Analysis of the Effect of a Heat Bath on a Rate-Independent System

M. Koslowski[1], M. Ortiz[2], T.J. Sullivan[2] and F. Theil[3]

[1] School of Mechanical Engineering, Purdue University, U.S.A.

[2] Graduate Aeronautical Laboratories, California Institute of Technology, U.S.A.

[3] Mathematics Institute, University of Warwick, U.K.

Energy-Driven Systems
Carnegie Mellon University
27–29 August 2009

Introduction

- Many (deterministic) physical processes are modeled by evolution laws that incorporate not just energetics, but also dissipation, i.e. loss of energy due to (say) frictional effects.
- Also, many physical processes are subject to some randomness, perhaps due to the presence of a heat bath, which supplies additional (uncorrelated) energy to the process.

Question

How does a dissipative system (in particular, a rate-independent system) behave when placed in contact with a heat bath?

To investigate this, we introduce a notion of thermalized gradient descent.

Introduction

- This talk will concentrate on the case of one-homogeneous dissipation, in which case the unthermalized dynamics are rate-independent.
- In our analysis, the thermalized dynamics turn out to be a nonlinear gradient descent; the thermalized dissipation potential is a "smoothing out" of the original one.
- As a toy model, consider a rough block sitting on a sandpaper table and subject to forces (springs, external loads, & *c*.) weaker than the frictional resistance of the sandpaper/block interface. Intuition suggests that
  - at "zero temperature", the block shouldn't move at all;
  - at "positive temperature" (shaking the table), the block might move — deterministically? randomly?

Gradient Descents

- A gradient descent in $\mathbb{R}^n$, say, is an evolutionary system described by two potentials: an energetic potential E and a dissipative potential Ψ.
- Typically, existence and uniqueness questions, as well as computation, are addressed using the Moreau–Yosida (implicit Euler) incremental formulation: given $x_i \approx x(t_i)$, find $x_{i+1} \approx x(t_{i+1})$ to minimize

$$\mathscr{W}_{i+1} \colon y \mapsto E(t_{i+1}, y) - E(t_i, x_i) + \Delta t_{i+1} \Psi \left(\frac{y - x_i}{\Delta t_{i+1}} \right). \qquad \text{(MY)}$$

- The idea now is to generate a thermalized gradient descent by seeking densities that minimize a functional in which (MY) competes with a entropy term.

Thermalized Gradient Descent

- Consider the following incremental problem for the PDF $\rho(t,\cdot)$ of X_t at discrete times $0 = t_0 < t_1 < \ldots < t_N = T$:
  - Consider the "prior" density $\rho_i \approx \rho(t_i,\cdot)$.
  - Find a new joint density $\rho_{i,i+1}(\cdot,\cdot)$, with first marginal ρ_i, that minimizes

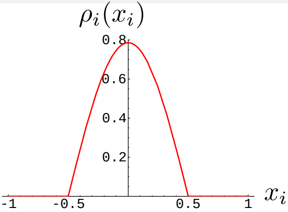
$$\tilde{\rho} \mapsto \iint \left[\mathscr{W}_{i+1}\tilde{\rho} + \varepsilon \tilde{\rho} \log \tilde{\rho} \right],$$

    where

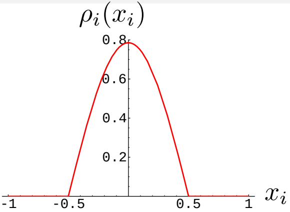
$$\mathscr{W}_{i+1}(x_i, x_{i+1}) = E(t_{i+1}, x_{i+1}) - E(t_i, x_{i+1}) + \Delta t_{i+1} \Psi \left(\frac{\Delta x_{i+1}}{\Delta t_{i+1}} \right)$$

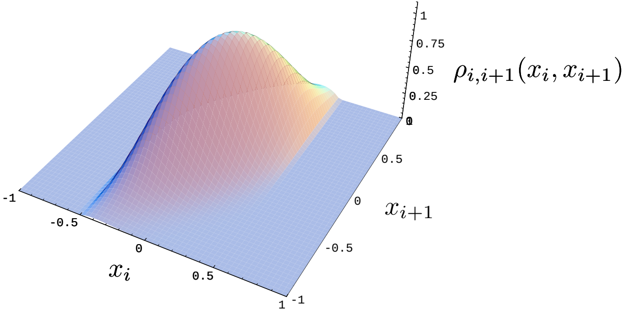
    is the "cost" of changing from state x_i to state x_{i+1}.
  - Integrate/marginalize over the first slot of $\rho_{i,i+1}(\cdot,\cdot)$ to get a new density ρ_{i+1} for time t_{i+1}.

Thermalized Gradient Descent

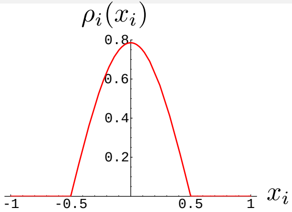


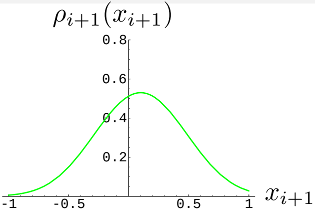
Thermalized Gradient Descent

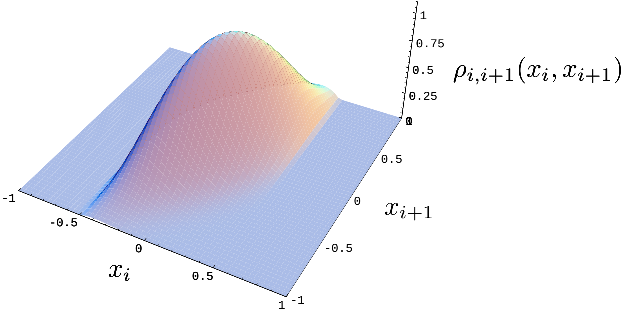




Thermalized Gradient Descent







Thermalized Gradient Descent

Lemma

Subject to mild regularity and growth conditions, the single-step increments follow a Gibbs-Boltzmann-type distribution with respect to the incremental cost function:

$$\rho_{i+1}(x_{i+1}|x_i) = \frac{1}{Z(x_i)} \exp\big(- \mathscr{W}_{i+1}(x_i, x_{i+1})/\varepsilon \big).$$

Definition

On a partition P of $[0, T]$, we will call the Markov chain $X^{(P)}$ so generated the (discrete-time) thermalized gradient descent in E and Ψ.

To do: take a continuous-time interpolation and examine the limit of $X^{(P)} \colon \Omega \times [0, T] \to \mathbb{R}^n$ as $\mathrm{mesh}(P) \to 0$.

Thermalized Gradient Descent

- For “nice” potentials E and 2-homogeneous Ψ, the incremental scheme makes sense. As the parition mesh tends to zero, the $X^{(P)}$ converges in law on path space to the solution Y of the Itō stochastic gradient descent

$$\dot{Y}(t) = -\nabla E(t, Y(t)) + \sqrt{\varepsilon}\,\dot{W}(t);$$

  i.e., the thermalized gradient descent scheme is a plausible model for “linear kinetics $+$ noise”.
- The discrete-time scheme also makes sense for 1-homogeneous Ψ — but what is the continuous-time limit as $\mathrm{mesh}(P) \to 0$?

Consider a "nice" rate-independent system in $\mathbb{R}^n$:

- an energetic potential $E \colon [0,T] \times \mathbb{R}^n \to \mathbb{R}$ — convex, time derivative in $W^{1,\infty}$, space derivative in $\mathcal{C}^1$;
- a dissipation potential $\Psi \colon \mathbb{R}^n \to [0,+\infty)$ — homogeneous of degree one and strictly convex (i.e. non-degenerate). Ψ is the convex conjugate of the characteristic function of a convex compact set $\mathscr{E} \subsetneq (\mathbb{R}^n)^*$ that has $0 \in \mathring{\mathscr{E}}$, the *elastic region*:

$$\Psi(v) = \sup\{\langle \ell, v\rangle \mid \ell \in \mathscr{E}\}.$$

Study the process $X^{(P)}$ by studying its increments:

$$\Delta X_i^{(P)} := X_i^{(P)} - X_{i-1}^{(P)}.$$

Definition

Define an effective dual dissipation potential

$$\widetilde{\Psi}^{\star} \colon (\mathbb{R}^n)^* \to \mathbb{R} \cup \{+\infty\}$$

by

$$\widetilde{\Psi}^{\star}(\ell) := \log \int_{\mathbb{R}^n} \exp\left(-\left(\langle \ell, z \rangle + \Psi(z)\right)\right) \, \mathrm{d}z.$$

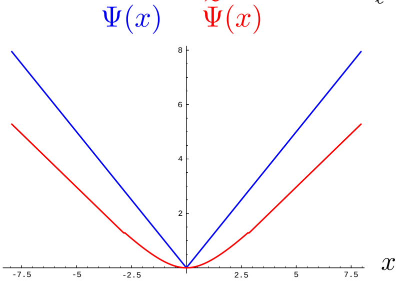
Define the effective dissipation potential by convex conjugation:

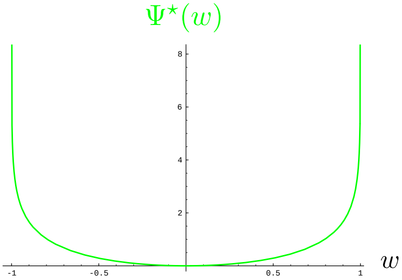
$$\widetilde{\Psi}(v) = \widetilde{\Psi}^{\star\star}(v) := \sup\left\{\langle \ell, v \rangle - \widetilde{\Psi}^{\star}(\ell) \,\middle|\, \ell \in (\mathbb{R}^n)^*\right\}.$$

Note that $\widetilde{\Psi}$ is determined purely by the dissipation functional Ψ (or, equivalently, the elastic region $\mathscr{E} \subsetneq (\mathbb{R}^n)^*$ associated to Ψ).

$$\widetilde{\Psi}^\star(\ell) := \log \int_{\mathbb{R}^n} \exp\big(-\big(\langle \ell, z\rangle + \Psi(z)\big)\big)\, \mathrm{d}z.$$

$$\widetilde{\Psi}(v) := \sup_\ell \big[\langle \ell, v\rangle - \widetilde{\Psi}^\star(\ell)\big].$$





The effective (dual) dissipation potential in dimension one for $\Psi(x) = |x|$. $\widetilde{\Psi}^\star$ blows up like the logarithm of the distance to the yield surface $\partial\mathscr{E}$. $\widetilde{\Psi}$ is smooth and is asymptotic to Ψ at infinity.

- The reason that $\widetilde{\Psi}$ is so important is that the change of variables $X_{i+1} \rightsquigarrow \Delta X_{i+1}/\varepsilon_{i+1}$ yields (modulo higher-order error terms):

$$\mathbb{E}\big[\Delta X_{i+1}\big|X_i = x_i\big] \approx -\varepsilon_{i+1}\mathrm{D}\widetilde{\Psi}^\star\big(\mathrm{D}E(t_{i+i}, x_i)\big);$$

$$\mathrm{Var}\big[\Delta X_{i+1}\big|X_i = x_i\big] \approx -\varepsilon_{i+1}^2 \left|\mathrm{D}^2\widetilde{\Psi}^\star\big(\mathrm{D}E(t_{i+i}, x_i)\big)\right| \ll \varepsilon_{i+1}.$$

- Hence, it looks like the continuous-time limit with $\varepsilon_{i+1} = \theta\Delta t_{i+1}$ should satisfy the deterministic ordinary differential equation

$$\dot{y}(t) = -\theta\mathrm{D}\widetilde{\Psi}^\star\big(\mathrm{D}E(t, y(t))\big)$$

$$\text{i.e., } \mathrm{D}\widetilde{\Psi}\left(-\frac{\dot{y}(t)}{\theta}\right) = \mathrm{D}E(t, y(t)).$$

- If Ψ is the weighted ℓ^1 norm $\Psi(z) := \sigma_1|z_1| + \ldots + \sigma_n|z_n|$, with weights $\sigma_i > 0$, then (up to an additive constant)

$$\widetilde{\Psi}^\star(\ell) = -\sum_{i=1}^{n} \log\left(\sigma_i^2 - (\ell \cdot e_i)^2\right).$$

- If Ψ is a multiple of the Euclidean norm, $\Psi(z) := \sigma|z|_2$, $\sigma > 0$, then

$$\begin{aligned}\widetilde{\Psi}^\star(\ell) &= \log \int_{\mathbb{S}^{n-1}} \frac{(n-1)!}{(\ell \cdot \omega + \sigma)^{-n}} \, \mathrm{d}\mathcal{H}^{n-1}\lfloor_{\mathbb{S}^{n-1}}(\omega) \\ &= -\frac{n+1}{2} \log\left(\sigma^2 - |\ell|_2^2\right).\end{aligned}$$

Convergence Result

Theorem (Convergence to Nonlinear Gradient Descent)

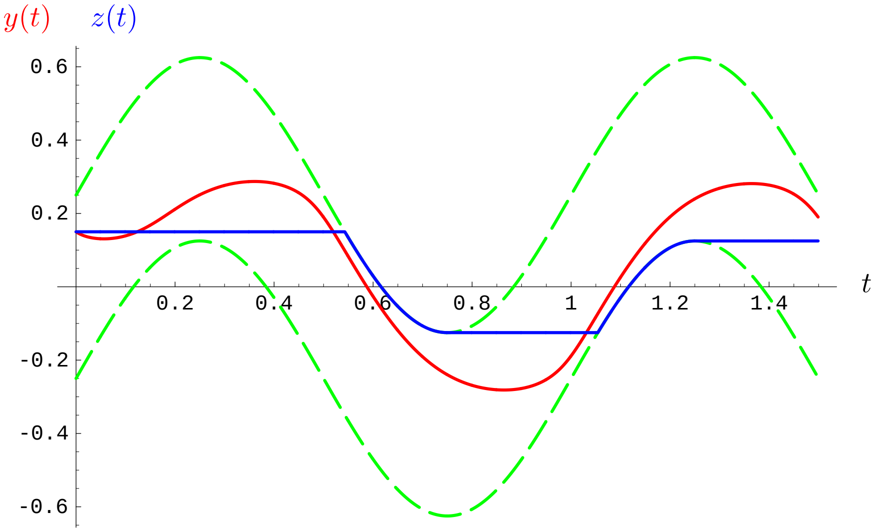
If $(t, x) \mapsto \widetilde{\Psi}^\star(\mathrm{D}E(t, x))$ *is convex in* x *for each* $t \in [0, T]$*, then the piecewise constant càdlàg interpolation* $\bar{X}^{(P)} \colon \Omega \times [0, T] \to \mathbb{R}^n$ *converges in probability as* $\operatorname{mesh}(P) \to 0$ *to the solution of*

$$\mathrm{D}\widetilde{\Psi}\left(-\frac{\dot{y}(t)}{\theta}\right) = \mathrm{D}E(t, y(t)).$$

More precisely, for any $\lambda > 0$*, as* $\operatorname{mesh}(P) \to 0$,

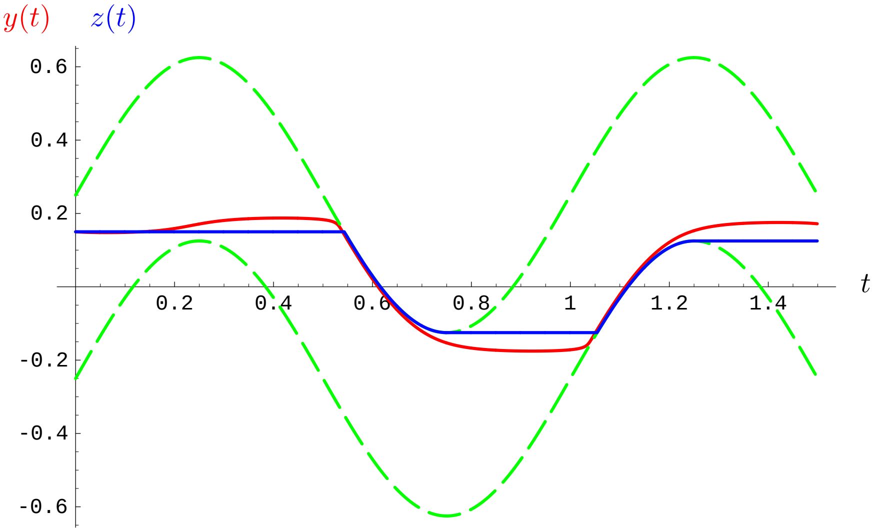
$$\mathbb{P}\left[\sup_{t \in [0,T]} \left|\bar{X}^{(P)}(t) - y(t)\right|_2 \geq \lambda\right] \in O\big(\operatorname{mesh}(P)^{1/2}\big).$$

(If $\mathrm{D}^2 E \equiv 0$*, then the order of convergence is* $O(\operatorname{mesh}(P))$*.)*



$\theta = 1$

Indicated in green is the frontier of the stable region,



$$\theta = \frac{1}{10}$$

Indicated in green is the frontier of the stable region,

Andrade's Creep Law

- In 1910, Andrade reported that as a function of time, t, the creep deformation, ξ, of soft metals at constant temperature and applied stress can be described by a power law $\xi(t) \sim t^{1/3}$.
- Similar behavior has been observed in many classes of materials, including non-crystalline materials.
- Morally, macroscopic creep should be observed as a change in the mean of the microscopic slip field.

Phase Field Model

- Consider the Koslowski–Cuitiño–Ortiz phase field model for a material sample along a single slip plane, thought of as the unit torus, $\mathbb{T}^2$.
- $u(x) \in \mathbb{R}$ is the slip (in multiples of the Burgers vector) at $x \in \mathbb{T}^2$.
- Dissipation is concentrated over small discs centred on a (random) set of obstacles, $\mathcal{O}$.
- KCO is a random, large finite-dimensional model, so we reduce by a mean field approximation $u \rightsquigarrow \xi \in \mathbb{R}$.
- The resulting model is of the form

$$E_{\mathrm{MF}}(t, \xi) = -\ell(t)\xi; \quad \Psi_{\mathrm{MF}}(\xi) = \sigma|\xi|.$$

Convergence Result

- In this case, the effective dual dissipation potential for the mean field is
$$\widetilde{\Psi}^{\star}_{\mathrm{MF}}(\ell) = -\log\left(\sigma^2 - \ell^2\right).$$
- Under the assumption of linear strain hardening (i.e. $\sigma = \sigma_0 \xi$) the resulting effective ordinary differential equation is
$$\dot{\xi} = \frac{2\theta\ell}{\sigma_0^2\xi^2 - \ell^2}.$$
- For constant θ and constant $0 < \ell \ll \sigma$, solutions grow in accord with Andrade's creep law, $\xi(t) \sim t^{1/3}$.

Conclusions

- The thermalized gradient descent scheme describes the effect of adding Itō noise to a system with 2-homogeneous dissipation.
- In the case of 1-homogeneous dissipation, neglecting inertia, the scheme yields a deterministic gradient flow in an effective dissipation potential $\widetilde{\Psi}$ that is a nonlinear transformation ("smoothing-out") of the original dissipation potential Ψ.
- This analysis can be used to derive Andrade's creep law.

Future Work

Extension to

- second-order equations of motion (inertial effects, non-Markovian processes)?
- infinite-dimensional state spaces?
- curved state spaces (manifolds)?