Thermalization of Rate-Independent Processes by Entropic Regularization

Tim Sullivan
California Institute of Technology, U.S.A.

Interplay of Analysis and Probability in Physics
Mathematisches Forschungsinstitut Oberwolfach
22–28 January 2012

Joint work with M. Koslowski (Purdue), F. Theil (Warwick) and M. Ortiz (Caltech).
Problem Setting

- **Gradient descent** on a connected Riemannian manifold \((Q, g)\) in an energetic potential \(E : [0, T] \times Q \to \mathbb{R}\) with respect to a dissipation potential \(\Psi : [0, T] \times TQ \to [0, +\infty)\):

\[
\partial \Psi(t, z(t), \dot{z}(t)) \ni -DE(t, z(t)). \quad (RI)
\]

- Each \(\Psi(t, x, \cdot)\) is 1-homogenous: the dissipation is a Finsler structure on \(Q\), continuous and non-degenerate w.r.t. \(g\). This makes the evolution rate-independent (a.k.a. quasi-static): the solution operator commutes with monotone reparametrizations of time.

- (RI) models stick-slip dynamics, dry friction, evolution of some material properties (e.g. the Barkhausen effect in magnetization).

- We analyse a positive-temperature perturbation of (RI). As an application, this model explains the creep effects shown by such systems at positive temperature.
The discrete time incremental formulation of (RI) is, given times 
\( \{t_i = ih \mid i = 0, \ldots, T/h\} \) and the state \( z_i \) at time \( t_i \), to find the 
state \( z_{i+1} \) at time \( t_{i+1} \) that minimizes 

\[
W(z_i, z_{i+1}) := E(t_{i+1}, z_{i+1}) - E(t_i, z_i) + h\Psi (\log_{z_i} (z_{i+1})/h).
\]
The discrete time incremental formulation of (RI) is, given times \( \{t_i = i h \mid i = 0, \ldots, T/h\} \) and the state \( z_i \) at time \( t_i \), to find the state \( z_{i+1} \) at time \( t_{i+1} \) that minimizes
\[
W(z_i, z_{i+1}) := E(t_{i+1}, z_{i+1}) - E(t_i, z_i) + h \Psi(\log_{z_i}(z_{i+1}) / h).
\]

To model the effect of a heat bath with power \( \theta > 0 \) (i.e. injects energy \( \theta h \) over \([t_i, t_{i+1}]\)), we posit that the random next state \( Z^h_{i+1} \) has probability distribution \( \rho(\cdot \mid z_i) \) dVol\(_g\) on \( Q \) that minimizes
\[
\int_Q \left[ W(z_i, \cdot) \rho(\cdot \mid z_i) + \theta h \rho(\cdot \mid z_i) \log \rho(\cdot \mid z_i) \right] \text{dVol}_g
\]
i.e.
\[
\rho(z_{i+1} \mid z_i) \propto \exp \left( -\frac{W(z_i, z_{i+1})}{\theta h} \right)
\]
and consider the Markov chain \( Z^h \) with such transition probabilities.

For 2-homogeneous \( \Psi \), this procedure corresponds to adding Itō noise. What is the continuous-time limit for 1-homogeneous \( \Psi \)?
Quick back-of-envelope calculations in $T_{z_i} Q$ yield

$$E[\log_{z_i}(Z_{i+1})|Z_i = z_i] \approx -\theta h D\tilde{\Psi}^*(t_i, z_i, DE(t_i, z_i)),$$

$$\nabla[\log_{z_i}(Z_{i+1})|Z_i = z_i] \approx (\theta h)^2 D^2\tilde{\Psi}^*(t_i, z_i, DE(t_i, z_i)).$$

**Conjecture**

The variance is essentially negligible, and so the limit process as $h \to 0$ is a deterministic flow along the vector field on the RHS of the expression for the mean:

$$\dot{y}(t) = -\theta D\tilde{\Psi}^*(t, y(t), DE(t, y(t))),$$

i.e.

$$D\tilde{\Psi}(t, y(t), -\theta^{-1}\dot{y}(t)) = DE(t, y(t)),$$

(If $\Psi$ is even)

$$D\tilde{\Psi}(t, y(t), \theta^{-1}\dot{y}(t)) = -DE(t, y(t)),$$

i.e. the non-linear $\tilde{\Psi}$-gradient descent in $E$. 
Effective Dissipation — Cramer Transform

Definitions

The effective dissipation potential \( \tilde{\Psi} \) on the previous slide is the Cramer transform of \( \Psi \), defined for each \( (t, x) \in [0, T] \times Q \) by

\[
\tilde{\Psi}^\ast(t, x, \ell) := \log \int_{T_x Q} \exp \left( - \left( \langle \ell, v \rangle + \Psi(t, x, v) \right) \right) dv, \quad \ell \in T^*_x Q,
\]

\[
\tilde{\Psi}(t, x, v) := \sup \left\{ \langle \ell, v \rangle - \tilde{\Psi}^\ast(t, x, \ell) \middle| \ell \in T^*_x Q \right\}, \quad v \in T_x Q.
\]

Example

\[
\Psi(v) := \sigma \|v\|_2 \text{ on } \mathbb{R}^n, \quad \sigma > 0,
\]

\[
\tilde{\Psi}^\ast(\ell) = -\frac{n + 1}{2} \log \left( \sigma^2 - \|\ell\|_2^2 \right)
\]
Convergence Theorem

**Theorem**

Under technical conditions, the piecewise-constant interpolants of the discrete time Markov chain $Z^h$ converge in probability as $h \to 0$ to $y$, the solution of (NL), i.e.

$$D\tilde{\Psi}(t, y(t), -\theta^{-1}\dot{y}(t)) = DE(t, y(t))$$

with the same initial condition. That is, for all $\delta > 0$,

$$\lim_{h \to 0} \mathbb{P}\left[ \sup_{t \in [0, T]} d(Q, g)(Z^h(t), y(t)) \geq \delta \right] = 0.$$

**Figure:** Comparison of the original rate-independent process $z$ (blue) that solves (RI) and the thermalized process $y$ (red) that solves (NL).

(a) $\theta = 1$

(b) $\theta = \frac{1}{10}$
Main technical condition (for the moment!): the vector field

\[ f(t, x) := -\nabla_{\tilde{\Psi}^*}(t, x, \nabla E(t, x)) \]

should admit a spacetime neighbourhood of the solution \( y \) in which, for any two initial conditions \((t, x)\) and \((t, x')\) and small enough \( h > 0 \),

\[ d(Q,g)(\text{Exp}_x(hf(t, x)), \text{Exp}_{x'}(hf(t, x'))) \leq d(Q,g)(x, x'). \]

This can be seen as a combination of two criteria:
- the vector field \( f \) should not be outward-pointing;
- the curvature of \((Q, g)\) should not be strongly positive.
Application: Andrade Creep

**Andrade’s creep law (1910)**

For soft metals under constant subcritical stress, strain grows initially $\sim t^{1/3}$ and later $\sim t$.

- Work on $Q = (0, +\infty)$ with energy gradient $DE(t, x) \equiv \ell$ and the Finsler dissipation $\Psi(t, x, v) = \sigma x |v|$, *i.e.* linear strain hardening.
- Solutions to the effective evolution (NL)

\[
y(0) = 1, \quad D\tilde{\Psi}(t, y(t), -\theta^{-1} \dot{y}(t)) = \ell
\]

i.e.
\[
\dot{y}(t) = \frac{2\theta \ell}{\sigma y(t)^2 - \ell^2}
\]

do indeed grow $\sim t^{1/3}$, in accordance with Andrade’s creep law:

\[
y(t) = \left( C + \frac{6\theta \ell t}{\sigma^2} \right)^{1/3}.
\]