GEOMETRY OF IRREVERSIBILITY

Alexander N. Gorban
Institute of Computational Modeling, Russian Academy of Sciences
660036 Krasnoyarsk, Russia
Institut des Hautes Études Scientifiques
91440 Bures-sur-Yvette, France
gorban@icm.krasn.ru

Iliya V. Karlin
ETH Zürich, Institute of Polymers
CH-8092 Zürich, Switzerland
ikarlin@ifp.mat.ethz.ch

Abstract

A general geometrical setting of nonequilibrium thermodynamics is developed. The approach is based on the notion of the natural projection which generalizes Ehrenfests' coarse-graining. It is demonstrated how derivations of irreversible macroscopic dynamics from the microscopic theories can be addressed through a study of stability of quasi-equilibrium manifolds.

Keywords: Nonequilibrium thermodynamics, quasi-equilibrium, natural projection, stability.

Introduction

The goal of this paper is to discuss the (nonlinear) problem of irreversibility, and how the nonequilibrium thermodynamics attempts to solve it. This problem has been intensively discussed in the past, and nice accounts of these discussions can be found in the literature [1]. Here, we intend to develop a more geometrical viewpoint on the subject. The paper consists of two parts. First, in section 1, we discuss in an informal way the origin of the problem, and demonstrate how the basic constructions arise. Second, in section 2, we give a consistent geometric formalization of these constructions. Our presentation is based on the notion of the natural projection introduced therein. We discuss in detail the formalization of Ehrenfest's ideas of coarse-graining. The novel approach developed in sections 2.4, 2.5, and 2.6 allows to go beyond limitations of the short memory approximations through a study of stability of the quasi-equilibrium manifold.

1. The problem of irreversibility

1.1. The phenomenon of the macroscopic irreversibility

The best way to demonstrate the problem of irreversibility is the following *Gedankenexperiment*: Let us watch the movie: It's raining, people are running, cars rolling. Let us now wind this movie in the opposite direction, and we will see a strange and funny picture: Drops of the rain are raising up to the clouds, which next condensate into the vapor on the pools, on the surfaces of rivers, people run with their backs forward, cars behave also quite strange, and so forth. This cannot be, and we "know" this for sure, we have never seen something like this in our life. Let us now imagine that we watch the same movie with a magnitude of $10^8 - 10^9$ so that we can resolve individual particles. And all of the sudden we discover that we cannot see any substantial difference between the direct and the reverse demonstration: Everywhere the particles are moving, colliding, reacting according to the laws of physics, and nowhere there is a violation of anything. We cannot tell the direct progressing of the time from the reversed. So, we have the irreversibility of the macroscopic picture under the reversibility of the microscopic one.

Rain, people, cars - this all is too complicated. One of the most simple examples of the irreversible macroscopic picture under the apparent reversibility of the microscopic picture (the "thermal ratchet") is given by R. Feynman in his lectures on the character of physical law [2]. We easily label it as self-evident the fact that particles of different colors mix together, and we would see it as a wonder the reverse picture of a spontaneous decomposition of their mixture. However, itself an appreciation of one picture as usual, and of the other as unusual and wonderful - this is not yet the physics. It is desirable to measure somehow this transition from order to disorder.

1.2. Phase volume and dynamics of ensembles

Let there be n blue and n white particles in a box, and let the box is separated in two halves, the left and the right. Location of all the particles in the box is described by the assembly of $2n$ vectors of locations of individual particles. The set of all the assemblies is a "box" in the $6n$-dimensional space. A point in this $6n$-dimensional box describes a configuration. The motion of this point is defined by equations of mechanics.

"Order" is the configuration in which the blue particles are all in the right half, and all the white particles are in the left half. The set of all such configurations has a rather small volume. It makes only $(1/2)^{2n}$ of the total volume of the $6n$-dimensional box. If $n = 10$, this is of the order of one per million of the total volume. It is practically unthinkable to get into such a configuration by a chance. It is also highly improbable that, by forming more or less voluntary the initial conditions, we can observe that the system becomes ordered by itself. From this standpoint, the motion goes from the states of "order" to the state of "disorder", just because there are many more states of "disorder".

However, we have defined it this way. The well known question of where is there more order, in a fine castle or in a pile of stones, has a profound answer: It depends on which pile you mean. If "piles" are thought as all configurations of stones which are not castles, then there are many more such piles, and so there is less order in such a pile. However, if these are specially and uniquely placed stones (for example, a garden of stones), then there is the same amount of order in such a pile as in the fine castle. *Not a specific configuration is important but an assembly of configurations embraced by one notion.*

This transition from single configurations to their assemblies (ensembles) play the pivotal role in the understanding of irreversibility: The irreversible transition from the ordered configuration (blue particles are on the right, white particles are on the left) to the disordered one occurs simply because there are many more of the disordered (in the sense of the volume). Here, rigorously speaking, we have to add also a reference to the Liouville theorem: The volume in the phase space which is occupied by the ensemble does not change in time as the mechanical system evolves. Because of this fact, the volume V is a good measure to compare the assemblies of configurations. However, more often the quantity $\ln V$ is used, this is called the entropy.

The point which represents the configuration, very rapidly leaves the small neighborhood and for a long time (in practice, never) does not come back into it. In this, seemingly idyllic picture, there are still two rather dark clouds left. First, the arrow of time has not appeared. If we move from the ordered initial state (separated particles) backwards in time, then everything will stay the same as when we move forward in time, that is, the order will be changing into the disorder. Second, let us wind the film backwards, let us shoot the movie about mixing of colored particles, and then let us watch in the reverse order their demixing. Then the initial configurations for the reverse motion will only seem to be disordered. Their "order" is in the fact that they were obtained from the separated mixture by letting the system to evolve for the time t. There are also very few such configurations, just the same number as of the ordered (separated particles) states. If we start with these configurations, then we obtain the ordered system after the time t. Then why this most obvious consequence of the laws of mechanics looks so improbable on the screen? Perhaps, it should be accepted that states which are obtained from the ordered state by a time shift, and by inversion of particle's velocities (in order to initialize the reverse motion in time), cannot be prepared by using macroscopic means of preparation. In order to prepare such states, one would have to employ an army of Maxwell's daemons which would invert individual velocities with sufficient accuracy (here, it is much more into the phrase "sufficient accuracy" but this has to be discussed separately and next time).

For this reason, we lump the distinguished initial conditions, for which the mixture decomposes spontaneously ("piles" of special form, or "gardens of stones") together with other configurations into *macroscopically definable ensembles*. And already for those ensembles the spontaneous demixing becomes improbable. This way we come to a new viewpoint: (i). We cannot prepare individual systems but only representatives of ensembles. (ii) We cannot prepare ensembles at our will but only " macroscopically definable ensembles". What are these macroscopically definable ensembles? It seems that one has to give some constructions, the universality of which can only be proven by the time and experience.

1.3. Macroscopically definable ensembles and quasi-equilibria

The main tool in the study of macroscopically definable ensembles is the notion of the macroscopic variables, and of the quasi-equilibria. In the dynamics of the ensembles, the macroscopic variables are defined as linear functionals (moments) of the density distribution of the ensemble. Macroscopic variables M usually include hydrodynamic fields, density of particles, densities of momentum, and density of the energy, also the list may include stress tensor, reaction rates and other quantities. In the present context, it is solely important that the list the macroscopic variables is identified for the system under consideration.

A single system is characterized by a single point x in the phase space. The ensemble of the systems is defined by the probability density F on the phase space. Density F must satisfy a set of restrictions, the most important of which are: Nonnegativity, $F(x) \geq 0$, normalization,

$$\int_X F(x)dV(x) = 1, \tag{1}$$

and that the entropy is defined, that is, there exists the integral,

$$S(F) = -\int_X F(x)\ln F(x)dV(x). \tag{2}$$

(Function $F \ln F$ is continuously extended to zero values of F: $0 \ln 0 = 0$). Here $dV(x)$ is the invariant measure (phase volume).

The quasi-equilibrium ensemble describes the "equilibrium under restrictions". It is assumed that some external forcing keeps the given values of the macroscopic variables M, with this, "all the rest" comes the corresponding (generalized) canonic ensemble F which is the solution to the problem:

$$S(F) \to \max,\ M(F) = M. \tag{3}$$

where $S(F)$ is the entropy, $M(F)$ is the set of macroscopic variables.

Hypothesis about the macroscopically definable ensembles. Macroscopically definable ensembles are obtained as the result of two operations:

(i). Bringing the system into the quasi-equilibrium state corresponding to either the whole set of the macroscopic variables M, or to its subset.

(ii). Changing the ensemble according to the microscopic dynamics (due to the Liouville equation) during some time t.

These operations can be applied in the interchanging order any number of times, and for arbitrary time segments t. The limit of macroscopically definable ensembles will also be termed the macroscopically definable. One always starts with the operation (i).

In order to work out the notion of macroscopic definability, one has to pay more attention to partitioning the system into subsystems. This involves a partition of the phase space X with the measure dV on it into a direct product of spaces, $X = X_1 \times X_2$ with the measure $dV_1 dV_2$. To each admissible ("macroscopic") partition into sub-systems, it corresponds the operation of taking a "partial quasi-equilibrium", applied to some density $F_0(x_1, x_2)$:

$$\begin{aligned} &S(F) \to \max, \\ &M(F) = M,\ \int_{X_2} F(x_1, x_2)dV_2(x_2) = \int_{X_2} F_0(x_1, x_2)dV_2(x_2). \end{aligned} \tag{4}$$

where M is some subset of macroscopic variables (not necessarily the whole list of the macroscopic variables). In Eq. (4), the state of the first subsystem is not changing, whereas the second subsystem is brought into the quasi-equilibrium. In fact, the problem (4) is a version of the problem (3) with additional "macroscopic variables",

$$\int_{X_2} F(x_1, x_2) dV_2(x_2). \tag{5}$$

The extended hypothesis about macroscopically definable ensembles allows to use also operations (4) with only one restriction: The initial state should be the "true quasi-equilibrium" that is, macroscopic variables related to all possible partitions into subsystems should appear only after the sequence of operations has started with the solution to the problem (3) for some initial M. This does not exclude a possibility of including operators (5) into the list of the basic macroscopic variables M. The standard example of such an inclusion are few-body distribution functions treated as macroscopic variables in derivations of kinetic equations from the Liouville equation.

Irreversibility is related to the choice of the initial conditions. The extended set of macroscopically definable ensembles is thus given by three objects:

(i). The set of macroscopic variables M which are linear (and, in an appropriate topology, continuous) mappings of the space of distributions on the phase space onto the space of values of the macroscopic variables;

(ii). Macroscopically admissible partitions of the system into subsystems;

(iii). Equations of microscopic dynamics (the Liouville equation).

The choice of the macroscopic variables and of the macroscopically admissible partitions is a distinguished topic. Here they are represented as formal elements of the construction, and the arbitrariness is removed only at solving specific problems.

1.4. Irreversibility and initial conditions

The choice of the initial state of the ensemble plays the crucial role in the hypothesis about the macroscopically definable ensembles. The initial state is always taken as the quasi-equilibrium distribution which realizes the maximum of the entropy for given values of the macroscopic variables. The choice of the initial state splits the time axis into two semi-axes: moving forward in time, and moving backward in time, the observed non-order increases (the simplest example is the mixing of the particles of different colors).

In some works, in order to achieve the "true nonequilibrium", that is, the irreversible motion along the whole time axis, the quasi-equilibrium initial condition is shifted into $-\infty$ in time. This trick, however, casts some doubts, the major being this: Most if the known equations of macroscopic dynamics which describe irreversible processes have solutions which can be extended backwards in time only for finite times (or cannot be extended at all). Such equations as the Boltzmann kinetic equation, diffusion equation, equations of chemical kinetics and like do not allow for almost all their solutions to be extended backward in time for indefinitely long. All motions have a "beginning" beyond which some physical properties of a solution will be lost (often, positivity of distributions), although formally solutions may even exist, as in the case of chemical kinetics.

1.5. Weak and strong tendency to equilibrium, shaking and short memory

One aspect of irreversibility is the special choice of initial conditions. Roughly speaking, the arrow of time is defined by the fact that the quasi-equilibrium initial condition was in the past.

This remarkably simple observation does not, however, exhaust the problem of transition from the reversible equations to irreversible macroscopic equations. One more aspect deserves a serious consideration. Indeed, distribution functions tend to the equilibrium state according to macroscopic equations in a strong sense: deviations from the equilibrium tends to zero in the sense of most relevant norms (in the L^1 sense, for example, or even uniformly). On the contrast, for the Liouville equation, tendency to equilibrium occurs (if at all) only in the weak sense: mean values of sufficiently "regular" functions on the phase space do tend to their equilibrium values but the distribution function itself does not tend to the equilibrium with respect to any norm, not even pointwise. This is especially easy to appreciate if the initial state has been taken as the equipartition over some finite subset of the phase space (the "phase drop"). This phase drop can mix over the phase space, but for all the times it will remain "the drop of oil in the water", the density will always be taking only two values, 0 and $p > 0$, and the volume of the set where the density is larger than zero will not be changing in time, of course. So, how to arrive from the weak convergence (in the sense of the convergence of the mean values), to the strong convergence (to the L^1 or to the uniform convergence, for example)? In order to do this, there are two basic constructions: The coarse-graining (shaking) in the sense of Ehrenfests', and the short memory approximation.

The idea of coarse-graining dates back to P. and T. Ehrenfests, and it has been most clearly expressed in their famous paper of 1911 [2].

Ehrenfests considered a partition of the phase space into small cells, and they have suggested to alter the motions of the phase space ensemble due to the Liouville equation with "shaking" - averaging of the density of the ensemble over the phase cells. In the result of this process, the convergence to the equilibrium becomes strong out of the weak. It is not difficult to recognize that ensembles with constant densities over the phase cells are quasi-equilibria; corresponding macroscopic variables are integrals of the density over the phase cells ("occupation numbers" of the cells). This generalizes to the following: alternations of the motion of the phase ensemble due to microscopic equations with returns to the quasi-equilibrium manifold, preserving the values of the macroscopic variables. It is precisely this construction which serves for the point of departure for many of the constructions below.

Another construction is the short memory approximation. The essence of it is the following: If one excludes microscopic variables and assumes quasi-equilibrium initial conditions, then it is possible to derive integro-differential equations with retardation for the macroscopic variables (the way to do this is not unique). The form of the resulting equations is approximately this:

$$M(t) = \int_0^t K(t,t')[M(t')]dt',$$

where $K(t,t')$ is an operator (generally speaking, nonlinear) acting on $M(t')$. Once this equation is obtained, one assumes that the kernels of these integro-differential equations decay at a sufficiently high rate into the past (for example, exponentially, as $\|K(t,t')[M(t')]\| \leq \exp\{-\lambda(t-t')\}\|M(t')\|$). This can be interpreted in the spirit of Ehrenfests': Every motion which has begun sufficiently recently (the "memory time" τ before now) can be regarded as being started from the quasi-equilibrium. Thus, after each time τ has elapsed, the system can be shaken in the sense of Ehrenfests - the result should not differ much.

1.6. The essence of irreversibility in two words

(i) The direction of the arrow of time is defined by the fact that only "macroscopically definable ensembles" can be taken as initial conditions, that is, only quasi-equilibrium ensembles and what can be obtained from them when they are exposed to the true microscopic dynamics, or when partial quasi-equilibria are taken in positive time. *We* are created in such a way that we prepare and control (in part) the present, and observe what happens in the future. (In a sense, this is a definition of the subjective time).

(ii) Microscopic dynamics can give only the weak convergence to the equilibrium, convergence of mean values. Macroscopic variables tend to the equilibrium in the strong sense. The passage from micro to macro occurs here with the help of Ehrenfests' coarse-graining procedure or its analogs.

One might feel uneasy about the second of these points because the procedure of coarse-graining is not the result of the equations of motion, and therefore it is somehow voluntary. The only hope to lift this arbitrariness is that it may well happen that, in the limit of a very large number of particles, the perturbation caused by the coarse-graining can be made arbitrary small, for example, by increasing the time interval between coarse-graining.

1.7. Equivalence between trajectories and ensembles in the thermodynamic limit

In the preceding sections we were speaking about the dynamics of ensembles. This apparently contradicts the fact that the dynamics of a classical system goes along a single trajectory. Two arguments make it possible to proceed from the trajectories to ensembles:

(i) High sensitivity of trajectories to external perturbations when the number of particles is large. Arbitrary weak noise results in the stochastization of the motion.

(ii) In the thermodynamic limit, it is possible to partition the system into an arbitrary large number of small but still macroscopic sub-systems. Initial conditions in the sub-systems are independent from one sub-system to another, and they cannot be assigned completely voluntary but are taken from some distribution with a fixed sum of mean values (an analog of the macroscopic definability of ensembles). For spatially inhomogeneous systems, such small but still macroscopic sub-systems are defined in small and "almost homogeneous" volumes.

1.8. Subjective time and irreversibility

In our discussion, the source of the arrow of time is, after all, the asymmetry of the subjective time of the experimentalist. We prepare initial conditions, and after that we watch what will happen in the future but not what happened in the past. Thus, we obtain kinetic equations for specifically prepared systems. How is this related to the dynamics of the real world? These equations are applicable to real systems to the extent that the reality can be modeled with systems with specifically prepared quasi-equilibrium initial conditions. This is anyway less demanding than the condition of quasi-stationarity of processes in classical thermodynamics. For this reason, versions of nonequilibrium thermodynamics and kinetics based on this understanding of irreversibility allowed to include such a variety of situations, and besides that, they include all classical equations of nonequilibrium thermodynamics and kinetics.

2. Geometrization of irreversibility

2.1. Quasi-equilibrium manifold

Let E be a linear space, and $U \subset E$ be a convex subset, with a nonempty interior $\mathrm{int}U$. Let a twice differentiable concave functional S be defined in $\mathrm{int}U$, and let S is continuous on U. According to the familiar interpretation, S is the entropy, E is an appropriate space of distributions, U is the cone of nonnegative distributions from E. Space E is chosen in such a way that the entropy is well defined on U.

Let L be a closed linear subspace of space E, and $m : E \to E/L$ be the natural projection on the factor-space. The factor-space E/L will further play the role of the space of macroscopic variables (in examples, the space of moments of the distribution).

For each $M \in \mathrm{int}U/L$ we define the quasi-equilibrium, $f^*_M \in \mathrm{int}U$, as the solution to the problem,

$$S(f) \to \max, \; m(f) = M. \tag{6}$$

We assume that, for each $M \in \mathrm{int}U/L$, there exists the (unique) solution to the problem (6). This solution, f^*_M, is called the quasi-equilibrium, corresponding to the value M of the macroscopic variables. The set of quasi-equilibria f^*_M forms a manifold in $\mathrm{int}U$, parameterized by the values of the macroscopic variables $M \in \mathrm{int}U/L$.

Let us specify some notations: E^T is the adjoint to the E space. Adjoint spaces and operators will be indicated by ^{T}, whereas notation * is earmarked for equilibria and quasi-equilibria.

Furthermore, $[l, x]$ is the result of application of the functional $l \in E^T$ to the vector $x \in E$. We recall that, for an operator $A : E_1 \to E_2$, the adjoint operator, $A^T : E_1^T \to E_2^T$ is defined by the following relation: For any $l \in E_2^T$ and $x \in E_1$,

$$[l, Ax] = [A^T l, x].$$

Next, $D_f S(f) \in E^T$ is the differential of the functional $S(f)$, $D^2 S(f)$ is the second differential of the functional $S(f)$. Corresponding quadratic functional $D^2 S(f)(x, x)$ on E is defined by the Taylor formula,

$$S(f + x) = S(f) + [D_f S(f), x] + \frac{1}{2} D_f^2 S(f)(x, x) + o(\|x\|^2). \tag{7}$$

We keep the same notation for the corresponding symmetric bilinear form, $D_f^2 S(f)(x,y)$, and also for the linear operator, $D_f^2 S(f) : E \to E^T$, defined by the formula,

$$[D_f^2 S(f)x, y] = D_f^2 S(f)(x,y).$$

Here, on the left hand side there is the operator, on the right hand side there is the bilinear form. Operator $D_f^2 S(f)$ is symmetric on E, $D_f^2 S(f)^T = D_f^2 S(f)$.

Concavity of S means that for any $x \in E$ the inequality holds, $D_f^2 S(f)(x,x) \leq 0$; in the restriction onto the affine subspace parallel to L we assume the strict concavity, $D_f^2 S(f)(x,x) < 0$ if $x \in L$, and $x \neq 0$.

A comment on the degree of rigor is in order: the statements which will be made below become theorems or plausible hypotheses in specific situations. Moreover, specialization is always done with an account for these statements in such a way as to simplify the proofs.

Let us compute the derivative $D_M f_M^*$. For this purpose, let us apply the method of Lagrange multipliers: There exists such a linear functional $\Lambda(M) \in (E/L)^T$, that

$$D_f S(f)|_{f_M^*} = \Lambda(M) \cdot m, \; m(f_M^*) = M, \tag{8}$$

or

$$D_f S(f)|_{f_M^*} = m^T \cdot \Lambda(M), \; m(f_M^*) = M. \tag{9}$$

¿From equation (9) we get,

$$m(D_M f_M^*) = 1_{(E/L)}, \tag{10}$$

where we have indicated the space in which the unit operator is acting. Next, using the latter expression, we transform the differential of the equation (8),

$$D_M \Lambda = (m(D_f^2 S)_{f_M^*}^{-1} m^T)^{-1}, \tag{11}$$

and, consequently,

$$D_M f_M^* = (D_f^2 S)_{f_M^*}^{-1} m^T (m(D_f^2 S)_{f_M^*}^{-1} m^T)^{-1}. \tag{12}$$

Notice that, elsewhere in equation (12), operator $(D_f^2 S)^{-1}$ acts on the linear functionals from $\mathrm{Im} m^T$. These functionals are precisely those which become zero on L (that is, on $\ker m$), or, which is the same, those which can be represented as functionals of macroscopic variables.

The tangent space to the quasi-equilibrium manifold in the point f^*_M is the image of the operator $D_M f^*_M$:

$$\mathrm{Im}\,(D_M f^*_M) = (D^2_f S)^{-1}_{f^*_M} \mathrm{Im} m^T = (D^2_f S)^{-1}_{f^*_M} \mathrm{Ann} L \tag{13}$$

where $\mathrm{Ann}L$ is the set of linear functionals which become zero on L. Another way to write equation (13) is the following:

$$x \in \mathrm{Im}\,(D_M f^*_M) \Leftrightarrow (D^2_f S)_{f^*_M}(x, y) = 0,\ y \in L \tag{14}$$

This means that $\mathrm{Im}\,(D_M f^*_M)$ is the orthogonal complement of L in E with respect to the scalar product,

$$\langle x|y\rangle_{f^*_M} = -(D^2_f S)_{f^*_M}(x, y). \tag{15}$$

The entropic scalar product (15) appears often in the constructions below. (Usually, this becomes the scalar product indeed after the conservation laws are excluded). Let us denote as $T_{f^*_M} = \mathrm{Im}(D_M f^*_M)$ the tangent space to the quasi-equilibrium manifold in the point f^*_M. An important role in the construction of quasi-equilibrium dynamics and its generalizations is played by the quasi-equilibrium projector, an operator which projects E on $T_{f^*_M}$ parallel to L. This is the orthogonal projector with respect to the entropic scalar product, $\pi_{f^*_M} : E \to T_{f^*_M}$:

$$\pi_{f^*_M} = (D_M f^*_M|_M)\, m = \left(D^2_f S|_{f^*_M}\right)^{-1} m^T \left(m \left(D^2_f S|_{f^*_M}\right)^{-1} m^T\right)^{-1} m. \tag{16}$$

It is straightforward to check the equality $\pi^2_{f^*_M} = \pi_{f^*_M}$, and the self-adjointness of $\pi_{f^*_M}$ with respect to entropic scalar product (15). Thus, we have introduced the basic constructions: Quasi-equilibrium manifold, entropic scalar product, and quasi-equilibrium projector.

2.2. Thermodynamic projector

The construction of the quasi-equilibrium allows for the following generalization: Almost every manifold can be represented as a set of minimizers of the entropy under linear constraints. However, in general, these linear constraints will depend on the point on the manifold.

So, let the manifold $\Omega = f_M \subset U$ be given. This is a parametric set of distribution function, however, now macroscopic variables M are not functionals on R or U but just parameters defining the point on the manifold. The problem is how to extend definitions of M onto a neighborhood of f_M in such a way that f_M will appear as the solution to the variational problem:

$$S(f) \to \max,\ m(f) = M. \tag{17}$$

For each point f_M, we identify $T_M \in E$, the tangent space to the manifold Ω in f_M, and subspace $L_M \subset E$, which depends smoothly on M, and which has the property, $L_M \bigoplus T_M = E$. Let us define $m(f)$ in the neighborhood of f_M in such a way, that

$$m(f) = M,\ \text{if}\ f - f_M \in L_M. \tag{18}$$

The point f_M will be the solution of the quasi-equilibrium problem (17) if and only if

$$D_f S(f)|_{f_M} \in \text{Ann}\ L_M. \tag{19}$$

That is, if and only if $L_M \subset \ker D_f S(f)|_{f_M}$. It is always possible to construct subspaces L_M with the properties just specified, at least locally, if the functional $D_f S(f)|_{f_M}$ is not identically equal to zero on T_M.

The construction just described allows to consider practically any manifold as a quasi-equilibrium. This construction is required when one seeks the induced dynamics on a given manifold. Then the vector fields are projected on T_M parallel to L_M, and this preserves intact the basic properties of the quasi-equilibrium approximations.

2.3. Quasi-equilibrium approximation

Let a kinetic equation be defined in U:

$$\frac{df}{dt} = J(f). \tag{20}$$

(This can be the Liouville equation, the Boltzmann equation, and so on, dependent on which level of precision is taken for the microscopic description.) One seeks the dynamics of the macroscopic variables M. If we adopt the hypothesis that the solutions of the equation (20) of interest for us begin on the quasi-equilibrium manifold, and stay close to it for all the later times, then, as the first approximation, we can take the quasi-equilibrium approximation. It is constructed this way: We regard f as the quasi-equilibrium, and write,

$$\frac{dM}{dt} = m\left(J\left(f_M^*\right)\right). \tag{21}$$

With this, the corresponding to M point on the quasi-equilibrium manifold moves according to the following equation:

$$\frac{df^*_{M(t)}}{dt} = (D_M f^*_M) m(J(f^*_M)) = \pi_{f^*_M} J(f^*_M), \tag{22}$$

where $\pi_{f^*_M}$ is the quasi-equilibrium projector (16). It is instructive to represent solutions to equations of the quasi-equilibrium approximation (22) in the following way: Let $T_\tau(f)$ be the shift operator along the phase flow of equation (20) (that is, $T_\tau(f)$ is solution to equation (20) at the time $t = \tau$ with the initial condition f at $t = 0$). Let us take the initial point $f_0 = f^*_{M_0}$, and set, $f_{1/2} = T_\tau(f_0)$, $M_1 = m(f_{1/2})$, $f_1 = f^*_{M_1}$, $\ldots$, $f_{n+1/2} = T_\tau(f_n)$, $M_{n+1} = m(f_{n+1/2})$, $\ldots$. The sequence f_n will be termed the *Ehrenfests' chain*. We set, $f_\tau(n\tau) = f_n$. Then, $f_\tau(t) \to f(t)$, where $f(t)$ is the solution to the quasi-equilibrium approximation (22), as $\tau \to 0$, $n \to \infty$, $n\tau = t$.

Let us notice that the way the entropy evolves in time according to the Ehrenfests' chain is defined in the limit $\tau \to 0$ solely by the way it evolves along trajectories of the kinetic equation (20). Indeed, f^*_M is the point of maximum of the entropy on the subspace defined by equation, $m(f) = M$. Therefore, for

$$S(f_{n+1/2}) - S(f_{n+1}) = o(\|f_{n+1/2} - f_{n+1}\|) = o(\tau),$$

it holds

$$\sum_n |S(f_{n+1/2}) - S(f_{n+1})| = o(n\tau) \to 0,$$

for $\tau \to 0$, $n \to \infty$, $n\tau = const$. This simple observation has a rather important implication: Let us denote as $dS(f)/dt$ the entropy production due to the original kinetic equation (20), and as $(dS(f^*_M)/dt)_1$ its derivative due to the quasi-equilibrium system (22). Then,

$$(dS(f^*_M)/dt)_1 = dS(f)/dt|_{f=f^*_M}. \tag{23}$$

Let us give a different formulation of the latter identity. Let us term function $S(M) = S(f^*_M)$ *the quasi-equilibrium entropy*. Let us denote as $dS(M)/dt$ the derivative of the quasi-equilibrium entropy due to the quasi-equilibrium approximation (21). Then,

$$\frac{dS(M)}{dt} = \left.\frac{dS(f)}{dt}\right|_{f=f^*_M}. \tag{24}$$

¿From the identity (23), it follows the theorem about preservation of the type of dynamics:

(i) If, for the original kinetic equation (20), $dS(f)/dt = 0$ at $f = f_M^*$, then the entropy is conserved due to the quasi-equilibrium system (22).

(ii) If, for the original kinetic equation (20), $dS(f)/dt \geq 0$ at $f = f_M^*$, then, at the same points, f_M^*, $dS(M)/dt \geq 0$ due to the quasi-equilibrium system (21).

The theorem about the preservation of the type of dynamics demonstrates that, if there was no dissipation in the original system (20) (if the entropy was conserved) then there is also no dissipation in the quasi-equilibrium approximation. The passage to the quasi-equilibrium does not introduce irreversibility (the reverse may happen, for example, there is no dissipation in the quasi-equilibrium approximation for hydrodynamic variables as obtained from the Boltzmann kinetic equation; though dissipation is present in the Boltzmann equation, it occurs in different points but on the quasi-equilibrium manifold of local Maxwellians the entropy production is equal to zero). The same statement also hold for the thermodynamic projectors described in Section 2.2.

Usually, the original dynamics (20) does not leave the quasi-equilibrium manifold invariant, that is, vector field $J(f)$ is not tangent to the quasi-equilibrium manifold in all its points f_M^*. In other words, the *condition of invariance*,

$$(1 - \pi_{f_M^*})J(f_M^*) = 0, \tag{25}$$

is not satisfied. The left hand side of the invariance condition (25) is of such an outstanding importance that it deserves a separate name. We call it the *defect of invariance*, and denote it as $\Delta_{f_M^*}$. It is possible to consider the invariance condition as an equation, and to compute corrections to the quasi-equilibrium approximation f_M^* in such a way as to make it "more invariant". In those cases where the original equation (20) is already dissipative, this route of corrections, supplemented by the construction of the thermodynamic projector as in Section 2.2, leads to an appropriate macroscopic kinetics [4].

However, here, we are mainly interested in the route "from the very beginning", from conservative systems to dissipative. And here solving of the invariance equation does not help since it will lead us to, while "more invariant", but still conservative dynamics. In all the approaches to this problem (passage from the conservative to the dissipative systems), dissipation is introduced in a more or less explicit fashion by various assumptions about the "short memory". The originating point of our constructions will be the absolutely transparent and explicit approach of Ehrenfests.

2.4. Natural projector

So, let the original system (20) be conservative, and $dS(f)/dt = 0$. The idea of Ehrenfests is to supplement the dynamics (20) by "shaking". Shaking are external perturbations which are applied periodically with a fixed time interval τ, and which lead to a "forgetting" of small scale (nonequilibrium) details of the dynamics. For us here the shaking is the replacement of f with the quasi-equilibrium distribution $f^*_{m(f)}$. In the particular case which has been considered in by Ehrenfests, the macroscopic variables $m(f)$ were the averages of f over cells in the phase space, while $f^*_{m(f)}$ has been the cell-homogeneous distribution with the same average density, as for f. As we have already mentioned it, in the limit $\tau \to 0$, one gets back the quasi-equilibrium approximation - and the type of the dynamics is preserved. In this limit we obtain just the usual projection of the vector field $J(f)$ (20) on the tangent bundle to the quasi-equilibrium manifold. So, the natural question appears: What will happen, if we will not just send τ to zero but will consider finite, and even large, τ? In such an approach, not the vector fields are going to be projected but segments of trajectories. We shall term this way of projecting the *natural*. Let us now pose the problem of the *natural projector* formally. Let $T_t(f)$ be the phase flow of the system (20). We have to derive a phase flow of the macroscopic system, $\Theta_t(M)$ (that is, the phase flow of the macroscopic system, $dM/dt = F(M)$, which we are looking for), such that, for any M,

$$m(T_\tau(f^*_M)) = \Theta_\tau(M). \tag{26}$$

That is, when moving along the macroscopic trajectory, after the time τ we arrive at the same values of the macroscopic variables as if we were moving along the true microscopic trajectory for the same time τ, starting with the quasi-equilibrium initial condition.

It is instructive to remark that, at finite τ, the entropy growth follows immediately from equation (26) because $S(f) < S(f^*_{m(f)})$. The difference is of the order $\|f - f^*_{m(f)}\|^2$, for the time τ, thus, the first non-vanishing order in the entropy production will be of the order of τ. Let us find it.

We shall seek F in terms of a series in τ. Let us expand F and *both* the sides of the equation (26) in the powers of τ up to second order, and find the expansion coefficients of F [4]:

$$T_\tau(f^*_0) = f_0 + df/dt|_{f_0}\tau + d^2f/dt^2|_{f_0}(\tau^2/2) + o(\tau^2),$$
$$\Theta_\tau(M_0) = M_0 + dM/dt|_{M_0}\tau + d^2M/dt^2|_{M_0}(\tau^2/2),$$
$$df/dt|_{f_0} = J(f_0),\ d^2f/dt^2|_{f_0} = D_fJ(f)|_{f_0}J(f_0),$$

$$dM/dt|_{M_0} = F(M_0),\ d^2M/dt^2|_{M_0} = D_M F(M)|_{M_0} F(M_0),$$
$$F(M) = F_0(M) + \tau F_1(M) + 0(\tau).$$

Using these expansions in the condition for natural projector (26), we get,

$$\begin{aligned} & f_0 = f^*_{M_0}, \\ & m(f_0) + \tau m(J(f_0)) + (\tau^2/2) D_f J(f)|_{f_0} J(f_0) + o(\tau^2) \\ & = M_0 + \tau F_0(M_0) + \tau^2 F_1(M_0) + (\tau^2/2) D_M F(M)|_{M_0} F(M_0) + o(\tau^2), \end{aligned}$$

whereupon,

$$\begin{aligned} & F_0(M) = m(J(f^*_M)), \\ & F_1(M) = (1/2)\left\{ m(D_f J(f)|_{f^*_M} J(f^*_M)) - D_M F_0(M)|_M F_0(M) \right\}. \end{aligned}$$

Thus, the approximation F_0 is the quasi-equilibrium, and using this fact in the expression for F_1, after some transformation, we derive,

$$\begin{aligned} F_1 &= (1/2)\left\{ m(D_f J(f)|_{f^*_M} J(f^*_M)) - D_M(m(J(f^*_M)))m(J(f^*_M)) \right\} \\ &= (1/2)\left\{ m(D_f J(f)|_{f^*_M} J(f^*_M)) - m(D_f J(f)|_{f^*_M} D_M f^*_M) m(J(f^*_M)) \right\} \\ &= (1/2) m\left(D_f J(f)|_{f^*_M} [J(f^*_M) - D_M f^*_M m(J(f^*_M))] \right) \\ &= (1/2) m\left(D_f J(f)|_{f^*_M} [1 - \pi_{f^*_M}] J(f^*_M) \right) \\ &= (1/2) m\left(D_f J(f)|_{f^*_M} \Delta_{f^*_M} \right). \end{aligned}$$

Thus, the final form of the equation for the macroscopic variables M may be written:

$$\frac{dM}{dt} = F(M) = m(J(f^*_M)) + (\tau/2) m(D_f J(f)|_{f^*_M} \Delta_{f^*_M}) + o(\tau^2). \quad (27)$$

It is remarkable the appearance of the defect of invariance in the second term (proportional to τ): If the quasi-equilibrium manifold is invariant with respect to the microscopic dynamics, then $F(M)$ is quasi-equilibrium.

Let us compute the production of the quasi-equilibrium entropy $S(M) = S(f^*_M)$ due to macroscopic equations (27), neglecting the higher-order term $o(\tau^2)$.

$$dS(f^*_M)/dt = (\tau/2) D_f S(f)|_{f^*_M} \pi_{f^*_M} D_f J(f)|_{f^*_M} \Delta_{f^*_M}.$$

We notice that,

$$D_f S(f)|_{f^*_M}\pi_{f^*_M} = D_f S(f)|_{f^*_M},$$

because $\pi_{f^*_M}$ is a projector, and also because

$$\ker\, \pi_{f^*_M} \subset \ker\, D_f S(f)|_{f^*_M}.$$

Next, by our assumption, system (20) conserves the entropy,

$$dS(f)/dt = D_f S(f)|_f J(f) = 0.$$

Let us differentiate the latter identity:

$$D^2_f S(f)|_f J(f) + D_f S(f)|_f D_f J(f)|_f = 0. \tag{28}$$

Thus, due to the right hand side of equation (27),

$$\begin{aligned}\frac{dS(f^*_M)}{dt} &= (\tau/2) D_f S(f)|_{f^*_M} D_f J(f)|_{f^*_M} \Delta_{f^*_M}\\ &= -(\tau/2)\left(D^2_f S(f)|_{f^*_M} J(f^*_M)\right)\Delta_{f^*_M}\\ &= (\tau/2)\langle J(f^*_M)|\Delta_{f^*_M}\rangle_{f^*_M},\end{aligned}$$

where we have used notation for entropic scalar product (15). Finally,

$$\Delta_{f^*_M} = (1-\pi_{f^*_M})J(f^*_M) = (1-\pi_{f^*_M})^2 J(f^*_M),$$

whereas projector $\pi_{f^*_M}$ is self-adjoint in the entropic scalar product (15). Thus, $\langle J(f^*_M)|\Delta_{f^*_M}\rangle_{f^*_M} = \langle \Delta_{f^*_M}|\Delta_{f^*_M}\rangle_{f^*_M}$, and

$$\frac{dS(f^*_M)}{dt} = (\tau/2)\langle \Delta_{f^*_M}|\Delta_{f^*_M}\rangle_{f^*_M}. \tag{29}$$

Thus, the quasi-equilibrium entropy demonstrates an increase due to equation of macroscopic dynamics (27) in those points of the quasi-equilibrium manifold where the defect of invariance is not equal to zero. This way we see it how the problem of the natural projector (projected are not vector fields but segments of trajectories) results in dissipative e-quations. For specific examples see [5] where the second term in equation (27) results in viscous terms in the Navier-Stokes equations, diffusion and other dissipative contributions. However, it remains the undetermined coefficient τ. Formula (29) gives the entropy production just proportion-al to the time interval between subsequent coarse-graining. Of course, this could be true only for small enough τ, whereas we are mostly inter-ested in the limit $\tau \to \infty$. It is only in this limit where one can get rid of the arbitrariness of in the choice of τ present in equations (27) and (29). In order to do this, we need to study more carefully the structure of the trajectories which begin on the quasi-equilibrium manifold.

2.5. One-dimensional model of nonequilibrium states

In the background of many derivations of nonequilibrium kinetic equations there is present the following picture: Above each point of the quasi-equilibrium manifold there is located a huge subspace of nonequilibrium distributions with the same values of the macroscopic variables, as in the quasi-equilibrium point. It is as if the motion decomposes into two projections, above the point on the quasi-equilibrium manifold, and in the projection on this manifold. The motion in each layer above the points is highly complicated, but fast, and everything quickly settles in this fast motion.

However, upon a more careful looking into the motions which start from the quasi-equilibrium points, we will observe that, above each point of the quasi-equilibrium manifold it is located just a single curve, and all the nonequilibrium (not-quasi-equilibrium) states which come into the game form just a one-dimensional manifold. It is namely this curve the construction of which we shall be dealing with in this section.

For each value of the macroscopic variables M, and for each time τ, we define $M_{-\tau}$ by the following equality:

$$m(T_\tau(f^*_{M_{-\tau}})) = M. \tag{30}$$

In other words, $M_{-\tau}$ are those values of macroscopic variables which satisfy $\Theta_\tau(M_{-\tau}) = M$ for the natural projector (26). Of course, it may well happen that such $M_{-\tau}$ exists not for every pair (M, τ) but we shall assume here that for every M there exists such $\tau_M > 0$ that there exists $M_{-\tau}$ for $0 < \tau < \tau_M$.

A set of distributions, $q_{M,\tau} = T_\tau(f^*_{M_{-\tau}})$, forms precisely the curve of nonequilibrium states with given values of M in question. Notice that, for each τ, it holds, $m(q_{M,\tau}) = M$. The set $\{q_{M,\tau}\}$ for all possible M and τ is positive invariant: If the motion of the system starts on it at some time t_0, it stays on it also at $t > t_0$. If the dependence $q_{M.\tau}$ is known, equations of motion in the coordinate system (M, τ) have a simple form:

$$\begin{aligned} \frac{d\tau}{dt} &= 1, \\ \frac{dM}{dt} &= m(J(q_{M,\tau})). \end{aligned} \tag{31}$$

The simplest way to study $q_{M,\tau}$ is through a consideration of a sequence of its derivatives with respect to τ at fixed M. The first derivative is readily written as,

$$\left.\frac{dq_{M,\tau}}{d\tau}\right|_{\tau=0} = J(f^*_M) - \pi_{f^*_M} J(f^*_M) = \Delta_{f^*_M}. \tag{32}$$

By the construction of the quasi-equilibrium manifold (we remind that $L = \ker m$), for any $x \in L$,

$$S(f^*_M + \tau x) = S(f^*_M) - \tau^2 \langle x|x\rangle_{f^*_M} + o(\tau^2).$$

Therefore,

$$S(q_{M,\tau}) = S(f^*_M) - \tau^2 \langle \Delta_{f^*_M}|\Delta_{f^*_M}\rangle_{f^*_M} + o(\tau^2).$$

Thus, to first order in τ, we have, as expected.

$$q_{M,\tau} = f^*_M + \tau \Delta_{f^*_M} + o(\tau).$$

Let us find $q_{M,\tau}$ to the accuracy of the order $o(\tau^2)$. To this end, we expand all the functions in equation (30) to the order of $o(\tau^2)$. With

$$M_{-\tau} = M - \tau m(J(f^*_M)) + \tau^2 B(M) + o(\tau^2),$$

where function B is yet unknown, we write:

$$f^*_{M_{-\tau}} = f^*_M - \tau D_M f^*_M m(J(f^*_M)) + \tau^2 D_M f^*_M B(M) + (\tau^2/2) A_2(M) + o(\tau^2),$$

where

$$A_2(M) = \left.\frac{d^2 f^*_{M+tm(J(f^*_M))}}{dt^2}\right|_{t=0}, \tag{33}$$

and

$$\begin{aligned}
T_\tau(x + \tau\alpha) &= x + \tau\alpha + \tau J(x) + \tau^2 D_x J(x)|_x \alpha \\
&\quad + (\tau^2/2) D_x J(x)|_x J(x) + o(\tau^2), \\
T_\tau(f^*_{M_{-\tau}}) &= f^*_M - \tau D_M f^*_M m(J(f^*_M)) + \tau^2 D_M f^*_M B(M) + (\tau^2/2) A_2(M) \\
&\quad + \tau J(f^*_M) - \tau^2 D_f J(f)|_{f^*_M} D_M f^*_M m(J(f^*_M)) \\
&\quad + (\tau^2/2) D_f J(f)|_{f^*_M} J(f^*_M) + o(\tau^2) \\
&= f^*_M + \tau \Delta_{f^*_M} + (\tau^2/2) A_2(M) + (\tau^2/2) D_f J(f)|_{f^*_M} (1 - 2\pi_{f^*_M}) J(f^*_M) \\
&\quad + \tau^2 D_M f^*_M B(M) + o(\tau^2).
\end{aligned}$$

The latter somewhat lengthy expression simplifies significantly under the action of m. Indeed,

$$\begin{aligned}
&m(A_2(M)) = d^2[M + tm(J(f^*_M))]/dt^2 = 0,\\
&m(1 - \pi_{f^*_M}) = 0,\\
&m(D_M f^*_M) = 1.
\end{aligned}$$

Thus,

$$m(T_\tau(f^*_{M-\tau})) = M + (\tau^2/2)m(D_f J(f)|_{f^*_M}(1-2\pi_{f^*_M})J(f^*_M)) + \tau^2 B(M) + o(\tau^2),$$

$$B(M) = (1/2)m(D_f J(f)|_{f^*_M}(2\pi_{f^*_M} - 1)J(f^*_M)).$$

Accordingly, to second order in τ,

$$\begin{aligned}
q_{M,\tau} &= T_\tau(f^*_{M-\tau}) \\
&= f^*_M + \tau\Delta_{f^*_M} + (\tau^2/2)A_2(M) \\
&\quad +(\tau^2/2)(1 - \pi_{f^*_M})D_f J(f)|_{f^*_M}(1 - 2\pi_{f^*_M})J(f^*_M) + o(\tau^2).
\end{aligned} \tag{34}$$

Notice that, besides the dynamic contribution of the order of τ^2 (the last term), there appears also the term A_2 (33) which is related to the curvature of the quasi-equilibrium manifold along the quasi-equilibrium trajectory.

Let us address the behavior of the entropy production in the neighborhood of f^*_M. Let $x \in L$ (that is, $m(x) = 0$). The production of the quasi-equilibrium entropy, $\sigma^*_M(x)$, equals, by definition,

$$\sigma^*_M(x) = D_M S(f^*_M) \cdot m(J(f^*_M + x)). \tag{35}$$

Equation (35) gives the rate of entropy change under the motion of the projection of the state onto the quasi-equilibrium manifold if the true trajectory goes through the point $f^*_M + x$. In order to compute the right hand side of equation (35), we use essentially the same argument, as in the proof of the entropy production formula (29). Namely, in the point f^*_M, we have $L \subset \ker D_f S(f)|_{f^*_M}$, and thus $D_f S(f)|_{f^*_M}\pi_{f^*_M} = D_f S(f)|_{f^*_M}$. Using this, and the fact that entropy production in the quasi-equilibrium approximation is equal to zero, equation (35) may be written,

$$\sigma^*_M(x) = D_f S(f)|_{f^*_M}(J(f^*_M + x) - J(f^*_M)). \tag{36}$$

To the linear order in x, the latter expression reads:

$$\sigma^*_M(x) = D_f S(f)|_{f^*_M} D_f J(f)|_{f^*_M} x. \tag{37}$$

Using the identity (28), we obtain in equation (37),

$$\sigma^*_M(x) = -D^2_f S(f)|_{f^*_M}(J(f^*_M), x) = \langle J(f^*_M)|x\rangle_{f^*_M}. \tag{38}$$

Because $x \in L$, we have $(1 - \pi_{f^*_M})x = x$, and

$$\begin{aligned}\langle J(f^*_M)|x\rangle_{f^*_M} &= \langle J(f^*_M)|(1-\pi_{f^*_M})x\rangle_{f^*_M} \\ &= \langle (1-\pi_{f^*_M})J(f^*_M)|x\rangle_{f^*_M} = \langle \Delta_{f^*_M}|x\rangle_{f^*_M}.\end{aligned}$$

Thus, finally, the entropy production in the formalism developed here, to the linear order reads,

$$\sigma^*_M(x) = \langle \Delta_{f^*_M}|x\rangle_{f^*_M}. \tag{39}$$

2.6. Stability of quasi-equilibrium manifolds

The notion of stability does not cause essential difficulties when it goes about an invariant manifold, it is stable if, for any $\epsilon > 0$, there exist such $\delta > 0$ that a motion which has started at $t = 0$ at the distance (in some appropriate sense) less than δ from the manifold will not go further than ϵ at any $t > 0$.

However, this is not so for a non-invariant manifold, and, probably, it is not possible to give a useful for all the cases formalization of the notion of *stability of the quasi-equilibrium manifold*, in the spirit of motions going not far away when started sufficiently close to the manifold (indeed, what is here “sufficiently close” and “not far”?). In spite of that, expression (34) gives an important opportunity to measure the stability. Indeed, let us consider how the entropy production depends on τ, that is, let us study the function,

$$\sigma_M(\tau) = \langle \Delta_{f^*_M}|q_{M,\tau}\rangle_{f^*_M}. \tag{40}$$

It is natural to expect that $\sigma_M(\tau)$ initially increases, and then it saturates to some limiting value. The question is, however, how does function $\sigma_M(\tau)$ behave at $t = 0$, is it concave or is it convex in this point? If function $\sigma_M(\tau)$ is concave, $d^2\sigma_M(\tau)/d\tau^2|_{\tau=0} < 0$, then the speed with which it grows reduces immediately, and one can even estimate the limiting value,

$$\sigma^*_M = \lim_{\tau\to\infty} \sigma_M(\tau),$$

using the first Padé approximate:

$$\begin{aligned}\sigma_M(\tau) &= a\tau/(1+b\tau) = a\tau - ab\tau^2 + \dots \qquad (41)\\ \sigma^*_M &= a/b = -\frac{2(d\sigma_M(\tau)/d\tau|_{\tau=0})^2}{d^2\sigma_M(\tau)/d\tau^2|_{\tau=0}}.\end{aligned}$$

Concavity of $\sigma_M(\tau)$ at $\tau = 0$ ($d^2\sigma_M(\tau)/d\tau^2|_{\tau=0} < 0$) is analogous to a soft instability: The motion does not run too far away, and it is possible to estimate where it will stop, see equation (41). However, if $d^2\sigma_M(\tau)/d\tau^2|_{\tau=0} > 0$, then this is analogous to a hard instability, and none of the estimates like (41) work. Thus, everything is defined by the sign of the scalar product,

$$\left.\frac{d^2\sigma_M(\tau)}{d\tau^2}\right|_{\tau=0} = \langle \Delta_{f^*_M} | A_2(M) + D_f J(f)|_{f^*_M}(1 - 2\pi_{f^*_M})J(f^*_M)\rangle_{f^*_M}. \quad (42)$$

If this expression is negative, then the Padé estimate (41) gives:

$$\sigma^*_M = -\frac{2\langle \Delta_{f^*_M} | \Delta_{f^*_M}\rangle^2_{f^*_M}}{\langle \Delta_{f^*_M} | A_2(M) + D_f J(f)|_{f^*_M}(1 - 2\pi_{f^*_M})J(f^*_M)\rangle_{f^*_M}}. \quad (43)$$

In the opposite case, if the sign of the expression (42) is *positive*, we call the quasi-equilibrium manifold *unstable*.

Equation (43) allows us to estimate the parameter τ in the equations of the method of natural projector. To this end, we make use of equation (29):

$$(\tau/2)\langle \Delta_{f^*_M} | \Delta_{f^*_M}\rangle_{f^*_M} = \sigma^*_M,$$

whereupon,

$$\tau \approx -\frac{4\langle \Delta_{f^*_M} | \Delta_{f^*_M}\rangle_{f^*_M}}{\langle \Delta_{f^*_M} | A_2(M) + D_f J(f)|_{f^*_M}(1 - 2\pi_{f^*_M})J(f^*_M)\rangle_{f^*_M}}, \quad (44)$$

if the denominator assumes negative values. In this case, there are no free parameters left in equation (27).

Above, the parameter τ, or the time of "leaving the initial quasi-equilibrium condition" , has been appearing explicitly in the equations. Except for the case of linear quasi-equilibrium manifolds where the formal limit $\tau \to \infty$ can be addressed to derive generalized fluctuation-dissipation relations [7], this may be not the best way to do in the general, nonlinear case. In a consequent geometric approach to the problem of constructing the one-dimensional model of nonequilibrium states it is sufficient to consider the entropic parameter, $\delta S = S^*(M) - S$. Within this parameterization of the one-dimensional curve of the nonequilibrium states, one has to address functions $\sigma_M(\Delta S)$, rather than $\sigma_M(\tau)$ (40), whereas their Padè approximates can be constructed, in turn, from expansions in τ. Specific examples of this construction will be addressed in a separate publication.

In order to give an example here, we notice that the simplest geometric estimate amounts to approximating the trajectory $q_{M,\tau}$ with a second order curve. Given $\dot{q}_{M,\tau}$ and $\ddot{q}_{M,\tau}$ (34), we construct a tangent circle (in the entropic metrics, $\langle | \rangle_{f^*_M}$, since the entropy is the integral of motion of the original equations). For the radius of this circle we get,

$$R = \frac{\langle \dot{q}_{M,0} | \dot{q}_{M,0} \rangle_{f^*_M}}{\sqrt{\langle \ddot{q}_{\perp\ M,0} | \ddot{q}_{\perp\ M,0} \rangle_{f^*_M}}}, \tag{45}$$

where

$$\begin{aligned}
\dot{q}_{M,0} &= \Delta_{f^*_M}, \\
\ddot{q}_{\perp\ M,0} &= \ddot{q}_{M,0} - \frac{\langle \ddot{q}_{M,0} | \Delta_{f^*_M} \rangle_{f^*_M} \Delta_{f^*_M}}{\langle \Delta_{f^*_M} | \Delta_{f^*_M} \rangle_{f^*_M}}, \\
\ddot{q}_{M,0} &= (1 - \pi_{f^*_M}) D_f J(f)|_{f^*_M} (1 - 2\pi_{f^*_M}) J(f^*_M) + \left(D_M \pi_{f^*_M} \right) m(J(f^*_M)).
\end{aligned}$$

This geometric estimate amounts to the following value of the microscopic time τ (different from the above estimate based on Padé approximation):

$$\tau \approx \frac{\pi}{2} \sqrt{\frac{\langle \Delta_{f^*_M} | \Delta_{f^*_M} \rangle_{f^*_M}}{\langle \ddot{q}_{\perp\ M,0} | \ddot{q}_{\perp\ M,0} \rangle_{f^*_M}}}. \tag{46}$$

This approximation of τ gives also the estimate for the relaxation time of the entropy production to its limiting value, $\sigma = (2\tau/\pi) \langle \Delta_{f^*_M} | \Delta_{f^*_M} \rangle_{f^*_M}$.

Acknowledgments

A.N.G. is thankful to G. Sh. Fridman for stimulating discussions and encouragement. A.N.G. thanks Institut des Hautes Études Scientifiques, where parts of this work have been developed, for hospitality. I.V.K. is thankful to the organizers of the First Mexican Meeting on Mathematical and Experimental Physics, where parts of this work have been presented, and especially to Leo García-Colín and Francisco Uribe for discussions. I.V.K. acknowledges stimulating discussions with Miroslav Grmela. Finally, it is our pleasure to thank Misha Gromov and Hans Christian Öttinger for stimulating discussions.

References

[1] See, for example: *Studies in Statistical Mechanics, V. IX*, Eds. E.W. Montroll and J.L. Lebowitz, (North-Holland, 1981); J.L. del Río-Correa and L.S. García-Colín, *Phys. Rev. E* **48** (1993) 891; B. Robertson, *Phys. Rev.* **144** (1966) 151; H. Grabert, *Projection Operator Techniques in Nonequilibrium Statistical Mechanics* (Springer, Berlin, 1982); D. Zubarev, V. Morozov, and G. Röpke, *Statistical Mechanics of Nonequilibrium Processes, V. 1* (Akademie Verlag, Berlin, 1996).

[2] R. Feynman, *The Character of Physical Law*, (Cox and Wyman, London, 1965). Lecture No. 5.

[3] P. Ehrenfest, *Collected Scientific Papers*, (North-Holland, Amsterdam, 1959), pp. 213-300.

[4] A. N. Gorban and I.V. Karlin, *Transport Theory and Stat. Phys.* **23**, 559 (1994).

[5] A.N. Gorban, I.V. Karlin, H.C. Öttinger, and L.L. Tatarinova, *Phys. Rev. E* **63**, 066124(1-6) (2001).

[6] The well known work of Lewis [R.M. Lewis, *J. Math. Phys.* **8**, 1448 (1967).] contains, unfortunately, an inconsistency: He expanded only the right hand side of equation (26), forgetting to do the same also with the left hand side.

[7] A.N. Gorban and I.V. Karlin, *Phys. Rev. E* **65**, 026116(1-5) (2002).