Remarks about metastability

Richard Jordan
Carnegie Mellon University

David Kinderlehrer

Felix Otto

Follow this and additional works at: http://repository.cmu.edu/math
REMARKS ABOUT METASTABILITY

Richard Jordan
Department of Mathematics
University of Michigan

David Kinderlehrer
Department of Mathematical Sciences
Carnegie Mellon Univ.

Felix Otto
Courant Institute
New York University

Research Report No. 97-NA-003

January 1997

Sponsors

U.S. Army Research Office
Research Triangle Park
NC 27709

National Science Foundation
4201 Wilson Boulevard
Arlington, VA 22230
The compelling mechanism we wish to bring forward here is that the gradient flux of the native thermodynamic energy is the Fokker-Planck Equation, when taken in a suitable metric, [8],[9],[10]. This is the Wasserstein metric and it defines the weak* topology on the admissible class of probability densities. This immediately provides a deep connection between variational principles and stochastic differential equations. Moreover, we are allowed a format to discuss metastable systems: evolution may be interpreted as a competition between energy and distance in the weak* topology.

In this note, which is a plan for research more than a statement of results, we take up some of these issues. Our attention is confined to the simplest case.

2. A SIMPLE PARADIGM

For illustration, consider a potential \( \psi(\xi) \), for a generalized snap spring or a shape memory element, for example, where \( \xi \) is the relative elongation of the spring or the shear length of the shape memory element, \( -\infty < \xi < +\infty \). Let \( n_{\xi} \) denote the number of elements of elongation \( \xi \in \{ \xi_1, \ldots, \xi_M \} \) so that

\[
E = \sum \psi(\xi)n_{\xi} \quad \text{and} \quad d = \sum \xi n_{\xi}
\]

are the total energy and elongation of the configuration. For such a system there is a configurational entropy or degeneracy given by Boltzmann's statistical definition

\[
S = -\log \left( \frac{1}{n_{\xi}} \right),
\]

where \( \binom{n_{\xi}}{l} \) is the number of ways of arranging \( 1 \, n_{\xi} \, 1 \) objects into \( M \) subsets with \( n_{\xi} \) elements in the \( j \)th subset. The average free energy of the system is, \( N = 1 \, n_{\xi} \, 1 \, \sigma > 0 \),
\[ l_n = \frac{1}{n_n} \sum \psi(\xi_n) n_k + \sigma \log \left( \frac{n_k}{n_{k+1}} \right) \]

perhaps subject to an imposed constraint like \( d = d_0 \) with \( d_0 \) fixed. The parameter \( \sigma \) plays the role of the temperature. The potential energy of independent layers or springs, given by \( E \), seeks to be minimum while the entropy seeks to be maximal by distributing elements evenly over the range. See Hou and Müller [5] for development of this sort of model in shape memory alloys. Passing to the limit as \( N,M \to \infty \), gives, typically, the functional, defined on probability densities \( \rho \),

\[ F_0(\rho) = \int_\mathbb{R} \psi \rho \, d\xi + \sigma \int_\mathbb{R} \rho \log \rho \, d\xi. \]

This is the type of functional we wish to consider.

A convex function of \( \rho \), it admits a unique minimum, the (stationary) Gibbs distribution

\[ \rho_\sigma(\xi) = \frac{1}{Z(\sigma)} e^{-\sigma \psi(\xi)} \text{, with } Z(\sigma) = \int_\mathbb{R} e^{-\sigma \psi(\xi)} \, d\xi. \]

We witness in this construction virtually the paradigm of classical statistical statistical mechanics and, in the example as a particular case, the derivation of a Young measure formulation of a variational problem coupled to an entropic stabilization. The relaxation or Young measure distribution of a nonconvex variational principle may be realized as the zero temperature limit of \( F_0 \). Namely,

\[ l(\rho) = \int_\mathbb{R} \psi \rho \, d\xi \]

represents the Young measure relaxation approach to a minimization problem and in our interpretation it has become ineluctably wedded to an entropy functional. Likewise, its driving force equation becomes linked to a Langevin Equation.

3. **Fokker-Planck Dynamics**

Were we to give an initial distribution of elements \( \rho_0 \) and ask how it relaxes to equilibrium, we might impose evolution of the probability flux or the Fokker-Planck Equation

\[ \frac{\partial \rho}{\partial t} = \sigma^2 \frac{\partial^2 \rho}{\partial \xi^2} + \frac{\partial}{\partial \xi} (\psi \rho), \quad -\infty < \xi < \infty, \quad t > 0, \quad (3.1) \]

\[ \rho(t,0) = \rho_0. \]

The motivation for this is that the solution of (3.1) satisfies

\[ \frac{d}{dt} F_0(\rho) \leq 0 \quad \text{for solutions } \rho \text{ of (3.1).} \]

This is equivalent to seeking the distribution of the relaxation of the driving force equation of the strain rate, with a small stochastic force, given by the Langevin Equation

\[ dX = -\psi\'(X(t)) \, dt + \sqrt{\sigma} \, dB(t), \quad (3.2) \]

where \( B(t) \) is standard Brownian motion and \( dB(t) \) represents white noise.

The compelling connection we wish to bring forward is that the implicit scheme

\[ \text{Determine } \rho^m \text{ that minimizes} \]

\[ \frac{1}{2} d(\rho,\rho^m)^2 + \int_\mathbb{R} \psi \rho \, d\xi + \sigma \int_\mathbb{R} \rho \log \rho \, d\xi, \quad (3.3) \]

where \( d \) is the Wasserstein metric, briefly described below, gives rise to a solution of the Fokker-Planck Equation, [8],[9],[10].
The Wasserstein distance, \([12]\), between two probability measures \(\mu_1\) and \(\mu_2\) on \(\mathbb{R}\) is
\[
d(\mu_1, \mu_2)^2 = \inf_{\mathcal{P}} \int_{\mathbb{R} \times \mathbb{R}} |x - y|^2 \, dp(x,y),
\]
(3.4)
\[
\mathcal{P} = \mathcal{P}(\mu_1, \mu_2) = \text{probability measures on } \mathbb{R} \times \mathbb{R} \text{ with first marginal } \mu_1 \text{ and second marginal } \mu_2.
\]
So for \(p \in \mathcal{P}\), \(p(A \times \mathbb{R}) = \mu_1(A)\) and \(p(\mathbb{R} \times B) = \mu_2(B)\). It is well known that \(d\) defines a metric equivalent to the weak* topology on probability measures with the property
\[
\int_{\mathbb{R}} |x|^2 \, d\mu(x) < \infty
\]
when appropriately defined as contained in a dual space. Equivalently,
\[
d(\mu_1, \mu_2)^2 = \inf E(|X - Y|^2),
\]
where \(E\) denotes the expectation of the random variable and the infimum is taken over random variables \((X, Y)\) where \(X\) has distribution \(\mu_1\) and \(Y\) has distribution \(\mu_2\). Since
\[
E(|X - Y|^2) = E(|X|^2) + E(|Y|^2) - 2E(XY)
\]
and \(E(|X|^2) = \int_{\mathbb{R}} |x|^2 \, d\mu_1(x),\ E(|Y|^2) = \int_{\mathbb{R}} |y|^2 \, d\mu_2(y),\)
calculating the Wasserstein distance consists in maximizing the correlation between \(X\) and \(Y\).

The variational problem (3.4) is an example of a Monge-Kantorovich mass transference problem with the cost function \(c(x, y) = |x - y|^2\), \([2],[3]\). Variational problems of this type have applications in many disciplines. A minimizer in (3.4) is called an optimal transference plan is easily shown to exist. In our situation, \(\mu_1\) and \(\mu_2\) will always be absolutely continuous with respect to Lebesgue measure and so we shall not distinguish between them and their densities, say, \(\rho_1\) and \(\rho_2\).

4. A SYSTEM EXHIBITING HYSTERESIS

Systems that exhibit hysteresis are only metastable. Here we illustrate an extremely simple example determined by a family of double well potentials of varying relative heights \(\psi(\xi, L)\). The two wells are at \(\pm 1\) and \(\psi\) is a step function of the parameter \(L\). In this elementary testbed, the solution of the Fokker-Planck Equation which gives the distribution of the \(\xi\) is simulated by a straightforward explicit scheme and implemented with Maple. The first moment, or average \(\xi\),

Figure 1 A hysteresis portrait determined by a Fokker-Planck Equation showing metastable states
is plotted as a function of the load parameter $L$. Although we lack the space to provide details, most of the outer loop does represent a distribution which is in equilibrium, which is the appropriate Gibbs distribution. But the inner segment and the portion of the outer loop from $0.5 < L < 0.8$, $x = -1$ is only metastable.

5. A BRIEF VIEW OF THE CONSTRAINED THEORY

In this section we give a schematic description of work with Richard James and Shlomo Ta'asan. The notion of the wiggly energy was first introduced by Abeyaratne, Chu, and James [1] in the shape memory CuAlNi system to interpret the hysteresis in evolution of the microstructure. A system similar to that of §2, governed by a Helmholtz free energy $W(\alpha)$ and an additional work of loading $T(\alpha)$ will have total energy, in its homogeneous Young measure form,

$$\int_{\mathbb{R}} (W(\alpha) + T(\alpha)) \, d\nu(\alpha).$$

Assume that the system is near equilibrium, which leads to the constraint

$$\text{supp} \, \nu \subset \{ W = \min W = 0 \}$$

and energy

$$E = \min_{\nu} \int_{\mathbb{R}} (W(\alpha) + T(\alpha)) \, d\nu(\alpha) = \min_{\nu} \int_{\mathbb{R}} T(\alpha) \, d\nu(\alpha).$$

Assume that the set of Young measures obeying (5.1) is a 1-parameter family $\nu^\xi$ depending on $\xi$. Then

$$E = \psi(\xi) = \min_{\xi} \int_{\mathbb{R}} T(\alpha) \, d\nu^\xi(\alpha),$$

leading to the driving force equation, where $\mu > 0$ is a parameter.

\[ \frac{d\xi}{dt} = -\mu \psi(\xi). \]

We take $\mu = 1$ in the sequel.

Owing to the accommodation of a finer scale structure, whose details we do not describe here, we are led to augment $\psi$ by a family $\psi_{\epsilon}$ which we take to be,

$$\psi_{\epsilon}(\xi) = \frac{1}{2} \xi^2 + \epsilon \psi(\frac{\xi}{\epsilon}), \epsilon > 0.$$ \[ \psi_{\epsilon} \text{ periodic of period 1, } ||\psi_{\epsilon}||_1 \leq \sigma, \text{ and } \int_{\mathbb{R}} \psi_{\epsilon}(y) \, dy = 0. \]

We arrive in this way at a Langevin Equation

$$dX_{\epsilon} = -\psi_{\epsilon}(X_{\epsilon}) \, dt + \sqrt{2} \, dB(t),$$

and corresponding Fokker-Planck Equation

$$\frac{\partial p}{\partial t} = \sigma \frac{\partial^2 p}{\partial \xi^2} + \frac{\partial}{\partial \xi} (\psi_{\epsilon}(p)), \quad -\infty < \xi < \infty, \quad t > 0.$$
This first order linear equation may be solved by differencing in time, as suggested by our discrete scheme, which leads to the approximating equation

\[
\frac{\partial}{\partial t} (\psi_t \rho) = \frac{1}{h} (\rho - \rho^{n-1}),
\]

or in weak form,

\[
-\int \zeta \psi_t \rho \, d\xi = \frac{1}{h} \int \zeta (\rho - \rho^{n-1}) \, d\xi.
\]

Choosing \( \zeta \) a suitable Heaviside function with jump at a given \( \xi \),

\[
\psi_t (\xi) \rho (\xi) = \frac{1}{h} (F(\xi) - F^{n-1}(\xi)),
\]

where \( F \) and \( F^{n-1} \), the distribution functions of the probability densities \( \rho \) and \( \rho^{n-1} \), converge pointwise as \( \varepsilon \to 0 \). From this and the fact that \( \psi_t \) has on the order of \( 2a/\varepsilon \) zeros in \([-a,a]\), we may infer the form of the limit transport equation

\[
\frac{\partial}{\partial t} \rho = \frac{\partial}{\partial x} (b \rho)
\]

where \( b (\xi) = \int_{1} (\xi + \psi_t (y))^{-1} \, dy \) if \( |\xi| > a \),

\[
b (\xi) = 0 \quad \text{if} \quad |\xi| \leq a.
\]

This is the form obtained in [1].

Acknowledgements

This research was supported by the ARO and the NSF through grants to the Center for Nonlinear Analysis. Additional support was from ARO DAAH04-96-0060, NSF/DMS 950578, and the Deutsche Forschungsgemeinschaft. We thank Chun Liu and Robert Sekera for helpful conversations.

References


Remarks about metastability

Richard Jordan  David Kinderlehrer  Felix Otto
Department of Mathematics  Department of Mathematical Sciences  Courant Institute
University of Michigan  New York University
Carnegie Mellon University

Processes that involve disparate length scales and which are only metastable are inherent to the investigation of mesoscopic and microscopic systems. We focus here on a mechanism we believe to be deeply intertwined with these properties. This is the competition between the thermodynamic energy and nearness in the weak* topology for the distribution of microscopic variables whose averages describe the evolution of the macroscopic system. Brief examples show metastable evolution and the possibility of accommodating additional fine scale variables.