On Gradient Descents in Random Wiggly Energies

Tim Sullivan[1] & Florian Theil[2]

[1]California Institute of Technology, USA. tjs@caltech.edu

[2]University of Warwick, UK. f.theil@warwick.ac.uk

Microstructures in Solids: From Quantum Models to Continua

Mathematisches Forschungsinstitut Oberwolfach

15 March 2010

Outline

A Toy Model for Rate-Independence and Plasticity

- Consider a block, thought of as a point mass, sliding down a rough plane inclined at angle θ to the horizontal. For small θ, the block sticks; for large θ, it slips.

A Toy Model for Rate-Independence and Plasticity

- Consider a block, thought of as a point mass, sliding down a rough plane inclined at angle θ to the horizontal. For small θ, the block sticks; for large θ, it slips.
- From the macroscopic viewpoint, this is due to friction.
- From the microscopic viewpoint, this is due to microstructural variation; there are lots of local energy minima in which the evolution can get stuck.

A Toy Model for Rate-Independence and Plasticity

- Consider a block, thought of as a point mass, sliding down a rough plane inclined at angle θ to the horizontal. For small θ, the block sticks; for large θ, it slips.
- From the macroscopic viewpoint, this is due to friction.
- From the microscopic viewpoint, this is due to microstructural variation; there are lots of local energy minima in which the evolution can get stuck.
- We "ought" to be able to mathematically derive the macroscopic friction coefficient from the statistical properties of the microstructure.

Moral/General Theme

Microstructural variations in the energy landscape "average out" to give a qualitative change in the dissipation potential.

Barkhausen Effect

A less toy-like example with many of the same features is the Barkhausen effect, which describes the rate independent evolution of a magnetic wall in a ferromagnetic material sample under a varying applied field:

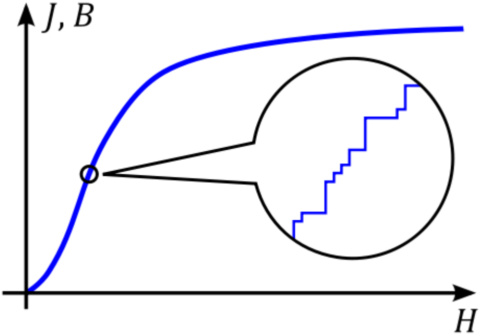

Figure: Magnetization (J) or flux density (B) as a function of applied magnetic field intesity (H). The inset shows Barkhausen jumps.

Gradient Descents — The Basics

- Many models for plastic evolutions are phrased in terms of a quantity/field of interest, $z \colon [0, T] \to \mathcal{Z}$, $\mathcal{Z}$ being some (suitably nice) linear space (*e.g.* Hilbert, Banach, $\mathrm{BV}(\Omega; \mathbb{R}^3)$, . . .).

Gradient Descents — The Basics

- Many models for plastic evolutions are phrased in terms of a quantity/field of interest, $z\colon [0,T] \to \mathcal{Z}$, $\mathcal{Z}$ being some (suitably nice) linear space (*e.g.* Hilbert, Banach, $\mathrm{BV}(\Omega;\mathbb{R}^3)$, . . .).
- The evolution of z is determined by an initial condition, an energetic potential $E\colon [0,T] \times \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ and a dissipation potential $\Psi\colon \mathcal{Z} \to [0,+\infty]$.

Gradient Descents — The Basics

- Many models for plastic evolutions are phrased in terms of a quantity/field of interest, $z\colon [0,T] \to \mathcal{Z}$, $\mathcal{Z}$ being some (suitably nice) linear space (*e.g.* Hilbert, Banach, $\mathrm{BV}(\Omega;\mathbb{R}^3)$, ...).
- The evolution of z is determined by an initial condition, an energetic potential $E\colon [0,T] \times \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ and a dissipation potential $\Psi\colon \mathcal{Z} \to [0,+\infty]$.

Example

In $\mathcal{Z} = \mathbb{R}^n$ with dissipation $\Psi = \frac{1}{2}|\cdot|^2$, we have the classical gradient descent

$$\dot{z}(t) = -\nabla E(t, z(t)).$$

Along a trajectory, the energy satisfies the energy balance

$$\frac{\mathrm{d}}{\mathrm{d}t} E(t, z(t)) = -|\dot{z}(t)|^2 + (\partial_t E)(t, z(t)).$$

Gradient Descents — Energetic Solutions

Definitions

$z\colon [0,T] \to \mathcal{Z}$ is said to be an energetic solution of the gradient descent problem in E and Ψ if z is absolutely continuous, satisfies the prescribed intitial condition, and, a.e. in $[0,T]$, the energy balance

$$\frac{\mathrm{d}}{\mathrm{d}t} E(t, z(t)) = -\big(\Psi(\dot{z}(t)) + \Psi^{\star}(\mathrm{D}E(t, z(t)))\big) + (\partial_t E)(t, z(t)),$$

where $\Psi^{\star}\colon \mathcal{Z}^{\star} \to \mathbb{R} \cup \{+\infty\}$ is the convex conjugate of Ψ:

$$\Psi^{\star}(\ell) := \sup\{\langle \ell, x\rangle - \Psi(x) \mid x \in \mathcal{Z}\}.$$

Much of this carries over to state spaces with no linear structure: see Ambrosio, Gigli & Savaré (2008), *Gradient Flows in Metric Spaces and in the Space of Probability Measures*.

Gradient Descents — Energy Inequality

- Often we work with the integrated form of the energy balance equation instead: for every $[a, b] \subseteq [0, T]$,

$$0 = E(b, z(b)) - E(a, z(a)) + \int_a^b \big(\Psi(\dot{z}(t)) + \Psi^\star(\mathrm{D}E(t, z(t))) - (\partial_t E)(t, z(t))\big)\,\mathrm{d}t.$$

Gradient Descents — Energy Inequality

- Often we work with the integrated form of the energy balance equation instead: for every $[a, b] \subseteq [0, T]$,

$$0 = E(b, z(b)) - E(a, z(a)) + \int_a^b \big(\Psi(\dot{z}(t)) + \Psi^\star(\mathrm{D}E(t, z(t))) - (\partial_t E)(t, z(t)) \big) \, \mathrm{d}t.$$

- In this equality, $\leq$ always holds, so it is enough to check whether or not the following energy inequality holds: for every $[a, b] \subseteq [0, T]$,

$$0 \geq E(b, z(b)) - E(a, z(a)) + \int_a^b \big(\Psi(\dot{z}(t)) + \Psi^\star(\mathrm{D}E(t, z(t))) - (\partial_t E)(t, z(t)) \big) \, \mathrm{d}t.$$

Rate Independent Processes

- A rate-independent evolution is one "with no time-scale of its own", one for which time-reparametrized solutions are solutions to the time-reparametrized problem. In terms of the above set-up, this corresponds to Ψ being homogeneous of degree one.

Rate Independent Processes

- A rate-independent evolution is one "with no time-scale of its own", one for which time-reparametrized solutions are solutions to the time-reparametrized problem. In terms of the above set-up, this corresponds to Ψ being homogeneous of degree one.
- In this case, $\Psi^\star$ only takes the values 0 and $+\infty$ and we can re-write the definition of an energetic solution in terms of an energy constraint and a stability constraint:

$$0 \geq E(b, z(b)) - E(a, z(a)) + \int_a^b \left(\Psi(\dot{z}(t)) - (\partial_t E)(t, z(t)) \right) \mathrm{d}t.$$

$$-\mathrm{D}E(t, z(t)) \in \mathscr{E} := \{ \ell \in \mathcal{Z}^\star \mid \Psi^\star(\ell) = 0 \}.$$

- We call $\mathscr{E}$ the elastic region and call $\mathcal{S}(t) := \{ x \mid -\mathrm{D}E(t, x) \in \mathscr{E} \}$ the (locally) stable region at time t.

Rate Independent Processes

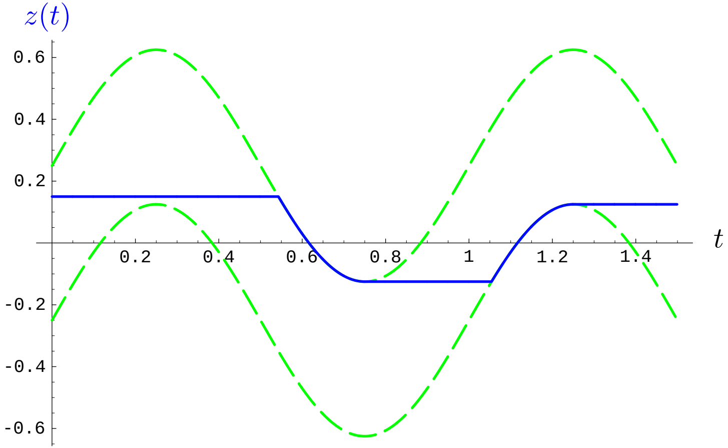

Figure: In blue, a typical rate-independent evolution in one dimension. The frontier of the stable region is shown in green.

What We Seek

We seek theorems of the following type:

Theorem ("Proto-theorem")

If E_ε is a suitable random (spatial) perturbation of E, then there exists a 1-homogeneous dissipation potential Ψ such that if z_ε solves the wiggly classical gradient descent

$$\dot{z}_\varepsilon(t) = -\frac{1}{\varepsilon}\nabla E_\varepsilon(t, z_\varepsilon(t)),$$

and z solves the rate-independent problem in E and Ψ,

$$\partial\Psi(\dot{z}(t)) \ni -\mathrm{D}E(t, z(t)),$$

then $z_\varepsilon \to z$ in some sense as $\varepsilon \to 0$.

We expect Ψ to depend on the structure of the perturbation $E_\varepsilon - E$.

Previous Results

- Abeyaratne–Chu–James 1996: in $n = 1$ with periodic perturbations, up to a subsequence,

$$z_\varepsilon \to z \text{ uniformly on } [0, T] \text{ and } \dot{z}_\varepsilon \xrightarrow{*} \dot{z} \text{ in } L^\infty([0, T]; \mathbb{R}).$$

Previous Results

- Abeyaratne–Chu–James 1996: in $n = 1$ with periodic perturbations, up to a subsequence,

$$z_\varepsilon \to z \text{ uniformly on } [0, T] \text{ and } \dot{z}_\varepsilon \overset{*}{\rightharpoonup} \dot{z} \text{ in } L^\infty([0, T]; \mathbb{R}).$$

- Menon 2002: in $n = 2$, periodic perturbations, same result as in $n = 1$ but with *caveats* — horrible grid effects and resonance zones.

Previous Results

- Abeyaratne–Chu–James 1996: in $n = 1$ with periodic perturbations, up to a subsequence,

$$z_\varepsilon \to z \text{ uniformly on } [0, T] \text{ and } \dot{z}_\varepsilon \overset{*}{\rightharpoonup} \dot{z} \text{ in } L^\infty([0, T]; \mathbb{R}).$$

- Menon 2002: in $n = 2$, periodic perturbations, same result as in $n = 1$ but with *caveats* — horrible grid effects and resonance zones.

Periodicity is a rather unnatural assumption to have to make and — as Menon's results show — it even introduces some undesirable features.

1-Dimensional Set-Up

- Consider the moving uniformly convex energy

$$E(t, x) := V(x) - \ell(t)x,$$

where $V \in \mathcal{C}^3(\mathbb{R}; \mathbb{R})$ is uniformly convex and $\ell\colon [0, T] \to \mathbb{R}^*$ is uniformly Lipschitz.

1-Dimensional Set-Up

- Consider the moving uniformly convex energy

$$E(t, x) := V(x) - \ell(t)x,$$

where $V \in \mathcal{C}^3(\mathbb{R}; \mathbb{R})$ is uniformly convex and $\ell \colon [0, T] \to \mathbb{R}^*$ is uniformly Lipschitz.

- The perturbed energy will be

$$E_\varepsilon(t, x) := E(t, x) + \varepsilon G(x/\varepsilon),$$

where

$$g := -G' \colon \Omega \times \mathbb{R} \to [-\sigma, +\sigma]$$

is $\mathbb{P}$-almost surely defined, continuous and surjective.

1-Dimensional Set-Up

- Consider the moving uniformly convex energy

$$E(t,x) := V(x) - \ell(t)x,$$

where $V \in \mathcal{C}^3(\mathbb{R};\mathbb{R})$ is uniformly convex and $\ell \colon [0,T] \to \mathbb{R}^*$ is uniformly Lipschitz.

- The perturbed energy will be

$$E_\varepsilon(t,x) := E(t,x) + \varepsilon G(x/\varepsilon),$$

where

$$g := -G' \colon \Omega \times \mathbb{R} \to [-\sigma, +\sigma]$$

is $\mathbb{P}$-almost surely defined, continuous and surjective.

- We will show that if G is "wiggly enough", then the wiggles "average out" as $\varepsilon \to 0$ to give the 1-homogeneous dissipation potential $\Psi := \sigma|\cdot|$.

How Wiggly is "Wiggly Enough"?

Definition

Fix $\sigma > 0$. For a continuous, surjective function $g\colon \mathbb{R} \to [-\sigma, +\sigma]$, define $D_0^+ \geq 0$ to be the least $x > 0$ such that $g(x) = -\sigma$; inductively define D_{n+1}^+ to be the least positive number such that g takes both values $-\sigma$ and $+\sigma$ in the interval

$$\left(\sum_{i=0}^{n} D_i^+, \sum_{i=0}^{n+1} D_i^+ \right];$$

and define $D_n^- \leq 0$ similarly. Then g is said to have property (✠) if

- $D_n^\pm$ exists and is finite for all n;
- $\sum_{n=0}^\infty D_n^\pm = \pm\infty$;
- $\lim_{n\to\infty} \left(D_{n+1}^\pm / \sum_{i=0}^n D_i^\pm \right) = 0$.

1-Dimensional Convergence Theorem

Theorem (S. & T. 2007)

Let E, E_ε, Ψ be as above, and

$$\dot{z}_\varepsilon(t) = -\frac{1}{\varepsilon} E'_\varepsilon(t, z_\varepsilon(t)),$$

$$\Psi(\dot{z}(t)) \ni -E'(t, z(t)).$$

Then $z_\varepsilon \to z$ in probability (and hence in distribution) in $\mathcal{C}^0([0,T];\mathbb{R})$ as $\varepsilon \to 0$ if, and only if, g has property (✠). That is, for any $\delta > 0$,

$$\mathbb{P}\left[\sup_{0 \le t \le T} |z_\varepsilon(t) - z(t)| \ge \delta\right] \to 0 \text{ as } \varepsilon \to 0.$$

Hence, up to subsequences, $z_\varepsilon \to z$ uniformly on $[0,T]$, $\mathbb{P}$-almost surely.

n-Dimensional Set-Up

- For simplicity, we consider a moving quadratic energy $E(t,x) := \frac{1}{2}x \cdot Ax - \ell(t) \cdot x$, $A \in \mathbb{R}^{n \times n}$ postive definite, ℓ Lipschitz.

n-Dimensional Set-Up

- For simplicity, we consider a moving quadratic energy $E(t,x) := \frac{1}{2}x \cdot Ax - \ell(t) \cdot x$, $A \in \mathbb{R}^{n \times n}$ postive definite, ℓ Lipschitz.
- We randomly "dent" E by adding to it the dent function

$$D(x;y,\varepsilon) := \frac{\sigma}{2}\left(\left|\frac{x-y}{\varepsilon}\right|^2 - 1\right)_-$$

for $y \in$ the points of a dilute Poisson point process $\mathcal{O}$ of intensity ε^{-p}; for technical reasons, we require that $p \in (n-1, n)$. Set

$$E_\varepsilon(t,x) := E(t,x) + \sum_{y \in \mathcal{O}} D(x;y,\varepsilon).$$

n-Dimensional Set-Up

- For simplicity, we consider a moving quadratic energy $E(t,x) := \frac{1}{2} x \cdot Ax - \ell(t) \cdot x$, $A \in \mathbb{R}^{n \times n}$ postive definite, ℓ Lipschitz.
- We randomly "dent" E by adding to it the dent function

$$D(x; y, \varepsilon) := \frac{\sigma}{2} \left(\left| \frac{x-y}{\varepsilon} \right|^2 - 1 \right)_-$$

for $y \in$ the points of a dilute Poisson point process $\mathcal{O}$ of intensity ε^{-p}; for technical reasons, we require that $p \in (n-1, n)$. Set

$$E_\varepsilon(t,x) := E(t,x) + \sum_{y \in \mathcal{O}} D(x; y, \varepsilon).$$

- Since the dents are isotropic, we expect that the dissipation potential for the hoped-for rate-independent limit will be isotropic as well; set $\Psi := \sigma | \cdot |$.

n-Dimensional Convergence Theorem

Theorem (S. & T. 2009)

Let E, E_ε, Ψ be as above, and

$$\dot{z}_\varepsilon(t) = -\frac{1}{\varepsilon} \nabla E_\varepsilon(t, z_\varepsilon(t)),$$

$$\Psi(\dot{z}(t)) \ni -\mathrm{D}E(t, z(t)).$$

Then $z_\varepsilon \to z$ in probability (and hence in distribution) in $\mathcal{C}^0([0,T];\mathbb{R}^n)$ as $\varepsilon \to 0$. That is, for any $\delta > 0$,

$$\mathbb{P}\left[\sup_{0 \leq t \leq T} |z_\varepsilon(t) - z(t)| \geq \delta\right] \to 0 \text{ as } \varepsilon \to 0.$$

Hence, up to subsequences, $z_\varepsilon \to z$ uniformly on $[0,T]$, $\mathbb{P}$-almost surely.

Strategy of the Proof

For $[a,b] \subseteq [0,T]$, define the energy surplus of $u \colon [a,b] \to \mathbb{R}^n$ by the L^∞-lower semicontinuous functional $\mathrm{ES}(-,[a,b]) \colon \mathrm{BV}([a,b];\mathbb{R}^n) \to \mathbb{R}$

$$\mathrm{ES}(u,[a,b]) := E(b,u(b)) - E(a,u(a)) + \int_a^b \left(\Psi(\dot{u}(t)) - (\partial_t E)(t,u(t))\right) \mathrm{d}t.$$

This is the amount by which the desired energy inequality fails to hold.

Strategy of the Proof

For $[a, b] \subseteq [0, T]$, define the energy surplus of $u \colon [a, b] \to \mathbb{R}^n$ by the L^∞-lower semicontinuous functional $\mathrm{ES}(-, [a, b]) \colon \mathrm{BV}([a, b]; \mathbb{R}^n) \to \mathbb{R}$

$$\mathrm{ES}(u, [a, b]) := E(b, u(b)) - E(a, u(a)) + \int_a^b \left(\Psi(\dot{u}(t)) - (\partial_t E)(t, u(t)) \right) \mathrm{d}t.$$

This is the amount by which the desired energy inequality fails to hold. We show that

- $(z_\varepsilon)_{\varepsilon > 0}$ is tight (has a uniformly convergent subsequence);
- $\liminf_{\varepsilon \to 0} \mathrm{ES}(z_\varepsilon, [0, T]) \leq 0$;
- any such uniform limit will satisfy stability;
- uniqueness results (*e.g.* Mielke–T. 2004) for rate-independent processes imply that the limit process must be z.

An Important Observation

- It follows from the set-up that if z_ε enters a dent $\mathbb{B}_\varepsilon(y)$, $y \in \mathcal{O}$, and that dent is stable is contained within the stable region, then z_ε cannot leave $\mathbb{B}_\varepsilon(y)$. Moreover, z_ε leaves $\mathbb{B}_\varepsilon(y)$ precisely at

$$\tau^{\mathrm{out}} = \inf\{t \mid \mathbb{B}_\varepsilon(y) \cap \mathcal{S}(t) = \emptyset\}.$$

An Important Observation

- It follows from the set-up that if z_ε enters a dent $\mathbb{B}_\varepsilon(y)$, $y \in \mathcal{O}$, and that dent is stable is contained within the stable region, then z_ε cannot leave $\mathbb{B}_\varepsilon(y)$. Moreover, z_ε leaves $\mathbb{B}_\varepsilon(y)$ precisely at

$$\tau^{\mathrm{out}} = \inf\{t \mid \mathbb{B}_\varepsilon(y) \cap \mathcal{S}(t) = \emptyset\}.$$

- This observation helps to keep everything under control: even though z_ε falls from one dent to another at speed $\sim \frac{1}{\varepsilon}$, it must then remain in a dent for a time period inversely proportional to the distance fallen, where it waits for $\partial\mathcal{S}(t)$ to "catch up".

Dent Entry and Exit Times

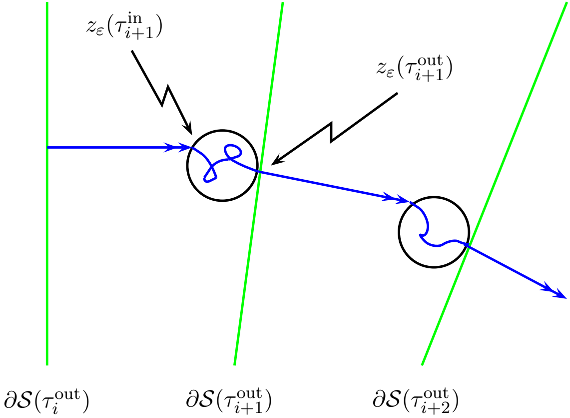

Figure: A “top-down” schematic illustration of z_ε (blue). The frontier of the stable region is shown in green at the three exit times; everything to the right of the green line is the stable region at that time. Dents are shown as black circles.

Dent Entry and Exit Times

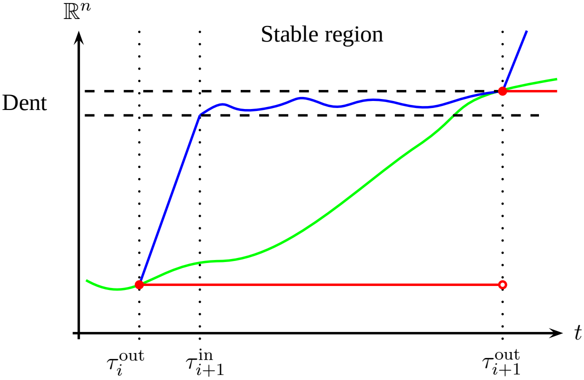

Figure: A "cross-sectional" schematic illustration of z_ε (blue). The frontier of the stable region is shown in green, and the piecewise-constant càdlàg solution to the Moreau–Yosida incremental formulation of the rate independent problem is shown in red.

Sketch of the Proof

- In what follows, for simplicity, it will be assumed that dents never overlap.
- In practice, overlaps can happen, and one must use statistical properties of the Poisson point process $\mathcal{O}$ to ensure that they do not happen "too often" and thereby ruin the total variation estimates.
- One could condition the process $\mathcal{O}$ to rule out overlaps (*e.g.* Matérn clustering and hard core processes), but would thereby lose explicit representation of the distance-to-nearest-neighbour distribution.

Sketch of the Proof

Asymptotic stability is easy to get, and tightness will follow from the energy estimates. The following lemma controls the energy surplus:

Lemma (Variation and energy surplus control)

If $z_\varepsilon|_{[a,b]}$ lies wholly outside all dents, then

$$\left|\mathrm{Var}_{[a,b]}(z_\varepsilon) - |z_\varepsilon(b) - z_\varepsilon(a)|\right| \leq C\left(\frac{|b-a|}{\|A\|} + \frac{|b-a|^2}{\varepsilon}\right),$$

and if $z_\varepsilon|_{[a,b]}$ lies wholly inside a dent, then

$$\mathrm{Var}_{[a,b]}(z_\varepsilon) \leq C\varepsilon.$$

Hence,

$$\mathrm{ES}(z_\varepsilon, [\tau_i^{\mathrm{out}}, \tau_{i+1}^{\mathrm{out}}]) \leq C\varepsilon + \frac{C'\sigma|\tau_{i+1}^{\mathrm{in}} - \tau_i^{\mathrm{out}}|^2}{\varepsilon}.$$

Sketch of the Proof

Armed with

$$\mathrm{ES}(z_\varepsilon, [\tau_i^{\mathrm{out}}, \tau_{i+1}^{\mathrm{out}}]) \leq C\varepsilon + \frac{C'\sigma|\tau_{i+1}^{\mathrm{in}} - \tau_i^{\mathrm{out}}|^2}{\varepsilon},$$

we just need to make sure that the rapid descents don't last too long, and that there are not so many of them that all these order ε errors will accumulate and ruin all our estimates as we take the limit $\varepsilon \to 0$. We get this control from the observation about waiting times and the distribution of the Poisson point process $\mathcal{O}$:

Sketch of the Proof

Armed with

$$\mathrm{ES}(z_\varepsilon, [\tau_i^{\mathrm{out}}, \tau_{i+1}^{\mathrm{out}}]) \leq C\varepsilon + \frac{C'\sigma|\tau_{i+1}^{\mathrm{in}} - \tau_i^{\mathrm{out}}|^2}{\varepsilon},$$

we just need to make sure that the rapid descents don't last too long, and that there are not so many of them that all these order ε errors will accumulate and ruin all our estimates as we take the limit $\varepsilon \to 0$. We get this control from the observation about waiting times and the distribution of the Poisson point process $\mathcal{O}$:

Proposition (Energy surplus goes to zero in mean square)

$$\mathbb{E}\big[\mathrm{ES}(z_\varepsilon, [0,T])\big] \leq CT\varepsilon^{p-n+1} \to 0,$$

$$\mathbb{V}\big[\mathrm{ES}(z_\varepsilon, [0,T])\big] \leq CT\varepsilon^{p-n+2} \to 0.$$

Conclusions and Outlook

To conclude, we have rigorously established a passage from a viscous evolution in a random energy landscape to a rate-independent evolution in the limit of the random landscape.

Conclusions and Outlook

To conclude, we have rigorously established a passage from a viscous evolution in a random energy landscape to a rate-independent evolution in the limit of the random landscape.

What's next?

- Anisotropic dents and dissipation potentials.
- Perturbations/dents without *a priori* bounds on $\nabla(E_\varepsilon - E)$.
- Extension to energies that are more general than quadratic forms? What if E is only uniformly convex? What about strictly convex, convex, or non-convex energies?
- Extension to infinite-dimensional spaces $\mathcal{Z}$?