

2-minute Research Presentation

Aaron Nung Kwan Yip

Dept. of Mathematics

Purdue University

Nov, 17, 2025

2-minute Research Presentation

Aaron Nung Kwan Yip

Research interests:

Applied Analysis, modeling
(Calculus of Variations,
Materials Science)

2-minute Research Presentation

Aaron Nung Kwan Yip

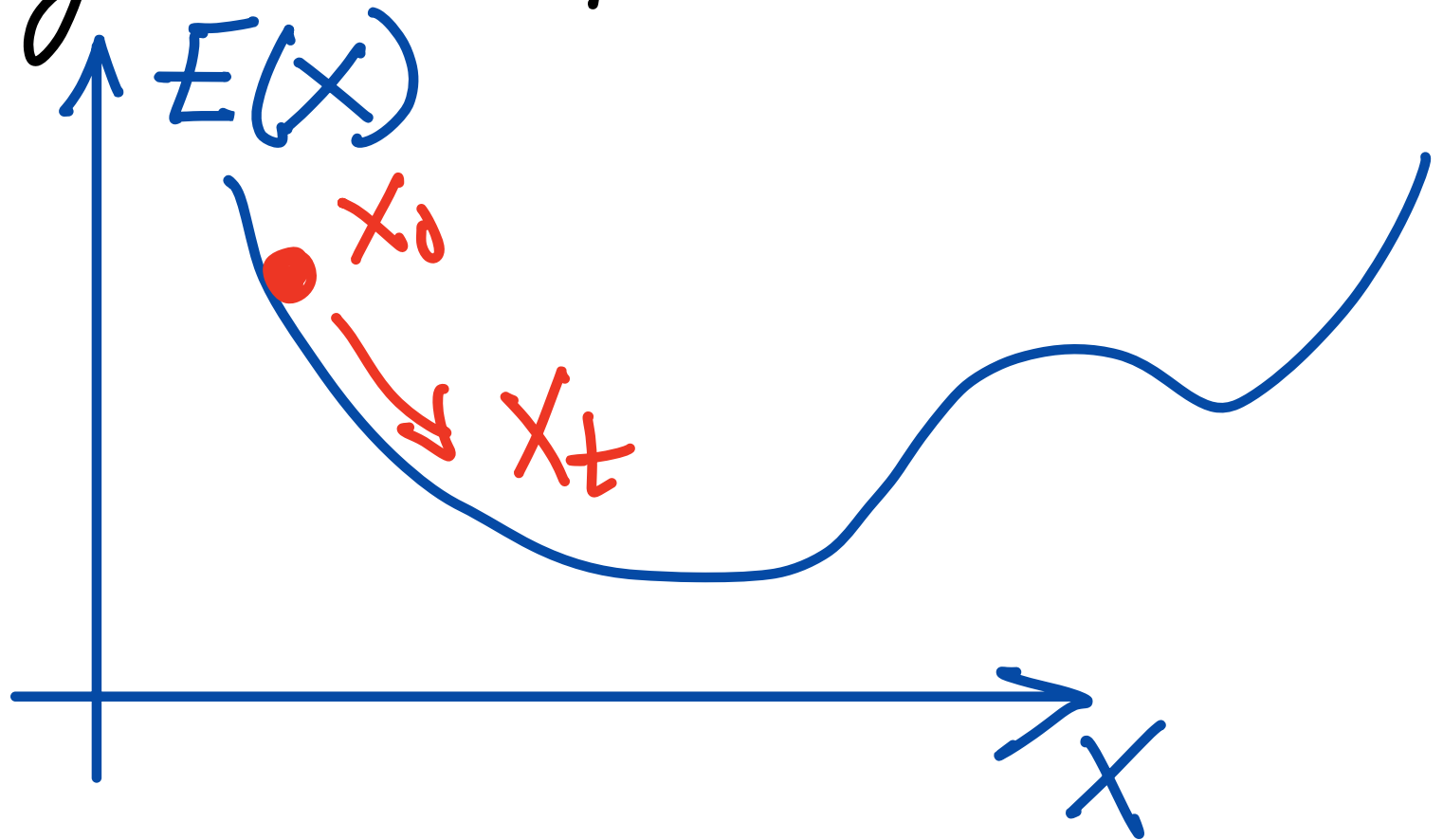
Preferred Interests from Students

Partial Differential Equations

Probability Theory

Differential Geometry

Many Facets of Gradient Descent



$$\frac{dX(t)}{dt} = -\nabla E(X(t))$$

Immediate Consequence of $\dot{x} = -\nabla E(x)$

- $E(x(t)) \downarrow \text{ in } t$

$$\begin{aligned}\frac{d}{dt} E(x(t)) &= \langle \nabla E(x(t)), \dot{x}(t) \rangle \\ &= -\|\nabla E(x(t))\|^2 \leq 0\end{aligned}$$

- $E(t) + \int_0^t \|\nabla E(x(s))\|^2 ds = E(x(0))$

Immediate Consequence of $\dot{x} = -\nabla E(x)$

$\dot{x} = -\nabla E(x)$ is in fact

equivalent to :

Sandier-Serfaty

$$E(x(t)) + \frac{1}{2} \int_0^t \|\nabla E(x(s))\|^2 ds + \frac{1}{2} \int_0^t \|\dot{x}(s)\|^2 ds \leq E(x(0))$$

Immediate Consequence of $\dot{x} = -\nabla E(x)$

$$\frac{d}{dt} E(x(t))$$

$$= \langle \nabla E(x(t)), \dot{x}(t) \rangle$$

$$\dot{x} \parallel -\nabla E(x(t))$$

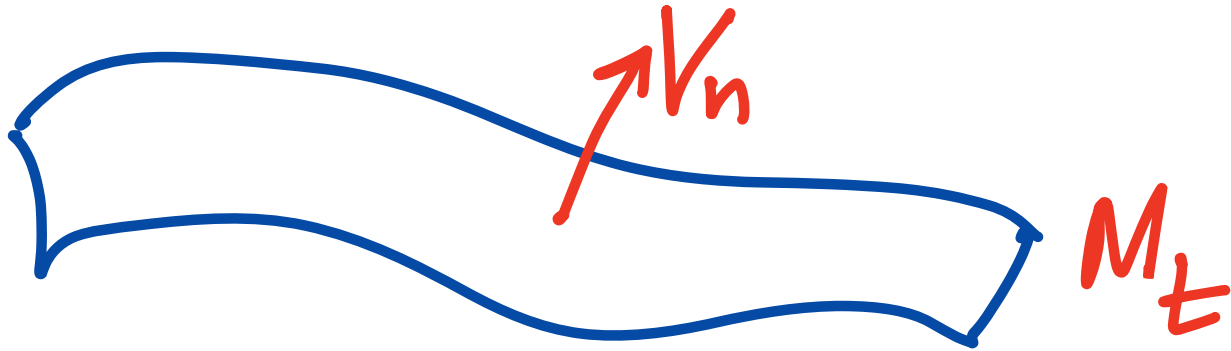
$$\leq -\|\nabla E(x(t))\| \|\dot{x}\|$$

$$\|\dot{x}\| = \|\nabla E(x)\|$$

$$= -\frac{1}{2} \|\nabla E(x(t))\|^2 - \frac{1}{2} \|\dot{x}\|^2$$

$$(ab = -|a||b| = -\frac{1}{2}|a|^2 - \frac{1}{2}|b|^2 \Leftrightarrow a = -b)$$

Gradient Flows and Geometric Evolutions



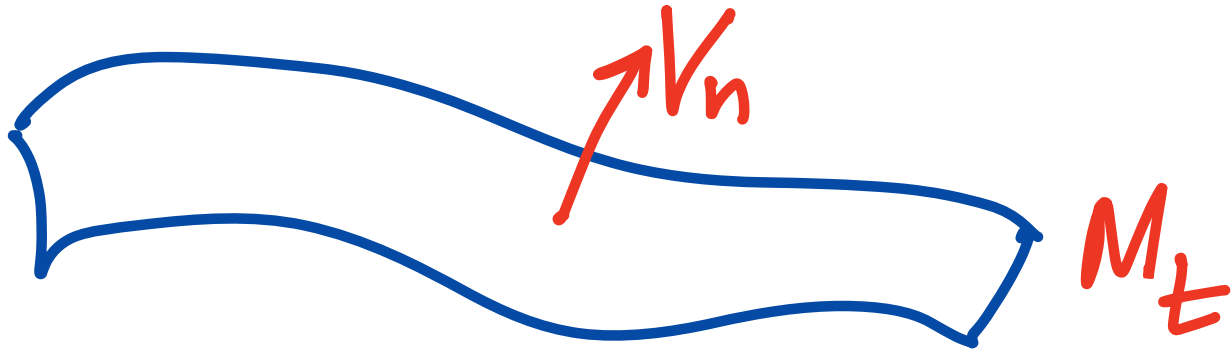
- Motion by Mean Curvature

$$V_n = H$$

- Surface Diffusion

$$V_n = -\operatorname{div}_S H$$

Gradient Flows and Geometric Evolutions



- Motion by Mean Curvature

$$V_n = H$$

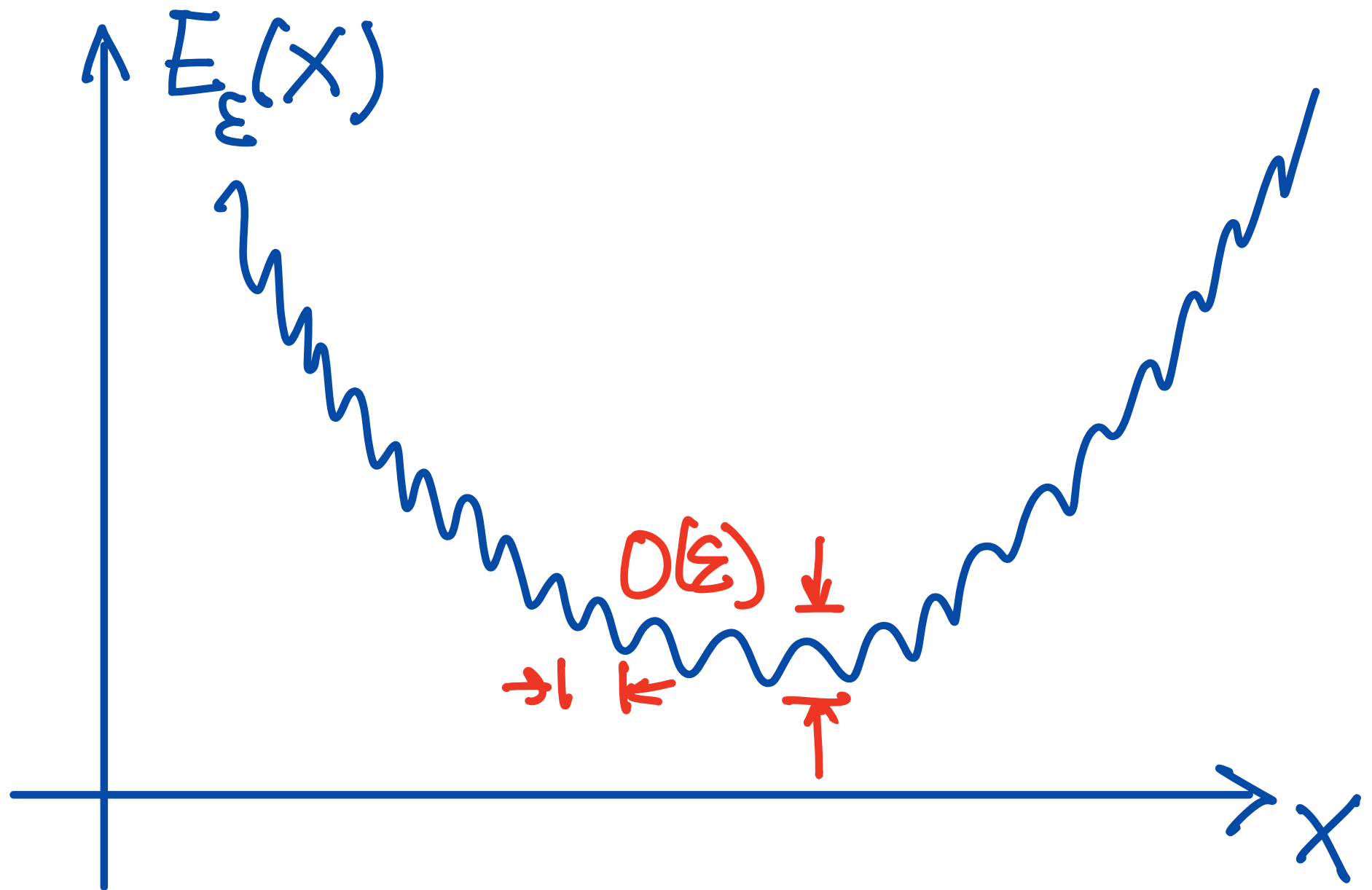
$$V_n = -\nabla_{L^2} E(M, H)$$

- Surface Diffusion

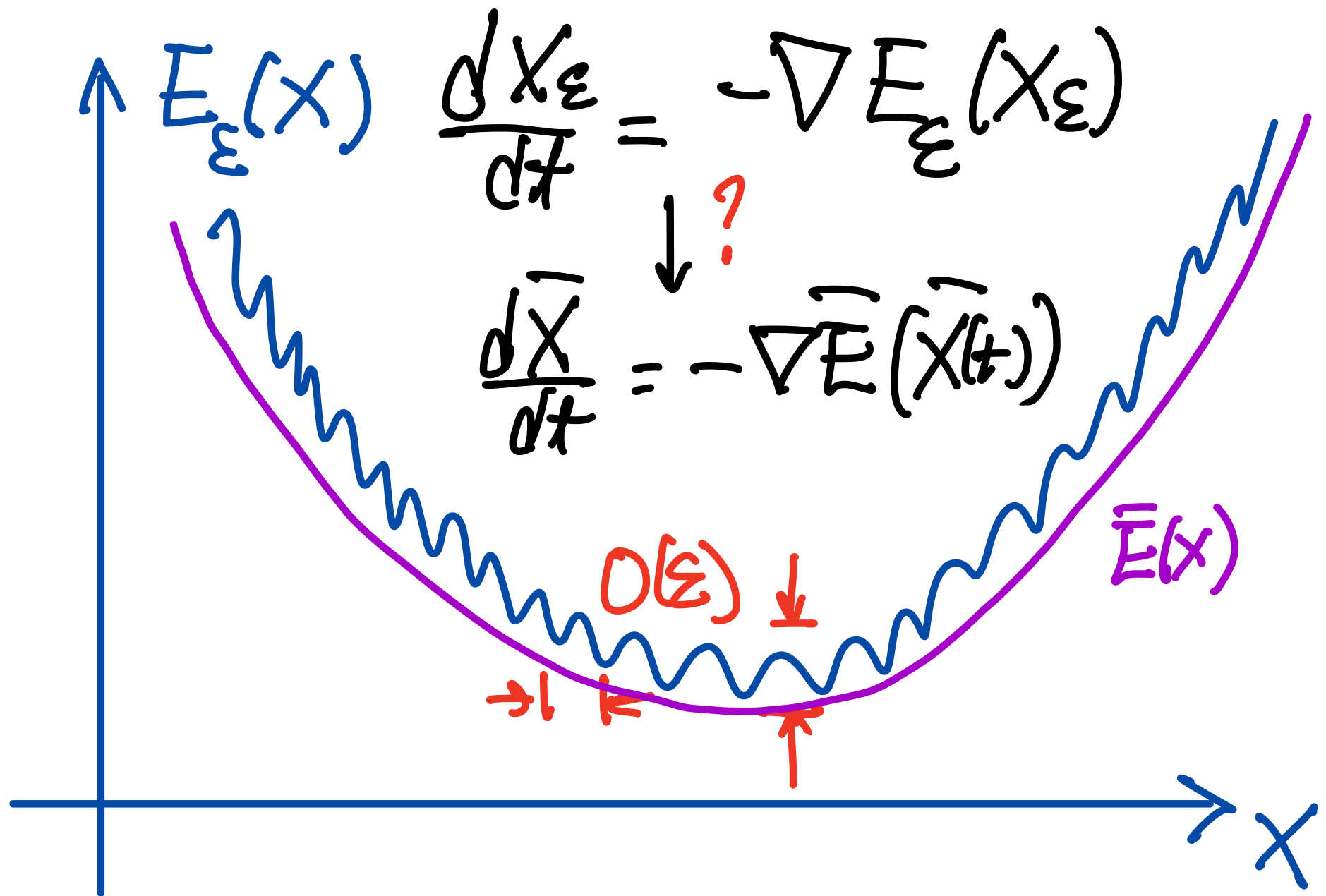
$$V_n = -\operatorname{div}_S H$$

$$V_n = -\nabla_H^{-1} E(M, H)$$

Gradient Flow in Inhomogeneous Medium



Gradient Flow in Inhomogeneous Medium



Gradient Flow in Inhomogeneous Medium

- Motion of Particle in Tilted Periodic Potential (Liang-Y.)

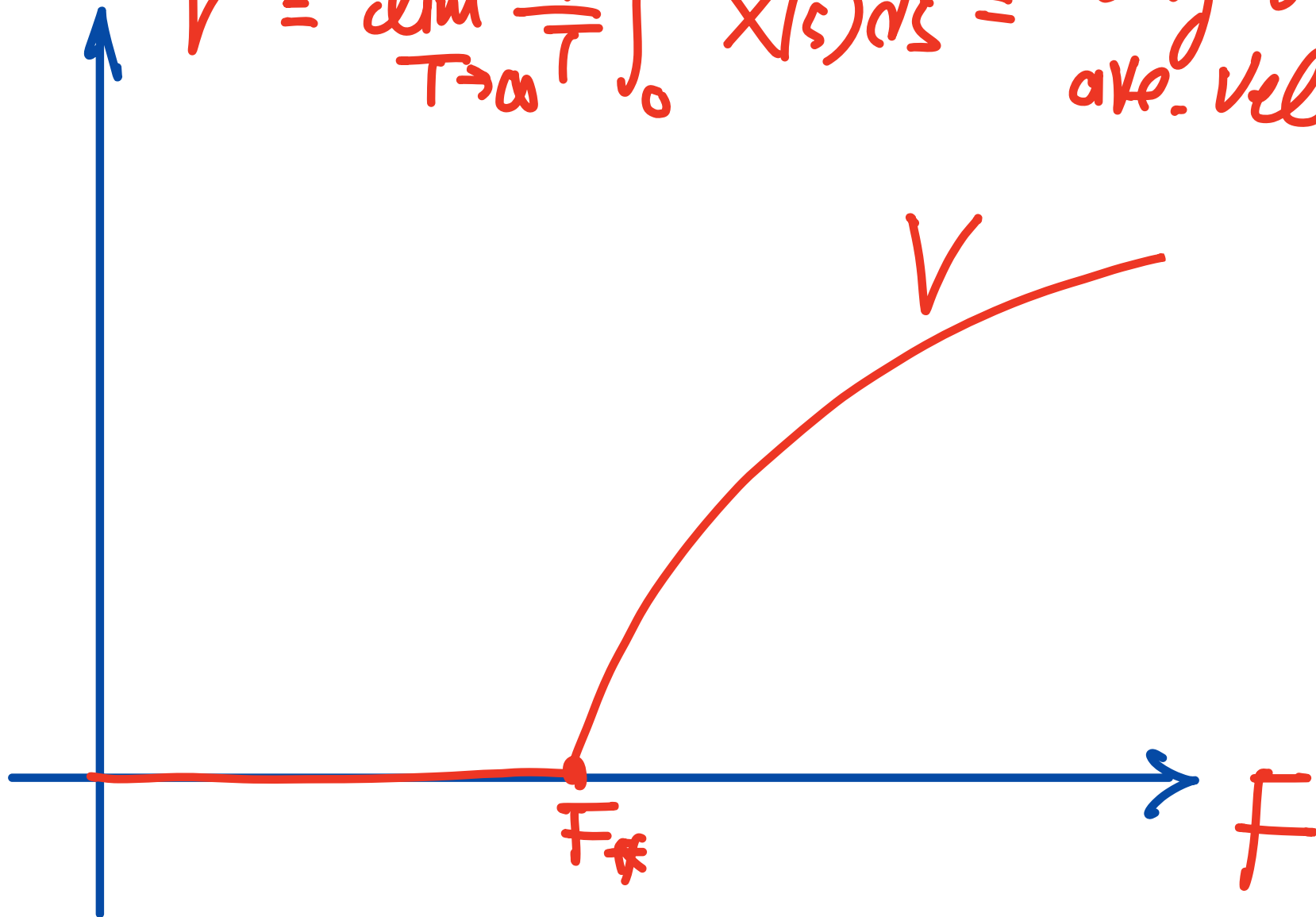
$$\dot{X}(t) = -\sin(X(t)) + F + \dot{W}(t)$$

- Motion by Mean Curvature in Inhomogeneous Medium (Dirr-Y.; Dirr-Karali-Y.)

$$u_t = \Delta u + f(x, u) + F$$

Gradient Flow in Inhomogeneous Medium

$$V = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T X(s) ds = \text{long time ave. velocity}$$



Gradient Flow in Probability Spaces

Heat equation

$$\rho_t = \Delta \rho$$

Gradient Flow in Probability Spaces

Heat equation

$$\rho_t = \Delta \rho$$



$$\rho_t = -\nabla_{L^2} E(\rho)$$

where

$$E(\rho) = \frac{1}{2} \int \|\nabla \rho\|^2 dx \quad (\text{Dirichlet energy})$$

Gradient Flow in Probability Spaces

$$\begin{aligned}\rho_t &= \Delta f \\ &= \nabla \cdot (\nabla f) \\ &= \nabla \cdot \left(f \frac{\nabla f}{f} \right) \\ &= \nabla \cdot (f \nabla \log f)\end{aligned}$$

$$-\nabla_W \text{Ent}(f)$$

- $\text{Ent}(f) = \int f \log f dx$

- $W = \text{Wasserstein Distance}$
(Optimal Transport)

Fokker-Planck Equation in Inhomogeneous Medium (Gao-Y.)

SDE

$$dX_t = -\nabla U_\varepsilon(X_t)dt + \sigma_\varepsilon(X_t) * dW_t$$

FPE

$$\rho_t^\varepsilon = \nabla \cdot (B_\varepsilon^{-1} \nabla \rho_t^\varepsilon) + \nabla \cdot (\rho_t^\varepsilon B_\varepsilon^{-1} \nabla U_\varepsilon)$$

Fokker-Planck Equation in Inhomogeneous Medium (Gao-Y.)

FPE

$$\rho_t^\varepsilon = \nabla \cdot (B_\varepsilon^{-1} \nabla \rho_t^\varepsilon) + \nabla \cdot (\rho_t^\varepsilon B_\varepsilon^{-1} \nabla U_\varepsilon)$$

$$= \nabla \cdot (\rho_t^\varepsilon B_\varepsilon^{-1} \nabla (\log \rho_t^\varepsilon + U_\varepsilon))$$

$$= -\nabla_{W_\varepsilon} E^\varepsilon(\rho_t^\varepsilon)$$

Fokker-Planck Equation in Inhomogeneous Medium (Gao-Y.)

FPE

Energy

$$\bar{E}_{\varepsilon}(\rho) = \int (\rho \log \rho + U_{\varepsilon}(x) \rho) dx$$

entropy

potential energy

Fokker-Planck Equation in Inhomogeneous Medium (Gao-Y.)

FPE

$$\rho_t^\varepsilon = -\nabla_{W_\varepsilon} E^\varepsilon(\rho_t^\varepsilon)$$

$\downarrow \varepsilon \rightarrow 0$

$$\bar{\rho}_t = -\nabla_{\bar{W}} (\bar{E}(\bar{\rho}_t))$$

Stochastic Optimization

Stochastic Gradient Descent

(with Pasupathy et.al.)

$$\rightarrow \underline{f(x) = E \phi(x)}$$

$$\underline{\min_x f(x) \text{ vs } \min_x F_n(x)}$$

$$F_n(x) = \underline{\phi(x, w_1) + \phi(x, w_2) + \dots + \phi(x, w_n)}$$

(empirical average) n

Stochastic Optimization

Stochastic Gradient Descent

(with Pasupathy et.al.)

$$X_{k+1} = X_k - \hat{c} \nabla f(x_k)$$

vs

$$X_{k+1} = X_k - \hat{c}_n \nabla F_n(x_k)$$

convergence
rate?

Thank You
for
Your Attention