frac{\partial \rho_\xi}{\partial x} \tag{1.20}$$

the solution can be formally written according to

$$\rho_\xi(x,t) = \mathcal{T}\exp\left(-\int_0^t \xi(s)\frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x}ds\right)\rho_0(x) \tag{1.21}$$

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

$$D_{H(x,t)} = \frac{\partial H}{\partial x}(x,t)J\frac{\partial}{\partial x}$$

at different times (we have not this problem if H is time independent)).

To compute the expectation value with respect all the noise realizations, we use the commutativity property of the $\mathcal{T}$ operator and the averaging computation.

This is possible since the Liouville equation depends linearly from the noise $\xi(t)$. By expanding the exponential operator we get terms of the form

$$\frac{1}{n!}\mathcal{T}\left\langle\left(\int_0^t -\xi(s)\frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x}ds\right)^n\right\rangle \tag{1.22}$$

We perform an explicit calculation in the case of a Gaussian process $\xi(s)$ where the stochastic fluctuation of the operator

$$D(t;\xi) = \int_0^t \xi(s)\frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x}\,ds \tag{1.23}$$

are Gaussian and satisfy the cumulant relations. The non Gaussian character can be considered when a Central Limit Theorem is applied to the fluctuations in the dynamics (this is possible for small noise) so that we recover the Gaussian distribution. Moreover according to eq. (**??**) $\xi(t)$ is a scalar random process so that it commutes with the derivative operator. Then the average over all the possible realizations and the time-ordering operator commutes and using the cumulant relations we get

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

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

$$\begin{aligned} B(x,t) &= \frac{1}{2}\left\langle (D(t,\xi))^2\right\rangle \\ &= \frac{1}{2}\int_0^t ds\int_0^t du\langle\xi(s)\xi(u)\rangle\frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x}\frac{\partial H}{\partial x}(x,u)J\frac{\partial}{\partial x} \end{aligned} \tag{1.24}$$

we get the equality (n is even otherwise we have no contribution)

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

In the case of a Gaussian process from the equation (**??**) we derive the relation

$$\langle\rho_\xi(x,t)\rangle = \exp\left(B(t)\right)\rho_0(x) \tag{1.25}$$

In the white noise limit, the operator $B(t)$ can be explicitly computed

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

and we have the formal identity

$$\langle\rho_\xi(x,t)\rangle = \exp\left(\frac{1}{2}\int_0^t \frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x}\frac{\partial H}{\partial x}(x,s)J\frac{\partial}{\partial x}ds\right)\rho_0(x) \tag{1.26}$$

The following *Lemma* holds:

Given a stochastic Liouville of the form

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

where $\xi(t)$ is a Gaussian noise with correlation function (**??**), in the limit of a white noise for the process $\xi(t)$, the average distribution $\bar{\rho}(x,t) = \langle \rho_\xi \rangle(x,t)$ satisfies the Fokker-Planck equation

$$\frac{\partial \bar{\rho}}{\partial t} = \frac{1}{2}\left[\frac{\partial H}{\partial x}(x,t)J\frac{\partial}{\partial x}\right]^2 \bar{\rho} \tag{1.27}$$

Eq. (**??**) is related by the backward equation for the evolution of the observable by an adjoint operation. From the equality

$$\int \frac{\partial}{\partial t} O(x_0,t)\rho_0(x_0)\, dx_0 = \int O(x)\frac{\partial}{\partial t}\langle \rho_\xi(x,t)\rangle\, dx$$

where we use the symplectic change of variable $\Phi^t_\xi(x_0) = x$, at $t = 0$ one gets

$$\begin{aligned}\int \left[D_{H_0} + \frac{1}{2}D_{H_1^2}\right] O(x_0,t)\rho(x_0,t)\, dx_0 &= \int O(x_0)\left[D_{H_0} + \frac{1}{2}D_{H_1^2}\right]^* \rho(x_0,t) \\ &= \int O(x_0)\frac{\partial}{\partial t}\rho(x_0,t)\, dx_0\end{aligned}$$

since $x = x_0$ and $\rho(x_0,t) = \rho_0(x_0)$ and the suffix $*$ denotes the adjoint operator. The choice $t = 0$ is generic: one can interpret the previous equality as the evolution from $t \to t + \Delta t$ in the limit $\Delta t \to 0$ starting from the condition at time t considered known. By definition one has

$$D_H^\dagger = \frac{\partial}{\partial x_i}J_{ij}\frac{\partial H}{\partial x_j} = \frac{\partial H}{\partial x_j}J_{ij}\frac{\partial}{\partial x_i} = -D_H$$

so that we recover the FP equation (**??**).
One has to remark is that the white noise limit for the process $\xi(t)$ is a singular limit: to keep finite the noise effect , the amplitude of the noise should diverge. The Liouville equation (**??**) is justified only when the noise correlation time γ^{-1} is fixed so that it is possible to consider an evolution time scale $\Delta t \ll \gamma^{-1}$. The white noise limit $\gamma^{-1} \to 0$ is an approximation of the solution of the stochastic Liouville equation when the correlation time scale is negligible with respect the evolution time scale. The FP equation (**??**) provides an effective way to approximate la solution of the stochastic Liouville equation averaged over all the possible noise realizations.
In the case of the Hamiltonian (**??**) a key role is played by the unperturbed dynamics $H_0(x)$: in the case of non-linear dynamics different initial condition can be related to independent evolution so that an average over a particle ensemble coincides with the average over the realizations. This is not only true if $H_0(x)$ has a chaotic dynamics and the correlation among different initial conditions depend on the Ljapunov exponents, but also if $H_0(x)$ is an integrable Hamiltonian and the orbit correlation depends on the nonlinear character and it has a power law decaying.

1.3 Remark on white noise limit

The white noise limit is consistent with the Stratonovich limit for the stochastic dynamics (**??**) that is related to the stochastic differential equation

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

associated to a FP equation (**??**) by the Ito calculus. The white noise limit can be useful to approximate the operator (**??**) when one has a fast decaying in the noise correlation. This means that we can applied the previous approach when the fluctuations are defined by a chaotic dynamics with a positive Ljapunov exponents. In such a case the different realizations are different chaotic trajectories (i.e. corresponding to different initial conditions) and their correlation decaying exponentially according to the sum of the positive Ljapunov exponents.
One could compute the effect of a finite noise correlation in a formal way using Gaussian stochastic operators $B(x,t)$ with an exponentially decaying correlation. The Gaussian character allows to compute the finite correlation effects of the noise, otherwise one needs to perform the white noise limit to recover the Gaussian character by means of the Central Limit Theorem.
To generalize the previous results to the Hamiltonian (**??**), we observe that the stochastic Liouville equation (**??**) can be reduced in the form required by the Lemma if one performs the symplectic change of variables

$$x(t,\xi) = \Phi_0^t(y(t,\xi))$$

where $\Phi_0^t(\cdot)$ is the phase flux associated to the deterministic Hamiltonian $H_0(x)$. As it is well known from Hamiltonian dynamics theory, in the new variables y, the new Hamiltonian is $\xi(t)H_1(\Phi_0^t(y))$ so that the corresponding Liouville has the form (**??**) and the average distribution $\langle \rho_\xi \rangle (y,t)$ satisfies the FP equation

$$\frac{\partial \langle \rho_\xi \rangle}{\partial t}(y,t) = \frac{1}{2} \frac{\partial H_1}{\partial y}(\Phi_0^t(y)) J \frac{\partial}{\partial y} \frac{\partial H_1}{\partial y}(\Phi_0^t(y)) J \frac{\partial}{\partial y} \langle \rho_\xi \rangle (y,t) \tag{1.28}$$

The Hamiltonian Liouville operator is covariant with respect the canonical changes of variables:

$$\frac{\partial H_1}{\partial y}(\Phi_0^t(y),t) J \frac{\partial}{\partial y}\Big|_{y=\Phi_0^{-t}(x)} = \frac{\partial H_1}{\partial x}(x) J \frac{\partial}{\partial x}$$

As a consequence the FP equation (**??**) in the original variables reads (recall that the symplectic character of the phase flow does not change the density measure)

$$\frac{\partial \rho}{\partial t} \langle \rho_\xi \rangle (\Phi_0^{-t}(x),t) = \frac{1}{2} \frac{\partial H_1}{\partial x} J \frac{\partial}{\partial x} \frac{\partial H_1}{\partial x} J \frac{\partial}{\partial x} \langle \rho_\xi \rangle (\Phi_0^{-t}(x),t)$$

where Φ_0^{-t} denotes the inverse phase flow. Letting $< \rho_\xi > (\Phi_0^{-t}(x),t) = \rho(x,t)$ an explicit calculation finally gives

$$\frac{\partial}{\partial t}\rho(x,t) = \frac{\partial \langle \rho_\xi \rangle}{\partial t}(\Phi_0^{-t}(x),t) + \frac{\partial H_0}{\partial x}(x) J \frac{\partial}{\partial x}\rho(x,t)$$

so that the final FP equation reads

$$\frac{\partial \rho}{\partial t}(x,t) = -\frac{\partial H_0}{\partial x} J \frac{\partial \rho}{\partial x}(x,t) + \frac{1}{2}\frac{\partial H_1}{\partial x} J \frac{\partial}{\partial x}\frac{\partial H_1}{\partial x} J \frac{\partial \rho}{\partial x}(x,t) \tag{1.29}$$

We have formally proved the following *proposition*:
the average distribution function of the stochastic Hamiltonian system (**??**) in the limit of a white Gaussian noise, satisfies the Fokker-Planck equation (**??**).
Remark: if one interprets the system (**??**) as a stochastic differential equation, the FP equation (**??**) is associated to the stochastic differential equation (cfr. eq (**??**))

$$dx = \left(\frac{\partial H_0}{\partial x} J + \frac{1}{2}\frac{\partial H_1}{\partial x} J \frac{\partial^2 H_1}{\partial x \partial x} J \right) dt + \frac{\partial H_1}{\partial x} J dw_t \tag{1.30}$$

is obtained if one follows the Stratonovich interpretation of the canonical stochastic differential equations (**??**). The previous stochastic equation has to be used instead of the usual canonical equation for stochastic Hamiltonian systems. As a the consequence of the symplectic nature of the Hamiltonian dynamics, the differential operator that defines the Fokker-Planck equation

$$\mathcal{F}_H = -\frac{\partial H_0}{\partial x}(x) J \frac{\partial}{\partial x} + \frac{1}{2}\left(\frac{\partial H_1}{\partial x} J \frac{\partial}{\partial x}\right)^2$$

is an self-adjoint operator (one can use the properties of the Lie-derivative) so that the eigenvalues are all real and the eigenvector are orthogonal with respect to the L^2-product. We define the drift and the diffusion coefficients by

$$a_j(x) = \sum_i \frac{\partial H_0}{\partial x_i} J_{ij} + \frac{1}{2}\sum_{ihl} \frac{\partial H_1}{\partial x_i} J_{ih} \frac{\partial^2 H_1}{\partial x_h x_l} J_{lj} \quad b_{jk}(x) = \sum_{ih} \frac{\partial H_1}{\partial x_i} J_{ij} \frac{\partial H_1}{\partial x_h} J_{hk} \tag{1.31}$$

We observe that the diffusion coefficient only depends on the perturbation Hamiltonian $H_1(x)$ and not on the dynamical properties of the unperturbed system $H_0(x)$.
Moreover the operator is semi-negative defined: indeed if one computes the quadratic form

$$\int \rho \mathcal{F}_H \rho dx = -\int \rho D_{H_0} \rho dx + \frac{1}{2}\int \rho D_{H_1}^2 \rho dx$$

Using the equality

$$D_{H_0}\rho^2 = 2\rho D_{H_0}\rho$$

and the property

$$\int_X D_H \rho \, dx = \int_X \sum_{ij} \frac{\partial H}{\partial x_i} J_{ij} \frac{\partial \rho}{\partial x_j} \, dx =$$
$$\int_X \sum_{ij} J_{ij} \frac{\partial}{\partial x_j}\frac{\partial H}{\partial x_i} \rho \, dx = \int_{\partial X} \sum_{ij} J_{ij} \frac{\partial H}{\partial x_i} \rho \, d\sigma_j = 0$$

assuming vanishing boundary conditions for ρ and its derivatives, then one gets

$$\int \rho \mathcal{F}_H \rho dx = -\frac{1}{2}\int (D_{H_1}\rho)^2 dx \leq 0 \tag{1.32}$$

The stationary solution is defined by the condition $\mathcal{F}_H \rho_s = 0$; according to the previous result ρ_s has to satisfy both the equations

$$D_{H_0}\rho_s = 0 \qquad D_{H_1}\rho_s = 0$$

which implies $\{H_0, H_1\} = 0$ (due to the Jacobi identity) which means that if H_1 is an integral of motion of the unperturbed Hamiltonian H_0. In such a case the surfaces $H_0(x) = E$ are invariant with respect to the stochastic dynamics (**??**): using the Ito calculus one gets (cfr. eq. (**??**)

$$\begin{aligned} dH_0 &= \sum_j \frac{\partial H_0}{\partial x_j}\left[\sum_i \frac{\partial H_0}{\partial x_i} J_{ij} + \frac{1}{2}\sum_{ihl} \frac{\partial H_1}{\partial x_i} J_{ih} \frac{\partial^2 H_1}{\partial x_h x_l} J_{lj} dt + \sum_i \frac{\partial H_1}{\partial x_i} J_{ij} dw_t\right] \\ &+ \frac{1}{2}\sum_{ij} \frac{\partial^2 H_0}{\partial x_i \partial x_j}\left[\frac{\partial H_1}{\partial x_h} J_{hi} \frac{\partial H_1}{\partial x_k} J_{kj}\right] dt \\ &= \frac{1}{2}\sum_{ijhl} \frac{\partial H_0}{\partial x_j}\frac{\partial}{\partial x_l}\left[\frac{\partial H_1}{\partial x_i} J_{ih} \frac{\partial H_1}{\partial x_h}\right] J_{lj} dt - \frac{1}{2}\sum_{ijhl} \frac{\partial H_0}{\partial x_j}\frac{\partial^2 H_1}{\partial x_i \partial x_l} J_{ih} \frac{\partial H_1}{\partial x_h} J_{lj} dt \\ &+ \frac{1}{2}\sum_{ij} \frac{\partial}{\partial x_i}\left[\frac{\partial H_0}{\partial x_j}\frac{\partial H_1}{\partial x_k}\right] J_{hi} \frac{\partial H_1}{\partial x_h} J_{kj} dt + \frac{1}{2}\sum_{ij} \frac{\partial H_0}{\partial x_j}\frac{\partial^2 H_1}{\partial x_i \partial x_k} J_{ih} \frac{\partial H_1}{\partial x_h} J_{kj} dt \\ &= \frac{1}{2}\sum_{ij} \frac{\partial}{\partial x_i}\left[\frac{\partial H_0}{\partial x_j} J_{kj} \frac{\partial H_1}{\partial x_k}\right] J_{hi} \frac{\partial H_1}{\partial x_h} dt = 0 \end{aligned}$$

Then the stationary distribution is related to a micro-canonical ensemble if the initial condition is on the invariant surface $H_0(x) = E$ and, if no other integrals of motion exist, the invariant measure is uniform on the surface.
Remark: the stochastic Liouville equation is time-reversible, whereas the FP equation is not: indeed the substitution $t \to -t$ and $H(x) \to -H(x)$ leaves invariant the equation (**??**), but not the equation (**??**), so that $\langle \rho_\xi \rangle(x, -t)$ is not the evolution of a physical distribution (see next section).
Remark: in the applications the equation (**??**) can be only approximately satisfied by real systems: one can consider the white noise approximation when the typical correlation time of the stochastic process $\xi(t)$ is much smaller that the typical evolution scale of the Hamiltonian dynamics $H_0(x)$. The justification of a stochastic approach understands the existence of two time scales: the noise time scale that implies a fast decorrelation for the realizations $\xi(t)$, but with the possibility of large fluctuations, and the slow evolution time scale of the system. The noise time scale can be determined by the chaotic character of the hidden degrees of freedom that has positive Ljapunov exponents that in a coarse grained description give an exponential decorrelation like for Markov processes. The presence of a single noise in the Hamiltonian (**??**) means that $\xi(t)$ is a description of the global effect of the hidden degrees.
Introducing the Lie-operator for Hamiltonian systems, the previous results can be written in the form

$$\left\langle \mathcal{T} \exp(\int_0^t D_{H_0} + \xi(t) D_{H_1} ds) \right\rangle \rho_0(x) = \exp\left(-t D_{H_0} + t/2 D_{H_1}^2\right) \rho_0(x)$$

in the white noise limit for the Gaussian process $\xi(t)$. The Fokker-Planck equation (**??**) can be written in the form

$$\frac{\partial \rho}{\partial t} = -D_{H_0}\rho + \frac{1}{2} D_{H_1}^2 \rho \tag{1.33}$$

(Since we assume H_0 and H_1 time independent, we do not need the time-ordering operator $\mathcal{T}$).
Remark: the limit $\xi(t)$ towards a white noise, can be formally stated in the following *Lemma*:
let $\xi(t)D(t)$ a stochastic process of linear operators that are distributed according to a Gaussian function with zero mean value, we consider the stochastic equation

$$\dot{x} = \xi(t)D(t)x$$

whose solution can be formally written as

$$x(t;\xi) = \mathcal{T}\exp\left(\int_0^t \xi(s)D(s)ds\right)x_0$$

where the initial condition x_0 is given. Let us define the average linear operator

$$B(t) = \frac{\sigma^2}{2}\int_0^t ds \int_0^t du\phi(\gamma(s-u)D(s)D(u)$$

where $\sigma^2\phi(\gamma s)$ is the correlation function, then the average evolution $X(t) =< x(t;\xi) >$ is given by

$$X(t) = \mathcal{T}\exp\left(B(t)\right)X_0$$

The proof is based on the fact that the stochastic operator

$$\int_0^t \xi(s)D(s)ds$$

is distributed according to a Gaussian function since $\xi(t)$ is Gaussian and the the time ordering operator $\mathcal{T}$ commutes with the averaging procedure. Then we consider the average exponential stochastic operator

$$\left\langle \mathcal{T}\exp(\int_0^t \xi(s)D(s)ds)\right\rangle =$$
$$\sum_{n\geq 0}^{even}\mathcal{T}\frac{1}{2^{2/n}(n/2)!}\left(\int_0^t\int_0^t \sigma^2\phi(\gamma(s-u))D(s)D(u)dsdu\right)^{n/2}$$

we have the identity

$$\left\langle \mathcal{T}\exp(\frac{1}{2}\int_0^t \xi(s)D(s)ds)\right\rangle = \mathcal{T}\exp(B(t))$$

From the definition we get

$$\frac{dB}{dt} = \frac{\sigma^2}{2}\int_0^t ds\phi(\gamma(t-s)(D(t)D(s)+D(s)D(t)$$

Then if one considers the formal solution of the differential equation

$$\dot{X}(t) = \frac{dB}{dt}(t)X(t)$$

one gets

$$X(t) = \sum_{n\geq 0}\int_0^t ds_1....\int_0^{s_n-1} ds_n\frac{dB}{ds_1}....\frac{dB}{ds_n}X_0 = \sum_{n\geq 0}\frac{1}{n!}\mathcal{T}\int_0^t ds_1....\int_0^t ds_n\frac{dB}{ds_1}....\frac{dB}{ds_n}X_0$$

or

$$X(t) = \mathcal{T}\exp(B(t))X_0$$

The function $X(t)$ is the formal solution of the differential equation

$$\dot{X} = \frac{\sigma^2}{2}\int_0^t ds\phi(\gamma(t-s)(D(t)D(s) + D(s)D(t)X$$

In the white noise limit one finally gets

$$\dot{X} = \frac{1}{2}D^2(t)X$$

The previous results can be extended to the case of the stochastic differential equation

$$\dot{x} = (C + \xi(t)D)x$$

One introduces the interaction vision using the variable $x = \exp(Ct)y$ which satisfies the equation

$$\dot{y} = \xi(t)\exp(-Ct)D\exp(Ct)y$$

Then the operator $B(t)$ and its derivatives are defined

$$\begin{aligned} B(t) &= \frac{\sigma^2}{2}\int_0^t ds\int_0^t du\phi(\gamma(s-u))e^{-Cs}De^{C(s-u)}De^{Cu} \\ \frac{dB}{dt} &= \frac{\sigma^2}{2}e^{-Ct}\int_0^t \phi(\gamma s)\left[De^{Cs}De^{-Cs} + e^{Cs}De^{-Cs}D\right]ds\, e^{Ct} \end{aligned}$$

Then the average evolution $Y(t)$ satisfies the linear equation

$$\dot{Y} = \frac{\sigma^2}{2}e^-Ct\int_0^t \phi(\gamma s)\left[De^{Cs}De^{-Cs} + e^{Cs}De^{-Cs}D\right]dse^{Ct}Y$$

By introducing the variable $X(t) = \exp(Ct)Y(t)$ we finally gets

$$\dot{X} = CX + \frac{\sigma^2}{2}\int_0^t \phi(\gamma s)\left[De^{Cs}De^{-Cs} + e^{Cs}De{-}CsD\right]Xds$$

and in the white noise limit

$$\dot{X} = CX + \frac{1}{2}D^2X$$

This equation corresponds to the backward equation of the observables.

In the case of the system (??) the physical problem is the evolution of an ensemble of particles in the phase space under the the environmental effect. We have two cases: each particle has a different noise realization $\xi(t)$ and since all the particles are identical and the particles are not interacting this is equivalent to average on the noise realizations; there is the same noise realization for all the particles (this is a the case when the noise is parametric). In the last case the unperturbed dynamics plays a fundamental role to justify the averaging procedure: in a linear case this approach may lead to wrong results, but in a generic case, the non-linear character of the dynamics introduces a decorrelation among the particle trajectories that can be the result of independent noise realization, but one should prove that the distribution of the noise realizations is the same if one consider a single particle or an ensemble on particles. The white noise approximation is justified if the experimental measures are not instantaneous but are obtained by a time-averaging procedure

$$\overline{f}_\xi(t) = \frac{1}{T}\int_t^{t+T} \hat{f}_\xi(t)dt \tag{1.34}$$

where the time interval T is sufficiently long with respect to the correlation time scale of the stochastic process $\xi(t)$. Then $\overline{f}_\xi(t)$ is almost independent from the realization $\xi(t)$ (in a probabilistic sense) and for $T \gg 1$ we have

$$\overline{f}(t) = \langle \hat{f}_\xi(t)\rangle = \int f(x)\langle \rho_\xi(x,t)\rangle dx$$

where we have introduced the expectation value of the distribution function $\rho_\xi(x,t)$ with respect all the possible noise realizations.

1.4 Stochastic symplectic maps

The introduction of a thermal bath for a symplectic map M using a stochastic perturbation of

1.5 Remark on the numerical integration

We consider the problem of integrating the stochastic Hamiltonian

$$H_0(x) + H_1(x)\xi(t)$$

in the white noise limit. We initially consider the reduce case

$$\dot{x} = D_{H_1} x\xi(t) \qquad x(\Delta t) = x + \int_0^{\Delta t} D_{H_1} x(s)\xi(s)\, ds$$

We have the expansion

$$D_{H_1} x(s)\xi(s) = D_{H_1} x\xi(s) + D^2_{H_1} x \int_0^s \xi(u)\, du + \frac{1}{2} D^3_{H_1} x \left(\int_0^s \xi(u)\, du\right)^2 + O(\Delta t^{3/2})$$

where the error has to be interpreted in a statistical sense. By substituting in the expansion we get

$$
\begin{aligned}
x(\Delta t) &= x + D_{H_1} x \int_0^{\Delta t} \xi(s)\, ds \\
&+ D_{H_1}^2 x \int_0^{\Delta t} \xi(s) \int_0^s \xi(u)\, du\, ds + \frac{1}{2} D_{H_1}^3 \int_0^{\Delta t} \xi(s) \left(\int_0^s \xi(u)\, du \right)^2 ds + O(\Delta t^2) \\
&= x + D_{H_1} x \Delta w_t + D_{H_1}^2 x \frac{\Delta w_t^2}{2} + \frac{1}{2} D_{H_1}^3 int_0^{\Delta t} \xi(s) \left(\int_0^s \xi(u)\, du \right)^2 ds + O(\Delta t^2)
\end{aligned}
$$

so that the solution can be written using the Lie-Transformation

$$
x(t + \Delta t) = \exp\left(\Delta w_t D_{H_1} \right) x(t) \tag{1.35}
$$

where $\Delta w_t = w(t + \Delta t) - w(t)$ is the increment of a Wiener process. In the generic case of a stochastic Hamiltonian we have

$$
\dot{x} = D_{H_0} x + \xi(t) D_{H_1} x \qquad x(\Delta t) = x + \int_0^{\Delta t} D_{H_0} x\, ds + \int_0^{\Delta t} D_{H_1} x(s) \xi(s)\, ds
$$

The problem is the non-commutativity of the operators D_{H_0} and D_{H_1} so that one cannot write the solution using an exponential operator. However we can approximate the solution by a leap-frog integration scheme with an error $O(\Delta t^2)$

$$
x(\Delta t) = \exp\left(\frac{\Delta t}{2} D_{H_0} \right) \exp(\Delta w_t D_{H_1}) \exp\left(\frac{\Delta t}{2} D_{H_0} \right) x + O(\Delta t^2) \tag{1.36}
$$

The leap-frog scheme reproduces correctly the expansion of the solution up to the mixed algebraic terms of second order of the operator D_{H_0} and D_{H_1} which means an error of order $O(\Delta t^2)$ (i.e. the same order of the stochastic operator (**??**). For a finite time integration, the final order of the scheme (**??**) is $O(\Delta t)$ on the mean L^2-norm. This implies a convergence of the distribution function $\rho(x, t)$ of the numerical scheme to the solution of the FP equation.
The advantage of the scheme (**??**) is the possibility of maintaining the symplectic character of the single particle solution solutions of small noise limit.
The operator (**??**) applies to the evolution of any observable. Let us consider the energy evolution $H_0(x(t))$ (i.e. the unperturbed energy of the system which is a integral of motion), an explicit computation gives

$$
H_0(x(t{+}\Delta t)) = \exp\left(\frac{\Delta t}{2} D_{H_0} \right) \exp(\Delta w_t D_{H_1}) H_0(x) = \exp(\Delta w_t D_{H_1}) H_0(x_0(t{+}\Delta t/2)
$$

where we use the symplectic properties of the evolution operator $\exp(\Delta t D_{H_0})$. In the limit $\Delta t \to 0$ the average value over the realizations gives

$$
\langle H_0(x(t + \Delta t)) \rangle = \exp\left(\frac{\Delta t}{2} D_{H_1}^2 \right) \langle H_0(x) \rangle
$$

that corresponds to the solution of the backward equation. Assume that the space is foliated into invariant surfaces $H_0(x) = const.$ and we have an elliptic fixed point x_* with $H_0(x_*) = E_*$, a direct calculation provides

$$
\frac{d}{dt} \int_{E_*}^0 H_0^2(x)\, dx = \int_{E_*}^0 H_0(x) D_{H_1}^2 H_0\, dx = -\int \left(D_{H_1} H_0 \right)^2 dx
$$

where the choice of the surface $H_0(x) = 0$ is arbitrary. Therefore we have a Ljapunov function and the equilibrium state in the set requires the condition $D_{H_1} H_0 = 0$. In a generic situation this is possible only if $H_0(x(t)) \to const.$, since the condition implies that H_1 is a first integral of motion of H_0. This means in the stationary solution for a stochastic Hamiltonian system with open boundary conditions has a trivial stationary state and the dynamics needs to study the transient states of the FP operator corresponding to the small eigenvalues.

Chapter 2

The Fokker-Planck equation

The probability distribution associated to a stochastic differential equation

$$\dot{x} = a(x) + b(x)\xi(t) \qquad \langle \xi \rangle = 0 \quad \langle \xi(t+\tau)\xi(t) \rangle = \frac{\gamma}{2} \exp(-\gamma\tau)$$

satisfies a FP equation in the white noise limit $\gamma \to \infty$

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x} a(x)\rho + \frac{1}{2}\frac{\partial}{\partial x} b(x) \frac{\partial}{\partial x} b(x)\rho \tag{2.1}$$

Remark: this is consistent with the Stratonovich interpretation of the stochastic differential equation that corresponds to the Ito equation

$$dx = a(x)dt + \frac{1}{2} b(x) \frac{\partial b}{\partial x} dt + b(x) dw_t \tag{2.2}$$

The Stratonovich interpretation is covariant: let $x = T(y)$ we have

$$dy = \frac{\partial y}{\partial x} \left[a(T(y))dt + \frac{1}{2} b(T(y)) \frac{\partial b}{\partial y}\frac{\partial y}{\partial x} dt + b(T(y))dw_t \right] + \frac{\partial^2 y}{\partial x^2} b^2(T(y))dt$$

We define

$$\hat{a}(y) = \frac{\partial y}{\partial x} a(T(y)) \qquad \hat{b}(y) = \frac{\partial y}{\partial x} b(T(y))$$

and the new equation has the form

$$dy = \hat{a}(y)dt + \hat{b}(y)dw_t + \frac{1}{2}\hat{b}(y)\frac{\partial \hat{b}}{\partial y} dt$$

where we use

$$\frac{\partial \hat{b}}{\partial y} = \frac{\partial b}{\partial y}\frac{\partial y}{\partial x} + \frac{\partial^2 y}{\partial x^2} b(T(y))$$

The stochastic equation corresponds to the white noise limit of

$$\dot{y} = \hat{a}(y) + \hat{b}(y)\xi(t)$$

that is the covariant transformation of the initial equation. Using the notation

$$dx = a(x)dt + b(x) \circ dw_t$$

to denote the equation (**??**), we observe that using a covariant form for the change of variable $x = T(y)$

$$dy = \frac{\partial y}{\partial x}(a(x)dt + b(x) \circ dw_t) = \hat{a}(y)dt + \hat{b}(y) \circ dw_t$$

we recover the Ito form of the stochastic equation associated with the new Stratonovich equation. The FP equation describes the evolution of the distribution function

$$\hat{\rho}(y,t) = \rho(T(y),t) \left| \frac{\partial x}{\partial y} \right|$$

and it has the same form as (**??**).
For stochastically perturbed Hamiltonian systems we have the form (**??**) for the evolution equation

$$dx = J\frac{\partial H_0}{\partial x}dt + \epsilon J\frac{\partial H_1}{\partial x} \circ dw_t = J\frac{\partial H_0}{\partial x}dt + \epsilon J\frac{\partial H_1}{\partial x}dw_t + \frac{\epsilon^2}{2}J\frac{\partial H_1}{\partial x}\frac{\partial}{\partial x}J\frac{\partial H_1}{\partial x}$$

Using the formalism of the Lie derivative the previous equation can be written

$$dx = D_{H_0}x\,dt + \epsilon D_{H_1}x dw_t + \frac{\epsilon^2}{2}D^2_{H_1}x\,dt = D_{H_0}x\,dt + \epsilon D_{H_1}x \circ dw_t$$

To perform a canonical change of variables $x = T(y)$ we can use the covariance nature of the Stratonovich formalism and we get

$$dy = D_{H_0(y)}y\,dt + \epsilon D_{H_1(y)}y \circ dw_t$$

where the Lie derivatives have to be computed in the new variables. In this way one performs a perturbation approach when H_0 is integrable.
We apply the action-angle variables transformation $(q,p) \to (\theta, I)$ so that $H_0(q,p) = H(I)$ and we have a new Hamiltonian

$$H_0(q,p)dt + \epsilon H_1(q,p) \circ dw_t = H_0(I) + \epsilon H_1(\theta, I) \circ dw_t$$

We remark that the stochastic differential equation in the action angle-variables (θ, I)

$$d\theta_j = \Omega_j(I)dt + \epsilon D_{H_1}\theta_j \circ dw_t = \Omega_j(I)dt + \epsilon\frac{\partial H_1}{\partial I_j}dw_t + \frac{\epsilon^2}{2}\left\{\frac{\partial H_1}{\partial I_j}, H_1\right\}dt$$

$$dI_j = \epsilon D_{H_1}I_j \circ dw_t = -\epsilon\frac{\partial H_1}{\partial \theta_j}dw_t - \frac{\epsilon^2}{2}\left\{\frac{\partial H_1}{\partial \theta_j}, H_1\right\}dt$$

We remark that the canonical form of the stochastic equations has to be interpreted. In the limit of small noise $\epsilon \ll 1$ the angles are fast variables and it is possible to average on the angle by considering them uniformly distributed. Applying an averaging principle on the angles variables means that we consider the angle variables as uncorrelated random variables independent of the fluctuations. In such a case the action dynamics reduces

$$dI_j = -\frac{\epsilon^2}{2}\left\langle\left\{\frac{\partial H_1}{\partial \theta_j}, H_1\right\}\right\rangle dt + \epsilon\sqrt{\frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k}}dw_k(t)$$

where we introduce an equivalent stochastic process with covariance matrix

$$C_{jk} = \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k}$$

using independent Wiener process. The FP equation has the form

$$\frac{\partial \bar\rho}{\partial t} = \frac{\epsilon^2}{2}\frac{\partial}{\partial I_j}\left[-\left\langle \left\{ \frac{\partial H_1}{\partial \theta_j}, H_1 \right\} \right\rangle \bar\rho + \bar\rho \frac{\partial}{\partial I_k}\left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k} \right\rangle + \left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k} \right\rangle \right] \frac{\partial \bar\rho}{\partial I_k}$$

We observe

$$\left\langle \left[\frac{\partial^2 H_1}{\partial \theta_j \partial \theta_k}\frac{\partial H_1}{\partial I_k} + \frac{\partial^2 H_1}{\partial I_k \partial \theta_k}\frac{\partial H_1}{\partial \theta_j} \right] \right\rangle = \left\langle \frac{\partial}{\partial \theta_k}\left[\frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial I_k} \right] \right\rangle = 0$$

and we get the self-adjoint FP equation

$$\frac{\partial \bar\rho}{\partial t} = \frac{\epsilon^2}{2}\frac{\partial}{\partial I_j}\left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_j} \right\rangle \frac{\partial \bar\rho}{\partial I_k} \tag{2.3}$$

where we have the approximation

$$\bar\rho(I,t) \simeq \frac{1}{(2\pi)^N}\int_0^{2\pi} \rho(\theta, I, t)\, d\theta$$

This is consistent if we apply the average to the FP equation

$$\frac{\partial \rho}{\partial t} = -\Omega(I)\frac{\partial \rho}{\partial \theta} + \frac{\epsilon^2}{2}\left[\frac{\partial}{\partial \theta}\frac{\partial H_1}{\partial I}\frac{\partial}{\partial \theta}\frac{\partial H_1}{\partial I} - \frac{\partial}{\partial \theta}\frac{\partial H_1}{\partial I}\frac{\partial}{\partial I}\frac{\partial H_1}{\partial \theta} \right.$$
$$\left. - \frac{\partial}{\partial I}\frac{\partial H_1}{\partial \theta}\frac{\partial}{\partial \theta}\frac{\partial H_1}{\partial I} + \frac{\partial}{\partial I}\frac{\partial H_1}{\partial \theta}\frac{\partial}{\partial I}\frac{\partial H_1}{\partial \theta} \right] \rho$$

When we average on θ the expression in the square bracket all the derivatives with respect to θ disappear. Moreover we have

$$-\frac{\partial}{\partial I}\frac{\partial H_1}{\partial \theta}\frac{\partial}{\partial \theta}\frac{\partial H_1}{\partial I} = -\frac{\partial^2}{\partial I \partial \theta}\frac{\partial H_1}{\partial \theta}\frac{\partial H_1}{\partial I} + \frac{\partial}{\partial I}\frac{\partial^2 H_1}{\partial \theta^2}\frac{\partial H_1}{\partial I}$$
$$\frac{\partial}{\partial I}\frac{\partial H_1}{\partial \theta}\frac{\partial}{\partial I}\frac{\partial H_1}{\partial \theta} = \frac{\partial}{\partial I}\left(\frac{\partial H_1}{\partial \theta}\right)^2\frac{\partial}{\partial I} + \frac{\partial}{\partial I}\frac{\partial H_1}{\partial \theta}\frac{\partial^2 H_1}{\partial I \partial \theta}$$

and summing the contributions we get the operator

$$-\frac{\partial^2}{\partial I \partial \theta}\frac{\partial H_1}{\partial \theta}\frac{\partial H_1}{\partial I} + \frac{\partial}{\partial I}\left[\frac{\partial}{\partial \theta}\left(\frac{\partial H_1}{\partial \theta}\frac{\partial H_1}{\partial I} \right) \right] + \frac{\partial}{\partial I}\left(\frac{\partial H_1}{\partial \theta}\right)^2 \frac{\partial}{\partial I}$$

Therefore assuming that the distribution $\rho(\theta, I, t) \simeq \rho(I, t)$ (i.e. the uniform distribution in the angle is invariant for the dynamics) the FP reduces to eq. (**??**). The uniform distribution in the angle depends on the dynamics: we expect that this assumption is valid for the Hamiltonian dynamics since it is related to the invariant measure, but it depends on he thermal bath and on the resonant condition on the unperturbed frequencies $\Omega(I)$.
The Stratonovich interpretation allows to develop a stochastic perturbation theory by changes of variables. The interesting case is when the unperturbed Hamiltonian $H_0(q, p; \lambda)$ depends on a parameter $\lambda = \mu t$. The limit $\mu \to 0$ is the so called adiabatic limit.

2.1 Thermodynamics point of view

In the case the stochastic Hamiltonian (**??**), the linear operator that defines the FP equation are related to the stochastic phase flow Φ^t_ξ. If the perturbation $\xi(t)$ is regular, the phase flow is symplectic but it depends on the noise realization. An averaging process will destroy the symplectic character of the dynamics since

$$J = \left\langle \frac{\partial \Phi^t_\xi}{\partial x}^T \frac{\partial \Phi^t_\xi}{\partial x} \right\rangle \neq \left\langle \frac{\partial \Phi^t_\xi}{\partial x}^T \right\rangle J \left\langle \frac{\partial \Phi^t_\xi}{\partial x} \right\rangle$$

so that the average Jacobian matrix is not symplectic.As a consequence the average probability distribution does not satisfy the Liouville equation. However the Hamiltonian character implies specific properties to the diffusion equation using the Stratonovich interpretation. Stochastic perturbation are considered to model the fast fluctuations of the environment and we have a parametric dependence on the noise

$$H(x;\zeta(t)) = H_0(x) + H_1(x;\zeta(t))$$

$H_1(x;\zeta(t))$ defines the fluctuating part and one assumes $\langle H_1(x;\zeta(t))\rangle = 0$. Then if the evolution time scales of the unperturbed part are slow with respect the fluctuation time scales. We introduce the interaction vision $x = \Phi^t(y)$ where $\Phi^t = \exp(tD_{H_0})$ is the phase flow associated to the unperturbed system. The new Hamiltonian reads

$$H(y,t;\zeta(t)) = H_1(\Phi^t(y);\zeta(t))$$

and the distribution function $\rho_\zeta(y,t)$ satisfies the stochastic Liouville equation

$$\frac{\partial \rho_\zeta}{\partial t} = -D_{H_1(y,t;\zeta)}\rho_\zeta$$

The Lie operators $D_{H_1(y,t;\zeta)}$ do not commute due to the explicit time dependence and the solution can be formally written

$$\rho_\zeta(y,t) = \mathcal{T}\exp\left(-\int_0^t D_{H_1(y,s;\zeta(s))\,ds}\right)\rho_0(y) \tag{2.4}$$

where $\mathcal{T}$ is the time ordering operator. We apply the white noise limit to the stochastic process $H_1(x,\zeta(t))$: i.e. this process can be approximated by a process $\hat{H}_1(x)\xi(t)$ where $\xi(t)$ tends to a white noise and $\hat{H}_1(x)$ defines the local variance of the fluctuations. We use the formal relation

$$\left\langle \mathcal{T}\exp\left(-\int_0^t D_{H_1(y,s;\zeta(s))\,ds}\right)\right\rangle \simeq \mathcal{T}\exp\left(\frac{1}{2}\int_0^t D^2_{\hat{H}_1(\Phi^s(y)}\,ds\right)$$

where we explicitly compute the average operator

$$\begin{aligned}\int_0^t \langle D^2_{H_1(y,s;\zeta(s))}\rangle ds &= \frac{\partial}{\partial y_i}\int_0^t J_{ik}\frac{\partial \hat{H}_1}{\partial y_k}(\Phi^s(y))J_{jh}\frac{\partial \hat{H}_1}{\partial y_h}(\Phi^s(y))\,ds\frac{\partial}{\partial y_j}\\ &= \frac{\partial}{\partial y_i}D_{ij}(\Phi^t(y))\frac{\partial}{\partial y_j}\end{aligned}$$

which define the FP equation for the evolution of the average distribution

$$\frac{\partial \bar{\rho}}{\partial t} = \frac{1}{2}\sum_{ij}\frac{\partial}{\partial y_i} D_{ij}(\Phi^t(y))\frac{\partial \bar{\rho}}{\partial y_j} \tag{2.5}$$

The r.h.s is a self-adjoint operator so that the corresponding stochastic process is reversible and all the eigenvalues are real. The stationary distribution is constant if the diffusion takes place in a compact region. Using the statistical properties of $\hat{H}_1(x)$ the diffusion coefficient can be written in the form

$$D_{ij}(\Phi^t(y)) = J_{ik}\frac{\partial \hat{H}_1}{\partial y_k}(\Phi^t(y))J_{jh}\frac{\partial \hat{H}_1}{\partial y_h}(\Phi^t(y))$$

This equation corresponds to the stochastic differential equation

$$\dot{y}_i = \sum_k J_{ik}\frac{\partial \hat{H}_1}{\partial y_k}(\Phi^t(y))\xi(t)$$

in the Stratonovich interpretation for the white noise. Since the equation has a canonical form, let $x = \Phi^t(y)$ we perform the change of variable using the covariance property of the Stratonovich interpretation

$$\dot{x}_i = \sum_k J_{ik}\left(\frac{\partial H_0}{\partial x_k}(x) + \frac{\partial \hat{H}_1}{\partial x_k}(x)\xi(t)\right) \tag{2.6}$$

and the corresponding FP equation reads

$$\frac{\partial \bar{\rho}}{\partial t} = -J_{ik}\frac{\partial H_0}{\partial x_i}(x)\frac{\partial \bar{\rho}}{\partial x_k} + \frac{1}{2}\sum_{ij}\frac{\partial}{\partial x_i}D_{ij}(x)\frac{\partial \bar{\rho}}{\partial x_j} \tag{2.7}$$

The self-adjoint nature of the operator implies that the stationary distribution of eq. (**??**) is constant in general (i.e. if the unpertburbed Hamiltonian H_0 and the perturbation Hamiltonian H_1 has no common integral of motion F). Then we have a Maximal Entropy solution without any constraint since the using the Gibbs Entropy we get

$$\frac{dS}{dt} = -\int \frac{\partial \bar{\rho}}{\partial t}\ln \bar{\rho}\, dx = -\int J\frac{\partial H_0}{\partial x}\frac{\partial \bar{\rho}}{\partial x}\, dx + \frac{1}{2}\int \frac{\partial \bar{\rho}}{\partial x}D(x)\frac{\partial \bar{\rho}}{\partial x}\, dx$$

and we prove $dS/dt > 0$ since the first term vanishes (it is a divergence of a vector field) and the second term is positive defined if $D(x)$ is a positive defined symmetric matrix so that $dS/dt = 0$ implies $\bar{\rho} = const.$.

2.2 Reversibility character of diffusion processes

The concept of stochastic reversibility can be formulated as follows: if $x(t)$ is a possible realization of a trajectory of a stochastic system in a stationary solution (i.e. the statistical properties of the system are described by the stationary distribution $\rho^s(x)$) then $x(-t)$ is a possible trajectory with the same probability: i.e. the transition probability $\mathcal{P}(x, t + \Delta t|y, t)$ satisfies the detailed balance condition

$$\mathcal{P}(x, \Delta t|y)\rho^s(y) = \mathcal{P}(y, \Delta t|x)\rho^s(x) \tag{2.8}$$

We recall that the conditional probability is the solution of the FP equation with initial condition $\delta(x_0 - y)$ and it can be formally written

$$\mathcal{P}(x, \Delta t|y) = \left\langle \delta(x - \Phi_\xi^{\Delta t}(y)\right\rangle \qquad \lim_{\Delta t \to 0} \mathcal{P}(x, \Delta t|y) = \delta(x - y)$$

where the expectation value is over all the realizations of the noie $\xi(t)$, and in the white noise limit we have

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int \mathcal{P}(x, \Delta t|y) \rho_0(y)\, dy = -\mathcal{L}_{FP}(x) \rho_0(x)$$

with $\mathcal{L}_{FP}$ the Fokker-Planck operator in the Stratonovich interpretation. The previous relations imply

$$\Pi^{\Delta t}(x, y) = \frac{1}{\sqrt{\rho^s(x)}} \mathcal{P}(x, \Delta t|y) \sqrt{\rho^s(y)} = \frac{1}{\sqrt{\rho^s(y)}} \mathcal{P}(y, \Delta t|x) \sqrt{\rho^s(x)} = \Pi^{\Delta t}(y, x) \tag{2.9}$$

which means that the matrix on the l.h.s. is symmetric;

$$\int f(x) \cdot \Pi^{\Delta t}(x, y) g(y)\, dy\, dx = \int [\Pi^{\Delta t}(y, x)]^T f(x) \cdot g(y)\, dx\, dy = \int \Pi^{\Delta t}(y, x) f(x) \cdot g(y)\, dx\, dy$$

Then the operator

$$\mathcal{S}_{FP} = \frac{1}{\sqrt{\rho^s}} \mathcal{L}_{FP} \sqrt{\rho^s}$$

has to be self-adjoint since the relation between the evolution matrix and the operator S_{FP} is understood by

$$\Pi^{\Delta t}(x, y) = \exp(-S_{FP}(x)\Delta t) \delta(x - y)$$

and the symmetry condition implies

$$\int f(x) \cdot \exp(-S_{FP}(x)\Delta t) g(x)\, dx = \int \exp(-S_{FP}(y)\Delta t) f(y) \cdot g(y)\, dy$$

so that $S_{FP}^\dagger = S_{FP}$.
Let us consider the case

$$\mathcal{L}_{FP} = \frac{\partial}{\partial x} \frac{\partial V}{\partial x} + T \frac{\partial^2}{\partial x^2}$$

where $\rho^s \propto \exp(-V(x)/T)$, an explicit computation gives

$$\frac{1}{\sqrt{\rho^s}} \mathcal{L}_{FP} \sqrt{\rho^s} = \frac{1}{\sqrt{\rho^s}} \frac{\partial}{\partial x} \left[\frac{\partial V}{\partial x} \sqrt{\rho^s} + T \frac{\partial}{\partial x} \sqrt{\rho^s} \right] =$$
$$\frac{1}{\sqrt{\rho^s}} \frac{\partial}{\partial x} \left[\frac{1}{2} \frac{\partial V}{\partial x} \sqrt{\rho^s} + T \sqrt{\rho^s} \frac{\partial}{\partial x} \right] = \frac{1}{2} \frac{\partial^2 V}{\partial x^2} - \frac{1}{T} \left(\frac{1}{2} \frac{\partial V}{\partial x} \right)^2 + T \frac{\partial^2}{\partial x^2}$$

and we get a self-adjoint operator. In a generic case if we know the stationary solution, we set

$$-T \ln \rho^s = \frac{\partial V}{\partial x} \quad \Rightarrow \quad a(x) = -\frac{\partial V}{\partial x} + a_{rot}(x)$$

The Kirkhoff condition implies div $J^s = 0$ for the stationary density currents

$$J^s(x) = -\frac{\partial V}{\partial x}\rho^s + a_{rot}\rho^s - T\frac{\partial \rho^s}{\partial x} = a_{rot}\rho^s \neq 0$$

so that

$$\frac{1}{\rho^s}\frac{\partial}{\partial x}a_{rot}\rho^s = \frac{\partial a_{rot}}{\partial x} - \frac{1}{T}\frac{\partial V}{\partial x}a_{rot} = 0$$

If if the gradient of the potential is orthogonal to the field a_{rot} then a_{rot} is a free divergence field and vice versa: this is the case of Hamiltonian systems. This means that $V(x)$ is an integral of motion for the vector field a_{rot} and it is possible to compute the stationary condition from the vector field. However a term a_{rot} destroys the adjoint properties of the FP operator since the term

$$\frac{1}{\sqrt{\rho^s}}\frac{\partial}{\partial x}a_{rot}\sqrt{\rho^s} = \frac{1}{2}\frac{\partial a_{rot}}{\partial x} + a_{rot}\frac{\partial}{\partial x}$$

is not self adjoint even in the case div $a_{rot} = 0$. However the free divergence condition implies

$$a_{rot}\frac{\partial V}{\partial x} = 0$$

i.e. the rotor field is orthogonal to the gradient of the internal energy. In this case the stochastic reversibility is recovered is we introduce the involution $x \to -x$: i.e. the reverse fluctuation at a state x observe in stationary condition corresponds to a direct fluctuation but at a state $-x$.

2.3 Correlation properties of fluctuations

The symmetry condition implies that the correlation among the fluctuations is symmetric if computed in the future or in the past. The correlation concept is related to definition of the expectation value for the fluctuations at different times in the stationary state. We consider a generic stochastic differential equation

$$\dot{x} = a(x) + \epsilon\xi(t)$$

without loss of generality we assume div $a(x) = 0$ so that the phase flow is volume preserving. Let $x(t)$ a solution of the unperturbed system with $x(0) = x_0$ we have the linearized dynamics

$$\delta\dot{x} = \frac{\partial a}{\partial x}(x(t))\delta x + \epsilon\xi(t) = A(t)\delta x + \epsilon\xi(t)$$

and we solve the linear dynamics using the principal matrix Φ_0^t

$$\delta x(t) = \epsilon\int_0^t \Phi_s^t\xi(s)\,ds \qquad \delta x(0) = 0$$

Then one considers the expectation value

$$\langle\delta x(t+\tau)\delta x(t)\rangle = C(\tau;t) \tag{2.10}$$

The integral converges if the Ljapunov exponent of the system are all negative: i.e. the matrix

$$\lim_{t\to\infty}\int_0^t \left(\Phi_0^s[\Phi_0^s]^T\right)ds$$

exists. In the white noise limit we compute

$$\langle \delta x_i(t+\tau)\delta x_j(t)\rangle = \epsilon^2 \int_0^{t+\tau}\int_0^t \left\langle [\Phi_s^{t+\tau}]_{ik}\xi_k(s)\,[\Phi_u^t]_{jh}\xi_h(u)\right\rangle dsdu$$
$$= \epsilon^2[\Phi_t^{t+\tau}]_{il}\int_0^t [\Phi_s^t]_{lk}\,[\Phi_s^t]^T_{kj}\,ds = \epsilon^2[\Phi_t^{t+\tau}]_{il}\Sigma_{lj}(t)$$

where $\Sigma(t)$ is the covariance matrix. When $t \gg 1$ the integral converges so the $c(\tau : t) \propto \Phi_t^{t+\tau}$ and when τ increases it decreases according to the maximum Ljapunov exponents. The linearization procedure requires $\epsilon \ll 1$ to be applied. The computation of (**??**) can be performed by using the solution of the FP equation

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x}A(t)x\rho + \frac{\epsilon^2}{2}\frac{\partial^2\rho}{\partial x^2}$$

The fundamental solution $\rho(\delta x, t\,|\,0)$ gives the probability to be in the state δx after a time t with initial condition $\lim_{t\to 0}\rho(\delta x, t\,|\,0) = \delta(0)$. By definition we have

$$\rho(\delta x, t\,|\,0) = \left\langle \delta\left(\delta x - \epsilon\int_0^t \Phi_s^t \xi(s)\,ds\right)\right\rangle$$

The correlation can be computed by

$$\langle \delta x_i(t+\tau)\delta x_j(t)\rangle = \int \delta x_i \delta y_j \rho(\delta x, t+\tau\,|\,\delta y, t)\rho(\delta y, t\,|\,0)\,d\delta x d\delta y$$

and we use the equality

$$\rho(\delta x, t+\tau\,|\,\delta y, t) = \left\langle \delta\left(\delta x - \Phi_t^{t+\tau}\delta y - \epsilon\int_t^{t+\tau}\Phi_s^{t+\tau}\xi(s)\,ds\right)\right\rangle$$

Then we compute

$$\langle \delta x_i(t+\tau)\delta x_j(t)\rangle = \int \left\langle \left([\Phi_t^{t+\tau}]_{il}\delta y_l - \epsilon\int_t^{t+\tau}[\Phi_s^{t+\tau}]_{il}\xi_l(s)\right)\right\rangle \delta y_j \rho(\delta y, t\,|\,0)\,d\delta y$$
$$= [\Phi_t^{t+\tau}]_{il}\int \delta y_l \delta y_j \rho(\delta y, t\,|\,0)\,d\delta y = [\Phi_t^{t+\tau}]_{il}\Sigma_{lj}(t)$$

recovering the previous result.
Let us consider the correlation with past events (i.e. we reverse the process)

$$\langle \delta x_i(t-\tau)\delta x_j(t)\rangle = \epsilon^2 \int_0^{t-\tau}\int_0^t \left\langle [\Phi_s^{t-\tau}]_{ik}\xi_k(s)\,[\Phi_u^t]_{jh}\xi_h(u)\right\rangle dsdu$$
$$= \epsilon^2\int_0^{t-\tau}[\Phi_s^{t-\tau}]_{lk}\,[\Phi_s^{t-\tau}]^T_{kl}\,ds[\Phi_{t-\tau}^t]^T_{lj} = \epsilon^2\Sigma_{il}(t-\tau)[\Phi_{t-\tau}^t]^T_{lj}$$

Therefore we have the relation

$$C_{ij}(\tau; t) = C_{ji}(-\tau; t+\tau)$$

Assuming the existence of a stationary limit $t \to \infty$ with $A(t) = A$ the reversibility condition is a symmetric condition

$$\exp(A\tau)\Sigma = \Sigma\exp(A^T\tau) \quad \Rightarrow \quad \Sigma^{-1/2}\exp(A\tau)\Sigma^{1/2} = \Sigma^{1/2}\exp(A^T\tau)\Sigma^{-1/2}$$

which means that $\Sigma^{-1/2}\exp(A\tau)\Sigma^{1/2}$ is a symmetric matrix. We recall that the stationary distribution is Gaussian with a covariance matrix Σ can be explicitly computed and the previous relation reads

$$\int_0^t \exp(A(t-s))\,\exp(A\tau)\exp(A^T(t-s))\,ds = \int_0^t \exp(A(t-s))\,\exp(A^T\tau)\exp(A^T(t-s))\,ds$$

that implies that $\exp(A\tau) = \exp(A^T\tau)$ and A is a symmetric matrix.
This condition generalized to nonlinear case and it follows that the field $a(x) = -\partial V/\partial x$.
The correlation among the fluctuations can be studied for any observable $I(x)$ by defining

$$I(t; I_0) = \int_{I(x_0)=I_0} \langle I(\Phi^t(x_0))\rangle\, dx_0$$

In this case the evolution depends on the adjoint operator $\mathcal{L}^{\dagger}_{FP}$ and for the stationary value is constant

$$I_s = \int I(x)\rho_s(x)\, dx$$

We compute

$$\begin{aligned}\langle \delta I(t+\tau)\delta I(t)\rangle &= \epsilon^2 \int_0^{t+\tau}\int_0^t \left\langle [\Phi_s^{t+\tau}]_{ik}\xi_k(s)\,[\Phi_u^t]_{jh}\xi_h(u)\right\rangle\, dsdu \\ &= \epsilon^2[\Phi_t^{t+\tau}]_{il}\int_0^t [\Phi_s^t]_{lk}\,[\Phi_s^t]^T_{kj}\,ds = \epsilon^2[\Phi_t^{t+\tau}]_{il}\Sigma_{lj}(t)\end{aligned}$$

*****

Chapter 3

Thermal Baths and Hamiltonian Systems

If the Hamiltonian (**??**) simulates the interaction dynamics between a test particle and the fluctuating nonhomogeneous environment, the missing of a dissipative term implies that the energy H_0 is not conserved in a average sense. The Hamiltonian (**??**) models the effects of a fluctuating potential in the environment, but there does not exist a stationary state due to average flux of energy between the system and the environment. If we assume that in a stationary the system energy has to be preserved in average, the previous equation (**??**) is not suitable to describe the evolution and a possible solution is to introduce a dissipative term in the equation of motion

$$\dot{x} = D_{H_0}x + \left(-\{H_0, H_1\} + \sqrt{2T}\xi(t)\right) D_{H_1}x \tag{3.1}$$

where $x = (q, p)$ and $\xi(t)$ is a white noise.
Remark: the previous equation is not Hamiltonian due to the presence of the dissipation: indeed the neglecting the fluctuations we have

$$\frac{dH_0}{dt} = -\{H_0, H_1\}D_{H_1}H_0 = -\{H_0, H_1\}^2 < 0$$

and the system collapses to the critical points $\partial H_0/\partial x = 0$. The physical meaning is that of the equation (**??**) is that there should be a relation between the energy fluctuations variance, defined by the $T(D_{H_1}H_0)^2 = T(\{H_0, H_1\})^2$ and the local dissipation field introduced by the term $\{H_0, H_1\}D_{H_1}H_0 = (D_{H_1}H_0)^2$: i.e. the variance of the fluctuations turns out to be proportional to the dissipation (fluctuation-dissipation relation). As an example if we set $H_1 = -\sqrt{m\gamma}q$ and $H_0 = p^2/2m + V(q)$, the stochastic equation (**??**) reads

$$\begin{aligned} \dot{q} &= \frac{p}{m} \\ \dot{p} &= -\frac{\partial V}{\partial q} - \gamma p + \sqrt{2mT\gamma}\xi(t) \end{aligned} \tag{3.2}$$

where T is the temperature and we satisfy the Einstein relation. The Fokker Planck equation (**??**) reads

$$\frac{\partial \rho}{\partial t} = -D_{H_0}\rho + \gamma\frac{\partial}{\partial p}\frac{p}{m}\rho + mT\gamma\frac{\partial^2}{\partial p^2}\rho$$

and we recover the usual stochastic model for a thermal bath.
The corresponding FP equation in the limit $\xi(t) \to$ white noise is written using the Stratonovich interpretation of the stochastic equation

$$\frac{\partial \rho}{\partial t} = -\left(D_{H_0} - D_{H_1}\{H_0, H_1\}\right)\rho + T D^2_{H_1}\rho = 0 \tag{3.3}$$

where we use the property

$$\frac{\partial}{\partial x_j} J_{jk} \frac{\partial H_1}{\partial x_k} f = J_{jk} \frac{\partial H_1}{\partial x_k} \frac{\partial f}{\partial x_j} = D_{H_1} f$$

so that

$$\frac{\partial}{\partial x}\{x, H_1\} \frac{\partial}{\partial x}\{x, H_1\}\rho = D^2_{H_1}\rho$$

The stationary solution can be computed in the form $\rho_s(H_0)$ so that:

$$0 = \left(-D_{H_0} + D_{H_1}\{H_0, H_1\} + T D^2_{H_1}\right)\rho_s(H_0) =$$
$$D_{H_1}\left(\{H_0, H_1\}\rho_s(H_0) + T D_{H_1}\rho_s(H_0)\right)$$

Then assuming a detailed balance (DB) condition and we reduce to the equation

$$\{H_0, H_1\} = -T\{H_0, H_1\}\frac{d}{dH_0}\ln(\rho_s(H_0))$$

whose solution is

$$\ln(\rho_s(H_0)) = -\frac{H_0}{T} + const. \qquad \Rightarrow \qquad \rho_s(H_0) \propto \exp\left(-\frac{H_0}{T}\right) \tag{3.4}$$

Remark: in a generic case of a FP equation

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x} a(x)\rho + T\frac{\partial^2 \rho}{\partial x^2} = -\frac{\partial}{\partial x} J(x,t)$$

for a system in a thermal bath: if the deterministic field can be factorized

$$a(x) = a_r(x) - \frac{\partial I}{\partial x}$$

where $I(x)$ is a first integral of motion for $a_r(x)$ and $a_r(x)$ is a zero-divergence field

$$a_r(x)\frac{\partial I}{\partial x} = 0 \qquad \frac{\partial a_r}{\partial x} = 0$$

Then stationary solution of the FP equation

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x}\left[a_r(x)\rho - \frac{\partial I}{\partial x}\rho - T\frac{\partial \rho}{\partial x}\right]$$

can be computed in the form $\rho_s(x) \propto \exp(-I(x)/T)$ using the relations

$$T\frac{\partial}{\partial x}\rho_s = -\frac{\partial I}{\partial x}\rho_s \qquad \frac{\partial}{\partial x} a_r(x)\rho_s = \frac{\rho_s}{T} a_r(x)\frac{\partial I}{\partial x}$$

We observe that

$$\mathrm{rot}\, a(x) = \mathrm{rot} a_r(x)$$

In this case the stationary currents do not vanish

$$J_s(x) = a(x)\rho_s(x) - T\frac{\partial \rho_s}{\partial x} = a_r(x)\rho_s(x)$$

but we have the relation

$$J_s(x)\frac{\partial \rho_s}{\partial x} = 0 \tag{3.5}$$

which means that the local entropy production vanishes (the DB condition is usually formulated $J_s(x) = 0$). A direct computation gives the entropy production

$$\frac{dS}{dt} = -\int \frac{J}{\rho}\frac{\partial \rho}{\partial x}\, dx$$

so that the condition (**??**) implies that the local entropy production vanishes at any point at the equilibrium condition.
We recover the Maxwell-Boltzmann equilibrium for the distribution function that is independent from the choice of H_1 where T is the temperature and the proportionality factor is inverse of the partition function

$$A(T) = \int \exp\left(-\frac{H_0(x)}{T}\right) dx$$

that relate the statistical mechanics with the thermodynamics formalism. In the case of a mechanical system (**??**) the stationary solution factorizes

$$\rho_s(q,p) \propto \exp\left(-\frac{p^2}{2mT}\right) \exp\left(-\frac{V(q)}{T}\right)$$

so that the definition of temperature follows

$$\frac{T}{2} = \left\langle \frac{p^2}{2m} \right\rangle$$

We introduce a thermodynamics formalism assuming that potential $V(q,\lambda)$ depends on a parameter λ so that the change in the parameter corresponds to the work performed on the system when λ is varying reads

$$\frac{đW}{d\lambda} = \int \frac{\partial V}{\partial \lambda}\rho(q,p,\lambda)\, dq\, dp = \int \frac{\partial H_0}{\partial \lambda}\rho(q,p,\lambda)\, dq\, dp \tag{3.6}$$

Using the FP operator

$$\mathcal{F}_{FP} = -D_{H_0(\lambda)} + \gamma\frac{\partial}{\partial p}p + T\gamma\frac{\partial^2}{\partial p^2}$$

for a given value of λ we describe the relaxation process

$$\frac{\partial \rho}{\partial t} = -D_{H_0}\rho - \frac{\partial}{\partial p}J(q,p)$$

by defining the probability currents

$$J(q,p) = \gamma\left[-p\rho - T\frac{\partial \rho}{\partial p}\right] = \gamma\left[-\frac{\partial H_0}{\partial p}\rho - T\frac{\partial \rho}{\partial p}\right]$$

Remark: the Hamiltonian evolution is reversible and cannot influence the relaxation process, but it changes the definition of probability currents. The energy change in the relaxation process is computed by the adjoint operator

$$\frac{đQ}{dt} = \int \mathcal{F}^{\star}_{FP} H_0(q,p,\lambda)\rho(q,p,\lambda)\, dq\, dp = \gamma \int \left[-p\frac{\partial H_0}{\partial p} + T\frac{\partial^2 H_0}{\partial p^2} \right] \rho(q,p,\lambda)\, dq\, dp$$

Then we get the relation

$$\frac{đQ}{dt} = \int \frac{\partial H_0}{\partial p} J(q,p)\, dq\, dp = -\int \frac{J^2(q,p)}{\rho}\, dq\, dp - T\int J(q,p)\frac{\partial \ln\rho}{\partial p}\, dq\, dp$$

The first term is always positive and it represent the dissipated energy that contributes to the total increases of entropy of the universe due to the non-reversible character of the transformation. The second term is the contribution to the internal energy. This is the heat exchanged by the system. In the case of adiabatic transformation $\lambda = \epsilon t$ $\epsilon \to 0$ we approximate

$$\rho(q,p,\lambda) = A(T,\lambda) \exp\left(-\frac{H_0(q,p,\lambda)}{T} \right)$$

In such a case $đQ = 0$ and we have

$$\frac{\partial}{\partial \lambda} \ln A(T,\lambda) = -T\int \frac{\partial V}{\partial \lambda} \rho_s(q,p,\lambda)\, dq\, dp = -\frac{đW}{d\lambda}$$

Then it follows that the change of the free energy corresponds to the work performed on the system in a adiabatic transformation the change

$$\frac{\partial F}{\partial \lambda} = \frac{đW}{d\lambda}$$

We consider the change of the Gibbs Entropy in the isothermal transformation (i.e. the entropy production)

$$\frac{dS}{dt} = -\int \frac{\partial \rho}{\partial t} \ln\rho\, dp\, dq = -\int \left[-D_{H_0}\rho + \gamma\frac{\partial}{\partial p} p\rho + T\gamma\frac{\partial^2 \rho}{\partial p^2} \right] \ln\rho\, dp\, dq$$

We observe that the natural boundary conditions implies

$$\int (D_{H_0}\rho) \ln\rho\, dx = -\int D_{H_0}\rho = 0$$

since the Hamiltonian fields have zero divergence, and the Lie derivative does not contribute to the entropy production. The remaining terms can be written using the density currents

$$\frac{dS}{dt} = \int \frac{\partial J}{\partial p} \ln\rho\, dx = -\int \frac{J}{\rho}\frac{\partial \rho}{\partial p}\, dp\, dq$$

Using

$$-\frac{\partial \rho}{\partial p} = \frac{J}{T\gamma} + \frac{p}{T}\rho$$

we get

$$T\frac{dS}{dt} = \gamma^{-1}\int \frac{J^2}{\rho}\, dp\, dq + \int Jp\, dp\, dq$$

The first term is always positive and can be interpreted as a change in the total entropy due to the transformation whereas the second term is related to the change of internal energy during the transformation or the relaxation process

$$\frac{dE}{dt} = \int H_0 \mathcal{L}_{FP} \rho \, dp \, dq = \int H_0 \frac{\partial}{\partial p} J \, dp \, dq = \int \frac{\partial H_0}{\partial p} J \, dp \, dq$$

i.e. the work performed by the system decreases the internal energy. This implies that if the initial distribution satisfies $\langle H_0 \rangle = E(T)$ (i.e. the asymptotic energy) the total work of the internal force has to vanish after the relaxation process, but the total entropy increases in any case.
To better explain the relaxation process one defines the non-equilibrium free energy

$$F = E - TS = \int (H_0(q,p) + T \ln \rho(q,p,t)) \, \rho(q,p,t) \, dp \, dq \tag{3.7}$$

which reduce to $-T \ln A(T)$ at equilibrium. Then it follows

$$\frac{dF}{dt} = \int \frac{\partial H_0}{\partial p} J \, dp \, dq - \gamma^{-1} \int \frac{J^2}{\rho} \, dp \, dq - \int J \frac{\partial H_0}{\partial p} \, dp \, dq = -\gamma^{-1} \int \frac{J^2}{\rho} \, dp \, dq$$

and in the relaxation process

$$\Delta F \leq 0 \tag{3.8}$$

which corresponds to the Second Principle of Thermodynamics. It is possible interpret the relaxation process using the relative entropy (Kullback-Leiber divergence) of two probability distribution

$$D_{KL}(\rho \parallel \rho_s) = \int \rho(q,p) \ln \frac{\rho(q,p)}{\rho_s(q,p)} \, dp \, dq \tag{3.9}$$

It is a measure of how much an approximating probability distribution ρ is different from a reference probability distribution ρ_s. We have the properties

$$\begin{aligned} & D_{KL}(\rho \parallel \rho_s) \geq 0 \\ & D_{KL}(\rho \parallel \rho_s) \neq D_{KL}(\rho \parallel \rho_s) \\ & D_{KL}(\rho \parallel \rho_s) = 0 \quad \Leftrightarrow \quad \rho_s = \rho \end{aligned}$$

The first inequality follows from $\ln x \geq 1 - 1/x$ when $x \geq 0$. Then

$$\int \rho_s(q,p) \ln \frac{\rho(q,p)}{\rho(q,p)} \, dp \, dq \geq \int \rho_s(q,p) \left(1 - \frac{\rho(q,p)}{\rho_s(q,p)} \right) \, dp \, dq = 0$$

By definition we have

$$\int \rho(q,p) \ln \frac{\rho(q,p)}{\rho_s(q,p)} \, dp \, dq = -S[\rho] - \frac{-F_s + \langle H_0 \rangle_\rho}{T} = \frac{\langle H_0 \rangle_\rho - TS(\rho) - F_s}{T}$$

where we recognize the non-equilibrium free energy. The minimum value is for $\rho = \rho_s$ and if $\langle H_0 \rangle_\rho = E_s$ it follows $S(\rho) \leq S(\rho_s)$ which corresponds to the maximum entropy principle. Then we compute

$$\frac{d}{dt} D_{KL}(\rho(q,p,t) \parallel \rho_s(q,p)) = \int \frac{\partial J}{\partial p} \ln \frac{\rho(q,p)}{\rho_s(q,p)} \, dp \, dq = - \int \frac{\partial J}{\partial p} \left[\ln \rho(q,p,t) - \ln \rho_s(q,p) \right] \, dp \, dq$$

$$= -\gamma^{-1} \int \frac{J^2}{\rho} \, dp \, dq - \int J \frac{\partial H_0}{\partial p} \, dp \, dq + \int J \frac{\partial H_0}{\partial p} \, dp \, dq = -\gamma^{-1} \int \frac{J^2}{\rho} \, dp \, dq < 0$$

The non-equilibrium behavior can be explained using a thermodynamics formalism: in the case we have a dependence $H_0(q,p,\lambda)$ and we simulate a thermal bath, we consider the quantity

$$\chi(q,p,\lambda) = \rho(q,p,\lambda)\,\langle \exp\left(-W(q,p,t)\right)\rangle_{(q,p)} \tag{3.10}$$

where $W(q,p,t)$ is the average work associated to a transformation $\lambda = \epsilon t$ for all the realizations passing through the state (q,p). Then we have the equation

$$\frac{\partial \chi}{\partial t} = \left[\mathcal{L}_{FP} - \frac{\dot{\lambda}}{T}\frac{\partial H_0}{\partial \lambda}\right]\chi$$

whose solution is simply

$$\chi(q,p,t) \propto \exp\left(-\frac{H_0(q,p,\lambda(t))}{T}\right)$$

according to the definition (**??**). Assuming an equilibrium initial condition

$$\rho_0(q,p,\lambda_0) = A^{-1}(T,\lambda_0)\exp\left(-\frac{H_0(x,\lambda_0)}{T}\right)$$

we have

$$\chi(x,t) = A^{-1}(T,\lambda_0)\exp\left(-\frac{H_0(x,\lambda_1)}{T}\right) = \frac{A(\lambda_1,T}{A(\lambda_0,T)}\rho_s(q,p,\lambda_1)$$

But using the definition (**??**) one gets the relation

$$\langle \exp\left(-W(q,p,\lambda_1)\right)\rangle = \int \chi(q,p,\lambda_1)\,dq\,dp = \frac{A(\lambda_1,T}{A(\lambda_0,T)} = \exp\left(-\frac{F(T,\lambda_1)-F(T\lambda_0)}{T}\right) \tag{3.11}$$

where we use the definition

$$\Delta F = -T\ln\frac{A(\lambda_1,T}{A(\lambda_0,T)}$$

This is the content of the **Jarzinski equality** that implies the 2nd Principle of Thermodynamics: using the Jensen inequality

$$\langle e^G\rangle \geq e^{\langle G\rangle}$$

it follows

$$\exp\left(-\frac{F(T,\lambda_1)-F(T\lambda_0)}{T}\right) \geq \exp\left(-\langle W(q,p,\lambda_1)\rangle\right)$$

$$\Delta F \leq T\langle W(q,p,\lambda_1)\rangle \tag{3.12}$$

3.0.1 Detailed balance condition and stochastic reversibility

The FP equation is a continuity equation for the probability density

$$\frac{\partial \rho}{\partial t} = \mathcal{L}_{FP}\rho$$

whose solutions can be interpreted as the transition probabilities among the microstates when the initial condition is $\delta(x-y)$

$$\exp\left(\mathcal{L}_{FP}(x)t\right)\delta(x-y) = \pi^t(x|y)$$

Remark we assume that the system is not explicitly time dependent. It is possible to prove that there always exists a stationary solution

$$\int \pi^t(x|y)\rho_s(y)\,dy = \rho_s(x)$$

since the matrix $\pi^t(x|y)$ is stochastic

$$\int \pi^t(x|y)\,dx = 1$$

The other eigenvalues satisfy $\mathrm{Re}\,\lambda < 0$ and define the relaxation process when the system is out of equilibrium: the study of the spectral properties of the operator $\mathcal{L}_{FP}$ is a key issue for non-equilibrium thermodynamics.
In the limit $t \to 0$ one gates the transition probability rates

$$\hat{\pi}(x|y) = \lim_{t\to 0}\frac{\pi^t(x|y)}{t} = \mathcal{L}_{FP}(x)\delta(x-y) \qquad x \neq y \tag{3.13}$$

If we define

$$\hat{\pi}(x|x) = \int \hat{\pi}(y|x)\,dy$$

The matrix

$$\mathcal{L}(x|y) = \hat{\pi}(x|x)\delta(x-y) - \hat{\pi}(x|y)$$

is a Laplacian matrix and the FP equation can be written in the form or a *master* equation

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= \int \mathcal{L}_{FP}(x)\delta(x-y)\rho(y,t)\,dy = \int \left[\hat{\pi}(x|y)\rho(y,t) - \hat{\rho}(y|x)\rho(x,y)\right]\,dy\\ &= -\int \mathcal{L}(x|y)\rho(y,t)\,dy\end{aligned} \tag{3.14}$$

The spectral properties of the FP operatr is a special case of the spectral properties of the Laplacian matrix. The master equation is a continuity equation and we define the density probability currents between two states

$$J(x,y,t) = \hat{\pi}(x|y)\rho(y,t) - \hat{\pi}(y|x)\rho(x,t)$$

A system satisfies the detailed balance (DB) condition if $J(x,y,t) = 0$ if the stationary currents vanish between any couple of states

$$\hat{\pi}(x|y)\rho_s(y) - \hat{\pi}(y|x)\rho_s(x) \tag{3.15}$$

The DB condition in a symmetry condition for the transition rate matrix. Using the FP operator we have

$$\mathcal{L}_{FP}(x)\delta(x-y)\rho_s(y) = \mathcal{L}_{FP}(y)\delta(y-x)\rho_s(x)$$

We observe that performing a change of variables $\rho \to \sqrt{\rho_s}\rho$the previous condition reads

$$\frac{1}{\sqrt{\rho_s(x)}}\mathcal{L}_{FP}(x)\delta(x-y)\sqrt{\rho_s(y)} = \frac{1}{\sqrt{\rho_s(y)}}\mathcal{L}_{FP}(y)\delta(y-x)\sqrt{\rho_s(x)}$$

that is a symmetry condition for the operator.
We have the *Lemma*: if the operator $\mathcal{L}(x)\delta(x-y)$ is symmetric then $\mathcal{L}(x)$ is self-adjoint.
Proof: a direct calculation implies

$$\mathcal{L}(x)f(x) = \int \mathcal{L}(x)\delta(x-y)f(y)\,dy = \int \mathcal{L}(y)\delta(y-x)f(y)\,dy$$
$$= \int \delta(y-x)\mathcal{L}^*(y)f(y)\,dy = \mathcal{L}^*(x)f(x)$$

where we use the symmetry condition. It follows $\mathcal{L}(x) = \mathcal{L}^*(x)$ ♯.
In a general case

$$\mathcal{L}_{FP} = \frac{\partial}{\partial x}\left[-a(x) + T\frac{\partial}{\partial x}\right]$$

the DB condition implies that the equilibrium distribution satisfies to

$$J_i^s(x) = \left[a_i(x) - T\frac{\partial}{\partial x_i}\right]\rho_s(x) = 0$$

This condition is an exact condition on the field $a(x)$

$$\frac{a_i(x)}{T} = \frac{\partial}{\partial x_i}\ln\rho_s(x)$$

which admits a potential: i.e. the $V(x) = \ln\rho_s(x)$. This implies that $a_i(x)$ defines the internal energy $V(x)$ and the equilibrium solution is

$$\rho_s(x) \propto \exp\left(-\frac{V(x)}{T}\right)$$

In the Hamiltonian case, we show that the operator

$$\frac{1}{\sqrt{\rho_s}}\frac{\partial}{\partial p}\left[p + T\frac{\partial}{\partial p}\right]\sqrt{\rho_s} \qquad \rho_s(p) \propto \exp\left(-\frac{p^2}{2T}\right)$$

is self-adjoint. We have

$$\frac{1}{\sqrt{\rho_s}}\frac{\partial}{\partial p}\left[-\frac{T}{2\sqrt{\rho_s}}\frac{\partial\rho_s}{\partial p} + T\sqrt{\rho_s}\frac{\partial}{\partial p}\right]$$
$$= \frac{1}{\sqrt{\rho_s}}\frac{\partial}{\partial p}\left[\frac{p}{2}\sqrt{\rho_s} + T\sqrt{\rho_s}\frac{\partial}{\partial p}\right] = \frac{1}{2} + \frac{p}{4\rho_s}\frac{\partial\rho_s}{\partial p} + T\frac{\partial^2}{\partial p^2}$$

However this is not true for the whole operator

$$\mathcal{L}_{FP} = -\frac{\partial}{\partial x}J\frac{\partial H_0}{\partial x} + \frac{\partial}{\partial p}\left[-\frac{T}{2}\frac{\partial\rho_s}{\partial p} + T\frac{\partial}{\partial p}\right]$$

since the Lie derivative D_{H_0} implies the existence of a non-zero stationary currents, but these currents are orthogonal to the gradient of equilprobability surfaces $\rho_s(x) = const.$. Then they do not contribute to the entropy production cucu

3.1 Perturbed Stochastic Hamiltonians

The Stratonovich interpretation allows to apply a covariant formalism to the stochastic equation (**??**) by changing the Hamiltonian function consistently. Assume that H_0 is almost-integrable we can introduce that action-angle variables and we have

$$H_0(I,\theta) = \hat{H}_0(I) + \epsilon \hat{H}(I,\theta)$$

Moreover the covariant properties of the Poisson bracket implies the stochastic system in the new variables can be written by computing $H_1 = H_1(\theta, I)$ and we get

$$\dot{\theta} = \Omega(I) + \epsilon \frac{\partial \hat{H}}{\partial I} + \left(\frac{1}{2}\Omega(I)\frac{\partial H_1}{\partial \theta} + \sqrt{T}\xi(t) \right) \frac{\partial H_1}{\partial I}$$
$$\dot{I} = -\epsilon \frac{\partial \hat{H}}{\partial \theta} - \left(\frac{1}{2}\Omega(I)\frac{\partial H_1}{\partial \theta} + \sqrt{T}\xi(t) \right) \frac{\partial H_1}{\partial \theta}$$

where $H_1 = H_1(\theta, I)$ and $\Omega(I) = \partial \hat{H}_0/\partial I$. The action evolution depends on the perturbation ϵ and the random fluctuations $\xi(t)$. The KAM theory shows that in most case the effect of the perturbation Hamiltonian is a local change of the action since the deterministic transport is very rare. If $\epsilon \ll \sqrt{T}$ we neglect the perturbation and without loss of generality we can set $\Omega(I) > 0$ (in a local region of the phase space) and the angle variables turn out to be fast variables and an averaging procedure can be applied (the averaging problem will be discussed later) so that the actions dynamics read

$$\dot{I}_j = -\frac{1}{2}\Omega_k(I) \left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k} \right\rangle + \sqrt{T \left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k} \right\rangle} \xi_k(t)$$

where $\langle\ \rangle$ denotes the angle average and we have introduced an effective noise for the average dynamics with variance

$$C_{jk} = \left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k} \right\rangle$$

According to eq. (**??**) the FP equation in the action variables read

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\frac{\partial}{\partial I_j} \left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k} \right\rangle \left[\Omega_k(I)\rho + T\frac{\partial \rho}{\partial I_k} \right]$$

We recover the equilibrium solution

$$-\frac{1}{T}\frac{\partial H_0}{\partial I_k} = \frac{\partial}{\partial I_k}\ln \rho(I) \quad \Rightarrow \quad \rho(I) \propto \exp\left(-\frac{H_0(I)}{T} \right)$$

which does not depends on the choice of H_1. Therefore $H_1(\theta, I)$ defines the relaxation time scale which is dependent on the system state. We consider the simple case

$$\left\langle \frac{\partial H_1}{\partial \theta_j}\frac{\partial H_1}{\partial \theta_k} \right\rangle = \delta_{jk} D_k I_k$$

and $\Omega_j(I) = \omega_j$ (linear case) and the equation. Then the action dynamics factorizes in the different components

$$\dot{I}_j = -\frac{1}{2}\omega_j I_j D_j + \sqrt{T I_j}\xi_j(t)$$

It is convenient to change variable $I_j = r_j^2/2$ so that the stochastic equation reduces

$$\dot{r} = -\frac{D}{4}\omega r + \sqrt{\frac{TD}{2}}\xi(t)$$

(we drop the index to simplify the notation) and we get the FP equation for the stochastic harmonic oscillator for the distribution $\rho(r,t)$

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial r}\left[\omega r \rho + T\frac{\partial \rho}{\partial r}\right] \tag{3.16}$$

where a time scaling $t \to D_j t/4$ is applied. The stationary distribution reads

$$\rho_{eq}(r,t) r\, dr \propto \exp\left(-\frac{\omega r^2}{2T}\right) r\, dr = \exp\left(-\frac{\omega I}{T}\right) dI$$

We look for the eigenvectors in the form $\Phi_\lambda = \rho_{eq}\phi_\lambda$ and we get the equation

$$\lambda \exp\left(-\frac{\omega r^2}{2T}\right)\phi_\lambda = T\frac{\partial}{\partial r}\left[\exp\left(-\frac{\omega r^2}{2T}\right)\frac{\partial \phi_\lambda}{\partial r}\right] \tag{3.17}$$

If we scale $r = z\sqrt{T/\omega}$ we get the equation for the Hermite polynomials

$$\frac{\lambda}{\omega}\exp\left(-\frac{z^2}{2}\right)\phi_\lambda = \frac{\partial}{\partial z}\left[\exp\left(-\frac{z^2}{2}\right)\frac{\partial \phi_\lambda}{\partial z}\right]$$

Then the eigenvalue are $\lambda = -n\omega$ where n is an integer. It follows that the relaxation time scale is independent on the temperature T and decreases with the frequency ω at the fixed point.

3.2 Statistical Physics point of view

The Hamiltonian systems in a thermal bath provide a mathematical tool to study non-equilibrium statistical physics if one assumes the possibility of modelling a thermal bath using the Wiener process. We consider the case of an Hamiltonian system $H_0(x,\lambda)$ depending on the parameter λ which describes the evolution of a system in a thermal bath associate to $H_1(x)$. The unperturbed Hamiltonian defines the internal energy

$$E(t;\lambda,T) = \int H_0(x,\lambda)\rho(x,t;\lambda)\, dx \tag{3.18}$$

where $\rho(x,t;\lambda)$ is the solution of the FP equation (see eq. **??**)

$$\frac{\partial \rho}{\partial t} = -\left(D_{H_0} - D_{H_1}\{H_0,H_1\}\right)\rho + T D_{H_1}^2\rho = \mathcal{L}_{FP}\rho$$

where the solution is dependent on the parameter λ. The equilibrium solution is the MB solution

$$\rho^s(x) = A^{-1}(\lambda, T)\exp\left(-\frac{H_0(x,\lambda)}{T}\right) \qquad A(\lambda, T) = \int \exp\left(-\frac{H_0(x,\lambda)}{T}\right) dx$$

If $\lambda(t)$ is time dependent we have an approximate solution $\rho(x, \lambda(t))$ that tends to the adiabatic solution when $\dot\lambda \ll 1$. If $x(t)$ is any stochastic trajectories, the change of the system energy is defined by $H_0(x(t), \lambda)$, so that if $\rho(x_0)$ is the distribution of the initial conditions we have

$$E(t; \lambda, T) = \int \langle H_0(x(t; x_0), \lambda)\rangle \, \rho(x_0)\, dx$$

The theory of stochastic systems implies

$$\frac{dE}{dt}(\lambda, T; t) = \int \mathcal{L}^*_{FP} H_0(x_t, \lambda)\rho(x_t, t)\, dx_t$$

where $\mathcal{L}^*_{FP}$ is the adjoint of the FP operator and x_t is the system state at time t. Then we have

$$\begin{aligned}\mathcal{L}^*_{FP} H_0(x, \lambda) &= \left(D_{H_0} - \{H_0, H_1\}D_{H_1}\right) H_0(x, \lambda) + TD^2_{H_1} H_0(x, \lambda)\\ &= -\{H_0, H_1\}^2 + T\{\{H_0, H_1\}, H_1\}\end{aligned}$$

If λ is fixed we define the exchanged heat with the environment during the relaxation process

$$\frac{đQ}{dt} = -\int \{H_0, H_1\}^2 \rho(x, t)\, dx + T\int \{\{H_0, H_1\}, H_1\}\rho(x, t)\, dx$$

If $\rho(x, t) = \rho^s(x) \propto \exp(-H_0/T)$ the MB distribution, we get $đQ = 0$ (stationary condition). We define the relaxation current density

$$J = -\{H_0, H_1\}\rho - TD_{H_1}\rho$$

and we have the relation

$$\begin{aligned}\frac{đQ}{dt} &= -\int \{H_0, H_1\}\left[\{H_0, H_1\}\rho + TD_{H_1}\rho\right] dx\\ &= \int \{H_0, H_1\}J\, dx = -\int \frac{J^2}{\rho}\, dx - T\int \frac{J}{\rho} D_{H_1}\rho\end{aligned}$$

We understand that exchanged heat per unit time is partially dissipated in the system and it is partially stored in the distribution. Out of equilibrium the system exchanges heat that creates the currents that change the internal distribution performing work and dissipate energy in the system. The energy dissipation vanishes if $J = 0$, the equilibrium condition that satisfies a detailed balance condition and it corresponds to a maximum entropy condition. In the relaxation process the system changes the distribution to a maximum entropy distribution but part of the energy is dissipated during the process. This effect is the basis of the second principle of thermodynamics since the adiabatic limit implies $J \to 0$. We observe that the FP operator cannot be written as a divergence of a density current due to the presence of the operator D_{H_0}.

If $\lambda(t)$ is time-dependent we define the work performed on the system

$$\frac{đW}{dt} = \int \dot{\lambda} \frac{\partial H_0}{\partial \lambda}(x,\lambda)\rho(x,t)\, dx$$

so that we get the first Principle of Thermodynamics. If we perform an adiabatic transformation the system distribution $\rho(x,t) \propto \exp(-H_0(x,\lambda)/T)$ so that $đQ = 0$ during the transformation. Using the partition function it follows

$$\frac{\partial}{\partial \lambda} \ln A(\lambda, T) = -\frac{1}{TA(\lambda,T)} \int \frac{\partial H_0}{\partial \lambda} \exp\left(-\frac{H_0(x,\lambda)}{T}\right) dx$$

so that in the case of adiabatic transformations we have

$$\frac{đW}{dt} = -T\dot{\lambda}\frac{\partial}{\partial \lambda} \ln A(\lambda, T) \quad \Rightarrow \quad W = -T[\ln A(\lambda_f, T) - \ln A(\lambda_i, T)]$$

The work performed by the system in a adiabatic transformation

$$W = F(\lambda_f, T) - F(\lambda_i, T) \qquad F(\lambda, T) = -T \ln A(\lambda T)$$

where we define the Helmholtz free energy $F(\lambda, T)$. In the case $\dot{\lambda}$ is finite we get the second principle of thermodynamics since

$$E_f - E_i = \int_{t_i}^{t_f} \left[-\int \frac{J^2}{\rho}\, dx - \int T\frac{J}{\rho} D_{H_1}\rho + \int \dot{\lambda}\frac{\partial H_0}{\partial \lambda}\rho\, dx \right] dt$$

cucu The second principle of thermodynamics implies $\Delta F \leq W$

We consider the classical system

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

where λ simulates the change of the energy due to the coupling with the environment. If one introduces a thermal bath the equations of motion have the form

$$\begin{aligned} dq &= p\, dt \\ dp &= \frac{\partial V}{\partial q}(q,\lambda)\, dt - \gamma p\, dt + \sqrt{2T\gamma}\, dw_t \end{aligned}$$

Any solution $q_w(t), p_w(t)$ represents a possible isothermal transformation of the system that relaxes toward a stationary stochastic process; in the relaxation process the system change its energy exchanging heat with the environment. We compute the exchange heat by

$$dH = -\gamma p^2\, dt + \sqrt{2T\gamma} p\, dw_t + T\gamma\, dt$$

and by averaging on the fluctuations we get the internal energy change

$$dU = (T - \langle p^2 \rangle)\gamma\, dt = đQ$$

where

$$\langle p^2 \rangle = \int p^2 \rho(q,p,t)\, dqdp$$

H_1 represents the environment effect using the statistical physics point of view: i.e. $H_1(x)$ represents the local fluctuations due to the environmental interactions that have to justified from physical laws. This also means that there are many physical system with the same equilibrium but with different coupling with the environment and different fluctuations. The spectral properties of the FP equation depends on the choice of H_1. The DB condition is related to the self-adjoint nature of the Fokker-Planck operator: the unperturbed Liouville operator $-D_{H_0}$ is not self-adjoint but we look for solutions in the kernel, so that we consider the operator

$$S_{FP} = \frac{1}{2}\exp\left(\frac{H_0}{2T}\right) D_{H_1}\left(\{H_0, H_1\} + TD_{H_1}\right)\exp\left(-\frac{H_0}{2T}\right)$$

and we change variable $\rho(x,t) \to \exp(-H_0/2T)u(x,t)$ to get a self-adjoint operator: a direct computation gives

$$\left(\{H_0, H_1\} + TD_{H_1}\right)\exp\left(-\frac{H_0}{2T}\right) = \frac{1}{2}\{H_0, H_1\}\exp\left(-\frac{H_0}{2T}\right) + T\exp\left(-\frac{H_0}{2T}D_{H_1}\right)$$

and we get

$$S_{FP} = \frac{1}{2}D^2_{H_1}H_0 - \frac{1}{4T}\{H_0, H_1\}^2 + TD^2_{H_1}$$

that is self-adjoint. The new FP equation reads

$$\frac{\partial u}{\partial t} = \left[-D_{H_0} + \frac{1}{2}D^2_{H_1}H_0 - \frac{1}{4T}\{H_0, H_1\}^2 + TD^2_{H_1}\right]u$$

A possible approach to the study of the spectral properties is to consider the interacting frame

$$u(x,t) = \exp(tD_{H_0})v(x,t) \qquad \frac{\partial u}{\partial t} = \exp(tD_{H_0})\left(\frac{\partial v}{\partial t} + D_{H_0}\right)v$$

and we look for a solution

$$\frac{\partial v}{\partial t} = \exp(-tD_{H_0})\left[\frac{1}{2}D^2_{H_1}H_0 - \frac{1}{4T}\{H_0, H_1\}^2 + TD^2_{H_1}\right]\exp(tD_{H_0})v$$

We analyze a statistical mechanics approach to a stochastic Hamiltonian system under the assumption that the distribution function that solves that F.P. equation (17). Let us introduce the Gibbs Entropy

$$S(\rho) = -\int \ln(\rho)\rho\, dx \tag{3.19}$$

for a given distribution $\rho(x,t)$. By using the Jensen's inequality one can prove that the Entropy is positive:

$$\exp\left(\int \ln(\rho)\rho dx\right) \geq \int \exp\left(\ln\rho\right)\rho dx = \int \rho^2 dx$$

Let us compute the time derivative of the Gibbs Entropy in the case **??**

$$\frac{dS}{dt} = -\int \ln\rho\left(-D_{H_0} + \frac{1}{2}D^2_{H_1}\right)\rho dx = -\int D_{H_0}\rho dx + \frac{1}{2}\int \frac{(D_{H_1}\rho)^2}{\rho}dx$$

and using the usual boundary condition on ρ, we get the Entropy Growth Principle

$$\frac{dS}{dt} = \frac{1}{2}\int \frac{(D_{H_1}\rho)^2}{\rho} dx > 0$$

Remarks:the zero Entropy production would imply $D_{H_1}\rho = 0$ (i.e. ρ is a first integral of motion for H_1), but this condition does not satisfy the eq. (**??**) so that one has a continuous Entropy production in the stationary state that is not an equilibrium state. On the contrary if one introduces the dissipative term (**??**) it is straightforward to check that the Maxwell-Boltzmann distribution is stationary distribution that maximize the Gibbs Entropy.

3.3 Coarse grained description

In a generic situation we have an Hamiltonian system $H(x)$ where x are the microstates that are assume equiprobable since the preserved measured is the usual Lebesgue measure. We have proved the following relations

$$F = -\beta^{-1}\sum_x \exp(-\beta H(x)) \qquad E = \beta\frac{\partial}{\partial\beta}\beta F \qquad S = \beta^2\frac{\partial F}{\partial\beta} = \beta(E-F)$$

we introduce the mesostates I and we consider the equilibrium distribution for each mesostate

$$\rho_{eq}(I) = \sum_{x\in I}\exp\left(-\beta(H(x)-F)\right) = \exp\left(-\beta(F_I - F)\right)$$

where we introduce the free energy of the state I. The definition is consistent with the conditional probability

$$\rho(x|I) =$$

*****

3.3.1 Master equation and Statistical Physics

The evolution of the master equation is written in the form

$$\dot\rho(x,t) = \sum_y \left[\pi_{xy}(\lambda)\rho(y,t) - \pi_{yx}(\lambda)\rho(x,t)\right] = -\sum_y \mathcal{L}_{xy}\rho(y,t) \tag{3.20}$$

where π_{xy} is the transition rate between the state $y \to x$ and $\mathcal{L}_{xy}$ the associated Laplacian matrix. In complex systems physics the structure of the Laplacian matrix defines the interaction among the mesostates of the system.The formal solution is written

$$\rho(x,t) = \sum_y \exp\left(-t\mathcal{L}_{xy}\right)\rho_0(y)$$

so that the transition probability reads

$$\pi^{\Delta t}_{xy} = \exp\left(-\Delta t\mathcal{L}_{xy}\right) \qquad y \neq x$$

This equation is related to a Markov stochastic process $x(t)$ whose realizations satisfy the condition

$$\mathcal{P}[x(t)] = \lim_{\Delta t \to 0} \prod_{k=1}^{n} \exp\left(-\Delta t \mathcal{L}_{x_k x_{k-1}}\right) \qquad x_k = x(k\Delta t) \quad t = n\Delta t \tag{3.21}$$

where $x(0) = x_0$ is the initial condition which is given. For finite n we have the normalizing condition

$$\sum_{\{x_k\}_{k=0}^n} \mathcal{P}[\{x\}_{k=0}^n] = 1$$

where x_0 is fixed. The continuous limit defines the path integral approach to Markov system. Let $\rho_s(x)$ the stationary solution, the DB condition can be formulated by

$$\exp\left(-\Delta t \mathcal{L}_{xy}\right) \rho_s(y) = \exp\left(-\Delta t \mathcal{L}_{yx}\right) \rho_s(x) \qquad \forall \Delta t$$

It follows

$$\mathcal{L}_{xy} \rho_s(y) = \mathcal{L}_{yx} \rho_s(x) \tag{3.22}$$

The DB condition (**??**) has two consequences: the Laplacian matrix defines the *internal energy* of the system by

$$\ln(\mathcal{L}_{xy}) - ln(\mathcal{L}_{yx}) = \beta(E(y) - E(x))$$

(β is a scaling factor to define the measure unit). Then $\rho_s(x)$ has the form of a MB distribution. We observe that the definition

$$\mathcal{L}_{xy} = \exp(\beta E(y)) \qquad x \neq y$$

means that the escape rate from the node y is exponentially small in the energy associated to the state $E(y)$ according to the Kramer transition rate theory, and it clearly satisfies the DB condition. This definition is consistent with a random walk dynamics on a landscape potential with many local minimum and each 'link' has the same weight. A different choice is

$$\mathcal{L}_{xy} = \exp\left(\frac{\beta}{2}(E(y) - E(x)\right) \qquad x \neq y$$

where the link between two state ahs a weight proportional to the energy difference (i.e. one chooses the nodes with lower energy in the dynamics). The second consequence is the stochastic reversibility of the dynamics: we define the reverse evolution by $y_k = x_n - k$ and we consider the probability

$$\mathcal{P}y(t)] = \lim_{\Delta t \to 0} \prod_{k=1}^{n} \exp\left(-\Delta t \mathcal{L}_{y_k y_{k-1}}\right) \rho_s(y_0)$$

where one assumes that the 'initial' condition is distributed according to the stationary distribution. Remark: we impose the same transition probabilities for the revrese process, so that we interpret the reverse process as a possible physical realization. Using the DB condition

$$\exp\left(-\Delta t \mathcal{L}_{y_k y_{k-1}}\right) \rho_s(y_{k-1}) = \exp\left(-\Delta t \mathcal{L}_{y_{k-1} y_k}\right) \rho_s(y_k) = \exp\left(-\Delta t \mathcal{L}_{x_{n-k+1} x_{n-k}}\right) \rho_s(x_{n-k})$$

we recover the probability of the forward sequence. We say that the system is stochastically time reversible since the statistical properties cannot distinguish the time arrow: any measure on the observables cannot distinguish if the evolution is time reversed at the stationary state.
The DB balance condition is strictly related to the maximum entropy principle but it implies consequence of the relaxation process. For a given initial distribution $\rho(x,0)$ the relaxation process is an irreversible isothermal transformation according to the equation

$$\frac{\partial \rho}{\partial t} = \sum_y \left[\pi_{xy}\rho(y,t) - \pi_{yx}\rho(x,t)\right]$$

Part of the exchanged heat is dissipated in the system or accumulated in the internal energy. We consider the entropy change during the relaxation process

$$\begin{aligned}\frac{dS}{dt} &= -\sum_{xy} \left[\pi_{xy}\rho(y,t) - \pi_{yx}\rho(x,t)\right] \ln \rho(x,t) \\ &= -\frac{1}{2}\sum_{xy} \left[\pi_{xy}\rho(y,t) - \pi_{yx}\rho(x,t)\right]\left[\ln \pi_{xy}\rho(y,t) - \ln \pi_{yx}\rho(x,t)\right] \\ &- \frac{1}{2}\sum_{xy} \left[\pi_{xy}\rho(y,t) - \pi_{yx}\rho(x,t)\right]\left[\ln \pi_{xy} - \ln \pi_{yx}\right]\end{aligned}$$

where we recognize the total entropy production

$$\frac{dS_{tot}}{dt} = \frac{1}{2}\sum_{xy} \left[\pi_{xy}\rho(y,t) - \pi_{yx}\rho(x,t)\right]\left[\ln \pi_{xy}\rho(y,t) - \ln \pi yx\rho(x,t)\right] \geq 0$$

which vanishes when the current density $\pi_{xy}\rho(y,t) - \pi_{yx}\rho(x,t)$ vanish. The second term is the work performed by the currents with respect the vector field $\ln \pi_{xy} - \ln \pi_{yx}$ associated to the internal interactions. The performed work can be written in the form

$$\dot{W} = -\sum_{xy} \pi_{xy}\rho(y,t)\ln \pi_{xy} - \sum_{yx} \pi_{yx}\rho(x,t)\ln \pi_{yx}$$

which can be interpreted as the difference between the forward entropy production and the backward entropy production when the trajectory is reverse. This different must be positive: i.e. the system tends to minimize the entropy production choosing the relaxation path.
If one diagonalizes the master equation using the Laplacian eigenvector

$$\rho(x,t) = \sum_\lambda c_\lambda v_x^\lambda \qquad \sum_x \|v_x^\lambda\|^2 = 1 \quad \sum_x v_x^\lambda = 0$$

where v^0 (the null eiegenvector) is the equilibrium solution and $c_0 = 1$. We have the dynamics for each eigenvector

$$\dot{c}^\lambda = -\lambda c_\lambda \quad \Rightarrow \quad c^\lambda = c^\lambda(0)e^{-\lambda t}$$

****** We observe that the DB condition implies

$$\frac{1}{\sqrt{\rho_s(x)}}\mathcal{L}_{xy}\sqrt{\rho_s(y)} = \frac{1}{\sqrt{\rho_s(y)}}\mathcal{L}_{yx}\sqrt{\rho_s(x)}$$

i.e. the matrix similar to $\mathcal{L}_{xy}$ i symmetric so that the eigenvalue od the Laplacian are real and positive whereas the corresponding eigenvectors are orthogonal with respect the measure ρ_s^{-1}. The spectral properties are related to the correlation function of the observable. Let $I_i(x)$ any set of observables we define the evolution of the average values

$$\dot{I}_i(x,t) = -\sum_y \mathcal{L}^{\dagger}_{xy} I_i(y,t)$$

using the adjoint operator which has a constant stationary eigenvalue. The correlation function between two observables is defined

$$\begin{aligned} K_{ij}(\tau) = \langle I_i(x_s(tau|y)) I_j(y)\rangle = \sum_y I_i(y,\tau) I_j(y)\rho_s(y) = \sum_{xy} I_i(x)\rho(x,\tau|y) I_j(y)\rho_s(y) \\ = \sum_{xy} I_i(x)\exp\left(-\mathcal{L}_{xy}\tau\right) I_j(y)\rho_s(y) \end{aligned} \tag{3.23}$$

where the suffix s means that the trajectories are computed at the stationary condition. By definition we have

$$\begin{aligned} K_{ij}(-\tau) = \langle I_i(x_s(t-\tau)) I_j(x_s(t))\rangle = K_{ji}(\tau) = \sum_{xy} I_j(x)\exp\left(-\mathcal{L}_{xy}\tau\right) I_i(y)\rho_s(y) \\ = \sum_{xy} I_i(y)\rho_s(y)\exp\left(-\mathcal{L}^{\dagger}_{yx}\tau\right) I_j(x) = K_{ji}(\tau) \end{aligned}$$

The stochastic reversibility assumes that $K_{ij}(\tau) = K_{ij}(-\tau) = K_{ji}(\tau)$ which implies

$$\exp\left(-\mathcal{L}_{xy}\tau\right)\rho_s(y) = \rho_s(y)\exp\left(-\mathcal{L}^{\dagger}_{yx}\tau\right) = \left[\exp\left(-\mathcal{L}_{xy}\tau\right)\rho_s(y)\right]^{\dagger}$$

and we recover the DB condition. The symmetry of the correlation $K_{ij}(\tau)$ can be related to the Onsager symmetry relations.
The DB balance condition implies constraint to the relaxation properties towards equilibrium since they are relate to the spectrum of the Laplacian operator. We assume that the Laplacian operator depend on a parameter $\lambda = \epsilon t$ $\mathcal{L}_{xy}(\lambda)$ such that the DB condition is satisfied for all the λ values $\in [\lambda_0, \lambda_1]$.The evolution of the system is described by two phases: if we introduce the energy $E(x,\lambda)$ through the DB condition for each value λ for each evolution of the system $x(t)$ we introduce the First Principle of Thermodynamics

$$\frac{dE}{dt}(x(t),\lambda) = \frac{\partial E}{\partial x}(x(t),\lambda)\dot{x} + \frac{\partial E}{\partial \lambda}(x(t),\lambda)\dot{\lambda}$$

The second term represents the work per unit time performs on the system when its state is $x(t)$ whereas the first term is the exchange heat with the environment and the transformation is isothermal.The first Principle follows by taking the average on all the possible states $x(t)$.
For each possible evolution $x(t)$ of the system we define

$$\hat{\pi}_{x_{k-1}x_k}(k+1) = \frac{\exp(-\beta E_{n-k}(x_{k-1}))}{\exp(-\beta E_{n-k}(x_k))}\pi_{x_k x_{k-1}}(n-k)$$

where we discretize the time evolution $t = k\Delta t$ with $k = 0, .., n$ (β is a parameter). We observe that the difference

$$\Delta W_{n-k} = E_{n-k}(x_{k-1}) - E_{n-k}(x_k)$$

is the work performed on the system when the state changes $x_k \to x_{k-1}$. The previous definition has the property

$$\sum_y \hat{\pi}_{yx}(k+1) = \sum_y \frac{\exp(-\beta E_{n-k}(y))}{\exp(-\beta E_{n-k}(x))} \pi_{xy}(n-k) = \sum_y \pi_{yx}(n-k) = 1$$

using the DB condition: i.e. it is a stochastic matrix. By definition we have

$$\langle \exp(\beta W[x(t)]) \rangle = \sum_{\{x_k\}_0^n} \prod_{k=1}^{n} \pi_{x_k x_{k-1}}(k) \frac{\exp(-\beta E_k(x_{k-1})}{\exp(-\beta E_{k-1}(x_{k-1})} \rho(x_0; 0)$$

that can be written in the form

$$\begin{aligned}\langle \exp(\beta W[x(t)]) \rangle &= \sum_{\{x_k\}_0^n} \prod_{k=0}^{n-1} \pi_{x_{n-k} x_{n-k-1}}(n-k) \frac{\exp(-\beta E_{n-k}(x_{n-k-1})}{\exp(-\beta E_{n-k-1}(x_{n-k-1})} \rho(x_0; 0) \\ &= \sum_{\{x_k\}_0^n} \prod_{k=0}^{n-1} \hat{\pi}_{x_{n-k-1} x_{n-k}}(k+1) \frac{\exp(-\beta E_{n-k}(x_{n-k}))}{\exp(-\beta E_{n-k-1}(x_{n-k-1})} \rho(x_0; 0)\end{aligned}$$

using the relation

$$\pi_{x_{n-k} x_{n-k-1}}(n-k) = \hat{\pi}_{x_{n-k-1} x_{n-k}}(k+1) \frac{\exp(-\beta E_{n-k}(x_{n-k}))}{\exp(-\beta E_{n-k}(x_{n-k-1}))}$$

Then we get

$$\begin{aligned}\langle \exp(W[x(t)]) \rangle &= \sum_{\{x_k\}_0^n} \frac{\exp(-\beta E_n(x_n))}{\exp(-\beta E_0(x_0)} \prod_{k=0}^{n-1} \rho(x_0; 0) \hat{\pi}_{x_{n-k-1} x_{n-k}}(k+1) \\ &= \frac{Z_n}{Z_0} \sum_{\{x_k\}_0^{n-1}} \prod_{k=0}^{n-1} \hat{\pi}_{x_{n-k-1} x_{n-k}}(k+1) \rho(x_n; 0) = \frac{Z_n}{Z_0}\end{aligned}$$

and we recover the Jazinski equality

$$\langle \exp(\beta W[x(t)]) \rangle = \exp(\beta \Delta F)$$

for the Markov systems.

****** If we define the function

$$\chi(x, \lambda) = \rho(x, \lambda) \exp(-(E(x, \lambda))$$

3.4 On numerical integration of stochastic equations

We consider the problem of introducing a numerical scheme for the eq. (??)

$$\dot{x} = D_{H_0} x + \left(-\frac{\gamma}{2} \{H_0, H_1\} + \sqrt{T\gamma} \xi(t) \right) D_{H_1} x$$

We divide the Hamiltonian distinguishing the deterministic and the stochastic part. The idea of a stochastic integrator is to substitute the process $x(t)$ with a new process $\hat{x}(t;\Delta t)$ such that

$$\mathcal{M}_2(x(t)-\hat{x}(t;\Delta t)) = \left\langle \|x(t)-\hat{x}(t;\Delta t)\|^2 \right\rangle = O(\Delta t^k) \tag{3.24}$$

for any finite $t = n\Delta t$. $k/2 > 0$ is the integration order of the numerical scheme; the average is computed over all the noise realizations. We observe that if we define $\Delta x = x - \langle x \rangle$

$$\|x-\hat{x}\|^2 = \|\Delta x - \Delta\hat{x} + (\langle x\rangle - \langle\hat{x}\rangle)\|^2 \leq \|\Delta x - \Delta\hat{x}\|^2 + (\langle x\rangle - \langle\hat{x}\rangle)^2$$

and if we prove that the average value and the mean square value of the processes $x(t)$ and $\hat{x}(t;\Delta t)$ differ by a quantity $O(\Delta t^{k/2})$ we get the estimate (**??**). We remark that there is not any requirement on the convergence of the single trajectories of the processes if we consider the same noise realization (strong convergence) since we are interested in the statistical properties of the process. We build a leap-frog numerical scheme: we approximates the deterministic dynamics by

$$\begin{aligned} x(t+\Delta t) &= \exp\left(\Delta t D_{H_0})x(t)\right)\exp\left(-\Delta t\frac{\gamma}{2}\{H_0,H_1\}D_{H_1}\right)x(t) \\ &= M^{\Delta t}x(t) \simeq \left(I + \Delta t D_{H_0} - \Delta t\frac{\gamma}{2}\{H_0,H_1\}D_{H_1} + O(\Delta t^2)\right)x(t) \end{aligned}$$

and we use a stochastic symplectic kick for the stochastic dynamics

$$x(t+\Delta t) = \exp\left(\sqrt{T\gamma}\Delta w_t D_{H_1}\right)x(t)$$

with Δw_t the increment of a Wiener process.
Remark: The scheme assumes the possibility of explicitly integrating the symplectic flow associated to $H_0(x)$ but it is enough an approximation with an error $O(\Delta t^2)$. The same is applied to the phase flow of the dissipation term $\gamma/2\{H_0,H_1\}D_{H_1}$: in this case we have not to satisfy a symplectic condition and the Euler approximation can be used

$$\exp\left(-\Delta t\frac{\gamma}{2}\{H_0,H_1\}D_{H_1}\right)x = x - \Delta t\frac{\gamma}{2}\{H_0,H_1\}\{x,H_1\} + O(\Delta t^2)$$

with $\Delta t\,\gamma \ll 1$. Then we define the stochastic process

$$x(t+\Delta t) = M^{\Delta t/2}\circ\exp\left(\sqrt{T\gamma}\Delta w D_{H_1}\right)\circ M^{\Delta t/2}x(t)$$

whose global error is expected $O(\Delta t^{3/2})$ in a statistical sense.
If $H_1 = q$ and $H_0 = p^2/2 + V(q)$ (the usual thermal bath) we have

$$\exp\left(-\Delta t\frac{\gamma}{2}\{H_0,H_1\}D_{H_1}\right)x = \exp\left(-\Delta t\gamma/2\; pD_q\right)x = \begin{cases} q \\ \exp\left(-\Delta t\frac{\gamma}{2}\right)p \end{cases}$$

$$\exp\left(\sqrt{T\gamma}\Delta w D_{H_1}\right)x = \begin{cases} q \\ p + \sqrt{T\gamma}\Delta w \end{cases}$$

As previously discussed the numerical scheme is order $O(\Delta t)$ when we integrate the stochastic equation for a finite time$t$.

3.5 The over-damped limit

If the unperturbed Hamiltonian can be written in the form $H_0(p,q) = p^2/2 + V(q)$ and $H_1 = -q$. We have the stochastic equation

$$\begin{aligned} dp &= -\frac{\partial V}{\partial q} dt + \left(-\gamma p\, dt + \sqrt{2T\gamma}\, dw_t\right) \\ dq &= p\, dt \end{aligned}$$

where the parameter γ allows to modulates the dissipation. We compute a limit $\gamma \gg 1$

$$\begin{aligned} \frac{dp}{\gamma} &= -\frac{\partial V}{\partial q}\frac{dt}{\gamma} + \left(-pdt + \sqrt{\frac{2T}{\gamma}}\, dw_t\right) \\ dq &= p\, dt \end{aligned}$$

Then we get the approximated dynamics in the slow time $\tau = t/\gamma$ by the reduced dynamics

$$dq = -\frac{\partial V}{\partial q} d\tau + \sqrt{2T} dw_\tau \tag{3.25}$$

The equilibrium distribution is the marginal distribution of the whole system on the q coordinates. Remark: to apply the over damped approximation the form of the Hamiltonian $H_0(p,q)$ and of the thermal bath $H_1(p,q) = -q$ cannot be modified. The over damped limit is coordinates dependent but the spectral properties of the FP operator associated to the stochastic equation(**??**) are related to the spectral properties of the whole FP equation.
We consider the relation with the averaging theorem: let the initial Hamiltonian being integrable such that using the action-angle variables (θ, I) we have $H_0 = H_0(I)$ and $H_1 = H_1(\theta, I)$. The thermal bath equations read

$$\begin{aligned} \dot\theta &= \Omega(I) + \gamma\Omega(I)\frac{\partial H_1}{\partial \theta}\frac{\partial H_1}{\partial I} + \sqrt{2T\gamma}\xi(t)\frac{\partial H_1}{\partial I} \\ \dot I &= -\gamma\Omega(I)\left(\frac{\partial H_1}{\partial \theta}\right)^2 - \sqrt{2T\gamma}\xi(t)\frac{\partial H_1}{\partial \theta} \end{aligned}$$

We consider the limit $\gamma \to 0$ (i.e. the small noise limit) where the angles θ are fast variables. The averaging principle considers a slow time $\tau = t\gamma$ and the white noise limit $\sqrt{T/\gamma}\xi(t)d\tau \to dw_\tau$ to derive the stochastic equations

$$\begin{aligned} d\theta &= \frac{\Omega(I)}{\gamma}\, d\tau + \Omega(I)\frac{\partial H_1}{\partial \theta}\frac{\partial H_1}{\partial I} d\tau + \sqrt{2T}\frac{\partial H_1}{\partial I}\, dw_\tau \\ dI &= -\Omega(I)\left(\frac{\partial H_1}{\partial \theta}\right)^2 d\tau - \sqrt{2T}\frac{\partial H_1}{\partial \theta}\, dw_\tau \end{aligned}$$

so that in the limit $\gamma \to 0$ the phase advance $\Omega(I)/\gamma$ distributes the angles randomly on a torus surface: i.e. one can apply the random phase approximation to the action dynamics, where the angle can be considered independent random variables uniformly distributed. This claim is not correct when the frequencies satisfies a resonance relation

$$\sum_i n_i \Omega(I) = 0 \tag{3.26}$$

In such a case even in the limit $\gamma \to 0$ the phases are not uniformly distributed in a N-dimensional torus but only on a subset whose dimension depends on the resonance order. However the angle dynamics is fast and in the case of an ensemble of particles with initial condition uniformly distributed on the torus, this condition is preserved by the evolution. The FP equation have the MB stationary solution (cfr.eq. (**??**)) which is uniform in the angle. Then we expect that there is a fast relaxation to the uniform distribution in the angle variables whereas the relaxation process in the action variables is a slow process. Then one can apply an averaging principle where the action stochastic dynamics

$$d\bar{I} = -\Omega(\bar{I})D(\bar{I})\,d\tau - \sqrt{2TD(\bar{I})}\,dw_\tau \tag{3.27}$$

where we define the quasi-linear diffusion coefficient

$$D(I) = \left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle$$

The corresponding FP equation reads

$$\frac{\partial \rho}{\partial \tau} = \frac{\partial}{\partial \bar{I}}\Omega(\bar{I})D(\bar{I})\rho(\bar{I},\tau) + T\frac{\partial}{\partial \bar{I}}D(\bar{I})\frac{\partial}{\partial \bar{I}}\rho(\bar{I},\tau)$$

and we recover the stationary distribution

$$\Omega(\bar{I})D(\bar{I})\rho_s(\bar{I}) + TD(\bar{I})\frac{\partial}{\partial \bar{I}}\rho_s(\bar{I}) = 0 \quad \Rightarrow \quad \rho_s(\bar{I}) \propto \exp\left(-\frac{H_0(\bar{I})}{T}\right)$$

The diffusion coefficient $D(I)$ is related to the relaxation process to the stationary distribution. The averaging principle cannot be applied in presence of separatrix surfaces in the phase space where the action-angle variables are singular and the particle distribution tends to concentrate near the hyperbolic fixed point in 2D Hamiltonian systems.

3.6 Susceptibility and Diffusion

The stochastic dynamics (**??**) can be interpreted as the perturbation effect of fast chaotic degrees of freedom on the dynamics of the Hamiltonian $H_0(x)$. The chaotic character is related to the correlation time γ^{-1} of the random perturbation $\xi(t)$ and γ is the maximum positive Ljapunov exponent. Let $x_0(t)$ a trajectory of the unperturbed system $H_0(x)$ we consider the problem to describe the effect of the noisy perturbation $\xi(t)H_1(x)$ in the limit of small white noise

$$\langle \xi(t+\tau)\xi(t) = \epsilon^2 \frac{\gamma}{2}\exp(-\gamma\tau)$$

The limit $\gamma \to \infty$ can be performed by keeping $\epsilon^2\gamma$ finite or one can introduce a slow time $s = t/\gamma$ and study the evolution in the slow time. We linearize the dynamics around the reference orbit

$$\dot{x} = \frac{\partial H_0}{\partial x}J + \xi(t)\frac{\partial H_1}{\partial x}J \quad \to \quad \to \delta\dot{x} = \frac{\partial^2 H_0}{\partial x^2}^T(x_0(t))\delta x J + \xi(t)\frac{\partial H_1}{\partial x}^T(x_0(t))J$$

where $\delta x = x(t) - x_0(t)$. This is a linear stochastic Hamiltonian

$$H(\delta x, t; \xi(t))) = \frac{1}{2}\delta x \frac{\partial^2 H_0}{\partial x^2}(x_0(t))\delta x + \xi(t)\frac{\partial H_1}{\partial x}^T(x_0(t))\delta x$$

that is valid only if $\|\delta x\| \ll 1$. We recall that the fundamental matrix Φ_0^t of the system with $\epsilon = 0$ gives the information on the Ljapunov exponents of the system $H_0(x)$ from the spectral properties of the matrix

$$\Lambda(t) = \frac{1}{2t}\ln([\Phi_0^t]^T\Phi_0^t)$$

and the limit $\lim_{t\to\infty}\Lambda(t)$ defines the Ljapunov exponents of the orbit. When t is small we have the local Ljapunov exponent that may have large fluctuations. The susceptibility of the system considers the solution

$$\delta x(t) = \int_0^t \xi(s)\Phi_s^t \frac{\partial H_1}{\partial x}(x_0(s))J\,ds$$

A direct calculation provides

$$\left\langle \|\delta x(t)\|^2 \right\rangle = \int_0^t\int_0^t \langle \xi(s_1)\xi(s_2)\rangle \left[\frac{\partial H_1}{\partial x}(x_0(s_1))J\right]^T \left[\Phi_{s_1}^t\right]^T \Phi_{s_2}^t \left[\frac{\partial H_1}{\partial x}(x_0(s_2))J\right] ds_1 ds_2$$

and performing the white noise limit one gets

$$\left\langle \|\delta x(t)\|^2 \right\rangle \simeq \epsilon^2 \int_0^t \left[\frac{\partial H_1}{\partial x}(x_0(s))J\right]^T \left[\Phi_s^t\right]^T \Phi_s^t \left[\frac{\partial H_1}{\partial x}(x_0(s))J\right] ds$$

The possibility of applying the white noise limit gives the possibility of choosing $t \ll 1$ and the behavior of the norm is defined by the average local Ljapunov exponent computed along the vectors $(\partial H_1/\partial x)J(x_0(t)$. The relation between the Ljapunov exponents and the average local Ljapunov exponent depends on how the orbit $x_0(t)$ explores the phase space: the problem is that the orbit can spend a long time trapped in regions with a very small local Ljapunov exponent (in principle it can be almost zero) where the time spent in a region with a great Ljapunov exponent is $\propto \exp(\lambda t)$. Assuming an averaging principle on an ergodic component Σ of the unperturbed dynamics a possible estimate could be

$$\left\langle \|\delta x(t)\|^2 \right\rangle \simeq \epsilon^2 \int_\Sigma \sigma(dx)\left[\frac{\partial H_1}{\partial x}J\right]\exp(2\Lambda(x,t)t)\frac{\partial H_1}{\partial x}J \tag{3.28}$$

where $x_0(t)$ has to explore the invariant set Σ ($\sigma(dx)$ is the dynamical measure that describes how the orbit explores the phase space at time t: when $t \to \infty$ the invariant measure of Hamiltonian systems is the Lebesgue measure, but this is not true for finite t). Then we expect that the positive Ljapunov directions dominates and we have an exponential linear increase of $\|\delta x(t)\|$ in the limit $\epsilon^2 \to 0$ which justifies the linearization procedure. The noise modulates the effect of the positive Ljapunov exponent and it can destroys local spatial correlation in the unperturbed dynamics. This is the susceptibility for a long term evolution. If we can perform the average for $t \ll 1$ we approximate

$$\Phi_s^t \simeq I + \frac{\partial^2 H_0}{\partial x^2}(x_0)J(t-s) + O(t^2)$$

and we get

$$\left\langle \|\delta x(t)\|^2 \right\rangle \simeq \epsilon^2 \left[\frac{\partial H_1}{\partial x}(x_0) J \right]^T \frac{\partial H_1}{\partial x}(x_0) J t + O(t^2)$$

so that

$$\lim_{t \to 0} \frac{\langle \|\delta x(t)\|^2 \rangle}{t} = \epsilon^2 \left[\frac{\partial H_1}{\partial x}(x_0) J \right]^T \frac{\partial H_1}{\partial x}(x_0) J = b(x_0) \tag{3.29}$$

(cfr. eq. **??**). ϵ has to be sufficiently small to justify the linear approach and the white noise limit has to hold (i.e. the evolution time has to be much slower than the noise correlation time scale to keep the effect of noise finite for small ϵ). In such a case we get information on the perturbation Hamiltonian $H_1(x_0)$ whereas the unperturbed dynamics is irrelevant, but we have the drift coefficient. This is the susceptibility of the short term evolution when the initial condition is known.

Chapter 4

Averaging principle for Stochastic Hamiltonians

The chaotic dynamics in non-integrable Hamiltonian system can be related to the existence of positive Ljapunov exponents in the evolution. The perturbation theory considers Hamiltonian system of the form

$$H(\theta, I) = H_0(I) + H_1(\theta, I) \qquad \|H_1\| \ll 1$$

but the perturbation term H_1 cannot be reduced to zero in generic cases but there are optimal estimate for the norm $\|H_1\|$. A rough estimate associates the maximum local Ljapunov exponent to the optimal estimate of H_1 (i.e. the Ljapunov exponent related to the hyperbolic points of non linear resonances that are ubiquitous in the phase space). The main assumption is that the effect of local Ljapunov exponents is to introduce fast fluctuations in the dynamics whose amplitude depend on the phase space point. This means that the fluctuations are everywhere and there is a uniform chaos in the space. The Ljapunov exponent is related to the decorrelation time of the fluctuation and this also may depend on the phase space point. However in the white noise limit we need to separate the evolution time scales of the variables and the fluctuations time scale. This will be possible in a slow time if the considered region has a positive Ljapunov exponent. In other word the diffusion approximation applies on an invariant subset of the phase space where we have a lower bound on the maximum Ljapunov exponent.
We consider a perturbed Hamiltonian system in the action-angle variables in presence of a weak chaotic dynamics in the phase space and it is possible to get a phenomenological description of the dynamics by mean of stochastically perturbed Hamiltonian system

$$H_0(I) + \xi(t) H_1(\theta, I) \tag{4.1}$$

where (θ, I) are action-angle variables and the noise realizations have an exponential decaying correlation. The justification of such models is the possibility to describe the effect of a coupling of an integrable Hamiltonian with an Hamiltonian chaotic environment that introduces fluctuations in the phase space. But one can also simulate the effect of weak chaotic layers in the phase space that introduce pseudo random kicks in the evolution of the action variables. To maintain the symplectic character of the dynamics we consider the regular (colored) stationary noise so that

$$\xi(t) = \int_\omega \hat{\xi}(\omega) e^{i\omega t} d\omega$$

where $\hat{\xi}(\omega)$ are random variables such that

$$\mathrm{E}(\hat{\xi}(\omega)) = 0 \qquad \mathrm{E}(\hat{\xi}(\omega)\bar{\hat{\xi}}(\omega')) = F(\omega)\delta(\omega - \omega')$$

where $F(\omega)$ is the power spectrum of the noise. We recall that

$$\mathrm{E}(\xi(t+\tau)\xi(t)) = \int_\omega \int_{\omega'} \mathrm{E}(\hat{\xi}(\omega)\bar{\hat{\xi}}(\omega')) e^{i(\omega-\omega')t + i\omega\tau} d\omega d\omega' = \int_\omega F(\omega) e^{i\omega\tau} d\omega$$

so that the correlation function is the Fourier transforms of the power spectrum $F(\omega)$. For Markov process one gets

$$\Phi(\tau) = \mathrm{E}(\xi(t+\tau)\xi(t)) \simeq \sigma^2 e^{-\gamma\tau}$$

where one can set $\sigma^2 = \gamma/2$ to get the white noise limit. Then

$$F(\omega) = \sigma^2 \int_0^\infty e^{-\gamma|\tau| - i\omega\tau} d\tau = \frac{\sigma^2\gamma}{\pi(\gamma^2 + \omega^2)}$$

A Fourier analysis of the orbits where we distinguish between a discrete spectrum related to the phantom of a regular orbit and a continuous like spectrum that is related to the chaotic behavior.The perturbation $H_1(\theta, I)$ describes the amplitude of the fluctuations whereas the "noise" $\xi(t)$ takes into account the correlation properties of the fluctuations: the noise realizations $\xi(t)$ depend on a probability space (this space can represent hidden degrees of freedom and the correlation of the related stochastic process can be related to the chaotic character of the dynamics). In the case of a weak chaotic region the realization depend on the initial condition and the system (**??**) is an effective description of the dynamics. Of course the measure of the chaotic region should be large to avoid the dynamical traps of Hamiltonian systems which are due to the stickiness of trajectories to some specific domains in phase space. In the case of a weak chaos the white noise approximation cannot be directly applied to the process $\xi(t)$, but one expects that $\|H_1\| \ll 1$ so that it is possible to introduce a slow diffusion time $\tau \propto \|H_1\|^2 t$ for the action variable at which the white noise approximation is justified. Indeed the "noise" depends on the angles variables that can be considered as a fast variable, which are independent in a time much shorter than the diffusion time scale. We introduce the slow variables

$$\phi = \theta - \Omega(I)t$$

and we have a new Hamiltonian

$$H(\phi, I) = \xi(t) H_1(\phi + \Omega(I)t, I)) \tag{4.2}$$

and we have the Fourier expansion

$$H_1(\phi + \Omega(I)t, I) = \sum_k h_k(I) \exp(ik(\phi + \Omega(I)t))$$

in the action angle variables of the unperturbed system $H_0(I)$.
Remark: the change of variables requires a covariance property for the equations so that we need to keep the Stratonovich interpretation in the white noise limit. To find an approximate solution of the stochastic dynamics (**??**) we consider the evolution of angle-action variables (ϕ, I) for a time T and we get the map

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

where $\Delta\phi = \phi(T) - \phi(0)$ and $\Delta I = I(T) - I(0)$. The norm $\|H_1\| \ll 1$ is the perturbation parameter and the time interval T should be sufficiently long to consider independent the noises $\xi(t)$ and $\xi(t+T)$ but it has to be short with respect the evolution time of the action and the slow phase. We observe that the fluctuations in the angle ϕ contains the effect of the secular term $\propto t\partial\Omega/\partial I$ so that the sensitivity to the fluctuations of the process $\xi(t)$ increases during the time interval T. In a perturbation approach where $\|H_1\| = \epsilon \ll 1$ the changes of the action is or order $|\Delta I| \propto \epsilon T\sqrt{\langle\xi^2(t)\rangle}$ whereas the secular terms provide the estimate

$$|\Delta\phi| \propto \epsilon \left\|\frac{\partial\Omega}{\partial I}\right\| T^2\sqrt{\langle\xi^2(t)\rangle} + O(T)$$

which means that if it is possible to choose $T \gg 1$ the angle fluctuations are dominated by the secular term assuming the non degeneracy condition

$$\left\|\frac{\partial\Omega}{\partial I}\right\| = O(1)$$

The stationary solution for a diffusion process in the angle is the uniform distribution, but if $\Delta\phi$ is small during a time interval T the existence of resonances $k\Omega(I) = 0$ does not allow to average the angles in the action dynamics. Let us estimate the variance the angle variables using the dominating term for an evolution time T_ϕ assuming that the action variables can be considered constant during such time interval (i.e. $T_\phi\epsilon \ll 1$). One gets

$$\mathrm{Var}[\Delta\phi]_{ij} \propto \int_0^{T_\phi}\int_0^{T_\phi} ts\frac{\partial H_1}{\partial\theta_k}(\phi+\Omega t,I))\frac{\partial H_1}{\partial\theta_m}(\phi+\Omega s,I))\frac{\partial\Omega_m}{\partial I_i}\frac{\partial\Omega_k}{\partial I_j}\Phi(t-s)\,dtds$$

and in the white noise limit (i.e. the maximum Ljapunov exponent defines a time scale $\lambda_{\max}^{-1} \ll T_\phi$ we get the following estimate

$$\|\mathrm{Var}[\Delta\phi]\|\,(T_\phi) \propto \epsilon^2\left\|\frac{\partial\Omega}{\partial I}\right\|^2\frac{T_\phi^3}{3} \tag{4.3}$$

The angles are defined in a n-dimensional torus $\mathbb{T}^n$ and the relaxation toward a uniform distribution has a characteristic time scale T_ϕ if

$$\sqrt{\|\mathrm{Var}[\Delta\phi]\|\,(T_\phi)} \simeq (2\pi)^n$$

Then T_ϕ has to satisfy

$$T_\phi^{3/2} \simeq \frac{(2\pi)^n}{\epsilon\,\|\partial\Omega/\partial I\|} \tag{4.4}$$

i.e. $T_\phi \propto \epsilon^{-2/3}$ and the requirement $T_\phi \epsilon \simeq O(\epsilon^{1/3}) \ll 1$ can be satisfied. At this time the angles can be considered uniformly distributed and an averaging principle can be applied. We recall also the condition $\lambda_{\max} \gg \epsilon^{2/3}$ on the local Ljapunov exponent to avoid long time trapping in small regions. On an ergodic component all the orbits have the same Ljapunov exponents and the previous requirement can be satisfied if the Ljapunov time scale (i.e. the convergence scale of the maximum Ljapunov exponent has to be much faster than $\epsilon^{2/3}$).
We remark that if we have an almost isochronous system

$$\left\| \frac{\partial \Omega}{\partial I} \right\| = O(\epsilon)$$

the previous estimate gives a relaxation time scale $T_\phi \simeq \epsilon^{-4/3}$. Let us consider a time interval T for the evolution of the actions

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

where it is understood $H_1(\phi + \Omega(I)t, I)$ and we neglect terms of order $o(\epsilon^2)$. We observe that an explicitly computation of the fluctuations contribution (i.e. the first term in the expansion) if the slow phases ϕ can be considered constant during T is performed by using the Fourier expansion

$$\begin{aligned}\int_0^T \frac{\partial H_1}{\partial \theta_j} \xi(s) ds = & \sum_k \int_\omega \int_0^T i k_j h_k(I) \exp(ik(\phi + \Omega(I))s + \omega s) \hat{\xi}(\omega) ds\, d\omega \\ = & \sum_k i k_j h_k(I) \exp(ik\phi) \int_\omega \frac{\exp(i(k\Omega(I) + \omega)T) - 1}{i(k\Omega(I) + \omega)} \hat{\xi}(\omega) d\omega\end{aligned}$$

where we observe the appearance of small denominators that provides a contribution to the fluctuations is of order $O(T)$ when $k\Omega(I) + \omega = 0$ (i.e. the non-linear resonance conditions). The variance per unit time of the fluctuating terms reads

$$\begin{aligned}D_{jl}(I, \phi) = & \frac{1}{T} \sum_{k,n} k_j n_l h_k(I) \bar{h}_n(I) e^{i(k-n)\phi} \\ & \int_\omega \frac{(e^{i(k\Omega(I)+\omega)T} - 1)(e^{-i(n\Omega(I)+\omega)T} - 1)}{(k\Omega(I) + \omega)(n\Omega(I) + \omega)} F(\omega) d\omega\end{aligned}$$

The integral can be written in the form

$$\begin{aligned}D_{jl}(I, \phi) = & \frac{4}{T} \sum_{k,n} k_j n_l h_k(I) \bar{h}_n(I) e^{i(k-n)(\phi + \Omega(I)T)} \\ & \int_\omega \frac{\sin((k\Omega + \omega)T/2) \sin((n\Omega + \omega)T/2)}{(k\Omega(I) + \omega)(n\Omega(I) + \omega)} F(\omega) d\omega\end{aligned}$$

However if $T \simeq T_\phi$ the phase can be considered relax and only the terms $k = n$ contribute. It is possible to prove that the limit

$$D_{jl}(I) = \lim_{T \to \infty} \frac{4}{T} \sum_k k_j k_l h_k(I) \bar{h}_k(I) \int_\omega \frac{\sin^2((k\Omega(I) + \omega)T/2)}{(k\Omega(I) + \omega)^2} F(\omega) d\omega \quad (4.5)$$

If the power spectrum is peaked at $\omega = \omega_*$ the variance of the action fluctuations is maximal at $I = I_*$ when $k\Omega(I_*) - \omega_* \simeq 0$. If $F(\omega) = const.$ (white noise limit), using the integral

$$\int_{-\infty}^{\infty} \frac{\sin^2(ax)}{x^2} dx = \pi a$$

one can prove that

$$D_{jl}(I) = \epsilon^2 \left\langle \frac{\partial \hat{H}_1}{\partial \theta_j} \frac{\partial \hat{H}_1}{\partial \theta_l} \right\rangle_\theta \tag{4.6}$$

that corresponds to the random phase limit.
Remark: the Fourier components of the fluctuations $F(\omega)$ depends on how the orbits explore the ergodic component; the condition $F(\omega) = const.$ is the white noise that implies large fluctuations (i.e. large Ljapunov exponent) but for a weak chaotic region $F(\omega)$ has a non trivial structure.
The result (**??**) holds only if we have separated relaxation scales for the slow phases and the actions. This condition means that we have not resonance structures in the phase space in the considered region. The local character of the action-angle variables opens the problem of performing a gluing procedure of the different charts in the phase space.
Ifthe relaxation time of the angles ϕ is much faster than the diffusion time of the action variables $\epsilon^2 T_\phi \ll 1$, the angles themselves can be treated as fast random variables that characterize the different noise realizations independent from the random perturbation $\xi(t)$, then we consider the action average dynamics with respect both the random perturbation and the angles ϕ

$$\begin{aligned}
\langle \Delta I_j \rangle(T) &= \frac{1}{2} \frac{\partial}{\partial I_k} \int_0^T \int_0^T \frac{\partial H_1}{\partial \theta_j}(\phi + \Omega t) \frac{\partial H_1}{\partial \theta_k}(\phi + \Omega s) \Phi(t-s) ds\, dt \\
&= \frac{1}{2} \frac{\partial}{\partial I_k} \sum_k k_j k_l |h_k(I)|^2 \int_0^T \int_0^T \exp(ik\Omega(t-s)) \Phi(t-s) ds\, dt
\end{aligned}$$

where we have neglected the contribution of the terms with zero mean value with respect to the angle variables. We remark that if $\Phi(\tau)$ is decaying sufficiently fast for $\tau \to \infty$ then it exists the limit

$$\langle \Delta I_j \rangle = \lim_{T \to \infty} \frac{1}{2T} \frac{\partial}{\partial I_k} \sum_k k_j k_l |h_k(I)|^2 \int_0^T \int_0^T \exp(ik\Omega(I)(t-s)) \Phi(t-s) ds\, dt$$

and in the white noise limit the expression reads

$$\langle \Delta I_j \rangle = \frac{1}{2T} \frac{\partial}{\partial I_k} \sum_k k_j k_l |h_k(I)|^2$$

The corresponding variance for the fluctuating terms reads

$$\langle \Delta I_j \Delta I_l \rangle = \frac{\partial}{\partial I_k} \sum_k k_j k_l |h_k(I)|^2 \int_0^T \int_0^T \exp(ik\Omega(t-s)) \Phi(t-s) ds\, dt + O(\epsilon^4)$$

and it coincides with (**??**) if divided by T. We introduce the diffusion time $\tau = \epsilon^2 T \ll 1$ to apply a diffusion limit $\epsilon \to 0$ $T_\phi \to \infty$. Then in the limit $\epsilon \to 0$, we choose a time step $T \gg \gamma^{-1}$ (i.e. the Ljapunov time scale of the fluctuations) so that the fluctuations of $\Delta\phi$ and ΔI can be considered independent at successive time intervals. If we consider an evolution time scale $T_\phi \simeq \epsilon^{-2/3}$ so that $\tau = O(\epsilon^{4/3}) \ll 1$, the slow phases relax at each time step and the action evolution can be described by a diffusion process whose diffusion and drift coefficients are computed averaging on the slow angle variables

$$\Delta I_j = -\sqrt{T_\phi \epsilon^2}\sqrt{D_{jl}(I)}\hat{\xi}_l + \frac{T_\phi \epsilon^2}{2}\frac{\partial}{\partial I_l}D_{jl}(I) \tag{4.7}$$

that is approximated by the stochastic differential equation

$$dI_j = \frac{1}{2}\sum_l \left(\frac{\partial D_{jl}}{\partial I_l}(I)d\tau + \sqrt{D_{jl}(I)}dw_l(\tau)\right) \tag{4.8}$$

where $D_{ij}(I)$ is defined by the equation (**??**)and $dw_l(\tau)$ are independent Wiener process in the slow time. 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

$$\frac{\partial \rho}{\partial \tau} = -\frac{1}{2}\frac{\partial}{\partial I_j}\left[\frac{\partial}{\partial I_k}D_{jk}(I)\right]\rho(I,\tau) + \frac{1}{2}\frac{\partial^2}{\partial I_j \partial I_k}D_{jk}(I)\rho(I,\tau)$$

After some algebraic calculations one gets

$$\frac{\partial \rho}{\partial \tau} = \frac{1}{2}\frac{\partial}{\partial I_j}D_{jk}(I)\frac{\partial}{\partial I_k}\rho(I,\tau) = \mathcal{D}_I \rho \tag{4.9}$$

In a generic case $D_{jk}(I)$ is strictly positive defined symmetric matrix and the stationary solution is trivial $\rho_s = const.$. It is easy to prove that the FP operator (**??**) is a self-adjoint operator with all non positive real eigenvalues. The operator (**??**) is defined on the space of probability distribution $\rho(I)$ which is an invariant space with the scalar product

$$(\rho_1, \rho_2) = \int \rho_1(I)\rho_2(I)\, dI$$

and it is easy to check

$$(\rho_1, \mathcal{D}_I \rho_2) = (\mathcal{D}_I \rho_1, \rho_2)$$

with natural boundary conditions (i.e. ρ and its derivatives vanish at the boundaries). Moreover we have

$$\frac{d}{dt}\|\rho\|^2 = \int dI\, \rho(I)\frac{\partial}{\partial I_j}D_{jk}(I)\frac{\partial \rho}{\partial I_k} = -\int dI \frac{\partial \rho}{\partial I_j}D_{jk}(I)\frac{\partial \rho}{\partial I_k} \le 0$$

so that $\|\rho\|^2(t)$ is a Ljapunov function for eq. (**??**) which is strictly decreasing except when $\partial\rho/\partial I = 0$. From the previous inequality it follows that all the eigenvalues are negative except $\lambda = 0$ and $|\lambda|^{-1}$ defines the relaxation time scale of the corresponding eigenvector. The eigenvectors define an orthonormal basis and we have the characteristic equation

$$\lambda \rho_\lambda = \frac{1}{2}\sum_{jk}\frac{\partial}{\partial I_j}D_{jk}(I)\frac{\partial}{\partial I_k}\rho_\lambda$$

The eigenvectors are orthogonal and if $\lambda \neq 0$ we have the condition

$$\int \rho_\lambda(I)\, dI = 0$$

i.e. the orthogonality with respect to the stationary condition. An interesting case is when $D(I)$ is an homogeneous polynomial of order 2 (quadrupolar noise): i.e. according to the definition H_1 is linear in the action and we have a linear multiplicative noise. A formal approach looks for solution ρ_λ in the form of homogeneous polynomials. The space of homogeneous polynomial is invariant for the FP operator so that we get a linear system in the coefficients. The eigenvalue λ has to be chosen in order that the solution is not trivial. The polynomials are not summabale so that the previous condition has to be interpreted.

4.1 Example of Physical Systems

We consider some physical examples: assume the unperturbed Hamiltonian is linear with a dipole noise

$$H(q,p) = \omega \frac{p^2+q^2}{2} + \epsilon \xi(t) q$$

The stochastic equations read

$$\begin{aligned} \dot q &= \omega p \\ \dot p &= -\omega q - \epsilon \xi(t) \end{aligned}$$

The solution of the system is a Gaussian random process in the white noise limit

$$\begin{pmatrix} q(t) & p(t) \end{pmatrix} = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} q_0 & p_0 \end{pmatrix} + \epsilon \int_0^t \begin{pmatrix} \cos(\omega s) & \sin(\omega s) \\ -\sin(\omega s) & \cos(\omega s) \end{pmatrix} \begin{pmatrix} 0 \\ \xi(s) \end{pmatrix} ds$$

The covariance of the process is increasing with t as

$$\frac{\epsilon^2}{2} \begin{pmatrix} t + \sin(2\omega t)/2\omega & (1-\cos(2\omega t)/\omega \\ (1-\cos(2\omega t)/\omega & t - \sin(2\omega t)/2\omega \end{pmatrix}$$

By averaging on the fast angle variable we have the FP equation

$$\frac{\partial \rho}{\partial t} = \epsilon^2 \frac{\partial}{\partial I} I \frac{\partial \rho}{\partial I}$$

in the action

$$I = \frac{p^2+q^2}{2}$$

A direct calculation gives

$$\langle I \rangle = \frac{\langle p^2 \rangle + \langle q^2 \rangle}{2} = \epsilon^2 t$$

Using the average FP equation we get

$$\frac{d\langle I \rangle}{dt} = \int I \frac{\partial \rho}{\partial t}\, dI = \epsilon^2 \int I \frac{\partial}{\partial I} I \frac{\partial \rho}{\partial I}\, dI = \epsilon^2 \int \rho\, I = \epsilon^2$$

where we integrate by part two times. Then we recover the correct result and we have a diffusion process. The second moment can be computed in the same way

$$\frac{d\langle I^2\rangle}{dt} = \epsilon^2 \int I^2 \frac{\partial}{\partial I} I \frac{\partial \rho}{\partial I}\, dI = 4\epsilon^2 \int I \rho\, I = 4\epsilon^2 \langle I\rangle$$

Then we get

$$\mathrm{Var}(I) = 2\epsilon^4 t^2 - \epsilon^4 t^2 = \epsilon^4 t^2$$

This result is consistent with a exponential distribution

$$\rho(I,t) \propto \epsilon^2 t \exp\left(-\frac{I}{\epsilon^2 t}\right)$$

where the average value and the rms coincide.
We consider the more general system (this system simulates a quadrupolar noise in a circular accelerator)

$$H(q,p;\xi) = \sum_i \omega_i \frac{p^2+q^2}{2} + \epsilon \frac{\xi(t)}{2} \sum_{ij} q_i \Omega_{ij} q_j$$

so that we have the motion equations

$$\begin{aligned} \dot{q}_i &= \omega_i p_i \\ \dot{p}_i &= -\omega_i q_i - \epsilon \xi(t) \sum_j \Omega_{ij} q_j \end{aligned}$$

where the Stratonovich interpretation is necessary to keep the physical interpretation. We observe that all the degrees of freedom feels the same noise $\xi(t)$ since the physical origin of the noise is a fluctuation in an external linear field that introduces a random coupling among the components. The solution of this problem is strictly related to the random matrix theory since the evolution can be approximated by a product of non-commuting random matrices, by discretizing the noise at a time scale Δt so that we get a stepwise linear system. As previously remarked in the stochastic case the time step Δt has a physical meaning since it is the correlation time scale of the noise if we assume independent fluctuations each time step. Let $\xi(t)$ a stochastic noise, its statistical properties are mainly characterized by the average value and the correlation (indeed these properties are related to the CL Theorem). The average is usually set $\langle \xi(t)\rangle = 0$ and the correlation reads (we assume a stationary condition)

$$\langle \xi(t+\tau)\xi(t)\rangle = c(\tau) \qquad c(\Delta t) \simeq 0$$

This implies that the continuous limit (i.e. the possibility of modeling the fluctuation using a Wiener noise) implies that the covariance has to change to maintain the fluctuation effect. The effective scaling law is that one has to increase the fluctuation variance proportionally to Δt^{-1} so that $c(\tau) \to \delta(\tau)$. An alternative is to consider the small noise limit $\epsilon \to 0$ and to introduce a slow diffusion time so that the slow time step $\epsilon^2 \Delta t \to 0$ even if Δt is finite.
Let us consider the problem to understand the stability time scale of the orbit: i.e. the estimate the positive Ljapunov exponent for the system. Using the linear action-angle variables

$$q_i = \sqrt{2I_i}\sin\theta_i \qquad p_i = \sqrt{2I_i}\cos(\theta_i)$$

one computes the diffusion coefficient from the definition of the average action dynamics

$$D_{ij} = \left\langle \frac{\partial H_1}{\partial \theta_i} \frac{\partial H_1}{\partial \theta_j} \right\rangle \qquad H_1(I,\theta) = \sum_{ij} \sqrt{I_i I_j} \sin\theta_i \Omega_{ij} \sin\theta_j$$

assuming the possibility of averaging on the fast angle variables (i.e. the distribution can be always considered relaxed in the angle variables). An explicit calculation gives

$$\begin{aligned} D_{ij}(I) &= 2\epsilon^2 \sum_{kh} \sqrt{I_i I_j I_k I_h} \Omega_{ik} \Omega_{jh} \langle \cos\theta_i \cos\theta_j \sin\theta_k \sin\theta_h \rangle \\ &= \frac{\epsilon^2}{2} \delta_{ij} \left[I_i \sum_k I_k \Omega_{ik} \Omega^T_{ki} - \frac{1}{2} I_i^2 \Omega_{ii}^2 \right] \end{aligned}$$

since the non-vanishing contributions require $i = j$ and $k = h$. The average FP equation (cfr eq. (**??**)) reads

$$\frac{\partial \rho}{\partial t} = \frac{\epsilon^2}{2} \frac{\partial}{\partial I_i} \left[I_i \sum_k I_k \Omega_{ik} \Omega^T_{ki} - \frac{1}{2} I_i^2 \Omega_{ii}^2 \right] \frac{\partial \rho}{\partial I_i}$$

where $\rho(I,t)$ is the distribution in the action variables. We compute the evolution of the average value $\langle I_h \rangle$ according to the relation

$$\frac{d}{dt} \langle I_h \rangle = \int I_h \frac{\partial}{\partial I_i} D_{ij}(I) \frac{\partial \rho}{\partial I_j} \, dI = \int \rho \frac{\partial}{\partial I_j} D_{hj}(I) \, dI = \left\langle \frac{\partial}{\partial I_j} D_{hj}(I) \right\rangle$$

and we get

$$\frac{d}{dt} \langle I_h \rangle = \frac{\epsilon^2}{2} \frac{\partial}{\partial I_h} \left[I_h \sum_k I_k \Omega_{hk} \Omega^T_{kh} - \frac{1}{2} I_h^2 \Omega_{hh}^2 \right] = \frac{\epsilon^2}{2} \sum_k \Omega_{hk} \Omega^T_{kh} \langle I_k \rangle$$

The eigenvalue of the matrix $\Omega \Omega^T$ are positive so that the solution $\langle I_h \rangle(t)$ increases exponentially. However we justified this approach in a nonlinear case where the relaxation process of the angle variables is related to the change of the frequencies with the actions. In a different approach it is convenient to use the moving variables (Q, P)

$$\begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} Q \\ P \end{pmatrix}$$

and we get

$$\begin{aligned} \dot{Q}_i &= \epsilon \xi(t) \sin(\omega_i t) \sum_j \Omega_{ij} (Q_j \cos(\omega_j t) + P_j \sin(\omega_j t)) \\ \dot{P}_i &= -\epsilon \xi(t) \cos(\omega_i t) \sum_j \Omega_{ij} (Q_j \cos(\omega_j t) + P_j \sin(\omega_j t)) \end{aligned}$$

The diffusion time scale depends on the spectrum and it scales as ϵ^{-2}. This result is justified in the white noise limit for the fluctuation $\xi(t)$ so that it is possible to average on the noise at a fixed value of the angles and then we use the fast angle dynamics assuming that the frequencies ω are non-resonant. The variables (Q,P) are both slow variables and using the diffusion time scale $\tau=\epsilon^2 t$ the system can be written in the form

$$\frac{dQ_i}{d\tau} = \hat{\xi}(\tau)\sin(\omega_i/\epsilon^2\tau)\sum_j \Omega_{ij}(Q_j\cos(\omega_j\epsilon^2\tau) + P_j\sin(\omega_j/\epsilon^2\tau))$$
$$\frac{dP_i}{d\tau} = -\hat{\xi}(\tau)\cos(\omega_i/\epsilon^2\tau)\sum_j \Omega_{ij}(Q_j\cos(\omega_j/\epsilon^2\tau) + P_j\sin(\omega_j/\epsilon^2\tau))$$

where $\hat{\xi}(\tau)$ tends to white noise in the slow time. We observe that any small change in τ corresponds to a large change in the phases $\theta_i=\omega_i/\epsilon^2\tau$ that can be considered random variables. The previous system can be approximated by a stochastic dynamics of the form

$$dZ = A(\theta)Z\,dw_\tau$$

where $A(\theta)dw_\tau$ are random matrices defining the increments of a Wiener process with $A(\theta)$ independent of w_τ. The evolution is described by a FP equation (**??**) with diffusion coefficient

$$D(Z) = \langle A(\theta)ZA(\theta)Z\rangle_\theta$$

Remark: the previous approach assume $\omega_i \simeq O(1)$; in a general situation (non linear coupling) the averaging procedure fails if the frequencies ω_i are almost resonant

$$\sum_i k_i\omega_i \simeq \epsilon^2$$

Since the average values vanish the stability study is related to the covariance matrix A direct calculation is possible using the evolution equation for the observables

$$\frac{d}{dt}\langle Q_iQ_j\rangle =$$

******

Given an initial condition I_0 the solution of the FP equation can be interpreted as the probability to realize a path that move from I_0 to a value I is a time t. By introducing the probability density

$$\pi^t(I\,|I_0) = \rho(I,t|I_0)$$

so that we get the Kolmogorov equation

$$\rho(I,t;\rho_0) = \int \pi^t(I\,|\,I_0)\rho_0(I_0)\,dI_0 \tag{4.10}$$

Moreover the transition probabilities satisfy the condition

$$\pi^t(I\,|\,I_0) = \int \pi^{t-s}(I\,|\,J)\pi^s(J\,|I_0)\,dJ \tag{4.11}$$

******

The particular form of the FP operator is strictly related to the action variables of the unperturbed Hamiltonian. The interpretation of the FP equation as a continuity equation introduces the concept of probability currents

$$J_j[\rho] = -\sum_k D_{jk}(I)\frac{\partial \rho}{\partial I_k}$$

and the stationary condition is written in the form $\text{div}\vec{J} = 0$. The DB condition would imply $J_j[\rho^s] = 0$ that is consistent with the condition $\rho^s = const.$ if D_{jk} does not have a null eigenvalue. The self-adjoint character of the operator implies that if $H(I)$ is any observable its evolution $H(I_0, t)$ satisfies the same equation

$$\frac{\partial H}{\partial \tau}(I_0, \tau) = \frac{1}{2}\frac{\partial}{\partial I_j} D_{jk}(I)\frac{\partial}{\partial I_k} H(I_0, \tau) \tag{4.12}$$

This is related to the Ito formula

$$dH = \frac{1}{2}\sum_{jl}\left(\frac{\partial D_{jl}}{\partial I_l}(I)d\tau + \sqrt{D_{jl}(I)}dw_l(\tau)\right)\frac{\partial H}{\partial I_j} + \frac{1}{4}\sum_{jl}\frac{\partial^2 H}{\partial I_j \partial I_l} D_{jl}(I)d\tau$$

Assuming we know a set of independent firat integral $H_k(I)$ of first integrals of motion for the unperturbed system H_0, we compute the fluctuations variance

$$\langle dH_k dH_h \rangle = \frac{1}{4}\sum_{jl}\frac{\partial H_k}{\partial I_j} D_{jl}(I)(\tau)\frac{\partial H_h}{\partial I_l} d\tau + O(d\tau^2)$$

and we get the relation

$$D^H_{kh} = \frac{\langle dH_k dH_h \rangle}{d\tau} = \frac{1}{4}\sum_{jl}\frac{\partial H_k}{\partial I_j} D_{jl}(I)\frac{\partial H_h}{\partial I_l}$$

The relation between D^H and D is defined by the Jacobian matrix $\partial H_k/\partial I_j$ that changes the metric matrix if it is not orthogonal. The spectrum problem is characterized by the condition

$$\max_{\|\Delta I\|=1}\sum_{jl}\Delta I_j D_{jl}\Delta I_l = \max_{\|\Delta I\|=1}\sum_{kh}\Delta I_k D^H_{kh}\Delta I_h$$

where the condition $\|\Delta I\| = 1$ has to be rewritten from the relation

$$\Delta I_j = \sum_k \frac{\partial I_j}{\partial H_k}\Delta H_k \quad \Rightarrow \quad \|\Delta I\|^2 = \sum_{jkh}\frac{\partial I_j}{\partial H_k}\frac{\partial I_j}{\partial H_h}\Delta H_k \Delta H_h$$

The metric matrix

$$G^H_{kh} = \frac{\partial I_j}{\partial H_k}\frac{\partial I_j}{\partial H_h}$$

defines a deformation of the unit sphere and it has to be estimated in the inequality

$$\|D\| \leq \|D^H\| \|G^H\|$$

that is related to the perturbation estimate H_1 whereas $\|D^H\|$ is affected by the definition of the integrals H.

4.2 Numerical computations of Diffusion Coefficient

The numerical computation of the diffusion coefficient has to cope with the problem that one does not known the unperturbed action I that reduce the Hamiltonian in the perturbation form

$$H(I, \theta) = H_0(I) + H_1(\theta, I)$$

where $\|H_1(\theta, I)\|$ is an optimal remainder of a perturbation theory in a weak chaotic region (possible homogeneous). The action-angle variables are the result of a perturbation approach but we have not an explicit definition; for any initial condition in a Poincaré section we assume that there exist a action value and that the angle variables have a fast dynamics due to the frequencies $\Omega(I) = \partial H_0/\partial I$. If one introduces the slow angle variable $\theta = \phi + \Omega(I)t$ the new Hamiltonian reads

$$H(I, \phi) = H_1(\phi + \Omega(I)t, I)$$

but we have secular terms in the angle dynamics

$$\dot{\phi} = -\frac{\partial H_1}{\partial I} - \frac{\partial H_1}{\partial \theta}\frac{\partial \Omega}{\partial I}t$$

We remark that if one integrates the dynamics $(\theta(t), I(t))$ to recover the dynamics $\phi(t)$ one can perform the reverse dynamics

$$\phi(t) = \theta(t) - \Omega(I(t))t$$

The chaotic region can be characterized by the presence of fast fluctuations in the dynamics of the slow variables (ϕ, I) due to local positive Ljapunov exponents. The local Ljapunov exponent is mainly related to the local structure of the phase (i.e. the presence of hyperbolic structures) whereas the Ljapunov exponent depends on how the orbit explores the phase space structure. The local fluctuations accumulate to induce an exponential decorrelation among the orbits but the stickiness phenomenon in Hamiltonian systems may introduce long term correlation in the orbit evolution. The existence of a positive Ljapunov exponent allows to consider different orbits as independent noise realizations in a numerical approach using a coarse grained description of the dynamics. For a given tessellation of the phase space $\{E_k\}$ we associate the matrix

$$\Phi^t_{hk} = \frac{\mu(\Phi^{-t}(E_n) \cup E_k)}{\mu(E_k)}$$

where $\mu(E)$ is the invariant measure (i.e. the Lebesgue measure for Hamiltonian system in the phase space) and Φ^t the phase flow. The matrix Φ^t_{hk} can be interpreted as a transition probability matrix that gives the probability to be in the 'state' E_n at time t conditioned to the information that the initial condition belongs to E_k. The partition $\{E_k\}$ covers an invariant set of the phase space (but one can introduce an absorbing barrier to study the stability properties) so that the condition

$$\sum_h \Phi^t_{hk} = 1$$

is satisfied. One can associate a Ljapunov exponent for any set E_k by considering the expectation values of the Ljapunov exponent for all the orbits with initial condition $\in E_k$ (possibly neglecting low probability outliers that do not affect the average evolution). We remark that the spatial scales play a fundamental role since they define what is local and global in the coarse grained dynamics: the Hamiltonian character introduces infinite time and spatial scales in the system (consider the stickiness phenomenon or the Arnold's diffusion) but the tessellation associates to the E_k the same dynamical features. This procedure requires a tolerance to define what is true in probability or with high probability) and what is exceptional in the evolution.
the existence of an average positive Ljapunov exponent λ for any set E_k allows to define a decorrelation time scale $T_l \simeq \lambda^{-1}$ and to approximate the evolution of a distribution function $\rho(h,t) = P(E_h,t)$ by a Markov process

$$\rho(h, t+\Delta t) = \sum_k \Phi^{\Delta t}_{hk} \rho(hk, t)$$

using $\Delta t \simeq \lambda^{-1}$. The coarse grained distribution describes the system evolution when each iteration the physical orbit that performs a transition $E_k \to E_h$ is substituted by another orbit chosen in the subset E_h for the next step. The chaotic behavior in each partition element allows to consider independent each iteration to justify the Markov character.
From the definition of transition matrix $\pi^{\Delta t}_{xy}$ of diffusion process, we have the relations with the drift and diffusion coefficient

$$a(y) = \frac{1}{\Delta t} \sum_x (x-y) \pi^{\Delta t}_{xy} + o(1)$$

$$D(y) = \frac{1}{\Delta t} \sum_x (x-y)^2 \pi^{\Delta t}_{xy} + o(1)$$

whereas no other contribution comes from the higher momenta (this is a consequence of the Central Limit Theorem). The problem is that one cannot perform a limit $\Delta t \to 0$ and the previous formula holds only if the evolution time scale of the variable y is much less than the Ljapunov time scale Δt. This is not the case for the angle variables θ.
The Markov property requires that most of the orbits starting from a set E_h has a local positive Ljapunov exponent, so that each time we shuffle the initial condition we have an exponential increasing from the initial orbit with high probability. Therefore using a coarse grained description we introduce a noise each time step Δt whose amplitude depends on the partition radius $R(E_h)$ (a continuum limit would requires $R(E_h) \ll 1$. The time interval Δt is dependent on the local Ljapunov exponent. If the Markov properties is justified a possible approach is to substitute the initial Hamiltonian with an effective Hamiltonian of the form

$$H(I,\theta) = H_0(I) + \xi(t) \hat{H}_1(\phi + \Omega(I)t, I)$$

where $\xi(t)$ is a stochastic noise with a correlation time Δt. The action dynamics gives

$$\dot{I} = -\xi(t) \frac{\partial \hat{H}_1}{\partial \theta}(\theta, I)$$

Let $(\theta(t), I(t))$ a given orbit, we linearize the dynamics

$$\delta \dot{I} = -\xi(t) \frac{\partial^2 \hat{H}_1}{\partial \theta \partial I}(\phi(t) + \Omega(I)t, I(t))\delta I = -H_{\theta I}(t)\delta I$$

To estimate a local Ljapunov exponent for the actions we assume the ansatz $(\phi(t), I(t))$ almost constant in the time interval Δt and we write the solution as

$$\delta I(\Delta t) \simeq \mathcal{T} \exp\left(-\int_0^{\Delta t} \xi(s) H_{\theta I}\, ds\right) \delta I_0$$

Using the estimate in the white noise limit

$$\left\langle \int_0^{\Delta t} ds \int_0^s du H_{\theta I}(s) H_{\theta I}(u) \xi(s) \xi(u) \right\rangle$$

Using an evolution time scale Δt we compute the drift and the diffusion coefficient for the action variable associating a single value for the coordinate in each set

$$a(I_k, \theta_k) = \frac{1}{\Delta t} \sum_h (I_h - I_k) \pi_{hk}^{\Delta t}$$

$$D(I_k, \theta_k) = \frac{1}{\Delta t} \sum_h (I_h - I_k)^2 \pi_{hk}^{\Delta t}$$

******

Chapter 5

Perturbation Theory and Stochastic Hamiltonians

The Stratonovich interpretation of the stochastic differential equation allows to develop a perturbation theory to the stochastic Hamiltonian dynamics using a covariant description of the motion equation

$$dx = D_H x\, dt + \sqrt{2T\gamma} D_{H_1} x \circ dw_t - \gamma\{H, H_1\} D_{H_1} x\, dt \qquad x(q,p)$$

in the limit of a white noise fluctuations $\xi(t)$. If $H(q,p)$ is a perturbed Hamiltonian one can perform the action-angle variable transformation to reduce the system in the form

$$H(q,p) = H_0(I) + \epsilon \hat{H}(I,\theta)$$

Let $x = \Phi(X)$ a canonical change of variables we have that the canonical equations $\dot{x} = D_H x$ are invariant in form if we use the new Hamiltonian $H(\Phi(X))$

$$D_{H(x)} x = D_{H(X)} X$$

The same is true for the term $D_{H_1} x$ assuming the Stratonovich interpretation for the stochastic differential. The Poisson bracket $\{H, H_1\}$ is a scalar and it can be computed using the invariance relation

$$\{H, H_1\}(\Phi(X) = \{H(\Phi(X), H_1(\Phi(X)\}$$

We get the differential equation

$$dX = D_H X\, dt + \sqrt{2T\gamma} D_{H_1} X \circ dw_t - \gamma\{H, H_1\}(X) D_{H_1} X\, dt$$

and we can apply a perturbation theory to study the solutions. If the initial Hamiltonian is integrable we can introduce the action angle variable and the previous equation takes the form

$$\begin{aligned} d\theta_k &= \Omega_k(I)\, dt + \sqrt{2T\gamma}\, dt \\ dI_k &= \frac{1}{2}\pi r^2 \end{aligned} \tag{5.1}$$

ciccio

In the physical applications we consider both the limit $\gamma, \epsilon \ll 1$ and we neglect the terms $O(\gamma\epsilon)$. In the action-angle variables assuming $H_1(I,\theta) = \sqrt{2I} f(\theta)$ we have

$$d\theta = \frac{\partial H_0}{\partial I} dt + \epsilon \frac{\partial \hat{H}}{\partial I} + \sqrt{\frac{T\gamma}{2I}} f(\theta) \circ dw_t - \frac{\gamma}{2} f(\theta) \frac{\partial f}{\partial \theta} \frac{\partial H_0}{\partial I} dt$$
$$dI = -\epsilon \frac{\partial \hat{H}}{\partial \theta} - \sqrt{2T\gamma I} \frac{\partial f}{\partial \theta} \circ dw_t + \gamma I \left(\frac{\partial f}{\partial \theta} \right)^2 \frac{\partial H_0}{\partial I} dt$$

5.1 Adiabatic Transformation in a Thermal bath

We consider a stochastic 1D Hamiltonian with a slowly modulated parameter $\lambda = \epsilon t$

$$H(x,t) = H_0(x, \lambda) \tag{5.2}$$

where $x = (q,p)$. In the case of physical system

$$H_0 = \frac{p^2}{2m} + V(q,\lambda)$$

the choice $H_1 = -\sqrt{m} q$ defines the usual thermal bath for a Boltzmann gas

$$dp = -\frac{\partial V}{\partial q} dt + \sqrt{2T\gamma m} dw_t - \gamma p \, dt$$

The arbitrariness in the choice of H_1 understands that the same MB equilibrium highlights to the universality of the MB distribution that allows to describe the equilibrium stochastic properties of a physical system whose elementary compo0nets perform different type of non-linear interactions simulated by the H_1 Hamiltonian. The FP equation for the density function $\rho(x,t)$ associated to the stochastic dynamics is computed using the Stratonovich interpretation

$$\frac{\partial \rho}{\partial t} = -D_{H_0}\rho + \gamma D_{H_1}\{H_0, H_1\}\rho + T\gamma D^2_{H_1}\rho = \mathcal{L}_{FP}\rho(x,\lambda) \tag{5.3}$$

The stationary solution can be computed by setting $\rho_{eq}(x) = \rho_{eq}(H_0(x))$ and we get the condition

$$\{H_0, H_1\}\rho_{eq}(H_0) + T D_{H_1}\rho_{eq}(H_0) = \{H_0, H_1\} \left[\rho_{eq}(H_0) + T \frac{d\rho_{eq}}{dH_0} \right] = 0$$

and we recover the MB distribution. Using the adjoint operator

$$\mathcal{L}^\dagger_{FP} = D_{H_0} - \gamma\{H_0, H_1\} D_{H_1} + T\gamma D^2_{H_1}$$

(we use the property $D^\dagger_H = -D_H$) one computes the evolution of any average observable $\langle K(x(t;w)) \rangle$ where $x = x(0; w(0))$

$$\frac{\partial K}{\partial t} = \mathcal{L}^\dagger_{FP} K = D_{H_0} K - \gamma\{H_0, H_1\} D_{H_1} K + T\gamma D^2_{H_1} K$$

The unperturbed energy evolves according to

$$dH_0 = \sqrt{2T\gamma} D_{H_1} H_0 dw_t - \gamma\{H_0, H_1\} D_{H_1} H_0 \, dt$$

and we get the fluctuation-dissipation relation

$$\langle dH_0^2 \rangle = 2T\gamma \left(D_{H_1} H_0\right)^2 = 2T\gamma(\{H_0, H_1\})^2 = -2T\langle dH_0 \rangle$$

If we define the current density

$$\mathcal{J} = -\{H_0, H_1\}\rho - T\{\rho, H_1\} \tag{5.4}$$

the FP has the continuity form

$$\frac{\partial \rho}{\partial t} = -D_{H_0}\rho - \gamma D_{H_1} \mathcal{J}$$

In the adiabatic limit $\epsilon \to 0$ and the frozen system solves the stochastic equation (**??**) for a given realization $\xi(t)$ by $x_\xi(x_0, t; \lambda)$ using λ as a parameter. The stochastic process $x_w(x_0, t; \lambda)$ relaxes to a stationary process for

$$\lim_{t\to\infty} x_w(x_0, t; \lambda) = x_w^\infty(\lambda) \qquad \forall\, x_0$$

with an exponential convergence in a weak sense. Then we call the adiabatic approximation of eq. (**??**) the solution $x_\xi^{ad}(x_0, t) = x_\xi^\infty(x_0, t; \epsilon t)$. We consider the evolution for a long time scale T such that $\epsilon T \to \lambda_1 - \lambda_0$ and the problem of adiabatic theory is to estimate the error of the adiabatic approximation with respect to the true solution $\lim_{\epsilon\to 0} x_w(x_0, t)$ when $t \in [0, T]$. The initial distribution of x_0 is assumed as the stationary distribution at the initial value of the parameter λ_0

5.1.1 Thermodynamics interpretation

The stochastic dynamics (**??**) describes the evolution of a statistical system where the particle interactions is described by the stochastic dynamics during a isothermal transformation. We consider the physical system

$$H_0(q, p, \lambda) = \frac{p^2}{2} + V(q, \lambda)$$

where the stochastic terms are associated to the Hamiltonian $H_1 = -\sqrt{2}q$. The frozen equilibrium distribution reads

$$\rho_{eq}(x, \lambda) = A(T, \lambda) \exp\left(-\frac{H_0(x, \lambda)}{T}\right) \qquad A(T, \lambda) = \int \exp\left(-\frac{H_0(x, \lambda)}{T}\right) dx$$

and we the temperature is defined as the expectation value of the kinetic energy

$$T = \left\langle \frac{p^2}{2} \right\rangle = \sqrt{\frac{1}{2\pi T}} \int \exp\left(-\frac{p^2}{2T}\right) dp$$

Then T does not depends on λ, whereas the average internal energy $E(T, \lambda) = \langle H_0(x, \lambda) \rangle$ follows from the definition of the Helmholtz Free Energy

$$F(T, \lambda) = -T \ln A(T, \lambda)$$

and using $\beta = T^{-1}$ we have

$$\langle H_0(x,\lambda)\rangle = -\frac{\partial}{\partial\beta}\ln A(\beta,\lambda) = \frac{1}{A(\lambda,\beta)}\int H_0(x,\lambda)\exp\left(-\beta H_0(x,\lambda)\right)\,dx$$

Analogously we compute

$$\frac{\partial}{\partial\lambda}\ln A(\beta,\lambda) = \frac{\beta}{A(\lambda,\beta)}\int \frac{\partial V}{\partial\lambda}(x,\lambda)\exp\left(-\beta H_0(x,\lambda)\right)\,dx = \beta X(\beta,\lambda)$$

where we define the generalized force X acting on the system to recover the equilibrium state when the parameter λ varies. We recover the relation

$$d\ln A = E(\beta,\lambda)\,d\beta - \beta^{-1}X(\beta,\lambda)\,d\lambda = d(E(\beta,\lambda)\beta) + \beta\,dE(\beta,\lambda) + \beta X(\beta,\lambda)\,d\lambda$$

so that

$$dE = T\,d(\ln(A(\beta,\lambda) - E(\beta,\lambda)\beta) - X(\beta\lambda)\,d\lambda$$

Recalling the first principle of thermodynamics

$$dE = T\,dS - X\,d\lambda$$

we get

$$S = \ln(A(\beta,\lambda) - \frac{1}{T}E(\beta,\lambda) \quad\Rightarrow\quad -T\ln(A(\beta,\lambda) = E - TS = F$$

using the definition of Free Energy and it follows

$$F(\lambda_1)-F(\lambda_0) = -\int_{\lambda_0}^{\lambda_1} X(\beta,\lambda)d\lambda = -\int_{\lambda_0}^{\lambda_1}\left\langle\frac{\partial V}{\partial\lambda}(x,\lambda)\right\rangle d\lambda = \langle V(x,\lambda_1)\rangle - \langle V(x,\lambda_0)\rangle$$

that is the works performed by the system. We remark that the previous calculations require that the adiabatic distribution describes the system evolution: i.e. the system crosses maximum entropy equilibrium states and the adiabatic distribution is an approximate solution of the FP equation.
In the adiabatic approximation the stationary distribution satisfies the DB condition $\mathcal{J}(x,\lambda) = 0$ we define the adiabatic distribution

$$\rho_{ad}(x,\epsilon t) = A^{-1}(\lambda,T)\exp\left(-\frac{H_0(x,\epsilon t)}{T}\right) \tag{5.5}$$

Remark: the condition $\mathcal{J}(x,\lambda) = 0$ is different from the usual DB condition for stochastic systems since the current density is a scalar and the condition means that

$$-\{H_0,H_1\} - T\{\ln\rho, H_1\} = 0$$

and it defines $\ln\rho$ uniquely if $\{H_0,H_1\}\neq 0$ since ρ is a first integral for H_0.
If we compute the entropy production using the solution of the FP equation (**??**)

$$\begin{aligned}\frac{dS}{dt} &= \int\left[D_{H_0}\rho + \frac{\gamma}{2}D_{H_1}\mathcal{J}\right]\ln\rho\,dx = -\frac{\gamma}{2}\int\frac{\mathcal{J}}{\rho}D_{H_1}\rho\,dx \\ &= \frac{\gamma}{2T}\int\frac{\mathcal{J}}{\rho}(\mathcal{J} + \{H_0,H_1\}\rho)\,dx = \frac{\gamma}{2T}\int\frac{\mathcal{J}^2}{\rho}\,dx + \frac{\gamma}{2T}\int\mathcal{J}\{H_0,H_1\}\,dx\end{aligned}$$

The parameter $\gamma/2T$ is a scaling factor in the entropy production and we distinguish the total entropy production and the dissipated work per unit time

$$\frac{dS^{tot}}{dt} = \frac{\gamma}{2T}\int \frac{\mathcal{J}^2}{\rho}\,dx \qquad W^{dis} = -\frac{\gamma}{2T}\int \mathcal{J}\{H_0, H_1\}\,dx$$

Since $dS^{tot/dt} \geq 0$ when the density current is $\neq 0$ (i.e. the transformation is not adiabatic), the dissipated work is always non negative and corresponds to the work performed by the current density which is dissipated and it does not contribute to the Free Energy change. The adiabatic condition $\mathcal{J} = 0$ during the transformation $\lambda_1 \to \lambda_2$ means that no work performed on (or by) the system is dissipated and it contributes to the change of the Free Energy. Due to the definition of thermal bath, $\mathcal{J} = 0$ in the equilibrium state and the total entropy production vanishes since it corresponds to a maximum entropy state. The condition can be also related to the Minimum Entropy Production Principle to characterize the equilibrium state.
In a general framework we consider the distribution function evolves according to a FP equation

$$\frac{\partial\rho}{\partial t}(x,t) = -\mathcal{L}_{FP}(x,\lambda)\rho(x,t)$$

where λ is considered a parameter in the frozen system. We evaluate the error of the adiabatic solution, by a perturbative approach

$$\rho(x,t) = \rho_{ad}(x,\lambda) + \epsilon\hat{\rho}(x,t;\lambda) + O(\epsilon^2) \qquad \lambda = \epsilon t$$

One has to prove that $\hat{\rho}$ is bounded on a time interval $\propto \epsilon^{-1}$ so that the error of the adiabatic approximation remains $O(\epsilon)$. A direct substitution in the Fokker-Planck equation provides

$$\frac{\partial\hat{\rho}}{\partial t} + \mathcal{L}_{FP}(x,\lambda)\hat{\rho} = -A^{-1}(\lambda)\left(\frac{1}{A(\lambda)}\frac{\partial A}{\partial\lambda} + \frac{1}{T}\frac{\partial H}{\partial\lambda}\right)\exp\left(-\frac{H(x,\lambda)}{T}\right)$$

with

$$\frac{\partial A}{\partial\lambda} = -\frac{1}{T}\int\frac{\partial H}{\partial\lambda}\exp\left(-\frac{H(x,\lambda)}{T}\right)dx = -\frac{A}{T}\left\langle\frac{\partial H}{\partial\lambda}\right\rangle$$

The average with respect to all the possible realizations of $\partial H/\partial\lambda$ in the adiabatic approximation, is the work performed on the system so that the term

$$\frac{\partial H}{\partial\lambda} - \left\langle\frac{\partial H}{\partial\lambda}\right\rangle$$

represents the difference between the adiabatic average work and the actual work due to the finite modulation velocity ϵ weighting each state as if it is at the equilibrium. Then the previous equation can be written in the form

$$\frac{\partial\hat{\rho}}{\partial t} - \mathcal{L}_{FP}(x,\lambda)\hat{\rho} = \frac{1}{T}\left(\frac{\partial H}{\partial\lambda} - \left\langle\frac{\partial H}{\partial\lambda}\right\rangle\right)\rho_{ad}(x,\lambda) + O(\epsilon) \qquad (5.6)$$

The solution satisfies

$$\int\hat{\rho}(x,\lambda)dx = 0$$

since

$$\int \left(\frac{\partial H}{\partial \lambda} - \left\langle \frac{\partial H}{\partial \lambda} \right\rangle \right) dx = 0$$

and the equation (**??**) is well defined since the r.h.s. has no component in the kernel of the operator $\mathcal{L}_{FP}$. All the eigenvalues of the operator $\mathcal{L}_{FP}$ has negative real part and one assumes that the fluctuations $\partial H/\partial\lambda - \langle \partial H/\partial\lambda \rangle$ are bounded for $\lambda \in [\lambda_a, \lambda_b]$. The adiabatic solution of the equation reads (we set $\hat{\rho} = 0$ for $t = 0$)

$$\begin{aligned}
&\hat{\rho}_{ad}(x,t) = \frac{1}{T}\int_0^t \exp(\mathcal{L}_{FP}(t-s)) \left(\frac{\partial H}{\partial \lambda}(x,\lambda) - \left\langle \frac{\partial H}{\partial \lambda} \right\rangle (\lambda) \right) \rho_{ad}(x,\lambda) ds \\
&= (I - \exp(\mathcal{L}_{FP}t))\mathcal{L}_{FP}^{-1} \left[\frac{1}{T} \left(\frac{\partial H}{\partial \lambda}(x,\lambda) - \left\langle \frac{\partial H}{\partial \lambda} \right\rangle (\lambda) \right) \rho_{ad}(x,\lambda) \right] \\
&\lambda = \epsilon t \qquad t = \frac{\lambda_b - \lambda_a}{\epsilon}
\end{aligned}$$

A direct calculation provides

$$\begin{aligned}
&\frac{\partial \hat{\rho}_{ad}}{\partial t} = -\exp(\mathcal{L}_{FP}t)) \left[\frac{1}{T} \left(\frac{\partial H}{\partial \lambda}(x,\lambda) - \left\langle \frac{\partial H}{\partial \lambda} \right\rangle (\lambda) \right) \rho_{ad}(x,\lambda) \right] + O(\epsilon) \\
&= \mathcal{L}_{FP}\hat{\rho} + \frac{1}{T} \left(\frac{\partial H}{\partial \lambda}(x,\lambda) - \left\langle \frac{\partial H}{\partial \lambda} \right\rangle (\lambda) \right) \rho_{ad}(x,\lambda) + O(\epsilon)
\end{aligned}$$

Remark: the error estimate requires to control the spectral properties of the operator

$$(I - \exp(\mathcal{L}_{FP}t))\mathcal{L}_{FP}^{-1}$$

if the quantity in the square bracket is bounded (we denote by $\mathcal{L}_{FP}^{-1}$ the pseudo-inverse of the operator). The eigenvalues of the operator $\mathcal{L}_{FP}^{-1}$ have a negative real part define the relaxation time scales of the unperturbed system and the previous estimates holds under the assumptions that the relaxation time scales are short with respect the modulation time scale ϵ^{-1} (also considering the λ dependence). In this case one has the limit

$$\lim_{t \to \epsilon^{-1}} \exp(\mathcal{L}_{FP}t) = 0$$

and the error depends on the inverse of the Fiedler eigenvalue λ_F of the operator (the eigenvalue of $\mathcal{L}_{FP}$ with the least negatove real part). If $\epsilon^{-1} \gg \lambda_F^{-1}$, the norm $\|\hat{\rho}_{ad}\|$ is bounded and the adiabatic approximation approximates the solution $\rho(x,t)$ with an error $O(\epsilon^2 t)$: we can estimate

$$\|\rho(x,t) - \rho_{ad}(x,t)\| \simeq O(\epsilon) \qquad t \in [0, \epsilon^{-1}] \tag{5.7}$$

We observe that $\rho_{ad}(x,t)$ changes by a quantity $O(1)$ during the evolution.
The current density (**??**) depends on $\hat{\rho}$

$$\mathcal{J} = -\epsilon \left[\{H_0, H_1\}\hat{\rho} + T\{\hat{\rho}, H_1\} \right]$$

We observe that the current density depends on the thermal bath realization H_1.

Given the realization $x_\xi(t)$ (i.e. a solution of the stochastic equation (**??**) for a given realization of the noise $\xi(t)$)

$$x_\xi(t) = \Phi^t_{t_a}(x_a; \xi, \lambda) \qquad \lambda = \epsilon t \quad t_{a,b} = \mu \lambda_{a,b}$$

where $\Phi^t_{t_a}(x_a; \xi, \lambda)$ is the stochastic phase flow for eq. (**??**) for the evolution from x_a at t_a to a state x at t. Considering an isotherm transformation between two states $x_a \to x_b$ in the equilibrium distribution, the change of the energy does not depends on the transformation

$$\Delta E_{ba} = \int H_0(x, \lambda_b)\rho_{ad}(x, \lambda_b)\, dx - \int H_0(x, \lambda_a)\rho_{ad}(x, \lambda_a)\, dx = \bar{E}(\lambda_b) - \bar{E}(\lambda_a)$$

If $\mu \ll 1$ is an adiabatic parameter, the system can be considered in an (almost) equilibrium states during all the transformations and we formally have

$$\begin{aligned} \frac{d\bar{E}}{d\lambda} &= \frac{d}{d\lambda} \int H_0(x, \lambda)\rho_{ad}(x, \lambda)\, dx \\ &= \int \left[\frac{\partial H_0}{\partial \lambda}(x, \lambda)\rho_{ad}(x, \lambda) + H_0(x, \lambda)\frac{\partial \rho_{ad}}{\partial \lambda}(x, \lambda) \right] dx \end{aligned} \tag{5.8}$$

The derivative represents the adiabatic evolution of the energy. Then $\Delta E_{ba} = 0$ if we perform an adiabatic isotherm distribution with λ constant. Using the partition function $A(\lambda, T)$ we define the Helmholtz Free Energy potential

$$F(\lambda) = -T \ln A(T, \lambda) = -T \ln \int \exp\left(-\frac{H_0(x, \lambda)}{T} \right) dx$$

which satisfies the relation

$$dF = dE - TdS$$

By definition we have

$$\frac{dF}{d\lambda} = \int \frac{\partial H_0}{\partial \lambda}(x, \lambda)\rho_{ad}(x, \lambda) dx = -\frac{\delta W}{d\lambda} \tag{5.9}$$

that can be defined as the performed work δW by an infinitesimal change $d\lambda$ in an adiabatic transformation. Then we have

$$\Delta F_{ba} = \int_{\lambda_a}^{\lambda_b} \int \frac{\partial H_0}{\partial \lambda}(x, \lambda)\rho_{ad}(x, \lambda)\, dx\, d\lambda = -\Delta W$$

In the adiabatic approximation the change of internal energy $dE = dF + TdS$ and we have

$$T\frac{dS}{d\lambda} = \int H_0(x, \lambda)\frac{\partial \rho_{ad}}{\partial \lambda}(x, \lambda)\, dx$$

that can be identified to the heat exchanged varying λ for an adiabatic transformation. This is consistent we the definition of Gibbs entropy

$$S = -\int \rho_{ad}(x, \lambda) \ln \rho_{ad}(x, \lambda)\, dx$$

since we compute

$$T\frac{dS}{d\lambda} = -T\int \frac{\partial \rho_{ad}}{\partial \lambda}\ln \rho_{ad}(x,\lambda)\,dx = \int \frac{\partial \rho_{ad}}{\partial \lambda}\left(H_0(x,\lambda) + T\ln A(\lambda,T)\right)\,dx =$$
$$= \int H_0(x,\lambda)\frac{\partial \rho_{ad}}{\partial \lambda}\,dx = \int H_0(x,\lambda)\frac{\partial \ln \rho_{ad}}{\partial \lambda}\rho_{ad}(x,\lambda)\,dx$$

using the equality

$$\int \frac{\partial \rho_{ad}}{\partial \lambda}\,dx = 0$$

We get

$$\frac{dE}{d\lambda} = \frac{dF}{d\lambda} + T\frac{dS}{d\lambda} \tag{5.10}$$

This is equivalent to the first principle of thermodynamics where the exchanged heat is performed at constant temperature and it is identified by $\delta Q = T\,dS$ in the adiabatic limit $\mu \to 0$. We observe that the exchanged heat depends on the entropy change which is due to the change of the distribution $\rho_{ad}(x,\lambda)$ when λ varies. Then if one computes the change of energy for the initial equilibrium distribution $\rho_{ad}(x_a,\lambda_a)$ we get the average performed work.
When μ is finite and we perform a isotherm transformation, we generalize previous definition by computing the work and the exchanged heat for a single realization $x_\xi(t)$

$$\begin{aligned} W_{ba}[x] &= -\int_{\lambda_a}^{\lambda_b} \frac{\partial H_0}{\partial \lambda}(\Phi_{t_a}^{\lambda/\mu}(x_a;\xi,\lambda),\lambda)\,d\lambda \\ Q_{ba}[x] &= \int_{\lambda_a}^{\lambda_b} H_0(\Phi_{t_a}^{\lambda/\mu}(x_a;\xi,\lambda),\lambda)\frac{\partial \ln \rho_{ad}}{\partial \lambda}(\Phi_{t_a}^{\lambda/\mu}(x_a;\xi,\lambda),\lambda)\,d\lambda \end{aligned} \tag{5.11}$$

where W_{ba} is positive if the system performs work on the environment. We observe that one has to consider the limit of the stochastic phase flow $\Phi_{t_a}^{\lambda/\mu}(x_a;\xi,\lambda)$ in the slow time $\lambda = \mu t$ when $\mu \to 0$ and $t \to \infty$, to prove that the adiabatic evolution of the system is described by the distribution $\rho_{ad}(x,\lambda)$.
The second principle of thermodynamics means that if μ is finite

$$\langle W_{ba}[x]\rangle = -\int_{\lambda_a}^{\lambda_b}\left\langle \frac{\partial H_0}{\partial \lambda}(\Phi_{\lambda_a}^{t/\mu}(x_a;\xi,\lambda),\lambda)\right\rangle d\lambda \le -\Delta F_{ba} \tag{5.12}$$

In the adiabatic limit $\mu \to 0$, then $x_\xi(t,\lambda)$ is a realization in the equilibrium distribution of the stochastic process at a fixed value of the parameter λ and the exchanged work read

$$\langle W_{ba}\rangle = -\int_{\lambda_a}^{\lambda_b}\int \frac{\partial H_0}{\partial \lambda}(x,\lambda)\rho_{ad}(x,\lambda)\,dx\,d\lambda = T\int_{\lambda_a}^{\lambda_b}\frac{\partial}{\partial \lambda}\ln A(\lambda,T)\,d\lambda = -\Delta_{ba}F$$

according to the thermodynamics formalism. We recall the relation

$$\left\langle \frac{\partial H_0}{\partial \lambda}(\Phi_{t_a}^{t}(x_a;\xi),\lambda)\right\rangle = \int \frac{\partial H_0}{\partial \lambda}(x,\lambda)\rho(x,t;\lambda)\,dx$$

is the evolution of the observable $\partial H_0/\partial\lambda(x,\mu t)$ where $\rho_\lambda(x,t)$ is the evolution of the distribution function at time t when $\lambda = \mu t$. In the adiabatic approximation $\rho(x,t;\lambda) = \rho_{ad}(x,\mu t)$.

Remark: in the adiabatic approximation $\mu \to 0$ the change of λ is so slow that the system can be considered in the equilibrium condition $\rho_a(x,\lambda)$ where the current probability vanishes and there is no entropy production. When μ is finite, the system crosses non-equilibrium states which contribute to tottal entropy production through the dissipation process and the performed work is reduced.

We define the work performed by the system along the realization of an isotherm transformation $\Phi_{t_a}^{t_b}(x_a;\xi,\lambda)$ (we set $\lambda_{a,b} = \mu t_{a,b}$) that connects two equilibrium states by

$$
\begin{aligned}
-W_{ba}[x_\xi] = H_0(\Phi_{t_a}^{t_b}(x_a;\xi,\lambda)) - T\ln(\rho_{ad}(\Phi_{t_a}^{t_b}(x_a;\xi,\lambda),\lambda_b) \\
- H_0(x_a,\lambda_a) + T\ln\rho_{ad}(x_a,\lambda_a)
\end{aligned}
$$

where the initial state x_a is distributed according to the equilibrium distribution with $\lambda = \lambda_a$. We compute

$$
\begin{aligned}
&\left\langle \exp\left(\frac{W_{ba}[x]}{T}\right)\right\rangle = \\
&= \left\langle \exp\left(-\frac{H_0(\Phi_{t_a}^{t_b}(x_a;\xi,\lambda)) - H_0(x_a,\lambda_a)}{T} - \ln(\rho_{ad}(\Phi_{t_a}^{t_b}(x_a;\xi,\lambda),\lambda_b) + \ln\rho_{ad}(x_a,\lambda_a)\right)\right\rangle \\
&= \frac{1}{A(\lambda_a)}\left\langle \frac{\exp\left(-H_0(\Phi_{t_a}^{t_b}(x_a;\xi,\lambda),\lambda_b)/T\right)}{\rho_{ad}(\Phi_{t_a}^{t_b}(x_a;\xi,\lambda),\lambda_b)}\right\rangle = \int \exp\left(-\frac{H_0(x,\lambda_b)}{T}\right)dx = \frac{A(\lambda_b)}{A(\lambda_a)}
\end{aligned}
$$

where we use the assumption that $x = \Phi_{t_a}^{t_b}(x_a;\xi,\lambda)$ can be considered distributed at equilibrium: this means that when the modulation $\lambda = \mu t$ ends at $t = t_b$ the system has to relax to the equilibrium state for fixed value λ_b. The relaxation process takes place at a fixed temperature T (i.e. there is a coupling with the environment) and there is an increase of entropy during that corresponds to a work dissipation performed by the system. This is at the base of the Second Principle of Thermodynamics where the arrow of time (i.e. the relaxation process) is related to the Maximum Entropy Principle for physical systems, taking into account the preservation of energy. In other words, the transformation is assumed to connect equilibrium states of the system in a non-reversible way.

Therefore we get the Jarzinski's relation

$$
-T\ln\left\langle \exp\left(\frac{W_{ba}[x])}{T}\right)\right\rangle = \Delta F_{ba} \tag{5.13}
$$

that holds for a non-adiabatic transformation between equilibrium states.

The equation (**??**) means that if one measures the performed work and the exchanged heat in many transformations a takes the average value, one gets a measure of the free energy ΔF_{ba}. Moreover the Jensen inequality implies

$$
\exp\left(\frac{\langle W_{ba}[x]\rangle}{T}\right) \leq \left\langle \exp\left(-\frac{W_{ba}[x]}{T}\right)\right\rangle = \exp\left(-\frac{\Delta F_{ba}}{T}\right)
$$

It follows

$$
\langle W_{ba}[x]\rangle \leq -\Delta F_{ba} \tag{5.14}
$$

This is equivalent to the formulation of the Second Principle of Thermodynamics for isotherm transformation where the work performed by the system is always less or equal to $-\Delta F_{ba}$.
Performing the white noise limit, we compute the change of the energy during an infinitesimal transformation at the state x we have

$$\begin{aligned} dH_0(x,\lambda) &= \sqrt{2T\gamma} D_{H_1} H_0(x,\lambda) \circ dw_t - \gamma (D_{H_1} H_0)^2(x,\lambda)\, dt + \mu \frac{\partial H_0}{\partial \lambda}(x,\lambda)\, dt \\ &= \sqrt{2T\gamma} D_{H_1} H_0(x,\lambda) dw_t + T\gamma D^2_{H_1} H_0(x,\lambda) \\ &- \gamma (D_{H_1} H_0)^2(x,\lambda)\, dt + \mu \frac{\partial H_0}{\partial \lambda}(x,\lambda)\, dt \end{aligned}$$

that implies the backward equation if the average on the noise realizations

$$\frac{\partial \bar{H}_0}{\partial t} = \gamma D^2_{H_1} \bar{H}_0(x,\lambda) - \gamma (D_{H_1} \bar{H}_0)^2(x,\lambda) + \mu \frac{\partial \bar{H}_0}{\partial \lambda}(x,\lambda) \tag{5.15}$$

where $\bar{H}_0 = \langle H_0 \rangle$ where x is the initial state. For a fixed value λ the backward equation (**??**) shows that the energy $\bar{H}_0$ evolves towards a constant value (the null eigenvector of the operator), which correspond to the equilibrium energy, but the relaxation depends on the spectrum of the operator. When λ is modulated the change of the energy can be $O(1)$, but it does not depend on the type of transformation we consider but only on the initial and final state. By varying the realization $\xi(t)$ we are considering all the possible final states at λ for an isotherm transformation, and in the adiabatic limit x are distributed according to a MB distribution. If we consider an instantaneous change of λ then the transformation happens without heat exchange but we have a relaxation in the final state which a increase of the total entropy and a dissipation of part of the performed work.

5.1.2 Perturbation approach for the 1 DoF case

Assuming that $E = H_0(x,\lambda)$ defines closed curves, one can introduce action-angle variables for any value of the parameter

$$\begin{aligned} I(E,\lambda) &= \oint_{H_0=E} p(q,E,\lambda) dq \\ \theta &= \left.\frac{\partial}{\partial I}\right|_q \int_{H_0=E}^q p(q,E,\lambda) dq \end{aligned} \tag{5.16}$$

Let $F(q,I,\lambda)$ the generating function

$$F(q,I,\lambda) = \int_{H_0=E}^q p(q,E(I,\lambda),\lambda) dq$$

We perform the change of variables (**??**) on the Hamiltonian system (**??**) and the new Hamiltonian reads

$$H_0(I,\lambda) + \mu \left.\frac{\partial F}{\partial \lambda}\right|_{(q,I)} (\theta, I, \lambda) \tag{5.17}$$

In the 1DoF cases the adiabatic theory uses the relation

$$\left.\frac{\partial F}{\partial \lambda}\right|_{q,I} = -\frac{1}{\Omega(E,\lambda)}\int^{\theta}\left(\left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p} - \left\langle \left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p}\right\rangle\right) d\theta \tag{5.18}$$

where $\Omega(E\,\lambda) \simeq O(1)$ is the frequency of the frozen dynamics $\mu = 0$. The relation points out as the θ-mean value of $\partial F/\partial\lambda$ vanishes. This remark is at the basis of the adiabatic invariance of the action variable I since the only resonance condition is at the separatrix where $\Omega \to 0$. A direct calculation shows that

$$\begin{aligned} I(t+T) - I(t) &= \frac{\mu}{\Omega}\int_t^{t+T}\left(\left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p} - \left\langle \left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p}\right\rangle\right) ds \\ &= \frac{\mu}{\Omega^2}\int_0^{2\pi}\left(\left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p} - \left\langle \left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p}\right\rangle\right) d\theta \simeq O\left(\frac{\mu^2}{\Omega^2}\right) \end{aligned}$$

where we explicitly use the estimate $\Delta I \simeq O(\mu)$ inside the integral. Given a time of order μ^{-1} one can apply the previous estimate fn times with $n = (T\mu)^{-1}$ and we get the adiabatic invariance estimate for the action variables

$$I(t+\mu^{-1}) - I(t) \simeq O\left(\frac{\mu}{\Omega^2}\right)$$

We observe that the ratio μ/Ω defines the adiabatic parameter since it compares the evolution time scale with the modulation time scale. The use of a perturbation approach allows to introduce improved adiabatic invariant by performing a change of variables $(\theta, I) \to (\phi, J)$ in the Hamiltonian (**??**). This can be performed by a generating function

$$G(\theta, J) = \theta J + \mu G_1(\theta, J, \lambda)$$

where the first perturbation order G_1 satisfies the homological equation

$$\Omega(J,\lambda)\frac{\partial G_1}{\partial \theta} + \mu\frac{\partial F}{\partial \lambda} = 0$$

The solution $G_1(\theta, J)$ exists thanks to the relation (**??**) and $G_1(J,\theta)$ describes the oscillation of the action variable I. The average value of $G_1(\theta, J)$ is not determined. The new Hamiltonian reads

$$H(J,\phi,\lambda) = H_0(J,\lambda) + O\left(\frac{\mu^2}{\Omega^2}\right)$$

It is possible to prove that the new action $J(q,p,\lambda)$ varies by a quantity $O(\mu^2)$ in a time $O(\mu^{-1})$ and it is an improved adiabatic invariant.

5.2 Application to physical systems

We consider the effect of a thermal bath assuming an initial Hamiltonian of a physical system

$$H_0(q,p,\lambda) = \frac{p^2}{2m} + V(q,\lambda) \tag{5.19}$$

and the stochastic dynamics (**??**)

$$\dot{x} = D_{H_0}x - \gamma\{H_0, H_1\}D_{H_1}x + \sqrt{2T\gamma}\xi(t)D_{H_1}x$$

where we choose $H_1 = -\sqrt{m}q$ (remark: H_1 has not the dimension of an Hamiltonian) to get the usual form of the thermal bath

$$\begin{aligned}\dot{q} &= p \\ \dot{p} &= -\frac{\partial V}{\partial q}(q,\lambda) - \gamma p + \sqrt{2T\gamma m}\xi(t)\end{aligned} \tag{5.20}$$

where we set $m = 1$ that is always possible scaling the potential and the temperature. The physical system is an ensemble of particles under the effect of a potential $V(q,\lambda)$ that interact by physical collisions satisfying the assumption of molecular chaos so that the Central Limit Theorem can be applied. The collision effect is described by the white noise and the Einstein fluctuation-dissipation relation. In general the fluctuation $\xi(t)$ is a vector of i.i.d. iMarkov processes. This allows a statistical interpretation of an adiabatic transformation $\lambda_a \to \lambda_b$ with $\lambda = \mu t$ $\mu \ll 1$. For a fixed λ the system relaxes to an equilibrium state

$$\rho_{ad}(q,p,\lambda) = A^{-1}(T,\lambda)\exp\left(-\frac{1}{T}\left(\frac{p^2}{2m} + V(q,\lambda)\right)\right) \tag{5.21}$$

where $A(T,\lambda)$ is the partition function

$$\begin{aligned}A(T,\lambda) &= \int \exp\left(-\frac{H_0(x,\lambda)}{T}\right)dx = \int \exp\left(-\frac{p^2}{2Tm}\right)dp\exp\left(-\frac{V(q,\lambda}{T}\right)dq \\ &= (\sqrt{2\pi T})^d A_q(T,\lambda)\end{aligned}$$

where d is the number of DoF. $A_q(T,\lambda)$ is well define if the potential confines the particles $V(q,\lambda) \to \infty$ for $|q| \to \infty$ but to compute the explicit dependence on the parameters can be difficult.
Remark: the relaxation process follows from the Markov property of the fluctuations that in the white noise limit describes the evolution of the distribution function using the FP equation

$$\frac{\partial \rho}{\partial t} = -\left[\frac{p}{m}\frac{\partial \rho}{\partial q} - \frac{\partial V}{\partial q}\frac{\partial \rho}{\partial p}\right] + \gamma\frac{\partial}{\partial p}p\rho + \gamma T m\frac{\partial^2 \rho}{\partial p^2} \tag{5.22}$$

and the relaxation process depends on the spectral property of the FP operator. The operator cannot be separated so that the solution $\rho(q,p,t)$ cannot be written as a product of two distributions in the momentum and in the configuration. This means that during the relaxation process the momentum and the coordinate evolution are connected through the eigenvectors of the operator. The factorization of the stationary distribution (i.e. the null eigenvector) suggests that there can be a limit for which the evolution time scales of the coordinates and momentum are separated and one can an averaging principle to reduce the dynamics dimensionality. This is the case when $\gamma \gg 1$ where the relaxation time of the momentum tends $\to 0$ and spectrum of the operator can be separated in two clusters: one cluster tends to ∞ related to the momentum relaxation whereas the other remains finite and describe the coordinates relaxation.

We remark that in the considered case we factorize the entropy into the sum of the kinetic entropy plus the configurational entropy

$$-\int \rho_{ad}(x,\lambda)\ln\rho_{ad}(x,\lambda)\,dx = \frac{d}{2}\ln T + \langle V(q,\lambda\rangle + d\frac{T}{2} + A_q(T,\lambda) = S_p(T) + S_q(T,\lambda)$$

where we define

$$\left\langle \frac{p^2}{2}\right\rangle = \int \frac{p^2}{2}\rho_{ad}(q,p,\lambda)\,dp\,dq = \frac{T}{2}$$

in the equilibrium state (this is the 'thermal energy' that remain constant during an isotherm transformation). This is the thermodynamics definition of temperature as the average value of the kinetic energy. The adiabatic distribution is factorized into the momentum Gaussian distribution and the coordinate distribution

$$\rho_{ad}(q,\lambda) = A_q^{-1}(T,\lambda)\exp\left(-\frac{V(q,\lambda)}{T}\right) \qquad A_q(T,\lambda) = \int \exp\left(-\frac{V(q,\lambda)}{T}\right)\,dq$$

where $A_q(\lambda)$ is the partition function that allows to study phase transition in multidimensional system. We recall that the introduction of a thermal bath describe the particle interaction under the assumption of molecular chaos so that the dynamics of each particle can be considered independent, whereas the potential $V(q)$ is an external potential for each particle depending only on its coordinates. The average field approach assumes the possibility of describing the particle interaction not included in the instantaneous collisions using a self-consistent external field, but this is not justified in many cases. The temperature T enters in the definition of $\rho_{ad}(q)$ but it is not the average value of the potential energy

$$\langle V(q)\rangle = A_q^{-1}(T)\int V(q)\exp\left(-\beta V(q)\right)\,dq = -\frac{\partial}{\partial\beta}\ln A(T,\lambda) \qquad \beta = T^{-1}$$

The exception is the quadratic potential $V(q) = \omega^2/2\,q^2$ when $A_q(T) = \sqrt{2\pi T}/\omega$: in case of linear force the average of kinetic energy equals the average of potential energy.

The dependence on λ introduces the possibility of performing isotherm transformations for the system for which a thermodynamics description can be applied however this requires the assumption that a quasi static description of the system is possible. The spectral properties of FP equation associated to the system (**??**) are not easily derived in the nonlinear cases. These properties would allow to study the non equilibrium and transient states of the system (this is one of the main goal of the stochastic thermodynamics that aims to characterize the thermodynamics properties of transient states). One possible approach is to distinguish between fast and slow relaxing variables and takes only the dynamics of the slow variables. The over damped limit $\gamma \to \infty$ (see section **??**) means that the momentum p are fast variables and the dynamics reduces to the Smoluchowski stochastic equation

$$dq = -\frac{\partial V}{\partial q}(q,\lambda) + \sqrt{2T}dw_\tau$$

which describes the relaxation in the configuration space, where $\tau = t/\gamma$ is a slow time. This limit implies that the spectrum of the FP operator (**??**)

divides into two parts one of which tends to ∞ ($\propto \gamma$) and corresponds to the relaxation process in the p variables, the other part tend to the real axis and remain finite. The corresponding eigenvectors can be factorized in the Gaussian distribution for the momentum and coordinate eigenvectors that describe the relaxation process in the coordinate. ****

The Fokker-Planck equation reads

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial q}\frac{\partial V}{\partial q}(q,\lambda)\rho + T\frac{\partial^2 \rho}{\partial q^2}$$

The stationary solution is $\rho_{ad}(q,\lambda)$ and the FP equation can be written in a self-adjont form: we look for solution in the form

$$\rho = \exp\left(-\frac{V}{T}\right)\hat{\phi}$$

and we get

$$\frac{\partial \hat{\phi}}{\partial t} = T\exp\left(\frac{V}{T}\right)\frac{\partial}{\partial q}\exp\left(-\frac{V}{T}\right)\frac{\partial \hat{\phi}}{\partial q}$$

The operator can be put is a self adjoint form by a change of variable

$$\hat{\phi} = \exp\left(\frac{V}{2T}\right)\phi \quad \Rightarrow \quad \rho = \exp\left(-\frac{V}{2T}\right)\phi$$

and we get the FP operator

$$T\exp\left(\frac{V}{2T}\right)\frac{\partial}{\partial q}\exp\left(-\frac{V}{T}\right)\frac{\partial}{\partial q}\exp\left(\frac{V}{2T}\right)$$

The characteristic equation

$$\eta\,\phi = T\exp\left(\frac{V}{2T}\right)\frac{\partial}{\partial q}\exp\left(-\frac{V}{T}\right)\frac{\partial}{\partial q}\exp\left(\frac{V}{2T}\right)\phi$$

defines orthogonal wave functions with negative real eigenvalues. We have a Schrödinger like operator

$$T\left[\frac{\partial}{\partial q} - \frac{1}{2T}\frac{\partial V}{\partial q}\right]\left[\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right]\phi = \eta\,\phi$$

We observe that assuming natural boundary condition

$$\begin{aligned}\eta\int \phi^2(q)\,dq &= T\int \phi\left[\frac{\partial}{\partial q} - \frac{1}{2T}\frac{\partial V}{\partial q}\right]\left[\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right]\phi\,dq \\ &= -T\int\left(\left[\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right]\phi\right)^2 dq\end{aligned}$$

so that $\eta < 0$ and we recover $\eta = 0$ when

$$\left[\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right]\phi = 0 \quad \Rightarrow \quad \phi(q) \propto \exp\left(-\frac{V}{2T}\right)$$

Remark: $\phi(q)$ can be normalized in the $\mathcal{L}^2$-norm and the general solution of the FP equation can be written in the form

$$\rho(q,t) = \rho_{ad}(q) + \sum_{\eta \neq 0} c_\eta e^{\eta t} \exp\left(-\frac{V(q)}{2T}\right) \phi_\eta(q)$$

We have also the properties

$$\int \exp\left(-\frac{V(q)}{2T}\right) \phi_\eta(q)\, dq = 0 \qquad \int \phi_\eta(q)\phi_\nu(q)\, dq = \delta_{\eta\nu}$$

and the coefficients c_η follow from the relation

$$c_\eta = \int \phi_\eta(q)\rho(q,0) \exp\left(\frac{V(q)}{2T}\right)\, dq$$

We observe that the FP operator has the form

$$\mathcal{L} = -T\mathcal{A}^\dagger \mathcal{A}$$

where

$$\mathcal{A} = \left[\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right]$$

can be interpreted as annihilation operator when the potential is linear $V(q) = \omega^2 q^2/2$. We have the commutator relations

$$[\mathcal{A}, \mathcal{A}^\dagger] = \left[\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right]\left[-\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right] - \left[-\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right]\left[\frac{\partial}{\partial q} + \frac{1}{2T}\frac{\partial V}{\partial q}\right] = \frac{\omega^2}{T}$$

The eigenvectors of the FP operator can be computed recursively starting from the stationary state $\mathcal{A}p_0(q) = 0$. If $p(q)$ is an eigenvector with eigenvalue η, we use the relation

$$\mathcal{L}\mathcal{A}^\dagger p(q) = -T\mathcal{A}^\dagger \mathcal{A}\mathcal{A}^\dagger p(q) = -T\mathcal{A}^\dagger\left[\mathcal{L} + \frac{\omega^2}{T}\right] p(q) = -\left[\nu T + \omega^2\right]\mathcal{A}^\dagger p(q)$$

Then we have the spectrum

$$\mathcal{L}\left(\mathcal{A}^\dagger\right)^n p_0(q) = -n\frac{\omega^2}{T}$$

We have a discrete real spectrum with negative eigenvalues that decreases as the linear force increases: i.e. if $\omega(\lambda) \to 0$ we have a singular behavior that has to be studied also considering if the separation between fast-slow variables is still valid. In a generic situation the possibility to introduce fast-slow variable depends on the dynamics. Using the Stratonovich interpretation, the MB distribution $\rho(x) \propto \exp(-H(x)/T)$ is a covariant object and we can use any set of canonical variables. In the case of integrable systems we have $\rho(I) \propto \exp(-H(I)/T)$ as a function of the action variables so that the angle distribution is uniform. The problem is if in the relaxation process the angle variable can be considered fast variables with respect to the action variables. In the case of linear systems we have the stochastic equations (Stratonovich form)

$$\dot\theta = \omega + \gamma \sin(2\theta) - \sqrt{\frac{T\gamma}{I\omega}} \sin(\theta)\xi(t)$$

$$\dot I = -2\gamma I \cos^2\theta + \sqrt{\frac{4IT\gamma}{\omega}} \cos\theta\xi(t)$$

If $\gamma \ll \omega$ the angle is a fast variable but the deterministic dynamics is not related to the relaxation process. However it is possible to consider the average dynamics if ω does not satisfies a resonance condition: if we compute the average on the angle variable of the stochastic equation using the slow angle $\theta = \phi + \omega t$ and introducing the slow time $\tau = \gamma t$ we get

$$\begin{aligned} \frac{d\phi}{d\tau} &= \sin(2\phi + 2\omega/\gamma\tau)1 - \sqrt{\frac{T}{I\omega}}\sin(\phi + \omega/\gamma\tau)\xi(\tau) \\ \frac{dI}{d\tau} &= -2I\cos^2(\phi + \omega/\gamma\tau) + \sqrt{\frac{4IT}{\omega}}\cos(\phi + \omega/\gamma\tau)\xi(\tau) \end{aligned}$$

We remark that we have a singularity in the angle dynamics for $I \to 0$ due to the polar variables not to the dynamics. The averaging principle assumes that if ω/γ is big, the phases in the coefficient of the action dynamics can be considered random variables extracted form the uniform distribution (i.e. the stationary distribution of the phase dynamics). This assumption is satisfies when the slow time scale evolution $\Delta\tau$ can be chosen such that $\omega/\gamma\ \Delta\tau \gg 1$. In this case the action dynamics tends to a stochastic dynamics but one has to take into account the correlation between the fluctuations in the action and angle variables since they are generated by the same $\xi(\tau)$. In particular the angle fluctuating term gives a contribution to the action dynamics

$$-\frac{T}{\omega}\left\langle \sin(\phi + \omega/\gamma\tau)\frac{\partial}{\partial\phi}\cos(\phi + \omega/\gamma\tau)\right\rangle = \frac{T}{\omega}\left\langle \sin^2(\phi + \omega/\gamma\tau)\right\rangle$$

whereas the variance of the fluctuating term reads

$$\frac{4IT}{\omega}\langle\cos^2\phi\rangle_\phi = \frac{2IT}{\omega}$$

We have the reduced average dynamics

$$\frac{dI}{d\tau} = \left(-I + \frac{T}{2\omega}\right) + \sqrt{\frac{2IT}{\omega}}\xi(\tau)$$

Remark: for a linear systems the averaging process cannot be justified by the relaxation of the slow angle ϕ (except for low values of the action$I$) but it follows form the unperturbed dynamics $\phi + \omega/\gamma\tau$ when $\gamma \to 0$. The phase dynamics could be used to check that the uniform distribution of the angle variables does not depend on the action evolution, so that it is possible to perform the phase average on the FP equation.
The average FP equation for the action distribution $\rho(I,t)$ reads

$$\frac{\partial\rho}{\partial\tau} = \frac{\partial}{\partial I}\left(I - \frac{T}{2\omega} + \frac{T}{\omega}\sqrt{I}\frac{\partial}{\partial I}\sqrt{I}\right)\rho = \frac{\partial}{\partial I}\left(I + \frac{T}{\omega}I\frac{\partial}{\partial I}\right)\rho$$

We see that the evolution of $\rho(I,t)$ is singular when $\omega \to 0$ and in the case of an adiabatic modulation $\omega(\lambda)$.
By varying the parameter λ we perform an isotherm on the system and the change of the potential energy $\langle V(q,\lambda)\rangle$ represents the amount of work performed on the system by the environment

$$\langle V(q,\lambda)\rangle = \int \frac{p^2}{2}\rho_{ad}(q,p,\lambda)\,dp\,dq$$

****

$$\frac{\partial H_0}{\partial t} = -\gamma p \frac{\partial H_0}{\partial p} + T\gamma \frac{\partial^2 H_0}{\partial p^2} = -\gamma(p^2 - T)$$

$$\frac{dS}{dt} = \int \frac{\partial \rho}{\partial t} \ln \rho \, dx = D_{H_0}\rho$$

The idea is to prove that in the adiabatic limit the system can be considered in equilibrium state (frozen system)

$$E(T,\lambda) = \langle H_0 \rangle (T,\lambda) = -\frac{\ln A(T,\lambda)}{\partial \beta}$$

that connect the partition function with the Helmholtz free energy

$$F = E - TS = \int H \rho_{ad} \, dp \, dq + T \int \rho_{ad} \ln \rho_{ad} \, dq \, dp = -T \ln A$$

We recall that for an adiabatic isotherm transformation

$$F_B - F_A = F(T, \lambda_b) - F(T, \lambda_a) = -W$$

where W is the performed work by the system. By varying λ, the change of the energy is defined by two contributions

$$\frac{\partial}{\partial \lambda} \int H(q,p,\lambda) \rho_{ad}(q,p,\lambda) \, dp \, dq =$$
$$\int \frac{\partial H}{\partial \lambda}(q,p,\lambda) \rho_{ad}(q,p,\lambda) \, dp \, dq + \int H(q,p,\lambda) \frac{\partial \rho_{ad}}{\partial \lambda}(q,p,\lambda) \, dp \, dq$$

A direct computation gives

$$\begin{aligned} -T \frac{\partial}{\partial \lambda} \int \rho_{ad} \ln \rho_{ad} \, dq \, dp &= -T \int \frac{\partial \rho_{ad}}{\partial \lambda} \ln \rho_{ad} \, dq \, dp \\ &= \int H \frac{\partial \rho_{ad}}{\partial \lambda} \, dq \, dp + T \ln A(T,\lambda) \int \frac{\partial \rho_{ad}}{\partial \lambda} \, dq \, dp \\ &= \int H \frac{\partial \rho_{ad}}{\partial \lambda} \, dq \, dp \end{aligned}$$

and we get

$$\frac{\partial E}{\partial \lambda} = \int \frac{\partial H}{\partial \lambda} \rho_{ad} \, dp \, dq + T \frac{\partial S}{\partial \lambda} \tag{5.23}$$

We indentify

$$W(T,\lambda) = -\int_{\lambda_a}^{\lambda_b} d\lambda \int \frac{\partial H}{\partial \lambda}(q,p,\lambda) \rho_{ad}(q,p,\lambda) \, dp \, dq$$

and we extend this definition to non-adiabatic isotherm transformation when $\lambda = \mu t$. More precisely we consider a solution of the stochastic equation (**??**) $x_\xi(t \,|\, x_0)$ for a given realization of the noise $\xi(t)$, $\lambda = \mu t$ and a given initial state x_0 (x_0 can be a random variable distributed according to an equilibrium state) we define

$$W_\xi(\lambda) = -\int_0^t dt \int \frac{\partial H}{\partial \lambda}(x_\xi(t \,|\, x_0), \lambda_a + \mu t) \rho_{ad}(x_0, \lambda_a) \, dx_0 \qquad \lambda = \lambda_a + \mu t$$

and $\langle W_\xi(\lambda)\rangle$ denotes the expectation value with respect all the realizations $\xi(t)$. The second principle implies for an irreversible transformation (i.e. μ is finite)

$$\langle W_\xi(\lambda)\rangle \leq -\Delta F$$

The idea is to prove that in the non-adiabatic transformation when μ finite this relation should be the consequence of the difference between the real distribution $\rho(q,p,t)$ and the adiabatic quasi-static distribution $\rho_{ad}(q,p,t)$. If we solve the stochastic equation to compute a single trajectory $x_\xi(t\,|\,x_0)$ the previous condition could be not verified.

5.3 Action-angle variables

Assuming $H_1 = -\sqrt{m}q$, we perform the action-angle variables using the covariance properties of the Stratonovich interpretation of eq. (**??**). We have to compute

$$\{H_0, H_1\} = -\Omega(I,\lambda)\frac{\partial H_1}{\partial \theta}(\theta, I, \lambda) = \sqrt{m}\left.\frac{\partial H_0}{\partial p}\right|_{q,\lambda}(\theta, I, \lambda)$$

and

$$D_{H_1} I = -\frac{\partial H_1}{\partial \theta} = \sqrt{m}\left.\frac{\partial I}{\partial p}\right|_{q,\lambda} = \frac{\sqrt{m}}{\Omega}\frac{\partial H_0}{\partial p}$$
$$D_{H_1}\theta = \frac{\partial H_1}{\partial I}$$

Remark: performing the change of variables in eq. (**??**) λ can be considered a parameter even if $\lambda = \mu t$, but in the computation of the differentials of the new variables one has to compute the variation with respect to λ. This can be performed maintaining the canonical form by changing the Hamiltonian H_0 according to

$$\hat{H}_0(\theta, I, \lambda) = H_0(I,\lambda) + \mu\left.\frac{\partial F}{\partial \lambda}\right|_{q,I}(\theta, I, \lambda)$$

where $F(q, I, \lambda)$ is the generating function for the action-angle change of variables.
The stochastic dynamics for the action-angle variables read

$$\dot{\theta} = \Omega(I,\lambda) + \mu\frac{\partial}{\partial I}\left.\frac{\partial F}{\partial \lambda}\right|_{q,I} - \gamma\Omega\frac{\partial H_1}{\partial \theta}\frac{\partial H_1}{\partial I} + \sqrt{2T\gamma}\frac{\partial H_1}{\partial I}\xi(t)$$
$$\dot{I} = -\mu\frac{\partial}{\partial \theta}\left.\frac{\partial F}{\partial \lambda}\right|_{q,I} - \gamma\Omega\left(\frac{\partial H_1}{\partial \theta}\right)^2 - \sqrt{2T\gamma}\frac{\partial H_1}{\partial \theta}\xi(t) \qquad (5.24)$$

where we perform a white noise limit for $\xi(t)$. The existence of a stochastic adiabatic invariance for the action variable is related to the possibility of averaging eq. (**??**) with respect the noise realizations and with respect to the angle variable. In the limit of white noise (Stratonovich interpretation) the average action evolution reads

$$
\begin{aligned}
\langle \Delta I \rangle = & \left[-\mu \frac{\partial}{\partial \theta} \frac{\partial F}{\partial \lambda}\bigg|_{q,I} - \gamma \Omega \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right] \Delta t \\
& - \gamma T \left[\frac{\partial^2 H_1}{\partial \theta^2} \frac{\partial H_1}{\partial I} - \frac{\partial^2 H_1}{\partial \theta \partial I} \frac{\partial H_1}{\partial \theta} \right] \Delta t + O(\Delta t^2, \mu^2 \Delta t)
\end{aligned}
$$

The last term in obtained by the stochastic fluctuations and it can be written in the form

$$
\frac{\partial^2 H_1}{\partial \theta^2} \frac{\partial H_1}{\partial I} - \frac{\partial^2 H_1}{\partial \theta \partial I} \frac{\partial H_1}{\partial \theta} = \frac{\partial}{\partial \theta} \left[\frac{\partial H_1}{\partial \theta} \frac{\partial H_1}{\partial I} \right] - \frac{\partial}{\partial I} \left(\frac{\partial H_1}{\partial \theta} \right)^2
$$

The average dynamics for the action can be written in the form

$$
\begin{aligned}
\langle \Delta I \rangle = & \left[-\mu \frac{\partial}{\partial \theta} \frac{\partial F}{\partial \lambda}\bigg|_{q,I} - \gamma \Omega \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right] \Delta t \\
& - \gamma T \left[\frac{\partial}{\partial \theta} \left(\frac{\partial H_1}{\partial \theta} \frac{\partial H_1}{\partial I} \right) - \frac{\partial}{\partial I} \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right] \Delta t + O(\Delta t^2, \mu^2 \Delta t)
\end{aligned}
$$

Then it is possible to consider the limit of the stochastic process $\bar{I}$ where we also consider the average on the angle variable since the angle is a fast variable and the uniform stationary distribution on the angle can be considered invariant during the evolution. The average stochastic process $\bar{I}$ satisfies the stochastic differential equation

$$
d\bar{I} = -\gamma \left[\Omega \left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta - T \frac{\partial}{\partial I} \left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta \right] dt + \sqrt{2T\gamma} \sqrt{\left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta} dw_t \tag{5.25}
$$

This equation holds if the uniform distribution of the angle is preserved in the evolution since the angle variable has a relaxation time faster than the action variable (this can be proven in the limit of small noise). The FP equation for $\bar{\rho}(I, t)$

$$
\begin{aligned}
\frac{\partial \bar{\rho}}{\partial t} & = \gamma \frac{\partial}{\partial I} \left[\Omega \left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta - T \frac{\partial}{\partial I} \left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta \right] \bar{\rho} + \gamma T \frac{\partial^2}{\partial I^2} \left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta \bar{\rho} \\
& = \gamma \frac{\partial}{\partial I} \left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta \left[\Omega + T \frac{\partial}{\partial I} \right] \bar{\rho}
\end{aligned} \tag{5.26}
$$

The equation (**??**) can describe the evolution of the action distribution far from the separatrix curves where $\Omega \to 0$. In the case of physical systems (cfr. eq. (**??**) and $H_1 = \sqrt{2mq}$), the diffusion coefficient can be written in the form

$$
\left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta = \frac{1}{\Omega^2} \left\langle \left(\frac{\partial H_0}{\partial p} \right)^2 \right\rangle_\theta = \frac{1}{\Omega^2} \left\langle \frac{p^2}{m} \right\rangle
$$

that highlights as we have a singularity when $\Omega \to 0$ at the separatrix. Far from the separatrix the $\bar{\rho}(I, t)$ is regular and for a fixed λ (frozen system) the equilibrium condition reads

$$\left[\Omega + T\frac{\partial}{\partial I}\right]\bar{\rho}_{eq}(I,\lambda) = 0 \quad \Rightarrow \quad \rho_{eq}(I) \propto \exp\left(-\frac{H_0}{T}\right)$$

but one should introduce the boundary condition when I increases. Moreover the angle dynamics is not fast near the separatrix and the averaging procedure is not justified.

5.4 Adiabatic character of the action distribution

If one considers the evolution of the distribution function $\rho(I,t)$ average on the angle variables we get the conjugated operator of eq. (**??**): i.e the average FP equation (see. eq. (**??**))

$$\frac{\partial\bar{\rho}}{\partial t} = \gamma\frac{\partial}{\partial I}\left\langle\left(\frac{\partial H_1}{\partial\theta}\right)^2\right\rangle_\theta\left[\Omega(I)\bar{\rho} + T\frac{\partial\bar{\rho}}{\partial I}\right] + O(\mu^2) \tag{5.27}$$

with an error $O(\mu^2)$ that we neglect if $\mu \ll \gamma$.
Remark: the dissipation γ^{-1} is a relaxation time scale so that we have to consider the relation between the adiabatic time scale μ^{-1} and the relaxation time scale. The adiabatic assumption $\mu^{-1} \gg \gamma^{-1}$ means that we have the adiabatic approximation for the evolution of the distribution function

$$\Omega\bar{\rho} + T\frac{\partial\bar{\rho}}{\partial I} \simeq 0 \quad \Rightarrow \quad \rho_{ad}(I,t) = A^{-1}(\mu t)\exp\left(-\frac{H_0(I,\mu t)}{T}\right)$$

and the partition function reads

$$A(\mu t) = \int \exp\left(-\frac{H_0(I,\mu t)}{T}\right) dI$$

In this case the distribution is independent from the initial condition and the system is in a thermal equilibrium. The adiabatic approximation assumes the quasi-stationary condition for the slowly modulated FP equation: the expected error can be computed in a perturbation way, let

$$\rho(I,t) = \rho_{ad}(I,\mu t) + \mu\Delta\rho(I,t) + O(\mu^2)$$

A direct calculation provides

$$\frac{\partial\rho}{\partial t} = \gamma\frac{\partial}{\partial I}D(I)\left[\Omega(I,\lambda)\Delta\rho + T\frac{\partial\Delta\rho}{\partial I}\right] - \mu\frac{\partial\rho_{ad}}{\partial\lambda}$$

where we set

$$D(I) = \left\langle\left(\frac{\partial H_1}{\partial\theta}\right)^2\right\rangle_\theta$$

We observe that by scaling the time we define the adiabatic parameter by the ratio μ/γ and we require the condition $\mu/\gamma \ll 1$, Using the relation

$$\frac{\partial}{\partial \lambda} A^{-1}(\lambda) \exp\left(-\frac{H_0(I,\lambda)}{T}\right) = \frac{1}{T}\left[\left.\frac{\partial H_0}{\partial \lambda}\right|_I - \int \left.\frac{\partial H_0}{\partial \lambda}\right|_I \rho_{ad}(I,\lambda)\, dI\right] \rho_{ad}(I,\lambda)$$
$$= \frac{1}{T}\left[\left.\frac{\partial H_0}{\partial \lambda}\right|_I - \left\langle \left.\frac{\partial H_0}{\partial \lambda}\right|_I \right\rangle\right] \rho_{ad}(I,\lambda)$$

we see that $\partial \rho_a / \partial \lambda$ has no component on the kernel of the FP operator. We recall the thermodynamics interpretation of the isotherm transform and we recognize in the previous formula the difference between the Helmholtz free energy (**??**) and the average work performed during a non adiabatic transformation of the action $I(t;\xi)$ (but we average on the angle variables assuming that this is the stationary condition). If we introduce the evolution operator

$$\Phi_0^t = \exp\left(t\gamma \frac{\partial}{\partial I} D(I)\left[\Omega(I,\lambda) + T\frac{\partial}{\partial I}\right]\right) = \exp\left(t\gamma \mathcal{L}_{FP}(I)\right)$$

which solve the FP equation (**??**) for fixed λ, we formally have

$$\Delta\rho(I,t) = -\mu \int_0^t \Phi_s^t\, ds \left[\frac{\partial \rho_{ad}}{\partial \lambda}(I,\lambda)\right] + O(\mu^2)$$

restricted on the invariant space where the operator is invertible. A direct calculation provides

$$\exp\left(\frac{H_0}{2T}\right) \mathcal{L}_{FP}(I) \exp\left(-\frac{H_0}{2T}\right) = T\frac{\partial}{\partial I} D \frac{\partial}{\partial I} + \frac{1}{2}\left[\frac{\partial}{\partial I} D\Omega - \Omega D \frac{\partial}{\partial I}\right] - \frac{\Omega^2}{4T} D$$

that is self-adjoint operator and eigenvalues of the FP operator $\mathcal{L}_{FP}$ are all real and negative except the null one. Then the evolution operator Φ_s^t is a contraction and we have the estimate

$$\lim_{t\to\infty} \|\Delta\rho(I,t)\| \le \lambda_{\min}^{-1} \left\|\frac{\partial \rho_{ad}}{\partial \lambda}\right\| + O(\mu t)$$

where $\lambda_{\min}$ is the Fiedler eigenvalue of the FP operator. If $\lambda_{min} = O(1)$ the previous results implies that $\Delta\rho$ is bounded $O(1)$ for a time $t \propto \mu^{-1}$ and the adiabatic stationary distribution approximates the real distribution with an error $O(\mu)$: i.e. it is an adiabatic invariant for the stochastic dynamics. This result does not hold near a sepatatrix curve and during the phenomenon of separatrix crossing the distribution changes.

When the phase space is divided into different region by the separatrix curves, the previous results apply to each region separately if the transition rate among the region is much smaller than the modulation time scale. This can be estimated by using the Kramer's Theory of transition rate.

To study the FP spectrum for fixed λ we consider solutions of the FP equation in the form (only for 1DoF case)

$$e^{\eta t} \rho_a(I) \phi_\eta(I)$$

which satisfy

$$\eta \rho_a(I)\phi_\eta(I) = \gamma \frac{\partial}{\partial I} \frac{1}{\Omega^2(I)} \left\langle \left(\frac{\partial H_0}{\partial p}\right)^2 \right\rangle \rho_a(I) T \frac{\partial \phi_\eta}{\partial I} = \gamma T \frac{\partial}{\partial I} D(I) \rho_a(I) \frac{\partial \phi_\eta}{\partial I}$$

where

$$D(I) = \left\langle \left(\frac{\partial I}{\partial p}\right)^2 \right\rangle = \left\langle \left(\frac{\partial H_0}{\partial p}\frac{\partial I}{\partial H_0}\right)^2 \right\rangle = \frac{1}{\Omega(I)^2}\left\langle \left(\frac{\partial H_0}{\partial p}\right)^2 \right\rangle$$

The eigenvalues scale as $\eta \to \eta/\gamma$ (i.e. increasing γ the relaxation time increases) and the operator

$$\rho_a^{-1}(I)\frac{\partial}{\partial I}D(I)\rho_a(I)\frac{\partial}{\partial I}$$

is self-adjoint with respect the scalar product

$$\phi \cdot \psi = \int \phi(I)\rho_a(I)\psi(I)\, dI$$

The eigenvectors $\phi_\eta(I)$ are orthogonal with respect to the scalar product and all the eigenvalues η are real negative.
Near a separatrix curve the adiabatic invariance is lost and we can observe the resonance trapping phenomenon in analogy with the deterministic dynamics. The action-angle variables are singular at the separatrix and we have different action-angle variables so that the FP equation (**??**) is a local equation in each region of the phase space. In the stochastic case the resonance trapping is rule by three different time scales: the angle relaxation time scale that is a consequence of the fast character of the angle variable due to the frequency dependence on the action $d\Omega/dI \neq 0$, the adiabatic modulation time scale ϵ^{-1} and the Kramer's time scale due to the transition of the particles between two resonance region. These time scales have to be separated in a increasing order to describe the adiabatic resonance trapping into different resonance region using the equilibrium solution of the FP equation (**??**). ********
The study of the spectral properties as a function of λ is a key issue to apply an adiabatic approximation. We recall that linearizing the dynamics near an elliptic point (linear noise approximation) the corresponding dynamics is driven by a quadratic Hamiltonian system

$$H_0(q,p;\lambda) = \sum_k \frac{p_k^2}{2} + \frac{1}{2}\sum_{hk} q_h V_{hk}(\lambda) q_k = \sum_k \omega_k(\lambda) I_k(\lambda)$$

plus a dissipative force and a stochastic term. Averaging on the angle variable we get the equation

$$\frac{\partial \rho}{\partial t} = \gamma \sum_k \frac{\partial}{\partial I_k}\omega_k^{-2}\left\langle \left(\frac{\partial H_0}{\partial p_k}\right)^2 \right\rangle \left[\omega_k \rho + T\frac{\partial \rho}{\partial I_k}\right]$$

where ω_K is the linear frequency at the elliptic point and by definition we have

$$\left\langle \left(\frac{\partial H_0}{\partial p_k}\right)^2 \right\rangle = \langle p_k^2 \rangle = \omega_k I_k$$

The averaging principle can be applied if $\vec{\omega}$ does not satisfy a resonance condition. The FP equation reads

$$\frac{\partial \rho}{\partial t} = \gamma \sum_k \frac{\partial}{\partial I_k} I_k \left[\rho + \frac{T}{\omega_k}\frac{\partial \rho}{\partial I_k}\right]$$

We observe that γ is a time scale and that the stationary solution is

$$\rho_s(I) = \frac{T}{\omega} \exp\left(-\frac{1}{T}\sum_k \omega_k I_k\right)$$

We look for the eigenvectors in the form

$$\exp\left(-\frac{1}{T}\sum_h \omega_h I_h\right) p_\eta(I)$$

where $p_\eta(I)$ is a polynomial. A direct substitution gives

$$\frac{\eta}{\gamma} p_\eta(I) = T \exp\left(\frac{1}{T}\sum_h \omega_h I_h\right) \sum_k \frac{\partial}{\partial I_k} I_k \exp\left(-\frac{1}{T}\sum_h \omega_h I_h\right) \frac{\partial p_\eta}{\partial I_k}$$

If $p_\lambda(I)$ is a polynomial of order n in the action we have the condition

$$\eta = -\gamma \sum_k n_k \omega_k \qquad \sum_k n_k = n$$

that shows as the eigenvalues are related to the linear combination of the frequencies ω_k and scale with γ. The polynomials $p_\lambda(I)$ are orthogonal with respect the scalar product associated to the invariant distribution and all the eigenvalues are real negative. The linear combination of the frequencies can be arbitrarily small for $n \gg 1$ so that λ are arbitrarily small but as the polynomial degree increase the coefficients of the expansion on the eigenvector base decrease so that the contribution of higher order polynomial can be smallw er than μ.
In the 1DoF case, the equation (**??**) is singular at a separatrix curve since $\Omega(I) \to 0$ and both the averaging procedure is not justified and the relaxation time scale diverges. This is the case when we have a non linear resonance in the phase space of $H_0(q,p;\lambda)$. However it is possible to derive an adiabatic theory assuming that $\rho(I;\mu t)$ is preserved with a small error is we are not near the separatrix and it has an absorbing boundary condition at the I_* the singular action value at the separatrix.

5.5 Stochastic adiabatic theory

The generalization of the stochastic adiabatic theory in presence of non-linear resonances in the phase space is based on the following picture (to be analytically justified). The phase space is divided into different areas where local action-angle variables can be introduced and we have a global diffusion dynamics due to a thermal bath. In the 1DoF case we have the FP equation

$$\frac{\partial \rho}{\partial t} = -D_{H_0}\rho + \gamma \frac{\partial}{\partial p}\rho + T\gamma \frac{\partial^2 \rho}{\partial p^2} \qquad H_0 = \frac{p^2}{2} + V(q,\lambda) \tag{5.28}$$

If $H_0(q,p;\mu t)$ is slowly modulated $\mu \ll 1$ the adiabatic distribution corresponds to the MB solution

$$\rho(q,p,\mu t) = A(\lambda)^{-1} \exp\left(\frac{H_0(q,p;\mu t}{T}\right)$$

Remark: the adiabatic distribution is Gaussian in the momentum p so that $\langle p^2 \rangle = 2T$ for each degree of freedom (this is the definition of temperature). However it is not possible to separate the relaxation time scale of the coordinate dynamics except in the over damped limit $\gamma \gg 1$.
We have seen that if the Hamiltonian is integrable and the frequencies $\Omega(I)$ do not satisfies resonance condition (in particular in the 1Dof case we have no separatrix in the phase space). Then if the phase space has a bifurcation phenomenon during modulation the adiabatic invariance of the distribution is nor guaranteed. In the case the phase space has a sepatatrix a possible extension of the adiabatic theory would be to use the action variable in each region of the space so that we expect that the adiabatic distribution in the action variable $\rho \propto \exp(-H_0(I,\lambda)/T)$ is almost constant for long time and to introduce the sepatarix as an absorbing boundary condition (in the action space) for the average FP equation. In the case it is possible to use the Kramer's Theory to estimate the probability to cross the sepatarix in the different region. The slow modulation effect is to move the separatrix in the phaase space so that the transition rates change. The theoretical problem is to show that if the modulation is slow the real distribution is well described by the adiabatic distribution $rho(I;\mu t)$ is each region since the relaxation time scale is faster than the Kramer's transition scale and the contribution of the action dynamics near the separatrix (where the angle are slow variables) is negligible. ***** One can introduce an improved action variable

$$I = J + \epsilon \frac{\partial G}{\partial \theta}(\theta, J, \lambda) \tag{5.29}$$

where the generating function $G(\theta, J, \lambda)$ satisfies the homological equation

$$\frac{\partial G}{\partial \theta} = \frac{1}{\Omega^2(J,\lambda)} \int^{\theta} \left(\left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p} - \left\langle \left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p} \right\rangle \right) d\theta$$

The solution is formally written as

$$G(\theta, J, \lambda) = \frac{1}{\Omega^2(J,\lambda)} \int^{\theta} \int^{\theta'} \left(\left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p} - \left\langle \left.\frac{\partial H_0}{\partial \lambda}\right|_{q,p} \right\rangle \right) d\theta' d\theta$$

where the integration defines a primitive function (we could require a further condition like that $G(\theta, J)$ vanishes at a given section in the phase space). Using the improved action variable the Hamiltonian (**??**) reads

$$H_0(J,\lambda) + \epsilon H_1(\phi, J, \lambda)\xi(t) + O(\epsilon^3)\xi(t) + O(\epsilon^4)$$

where the new angle ϕ is implicitly defined by

$$\phi = \theta + \epsilon \frac{\partial G}{\partial J}(\theta, J, \lambda)$$

We remark as all the transformation are singular as $\Omega(I,\lambda) \to 0$ that corresponds to crossing a nonlinear resonance.
The stochastic equations of motion are

$$\begin{aligned} \dot{\phi} &= \Omega_0(J,\lambda) + \epsilon \frac{\partial H_1}{\partial J}\xi(t) \\ \dot{J} &= -\epsilon \frac{\partial H_1}{\partial \phi}\xi(t) \end{aligned}$$

where we neglect the error terms. In the diffusion time scale $\tau = \epsilon^2 t$ we can describe the evolution of the improved action as a diffusion process according to the diffusion equation

$$\frac{\partial \rho}{\partial \tau} = \frac{1}{2}\frac{\partial}{\partial J} D(J,\tau)\frac{\partial \rho}{\partial J} \tag{5.30}$$

where the diffusion coefficient is computed from the expectation value

$$D(J,\lambda) = \left\langle \left(\frac{\partial H_1}{\partial \theta}\right)^2 \right\rangle_\theta \int_0^T \int_0^T \langle \xi(s)\xi(t)\rangle dt ds$$

assuming that the angle are relaxed to the uniform distribution which is invariant at the diffusion time scale. In the limit of a fast decorrelation of the noise at the diffusion time scale

$$D(J,\tau) = \left\langle \left(\frac{\partial H_1}{\partial \theta}(\theta, J, \lambda)\right)^2 \right\rangle_\theta \tag{5.31}$$

The distribution function $\rho(J,\tau)$ form the equation (**??**) differs from the distribution $\rho(I,\tau)$ by terms of order $O(\epsilon^2)$. Therefore the evolution of the action distribution at the diffusion time scale is not affected by the presence of a slow modulation $\lambda = \epsilon^2 t$ in the Hamiltonian up to an error of order $O(\epsilon)$.

5.6 Remarks on perturbation theory

The perturbation theory for Hamiltonian systems is strictly connected with the definition of Arnold-Liouville integrability

In the case of a Nekhoroshev like estimate for the diffusion coefficient

Appendix A

The two degrees of freedom

Remark: in a two dimensional case one expects perturbation expansions in action variables

$$\hat{H}(I,\theta) = \sum h_{kh} \left(\frac{I_1}{I_1^*}\right)^{|k|} \left(\frac{I_2}{I_2^*}\right)^{|h|} e^{ik\theta_1 + ih\theta_2}$$

so that a scaling of the variables $I_i' = I_i/I_i^*$ (the transformation is not canonical) changes the equations of motion according to

$$\dot{I}_i' = -\frac{1}{I_i^*}\frac{\partial \hat{H}}{\partial \theta_i}(I',\theta)$$

We consider a diffusion coefficient of the form

$$h_{jk}(I) = \delta_{jk}\sigma_j^2 D\,(I_1 + I_2)$$

that corresponds to a Laplacian operator

$$\frac{1}{2}\sum_j \frac{\partial}{\partial I_j}\sigma_j^2 D(I_1 + I_2)\frac{\partial}{\partial I_j}$$

We introduce the new actions

$$\begin{aligned} J_1 &= I_1 + I_2 \\ J_2 &= I_1 - I_2 \end{aligned}$$

and we use the Ito calculus to change variables

$$\begin{aligned} dJ_1 &= dI_1 + dI_2) = \frac{1}{2}(\sigma_1^2 + \sigma_2^2)\frac{\partial D}{\partial J_1}dt + \sqrt{D}(\sigma_1 dw_1 + \sigma_2 dw_2) \\ dJ_2 &= (dI_1 - dI_2) = \frac{1}{2}(\sigma_1^2 - \sigma_2^2)\frac{\partial D}{\partial J_1}dt + \sqrt{D}(\sigma_1 dw_1 - \sigma_2 dw_2) \end{aligned}$$

The new Laplacian operator reads

$$\begin{aligned} &\frac{\sigma_1^2 + \sigma_2^2}{2}\frac{\partial}{\partial J_1}\frac{\partial D}{\partial J_1} + \frac{\sigma_1^2 - \sigma_2^2}{2}\frac{\partial D}{\partial J_1}\frac{\partial}{\partial J_2} + \\ &\frac{\sigma_1^2 + \sigma_2^2}{2}\left(\frac{\partial^2}{\partial J_1^2}D(J_1) + D(J_1)\frac{\partial^2}{\partial J_2^2}\right) + (\sigma_1^2 - \sigma_2^2)\frac{\partial}{\partial J_1}\frac{\partial}{\partial J_2}D(J_1) \end{aligned}$$

We remark that there always exists a solution of the FP equation of the form $\rho(J,\tau) = \rho(J_1,\tau)$ such that

$$\frac{\partial \rho}{\partial \tau} = \frac{\sigma_1^2+\sigma_2^2}{2}\frac{\partial}{\partial J_1}\frac{\partial D}{\partial J_1}\rho + \frac{\sigma_1^2+\sigma_2^2}{2}\frac{\partial^2}{\partial J_1^2}D(J_1)\rho = \frac{\sigma_1^2+\sigma_2^2}{2}\frac{\partial}{\partial J_1}D(J_1)\frac{\partial}{\partial J_1}\rho$$

and that the Laplacian operator corresponds to a skew system if $\sigma_1 = \sigma_2$ (see appendix). We have also to take into account the boundary conditions that can be not invariaant with respect to the change of variables.
From a perturbative point of view, the assumption that

$$D(I) = D\left(\sigma_1 I_1 + \sigma_2 I_2\right) \qquad I_1, I_2 \geq 0$$

means that the phase space is foliated into surfaces

$$\sigma_1 I_1 + \sigma_2 I_2 = const.$$

where the diffusion dynamics is the same.

A.1 Stochastic symplectic maps

An analogous approach can be performed to study the perturbed symplectic maps of the form

$$\begin{pmatrix}\theta_{n+1}\\ I_{n+1}\end{pmatrix} = \begin{pmatrix}\theta_n + \epsilon\xi_n M_\theta(\theta_n, I_n;\epsilon)\\ I_n + \epsilon\xi_n M_I(\theta_n, I_n;\epsilon)\end{pmatrix} \circ \begin{pmatrix}\theta_n + \Omega(I_n)\\ I_n\end{pmatrix}$$

where the stochastic perturbation is a symplectic maps tangent to the identity and ξ_n are independent identical distributed random variables with zero mean value.
Remark: the random variables change any unit time so that we have a correlation inside each time interval. This means that the random perturbation cannot be described by a white $\delta-$correlated noise e one has to consider the perturbation effect.
Remark: without loss of generality we assume that the perturbation map can be represented by a Lie transformation with Hamiltonian $H_1(\theta, I)$

$$\begin{pmatrix}\theta_n + \epsilon\xi_n M_\theta(\theta_n, I_n;\epsilon)\\ I_n + \epsilon\xi_n M_I(\theta_n, I_n;\epsilon)\end{pmatrix} = \exp(\epsilon\xi_n D_{H_1}(\theta, I)) \begin{pmatrix}\theta_n\\ I_n\end{pmatrix}$$

Let R the integrable map and M_n the stochastic map, the evolution is the result of the composition of the maps

$$x_{n+1} = M_n \circ R \circ M_{n-1} \circ R..... \circ M_0 \circ R(x_0)$$

from a given initial condition and a realization of the ξ_i variables. It is convenient to perform a time dependent change of variables using the slow angle

$$\begin{aligned}\phi_n &= \theta_n - n\Omega(I)\\ J_n &= I_n\end{aligned} \tag{A.1}$$

Then $y_n = R^{-n} x_n$ and we have

$$\begin{aligned} y_{n+1} &= R^{-n-1} \circ M_n \circ R \circ M_{n-1} \circ R ..... \circ M_0 \circ R(y_0) \\ &= R^{-n-1} \circ M_n \circ R^{n+1} \circ R^{-n} \circ M_{n-1} \circ R^n .....R^{-1} \circ M_0 \circ R(y_0) \end{aligned}$$

We use the property of the Lie transformation (R^k is a canonical change of variables)

$$R^{-k} \circ \exp(\epsilon \xi_k D_{H_1}) \circ R^k = \exp(\epsilon \xi_n D_{H_1 \circ R^k})$$

where

$$H_1 \circ R^k = H_1(\phi + k\Omega(J), J)$$

The main idea to apply a diffusive approach to the stochastic dynamics is to show that an averaging principle holds on the angle variable for $\epsilon \ll 1$. In the diffusion time scale $\tau = \epsilon^2 k$ one can perform a continuous limit $\epsilon = \sqrt{\Delta\tau}$ and one interprets the maps

$$\exp(\epsilon \xi_n D_{H_1 \circ R^k})$$

as a stochastic discrete dynamics according to

$$\begin{aligned} \theta_{k+1} &= \theta_k + \epsilon \xi_k \frac{\partial H_1}{\partial I}(\theta_k + k\Omega(I_k), I_k) + \epsilon \xi_k k \frac{\partial H_1}{\partial \theta}(\theta_k + k\Omega(I_k), I_k) \frac{\partial \Omega}{\partial I}(I_k) + O(\epsilon^2) \\ I_{k+1} &= I_k - \epsilon \xi_k \frac{\partial H_1}{\partial I}(\theta_k + k\Omega(I_k), I_k) + O(\epsilon^2) \end{aligned}$$

Remark: the average value of the term $O(\epsilon^2)$ should be explicitly computed, but we know apriori that the diffusion limit will provide a Fokker-Planck equation in the conservative form. Then the drift coefficient is related to the diffusion coefficient determined by the fluctuating terms $O(\epsilon)$.
Let $\epsilon^{-1} \gg N \gg 1$ we consider an average approach over N steps

$$\begin{aligned} \theta_N &= \theta_0 + \epsilon\sqrt{N} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \xi_k \frac{\partial H_1}{\partial I}(\theta_k + k\Omega(I_k), I_k) \\ &+ \epsilon N^{3/2} \frac{1}{N^3/2} \sum_{k=0}^{N-1} \xi_k k \frac{\partial H_1}{\partial \theta}(\theta_k + k\Omega(I_k), I_k) \frac{\partial \Omega}{\partial I}(I_k) + O(N\epsilon^2) \\ I_{k+1} &= I_k - \epsilon\sqrt{N} \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \xi_k \frac{\partial H_1}{\partial \theta}(\theta_k + k\Omega(I_k), I_k) + O(N\epsilon^2) \end{aligned}$$

The limit $N \to \infty$ with $\epsilon N^3/2 \propto \sqrt{\Delta\tau}$ (τ slow time) implies that the actions can be considered constant whereas the angles tends to a diffusion process with diffusion coefficient

$$D_\theta(\theta, I) = \frac{1}{N^3} \sum_{n,k=0}^{N-1} kn \frac{\partial H_1}{\partial \theta}(\theta + k\Omega(I), I) \frac{\partial \Omega}{\partial I} \frac{\partial H_1}{\partial \theta}(\theta + n\Omega(I), I) \frac{\partial \Omega}{\partial I}(I) E(\xi_k \xi_n)$$

Remark: $D_\theta(I)$ depends on the derivate $\partial\Omega/\partial I$ that is assumed of order $O(1)$ and it vanishes for linear systems.
Remark: we have a dependence on θ if resonance conditions hold $\vec{h} \cdot \Omega = 2\pi n$. The relaxation time scale in the angle variables is of order $(\epsilon^2 N^3)^{-1}$ so that when the iteration $N \simeq \epsilon^{-2/3}$ the angle relaxes but the actions change of order

$O(\epsilon^{2/3})$ and can be considered constant. On the contrary in the diffusion time scale $\epsilon\sqrt{N} \propto \Delta\tau$ (τ is the diffusion time) the angles can be considered relaxed and the action diffusion coefficient reads

$$D_I(I) = \left\langle \frac{1}{N} \sum_{k,n=0}^{N-1} \frac{\partial H_1}{\partial \theta}(\theta + k\Omega(I), I) \frac{\partial H_1}{\partial \theta}(\theta + n\Omega(I), I) E(\xi_k \xi_n) \right\rangle_\theta$$

where the average $<>_\theta$ is justified since we have the relaxation of the angle variable at a time scale $N \propto \epsilon^{-2}$. In the case of a δ-correlated noise $E(\xi_k \xi_n)$, the diffusion coefficient reduces to the random pahse approximation

$$D_I(I) = \langle \left(\frac{\partial H_1}{\partial \theta}(\theta, I) \right)^2 \rangle_\theta \tag{A.2}$$

This equation requires N sufficiently big so ϵ has to be small. (cfr. Averaging Principle section).

Appendix B

1D Stochastic Hamiltonian

In the 1D case the diffusion equation for a stochastically perturbed Hamiltonian system has the form

$$\frac{\partial \rho}{\partial \tau} = \frac{1}{2}\frac{\partial}{\partial I} h^2(I) \frac{\partial \rho}{\partial I}$$

where

$$h^2(I) = D(I) = \frac{1}{\|H_1\|^2} \left\langle \left(\frac{\partial H_1}{\partial \theta} \right)^2 \right\rangle_\theta$$

and $\tau = \|H_1\|^2 t$ is the diffusion time. Then the scaling parameter in the definition of the diffusion time allows to normalize the diffusion coefficient

$$\int_0^{I_{abs}} h^2(I) dI = 1$$

Moreover if the diffusion coefficient can be written in the form

$$D(I) = D\left(\frac{I}{I_0}\right)$$

the dependence of the parameter I_0 can be absorbed by a scaling on the action variable $I/I_0 \to I'$ that corresponds to a scaling in the time scale $\tau/I_0^2 \to \tau'$. Of course one has to consider the change in the boundary conditions. To study the 1D case, we consider a FP (**??**) of the form

$$\frac{\partial \rho}{\partial \tau} = -\frac{1}{2}\frac{\partial}{\partial J}\frac{dh}{dI}\rho + \frac{1}{2}\frac{\partial^2 \rho}{\partial J^2}$$

We can define an effective potential $V(J)$

$$V(J) = -\int \frac{dh}{dI} dJ = -\int \frac{1}{h}\frac{dh}{dI} dI = -\ln h\left(I(J)\right)$$

that introduces a drift term towards the boundary condition at $J = 0$ if we set

$$J(I) = -\int_I^{I_a} \frac{dI}{h(I)}$$

so that $\lim_{I\to 0} J(I) = -\infty$. The FP equation assumes the standard form

$$\frac{\partial \hat{\rho}}{\partial \tau} = \frac{1}{2}\frac{\partial}{\partial J}\frac{dV}{dJ}\hat{\rho} + \frac{1}{2}\frac{\partial^2 \hat{\rho}}{\partial J^2} \tag{B.1}$$

$V(J)$ is a decreasing potential as $J \to 0^-$ so that it introduces a drift force towards the absorbing boundary condition. We use the eq. (**??**) to study the diffusion of weakly chaotic Hamiltonian system in presence of small stochastic perturbations. According to a Nekhoroshev-like estimate for the perturbation term in the Hamiltonian (**??**) we assume

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

where I_* depends on the nonlinear terms and the exponent α is usually related to the dimension of the system. Then we remark that the diffusion coefficient $h^2(I)$ tends to zero exponentially fast as $I_*/I \to 0$ so that the relaxation time scale depends strongly on the initial condition. We are interested in the evolution of an initial distribution with an absorbing boundary condition at $I_a \ll I_*$. The main problem is to estimate the current at the boundary condition and its dependence from the parameters of the diffusion coefficient. The parameter ϵ^2 has the dimension of a diffusion coefficient $[I^2/t]$ but can be scaled introducing a slow time $\tau = \epsilon^2 t$. The FP equation (**??**) can be associated to a stochastic differential equation in the slow time $\tau = \epsilon^2 t$

$$dI = \frac{\alpha}{I}\left(\frac{I_*}{I}\right)^\alpha \exp\left[-2\left(\frac{I_*}{I}\right)^\alpha\right] d\tau + \exp\left[-\left(\frac{I_*}{I}\right)^\alpha\right] dw_\tau$$

Introducing the adimensional action $x = I/I_*$ we have

$$dx = \frac{\alpha}{I_*^2 x^{\alpha-1}} \exp\left[-2\left(\frac{1}{x}\right)^\alpha\right] d\tau + \exp\left[-\left(\frac{1}{x}\right)^\alpha\right] dw_\tau$$

The change of variable

$$y = -\int_x^{x_a} \exp\left[\left(\frac{1}{x}\right)^\alpha\right] dx \qquad x_a = I_a/I_*$$

reduces the evolution to the form

$$dy = \frac{\alpha}{I_*^2} a(y) d\tau + dw_\tau$$

with an absorbing boundary condition at $y = 0$ and

$$a(y) = \frac{1}{x^{\alpha-1}(y)} \exp\left[-\left(\frac{1}{x(y)}\right)^\alpha\right]$$

Remark: $a(y) > 0$ represents a external forcing towards the boundary that vanishes when $x \to 0$ (i.e. $y \to -\infty$).
If one neglects the effect drift field $a(y)$, $y(t)$ is a Wiener process with an absorbing boundary condition and the fundamental solution is

$$\rho(y,t) = \frac{1}{\sqrt{2\pi t}}\left[\exp\left(-\frac{(y+y_0)^2}{2t}\right) - \exp\left(-\frac{(y-y_0)^2}{2t}\right)\right] \qquad y \le 0$$

with $\rho(0,y) = \delta(y + y_0)$. Then we compute the probability current $J(t)$ at the absorbing barrier is

$$J(t) = \frac{1}{2}\left.\frac{\partial \rho}{\partial y}\right|_{y=0} = \frac{y_0}{\sqrt{2\pi} t^{3/2}} \exp\left(-\frac{y_0^2}{2t}\right)$$

$J(t)$ has to be an underestimate of the real current since $a(x) > 0$. In the initial variable we have

$$J(t) = \frac{y_0}{2\sqrt{2\pi}t^{3/2}} \exp\left(-\frac{y_0^2}{2t}\right) \qquad y_0 = -\frac{1}{I_*}\int_{I_0}^{I_a} \exp\left[\left(\frac{I_*}{I}\right)^{\alpha}\right] dI \tag{B.2}$$

Given a generic initial condition $\rho(I_0)$ in the new variable we have

$$\rho(I_0)\frac{dI_0}{dy} = I_*\rho(I_0)\exp\left[\left(-\frac{I_*}{I_0}\right)^{\alpha}\right]$$

where $I_0 = I(y_0)$ an the total current is

$$J(t) = \frac{1}{2\sqrt{2\pi}t^{3/2}} \int_{-\infty}^{0} y(I_0)\exp\left(-\frac{y^2(I_0)}{2t}\right)\rho(I_0)dI_0 \tag{B.3}$$

We observe that the current vanishes at $t = 0$ and increases up to a characteristic time of order $t \simeq y_0^2$: around $t \simeq y_0^2$ the current is stationary since the function

$$f(y) = ye^{-y^2/2}$$

has a critical value at $y = 1$ and the current is stationary at a local maximum at $y_0^2/t \simeq 1$. Due to the sensible dependence in the exponential one can estimate the integral by using the biggest term

$$\int_{I_0}^{I_a} \exp\left[\left(\frac{I_*}{I}\right)^{\alpha}\right] dI \simeq C\exp\left[\left(\frac{I_*}{I_0}\right)^{\alpha}\right]$$

where C is a suitable constant. Then after a transient time, one expects a constant negative current at the barrier due to the particles distributed near the position I_0

$$J(I_0) \simeq C\exp\left[-2\left(\frac{I_*}{I_0}\right)^{\alpha}\right]$$

We propose two different approaches assuming that after a transient time the system relaxes in a condition where the evolution of an ensemble of particle at a position $I = I_0$ mainly depends from the value I_0. This assumption is satisfied when the derivative $dh/dI > 0$. One can estimate the average current measured at the absorbing barrier during time under the assumption that, during the evolution, the particle distribution is divided by a small region into two parts with different relaxation times. This means the time $\tau(I)$ is strongly dependent from the point I (it is true when $d\tau/dI \gg 1$). In a quasi stationary condition the escape rate of a particle at the position I in the interval $[\tau, \tau + d\tau]$ is estimated by $1/\tau(I)$ since the main contribution to the average first passage time is the time spent near the position $x = I$, then we have

$$\frac{\partial\rho}{\partial\tau}(I,\tau) = -\frac{\rho(I,\tau)}{\tau(I)}$$

since one neglects both the incoming probability at the position I in the time $d\tau$ and the surviving probability once the particle has left that position. The evolution of the distribution function is approximated by

$$\rho(I,\Delta t) = \rho(I,\tau)\exp\left(-\frac{\Delta t}{\tau(I)}\right) \tag{B.4}$$

In the case of a Nekhoroshev like diffusion coefficient

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

the average first passage times reads (see Appendix A)

$$\tau(I) = 2\epsilon^2 \int_I^{I_a} x \exp\left[2\left(\frac{I_*}{x}\right)^\alpha\right] dx = 2\epsilon^2 I_*^2 \int_{I_a/I^*}^{u} \frac{e^{2u^\alpha}}{u^3} du$$

where $u = I^*/x$ and we set $u^* = I^*/I_a \simeq 1$. A crude estimate for the integral when $u \gg 1$ is

$$\int_{u^*}^{u} \frac{e^{2u^\alpha}}{u^3} du = \frac{1}{2\alpha} \left.\frac{e^{2u^\alpha}}{u^{4-\alpha}}\right|_{u^*}^{u} - \frac{4-\alpha}{2\alpha} \int_{u^*}^{u} \frac{e^{2u^\alpha}}{u^{5-\alpha}} du$$

The first term gives the main contribution to the integral when $u \gg 1$

$$\int_1^u \frac{e^{2u^\alpha}}{u^3} du \simeq \frac{e^{2u^\alpha}}{u^{4-\alpha}}$$

so that

$$\tau(I) \simeq 2\epsilon^2 \frac{I^{4-\alpha}}{I_*^{2-\alpha}} \exp\left[2\left(\frac{I_*}{I}\right)^\alpha\right]$$

In the previous case $\tau(I)$ has a weak dependence from the position of the absorbing barrier. In an analogous way one can compute the stationary solution $\rho_s(I)$ when a source is set at $I = I_s$ and an absorbing boundary condition at $I = I_a$

$$\frac{\partial}{\partial I} h^2(I) \frac{\partial}{\partial I} \rho_s(I) = -2J\delta(I - I_s) \tag{B.5}$$

We have

$$\rho_s(I) = 2J \int_I^{I_a} \Theta(y - I_s) \frac{dy}{h^2(y)}$$

where J is the incoming particles flow at the source. The escape rate $k(I_s)$ can be estimated by the integral

$$N(I_s) = 2J \int_{I_s}^{I_a} \int_y^{I_a} \Theta(z - I_s) \frac{dz}{h^2(z)} dy = 2J\,(I - I_s) \int_I^{I_a} \left.\frac{dz}{h^2(z)}\right|_{I_s}^{I_a} + 2J \int_{I_s}^{I_a} (I - I_s) \frac{dI}{h^2(I)}$$

and we have

$$k^{-1}(I_s) = \frac{N(I_s)}{J} = 2\int_{I_s}^{I_a} (I - I_s) \frac{dI}{h^2(I)} \propto \frac{(I_a - I_s)^2}{\epsilon^2} \exp\left[2\left(\frac{I_*}{I_s}\right)^\alpha\right] \tag{B.6}$$

and we recover an the evolution equation similar to (**??**). All the previous approaches suggest the current the absorbing barrier is proportional to

$$J \propto \exp\left[-2\left(\frac{I_*}{I_s}\right)^\alpha\right]$$

for a Nekhoroshev's like diffusion coefficient with a strong dependence from the initial condition.

Appendix C

Evaluation of the first non-zero eigenvalue

Let us consider the 1d-Smoluchowski equation

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\frac{dV}{dx}\rho + D\frac{\partial^2 \rho}{\partial x^2} \tag{C.1}$$

It is possible to change the diffusion operator to a self-adjoint form by introducing

$$\rho(x,t) = \exp\left(-\frac{V(x)}{2D}\right) p(x,t)$$

Remark: the operator

$$\mathcal{L}_{FP}\frac{\partial}{\partial x}\frac{dV}{dx} + D\frac{\partial^2}{\partial x^2}$$

satisfies the detailed balance condition for the stationary state and it is self-adjoint for the scalar product

$$(f,g) = \int f(x)\exp\left(\frac{V(x)}{D}\right) g(x)dx$$

Indeed a direct check shows that

$$(f,\mathcal{L}_{FP}g) = \int f(x)\exp\left(\frac{V(x)}{D}\right)\mathcal{L}_{FP}g(x)dx = \int (\mathcal{L}_{FP}f(x))\exp\left(\frac{V(x)}{D}\right) g(x)dx$$

Then the transformation reduces the scalar product to the usual form: this is possible only if the detailed balance condition is satisfied.
A direct calculation provides

$$\exp\left(-\frac{V(x)}{2D}\right)\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial}{\partial x}\frac{dV}{dx}\exp\left(-\frac{V(x)}{D}\right)p + D\frac{\partial}{\partial x}\exp\left(-\frac{V(x)}{2D}\right)\frac{\partial p}{\partial x}$$

and finally

$$\frac{\partial p}{\partial t} = \frac{1}{2}\left[\frac{d^2V}{dx^2} - \frac{1}{2D}\left(\frac{dV}{dx}\right)^2\right]p + D\frac{\partial^2 p}{\partial x^2} \tag{C.2}$$

The eigenfunctions $\phi_\lambda(x)$ of the equation (**??**) satisfy

$$\frac{1}{D}\left[a(x)-\lambda\right]\phi_\lambda(x)=\frac{d^2\phi_\lambda}{dx^2}$$

where

$$a(x)=\frac{1}{2}\left[\frac{1}{2D}\left(\frac{dV}{dx}\right)^2-\frac{d^2V}{dx^2}\right]$$

The solution is expanded in the form

$$p(x,t)=\sum_\lambda c_\lambda(t)\phi_\lambda(x)$$

where

$$c_\lambda(t)=c_\lambda(0)e^{-\lambda t}$$

and we have the orthogonality and completeness condition

$$\int dx\phi_\lambda(x)\phi_{\lambda'}(x)=\delta(\lambda-\lambda') \qquad \sum_\lambda \phi_\lambda(x)\phi_\lambda(x')=\delta(x-x')$$

Using the initial condition

$$\rho(x,0)=\exp\left(-\frac{V(x)}{2D}\right)p(x,0)=\exp\left(-\frac{V(x)}{2D}\right)\sum_\lambda c_\lambda(0)\phi_\lambda(x)$$

we get

$$c_{\lambda'}(0)=\int dx\exp\left(\frac{V(x)}{2D}\right)\rho(x,0)\phi_{\lambda'}(x)$$

In the case $\rho(x,0)=\delta(x-x_0)$ the final solution reads

$$\rho(x,t)=\exp\left(-\frac{(V(x)-V(x_0))}{2D}\right)\sum_\lambda e^{-\lambda t}\phi_\lambda(x_0)\phi_\lambda(x)$$

If we set an absorbing boundary condition at $x=0$ ($\rho(0,t)=0$), then we require $\phi_\lambda(0)=0$ for the eigenfunctions and the current at the barrier reads

$$J(t)=D\frac{\partial\rho}{\partial x}(0,t)=D\exp\left(-\frac{(V(0)-V(x_0))}{2D}\right)\sum_\lambda e^{-\lambda t}\phi_\lambda(x_0)\frac{\partial\phi_\lambda}{\partial x}(0) \quad \text{(C.3)}$$

We see the existence of different decaying scales related to λ^{-1} and to the initial condition.
Let us consider the case $V(x)=-\nu x$ (i.e. we have a constant drift toward the barrier) with an absorbing boundary condition at $x=0$, then $a(x)=\nu^2/4D$ and the self-adjoint operator reads

$$-\frac{\nu^2}{4D}+D\frac{\partial^2}{\partial x^2}$$

and it is negative defined, so that all the eigenvalues are ≤ 0; indeed $D\partial^2/\partial x^2$ is also a negative defined symmetric operator and we have an upper bound to the spectrum (except the zero eigenvalue) $\leq -\nu^2/4D$. The eigenvector equation

$$-\frac{1}{D}\left[\lambda-\frac{\nu^2}{4D}\right]\phi_\lambda(x)=\frac{d^2\phi_\lambda}{dx^2}$$

with the boundary condition $\phi_\lambda(0) = 0$. If we set

$$k_\lambda = \sqrt{\lambda - \frac{\nu^2}{4D}}$$

we have non-trivial solutions with

$$\phi_\lambda(x) = \sqrt{\frac{1}{\pi}} \sin\left(\frac{k_\lambda}{\sqrt{D}} x\right)$$

(the factor is a normalizing condition). The zero-eigenvalue corresponds to a trivial solution and we have an upper limit for the negative eigenvalues from the condition

$$\frac{\nu^2}{4D} - \lambda > 0 \qquad \Rightarrow \qquad \lambda_{min} = \frac{\nu^2}{4D}$$

λ_{min}^{-1} defines the decaying characteristic time for Fiedler's eigenvector. Therefore the existence of a constant drift field implies an upper bound to the Fiedler's eigenvalue. As an example we compute the current at the absorbing barrier for $\nu = 0$ so that $\phi_\lambda(x) \propto \sin \sqrt{(\lambda/D)}x$, then from eq. (**??**) we have

$$J(t) = \frac{D}{\pi} \int_0^\infty e^{-\lambda t} \sin\left(\sqrt{\frac{\lambda}{D}} x_0\right) \sqrt{\frac{\lambda}{D}} d\left(\sqrt{\frac{\lambda}{D}}\right) = \frac{D}{\pi} \int_0^\infty e^{-u^2 Dt} \sin(u x_0)\, u du$$

where $\lambda = u^2 D$. The integral reduces to a Gaussian integral since

$$\exp\left(-u^2 Dt\alpha \pm iux_0\right) = \exp\left[-Dt\alpha\left(u \pm i\frac{x_0}{2Dt\alpha}\right)^2\right] \exp\left(-\frac{x_0^2}{4Dt\alpha}\right)$$

and

$$\begin{aligned} & \frac{1}{2i} \int_0^\infty \exp\left[-Dt\left(u - i\frac{x_0}{2Dt}\right)^2\right] udu - \int_0^\infty \exp\left[-Dt\left(u + i\frac{x_0}{2Dt}\right)^2\right] udu \\ = & \int_{-\infty}^\infty \exp\left[-Dt\left(u - i\frac{x_0}{2Dt}\right)^2\right] udu = \frac{x_0}{4}\sqrt{\frac{\pi}{(Dt)^{3/2}}} \end{aligned}$$

Then we get

$$J(t) = \frac{D}{2\pi i} \exp\left(-\frac{x_0^2}{4Dt}\right) \int_{-\infty}^\infty \exp\left[-\frac{2Dt}{2}\left(u - i\frac{x_0}{2Dt}\right)^2\right] udu = \exp\left(-\frac{x_0^2}{4Dt}\right) \frac{x_0}{\sqrt{2\pi D}(2t)^{3/2}}$$

that coincides with the results (**??**). The presence of a drift field modifies the previous calculation and we define $\lambda - \nu^2/(4D) = k_\lambda^2 = u^2 D$ and we have the relation

$$\lambda t \pm i\frac{k_\lambda}{\sqrt{D}} x_0 = \left(u^2 D + \frac{\nu^2}{4D}\right) t \pm iux_0 = \frac{\nu^2}{4D} t + \frac{2Dt}{2}\left(u \pm i\frac{x_0}{2Dt}\right)^2 + \frac{x_0^2}{4Dt}$$

we get the Gaussian

$$\exp\left[-\frac{\nu^2}{4D} t - \frac{x_0^2}{4Dt}\right] \exp\left[-\frac{2Dt}{2}\left(u \pm i\frac{x_0}{2Dt}\right)^2\right)$$

Let $V(x) = -\nu x$, the current reads

$$J(t) = \exp\left(-\frac{\nu x_0}{2D} - \frac{\nu^2}{4D}t - \frac{x_0^2}{4Dt}\right)\frac{x_0}{\sqrt{2\pi D}(2t)^{3/2}} = \exp\left(-\frac{(x_0+\nu t)^2}{2Dt}\right)\frac{x_0}{\sqrt{2\pi D}(2t)^{3/2}}$$

and we have a current describes by a moving Gaussian towards the barrier which is diffusing. In an adiabatic approximation it is possible to approximate the current for a slowly varying drift in the form

$$J(t|x_0) \simeq \exp\left(-\frac{(\Phi^t(x_0))^2}{2Dt}\right)\frac{x_0}{\sqrt{2\pi D}(2t)^{3/2}}$$

where $\Phi^t(x_0)$ is the phase flow of the drift force

$$\frac{dx}{dt} = -\frac{dV}{dx}$$

This approximation holds if the drift dynamics of the distribution is well approximated by the dynamics of the initial condition x_0 (average field approximation). $J(t) = dP/dt$ has a physical dimension $[t]^{-1}$ so that it is a scalar with respect a change of variables. For a generic initial distribution $\rho_0(x_0)$ we have

$$J(t) \simeq \frac{1}{\sqrt{2\pi D}(2t)^{3/2}}\int_0^\infty x_0 \exp\left(-\frac{(\Phi^t(x_0))^2}{2Dt}\right)\rho_0(x_0)dx_0$$

We consider an explicit example with $h(I) = \eta(I/I_*)^n$, we have the FP equation

$$\frac{\partial\rho}{\partial t} = \frac{\eta^2}{2}\frac{\partial}{\partial I}\left(\frac{I}{I_*}\right)^{2n}\frac{\partial\rho}{\partial I}$$

To simplify the formulas we introduce the scaled time $\tau = t\eta^2/(2I_*^2)$ and the scaled action $u = I/I_*$ and the FP equation reads

$$\frac{\partial\rho}{\partial\tau} = \frac{\partial}{\partial u}u^{2n}\frac{\partial\rho}{\partial u}$$

and the change of variable (**??**) reads

$$J = -\int_u^{u_a}\frac{du}{u^n} = \frac{1}{u_a^{n-1}(n-1)}\left[1-\left(\frac{u_a}{u}\right)^{n-1}\right]$$

In the new variable the FP reduces

$$\frac{\partial\hat\rho}{\partial\tau} = -\frac{\partial}{\partial J}nu^{n-1}(J) - \frac{\partial^2\hat\rho}{\partial^2 J}$$

where

$$u(J) = \frac{u_a}{\left(1-(n-1)u_a^{n-1}J\right)^{1/(n-1)}}$$

The fundamental solution depends from the initial condition $J_0(u_0)$, $u_0 \in [0, I_a/I_*]$ and we integrate the drift field

$$\frac{dJ}{d\tau} = \frac{du}{d\tau}\frac{dJ}{du} = nu^{n-1}$$

so that

$$\frac{du}{d\tau} = nu^{2n-1} \quad \Rightarrow \quad \frac{1}{u_0^{2(n-1)}} - \frac{1}{u^{2(n-1)}} = 2n(n-1)\tau$$

and

$$J(\tau) = \frac{1}{(n-1)}\left[\frac{1}{u_a^{n-1}} - \frac{1}{u_0^{(n-1)}}\sqrt{1-2n(n-1)u_0^{2(n-1)}\tau}\right] \qquad J \leq 0$$

Remark: the previous expression is singular in τ, but our approximation holds only when $2n(n-1)u_0^{2(n-1)}\tau \ll 1$ so that

$$\frac{1}{u_0^{n-1}}\sqrt{1-2n(n-1)u_0^{2(n-1)}\tau} \simeq \frac{1}{u_0^{(n-1)}} - u_0^{n-1}n(n-1)\tau$$

and

$$J(\tau) \simeq \frac{1}{(n-1)}\left[\frac{1}{u_a^{n-1}} - \frac{1}{u_0^{(n-1)}} + u_0^{(n-1)}n(n-1)\tau\right] = J_0 + nu_0^{(n-1)}\tau$$

This condition defines a time scale in the initial variables

$$t_s \ll \frac{I_*^2}{\eta^2 n}\left(\frac{I_*}{I_0}\right)^{2(n-1)}$$

as a function of the initial condition I_0: the r.h.s. defines the dynamic aperture time scale at the initial condition I_0.

An approximation for the current at the absorbing barrier for $\rho_0(I) = \delta(I - I_0)$ is given by

$$J_c(\tau) \simeq \frac{J_0}{\sqrt{2\pi}(2\tau)^{3/2}} \exp\left(-\frac{\left(J_0 + nu_0^{(n-1)}\tau\right)^2}{2\tau}\right)$$

Finally in the initial time $\tau = t\eta^2/(2I_*^2)$ and using the initial action $I = uI_*$ we have

$$J_c(t) \simeq \frac{2J_0 I_*^3}{\sqrt{\pi}\eta^3 t^{3/2}} \exp\left(-\frac{\left(2I_*^2 J_0 + (I_0/I_*)^{(n-1)} n\eta^2 t\right)^2}{4I_*^2\eta^2 t}\right) \tag{C.4}$$

where

$$J_0 = \frac{1}{(n-1)}\left[\frac{I_*^{n-1}}{I_a^{n-1}} - \frac{I_*^{n-1}}{I_0^{n-1}}\right]$$

The same procedure can be applied in the case of a Nekhoroshev's like diffusion coefficient

$$h\left(\frac{I}{I_*}\right) = \eta \exp\left(-\frac{I_*}{I}\right)^{\alpha}$$

We perform the change of variable

$$J = -\int_u^{u_a} \exp\left(u^{-\alpha}\right) du \simeq (u - u_a)\exp\left(u^{-\alpha}\right)$$

where $u = I/I_*$ and we use the scaled time $\tau = t\eta^2/(2I_*^2)$. The drift field for the new FP equation is

$$\frac{dJ}{d\tau} = \alpha u^{-\alpha-1} \exp(-u)^{-\alpha} \simeq \alpha u_0^{-\alpha-1} \exp(-u_0)^{-\alpha}$$

assuming an average field approximation. Then we get the expression for the out-going current in the initial time $\tau = t\eta^2/(2I_*^2)$

$$J_c(t) \simeq \frac{2J_0 I_*^3}{\sqrt{\pi}\eta^3 t^{3/2}} \exp\left(-\frac{\left(2J_0 I_*^2 + \alpha \left(\frac{I_0}{I_*}\right)^{-\alpha-1} \exp - \left(\frac{I_0}{I_*}\right)^{-\alpha} \eta^2 t \right)^2}{4I_*^2 \eta^2 t} \right) \tag{C.5}$$

Remark: the different between the expressions (**??**) and (**??**) is the drift field we consider constant, therefore the quantification of the out-going current depends on the value J_0. Indeed if one introduces $x = J_0/\sqrt{\tau}$ and we neglect the effect of the drift, the both expressions are

$$J_c(x) \simeq J_0^{-2} \frac{x^3}{4\sqrt{\pi}} \exp\left(-\frac{x^2}{2}\right)$$

Therefore J_0^{-2} is the scaling factor for the current and its evolution time scale is

$$t_c \simeq \frac{2I_*^2}{\eta^2} J_0^2$$

A crude estimate gives

$$J_0^2 \simeq \frac{1}{(n-1)^2} \left(\frac{I_*}{I_0}\right)^{2(n-1)}$$

for the power law, whereas we have

$$J_0^2 \simeq \left(\frac{I_a}{I_*}\right)^2 \exp\left(2\frac{I_*}{I_0}^{\alpha}\right)$$

in the Nekhoroshev-like diffusion. Then we expect a different scaling of the current as we move the initial condition I_0: in the first case the current decreases as a power law and the diffusion time scale increases by the same factor; in the second case the current decreases exponentially

$$\propto \left(\frac{I_*}{I_a}\right)^2 \exp\left(-2\frac{I_*}{I_0}^{\alpha}\right)$$

Taking into account that the diffusion time is also scaled, the time to loose a fixed ΔP fraction for a $delta(I - I_0)$ initial condition is computed by integrating the current

$$\Delta P = \int J_c(t)dt \simeq J_0^{-2} \int_0^{x_P} \frac{x^3}{4\sqrt{\pi}} \exp\left(-\frac{x^2}{2}\right) \frac{dt}{dx} dx = -\frac{I_*^2}{2\sqrt{\pi}\eta^2} \int_0^{x_P} x \exp\left(-\frac{x^2}{2}\right) dx$$

where we use

$$t = \frac{2J_0^2 I_*^2}{\eta^2 x} \qquad \Rightarrow \qquad \frac{dt}{dx} = -\frac{2J_0^2 I_*^2}{\eta^2 x^2}$$

Then we get a relation between ΔP and x_p

$$\Delta P = -\frac{I_*^2}{2\sqrt{\pi}\eta^2}\left(1 - \exp\left(-\frac{x_P^2}{2}\right)\right) \tag{C.6}$$

and we estimate the dynamics aperture time scale

$$t_{ap} = \frac{2J_0^2 I_*^2}{\eta^2 x_P} \tag{C.7}$$

that is consistent with t_c.

Appendix D

First passage time and transition rate escaping from a barrier

We first consider the generic diffusion problem

$$\frac{\partial \rho}{\partial t}(x,t) = \mathcal{L}_{FP}\rho(x,t)$$

where $\mathcal{L}_{FP}$ is a generic Fokker-Planck operator and we assume that $x \in \Omega$ and $\partial\Omega$ is an absorbing boundary condition $\rho(x,t) = 0$ at $x \in \partial\Omega \quad \forall t$. By definition the probability to be in Ω at time t for an initial condition $\delta(x-a)$ is

$$P_\Omega(t|a) = \int_\Omega \rho(x,t|a)dx$$

Let t_a the first passage time at $\partial\Omega$ for a realization starting from a at $t = 0$ and let $p(t|a$ its probability distribution, we have by definition

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

since $\partial\Omega$ is an absorbing boundary condition. Then

$$p(t|a) = \int_\Omega \frac{\partial}{\partial t}\rho(x,t|a)dx = \int_\Omega \mathcal{L}_{FP}\rho(x,t|a)dx \tag{D.1}$$

To get an equation for the average first passage time $\tau(a)$

$$\tau(a) = \int_0^\infty tp(t|a)dt = \int_\Omega \int_0^\infty t\frac{\partial}{\partial t}\rho(x,t|a)dt\,dx = -\int_0^\infty \int_\Omega \rho(x,t|a)dx\,dt$$

we use the adjoint operator $L^\dagger_{FP}$ and taking into account the boundary conditions

$$L^\dagger_{FP}\tau(a) = -\int_0^\infty \int_\Omega L^\dagger_{FP}\rho(x,t|a)dx\,dt$$

Remark: $L_{FP}^{\dagger}$ acts on the initial state a. From the Kolmogorov relation for Markov processes we get

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

so that

$$\int_{\Omega} dx_0\rho(x_0,s|a,0)\frac{\partial}{\partial s}\rho(x,t|x_0,s) + \rho(x,t|x_0,s)\frac{\partial}{\partial \tau}\rho(x_0,s|a,0)$$
$$= \int_{\Omega} dx_0\rho(x_0,s|a,0)\frac{\partial}{\partial s}\rho(x,t|x_0,s) + \rho(x,t|x_0,s)\mathcal{L}_{FP}\rho(x_0,s|a,0)$$

Then

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

and since we consider a stationary process

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

Letting $s \to 0$, one gets the adjoint equation

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

According to the previous result, the average first passage time reads

$$L_{FP}^{\dagger}\tau(a) = -\int_0^{\infty}\int_{\Omega} L_{FP}^{\dagger}\rho(x,t|a)dx\,dt = \int_{\Omega}\int_0^{\infty}\frac{\partial}{\partial t}\rho(x,t|a,0)dt\,dx = -\int_{\Omega}\delta(x-a)dx = -1$$

since $\lim_{t\to\infty}\rho(x,t|a) = 0$ due to the presence of an absorbing boundary condition. In an explicit form we have the equation

$$L_{FP}^{\dagger}\tau(a) = -1 \tag{D.2}$$

Similar equation can be obtained for the variance of the distribution $p(t|a)$

$$\sigma^2(a) = \int_0^{\infty} t^2 p(t|a)dt - \tau^2(a) = -2\int_0^{\infty}\int_{\Omega} t\rho(x,t|a)dx\,dt - \tau^2(a)$$

Then we have

$$L_{FP}^{\dagger}\sigma^2(a) = -2\int_0^{\infty}\int_{\Omega} tL_{FP}^{\dagger}\rho(x,t|a)dx\,dt - L_{FP}^{\dagger}\tau^2(a)$$

Using the relation

$$2\int_{\Omega}\int_0^{\infty} t\frac{\partial}{\partial t}\rho(x,t|a,0)dt\,dx = -2\int_{\Omega}\int_0^{\infty}\rho(x,t|a,0)dt\,dx = -2\int_0^{\infty} tp(t|a)dt = -2\tau(a)$$

we finally get the equation

$$L_{FP}^{\dagger}\sigma^2(a) = -2\tau(a) - L_{FP}^{\dagger}\tau^2(a) \tag{D.3}$$

We specify the previous relation in the self-adjoint case (1D-Hamiltonian case)

$$\mathcal{L}_{FP} = \frac{1}{2}\frac{\partial}{\partial x}D(x)\frac{\partial}{\partial x}$$

one gets

$$\int_{\Omega} \mathcal{L}_{FP}\rho(x,t|a)dx = -\int_{\partial\Omega} D(x)\frac{\partial \rho}{\partial x}\cdot d\sigma$$

that defines the outgoing current through the boundary $\partial\Omega$ at time t. It is possible to evaluate an average value $\tau(a)$ for the first passage time starting form a point a

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

so that

$$\tau(a) = 2\int_a^{x_a} \frac{x}{D(x)}dx \tag{D.4}$$

where we set the boundary condition $\tau(x_a) = 0$ at the absorbing boundary condition x_a. Then from the relation

$$\frac{1}{2}\frac{\partial}{\partial x}D(x)\frac{\partial}{\partial x}\tau^2(x) = \frac{\partial}{\partial x}D(x)\tau(x)\frac{\partial \tau}{\partial x} = -2\tau(x) + D(x)\left(\frac{\partial \tau}{\partial x}\right)^2$$

we get

$$\frac{1}{2}\frac{\partial}{\partial a}D(a)\frac{\partial \sigma^2}{\partial a} = -D(a)\left(\frac{\partial \tau}{\partial a}\right)^2 = -\frac{4a^2}{D(a)} \tag{D.5}$$

and

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

We remark that the computation of the average first passage time requires the knowledge of the evolution at all times, however this result can be related to the escape probability rate. Let us apply the first average passage time to the computation of the Kramer probability escape from a potential well. We start from the Fokker-Planck equation

$$\frac{\partial}{\partial x}\frac{\partial V}{\partial x}\rho + D\frac{\partial^2 \rho}{\partial x^2} = \frac{\partial \rho}{\partial t}$$

where the potential $V(x)$ has a minimum at $x = x_0$ and a saddle point at $x = x_1$. The average first time $\tau(x)$ is written

$$-\frac{\partial V}{\partial x}\frac{\partial \tau}{\partial x} + D\frac{\partial^2 \tau}{\partial x^2} = -1$$

and we get

$$\frac{\partial}{\partial x}\exp\left(-\frac{V(x)}{D}\right)\frac{\partial \tau}{\partial x} = -\frac{1}{D}\exp\left(-\frac{V(x)}{D}\right)$$

The previous equation can be integrated

$$\frac{d\tau}{dx} = -\frac{1}{D}\exp\left(\frac{V(x)}{D}\right)\int_{-\infty}^{x} dz\exp\left(-\frac{V(z)}{D}\right)$$

where we assume $\lim_{x\to-\infty} V(x) = \infty$. The final result is

$$\tau(x) = \frac{1}{D}\int_x^{x_a} \exp\left(\frac{V(y)}{D}\right)\int_{-\infty}^{y} \exp\left(-\frac{V(z)}{D}\right) dzdy \tag{D.6}$$

Then the Kramer's estimate follows by approximating the integral at the critical points of the potential (x_a is not set near the local maximum x_1) so that the result independent from x. The transition rate is defined

$$k_c = \frac{1}{\tau}$$

It is critical to understand the underlying assumptions for the previous estimate: the fact that k_c does not depend on x is due to the fast relaxation scale time inside the potential well with respect to the escape time scale from the potential well and to the quasi-stationary distribution that concentrates the particles near the local minimal point where the effect of the potential can be approximated by a parabolic potential: then we estimate $\tau \simeq \tau(x)$ where $x \simeq x_0$.
In a analogous way one considers the stationary solution in presence of a local source at $x_s \ll x_0$ and the absorbing barrier at $x_1 \ll x_a$. The stationary solution obeys to the differential equation

$$\frac{\partial}{\partial x}\frac{\partial V}{\partial x}\rho(x) + D\frac{\partial^2 \rho}{\partial x^2} = -J\delta(x - x_s)$$

that can be integrated

$$\rho(x) = \frac{J}{D}\exp\left(\frac{-V(x)}{D}\right)\int_x^{x_a} \exp\left(\frac{V(y)}{D}\right)\Theta(y - x_s)dy$$

The parameter J defines the current at the absorbing barrier and the transition rate can be computed as the ratio between J and the number of particles in the potential well, $k_c = J/N$ where

$$N = \frac{J}{D}\int_{-\infty}^{x_1} \exp\left(\frac{-V(x)}{D}\right)\int_x^{x_a} \exp\left(\frac{V(y)}{D}\right)\Theta(y - x_s)dy\, dx \tag{D.7}$$

and the Kramer's estimate follows by approximating the integral (**??**). Assuming that $V(x)$ has a local minimum A and a local maximum B $x_s < x_A < x_B$
We apply the previous results in the Hamiltonian case when the operator $\mathcal{L}_{FP}$ is a self-adjoint operator and the first passage time equation (**??**) is

$$\frac{\partial}{\partial a}h^2(a)\frac{\partial}{\partial a}\tau(a) = -2$$

with $\lim_{a\to 0} h(a) = 0$. In the 1d case with an absorbing boundary condition at x_a, the equation reads

$$\frac{\partial}{\partial a}\tau(a) = -\frac{2a}{h^2(a)}$$

so that

$$\tau(a) = \int_a^{x_a} \frac{2x}{h^2(x)}dx \qquad a \geq 0 \tag{D.8}$$

Appendix E

Gaussian Processes

Given x_j $j = 1, .., n$ Gaussian variables we have the join distribution

$$\rho(x_1, .., x_n) = [(2\pi)^n \det \Sigma]^{-1/2} \exp\left(-\frac{1}{2} x^T \Sigma^{-1} x\right)$$

where Σ is the correlation matrix. It is clear that the second moments have to define all the higher order moments of the distribution. It exists an orthogonal matrix O such that $O^T \Sigma O = \Lambda$ where Λ is a diagonal matrix and we can introduce the independent variables $u = O^t x$ with variance Λ. Then we compute

$$\begin{aligned} < x_{i_1}..x_{i_{2k}} > &= \langle O_{i_1,j_1} u_{j_1} .....O_{i_{2k},j_{2k}} u_{j_{2k}} \rangle \\ &= \sum_{\{(r_1,s_1),..,(r_k,s_k)\}} O_{r_1,j_1} \Lambda_{j_1} O_{s_1,j_1} .....O_{r_k,j_k} \Lambda_{j_k} O_{s_k,j_k} \\ &= \sum_{\{(r_1,s_1),..,(r_k,s_k)\}} \Sigma_{r_1,s_1} .....\Sigma_{r_k,s_k} \end{aligned} \tag{E.1}$$

where the sum runs over all the possible different k couples (not ordered) that can be obtained from the sequence of index $i_1, ..., i_{2k}$ and we have used the definition $\Sigma = O\Lambda O^T$. In case the previous expression is independent from the couples $\{(r_1, s_1), .., (r_k, s_k)\}$, we have

$$< x_{i_1}..x_{i_{2k}} >= (2k-1)!!\Sigma_{i_1,i_2} .....\Sigma_{i_{2k-1},s_{2k}}$$

which generalizes the relations among the cumulants for Gaussian variables.
Let us consider a Örstein-Uhlenbeck process $x(t)$ so that

$$< x(t)x(t+\tau) >= \sigma^2 \exp(-\gamma|\tau|)$$

Given n-times $t_1 < t_2 < ... < t_n$ we can construct n Gaussian variables $x_i = x(t_i)$ whose correlation matrix reads

$$C_{ij} = \sigma^2 \exp(\gamma|t_i - t_j|)$$

and the joint probability function is

$$\rho(x_1, .., x_n) = [(2\pi)^n \det C]^{-1/2} \exp\left(-\frac{1}{2} x^T C^{-1} x\right)$$

The inverse matrix C^{-1} has a very simple tri-diagonal structure

$$C^{-1} = \sigma^{-2} \begin{pmatrix} d_1 & -e_1 & 0 & .... & 0 & 0 & 0 \\ -e_1 & d_2 & -e_2 & .... & 0 & 0 & 0 \\ & & & .... & & & \\ 0 & 0 & 0 & .... & -e_{n-2} & d_{n-1} & -e_{n-1} \\ 0 & 0 & 0 & .... & 0 & -e_{n-1} & d_n \end{pmatrix}$$

where introducing

$$r_i = \begin{cases} 0 & i = 0 \\ \exp[-\gamma(t_{i+1} - t_i)] & 1 \leq i \leq n-1 \end{cases}$$

we define

$$e_i = \begin{cases} 0 & i = 0 \\ r_i/(1 - r_i^2) & 1 \leq i \leq n-1 \\ 0 & i = n \end{cases}$$

and

$$d_i = 1 + r_i e_i + r_{i-1} e_{i-1}$$

Let us explicitly compute

$$x^T C^{-1} x = \sigma^{-2} \sum_{i=1}^{n} (d_i x_i^2 - 2 e_{i-1} x_i x_{i-1})$$

then a direct substitution of the definitions give

$$\sigma^{-2} \sum_{i=1}^{n} \left[(1 + r_{i-1} e_{i-1}) x_i^2 - 2 e_{i-1} x_i x_{i-1}\right] + \sigma^{-2} \sum_{i=1}^{n} r_i e_i x_i^2 =$$
$$\sigma^{-2} \sum_{i=1}^{n} \left[(1 + r_{i-1} e_{i-1}) x_i^2 - 2 e_{i-1} x_i x_{i-1} + r_{i-1} e_{i-1} x_{i-1}^2\right]$$

and finally we get the expression

$$\sum_{i=1}^{n} \frac{(x_i - r_{i-1} x_{i-1})^2}{\sigma^2 (1 - r_{i-1}^2)}$$

so that the joint distribution can be written in the form

$$\rho(x_1, .., x_n) = [(2\pi)^n \det C]^{-1/2} \exp\left(-\sum_{i=1}^{n} \frac{(x_i - r_{i-1} x_{i-1})^2}{2\sigma^2 (1 - r_{i-1}^2)}\right)$$

Remark that the distribution is singular if $r_i = 1$ (i.e. $t_{i+1} = t_i$). After some calculation we also get

$$\det C^{-1} = \prod_{i=1}^{n} \frac{1}{\sigma^2 (1 - r_{i-1}^2)}$$

Let us estimate the expectation value

$$\int \prod_{i=1}^{n} x_i \rho(x) dx$$

(n should be even otherwise the expression vanishes). According to the result (A.1) we ave to consider all the possible $n/2$ couples obtained from the indexes $1, ..., n$ and compute the corresponding correlation coefficients. Starting from the case $(n, n-1), (n-2, n-3), .., (2, 1)$ we have a contribution

$$\sigma^n \exp\left(-\gamma \sum_{i=1}^{n/2} (t_{2i} - t_{2i-1})\right)$$

and all the other cases can be obtained by exchanging indexes among the couples. Any exchange between the indexes j, k $(j > k)$ implies a factor

$$\exp(-2\gamma(t_j - t_k)) \qquad t_j > t_k$$

which multiplies the previous contribution. Therefore in the white noise limit $\gamma, \sigma \to \infty$ and σ^2/γ finite, we have the weak limit

$$\int \prod_{i=1}^{n} x_i \rho(x) dx \to \prod_{i=1}^{n/2} \delta(t_{2i} - t_{2i-1})$$

Finally we consider the expression

$$\int_0^t dt_n .... \int_0^{t_2} dt_1 \langle \prod_{i=1}^{n} x(t_i) \rangle = \frac{1}{2^{n-1}} \left\langle \left(\int_0^t x(s) ds \right)^n \right\rangle$$

since we can exchange the integration indexes. $\int x(s)ds$ is a Gaussian variable and we have the relation

$$\int_0^t dt_n .... \int_0^{t_2} dt_1 \langle \prod_{i=1}^{n} x(t_i) \rangle = \sigma^n \frac{(n-1)!!}{2^{n-1}} \left(\int_0^t \int_0^t \exp(-\gamma|s-u|) ds du \right)^{n/2}$$

This result has to be consistent with the relation (A.1) so that we get the equality

$$\int_0^t dt_n .... \int_0^{t_2} dt_1 \langle \prod_{i=1}^{n} x(t_i) \rangle = \sigma^n \frac{(n-1)!!}{2^{n/2-1}} \left(\int_0^t \int_0^s \exp(-\gamma|s-u|) ds du \right)^{n/2}$$

For a Gaussian variable x of average value $< x >$ and variance σ^2, we have the following equality

$$< \exp(x) >= \exp\left(< x > + \frac{\sigma^2}{2}\right) \tag{E.2}$$

For the proof we use the identity

$$\begin{aligned} < x^n > &= < ((x- < x >) + < x >)^n >= \sum_{j=0}^{n} \frac{n!}{j!(n-j)!} < (x- < x >)^j >< x >^{n-j} \\ &= \sum_{k=0}^{n/2} \frac{n!}{2k!(n-2k)!} (2k-1)!! \sigma^{2k} < x >^{n-2k} \end{aligned}$$

then we get

$$< x^n >= \sum_{k=0}^{n/2} \frac{n!}{2^k k!(n-2k)!} \sigma^{2k} < x >^{n-2k}$$

Then we consider the exponential expansions

$$\exp(< x >) \exp\left(\frac{\sigma^2}{2}\right) = \sum_{m\geq 0}\sum_{k\geq 0} \frac{< x >^m}{m!}\frac{\sigma^{2k}}{2^k k!} = \sum_{n\geq 0}\sum_{k=0}^{n/2} \frac{< x >^m}{(n-2k)!}\frac{\sigma^{2k}}{2^k k!}$$

$(n = m + 2k)$ and

$$\sum_{n\geq 0} \frac{< x^n >}{n!} = \sum_{n\geq 0}\sum_{k=0}^{n/2} \frac{1}{2^k k!(n-2k)!} \sigma^{2k} < x >^{n-2k}$$

The comparison between the two expressions gives the equation (A.2). We remark that the algebraic properties of the exponential function are fundamental for the proof.

The evolution of Gaussian processes is completely defined by the first and second moment of the distribution and it is described by the Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = \frac{D_{ij}}{2} \frac{\partial^2 \rho}{\partial x_i \partial x_j} \tag{E.3}$$

(D_{ij} is a strictly positive defined symmetric matrix), with an initial condition $\rho(x, 0) = \rho_0(x)$ and suitable boundary conditions. The equation (A.3) allows to compute directly the evolution of the distribution moments. In particular the mean value is zero and the covariance is

$$\frac{d < x_h x_k >}{dt} = \int x_h x_k \frac{\partial \rho}{\partial t} dx = \int D_{hk} \rho dx = D_{hk}$$

Then the solution of eq. (A.3) with initial condition $\rho_0(x) = \delta(x - x_0)$ reads

$$\rho(x, t) = \frac{1}{(2\pi)^{n/2}\sqrt{\mathrm{Det} D t}} \exp\left(-\frac{((x - x_0)D^{-1}(x - x_0))}{2t}\right)$$

with open boundary conditions. Since the equation is linear, any superposition of the previous fundamental distribution, is still a solution of the FP equation. In the one degree of freedom case, if one introduces a reflecting boundary condition at $x = 0$ we can find a solution for $x > 0$ with initial condition $\rho_0(x) = \delta(x-x_0)$ in the form

$$\rho(x, t) = \frac{1}{\sqrt{2\pi D t}} \exp\left(-\frac{(x - x_0)^2}{2Dt}\right) + \frac{1}{\sqrt{2\pi D t}} \exp\left(-\frac{(x + x_0)^2}{2Dt}\right)$$

Analogously the presence of an absorbing boundary condition at $x = 0$ can be realized by the superposition

$$\rho(x, t) = \frac{1}{\sqrt{2\pi D t}} \exp\left(-\frac{(x - x_0)^2}{2Dt}\right) - \frac{1}{\sqrt{2\pi D t}} \exp\left(-\frac{(x + x_0)^2}{2Dt}\right)$$

We remark as $\rho(x = 0, t) = 0$ and it is possible to evaluate the surviving probability

$$\frac{dP}{dt} = \int_0^\infty \frac{\partial \rho}{\partial t} dx = \frac{D}{2} \int_0^\infty \frac{\partial^2 \rho}{\partial x^2} dx = -\frac{D}{2} \frac{\partial \rho}{\partial x}(0, t)$$

that implies

$$P(t) = 1 - \frac{D}{2} \int_0^t \frac{\partial \rho}{\partial x}(0, t) dt$$

A direct computation gives

$$\frac{D}{2} \frac{\partial \rho}{\partial x}(0, t) = \frac{x_0}{\sqrt{2\pi D t^3}} \exp\left(-\frac{x_0^2}{2Dt}\right)$$

If one considers an initial population density $\mu(x_0)$, the probability flow on the barrier is estimated

$$J(t) = \int \int_0^{Dt} \frac{x_0}{\sqrt{2\pi u^3}} \exp\left(-\frac{x_0^2}{2u}\right) du \mu(x_0) dx_0$$

Let $s = 1/u$ and $y = \sqrt{s}$ we get the integrals

$$\int_{1/Dt}^\infty \frac{x_0}{\sqrt{2\pi s}} \exp\left(-\frac{x_0^2 s}{2}\right) ds = \int_{1/\sqrt{Dt}}^\infty \frac{2x_0}{\sqrt{2\pi}} \exp\left(-\frac{x_0^2 y^2}{2}\right) dy$$

and performing the final scaling $v = y x_0$ we compute

$$J(t) = \frac{2}{\sqrt{2\pi}} \int \int_{x_0/\sqrt{Dt}}^\infty \exp\left(-\frac{v^2}{2}\right) dv \mu(x_0) dx_0 = 2 \int \left(1 - \operatorname{Erf}\left(\frac{x_0}{\sqrt{Dt}}\right)\right) \mu(x_0) dx_0 \tag{E.4}$$

This result is related to the estimate of the first passage time through a barrier $x = b$. For the usual Brownian motion we have the definition

$$T_b = \inf \{t \;/\; w_t = b\}$$

with $w_0 = 0$. Using the Markov property we get

$$\mathcal{P}(w_t > b \;/\; T_b < t) = \mathcal{P}(w_{t-T_b} > 0)$$

i.e. since $w_{T_b} = b$ one evaluates the probability that the trajectory moves left or right. By definition we get

$$\mathcal{P}(w_t > b \;/\; T_b < t) = \frac{\mathcal{P}(w_t > b, T_b < t)}{\mathcal{P}(T_b < t)} = \frac{\mathcal{P}(w_t > b)}{\mathcal{P}(T_b < t)}$$

In a random walk we have

$$\frac{\mathcal{P}(w_t > b)}{\mathcal{P}(w_t > b)} = \mathcal{P}(w_{t-T_b} > 0) = \frac{1}{2}$$

and we get the relation

$$\mathcal{P}(T_b < t) = 2\mathcal{P}(w_t > b)$$

Given the probability distribution $\rho(x,t)$ we get

$$\mathcal{P}(w_t > b) = \int_b^\infty \rho(x,t)dx$$

which implies for the probability density

$$p_b(t) = 2\int_b^\infty \frac{\partial \rho}{\partial t}(x,t)dx = -D\partial\rho\partial x(b,t)$$

Therefore the probability density of the first passage time al b is two times the density current at b.
In the case of an absorbing barrier at $x = 0$ one can solve the FP equation with open boundary condition

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} a(|x|)\rho + \frac{D}{2}\frac{\partial^2 \rho}{\partial x^2}$$

(the solution at $x = 0$ is computed by continuity) and the required solution is the difference between two solutions with symmetric initial conditions with respect to the barrier. The probability flow at the barrier is twice the flow from the right at $x = 0$

$$\frac{dp}{dt} = -2\left(a(x)\rho_+ + \frac{D}{2}\frac{\partial \rho_+}{\partial x}\right)$$

When $a(|x|) > 0$ we can use a stationary solution to estimate the current

$$a(x)\rho_s + \frac{D}{2}\frac{\partial \rho_s}{\partial x} = 0 \qquad \Rightarrow \qquad \rho_s(x) \propto \exp\left(-\frac{2V(x)}{D}\right)$$

where

$$V(x) = \int a(x)dx$$

and the normalization constant change in time according to $1 - p(t)$: this approximation is correct if the exit flow is sufficiently small to allow a relaxation of the distribution.

E.1 Time ordering operator and averaging

Let $\xi(s)D(s)$ a stochastic processes of linear time dependent operators, we consider the linear equation

$$\dot{x} = \xi(t)D(t)x$$

whose formal solution reads

$$x(t) = \mathcal{T}\exp\left(\int_0^t \xi(s)D(s)ds\right)x_0$$

where $\mathcal{T}$ is the time ordering operator

$$\mathcal{T}D(t_1)D(t_2) = \begin{cases} D(t_1)D(t_2) & \text{if} \quad t_1 \geq t_2 \\ D(t_2)D(t_1) & \text{if} \quad t_2 \geq t_1 \end{cases}$$

Indeed

$$\mathcal{T}\left(\int_0^t D(s;\xi)ds\right)^n = \sum_{(j_1,\dots,j_n)} \int_0^t ds_1 \int_0^{s_1} ds_2 .... \int_0^{s_{n-1}} ds_n \xi(s_1)....\xi(s_n)\mathcal{T}D(s_{j_1})...D(s_{j_n})$$

$$= \sum_{(j_1,\dots,j_n)} \int_0^t ds_1 \int_0^{s_1} ds_2 .... \int_0^{s_{n-1}} ds_n \xi(s_1)....\xi(s_n)D(s_1)...D(s_n)$$

where the sum runs over all the possible permutation of the indexes $(1,...,n)$, and we get

$$\mathcal{T}\exp\left(\int_0^t \xi(s)D(s)ds\right) = \sum_{n\geq 0} \int_0^t ds_1 \int_0^{s_1} ds_2 .... \int_0^{s_{n-1}} ds_n \xi(s_1)....\xi(s_n)D(s_1)...D(s_n)$$

The averaging of the evolution operator gives the expression

$$\sum_n \frac{1}{n!}\left\langle \mathcal{T}\left(\int_0^t \xi(s)D(s)ds\right)^n\right\rangle$$

and it is possible to proof that

$$\left\langle \mathcal{T}\left(\int_0^t \xi(s)D(s)ds\right)^n\right\rangle = \mathcal{T}\left\langle \left(\int_0^t \xi(s)D(s)ds\right)^n\right\rangle \tag{E.5}$$

We proceed using an explicit computation

$$\left\langle \mathcal{T}\prod_{j=1}^n \left(\int_0^t \xi(s_j)D(s_j)ds_j\right)\right\rangle = n! \int_0^t ds_1 .... \int_0^{s_{n-1}} ds_{n-1} \left\langle \xi(s_1)...\xi(s_n)\right\rangle D(s_1)...D(s_n)$$

and analogously

$$\begin{aligned}\mathcal{T} \quad & \left\langle \prod_{j=1}^n \left(\int_0^t \xi(s_j)D(s_j)ds_j\right)\right\rangle \\ = \quad & \sum_{(j_1,\dots,j_n)} \int_0^t ds_1 .... \int_0^{s_{n-1}} ds_{n-1} < \xi(s_1)....\xi(s_n) > D(s_1)...D(s_n)\end{aligned} \tag{E.6}$$

The equation (**??**) follows.

There is a strict relation between Gaussian processes and the linear Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2} D_{ij}(t) \frac{\partial^2 \rho}{\partial x_i \partial x_j}$$

where $D(t)$ is a positive defined symmetric matrix. Let

$$\Sigma(t) = \int_0^t D(s)ds \tag{E.7}$$

the solution can be written in the form

$$\rho(x,t) = \frac{1}{(2\pi)^{n/2}\sqrt{\det\Sigma(t)}} \exp\left(-\frac{(x-x_0)\Sigma^{-1}(t)(x-x_0)}{2}\right) \tag{E.8}$$

To prove the assertion one considers the stochastic differential equation

$$dx = \sqrt{D(t)}dw$$

where dw are differential of Wiener processes and $\sqrt{D(t)}$ is well defined since $D(t)$ is symmetric positive defined. It follows that $x(t)$ is a Gaussian process since it is a linear transformation of the Wiener process and, using the Ito formula, its covariance $c_{ij=} < x_i x_j >$ satisfies the equation

$$dc_{ij} =< dx_i dx_j >= D_{ij}(t)$$

Then we recover eq. (**??**) and the distribution (**??**) is the probability density of $x(t)$ with the initial condition $x(0) = x_0$. However the direct substitution of (**??**) in the Fokker-Planck equation shows non-trivial identities: we have

$$\frac{1}{2}D_{ij}(t)\frac{\partial^2\rho}{\partial x_i \partial x_j} = \frac{1}{2}x\Sigma^{-1}D\Sigma^{-1}x\rho - \frac{1}{2}\mathrm{Tr}(D\Sigma^{-1})\rho$$

and

$$\frac{\partial\rho}{\partial t} = -\frac{1}{2}xD^{-1}x\rho - \frac{1}{\det\Sigma}\frac{d}{dt}\det\Sigma$$

The first terms coincide since

$$\frac{d}{dt}\Sigma\Sigma^{-1} = D\Sigma^{-1} + \Sigma D^{-1} = 0$$

that implies

$$\Sigma^{-1}D\Sigma^{-1} = -D^{-1}$$

To compute $\mathrm{Tr}(D\Sigma^{-1})$ we use the relation $\Sigma(t) = O^T(t)\Lambda(t)O(t)$ where Λ is a diagonal matrix and $O(t)$ an orthogonal matrix, then

$$D(t) = \frac{dO^T}{dt}(t)O(t)\Sigma(t) + \Sigma(t)O^T(t)\frac{dO}{dt}(t) + O^T(t)\frac{d\Lambda}{dt}(t)O(t)$$

and we recall that

$$A(t) = \frac{dO^T}{dt}(t)O(t)$$

is an antisymmetric matrix. Finally we have

$$\mathrm{Tr}(D\Sigma^{-1}) = \mathrm{Tr}(A(t)\Sigma(t)\Sigma^{-1}(t)) - \mathrm{Tr}\Sigma^{-1}(t)\Sigma(t)A(t) + \mathrm{Tr}O^T(t)\frac{d\Lambda}{dt}(t)O(t)\Sigma^{-1}$$

so that since $\mathrm{Tr}A(t) = 0$ we have

$$\mathrm{Tr}(D\Sigma^{-1}) = \mathrm{Tr}\left(O^T(t)\frac{d\Lambda}{dt}(t)\Lambda^{-1}O(t)\right) = \frac{d\Lambda}{dt}(t)\Lambda^{-1} = \sum_i \frac{1}{\lambda_i}\frac{d\lambda_i}{dt}$$

This expression has to be compared with

$$\frac{1}{\det\Sigma}\frac{d}{dt}\det\Sigma = \frac{d}{dt}\log\det\Sigma = \frac{d}{dt}\sum_i \log(\lambda_i(t))$$

and the result follows.
The usual form of Gaussian distribution is related to the boundary conditions at $\pm\infty$. In the case of an angle variable on a torus one has to impose periodic boundary conditions and we get

$$\rho(\theta) = \frac{1}{\sqrt{2\pi\sigma}} \sum_{k=-\infty}^{\infty} \exp\left(-\frac{(\theta - 2\pi k)^2}{2\sigma}\right)$$

The normalizing condition requires

$$\int_0^{2\pi} \rho(\theta) d\theta = \frac{1}{\sqrt{2\pi\sigma}} \sum_{k=-\infty}^{\infty} \int_{2\pi k}^{2\pi(k+1)} \exp\left(-\frac{\theta^2}{2\sigma}\right) d\theta = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \exp\left(-\frac{\theta^2}{2\sigma}\right) d\theta = 1$$

In this way one recover the fundamental solution of the diffusion equation

$$\frac{\partial \rho}{\partial t} = \frac{D}{2} \frac{\partial^2 \rho}{\partial \theta^2} \qquad \theta \in [0 : 2\pi]$$

on the 1D torus with initial condition $\delta_{2\pi}(\theta)$

$$\rho(\theta, t) = \frac{1}{\sqrt{2\pi Dt}} \sum_{k=-\infty}^{\infty} \exp\left(-\frac{(\theta - 2\pi k)^2}{gt}\right)$$

An alternative approach considers the Fourier development of the solution

$$\rho(\theta, t) = \sum_{k=-\infty}^{k=\infty} c_k(t) \cos(k\theta)$$

that inserted in the diffusion equation gives

$$\frac{dc_k(t)}{dt} = -\frac{Dk^2}{2} c_k(t) \qquad c_k(t) = \frac{1}{2\pi} \exp\left(-\frac{Dk^2}{2} t\right)$$

according to the initial condition. Then we get

$$\rho(\theta, t) = \sum_{k=-\infty}^{k=\infty} \exp\left(-\frac{Dk^2}{2} t\right) \cos(k\theta)$$

which implies the remarkable identity

$$\frac{1}{\sqrt{2\pi Dt}} \sum_{k=-\infty}^{\infty} \exp\left(-\frac{(\theta - 2\pi k)^2}{2Dt}\right) = \frac{1}{2\pi} \sum_{k=-\infty}^{k=\infty} \exp\left(-\frac{Dk^2}{2} t\right) \cos(k\theta)$$

The distribution tends exponentially to a stationary constant distribution.
An interesting example is a skew Hamiltonian system of the form

$$\begin{aligned} d\theta &= \alpha I dt \\ dI &= \epsilon dw_t \end{aligned}$$

The solution is a Gaussian distribution with covariance matrix

$$\Sigma(t) = \epsilon^2 \begin{pmatrix} \alpha^2 t^3/3 & \alpha t^2/2 \\ \alpha t^2/2 & t \end{pmatrix}$$

with periodic boundary conditions. It can be written in the form

$$\rho(\theta, I, t) = \frac{1}{\sqrt{2\pi \mathrm{det}\Sigma(t)}} \sum_{k=-\infty}^{\infty} \exp\left[-\frac{6}{\alpha^2\epsilon^2 t^3} \left(\theta - 2\pi k - \frac{\alpha I}{2} t \right)^2 - \frac{I^2}{2\epsilon^2 t} \right]$$

where we set the initial condition $I(0) = 0$. The corresponding F-P equation reads

$$\frac{\partial \rho}{\partial t} = -\alpha I \frac{\partial \rho}{\partial \theta} + \frac{\epsilon^2}{2} \left(t \frac{\partial^2 \rho}{\partial I^2} + \alpha t^2 \frac{\partial^2 \rho}{\partial \theta \partial I} + \frac{\alpha^2 t^3}{3} \frac{\partial^2 \rho}{\partial I^2} \right)$$

We introduce the diffusion time $\tau = \epsilon^2 t$ so that the previous equation reads

$$\frac{\partial \rho}{\partial \tau} = -\frac{\alpha}{\epsilon^2} I \frac{\partial \rho}{\partial \theta} + \frac{1}{2} \left(\frac{\partial^2 \rho}{\partial I^2} + 2\frac{\alpha}{\epsilon^2} \tau \frac{\partial^2 \rho}{\partial \theta \partial I} + \frac{\alpha^2}{\epsilon^4} \tau^2 \frac{\partial^2 \rho}{\partial \theta^2} \right)$$

whose solution is

$$\rho(\theta, I, \tau) = \frac{1}{\sqrt{2\pi \mathrm{det}\Sigma(\tau)}} \sum_{k=-\infty}^{\infty} \exp\left[-\frac{6\epsilon^4}{\alpha^2 \tau^3} \left(\theta - 2\pi k - \frac{\alpha I}{2\epsilon^2} \tau \right)^2 - \frac{I^2}{2\tau} \right]$$

In the diffusion scale time ($\epsilon^2/\alpha \ll 1$) the solution is well approximated by

$$\rho(I, \tau) \frac{1}{\sqrt{2\pi\tau}} = \exp\left(-\frac{I^2}{2\tau} \right)$$

since the angle relaxes quickly to a uniform distribution (the variance is proportional to α^2/ϵ^4). This result proves an averaging principle for the F-P equation which admits a solution $\rho(\theta, I, \tau) = \rho(I, \tau)$ in the diffusion time scale.

Appendix F

Fokker-Planck equation and Stochastic dynamical systems

The statistical properties of stochastic dynamical systems

$$\dot{x} = a(x) + \sigma(x)\xi(t)$$

that satisfy the stochastic by a Fokker-Planck equation in the Stratonovich interpretation

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x} a(x)\rho + \frac{1}{2}\frac{\partial}{\partial x}\sigma(x)\frac{\partial}{\partial x}\sigma(x)\rho \tag{F.1}$$

performing the white noise limit for the fluctuation $\xi(t)$. In a multidimensional case the FP equation (**??**) reads

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x_k} a_k(x)\rho + \frac{1}{2}\frac{\partial}{\partial x_k}\sigma_{kj}(x)\frac{\partial}{\partial x_h}\sigma_{hj}(x)\rho$$

We consider the particular case when the differential form

$$dy_k = \sigma_{kh}(x)dx_h$$

is exact (at least closed). Then we have the condition

$$\frac{\partial \sigma_{kh}}{\partial x_l} - \frac{\partial \sigma_{kl}}{\partial x_h} = 0$$

In such a case we can perform a change of variables ************** We introduce a change of variable $x = x(y)$ such that

$$\frac{dx}{dy} = \sigma(x)$$

Remark: $\sigma(x) \neq 0$ is an assumption to avoid singularity in the solution $\rho(x,t)$. Then we have the relation

$$\rho(x,t)dx = \rho(x,y)\sigma(x)dy = \hat{\rho}(y,t)dy$$

and we write the FP equation in the form

$$\frac{\partial \hat{\rho}}{\partial t} = -\frac{\partial}{\partial y}\frac{a(y)}{\sigma(y)}\hat{\rho} + \frac{1}{2}\frac{\partial^2 \hat{\rho}}{\partial y^2} \qquad (C.2)$$

We consider the simple case $a(y) = 0$, then we have the Gaussian solution

$$\hat{\rho}(y,t) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-y_0)^2}{2t}\right)$$

and in the initial variables

$$\rho(x,t) = \frac{1}{\sqrt{2\pi\sigma^2(x)t}} \exp\left(-\frac{(y(x)-y_0)^2}{2t}\right)$$

where

$$y(x) = \int^x \frac{dx}{\sigma(x)}$$

Interestingly average value of y is constant $< y >= y_0$ and we have no drift term, but the average value $< x >$ can be computed from the formula

$$dx = \frac{dx}{dy}dy + \frac{1}{2}\frac{d^2x}{dy^2}dy^2$$

where from our assumptions we have $dy = dw_t$. Therefore we get the stochastic differential equation

$$dx = \sigma(x)dw_t + \frac{1}{2}\sigma(x)\frac{d\sigma}{dx}dt$$

The average dynamics reads

$$< \dot{x} >= \frac{1}{2}\left\langle \sigma(x)\frac{d\sigma}{dx}\right\rangle$$

and in the mean field approximation one gets

$$< \dot{x} >= \frac{1}{2}\sigma(< x >)\frac{d\sigma}{dx}(< x >)$$

A direct computation shows that

$$< \dot{x} >= \int x(y)\frac{\partial \hat{\rho}}{\partial t}(y,t)dy = -\frac{1}{2}\int \sigma(y)\frac{\partial \hat{\rho}}{\partial y}dy = \frac{1}{2}\int \sigma(x)\frac{d\sigma}{dx}\rho(x,t)dx$$

The mean field approximation is valid when the distribution is peaked around the mean value, the given the distribution $\hat{\rho}(y,t)$ which solves the FP equation (C.2) we can approximate the evolution as

$$\hat{\rho}(y,t+\Delta t) = \frac{1}{\sqrt{2\pi\Delta t}}\int \exp\left(-\frac{y-(\Phi^{\Delta t}(x))^2}{2\Delta t}\right)\hat{\rho}(x,t)dx$$

where $\Phi^t(x)$ is the phase flow associated to the average dynamics

$$< \dot{y} >= \frac{a(< y >)}{\sigma(< y >)}$$

(mean field approximation). Indeed for $\Delta t \ll 1$ we have

$$\Phi^{\Delta t}(x) \simeq x + \frac{a(x)}{\sigma(x)}\Delta t = x + \hat{a}(x)\Delta t$$

We recall that the Gaussian

$$\frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{(y-\bar{x})^2}{2\Delta t}\right)$$

is the solution of the FP equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2}\frac{\partial^2 \rho}{\partial y^2}$$

for any value $\bar{x}$ with initial condition $\lim_{\Delta t \to 0} \rho = \delta(\bar{x} - x)$, therefore we get

$$\frac{1}{\sqrt{2\pi\Delta t}} \exp\left(-\frac{(y-\bar{x})^2}{2\Delta t}\right) \simeq \left[\delta((y-\bar{x}) + \frac{1}{2}\frac{\partial^2}{\partial y}\delta((y-\bar{x})\Delta t\right]$$

Let $\bar{x} = \Phi^{\Delta t}(x) \simeq x + \hat{a}(x)\Delta t$, then we have to evaluate the integral

$$\int \delta(y - x - \hat{a}(x)\Delta t)\left[1 + \frac{1}{2}\frac{\partial^2}{\partial x^2}\Delta t\right]\hat{\rho}(x,t)dx =$$
$$\int \delta((y-u)\left[\hat{\rho}(u - \hat{a}(u)\Delta t, t) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\Delta t\hat{\rho}(u,t)\right]\left(1 - \frac{\partial \hat{a}}{\partial x}(u)\Delta t\right) du$$

where $u = x + a(x)\Delta t$. Performing the integral, the final results is

$$\hat{\rho}(y, t+\Delta t) = \hat{\rho}(y,t) - \frac{\partial \hat{\rho}}{\partial y}\hat{a}(y)\Delta t - \hat{\rho}(y,t)\frac{\partial \hat{a}}{\partial y}\Delta t + \frac{1}{2}\frac{\partial^2 \hat{\rho}}{\partial y^2}\Delta t$$

so that performing the limit $\Delta t \to 0$ we recover the FP equation (C.2). In the mean field approximation can be applied, an approximate solution of the FP equation (C.2) is written in the form

$$\hat{\rho}(y,t) = \frac{1}{\sqrt{2\pi t}} \int \exp\left(-\frac{y - (\Phi^t(x))^2}{2t}\right)\hat{\rho}(x)dx$$

This approximation may be applied if the drift term $a(x)/\sigma(x)$ is small or the effect of diffusion is limited. A better approximation is achieved by considering a local linear approximation of the drift term

$$\frac{a(x)}{\sigma(x)} \simeq c_0(x_0) + c_1(x_0)(x - x_0)$$

Then the solution is Gaussian with a local variance $\sigma(t; x_0)$

$$\frac{d\sigma}{dt} = c_1(x_0)\sigma(t) + 1 \qquad \Rightarrow \qquad \frac{2}{c_1(x_0)}(e^{c_1 t} - 1)$$

and the formula (**??**) changes as

$$\hat{\rho}(y,t) = \frac{1}{\sqrt{2\pi t}} \int \exp\left(-\frac{y - (\Phi^t(x))^2}{2\sigma(t;x)}\right)\hat{\rho}(x)dx \qquad \text{(F.2)}$$

F.1 Change of variables for the Laplacian operator: polar coordinates

We consider a 2D Laplacian operator of the form

$$\frac{1}{2}\frac{\partial}{\partial r_i}\sigma_{ij}D(r\Sigma^{-1}r)\frac{\partial}{\partial r_j} \tag{F.3}$$

where $\Sigma = \sigma_{ij}$ is a positive defined symmetric matrix and

$$r\Sigma^{-1}r = \sum_{hk} r_h \sigma^{-1}_{hk} r_k$$

By performing an orthogonal transformation we can diagonalize the matrix Σ and the operator (**??**) reads

$$\frac{1}{2}\sum_i \frac{\partial}{\partial r_i}\sigma_i^2 D\left(\frac{r_1^2}{\sigma_1^2}+\frac{r_2^2}{\sigma_2^2}\right)\frac{\partial}{\partial r_i} \tag{F.4}$$

Then we scale the variables $r_i \to r_i/\sigma_i$ and the new operator (**??**) is

$$\frac{1}{2}\sum_i \frac{\partial}{\partial r_i}D(r_1^2+r_2^2)\frac{\partial}{\partial r_i} \tag{F.5}$$

We perform the change of variables

$$J = \frac{r_1^2+r_2^2}{2} \qquad \theta = \text{atan}\,\frac{r_1}{r_2}$$

on the Langevin dynamics associated to the Fokker-Planck equation with Laplacian operator (**??**)

$$dr_i = \frac{1}{2}\frac{\partial D}{\partial r_i}dt + \sqrt{D(J)}dw_i \tag{F.6}$$

By a direct calculation using the Ito formula, we get

$$\begin{aligned} dJ &= r_1 dr_1 + r_2 dr_2 + D(I)dt \\ d\theta &= \frac{r_2}{2J}dr_2 - \frac{r_1}{2J}dr_2 \end{aligned}$$

and the FP operator reads

$$-\frac{1}{2}\frac{\partial}{\partial J}\left(r_1\frac{\partial D}{\partial r_1}+r_2\frac{\partial D}{\partial r_2}+2D(J)\right)\rho - \frac{1}{2}\frac{\partial}{\partial\theta}\left(\frac{r_2}{2J}\frac{\partial D}{\partial r_1}-\frac{r_1}{2J}\frac{\partial D}{\partial r_2}\right)\rho$$
$$+\frac{\partial^2}{\partial J^2}JD(J)\rho + \frac{1}{4}\frac{\partial^2}{\partial\theta^2}\rho\frac{D(J)}{J}$$

According to the definitions we have in the drift term

$$\frac{I_1^2}{2}\frac{\partial D}{\partial J}+\frac{I_2^2}{2}\frac{\partial D}{\partial J}+D(J) = \frac{\partial}{\partial J}JD(J)$$

and the angular drift term cancels. We obtain a FP equation of the form

$$\frac{\partial\rho}{\partial\tau} = -\frac{\partial}{\partial J}\frac{\partial}{\partial J}(JD(J))\rho + \frac{\partial^2}{\partial J^2}JD(J)\rho + \frac{D(J)}{4J}\frac{\partial^2\rho}{\partial\theta^2}$$

that can be averaged on the angle variable

$$\frac{\partial <\rho>}{\partial\tau} = \frac{\partial}{\partial J}JD(J)\frac{\partial <\rho>}{\partial J} \tag{F.7}$$