Stochastic variational principles for dissipative and conservative systems

Francesco Guerra
Dipartimento di Fisica
Università di Roma "La Sapienza"
Istituto Nazionale di Fisica Nucleare
Sezione di Roma
Piazzale A. Moro 5, 00185 Roma, Italy

Granada, September 2012

In the frame of the method of statistical ensembles, relaxation to thermodynamic equilibrium can be efficiently described by using stochastic differential equations, according to the Langevin scheme.

This scheme is intrinsically non-invariant with respect to time reversal.

However, we can consider an additional dynamical variable, called the importance function, whose meaning and motivation arises from the neutron diffusion theory.

The resulting scheme is now time reversal invariant, with a complete symplectic structure. The equations for the density and the importance function are Hamilton equations for a properly chosen Hamiltonian, and obey stochastic variational principles.

On the other hand, we can consider the formulation of quantum mechanics, according to the stochastic scheme devised by Edward Nelson many years ago, and further developed.

In this frame a stochastic variational principle can be easily introduced, by an extension of the Lagrange principle of classical mechanics.

The theory is intrinsically time reversal invariant.

We have still a symplectic structure, and canonical Hamilton equations. Here the conjugated variables are the quantum mechanical probability density and the phase of the wave function.

The purpose of the talk is to give a synthetic description of the two schemes, by pointing out their structural similarity, and their deep physical difference.

We limit our exposition to the main ideas, by keeping to a minimum all technical details.

The (Guerra-Morato) variational principle of Nelson Stochastic Mechanics

We find convenient to reverse the order and start from stochastic mechanics.

This is a formulation of quantum mechanics in terms of stochastic processes, completely equivalent to the operatorial formulation.

There is no "stochastic interpretation" of quantum mechanics.

We consider a mechanical system, for example a massive particle in a potential, and call x the points in configuration space. For example, $x = (x_1, x_2, x_3) \in \mathbb{R}^3$.

We consider a class of stochastic processes. Call $q(t)$ the position of one of these processes, evolving in time t. Each has a density $\rho(x, t)$, so that for the expectations we have

$$\mathbb{E}(F(q(t)) = \int dx F(x)\rho(x, t),$$

for any smooth function F.

It is assumed that the time evolution of $q(t)$ is given by a stochastic differential equation, which I write in the following form, in order to avoid any misunderstanding.

Starting from time t_0, for a given small time increment $\Delta t > 0$, consider the sequence of times $t_0, t_0 + \Delta t, t_0 + 2\Delta t, \ldots$.

Then, the corresponding random increments for $q(t)$ are $\Delta q(t) = q(t + \Delta t) - q(t)$. They evolve under the influence of some guiding field $v(x, t)$, the drift field, which should be

considered as a dynamical variable of the theory, replacing in a sense the velocity, or momentum, variables of classical mechanics.

There is also a disturbing field $w(t)$, whose increments in each small time interval, $\Delta w(t) = w(t+\Delta t) - w(t)$, have centered Gaussian distribution, independent in each time interval, with variance

$$\mathbb{E}(\Delta w_i(t) \Delta w_j(t)) = 2\nu \delta_{ij} \Delta t,$$

where ν is a diffusion coefficient, to be fixed in the following.

I write the evolution of $q(t)$ in the form

$$\Delta q(t) = v(q(t), t)\Delta t + \Delta w(t) + \mathcal{O}((\Delta t)^{\frac{3}{2}}).$$

The symbols $\mathcal{O}((\Delta t)^p)$, with p integer or half-integer, obey the property

$$\mathcal{O}((\Delta t)^p)\mathcal{O}((\Delta t)^{p'}) = \mathcal{O}((\Delta t)^{p+p'}).$$

They have the following meaning. We say that a random variable $a(\Delta t)$ is $\mathcal{O}((\Delta t)^p)$ if for its averages, or conditional averages, we have

$$\mathbb{E}(a(\Delta t)) = \mathcal{O}((\Delta t)^{p+\epsilon}),$$

where $\epsilon = 0$ if p is integer, $\epsilon = \frac{1}{2}$ if p is half-integer. There is always a rounding up to the next integer.

In the previous formula, $\mathcal{O}((\Delta t)^p)$, p integer, for non-random variables, has the usual meaning. For a non-random $a(\Delta t)$,

$$a(\Delta t) = \mathcal{O}((\Delta t)^p)$$

means

$$|a(\Delta t)| \leq c(\Delta t)^p,$$

for some constant c, and all Δt small enough.

Therefore, we see that $v(q(t),t)\Delta t = \mathcal{O}(\Delta t)$, while $\Delta w(t) = \mathcal{O}((\Delta t)^{\frac{1}{2}})$.

Therefore, the meaning of the evolution

$$\Delta q(t) = v(q(t),t)\Delta t + \Delta w(t) + \mathcal{O}((\Delta t)^{\frac{3}{2}})$$

is clear. Starting from a $q(t_0)$, distributed according to a given $\rho(.,t_0)$, by neglecting the $\frac{3}{2}$ term, we can construct recursively the positions at the mesh of times $t_0, t_0+\Delta t, t_0+2\Delta t, \ldots$.

The process we are interested in will be obtained in the limit $\Delta t \to 0$, as a rigorous

analysis, not reported here, would show. It is necessary to assume some smoothness for $v(.,.)$. But also singular v's are allowed, through a suitable limiting procedure.

We have written the time evolution in the quite peculiar form

$$\Delta q(t) = v(q(t),t)\Delta t + \Delta w(t) + \mathcal{O}((\Delta t)^{\frac{3}{2}}).$$

In the mathematical literature, it would be written as Ito stochastic differential equation

$$dq(t) = v(q(t),t)dt + dw(t).$$

This is perfectly all right. However, the stochastic differential equation can not be taken to the square. In fact

$$(dq(t))^2 = (v(q(t),t)dt + dw(t))^2$$

is completely wrong.

In fact, by taking the square in the equation with the $\frac{3}{2}$ correction, we see that there is a resonance between the $\frac{3}{2}$ term and the $\frac{1}{2}$ term. Therefore in evaluating for example $\Delta q(t).\Delta q(t)$, up to the order $(\Delta t)^2$, we

must take into account the $\frac{3}{2}$ terms, easily calculated in explicit form.

Now we build the stochastic variational principle.

In the classical case, we would consider a Lagrange variational principle. I give the following version.

Consider a given time interval $[t_0, t_1]$. Assume a trial velocity $v : t \to v(t)$, acting as

control on the position, so that

$$\dot{q}(t) = v(t),\ q(t) = q(t_0) + \int_{t_0}^{t} v(t)dt,$$

for some initial position $q(t_0)$. Introduce the action functional

$$\mathcal{A}(v, q(t_0)) =$$

$$\int_{t_0}^{t_1} \Big(\frac{1}{2} m v(t)^2 - V(q(t)\Big) dt.$$

Consider variations $\delta v(t)$, and *fixed* initial position for q. Then we have

$$\delta q(t) = \int_{t_0}^{t} \delta v(t')dt',$$

and

$$\delta\mathcal{A}(v, q(t_0)) =$$

$$\int_{t_0}^{t_1}\Big(mv(t)\delta v(t) - (\nabla V)(q(t))\int_{t_0}^{t}\delta v(t')dt'\Big)dt =$$

$$\int_{t_0}^{t_1}\Big(mv(t) + \int_{t}^{t_1}(-\nabla V)(q(t'))\ dt'\Big)\delta v(t)\ dt.$$

The variational principle, in our version, requires stationarity of the action under variations of the trial velocity vanishing at t_1.

Then we have

$$mv(t) + \int_t^{t_1} (-\nabla V)(q(t'))\ dt' = mv(t_1),$$

where $mv(t_1)$ is the Lagrange multiplier associated to the constraint $\delta v(t_1) = 0$.

Immediately we recognize that the optimal v obeys Newton equations, under the force $(-\nabla V)$, *i.e.*

$$m\dot{v}(t) = (-\nabla V)(q(t)).$$

Notice that the boundary conditions are the position at the initial time, and the velocity at the final time.

In the usual formulation, there is an integration by part, and the boundary conditions are fixed position at the initial and the final time.

But the two formulations are physically equivalent.

In any case we produce the dynamical content of classical mechanics through the Lagrange variational principle.

In the stochastic case, the procedure is essentially the same, but mediated with respect the stochastic nature of the process. So we assume the drift field $v(.,.)$ as a trial field.

Our trial Lagrangian functional would be the following, as a simple generalization of the classical case.

$$\int_{t_0}^{t_1} \mathbb{E}\Big(\frac{1}{2}m(\frac{\Delta q}{\Delta t})^2 - V(q(t)\Big)dt,$$

to be made stationary under small variations of the drift field, by keeping for example fixed the densities at the initial and final times.

Of course, there is a difficulty, because $\frac{(\Delta q)^2}{(\Delta t)^2}$ is infinite in the limit $\Delta t \to 0$. However, in a very beautiful paper, Nelson has shown that one can subtract an infinite constant, independent from the trial field, and get the variational principle, exactly in the form guessed by Guerra-Morato.

Let us write $v(x,t)$ in the form $v_+(x,t)$, in order to recall that the evolution in $q(t)$ is

ruled by $v_+(x,t)$ in the given construction, with t continuously increasing by steps $\Delta t > 0$.

Of course, we have

$$v_+(x,t) =$$

$$\Delta t^{-1}\mathbb{E}(q(t+\Delta t) - q(t)|q(t) = x) + \mathcal{O}((\Delta t)),$$

where we have conditioned with respect to the starting position.

But we can also condition with respect to the ending position, and define, always for

$\Delta t > 0$,

$$v_-(x,t) =$$

$$\Delta t^{-1}\mathbb{E}(q(t) - q(t-\Delta t)|q(t) = x) + \mathcal{O}((\Delta t)).$$

It is easy to express $v_-(x,t)$ in terms of $v_+(x,t)$ and the density (Bayes rule).

We can also consider the time inverted process $t \to t' = -t, q \to q', q(t) = q'(t')$ (recall that q is a position!).

The time inverted process will satisfy a stochastic evolution, forward in the (inverted) time, analogous to that for the original process, but now driven by the $v_-(x,t)$ field.

The drift fields v_+ and v_- transform into each other under time reversal.

The phenomenological choice of Guerra-Morato for the Lagrangian, fully confirmed by the treatment given by Nelson, is the symmetric one

$$\int_{t_0}^{t_1} \mathbb{E}\Big(\frac{1}{2} m v_+(q(t),t) v_-(q(t),t) - V(q(t)\Big) dt.$$

The stochastic variational principle requires that the functional must be stationary under small variations of the drift field, with fixed densities at the time boundaries.

It is quite surprising to see that as a consequence of the variational principle the Schrödinger

equation emerges, provided the diffusion constant ν is chosen appropriately in function of the Planck constant.

The phase of the wave function appears as conjugated variable of the density. The whole structure is canonical. An Hamiltonian, as a functional of the density and the phase, can be easily introduced. A symplectic structure is present, as a generalization of the symplectic structure of classical mechanics.

The time evolution for the density and the phase are Hamilton equations in this formulation.

The whole scheme is time reversal invariant. This is due to the fact that drifts are to be considered as dynamical variables, subject to the transformations of time reversal.

The transformation $v_+ \to v_-$ is the stochastic generalization of the classical one for the velocity $v \to -v$.

The whole scheme can be generalized to the case where the configuration space is a Riemann manifold. Then there is a very nice blend between stochastic differential equations and differential geometry.

Also the generalization to systems with an infinite number of degrees of freedom (stochastic field theory) is very interesting. Here we find a connection between stochastic mechanics, in the form of stochastic field theory, and Euclidean Quantum Field Theory.

In particular, for interacting boson fields, the process associated to the ground state is the same as the process constructed in Euclidean Quantum Field theory, through a Wick rotation to imaginary time, starting from Wightman Quantum Fields.

In the stochastic formulation no Wick rotation is necessary, and one can work directly in the physical space-time. A result very difficult to swallow.

In a forthcoming paper, we will show the historical and conceptual road taking from

Fermi statistical theory of the electronic structure of the atom (the so called Thomas-Fermi statistical model) up to Nelson stochastic mechanics (through Feynes and Weizel).

Dissipative processes

Consider a mechanical system with phase space $(x, y) \in \Gamma = \mathbb{R}^n \times \mathbb{R}^n$. We denote by x the configurations, and by y the momenta. For the sake of simplicity, we assume an Hamiltonian of the form

$$H(x, y) = \frac{y^2}{2m} + V(x).$$

The physical trajectories

$$t \to (q(t), p(t)) \in \Gamma$$

obey the Hamilton equations

$$\dot{q}(t) = p(t)/m$$

$$\dot{p}(t) = -(\nabla_x V)(q(t)),$$

which are invariant under time reversal

$$t \to t' = -t,\ x \to x' = x,\ y \to y' = -y,$$

$$\nabla_x \to \nabla_{x'} = \nabla_x,\ \nabla_y \to \nabla_{y'} = -\nabla_y.$$

In the frame of the theory of statistical ensembles, if the system is in thermal con-

tact with a reservoir at absolute temperature T, it is assumed that $(q(t), p(t))$ are upgraded to stochastic processes, with distribution $\rho(x, y; t)$, evolving in time according to the Ito stochastic differential equations

$$dq(t) = \frac{p(t)}{m} dt,$$

$$dp(t) = -(\nabla_x V)(q(t))\ dt - \beta p(t) dt + dw(t),$$

where $dt > 0$, $dw(t) = w(t + dt) - w(t)$ does not depend on $(q(t), p(t))$. It is centered, with covariance

$$\mathbb{E}_t(dw(t) dw(t)) = 2\nu I dt.$$

Vector indices are suppressed in the notation, I is the identity matrix, and ν a properly chosen diffusion coefficient.

Now we have that the usual Hamilton equations have been complemented by a friction term, dependent on a parameter β, and a fluctuating term ruled by ν.

If the following fluctuation-dissipation relation is satisfied

$$\nu = Km\beta T,$$

then the dynamical evolution satisfies the second principle of thermodynamics (free energy never increasing), and the system relaxes to the Boltzmann equilibrium distribution

$$\bar{\rho}(x,y) = Z^{-1} \exp(-\frac{H(x,y)}{KT}),$$

where K is the Boltzmann constant, and T the absolute temperature of the thermal reservoir.

Notice that in the stochastic differential equation (Langevin equations) the drift field, act-

ing on $dp(t)$, reflects a law of nature, by describing the action of the thermostat on the system. This is in marked contrast with the scheme of stochastic mechanics, where drifts have the status of dynamical variables.

Now the evolution is intrinsically irreversible. In fact, under time reversal $q'(t') = q(t), p'(t') = -p(t)$, we have the reversal of the friction term, with absurd consequences.

The time inverted process no longer satisfies Langevin equation.

It looks like there is an arrow of time.

However, we can enlarge the scheme, and reach a situation where time reversal is restored.

For this, we need to introduce the "importance" function, and an idea of smooth final conditioning on the process. These concepts are borrowed from the theory of neutron diffusion and nuclear interaction.

Consider the transition probability density for the process

$$p(x, y, t; x', y', t'),$$

for $t \geq t'$. This is the distribution of the process at time t, under the assumption that it starts from the point x', y', at a preceding time t'. It can be easily calculated through a Fokker-Planck equation.

Obviously, the density evolves according to

$$\rho(x, y, t) = \int p(x, y, t; x', y', t')\rho(x', y', t')dx'dy'.$$

Consider some interval of observation $[t_0, t_1]$. For a positive given function $F(x_1, y_1, t_1)$, at the final time t_1, let us define the adjoint function, for $t_0 \leq t \leq t_1$,

$$F(x, y, t) =$$

$$\int dx_1 dy_1 F(x_1, y_1, t_1) p(x_1, y_1, t_1; x, y, t)$$

(technically a martingale for the stochastic process).

Notice the different roles of the phase space integration in the two expressions of ρ and F.

The density ρ is propagated forward in time, the importance F is propagated backward.

We can define a smooth final conditioning on the process (q, p), by defining a new process $(\bar{q}, \bar{p})$, through a deformation on the measure on trajectories, such that for the averages we have

$$\mathbb{E}(G(\bar{q}(s_1), \bar{p}(s_1); \bar{q}(s_2), \bar{p}(s_2); ...)) =$$

$$\mathbb{E}(G(q(s_1), p(s_1); q(s_2), p(s_2); ...)F(q(t_1), p(t_1); t_1)),$$

for any smooth function G, and any sequence of times $s_1, s_2, \ldots$.

Each trajectory is sampled and receives a modification of its stochastic weight, according to the point reached at the final time.

It is immediate to check that $(\bar{q}, \bar{p})$ satisfies a stochastic differential equation analogous to that for (q, p), with a modification of the drift. The density is changed according to

$$\bar{\rho}(x, y; t) = \rho(x, y; t) F(x, y; t).$$

From this we see that F must satisfy an appropriate normalization condition, in order to

have a normalized $\bar{\rho}$. It is enough to impose the normalization at some time, then it will automatically hold at all times.

We can easily realize that a kind of miracle here appears as far as time reversal is concerned.

Consider the time reversal transformation, and build up the new $(\bar{q}, \bar{p})$ process. Then we can immediately check that this time inverted $(\bar{q}, \bar{p})$ process can be considered as the

smooth conditioning, for an appropriate importance function, of a process obeying the original Langevin equations.

The whole theory is time reversal invariant, even if dissipative.

The arrow of time seems to have disappeared.

Local stochastic variational principles for the density and the importance function

The density and the importance function obey local stochastic variational principles, which are very important in the applications, for example to find approximations for the dynamics. They are strictly connected with the physical properties of entropy production.

In order to give a flavor of these principles, we will work in a simplified scheme, by taking into account only the configuration space

(Einstein-Smoluchowsky frame). But everything can be generalized to full phase space.

Consider a stochastic differential equation in $\mathbb{R}^n$

$$dq(t) = v(q(t))dt + dw(t),$$

with

$$dw(t)dw(t) = 2\nu dt.$$

The importance function is now

$$F(x,t) = \int dx_1 F(x_1,t_1)p(x_1,t_1;x,t)$$

(integration at the final time).

For a generic control field v', introduce the corresponding stochastic process q'.

Define the trial functional

$$I(x,t;v') =$$

$$\mathbb{E}(-\int_t^{t_1} dt' \mathcal{L}(q'(t'),t';v') + \log F(q'(t_1),t_1)|q'(t)=x).$$

Here the averages are made with respect to all trajectories starting from x at time t.

The Lagrangian function is defined by

$$\mathcal{L}(x,t;v') = \frac{1}{4\nu}(v'-v)^2(x,t).$$

Therefore, it depends on how much the trial drift is different from the original drift.

An easy calculation gives immediately that the importance function, with the given final condition, is expressed through the optimal variational principle

$$\log F(x,t) = \sup_{v'} I(x,t;v').$$

Moreover, the optimal v' is precisely the drift of the smooth conditioned process.

Therefore, we see that stochastic variational principles and smooth conditioning go very well together.

A further expansion of the theory gives immediately a symplectic structure on functionals of the density and the importance.

Moreover, one can introduce an Hamiltonian functional, so that the time evolution equations for the density and the importance are Hamilton equations.

Conclusions

We have considered two basic dynamical schemes, ruled by stochastic differential equations.

One is related to Nelson stochastic mechanics, a stochastic formulation of quantum mechanics, completely equivalent to the operator formulation.

The second is related to relaxation of thermodynamic systems in a thermal bath.

Time reversal invariance is implemented in a completely different way in the two cases.

In Nelson stochastic mechanics, time reversal follows from the nature of the drift fields, which have the status of dynamical variables, with appropriate transformation properties under time reversal.

In the case of the dissipative processes, time reversal is implemented by taking into account the observational modifications given by the importance function.

These concepts are suggested by the theory of neutron diffusion and nuclear interaction.

In both schemes there is an emergent symplectic structure.

The conjugated variables are the density and the phase function in stochastic mechanics, and the density and the importance function in the dissipative case.

In both schemes time evolution can be written as Hamiltonian evolution, according to

the underlying respective symplectic structures.

There are also stochastic variational principles.

In stochastic mechanics, the Lagrange variational principle is the same as the classical one, provided that an irrelevant infinite term is discarded, in the spirit of renormalization theory.

In the dissipative case, the stochastic variational principles reflect the physical nature of entropy production.

These two schemes seems to comprehend a very large area of physical applications.