

# On the behavior of dissipative systems in contact with a heat bath: Application to Andrade creep

T. Sullivan, M. Koslowski, F. Theil and M. Ortiz

March 15, 2009

## Abstract

We develop a theory of statistical mechanics for dissipative systems governed by equations of evolution that assigns probabilities to individual trajectories of the system. The theory is made mathematically rigorous and leads to precise predictions regarding the behavior of dissipative systems at finite temperature. Such predictions include the effect of temperature on yield phenomena and rheological time exponents. The particular case of an ensemble of dislocations moving in a slip plane through a random array of obstacles is studied numerically in detail. The numerical results bear out the analytical predictions regarding the mean response of the system, which exhibits Andrade creep.

## 1 Introduction

The fundamental question addressed in this work is: *How does a dissipative system behave when it is placed in contact with a heat bath?* For Hamiltonian systems this question can be addressed within the framework of the classical canonical ensemble of statistical mechanics, and the resulting probability of observing a state of the system is given by the Gibbs distribution. Alternatively, the Gibbs distribution follows from Bogoliubov's variational principle (cf, e.g., [8]), which can also be taken as a basis for formulating approximate mean-field theories. The work presented in this paper endeavors to develop a similar framework for dissipative systems governed by equations of evolution. Remarkably, this development requires an extension of classical statistical mechanics. More precisely, whereas classical statistical mechanics assigns probabilities to states of the system, the statistical mechanics of dissipative systems developed here assigns probabilities to state increments or rates and, by extension, to entire trajectories of the system. The probability distributions that result from the present theory may be regarded as generalizations of the Gibbs distribution of classical mechanics to trajectories. In particular, the computation of the probability of a particular state at a given instant of time requires the evaluation of a path integral over the trajectories of the system.

For definiteness, we confine our attention to systems whose state is described by an  $N$ -dimensional array  $x \in \mathbb{R}^N$  of generalized coordinates. The energetics of the system are described by an energy function  $E(t, x)$ . The explicit time dependence of  $E$  may arise, for example, as a result of the application to the system of a time-dependent external field. In addition, the system is assumed to be *dissipative*, and, therefore, the equilibrium equations are of the form:

$$0 \in \partial_x E(t, x) + \partial \Psi(\dot{x}) \quad (1.1)$$

where  $\Psi$  is a dissipation potential. These equilibrium equations define a set of ordinary differential equations which, given appropriate conditions at, say,  $t = 0$ , can be solved uniquely for the trajectory  $x(t)$ ,  $t \geq 0$ . If  $\Psi(z) = \frac{1}{2}|z|^2$ , then one obtains the classical gradient flow in the potential  $E$ . If  $\Psi$  is homogeneous of degree one, then the system becomes rate-independent: i.e., if  $x(t)$  solves (1.1) for  $E$  and  $\varphi$  is a strictly monotone reparametrization of time, then  $\tilde{x}(t) = x(\varphi(t))$  solves (1.1) for  $\tilde{E}(t, \cdot) = E(\varphi(t), \cdot)$ . To include the effect of thermal noise we leave the purely deterministic framework and consider stochastic systems in which the trajectories are no longer unique but random and define a probability measure on path space. More precisely, the time continuum  $\{t \geq 0\}$  is replaced by a discrete set of times  $0 = t_0 < t_1 < \dots$  and we consider the Markov chain with transition probability densities proportional to

$$\exp\left(-\left(E(t_{n+1}, x_{n+1}) + |t_{n+1} - t_n| \Psi\left(\frac{x_{n+1} - x_n}{t_{n+1} - t_n}\right)\right) / \varepsilon_n\right).$$

The parameters  $\varepsilon_n$  measure the strength of the thermal fluctuations. The second step is to determine the limiting process which is obtained when the fineness of the time discretization  $h := \sup_n(t_{n+1} - t_n)$  tends to 0 and  $\varepsilon_n$  is scaled accordingly. In the classical case of linear kinetics, in which  $\Psi(z) = \frac{1}{2}|z|^2$ , the limiting path measure is generated by well-known stochastic differential equations. This paper will mostly focus on the rate-independent case, in which  $\Psi$  is 1-homogeneous. In this case, the limiting measure is generated by ordinary differential equations, i.e. the evolution is deterministic. However, the differential equation depends in a nontrivial way on  $E$  and  $\Psi$ : the evolution takes the form

$$\dot{x} = -DF^0(DE(t, x)),$$

where  $F^0$  is determined by  $\Psi$  and plays the rôle of an *effective dual dissipation potential*. This object appears to be a novel contribution to the literature, and will be introduced and analyzed in section 3.

As a representative example of application, the theory is applied to the case of Andrade creep in metals; the theory shows that Andrade creep represents the mean-field behavior of an ensemble of linear elastic dislocations moving on a slip plane through a random array of obstacles under the action of an applied resolved shear stress at finite temperature. In order to establish this connection, the energetics and dynamics of the dislocation ensemble are described by means of the Koslowski–Cuitiño–Ortiz phase-field model [19]. The resulting governing equations are simplified by means of a mean-field approximation. The analysis then shows that the slip strain grows as a 1/3 power of time, i.e. the system exhibits Andrade creep.

## 1.1 Outline of the paper

In section 2, the discrete time Markov chain is obtained via a minimization principle where  $\varepsilon$  appears as a parameter of an interior-point regularization. A simple calculation shows that in the case of linear kinetics, in which the friction potential  $\Psi$  is a 2-homogeneous function of the velocity, the continuous-time limit is given by an Itô stochastic differential equation.

In section 3, the limiting continuous-time process is identified in the case where the potential energy  $E$  is quadratic and  $\Psi$  is 1-homogeneous. The linear scaling of  $\Psi$  with the velocity is an essential feature of rate-independent behavior. However, the coupling to the heat bath destroys the rate-independence in a controlled way. The key result is that the limiting dynamics is a generalized gradient flow generated by an effective dual dissipation potential which is given by the logarithm of a nontrivial partition function. Furthermore, this partition function is an intrinsic object that is determined by the geometry of the elastic region associated to  $\Psi$ . In several cases that are important for applications a representation of the partition function in closed form can be found. Mathematical rigor is claimed in this section.

In section 4, this framework is applied to dislocation dynamics. An infinite-dimensional model, which was introduced in [19], is reduced to a finite-dimensional system involving only the elastic displacement on the pinning sites using a formula from section 3. A further reduction step results in a scalar ordinary differential equation which can be shown to create the  $t^{1/3}$  law of Andrade creep if linear strain hardening is assumed.

The complete proofs of the theorems in section 3 are given in the appendix (section 6), along with some remarks about the convergence of random variables.

## 2 Heuristic derivation and examples

We proceed to introduce the basic scheme that accords probabilities to the trajectories of a dissipative system and to illustrate the scheme by means of an elementary example. The first step in the derivation is to discretize the equations of evolution in time and to reinterpret the time-discretized equations as defining the optimal transport of probability measures. An interior-point regularization of the transport problem then effectively places the system in contact with a thermal bath. Conveniently, the regularized incremental problem can be solved explicitly in closed form for the transition probability distribution. Iterating the solution via the Chapman–Kolmogorov equation for Markovian processes and formally taking the limit of vanishingly small time steps finally yields the sought trajectory probabilities.

### 2.1 Heuristic derivation

If the system is conservative, i.e.  $\Psi(\dot{x}) \equiv 0$ , then the instantaneous state  $x(t)$  of the system at time  $t$  follows directly from energy minimization, i.e., from the problem:

$$E(t, x(t)) = \inf_{y \in \mathbb{R}^N} E(t, y). \quad (2.1)$$

This variational framework can be extended to dissipative systems by time discretization [24, 25, 27]. Specifically, consider a discrete time incremental process consisting of a sequence of states  $x_n \in \mathbb{R}^N$  at times  $t_n = nh$ ,  $h > 0$  fixed, and introduce the incremental work function  $W$ :

$$W(x_n, x_{n+1}) = E(t_{n+1}, x_{n+1}) - E(t_n, x_n) + \inf_{\text{paths}} \int_{t_n}^{t_{n+1}} \Psi(\dot{x}(t)) dt \quad (2.2)$$

where the infimum is taken over all almost-everywhere-differentiable paths  $x: [t_n, t_{n+1}] \rightarrow \mathbb{R}^N$  such that  $x(t_n) = x_n$  and  $x(t_{n+1}) = x_{n+1}$ ; however, this work function will in much of the sequel be approximated by

$$C(x_n, x_{n+1}) := E(t_{n+1}, x_{n+1}) - E(t_n, x_n) + |t_{n+1} - t_n| \Psi \left( \frac{x_{n+1} - x_n}{t_{n+1} - t_n} \right). \quad (2.3)$$

The fundamental property of the incremental work function  $W$  is that it acts as a potential for the net forces  $f_{n+1}$  at time  $t_{n+1}$ , i.e.,

$$f_{n+1} = \partial_{x_{n+1}} W(x_n, x_{n+1}). \quad (2.4)$$

From this property it follows that the equilibrium equation  $f_{n+1} = 0$  is the Euler–Lagrange equation corresponding to the minimum principle:

$$\inf_{x_{n+1}} W(x_n, x_{n+1}). \quad (2.5)$$

It should be carefully noted that  $W(x_n, x_{n+1})$  is determined by both the *energetics* and the *kinetics* of the system and depends on both the initial and the final configuration of the system over the time step.

The incremental problem above can be re-interpreted as an *optimal transport* problem in which the incremental work function  $W$  as defined by (2.2) is regarded as a *cost function*, i.e., as the cost of rearranging the system from one configuration,  $x_n$ , to another one,  $x_{n+1}$ . Consider the probability density functions  $\rho_n: \mathbb{R}^N \rightarrow \mathbb{R}$  describing the state of the system at each time step  $t_n = nh$ . The corresponding optimal transport problem for the  $(n+1)^{\text{th}}$  time step is:

$$\inf \mathcal{F}_h[u], \quad (2.6a)$$

$$\text{subject to: } \int_{\mathbb{R}^N} u(x_n, x_{n+1}) dx_{n+1} = \rho_n(x_n), \quad (2.6b)$$

$$\iint_{\mathbb{R}^N} u(x_n, x_{n+1}) dx_n dx_{n+1} = 1, \quad (2.6c)$$

where  $u(x_n, x_{n+1}) = \rho_{n+1}(x_{n+1}|x_n)\rho_n(x_n)$  is the joint probability density of two states  $x_n$  and  $x_{n+1}$  at successive time steps and

$$\mathcal{F}_h[u] := \iint_{\mathbb{R}^N} W(x_n, x_{n+1}) u(x_n, x_{n+1}) dx_n dx_{n+1}. \quad (2.7)$$

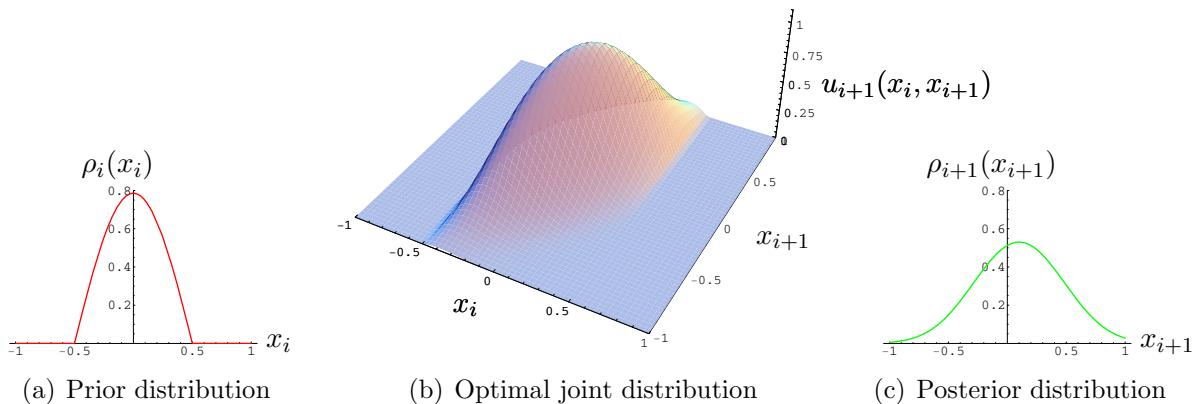


Figure 2.1: A schematic example of a single step in the optimal transport scheme: (a) a prior probability density  $\rho_i$ ; (b) the joint probability density  $u_{i+1}$  minimizing the expected-cost-with-interior-point-regularization functional and having  $\rho_i$  as its first marginal; (c) the posterior probability density  $\rho_{i+1}$ , equal to the second marginal of  $u_{i+1}$ .

In addition, it follows that

$$\rho_{n+1}(x_{n+1}) = \int_{\mathbb{R}^N} u(x_n, x_{n+1}) dx_n. \quad (2.8)$$

Note that the problem (2.6) differs from Kantorovich’s formulation of the optimal transport problem [18] in the following respect: in the Kantorovich problem on a state space  $\mathbb{X}$ , the objective is to minimize the expected cost with respect to probability measures on  $\mathbb{X}^2$  with specified first and second marginals on each copy of  $\mathbb{X}$ ; in the problem (2.6), only the first marginal is specified. The problem is thus more akin to a generalized gradient flow. See [1] for an introduction to optimal transport theory and the theory of gradient flows in spaces that do not have a linear structure.

Problem (2.6) is a linear programming problem. Introducing Lagrange multipliers  $\lambda: \mathbb{R}^N \rightarrow \mathbb{R}$  and  $\mu \in \mathbb{R}$ , define the Lagrangian

$$\mathcal{L}_h[u] := \mathcal{F}_h[u] + \iint_{\mathbb{R}^N} u(x_n, x_{n+1}) \lambda(x_n) dx_n dx_{n+1} + \mu \int_{\mathbb{R}^N} u(x_n, x_{n+1}) dx_n dx_{n+1}. \quad (2.9)$$

In order to model the effect of a heat bath of “temperature”  $\varepsilon > 0$ , additionally introduce the interior-point regularization:

$$\mathcal{L}_h^\varepsilon[u] := \mathcal{L}_h[u] + \varepsilon \int_{\mathbb{R}^N} u(x_n, x_{n+1}) \log u(x_n, x_{n+1}) dx_n dx_{n+1}. \quad (2.10)$$

The second term on the right-hand side is the negative of the well-known Gibbs–Boltzmann entropy functional for the probability density  $u$ , and has the effect of heavily penalizing

deterministic, coherent evolutions. The stationary point of the extended Lagrangian (2.10) is

$$u(x_n, x_{n+1}) = \exp\left(-\frac{W(x_n, x_{n+1}) + \lambda(x_n) + \mu + \varepsilon}{\varepsilon}\right). \quad (2.11)$$

Using the constraints (2.6b) and (2.6c) this becomes

$$u(x_n, x_{n+1}) = \frac{\rho_n(x_n)}{Z(x_n)} \exp\left(-\frac{W(x_n, x_{n+1})}{\varepsilon}\right), \quad (2.12)$$

where the partition function  $Z$  is given by

$$Z(x_n) := \int_{\mathbb{R}^N} \exp\left(-\frac{W(x_n, x_{n+1})}{\varepsilon}\right) dx_{n+1}. \quad (2.13)$$

Finally, (2.8) yields the explicit representation

$$\rho_{n+1}(x_{n+1}) = \int_{\mathbb{R}^N} \frac{\rho_n(x_n)}{Z(x_n)} \exp\left(-\frac{W(x_n, x_{n+1})}{\varepsilon}\right) dx_n. \quad (2.14)$$

Note that, in order for (2.14) to make sense, the work function  $W(x_n, x_{n+1})$  must grow rapidly enough as  $|x_{n+1}| \rightarrow +\infty$  to ensure that  $Z(x_n)$  is finite for each  $x_n \in \mathbb{R}^N$ . Physically, setting physical temperature  $\Theta = \varepsilon/k_B$  in terms of Boltzmann's constant  $k_B$ ,  $\varepsilon$  can be interpreted as the temperature of heat bath with which the system is coupled. Note the similarity to the unconstrained minimization principle for  $u$  given in [17]. The above derivation is valid subject to mild assumptions on  $E$  and  $\Psi$ : for example, being continuous and bounded below, and with  $e^{-W(x_n, \cdot)}$  integrable for every  $x_n$ . It is an easy application of Fubini's theorem and Lebesgue's dominated convergence theorem to extend the result to more general prior probability distributions (which might not have a densities with respect to Lebesgue measure).

Recall that  $u(x_n, x_{n+1})$  is the joint probability density for two states  $x_n$  and  $x_{n+1}$  at successive time steps  $t_n$  and  $t_{n+1}$  respectively. Iterating (2.14) via the Chapman–Kolmogorov equation yields

$$\rho_n(x_n) = \int \cdots \int_{\mathbb{R}^N} \left( \prod_{i=0}^{n-1} \frac{e^{-W(x_i, x_{i+1})/\varepsilon}}{Z(x_i)} \right) \rho_0(x_0) dx_{n-1} \cdots dx_0 \quad (2.15)$$

In the limit of  $h \rightarrow 0$  and  $n \rightarrow \infty$  at fixed  $t = nh$ , (2.15) defines a path integral of the form

$$u(x_0, x_t) = \int \cdots \int \exp\left(-\mathcal{I}[x]/\varepsilon\right) \mathcal{D}x \quad (2.16)$$

for some action functional  $\mathcal{I}$  (see, for example, [30]). This path integral gives the joint probability  $u(x_0, x_t)$  of finding the system in state  $x_t$  at time  $t \geq 0$  and that the system started in state  $x_0$  at time  $t = 0$ . The integrand of (2.16) may be interpreted as relating

to the probability of individual trajectories  $x(t)$ . More precisely, given two trajectories  $x(t)$  and  $y(t)$ , their relative probabilities are:

$$\frac{p(x)}{p(y)} = \frac{\exp(-\mathcal{I}[x]/\varepsilon)}{\exp(-\mathcal{I}[y]/\varepsilon)}$$

The limit  $\varepsilon \downarrow 0$  characterizes the trajectories of the system as the interior-point regularization (2.10) is removed. In the limit, the measure under the integral in (2.16) is expected to be concentrated on the *minimal paths*, i.e., on the solutions of the minimum principle

$$\inf_x \mathcal{I}[x]$$

This in turn supplies a minimum-principle characterization of the trajectories of problem (1.1).

## 2.2 An elementary example

The nature of the framework just outlined may be illustrated by means of a first elementary example. For simplicity,  $N$  is taken to be 1 (although the same analysis applies for general  $N \in \mathbb{N}$ ) and the system consists of an energetic potential and dissipative potential of the form

$$E(t, x) = -\ell(t)x; \tag{2.17a}$$

$$\Psi(\dot{x}) = \frac{\eta}{2}\dot{x}^2; \tag{2.17b}$$

where  $x \in \mathbb{R}$ ,  $\eta > 0$  is the dashpot viscosity, and  $\ell$  is a time-dependent applied load. Note that the dissipative potential  $\Psi$  is positively homogeneous of degree two. In this case, the equilibrium equation (1.1) is

$$0 = -\ell(t) + \eta\dot{x}(t) \tag{2.18}$$

and the approximate incremental work function is given by

$$C(t_n, x_n, t_{n+1}, x_{n+1}) = -\ell(t_{n+1})x_{n+1} + \ell(t_n)x_n + \frac{\eta}{2h}(x_{n+1} - x_n)^2.$$

In this case, the density for the conditional transition probability is

$$\rho_{n+1}^{h,\varepsilon}(x_{n+1}|x_n) = \frac{\exp\left(-\frac{\eta}{2h\varepsilon}\left(x_{n+1} - x_n - \frac{h}{\eta}\ell(t_{n+1})\right)^2\right)}{\int_{\mathbb{R}} \exp\left(-\frac{\eta}{2h\varepsilon}\left(x_{n+1} - x_n - \frac{h}{\eta}\ell(t_{n+1})\right)^2\right) dx_{n+1}}, \tag{2.19}$$

which is the probability density function for a normally distributed random variable with mean  $x_n + \frac{h}{\eta}\ell(t_{n+1})$  and variance  $\frac{h\varepsilon}{\eta}$ . Note (in contrast to what will follow in the 1-homogeneous case) that both the mean and variance are of the same order in  $h$ . On the

basis of this information about the single-step increment, it is not difficult to show that in the limit as  $h \downarrow 0$ , the piecewise affine interpolation  $\bar{X}^{h,\varepsilon}$  of  $X^{h,\varepsilon}$  (see (3.3)) converges in law to the solution  $Y$  of the Itô stochastic differential equation

$$dY_t = \frac{\ell(t)}{\eta} dt + \sqrt{\frac{\varepsilon}{\eta}} dB_t$$

with appropriate initial conditions, where  $B$  denotes a 1-dimensional Wiener process (Brownian motion). In particular, in the case  $\ell(t) \equiv 0$ ,  $Y$  is a Brownian motion. As intuition suggests, the variance of this trajectory decreases to zero as  $\varepsilon$  (“temperature”) tends to zero and/or  $\eta$  (viscosity) tends to  $+\infty$ . Note that, given information only on the initial position  $Y_0$ , the expected path satisfies

$$\mathbb{E}[Y_t | Y_0 = y_0] = y_0 + \frac{1}{\eta} \int_0^t \ell(s) ds. \quad (2.20)$$

In other words, the expected path satisfies the equilibrium equation (2.18). Furthermore, by the Itô isometry, the variance about this mean path is given by

$$\text{Var}[Y_t | Y_0 = y_0] = \frac{\varepsilon t}{\eta}.$$

This example can also be considered from the point of view of a path integral. Substituting the transition density (2.19) into the Chapman–Kolmogorov equation (2.15) yields

$$\begin{aligned} & \rho_n^{h,\varepsilon}(x_n) \\ &= \int \cdots \int_{\mathbb{R}} \left( \prod_{j=0}^{n-1} \frac{\exp\left(-\frac{\eta}{2h\varepsilon} \left(x_{j+1} - x_j - \frac{h}{\eta} \ell(t_{j+1})\right)^2\right)}{\sqrt{2\pi h\varepsilon/\eta}} \right) \rho_0^{h,\varepsilon}(x_0) dx_0 \cdots dx_{n-1} \\ &= \left(\frac{2\pi h\varepsilon}{\eta}\right)^{-n/2} \int \cdots \int_{\mathbb{R}} \exp\left(-\frac{1}{\varepsilon} \sum_{j=0}^{n-1} \frac{\eta}{2h} \left(x_{j+1} - x_j - \frac{h}{\eta} \ell(t_{j+1})\right)^2\right) \rho_0^{h,\varepsilon}(x_0) dx_0 \cdots dx_{n-1}. \end{aligned}$$

Formally, for  $t > 0$  and  $n = \lfloor t/h \rfloor$ , the sum in exponential above converges as  $h \downarrow 0$  to the integral

$$\mathcal{I}[x] := \frac{1}{2\eta} \int_0^t |\eta \dot{x}(t') - \ell(t')|^2 dt'.$$

Thus, it is observed that the work functional  $\mathcal{I}$  integrates in time the square of the equilibrium equation (2.18), reduced to units of power by means of the factor  $1/\eta$ . The functional  $\mathcal{I}$  has as its Euler-Lagrange equation the second-order ordinary differential equation

$$\eta \ddot{x}(t) - \dot{\ell}(t) = 0, \quad (2.21)$$

which is the equilibrium equation (2.18) differentiated once with respect to time,  $t$ .

At this point, it is appropriate to note that  $\mathbb{E}[Y_t]$  (as given by (2.20)) also satisfies (2.21), i.e., it is a *critical path*. Evidently, this is the path that minimizes the work functional  $\mathcal{I}$ . Indeed, in the limit of infinite viscosity or zero temperature, the *only* path which contributes to the path integral is the critical path, which in that limit coincides with the deterministic trajectory, i.e., with the solution of (2.18). For finite viscosity and finite temperature, however, all paths contribute to the path integral to varying degrees, the contributions becoming increasingly weaker as the trajectories depart from the critical path.

On the other hand, it is not in general true that a solution of the second-order equation (2.21) is also a solution of the first-order equation (2.18), unless compatible position *and* velocity data are supplied.

### 3 Identification of the limiting path measure

In this section, we concentrate on the rate-independent case where  $\Psi$  is 1-homogeneous and analyze the limiting behavior of the path-measure as the step-size  $h$  tends to 0. First, some remarks on the notation:

- As is usual in the probabilistic literature, random variables/stochastic processes defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  will be denoted using upper-case letters  $X, Y, Z, \dots$  while specific possible values of such functions will be denoted using lower-case letters  $x, y, z, \dots$ . Lower-case letters will also be used to denote deterministic quantities.
- $\Delta$  will denote the backward difference operator:

$$\Delta X_n := X_n - X_{n-1} \text{ for } n \in \mathbb{N}, \quad (3.1)$$

with the telescopic reconstruction formula

$$X_n = X_0 + \sum_{i=1}^n \Delta X_i \text{ for } n \in \mathbb{N}. \quad (3.2)$$

- For a discrete-time stochastic process (or deterministic sequence)  $X^h: \mathbb{N}_0 \times \Omega \rightarrow \mathbb{R}^N$ ,  $h > 0$ ,  $\bar{X}^h: [0, +\infty) \times \Omega \rightarrow \mathbb{R}^N$  will denote the continuous-time piecewise affine interpolation of  $X^h$ :

$$\bar{X}_t^h := X_{[t/h]}^h + \frac{t - h[t/h]}{h} (X_{[t/h]+1}^h - X_{[t/h]}^h) \text{ for } t \geq 0. \quad (3.3)$$

For some remarks on the convergence of random variables, see appendix 6.1 or [4, 5].

The heuristics of section 2 outline a procedure whereby the probability density function  $\rho_n^{h,\varepsilon}$  of the system at time  $t_n = nh$  is transported to a new probability density function  $\rho_{n+1}^{h,\varepsilon}$  of the system at time  $t_{n+1}$  given by a minimization problem with interior-point regularization.

This procedure can be made rigorous using Fubini's theorem and Lebesgue's dominated convergence theorem, yielding a well-defined discrete-time Markov chain

$$X = X^{h,\varepsilon}: \mathbb{N}_0 \times \Omega \rightarrow \mathbb{R}^N$$

with transition kernels  $\mathbb{P}[X_{n+1}|X_n = x_n]$  having densities  $\rho_{n+1}^{h,\varepsilon}(\cdot|x_n): \mathbb{R}^N \rightarrow \mathbb{R}$  given by

$$\rho_{n+1}^{h,\varepsilon}(x_{n+1}|x_n) = \frac{e^{-W(x_n, x_{n+1})/\varepsilon}}{Z(x_n)} \equiv \frac{e^{-W(x_n, x_{n+1})/\varepsilon}}{\int_{\mathbb{R}^N} e^{-W(x_n, y)/\varepsilon} dy}.$$

By analogy with the Kantorovich optimal transport problem, such a Markov chain may be termed an *optimal transport chain*, but will simply be referred to as “the Markov chain” in the sequel.

This set-up is simply one example of a broader class of discrete time models. This case is the model for a first-order differential inclusion in contact with a heat bath, and the chain  $X: \mathbb{N}_0 \times \Omega \rightarrow \mathbb{R}^N$  is strictly Markovian; one could also consider, for example, the *doubly nonlinear problem* of [10], or systems with inertia, in which case the next state of the chain  $X$  would depend on the two previous states. Thus, second-order processes involving inertial effects may also be studied using this general approach, although these lie beyond the scope of this paper.

It may also be helpful to note that although the heuristics of section 2 suggest that  $\varepsilon$  plays the role of a temperature parameter (up to units and constants), it is better to think of  $\varepsilon$  simply as an abstract regularization parameter. In particular,  $\varepsilon$  may depend on the time step  $h$ .

Associated to the Markov chain there is both a probability measure at each discrete time step  $t_n$ , and also a probability measure  $(X^{h,\varepsilon})_*(\mathbb{P})$  on the space of all discrete-time Markov chains taking values in  $\mathbb{R}^N$ . Hence, there is an induced probability measure  $(\bar{X}^{h,\varepsilon})_*(\mathbb{P})$  on the space of all continuous functions from  $[0, +\infty)$  into  $\mathbb{R}^N$ . In this setting, the objective of the procedure in this paper is to extract a suitable limit of the law of  $\bar{X}^{h,\varepsilon}$  as  $h \downarrow 0$ .

In the previous example of viscous (2-homogeneous) dissipation, in which the limiting process was an Itô diffusion, both the mean and variance of a single time step increment were of the same order,  $h$ . In contrast to this, as shown in the following section, when  $\Psi$  is 1-homogeneous, the mean and variance are of different orders in  $\varepsilon$  and  $h$  does not appear *a priori*. Thus, it proves necessary to take a diagonal limit in which  $\varepsilon$  is proportional to  $h$ , say  $\varepsilon = \theta h$ . From the point of view of the analysis,  $\theta > 0$  is simply a parameter; however, it can be interpreted as a “temperature” and its effect on the process  $X$  is to scale time.

The main point to be noted about the case of 1-homogeneous  $\Psi$  is that, although the processes  $\bar{X}^{h,\varepsilon}$  each have non-singular law, the laws collapse to a Dirac measure on  $\mathcal{C}^0([0, T]; \mathbb{R}^N)$  in the limit as  $h \downarrow 0$ . Thus, in a nutshell, *the optimal transport procedure for 1-homogeneous dissipation determines a unique continuous-time system trajectory for each  $\theta$  and each (deterministic) initial condition  $x_0$ .*

To describe this evolution, two further objects are required. The first is a notion of *stability* that is already well-established in the literature — see, for example, [21, 22]. The

intuitive model for this notion of stability is that a state  $x$  is stable if the net load at  $x$  is weaker than the dissipation; such states would be fixed points of the evolution in the absence of the heat bath, and so it is not surprising that the presence of the heat bath produces an interesting change in the dynamics on the stable set. This paper concentrates on the case in which the energetic potential is a quadratic form and singles out two particular sub-cases: the *flat case* in which the elasticity matrix is identically zero and the stability criterion can be expressed purely in terms of the applied load; and the *parabolic case* in which the elasticity matrix is symmetric and positive-definite and the stability criterion is stated in a form closer to that of [21, 22].

The second object is an *effective dual dissipation potential*, a logarithmic potential on load space (i.e. the dual variables), defined in (3.12). This potential encodes information about both the Markov chain  $X^{h,\varepsilon}$  and the ordinary differential equation that  $\bar{X}^{h,\varepsilon}$  satisfies in the limit as  $\varepsilon = \theta h \rightarrow 0$ . Its convexity properties are key to the proofs of these limit theorems, in particular theorem 3.4.

This paper contains two restriction that are unnatural from an engineering standpoint: first, the system is finite-dimensional; second, the energy is a convex quadratic form. Extension of the framework introduced in this paper will be the topic of future work, cf. [29].

### 3.1 Stability and the effective dual dissipation potential

The assumption that the dissipation potential  $\Psi: \mathbb{R}^N \rightarrow \mathbb{R}$  is non-negative and 1-homogeneous is equivalent to the assumption that  $\Psi$  is the convex conjugate  $\chi_{\mathcal{E}}^*$  of the convex characteristic function  $\chi_{\mathcal{E}}$  of  $\mathcal{E}$ , i.e.

$$\Psi(v) = \sup\{w \cdot v \mid w \in \mathcal{E}\},$$

where the *elastic region*  $\mathcal{E} \subseteq (\mathbb{R}^N)^*$  is compact and convex; the non-degeneracy of  $\Psi$  is equivalent to requiring that 0 be an interior point of  $\mathcal{E}$ . Given  $\Psi$  (or  $\mathcal{E}$ ) and an energetic potential  $(t, x) \mapsto E(t, x)$ , the elastic region  $\mathcal{E}$  can be “pulled back” to give a set of *stable states*  $\mathcal{S}(t) \subseteq \mathbb{R}^N$  for each time  $t \in [0, T]$ :

$$\mathcal{S}(t) := \{x \in \mathbb{R}^N \mid -DE(t, x) \in \mathcal{E}\}.$$

For example, if  $E(t, x) = \frac{1}{2}x \cdot Ax - \ell(t) \cdot x$  and  $A \in \mathbb{R}^{N \times N}$  is a symmetric and non-negative matrix, then  $\mathcal{S}(t)$  is the convex set  $A^{-1}(\ell(t) + \mathcal{E})$ . In general,  $\mathcal{S}(t)$  consists of those points  $x \in \mathbb{R}^N$  that are stable in the sense that, in the absence of the heat bath, the dissipation would outweigh the net force imparted by the gradient of  $E$ , and may fail to be convex. When  $E(t, \cdot)$  is convex,

$$\mathcal{S}(t) = \{x \in \mathbb{R}^N \mid \forall y \in \mathbb{R}^N, E(t, x) \leq E(t, y) + \Psi(y - x)\},$$

i.e. those states  $x$  that are almost minimizers of  $E(t, \cdot)$  in the sense that more energy would be dissipated by moving to any other state  $y$  than would be released by the potential difference  $E(t, x) - E(t, y)$ . In [21], the former version of stability is referred to as *local stability* and the latter as *global stability*.

We now introduce an *effective dual dissipation potential*, a function on  $(\mathbb{R}^N)^*$  that is determined entirely by the dissipation potential  $\Psi$  (i.e. by the geometry of the elastic region  $\mathcal{E}$ ).

**Definition 3.1.** Given a compact, convex set  $\mathcal{E} \subseteq (\mathbb{R}^N)^*$  having 0 in its interior and  $\Psi := \chi_{\mathcal{E}}^*$ , define the associated *effective dual dissipation potential*  $F^0: (\mathbb{R}^N)^* \rightarrow [0, +\infty]$  by

$$F^0(w) := \log \int_{\mathbb{R}^N} \exp(- (w \cdot z + \Psi(z))) dz. \quad (3.4)$$

$F^0$  takes its non-trivial values on the interior  $\overset{\circ}{\mathcal{E}}$  of the elastic region  $\mathcal{E}$ , reflected in the origin (in many applications, the elastic region is symmetric, so  $\mathcal{E} = -\mathcal{E}$  and  $F^0$  is non-trivial on  $\overset{\circ}{\mathcal{E}}$ ). It follows quickly from the definition of  $\Psi$  as  $\chi_{\mathcal{E}}^*$  that

$$F^0(w) \in \begin{cases} (0, +\infty), & \text{if } w \in -\overset{\circ}{\mathcal{E}}; \\ \{+\infty\}, & \text{otherwise.} \end{cases}$$

Lebesgue's dominated convergence theorem implies that  $F^0$  is continuous and (in the form of differentiation under the integral sign) that  $F^0$  is infinitely differentiable on  $\overset{\circ}{\mathcal{E}}$ . The derivative  $DF^0: (\mathbb{R}^N)^* \rightarrow (\mathbb{R}^N)^{**} \cong \mathbb{R}^N$  is a monotone operator, i.e. for all  $v, w \in -\overset{\circ}{\mathcal{E}}$ ,

$$(DF^0(v) - DF^0(w)) \cdot (v - w) \geq 0;$$

this follows from a direct computation and application of Hölder's inequality and Kachurovskii's theorem [28, proposition 7.4].

In some cases of interest, the elastic region, dissipation and effective dual dissipation potential can be written in closed form.

- In the Euclidean case, the elastic region is the round ball

$$\mathcal{E} := \{w \in (\mathbb{R}^N)^* \mid |w|_2^2 := w_1^2 + \dots + w_n^2 \leq \sigma^2\}$$

for some  $\sigma > 0$ ; the associated dissipation potential is  $\Psi(v) = \sigma|v|_2$ . A direct calculation using spherical polar coordinates yields that, up to an additive constant,

$$F^0(w) = -\frac{N+1}{2} \log(\sigma^2 - |w|_2^2). \quad (3.5)$$

- Another case of interest, used in the sequel, is that in which the elastic region is the cuboid

$$\mathcal{E} := \{w \in (\mathbb{R}^N)^* \mid |w_i| \leq \sigma_i \text{ for } i = 1, \dots, N\},$$

where  $\sigma_i > 0$  for each  $i$ ; the associated dissipation potential is the weighted  $\ell^1$  "Manhattan" norm

$$\Psi(v) = \sigma_1|v_1| + \dots + \sigma_N|v_n|.$$

A direct calculation using iterated integrals yields that, up to an additive constant,

$$F^0(w) = -\sum_{i=1}^N \log(\sigma_i^2 - |w_i|^2). \quad (3.6)$$

In the sequel, we will need not just that the effective dual dissipation potential  $F^0$  be convex (which is always the case), but that the vector field on  $\mathcal{S}(t)$  given by

$$x \mapsto -DF^0(DE(t, x))$$

be a monotonically decreasing vector field for every  $t \in [0, T]$ . For example, this is the case if  $E(t, x) = \frac{1}{2}x \cdot Ax - \ell(t) \cdot x$  and  $A \in \mathbb{R}^{N \times N}$  is symmetric and non-negative.

### 3.2 Flat energetic landscape

As an introductory case in which the analysis is significantly simpler, consider the so-called *flat case* of an energetic potential with zero Hessian. In this case, the energetic potential  $E: [0, +\infty) \times \mathbb{R}^N \rightarrow \mathbb{R}$  is given by

$$E(t, x) = -\ell(t) \cdot x; \tag{3.7}$$

for a Lipschitz-continuous external loading  $\ell: [0, +\infty) \rightarrow (\mathbb{R}^N)^*$ . The dissipative potential  $\Psi: \mathbb{R}^N \rightarrow \mathbb{R}$  is assumed to be positive-definite and positively homogeneous of degree one. Define the Markov chain  $X$  associated to the approximate cost function  $C$  as above.

The analysis in this case is simplified greatly by the fact that the approximate dual dissipation potential (3.12) that will be needed in the non-degenerate case equals  $F^0$  for every  $\varepsilon > 0$ . Furthermore, as a function of  $t$  and  $x$ , the effective dual dissipation potential  $F^0(DE(t, x))$  is independent of  $x$ :

$$F^0(DE(t, x)) = \log \int_{\mathbb{R}^N} \exp(-(\Psi(z) - \ell(t) \cdot z)) \, dz.$$

As noted in the previous subsection, this makes sense only if  $\ell$  satisfies a suitable stability criterion with respect to  $\Psi$ , i.e. that  $\ell$  does not get arbitrarily close to the yield surface  $\partial\mathcal{E}$ .

To save space, denote by  $f^0: [0, T] \rightarrow \mathbb{R}^N$  the vector field given by the negative of the derivative of  $F^0$  evaluated at  $DE(t, x)$ , i.e.

$$f^0(t) := -DF^0(DE(t, x)) = \frac{\int_{\mathbb{R}^N} z \exp(-(\Psi(z) - \ell(t) \cdot z)) \, dz}{\int_{\mathbb{R}^N} \exp(-(\Psi(z) - \ell(t) \cdot z)) \, dz}.$$

Based on lemma 6.2, the heuristic argument for this case is that the discrete-time Markov chain  $X^{h,\varepsilon}$  has independent increments with mean  $\varepsilon f^0(t_{i+1})$  and variance at most  $C\varepsilon^2$ , where the constant depends on  $\ell$ ,  $\Psi$ ,  $N$ , &  $c$ . Thus, a naïve application of the central limit theorem yields that “ $\bar{X}_t - \bar{X}_0$ ” is approximately normally distributed with mean  $\sum_{i=0}^{\lfloor t/h \rfloor} \varepsilon f^0(t_{i+1})$  and variance at most  $C\lfloor t/h \rfloor \varepsilon^2$ ; in the limit as  $\varepsilon = \theta h \downarrow 0$ , this gives the deterministic limit

$$“\bar{X}_t - \bar{X}_0” = \theta \int_0^t f^0(s) \, ds,$$

i.e. “ $\dot{\bar{X}}_t$ ” =  $\theta f^0(\ell(t))$ . This heuristic argument is made rigorous by the following theorem:

**Theorem 3.2** (Flat case). *Let  $\Psi$  be continuous, 1-homogeneous and coercive, i.e. there exists  $c > 0$  such that  $\Psi(z) \geq c|z|_2$  for all  $z \in \mathbb{R}^N$ . Suppose that  $\ell: [0, T] \rightarrow (\mathbb{R}^N)^*$  is uniformly Lipschitz and satisfies the uniform stability criterion*

$$\inf_{t \in [0, T]} \inf_{z \in \mathbb{R}^N \setminus \{0\}} (\Psi(z) - \ell(t) \cdot z) > 0,$$

*i.e. that  $\inf_{t \in [0, T]} \text{dist}(\ell(t), \partial \mathcal{E}) > 0$ . Then, as  $h \downarrow 0$ , the process  $\bar{X}^{h, \theta h}$  converges in probability, uniformly on the compact interval  $[0, T]$ , to the differentiable, deterministic process  $y = y^\theta$  satisfying*

$$\dot{y}_t = \theta f^0(t) \quad (= -\theta DF^0(DE(t, y_t))), \quad (3.8)$$

*with the same initial condition  $X_0^{h, \theta h} = y_0^\theta = x_0$ . More precisely, for any  $T > 0$ ,  $\lambda > 0$ ,*

$$\mathbb{P} \left[ \sup_{0 \leq t \leq T} |\bar{X}_t^{h, \theta h} - y_t^\theta|_2 \geq \lambda \right] \in O(h) \text{ as } h \downarrow 0. \quad (3.9)$$

*A fortiori, the limiting probability measure on path space  $\mathcal{C}^0([0, T]; \mathbb{R}^N)$  is a Dirac measure supported on  $y^\theta$ .*

Since the Markov chain  $X^{h, \varepsilon}$  was defined using a minimization principle, it is natural to ask whether  $y^\theta = \mathbb{P}\text{-}\lim_{h \downarrow 0} \bar{X}^{h, \theta h}$  also satisfies a minimization principle. The answer is affirmative: via the Brézis–Ekeland variational principle [6, 7], theorem 3.2 has the following immediate corollary:

**Corollary 3.3** (Variational principle). *The process  $\bar{X}^{h, \theta h}$  converges in probability, uniformly on compact intervals of time  $[0, T]$ , to the unique minimizer of the action functional  $\mathcal{S}: \{u \in \text{AC}([0, T]; \mathbb{R}^N) \mid u(0) = x_0\} \rightarrow [0, +\infty]$  given by*

$$\mathcal{S}[u] := \int_0^T \frac{1}{2} |v - \theta f^0(t)|_2^2 dt; \quad (3.10)$$

Note that the action  $\mathcal{S}$  in (3.10) is inequivalent to the path integral exponent  $\mathcal{I}$  in (2.16), which in this case is the integral

$$\mathcal{I}[u] = \int_0^t (\Psi(\dot{u}(s)) - \ell(s) \cdot \dot{u}(s)) ds.$$

Since, by assumption,  $z \mapsto \Psi(z) - \ell(t) \cdot z$  is positive-definite,  $\mathcal{I}[u] \geq 0$  and

$$\mathcal{I}[u] = 0 \iff \dot{u}(t) \cdot \ell(t) = \Psi(\dot{u}(t)) \text{ for almost all } t.$$

This condition is not generally satisfied by minimizers of the action (3.10); e.g., in dimension  $N = 1$  with  $\Psi(z) = \sigma|z|_2$  and  $\theta = 1$ , the minimizer of the action (3.10) satisfies  $\dot{u}(t) = f^0(t)$  but

$$f^0(t) \cdot \ell(t) = \frac{2|\ell(t)|^2}{\sigma^2 - |\ell(t)|^2} \neq \frac{2\sigma|\ell(t)|}{\sigma^2 - |\ell(t)|^2} = \Psi(f^0(t)).$$

Note also that the deterministic limiting system  $y^\theta = \mathbb{P}\text{-}\lim_{h \downarrow 0} \bar{X}^{h, \theta h}$  responds almost linearly to the applied load  $\ell$  when  $\ell$  is small. However, it exhibits a strongly non-linear yielding response to applied loads close in magnitude to the critical load for  $\Psi$  (when  $\Psi(\dot{x}) = \sigma|\dot{x}|_2$ , this is the quantity  $\sigma > 0$ ). The effect of the “temperature-like” parameter  $\theta$  is to soften the system response in proportion to  $\theta$  as  $\theta$  increases. See figure 3.1 for an illustration of the response of  $y$  to a cyclic applied load  $\ell$  in the one-dimensional case.

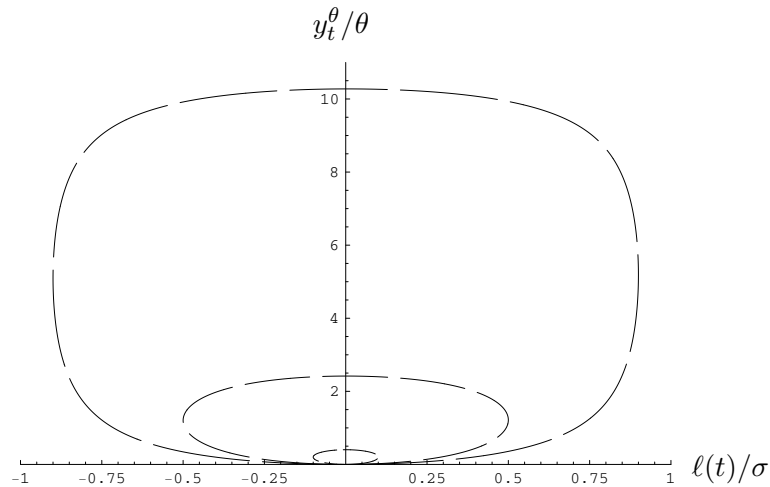


Figure 3.1: The nonlinear yielding response of the effective dynamics  $y^\theta$  (given by (3.8)) to an applied cyclic load  $\ell(t) = \ell_0 \sin t$  for  $\ell_0 = 0.1\sigma$  (short dashes),  $0.5\sigma$  (medium-length dashes) and  $0.9\sigma$  (long dashes).

An application of this analysis to Andrade creep under linear strain hardening will be given in greater depth in section 4.

### 3.3 Parabolic energetic landscape

In this section, the energetic potential  $E$  is a strictly convex parabola given by

$$E(t, x) = \frac{1}{2}x \cdot Ax - \ell(t) \cdot x; \tag{3.11}$$

for a uniformly Lipschitz external loading  $\ell: [0, +\infty) \rightarrow (\mathbb{R}^N)^*$  and a symmetric, positive-definite matrix  $A \in \mathbb{R}^{N \times N}$ . The dissipative potential  $\Psi: \mathbb{R}^N \rightarrow \mathbb{R}$  is again assumed to be positive-definite and positively homogeneous of degree one.

An heuristic model for this case is that the limiting process  $\mathbb{P}\text{-}\lim_{h \downarrow 0} \bar{X}_t^{h, \theta h}$  represents the position at time  $t$  of a light (inertialess) slider restrained by an elastic spring with spring constant/elasticity matrix  $A$  and subject to the time-dependent loading  $\ell$  as well as the action of the heat bath (which might be thought of as “shaking the table” on which the slider moves).

Once again, the result is convergence in probability to a deterministic path on compact intervals of time  $[0, T]$ . However, the characterization of the domain of the limiting evolution

requires a more sensitive analysis since the effect of the elasticity matrix  $A$  is to partition the state space  $\mathbb{R}^N$  into two regions: the bounded, convex stable set  $\mathcal{S}$  and its unbounded complement. The means, variances & c. of the increments of the Markov chain blow up on the boundary of the stable set; therefore, much of the mathematical work lies in keeping the process away from this boundary.

The fundamental object in the study of the limiting dynamics is again the scalar potential  $F^0: (\mathbb{R}^N)^* \rightarrow \mathbb{R}$  given by (3.4). In this case, however,  $F^0$  itself does not directly control the behaviour of the increments of the Markov chain; instead,  $F^0$  arises as the limit  $F^0 = \lim_{\varepsilon \rightarrow 0} F^\varepsilon$ , where the *approximate dual dissipation potential*  $F^\varepsilon: (\mathbb{R}^N)^* \rightarrow \mathbb{R}$  is defined by

$$F^\varepsilon(w) := \log \int_{\mathbb{R}^N} \exp \left( - (w \cdot z + \Psi(z) + \frac{\varepsilon}{2} z \cdot Az) \right) dz. \quad (3.12)$$

The potentials  $F^\varepsilon$  and  $F^0$  and their derivatives characterize the time-discrete and the limiting time-continuous evolution respectively.  $F^\varepsilon$  is a “nicer” object than  $F^0$  since the “ $+\frac{\varepsilon}{2} z \cdot Az$ ” in the exponent enlarges its essential domain beyond  $-\mathring{\mathcal{E}}$ ; in the case in which  $A$  is positive-definite,  $F^\varepsilon$  takes finite values on all of  $\mathbb{R}^N$ . Like  $F^0$ ,  $F^\varepsilon$  is a convex function, and so its derivative is a monotonically increasing operator, i.e. for all  $v, w \in (\mathbb{R}^N)^*$ ,

$$(DF^\varepsilon(v) - DF^\varepsilon(w)) \cdot (v - w) \geq 0.$$

As before, to save space, denote by  $f^0: [0, T] \times \mathbb{R}^N \rightarrow \mathbb{R}^N$  the time-dependent vector field given by evaluating the negative of the derivative of the effective dual dissipation potential  $F^0$  at the potential gradient  $DE(t, x)$ :

$$f^\varepsilon(t, x) := -DF^\varepsilon(DE(t, x)). \quad (3.13)$$

One key property of  $f^\varepsilon$  is that when  $DE(t, x) = Ax - \ell(t)$  with  $A \in \mathbb{R}^{N \times N}$  symmetric and positive-definite (as in the case (3.11)),  $f^\varepsilon(t, \cdot)$  is a monotonically decreasing vector field, i.e. for all  $x, y$  in the essential domain of  $F^\varepsilon$ ,

$$(f^\varepsilon(t, x) - f^\varepsilon(t, y)) \cdot (x - y) \leq 0.$$

The most important property of the energy  $F^\varepsilon$  is that, by lemma 6.2,

$$\mathbb{E}[X_{i+1}^{h,\varepsilon} | X_i^{h,\varepsilon} = x_i \in \mathcal{S}(t_{i+1})] = x_i + \varepsilon f^\varepsilon(t_{i+1}, x_i),$$

and, moreover, the  $k^{\text{th}}$  conditional central moment of  $X_{i+1}^{h,\varepsilon}$  is in  $O(\varepsilon^k)$ . In view of these properties and by analogy with theorem 3.2, the natural conjecture is that, as  $\varepsilon = \theta h$  tends to zero,  $\bar{X}^{h,\theta h}$  should converge to a deterministic process satisfying the deterministic gradient flow in  $F^0$ ; this is the content of theorem 3.4.

**Theorem 3.4** (Parabolic case). *Let  $\Psi$  be continuous, 1-homogeneous and coercive. For each  $\theta \geq 0$ ,  $T > 0$ , as  $h \downarrow 0$ ,  $\bar{X}^{h,\theta h}$  converges in probability in  $\mathcal{C}^0([0, T]; \mathbb{R}^N)$  to the differentiable, deterministic process  $y = y^\theta$  satisfying*

$$\dot{y}_t = \theta f^0(t, y_t) \quad (= -\theta DF^0(DE(t, y_t))), \quad (3.14)$$

with the same initial datum  $X_0^{h,\theta h} = y_0^\theta = x_0$  in the interior of  $\mathcal{S}(0)$ . More precisely, for any  $T > 0$ ,  $\lambda > 0$ ,

$$\mathbb{P} \left[ \sup_{0 \leq t \leq T} |\bar{X}_t^{h,\theta h} - y_t^\theta|_2 \geq \lambda \right] \in O(\sqrt{h}) \text{ as } h \downarrow 0. \quad (3.15)$$

A fortiori, the limiting measure on path space  $\mathcal{C}^0([0, T]; \mathbb{R}^N)$  is a Dirac measure supported on  $y^\theta$ .

Note that, for  $A = 0$ , one obtains theorem 3.2. As before, the corresponding variational characterization of the deterministic limiting trajectory is an easy corollary:

**Corollary 3.5** (Variational principle). *The process  $\bar{X}^{h,\theta h}$  converges in probability, uniformly on compact intervals of time  $[0, T]$ , to the unique minimizer of the action functional  $\mathcal{S}: \{u \in \text{AC}([0, T]; \mathbb{R}^N) \mid u(0) = x_0\} \rightarrow [0, +\infty]$  given by*

$$\mathcal{S}[u] := \int_0^T \mathcal{L}(t, u(t), \dot{u}(t)) dt; \quad (3.16)$$

### 3.4 Degenerate dissipation potentials

To conclude this section, we note that in many applications, the dissipation potential  $\Psi$  may not be coercive, in which case  $F^0$  may be infinite everywhere. However, formula (3.4) offers a very simple method of dimension reduction if  $\Psi$  is a weighted  $\ell^1$ -norm and the matrix  $A$  is such that  $F^\varepsilon$  is finite somewhere.

Indeed, assume for the sake of simplicity that  $x = (\xi, \eta) \in \mathbb{R}^N = \mathbb{R}^n \times \mathbb{R}^m$  and

$$\Psi(x) := \sum_{i=1}^n \sigma_i |x_i|$$

with  $\sigma_i > 0$  for all  $i \in \{1, \dots, n\}$ . Let  $E$  be quadratic as before. Writing the affine operator  $\mathcal{A} \equiv DE$  in block form as

$$\mathcal{A}x = \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix} \quad (3.17)$$

yields that if the approximate dual dissipation potential  $F^\varepsilon$  is finite, then  $\eta = C^{-1}(b - B^\top \xi)$ . This corresponds to instantaneous relaxation to the minimum of  $E$  along the coordinate directions that do not experience friction. Now let

$$\rho^\varepsilon(z_1) := \int_{\mathbb{R}^{N-n}} \exp \left( -\frac{\varepsilon}{2} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \cdot \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} - \Psi(z_1, 0) - (A\xi + B\eta - a) \cdot z_1 \right) dz_2,$$

so that

$$\partial_a F^\varepsilon(\mathcal{A}x) = \frac{\int_{\mathbb{R}^n} z_1 \rho^\varepsilon(z_1) dz_1}{\int_{\mathbb{R}^n} \rho^\varepsilon(z_1) dz_1}. \quad (3.18)$$

This yields the following explicit formula:

$$\rho^\varepsilon(z_1) = \frac{1}{\varepsilon^{m/2} \sqrt{\det(C)}} \exp\left(-\left((A - B^\top C^{-1} B)\xi - (a - BC^{-1}b)\right) \cdot z_1 - \Psi(z_1, 0) + O(\varepsilon)\right).$$

Unsurprisingly, the matrix  $\tilde{A} = A - BC^{-1}B^\top$  and the vector  $\tilde{a} = a - BC^{-1}b$  form the Schur complement of the affine operator  $\mathcal{A}$  in (3.17); in other words, the  $\xi$ -components experience the reduced energy

$$E_{\text{red}}(t, x) := \frac{1}{2}\xi \cdot \tilde{A}\xi - \tilde{a}(t) \cdot \xi.$$

Since the term  $\varepsilon^{m/2} \sqrt{\det(C)}$  does not depend on  $z_1$  it cancels in formula (3.18). This shows that the  $\xi$ -projection of the discrete Markov chain in the situation where  $\Psi$  is a weighted  $\ell^1$ -norm with degenerate weights converges to the solution of the reduced ordinary differential equation that is generated by the reduced dual dissipation potential (3.6)

$$F_{\text{red}}^0(w) := -\sum_{i=1}^n \log(\sigma_i^2 - (w \cdot e_i)^2)$$

applied to the energy gradient  $w = DE_{\text{red}}(t, \xi) = \tilde{A}\xi - \tilde{a}(t)$  as derived above.

This result may be summarized as follows:

**Corollary 3.6** (Degenerate friction potential  $\Psi$ ). *Let  $n \leq N$  be such that  $\Psi(z) = \sum_{i=1}^n \sigma_i |z_i|$  and  $\sigma_i > 0$  for all  $i \in \{1 \dots n\}$ . If the matrix  $C \in \mathbb{R}^{(N-n) \times (N-n)}$  defined in the induced decomposition (3.17) is positive definite, then, for each  $\theta \geq 0$ ,  $T > 0$ , as  $h \downarrow 0$ ,  $\bar{\xi}^{h, \theta h}$  converges in probability in  $C^0([0, T]; \mathbb{R}^n)$  to the differentiable, deterministic process  $y = y^\theta$  satisfying*

$$\dot{y}_t = \theta DF_{\text{red}}^0(DE_{\text{red}}(t, y_t)), \quad (3.19)$$

with the same initial stable datum  $\xi_0^{h, \theta h} = y_0^\theta = \xi_0$ .

This result is based on the fact that the zero-set of  $\Psi$  is a linear space. Note, however, that for general 1-homogeneous functions the zero set has in general a more complicated structure.

## 4 Application to crystallographic slip and creep

In this section, the preceding formalism and analysis are applied to *Andrade creep*, also known as  $\beta$ -creep. In 1910, Andrade [2, 3] reported that as a function of time,  $t$ , the creep deformation,  $y$ , of soft metals at constant temperature and applied stress can be described by a power law  $y(t) \sim t^{1/3}$ . Similar behavior has been observed in many classes of materials, including non-crystalline materials. Recently, it has been suggested that Andrade creep is an example of critical behavior [12, 13, 14]. In these theories, creep is considered as an instance of critical avalanche formation analogous to sandhill models introduced to explain

self-organized criticality. However, creep differs from sandhill behavior in two notable respects: the avalanches are nucleated by thermal activation; and the system exhibits strain hardening. Several numerical simulations are able to predict the power law behavior of the plastic deformation at constant temperature and applied stress in agreement with experimental observation [9, 14, 23]. Those models also predict a transition to steady state creep at later times, which results from recovery.

The analysis will proceed in four steps: the first is a presentation of an infinite-dimensional phase-field model for dislocation dynamics on a single slip plane (thought of as the 2-torus,  $\mathbb{T}^2$ ). This model has been proposed by Koslowski, Cuitiño and Ortiz in [19] and a mathematically rigorous analysis of the line-tension limit can be found in [15, 16]. Secondly, this infinite-dimensional evolution will be reduced to a finite-dimensional one that takes account of the values of the phase field at certain “obstacle sites”. Thirdly, a mean-field approximation will be derived to reduce this large finite-dimensional evolution to a one-dimensional model of the energetics-and-dissipation type analyzed earlier. It will be shown that Andrade’s  $t^{1/3}$  creep law follows from the reduced model when linear hardening is assumed. Finally, numerical simulations of phase-field model for dislocation dynamics are carried out. These simulations show excellent agreement with the reduced model.

The mean field approximation provides a convenient framework to illustrate the influence of temperature, obstacle density and strain rate during cyclic loading in the phase field model of dislocations.

## 4.1 Derivation of the mean-field model

So far, we have regarded the energetic potential as a function of time and state; we now change conventions and regard  $E$  as a function of an externally applied load and the state. The simplest setting is obtained if the reference configuration of the material is given by the unit torus  $\mathbb{T}^2 = [0, 1]^2$  and the potential energy of a phase-field  $u$ , which represents the amount of slip in units of the Burgers vector, is given by

$$E(s, u) := \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} \frac{\mu b^2}{4} \frac{1}{1/K + d/2} |\hat{u}(k)|^2 - b \cdot s \int_{\mathbb{T}^2} u(x) dx, \quad (4.1)$$

where  $k$  is the wavenumber,

$$K := \frac{1}{1 - \nu} |k|_2 - \frac{\nu}{1 - \nu} \frac{k_2^2}{|k|_2}, \quad (4.2)$$

$\hat{u}(k) := \int_{\mathbb{T}^2} e^{2\pi i x \cdot k} u(x) dx$  is the  $k^{\text{th}}$  Fourier component of  $u$ ,  $b$  is the Burgers vector,  $s$  is the applied shear stress,  $\mu$  is the shear modulus,  $\nu$  is Poisson’s ratio and  $d$  is the interplanar distance. The first term in the energy  $E$  is invariant under translations, i.e.

$$E(s, u + u') = E(s, u) + s \cdot b u' \quad (4.3)$$

if  $u'$  is a constant field. The obstacles are assumed to be disk-like subsets  $\mathbb{B}_r(p_i) \subsetneq \mathbb{T}^2$  where  $p_1, \dots, p_n \in \mathbb{T}^2$  are the centers of the obstacles and  $\mathbb{B}_r(p_i)$  is the disk with radius  $0 < r \ll 1$

and center at  $p_i$ . Typically it will be assumed that the positions  $p_i$  are random and that the phase field  $u$  is constant within each obstacle, i.e.  $u(x) = \xi_i$  if  $x \in \mathbb{B}_r(p_i)$ . The dissipation functional  $\Psi$  depends on the obstacles and is defined as follows:

$$\Psi(\dot{u}) = \sum_{p \in \mathcal{O}} \kappa \int_{\mathbb{B}_r(p)} |\dot{u}(x)| \, dx. \quad (4.4)$$

Note that in this formulation the friction functional  $\Psi$  vanishes on an infinite dimensional set, hence the assumptions of theorem 3.4 are not satisfied. However, since  $\Psi$  is a weighted  $L^1$ -norm, corollary 3.6 implies formally that the projection of the infinite dimensional Markov chain onto the values of the elastic field on the obstacles converges to the solution of a reduced differential equation.

This differential equation is determined by constructing the appropriate Schur complement of the elastic energy, i.e. by minimizing out the the degrees of freedom that do not experience friction. For each  $\xi \in \mathbb{R}^{\mathcal{O}}$ , define the reduced elastic energy by

$$E_{\text{red}}(s, \xi) := \inf \{ E(s, u) \mid u \in H^{1/2}(\mathbb{T}^2) \text{ and } u(x) = \xi_i \text{ if } x \in \mathbb{B}_r(p_i) \} \quad (4.5)$$

and the reduced dissipative potential by

$$\Psi_{\text{red}}(\dot{\xi}) := \kappa |\mathbb{B}_r| \sum_{p \in \mathcal{O}} |\dot{\xi}_p|. \quad (4.6)$$

Note that  $E_{\text{red}}$  is a quadratic form, i.e., there exists a matrix  $G \in \mathbb{R}^{\mathcal{O} \times \mathcal{O}}$ , a vector  $\tau \in \mathbb{R}^{\mathcal{O}}$  and a scalar  $h \in \mathbb{R}$  such that  $E_{\text{red}}(s, \xi) = \frac{1}{2} \xi \cdot G \xi - (s \cdot b) \tau \cdot \xi + \frac{h}{2} (s \cdot b)^2$ . The coefficients of  $G$  and  $\tau$  and the value of  $h$  are random and depend on the realization of the obstacle positions. The resulting dual dissipation potential is

$$F(DE_{\text{red}}(\tau, \xi)) = - \sum_{p \in \mathcal{O}} \log (\kappa^2 |\mathbb{B}_r|^2 - ((G\xi - (s \cdot b)\tau) \cdot e_p)^2), \quad (4.7)$$

where  $e_p \in \mathbb{R}^{\mathcal{O}}$  denotes the unit dislocation over the obstacle site  $p \in \mathcal{O}$ . Since the number of obstacles is typically large the differential equation (3.19) is difficult to analyze. The situation simplifies when one considers the asymptotic behavior of the solutions in the limit where the number of obstacles  $|\mathcal{O}|$  tends to infinity and  $r$  tends to zero such that  $\kappa |\mathcal{O}| |\mathbb{B}_r| = \sigma$ .

Observe first that, for each realization of the obstacle positions, the translation invariance (4.3) implies that, for the constant vector  $\underline{1} = \sum_{p \in \mathcal{O}} e_p \in \mathbb{R}^{\mathcal{O}}$ ,

$$G \underline{1} = 0 \text{ and } \tau \cdot \underline{1} = \bar{\tau} = E(\frac{1}{|b|^2} b, \underline{0}) - E(\frac{1}{|b|^2} b, \underline{1}). \quad (4.8)$$

It can be shown that  $\bar{\tau}$  converges to a deterministic value when the number of obstacles tends to infinity. In view of (4.8), splitting  $\xi$  into its mean value and the fluctuation around the mean value,

$$\bar{\xi} := \frac{1}{|\mathcal{O}|} \xi \cdot \underline{1}, \quad \hat{\xi} := \xi - \bar{\xi} \underline{1}, \quad (4.9)$$

yields the expressions

$$\begin{aligned} E_{\text{red}}(s, \xi) &= \frac{1}{2} \hat{\xi} \cdot G \hat{\xi} - (s \cdot b) \bar{\tau} \bar{\xi} - (s \cdot b) \hat{\tau} \cdot \hat{\xi} + \frac{1}{2} (b \cdot s)^2, \\ \Psi_{\text{red}}(z) &= \Psi_{\text{red}}(\hat{z} + \bar{z} \underline{1}), \end{aligned}$$

where  $\hat{\tau} := \tau - \frac{\bar{\tau}}{|\mathcal{O}|} \underline{1}$ ,  $\bar{z} := \frac{1}{|\mathcal{O}|} z \cdot \underline{1}$ , and  $\hat{z} := z - \bar{z} \underline{1}$ . Since  $\hat{z} \cdot \underline{1} = 0$ , the dissipation function  $\Psi_{\text{red}}$  satisfies the inequality

$$\Psi_{\text{red}}(\bar{z} \underline{1}) \leq \Psi_{\text{red}}(\bar{z} \underline{1} + \hat{z}) \leq \Psi_{\text{red}}(\bar{z} \underline{1}) + \Psi_{\text{red}}(\hat{z}). \quad (4.10)$$

Corollary 3.6 implies that  $\xi$  solves the differential equation

$$\frac{d\xi}{dt} = -DF_{\text{red}}^0(DE_{\text{red}}(\tau, \xi)). \quad (4.11)$$

As the number of obstacles  $|\mathcal{O}|$  is very large, it can be shown that  $\Psi_{\text{red}}(\hat{z})$  becomes small, i.e. the fluctuations of  $\xi$  around the mean value  $\bar{\xi}$  experience very little friction. A rigorous proof of this convergence is possible but beyond the scope of this paper. If  $\Psi_{\text{red}}(\bar{z} \underline{1} + \hat{z})$  is replaced by

$$\Psi_{\text{MF}}(\bar{z}) := \kappa |\mathbb{B}_r| |\mathcal{O}| |\bar{z}| = \sigma |\bar{z}|, \quad (4.12)$$

and  $E_{\text{red}}$  is replaced by the mean-field energy

$$E_{\text{MF}}(\bar{\tau}, \bar{\xi}) := -(s \cdot b) \bar{\tau} \bar{\xi}, \quad (4.13)$$

we obtain the corresponding effective dual dissipation potential that controls the evolution of the mean field  $\bar{\xi}$ :

$$F_{\text{MF}}^0(DE_{\text{MF}}(\bar{\tau}, \bar{\xi})) = -\log(\sigma^2 - (s \cdot b)^2 |\bar{\tau}|^2). \quad (4.14)$$

Note that  $F_{\text{MF}}^0$  is always a lower bound for  $F_{\text{red}}^0$ , even if the number of obstacles is finite.

## 4.2 Derivation of the Andrade creep law

These reduction steps yield a scalar differential equation

$$\frac{d\bar{\xi}}{dt} = \theta DF_{\text{MF}}^0(DE_{\text{MF}}(\bar{\tau}, \bar{\xi})) = \frac{2\theta\bar{\tau}}{\sigma^2 - \bar{\tau}^2} \quad (4.15)$$

with appropriate initial conditions; in order to simplify the notation it is assumed that  $s \cdot b = 1$ . Assume a constant applied load  $\bar{\tau}$  and that the resistance is  $\sigma$ , with  $0 \leq |\bar{\tau}| < \sigma$  — once again, this is the stability criterion that the applied load should remain in the interior of the elastic region. At least formally,  $\sigma$  may vary provided that the inequality  $\sigma > |\bar{\tau}| \geq 0$  still holds, and under the additional assumption of linear strain hardening, i.e.  $\sigma = \sigma_0 \bar{\xi}$ , the previous equation becomes

$$\frac{d\bar{\xi}}{dt} = \frac{2\theta\bar{\tau}}{\sigma_0^2 \bar{\xi}(t)^2 - \bar{\tau}^2}. \quad (4.16)$$

Clearly, the behaviour of  $\bar{\xi}(t)$  for small  $t$  will depend upon the magnitude of  $\bar{\tau}$  as compared to  $\sigma_0$ . However, under the assumption that the effective applied stress  $\hat{\tau}$  is small in comparison to the frictional resistance  $\sigma = \sigma_0 \bar{\xi}$ ,

$$\frac{d\bar{\xi}}{dt} = \frac{2\theta\bar{\tau}}{\sigma_0^2 \bar{\xi}(t)^2}. \quad (4.17)$$

It is noteworthy that, by a different approach, we arrive at the same differential equation as that proposed by Cottrell in [12, 13]. Integration of this equation yields

$$\bar{\xi}(t) \approx \left( C_1 + \frac{6\theta\bar{\tau}t}{\sigma_0^2} \right)^{1/3} \sim t^{1/3} \text{ as } t \rightarrow +\infty, \quad (4.18)$$

where  $C_1$  is a constant of integration, by the general solution

$$\frac{dy}{dt} = \frac{k\theta}{(\sigma_0\theta^\alpha y(t)^\beta)^2} \implies y(t) = \left( (1 + 2\beta) \left( C_1 + \frac{k\theta^{1-2\alpha}t}{\sigma_0^2} \right) \right)^{\frac{1}{1+2\beta}}. \quad (4.19)$$

Thus, as illustrated by the plots of numerical solutions to (4.16) in figure 4.1, Andrade's  $t^{1/3}$  creep law follows as a straightforward consequence of the limiting dynamics proved in section 3. Furthermore, the theory also predicts more rapid creep for larger values of the temperature-like parameter  $\theta$ , as intuition would suggest.

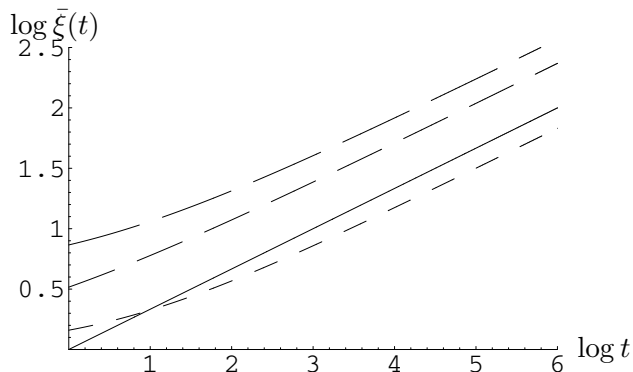


Figure 4.1: A log-log plot of  $\bar{\xi}(t)$  (dashed curves) against  $t$  for (4.16), illustrating Andrade's  $t^{1/3}$  Andrade creep law behaviour by comparison with  $t \mapsto t^{1/3}$  (solid curve). Parameters: initial condition  $\bar{\xi}(0) = 1$ ;  $\theta = 1$ ;  $\sigma_0 = 1$ ;  $\bar{\tau} = 0.1$  (short dashes), 0.5 (medium-length dashes), 0.9 (long dashes).

### 4.3 Numerical simulations

This subsection describes numerical simulations of the phase-field dislocation theory proposed by Koslowski, Cuitiño and Ortiz in [19]. Proceeding to the numerical solution of

the phase-field dislocation model, the slip plane is discretized in a square of 50 by 50 grid points with periodic boundary conditions and a random array of obstacles. The potential energy is computed in Fourier space following (4.1). The irreversible dislocation obstacle interactions are built into the framework developed in section 2 by the introduction of the incremental work function (2.2). The simulations are performed with a Metropolis Monte Carlo algorithm using the transition probability (2.2) from state  $u_n$  to  $u_{n+1}$  as derived in section 2.1 and the frictional potential (4.4). Given an initial state  $u_0(x)$ , the phase field is set to be constant within each obstacle, i.e.  $u_0(x) = \xi_i^0$  if  $x \in \mathbb{B}_r(p_i)$ . The system is allowed to relax for a given applied stress  $\bar{\tau}$ . After the system reaches a relaxed state the values of the phase field at the obstacles points are updated with the new values. These values are used to compute the friction at the obstacles for the next time step. In calculations we allow for a general frictional law of the form

$$\sigma = \sigma_0 \bar{u}(t)^\beta \quad (4.20)$$

where  $\sigma$  is the frictional resistance,  $\sigma_0$  is a flow stress,  $\beta$  is a hardening exponent and  $\bar{u}$  is the average slip. This allows an elucidation of the separate effects of the flow stress and the hardening exponent on the creep behavior of the material. For instance, the models Cottrell [13] and Daehn [14] predict that steady, Andrade and logarithmic creep result from increasing hardening rates, respectively, in agreement with our model.

Figure 4.2 shows the numerical simulations of the phase field dislocation model for different values of hardening coefficient at an applied stress  $\bar{\tau} = 0.5$  and  $\theta = 1$ . The ideally plastic case,  $\beta = 0$ , exhibits a power-law dependence on  $t$  of the form given in equation 4.19 with  $\frac{1}{1+2\beta} \sim 0.94 \pm 0.02$ . Figure 4.2 also shows the simulations for linear hardening, i. e.,  $\beta = 1$ . The simulated average slip follows a power-law relation with respect to time with exponents  $\frac{1}{1+2\beta} \sim 0.27 \pm 0.01$  and  $\frac{1}{1+2\beta} \sim 0.28 \pm 0.01$  for flow stresses  $\sigma_0 = 1$  and  $\sigma_0 = 2$  respectively, in agreement with Andrade creep.

## 5 Summary and conclusion

We have developed a theory of statistical mechanics for dissipative systems governed by equations of evolution that assigns probabilities to individual trajectories of the system. Important consequences of this theory can be analyzed rigorously which leads to precise predictions regarding the behavior of dissipative systems at finite temperature. Such predictions include the effect of temperature on yield phenomena and rheological time exponents. The particular case of an ensemble of dislocations moving in a slip plane through a random array of obstacles has been studied numerically in detail. The numerical results bear out the analytical predictions regarding the mean response of the system, which exhibits Andrade creep.

We conclude by pointing out some possible extensions and additional applications of the theory.

The theory developed in this paper may be viewed as fitting within current efforts to develop a calculus of variations for evolutionary systems and, eventually, for systems ex-

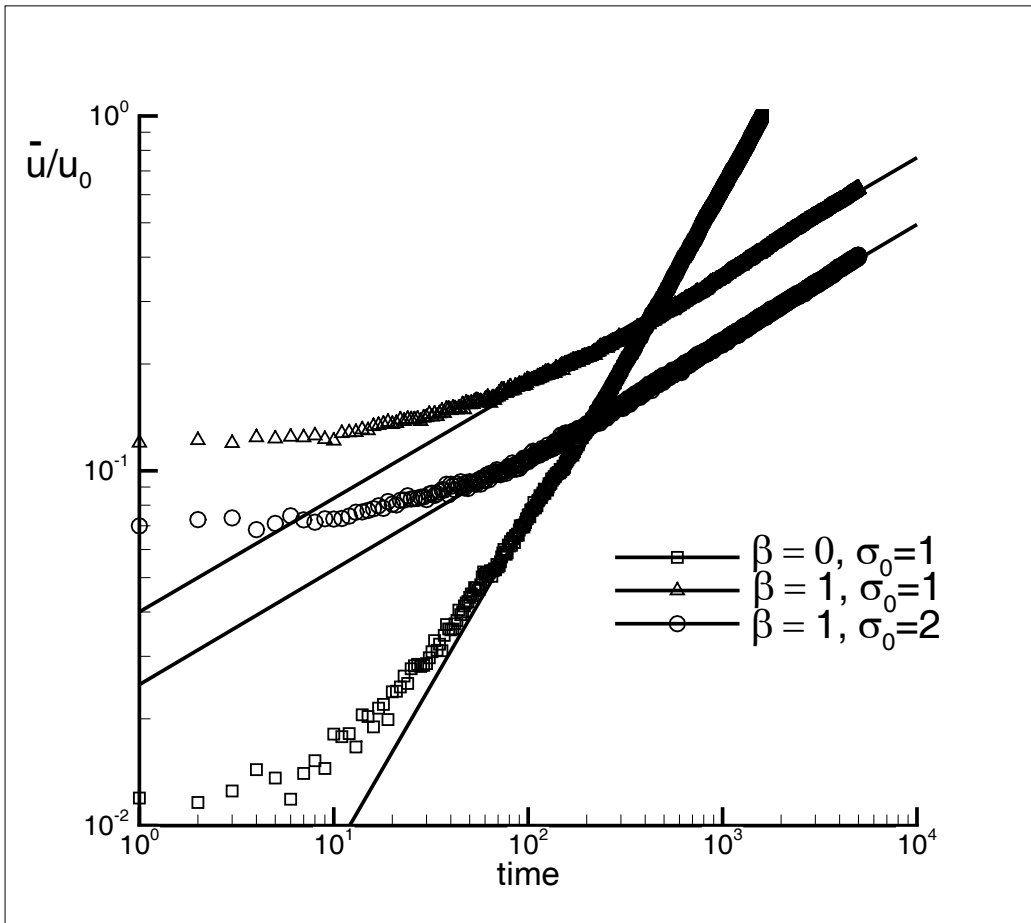


Figure 4.2: A log-log plot of  $\bar{u}(t)/u_0$  vs.  $t$ . Numerical simulations of the phase-field dislocation model: no-hardening ideally-plastic case ( $\beta = 0$ ); linear-hardening case ( $\beta = 1$ ) at yield stress  $\sigma_0 = 1$ ; and linear-hardening case ( $\beta = 1$ ) at yield stress  $\sigma_0 = 2$ .

hibiting evolving microstructure (cf, e. g., [11, 20]). The aim of these efforts is to formulate variational principles that characterize the evolution of dissipative systems. These principles may in turn be regarded as selection criteria that pick, among all possible trajectories of the system, those that are expected to be realized physically. In the present work, the selection criterion is based on probability, i. e., we seek to determine the trajectories that are most likely to occur for given conditions. The device of placing the system in contact with a thermal bath allows the system to explore all possible trajectories and us to determine which trajectories are traversed most frequently.

This finite-temperature “phase-space exploration” is expected to be most revealing when the energy landscape of the system is rough and the system tends to develop microstructure. Of particular interest are continuum systems whose states are described by fields in infinite-dimensional functional spaces and governed by energy functionals lacking lower semi-

continuity. In such cases, the preferred or most likely trajectories may be viewed as defining the effective or macroscopic energetics and kinetics of the system. Such generality, however, is beyond the scope of the present work, which is restricted to the finite-dimensional setting, and should be the focus of future studies, cf. [29].

Finally, we remark that the heuristic derivation of the theory presented in section 2 carries over to dynamics, e. g., by recourse to the incremental minimum principles put forth in [27]. However, a complication introduced by dynamics is that the resulting processes are not Markovian, owing to the fact that the transition probabilities depend on two preceding states. This non-Markovian character of the processes compounds the analysis and will require significant extensions thereof.

## 6 Appendices

### 6.1 Convergence of random variables

This appendix gives a brief review of the theory of random variables and the various associated notions of convergence. For a thorough treatment, see, for example, [4, 5, 26]. Throughout this section,  $(\Omega, \mathcal{F}, \mathbb{P})$  will denote a (fixed) probability space, and  $(M, d)$  will denote a separable metric space (separability ensures that the metric is a Borel-measurable function).

An  $M$ -valued *random variable* is a function  $X: \Omega \rightarrow M$  that is measurable with respect to the  $\sigma$ -algebra  $\mathcal{F}$  on  $\Omega$  and the Borel  $\sigma$ -algebra  $\mathcal{B}(M)$  on  $M$ . The *distribution* or *law* of  $X$  is the push-forward probability measure  $X_*(\mathbb{P})$  on  $M$  defined by

$$X_*(\mathbb{P})(A) = \mathbb{P}[X \in A] := \mathbb{P}(X^{-1}(A)) \text{ for all } A \in \mathcal{B}(M).$$

A sequence of Borel probability measures  $(\mu_k)_{k \in \mathbb{N}}$  on  $M$  is said to *converge weakly* as  $k \rightarrow \infty$  to a probability measure  $\mu$  on  $M$ , denoted  $\mu_k \rightharpoonup \mu$ , if

$$\lim_{k \rightarrow \infty} \int_M \phi \, d\mu_k = \int_M \phi \, d\mu \text{ for all bounded, continuous } \phi: M \rightarrow \mathbb{R}. \quad (6.1)$$

A sequence of random variables  $(X_k)_{k \in \mathbb{N}}$  is said to *converge weakly/in distribution/in law* as  $k \rightarrow \infty$  to a random variable  $X$ , denoted  $X_k \rightharpoonup X$ , if the distributions  $(X_k)_*(\mathbb{P})$  converge weakly to  $X_*(\mathbb{P})$ .  $(X_k)_{k \in \mathbb{N}}$  is said to *converge in probability* to  $X$ , denoted  $X_k \xrightarrow{\mathbb{P}} X$ , if, for every  $\lambda > 0$ ,

$$\lim_{k \rightarrow \infty} \mathbb{P}[d(X_k, X) \geq \lambda] = 0. \quad (6.2)$$

Heuristically, convergence in probability implies that, for large  $k$ ,  $X$  and  $X_k$  are strongly correlated. This is a stronger condition than convergence in law, since correlations are “invisible” when one considers only the laws of the random variables. As a simple example, take  $\Omega = M = \{-1, +1\}$ ,  $\mathbb{P} =$  uniform measure, the variables  $X_k$  to be independent copies of the identity map, and  $X: \pm 1 \mapsto \mp 1$ . Then  $X$  and all the  $X_k$  are equal in law (so

certainly  $X_k \rightarrow X$ ) but  $X$  and any  $X_k$  are *anti*-correlated with  $|X - X_k| \equiv 2$ , so there is no convergence in probability.

Given  $p > 0$ ,  $(X_k)_{k \in \mathbb{N}}$  is said to *converge in  $p^{\text{th}}$  mean* to  $X$  if

$$\lim_{k \rightarrow \infty} \mathbb{E}[d(X_k, X)^p] = 0. \quad (6.3)$$

When  $M$  is a normed space, convergence in  $p^{\text{th}}$  mean is simply strong convergence in the normed space  $L^p(\Omega, \mathcal{F}, \mathbb{P}; M)$  of random variables with finite  $p^{\text{th}}$  moment. Finally,  $(X_k)_{k \in \mathbb{N}}$  is said to *converge almost surely* to  $X$  if

$$\mathbb{P}\left[\lim_{k \rightarrow \infty} X_k = X\right] = 1. \quad (6.4)$$

Almost sure convergence implies convergence in probability, which, in turn, implies convergence in law. The converse implications are more delicate: if  $X_k \rightarrow X$  and  $\mathbb{P}[X = x] = 1$  for some  $x \in M$ , then  $X_k \xrightarrow{\mathbb{P}} X$ ; and  $X_k \xrightarrow{\mathbb{P}} X$  if, and only if, every subsequence of  $(X_k)$  has a further subsequence that converges to  $X$  almost surely. Further, by Chebyshev's inequality, convergence in  $p^{\text{th}}$  mean for any  $p > 0$  implies convergence in probability.

The following maximal inequality, known as Doob's submartingale inequality, bounds the probability that a sequence of random variables exceeds some value  $\lambda$  in terms of the expected value of the last term of the sequence. It includes as special cases the well-known inequalities of Chebyshev and Kolmogorov.

**Theorem 6.1** (Doob's submartingale inequality). *Let  $(X_k)_{k \in \mathbb{N}}$  be a sequence of non-negative  $\mathbb{R}$ -valued random variables with  $\mathbb{E}[X_{k+1} | X_1, \dots, X_k] \geq X_k$  for each  $k \in \mathbb{N}$ . Then, for any  $n \in \mathbb{N}$ ,  $p \geq 1$  and  $\lambda > 0$ ,*

$$\mathbb{P}\left[\max_{1 \leq k \leq n} X_k \geq \lambda\right] \leq \frac{1}{\lambda^p} \mathbb{E}[X_n^p]. \quad (6.5)$$

*In particular, taking  $X_k = |Y_1 + \dots + Y_k|^2$  for a sequence of mean-zero, independent  $\mathbb{R}$ -valued random variables  $(Y_k)_{k \in \mathbb{N}}$ , (6.5) implies Kolmogorov's inequality:*

$$\mathbb{P}\left[\max_{1 \leq k \leq n} |Y_1 + \dots + Y_k| \geq \lambda\right] \leq \frac{1}{\lambda^2} \text{Var}[Y_1 + \dots + Y_n] \equiv \frac{1}{\lambda^2} \sum_{k=1}^n \text{Var}[Y_k]; \quad (6.6)$$

*and also implies Chebyshev's inequality for a single random variable  $Y$ :*

$$\mathbb{P}[|Y| \geq \lambda] \leq \frac{1}{\lambda^p} \mathbb{E}[|Y|^p]. \quad (6.7)$$

## 6.2 Proofs and auxiliary results

**Lemma 6.2** (Formulae for the expected value and higher moments). *For the approximate incremental cost function (2.3) with  $E$  of the form (3.11) and  $\Psi$  positively homogeneous of degree one, and any  $x_n \in \mathbb{R}^N$ ,  $h, \varepsilon > 0$ ,*

$$\mathbb{E}[X_{i+1} | X_i] = X_i + \varepsilon \frac{I_1^\varepsilon(X_i)}{I_0^\varepsilon(X_i)},$$

and, for  $k \in \mathbb{N}$ ,

$$\mathbb{E} \left[ |X_{i+1} - \mathbb{E}[X_{i+1}|X_i]|^k \middle| X_i \right] \leq C_k(X_i)\varepsilon^k,$$

where

$$I_k^\varepsilon(x_i) := \int_{\mathbb{R}^N} z^k \exp \left( - \left( (Ax_i - \ell(t_{i+1})) \cdot z + \Psi(z) + \frac{\varepsilon}{2} z \cdot Az \right) \right) dz. \quad (6.8)$$

Further,  $C_k(x)$  can be chosen to be continuous in  $x$ , and even independent of  $x$  if  $A = 0$ .

*Proof.* The claims follow directly after re-writing the cost function  $C$  as

$$\begin{aligned} C(x_i, x_{i+1}) &= (Ax_i - \ell(t_{i+1})) \cdot (x_{i+1} - x_i) + \Psi(x_{i+1} - x_i) + \frac{1}{2}(x_{i+1} - x_i) \cdot A(x_{i+1} - x_i), \end{aligned}$$

and defining  $z$ , the dummy variable of integration, by  $z := (x_{i+1} - x_i)/\varepsilon$ . That  $C_k(x)$  does not depend on  $x$  for  $A = 0$  is obvious; as usual, the continuous dependence on  $x$  follows from Lebesgue's dominated convergence theorem.  $\square$

*Proof of theorem 3.2.* Fix  $T > 0$  and, for brevity, let  $\mathcal{C}^0 = \mathcal{C}^0([0, T]; \mathbb{R}^N)$  with its usual uniform norm. Define a deterministic sequence  $y^{h,\varepsilon}$ , approximating  $y^\theta$ , by the implicit Euler scheme:

$$y_0^{h,\varepsilon} := x_0; \quad (6.9a)$$

$$y_i^{h,\varepsilon} := y_{i-1}^{h,\varepsilon} + \varepsilon f^0(t_i); \quad (6.9b)$$

i. e.  $\Delta y_i^{h,\varepsilon} = \varepsilon f^0(t_i)$ . Let  $Z_i^{h,\varepsilon} := X_i^{h,\varepsilon} - y_i^{h,\varepsilon}$ . By classical results concerning the global error associated to an implicit Euler scheme,  $\|\bar{y}^{h,\theta h} - y^\theta\|_{\mathcal{C}^0} \in O(h)$ ; thus, it is enough to show that, for all  $\lambda > 0$ ,

$$\mathbb{P} \left[ \|\bar{Z}^{h,\theta h}\|_{\mathcal{C}^0} \geq \lambda \right] \in O(h) \text{ as } h \downarrow 0. \quad (6.10)$$

Clearly, (6.10) is implied by the condition that, for all  $\lambda > 0$ ,

$$\mathbb{P} \left[ \max_{1 \leq i \leq \lceil T/h \rceil} |Z_i^{h,\theta h}|_2 \geq \lambda \right] \in O(h) \text{ as } h \downarrow 0. \quad (6.11)$$

By lemma 6.2, for  $k \geq 1$ ,

$$\mathbb{E}[\Delta X_i^{h,\varepsilon} | X_{i-1}^{h,\varepsilon}] = \varepsilon f^0(t_i); \quad (6.12a)$$

$$\mathbb{E} \left[ |\Delta X_i^{h,\varepsilon} - \mathbb{E}[\Delta X_i^{h,\varepsilon} | X_{i-1}^{h,\varepsilon}]|^k \middle| X_{i-1}^{h,\varepsilon} \right] \leq C_k \varepsilon^k. \quad (6.12b)$$

(Note that  $C_k$  is independent of  $i$ ,  $h$  and  $\varepsilon$ .) In fact, in this case, the conditional expectation and variance do not depend on the prior state  $X_{i-1}^{h,\varepsilon}$ , and so are equal to the unconditional expectation and variance respectively. Thus,

$$\mathbb{E}[\Delta Z_i^{h,\theta h}] = 0; \quad (6.13a)$$

$$\mathbb{E}[|\Delta Z_i^{h,\theta h}|_2^k] \leq C_k \theta^k h^k. \quad (6.13b)$$

By Kolmogorov's inequality (6.6), for some constant  $C'$  depending on  $N$ ,

$$\mathbb{P} \left[ \max_{1 \leq i \leq \lceil T/h \rceil} |Z_i^{h, \theta h}|_2 \geq \lambda \right] \leq \frac{C'}{\lambda^2} \sum_{i=1}^{\lceil T/h \rceil} \mathbb{E} [|\Delta Z_i^{h, \theta h}|_2^2] \leq \frac{hC'\theta^2}{\lambda^2} \sum_{i=1}^{\lceil T/h \rceil} hC_2.$$

The sum on the far right-hand side converges as  $h \downarrow 0$  to  $TC_2 < +\infty$ . Thus, the right-hand side is  $O(h)$ , completing the proof.  $\square$

**Lemma 6.3** (Uniform convergence and stability). *Let  $E$  as in (3.11),  $F^\varepsilon$  be as in (3.12), and let  $f^\varepsilon$  be as in (3.13). Then, for all,  $T \in (0, +\infty)$ , all loads  $\ell: [0, T] \rightarrow (\mathbb{R}^N)^*$  in  $W^{1, \infty}$  and all continuously varying compact  $K(t) \subset\subset \mathcal{S}(t)$ ,*

$$\sup_{t \in [0, T]} \sup_{x \in K(t)} |f^\varepsilon(t, x) - f^0(t, x)|_2 \in O(\sqrt{\varepsilon}) \text{ as } \varepsilon \downarrow 0. \quad (6.14)$$

Furthermore, over any compact interval of time  $[0, T]$ , given  $x_0 \in \dot{\mathcal{S}}(0)$ , the forward orbits  $y^\varepsilon$  and  $y^0$  of  $x_0$  under  $f^\varepsilon$  and  $f^0$  respectively satisfy

$$\sup_{t \in [0, T]} |y^\varepsilon(t) - y^0(t)|_2 \in O(\sqrt{\varepsilon}) \text{ as } \varepsilon \downarrow 0. \quad (6.15)$$

*Proof.* Since the effect of  $\ell$  is only to “shift” the potential  $F^\varepsilon$  in space, the uniformity in time is obvious; it suffices to show uniform convergence in space for time  $t = 0$ . Neglecting load/time dependence, write  $f^\varepsilon(x) = I_1^\varepsilon(x)/I_0^\varepsilon(x)$  as in (6.8). Then, for  $x \in \mathcal{S}$ ,

$$\left| \frac{I_1^\varepsilon(x)}{I_0^\varepsilon(x)} - \frac{I_1^0(x)}{I_0^0(x)} \right|_2 \leq \frac{1}{I_0^0(x)} |I_1^\varepsilon(x) - I_1^0(x)|_2 + \frac{1}{I_0^0(x)I_0^\varepsilon(x)} |I_1^0(x)|_2 |I_0^\varepsilon(x) - I_0^0(x)|.$$

Hence, letting  $K \subset\subset \mathcal{S}$ , since all the denominators are bounded away from 0,

$$\|f^\varepsilon - f^0\|_{C^0(K)} \leq c_1 \|I_1^\varepsilon - I_1^0\|_{C^0(K)} + c_2 \|I_0^\varepsilon - I_0^0\|_{C^0(K)}.$$

Let  $\kappa$  be greater than or equal to the largest eigenvalue of  $A$ , and, for  $x \in \mathcal{S}$ , let

$$m(x) := \inf_{|\omega|=1} \Psi(\omega) + (Ax - \ell) \cdot \omega > 0.$$

The strategy now is to use the bound  $e^{-\varepsilon z \cdot Az} \geq e^{-\varepsilon \kappa |z|^2/2}$  and to split the integral

$$I_k^0(x) - I_k^\varepsilon(x) = \int_{\mathbb{R}^N} (1 - \exp(-\frac{\varepsilon}{2} z \cdot Az)) \exp(-((Ax - \ell) \cdot z + \Psi(z))) dz$$

into two terms: an integral over a ball of radius  $r_*$  about 0, where  $1 - e^{-\varepsilon z \cdot Az} \leq a \in (0, 1)$ , and the rest, where  $1 - e^{-\varepsilon z \cdot Az} \leq 1$ . Using spherical polar coordinates yields

$$|I_k^\varepsilon(x) - I_k^0(x)| \leq c_N a \int_0^{r_*} r^{k+N-1} e^{-m(x)r} dr + c_N \int_{r_*}^{+\infty} r^{k+N-1} e^{-m(x)r} dr.$$

Specifically, take  $r_* = -\frac{2}{\varepsilon\kappa} \log(1-a)$  and  $a = \sqrt{\varepsilon}$ . Then, by l'Hôpital's rule,  $r_* \rightarrow +\infty$  as  $\varepsilon \rightarrow 0$ . Hence,  $e^{-m(x)r_*} \rightarrow 0$  exponentially fast as  $\varepsilon \rightarrow 0$ . Further calculation yields

$$|I_k^\varepsilon(x) - I_k^0(x)| \leq c_1 \frac{1}{r_*^{k+N}} e^{-m(x)r_*} + c_2 \sqrt{\varepsilon}.$$

Since  $m(x)$  is uniformly bounded away from 0 for  $x \in K$ , (6.14) follows.

Now let  $x_0$  in the interior of  $\mathcal{S}(0)$  be arbitrary and let  $y^\varepsilon$  be the flow in  $f^\varepsilon$  starting at  $x_0$ ; let  $z^\varepsilon(t) := y^\varepsilon(t) - y^0(t)$ . Since  $l$  is essentially bounded and the potential  $F^0(DE(t, \cdot))$  blows up on  $\partial\mathcal{S}(t)$ ,  $y^0$  is bounded away  $\partial\mathcal{S}(\cdot)$ ; indeed, if  $l$  is constant, then  $F^0(DE(t, y^0(t)))$  decreases with time. Since  $F^\varepsilon \rightarrow F^0$ , the same conclusion holds for  $y^\varepsilon$  with  $\varepsilon$  small enough. Hence, choose  $\delta > 0$  small enough that

$$K(\cdot) := \{x \in \mathcal{S}(t) \mid \text{dist}(x, \partial\mathcal{S}(\cdot)) > \delta\}$$

contains the forward orbits of  $y^0$  and  $y^\varepsilon$  (for small enough  $\varepsilon$ ), i.e.  $y^\varepsilon(t) \in K(t)$  for all  $t \in [0, T]$ . Then

$$\begin{aligned} |z^\varepsilon(t)|_2 &\leq \int_0^t (|f^\varepsilon(s, y^\varepsilon(s)) - f^\varepsilon(s, y^0(s))|_2 + |f^\varepsilon(s, y^0(s)) - f^0(s, y^0(s))|_2) \, ds \\ &\leq \int_0^t \text{Lip}(f^\varepsilon|_{K(s)}) |z^\varepsilon(s)|_2 \, ds + T \|f^\varepsilon(0, \cdot) - f^0(0, \cdot)\|_{C^0(K(0))} \\ &\leq L \int_0^t |z^\varepsilon(s)|_2 \, ds + T \|f^\varepsilon(0, \cdot) - f^0(0, \cdot)\|_{C^0(K(0))}, \end{aligned}$$

where  $L = \text{Lip}(f^\varepsilon|_{K(s)}) < +\infty$  for any (all)  $s \geq 0$ . By Grönwall's inequality and (6.14), (6.15) follows.  $\square$

*Proof of theorem 3.4.* Without loss of generality, simplify notation by taking  $\theta = 1$ ,  $\varepsilon = h$  and dropping the upper indices from  $X^{h,\varepsilon}$  &  $c$ . By the  $O(\sqrt{h})$  estimate of lemma 6.3 and the standard  $O(h)$  global error estimate for the Euler method with time step  $h$  in the vector field  $f^h$ , it more than enough to show that

$$\mathbb{P} [|X_i^h - y_i^h|_2 \geq \lambda \text{ for some } 0 \leq i \leq T/h] \in O(h) \text{ as } h \downarrow 0; \quad (6.16)$$

an  $O(\sqrt{h})$  bound would suffice to prove the theorem. Let

$$K(t) := \{x \in \mathcal{S}(t) \mid \text{dist}(x, \partial\mathcal{S}(t)) > \delta\}$$

with  $\delta > 0$  smaller than the modulus of continuity for  $l$  with step  $h$  and also small enough that  $y_t \in K(t)$  for all  $t \in [0, T]$ . Write

$$\begin{aligned} X_{i+1} &= X_i + h f^h(t_{i+1}, X_i) + \Xi_{i+1}(X_i), \\ y_{i+1} &= y_i + h f^h(t_{i+1}, y_i). \end{aligned}$$

By lemma 6.2, for each  $x$ ,  $\Xi_{i+1}(x)$  is a random variable with mean 0 and  $k^{\text{th}}$  central moment at most  $C_k(x)h^k$ . The deviations  $Z$  satisfy

$$Z_{i+1} = Z_i + h(f^h(t_{i+1}, X_i) - f^h(t_{i+1}, y_i)) + \Xi_{i+1}(X_i). \quad (6.17)$$

Recall that the vector field  $f^h(t_{i+1}, \cdot)$

(M) is a monotonically decreasing vector field on  $\mathcal{S}(t_{i+1})$ ;

(B) is bounded on compactly-embedded subsets of  $\mathcal{S}(t_{i+1})$ .

Let  $\mathcal{K}_i$  be (the  $\sigma$ -algebra generated by) the event that  $X_j \in K(t_{j+1})$  for  $0 \leq j \leq i$ . Applying the conditional expectation operator  $\mathbb{E}[-|\mathcal{K}_i]$  (which is never conditioning on an event of zero probability) to the square of (6.17) yields:

$$\begin{aligned} & \mathbb{E}[|Z_{i+1}|_2^2 | \mathcal{K}_i] - \mathbb{E}[|Z_i|_2^2 | \mathcal{K}_i] \\ &= 2h\mathbb{E}[(f^h(t_{i+1}, X_i) - f^h(t_{i+1}, y_i)) \cdot Z_i | \mathcal{K}_i] \leq 0 \text{ by (M)} \\ &+ 2\mathbb{E}[Z_i \cdot \Xi_{i+1}(X_i) | \mathcal{K}_i] = 0 \text{ since } \mathbb{E}[\Xi_{i+1}(x)] \equiv 0 \\ &+ 2h\mathbb{E}[(f^h(t_{i+1}, X_i) - f^h(t_{i+1}, y_i)) \cdot \Xi_{i+1}(X_i) | \mathcal{K}_i] = 0 \text{ for the same reason} \\ &+ h^2\mathbb{E}[|f^h(t_{i+1}, X_i) - f^h(t_{i+1}, y_i)|_2^2 | \mathcal{K}_i] \leq Ch^2 \text{ by (B)} \\ &+ \mathbb{E}[|\Xi_{i+1}(X_i)|_2^2 | \mathcal{K}_i] \leq Ch^2 \text{ by lemma 6.2} \\ &\leq Ch^2. \end{aligned}$$

Applying the unconditional expectation operator to both sides yields the following uniform bound for the second moment of the deviations:

$$\max_{0 \leq i \leq T/h} \mathbb{E}[|Z_i|_2^2] \leq CTh. \quad (6.18)$$

The inequality (6.18) is insufficient to establish (6.16), but can be used to “bootstrap” a similar inequality for the fourth moments. Define a tetralinear form  $\langle \cdot, \cdot, \cdot, \cdot \rangle: \mathbb{R}^N \rightarrow \mathbb{R}$  by

$$\langle w, x, y, z \rangle := (w \cdot x)(y \cdot z), \quad (6.19)$$

so that  $|x|_2^4 = \langle x, x, x, x \rangle$ . This tetralinear form is invariant under arbitrary compositions of the following interchanges of entries: (1, 2), (3, 4) and (1, 3)(2, 4). The Cauchy–Schwarz inequality for the Euclidean inner product implies a corresponding Cauchy–Schwarz-type inequality for this tetralinear form: for all  $w, x, y, z \in \mathbb{R}^N$ ,

$$|\langle w, x, y, z \rangle| \leq |w|_2 |x|_2 |y|_2 |z|_2. \quad (6.20)$$

Hence,  $\mathbb{E}[|Z_{i+1}|_2^4] \equiv \mathbb{E}[\mathbb{E}[|Z_{i+1}|_2^4 | \mathcal{K}_i]]$  can be expanded using the tetralinear form (6.19) and (6.17) and each term estimated as in the derivation of (6.18). Those terms containing precisely one  $\Xi_{i+1}(X_i)$  have zero expectation; the terms of the form

$$\mathbb{E}[\langle Z_i, Z_i, Z_i, h(f^h(t_{i+1}, X_i) - f^h(t_{i+1}, y_i)) \rangle | \mathcal{K}_i]$$

are non-positive by (M); the remaining terms can all be estimated using (B), (6.18), (6.20) and lemma 6.2, with the worst bound being  $O(h^3)$ . Thus, the following uniform bound for the fourth moment of the deviations holds:

$$\max_{0 \leq i \leq T/h} \mathbb{E}[|Z_i|_2^4] \leq CT h^2. \quad (6.21)$$

Hence, for  $\lambda > 0$ ,

$$\begin{aligned} & \mathbb{P}[|Z_i|_2 \geq \lambda \text{ for some } 0 \leq i \leq T/h] \\ & \leq \sum_{i=0}^{T/h} \mathbb{P}[|Z_i|_2 \geq \lambda] \\ & \leq \sum_{i=0}^{T/h} \lambda^{-4} \mathbb{E}[|Z_i|_2^4] \text{ by Chebysh\u00e9v's inequality} \\ & \leq \lambda^{-4} CT^2 h \text{ by (6.21),} \end{aligned}$$

which establishes (6.16) and completes the proof.  $\square$

## Acknowledgments

The support of the DoE through Caltech's PSAAP Center for the Predictive Modeling and Simulation of High-Energy Density Dynamic Response of Materials is gratefully acknowledged. TJS' research is supported by the EPSRC and University of Warwick Research Development Fund.

## References

- [1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Z\u00fcrich. Birkh\u00e4user Verlag, Basel, 2005.
- [2] E. N. da C. Andrade. *Proceedings of the Royal Society of London*, A(84):1, 1910.
- [3] E. N. da C. Andrade. *Proceedings of the Royal Society of London*, A(90):329, 1914.
- [4] P. Billingsley. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition, 1995. A Wiley-Interscience Publication.
- [5] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

- 
- [6] H. Brézis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas dépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(20):Ai, A1197–A1198, 1976.
- [7] H. Brézis and I. Ekeland. Un principe variationnel associé à certaines équations paraboliques. Le cas indépendant du temps. *C. R. Acad. Sci. Paris Sér. A-B*, 282(17):Aii, A971–A974, 1976.
- [8] P. M. Chaikin and T. C. Lubensky. *Principles of Condensed Matter Physics*. Cambridge University Press, 1995.
- [9] D. C. Chrzan and M. J. Mills. Criticality in the plastic deformation of  $\text{Ni}_3\text{Al}$ . *Physical Review Letters*, 69(19):2795–2798, 1992.
- [10] P. Colli and A. Visintin. On a class of doubly nonlinear evolution equations. *Comm. Partial Differential Equations*, 15(5):737–756, 1990.
- [11] S. Conti and M. Ortiz. Minimum principles for the trajectories of systems governed by rate problems. *J. Mech. Phys. Solids*, 56(5):1885–1904, 2008.
- [12] A. H. Cottrell. Criticality in andrade creep. *Philosophical Magazine A*, 74(4):1041–1046, 1996.
- [13] A. H. Cottrell. Strain hardening in andrade creep. *Philosophical Magazine Letters*, 74(5):375–379, 1996.
- [14] G. S. Daehn. Primary creep transients due to non-uniform obstacle sizes. *Materials Science and Engineering*, A(319-321):765–769, 2001.
- [15] A. Garroni and S. Müller.  $\Gamma$ -limit of a phase-field model of dislocations. *SIAM J. Math. Anal.*, 36(6):1943–1964 (electronic), 2005.
- [16] A. Garroni and S. Müller. A variational model for dislocations in the line tension limit. *Arch. Ration. Mech. Anal.*, 181(3):535–578, 2006.
- [17] R. Jordan and D. Kinderlehrer. An extended variational principle. In *Partial differential equations and applications*, volume 177 of *Lecture Notes in Pure and Appl. Math.*, pages 187–200. Dekker, New York, 1996.
- [18] L. V. Kantorovich. On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, 37:199–201, 1942.
- [19] M. Kosłowski, A. M. Cuitiño, and M. Ortiz. A phase-field theory of dislocation dynamics, strain hardening and hysteresis in ductile single crystals. *J. Mech. Phys. Solids*, 50(12):2597–2635, 2002.

- 
- [20] A. Mielke and M. Ortiz. A class of minimum principles for characterizing the trajectories and the relaxation of dissipative systems. *ESAIM: Control, Optimisation and Calculus of Variations*, 14:494516, 2008.
- [21] A. Mielke and F. Theil. On rate-independent hysteresis models. *NoDEA Nonlinear Differential Equations Appl.*, 11(2):151–189, 2004.
- [22] A. Mielke, F. Theil, and V. I. Levitas. A variational formulation of rate-independent phase transformations using an extremum principle. *Arch. Ration. Mech. Anal.*, 162(2):137–177, 2002.
- [23] M. C. Miguel, A. Vespignani, M. Zaiser, and S. Zapperi. Dislocation jamming and Andrade creep. *Physical Review Letters*, 89(16):165501–1–4, 2002.
- [24] M. Ortiz and E. A. Repetto. Nonconvex energy minimization and dislocation structures in ductile single crystals. *Journal of the Mechanics and Physics of Solids*, 47(2):397–462, 1999.
- [25] M. Ortiz and L. Stainier. The variational formulation of viscoplastic constitutive updates. *Computer Methods in Applied Mechanics and Engineering*, 171(3-4):419–444, 1999.
- [26] D. Pollard. *Convergence of stochastic processes*. Springer Series in Statistics. Springer-Verlag, New York, 1984.
- [27] R. A. Radovitzky and M. Ortiz. Error estimation and adaptive meshing in strongly nonlinear dynamic problems. *Computer Methods in Applied Mechanics and Engineering*, 172:203–240, 1999.
- [28] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.
- [29] T. J. Sullivan. *Analysis of Evolutions in Random Environments*. PhD thesis, University of Warwick, Coventry, U.K., 2009. In preparation.
- [30] F. W. Wiegell. *Introduction to Path-Integral Methods in Physics and Polymer Science*. World Scientific, 1986.