Towards a Geometry and Analysis for Bayesian Mechanics

07 March 2022

Dalton A R Sakthivadivel

Formal structures in physics:

Geometry – study of spaces and their properties, such as the state spaces of dynamical systems

Analysis – study of functions and their properties, such as the equations of motion of dynamical systems geometric

Motivation: relate gauge symmetries and gauge

analysis

forces to inference and complex systems,

especially in the context of the free energy principle

Bayesian mech.

Approach:

1. Relate the FEP to maximum entropy
2. Show max ent is actually a gauge theory
3. Profit ?

Overview of some recent results

slides can be found later at darsakthi.github.io/talks ; preprints to follow (soon)

Find p $\arg\max_p \left(- \int \ln p \; p \right)$

subject to $\int \underbrace{\int_{z}(x) \, p(x) \, dx}_{\mathbb{E}[\int(X)]} = C$

$= C$

$$\left\{ \text{Find } p \quad -\int \ln p \; p - \lambda\left(\int g\, p + c\right) \right.$$

is max

$$\frac{\partial}{\partial p}\left(-\ln p \; p - \lambda g\, p\right) = 0$$

$$-\ln p - \lambda g = 0$$

$$p = \exp\{-\lambda g\}$$

$$\lambda g - \lambda g = 0 \quad \checkmark$$

$$e^{-\beta V(x)}$$

$$\partial_x \left(\partial_x V p \right) - \beta \partial_{xx} p = 0$$

$V(x)$: a potential

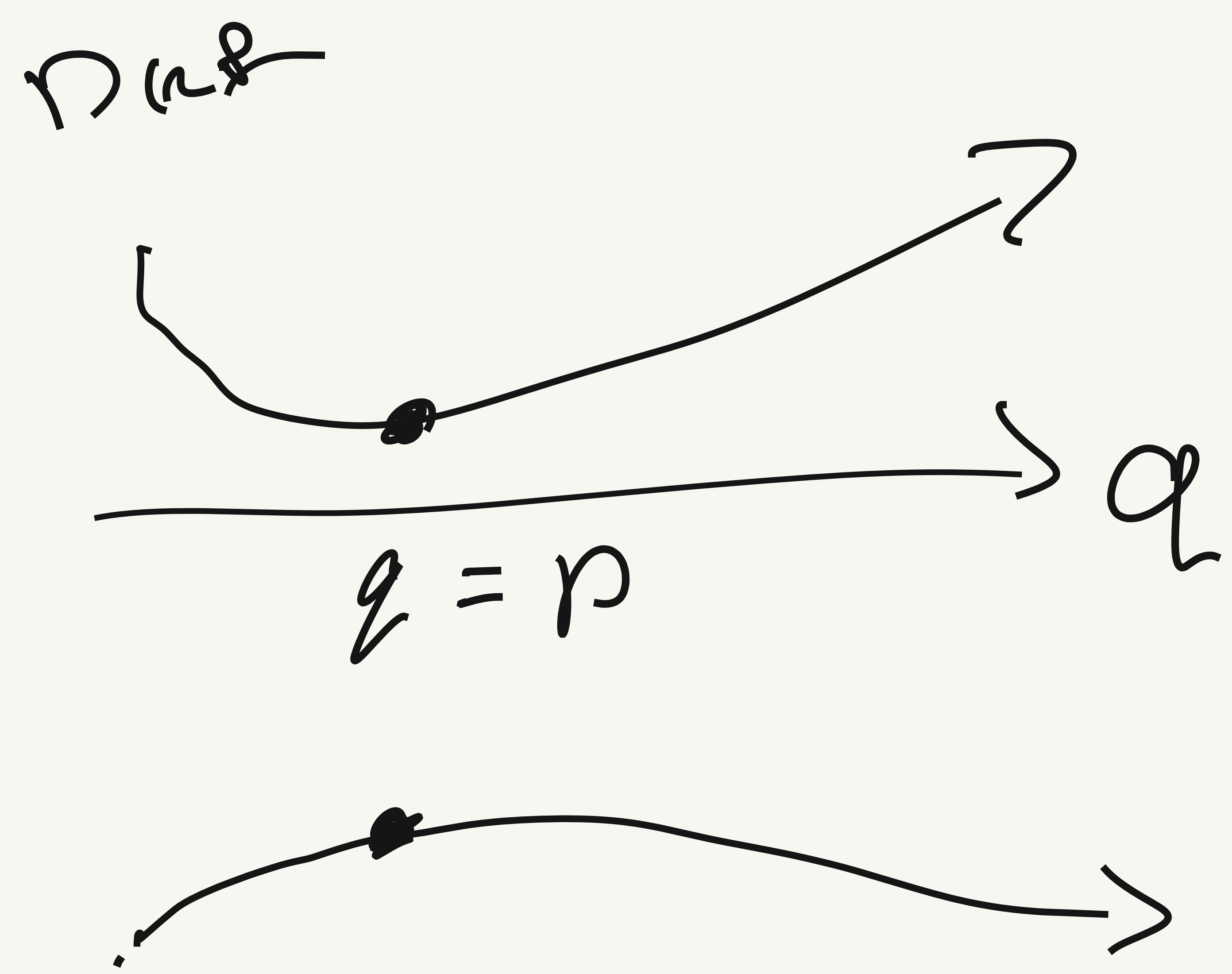
$$V(x) = \int^{(x)}$$

$$D_{KL}(q||p) = \mathbb{E}_q[\ln q] - \mathbb{E}_q[\ln p]$$

$$\arg\min_q (D_{KL}) \Rightarrow q = p$$

$$\longrightarrow \arg\max_q (-D_{KL}) \Rightarrow q = p$$

$$-\mathbb{E}_q \ln q + \mathbb{E}_q \ln p$$



$$-\int \ln q \; q + \int \ln p \; q - 0$$

$$\underbrace{-\int \ln q \; q}_{\substack{\text{entropy} \\ q}} - \underbrace{\int - \ln p \; q}_{E[-\ln p] = 0} + 0$$

$$-\int \ln q \; q \quad - \left(\int -\ln p(n \mid \mu, s)\, q + \ln p(\mu, s) \right)$$

$$\mathbb{E}\left[-\ln p(n \mid \mu, s) \right]$$

$$= -\ln p(\mu, s)$$

Thm. For an appropriate set of constraints,

FEP $\Leftrightarrow$ Max Ent.

Proof. $\arg\max_q (-D_{KL}) = \arg\min_q (D_{KL})$.

$$q^* = p$$

$$q^* = Z \cancel{\exp\{-(-\ln p)\}}$$

$$q^*_{i\nu} = p$$

Symmetry of Markov blanket

$$\text{a hidden assumption} \quad \begin{cases} \sigma(\bar{\mu}) = \bar{\eta}. \\ \sigma^{-1}(\bar{\eta}) = \bar{\mu}. \end{cases}$$

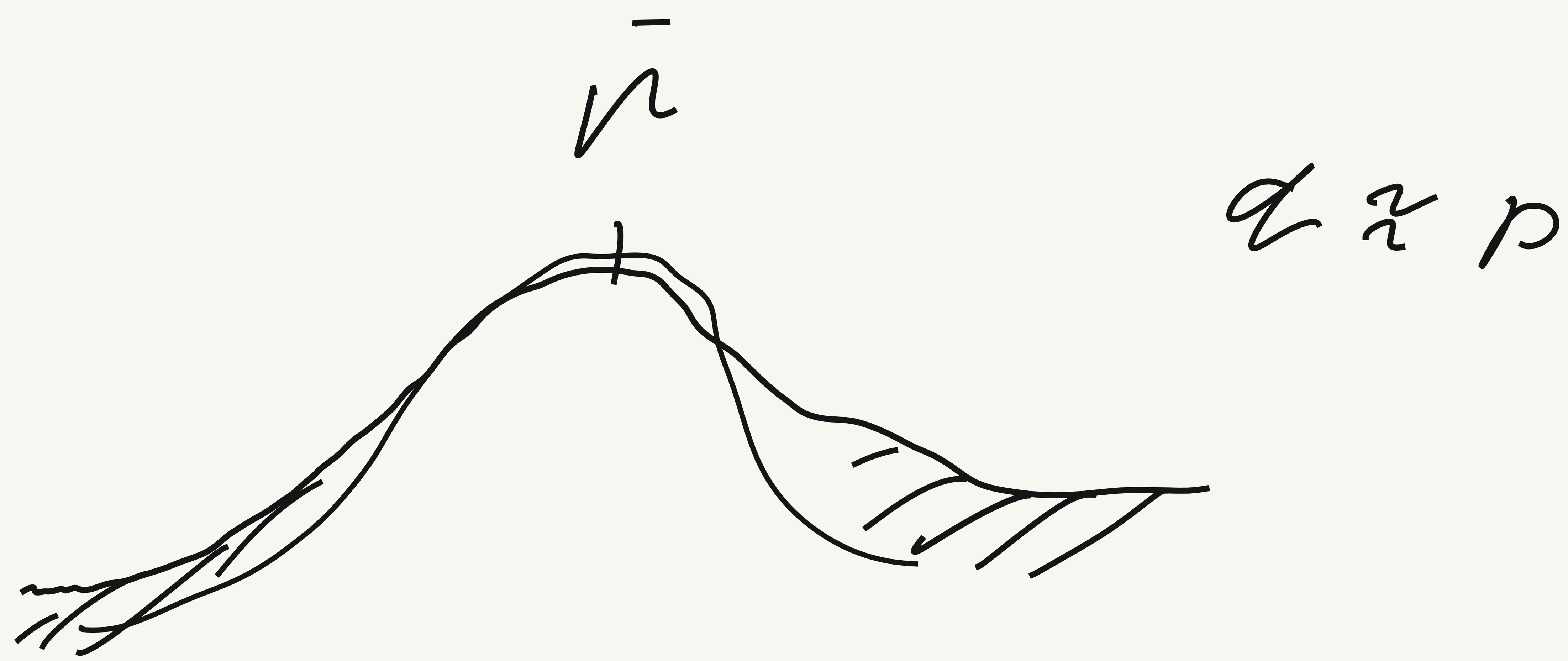
If $q = q(r; \sigma(\bar{\mu}))$ then q minimises D_{KL}.

Proof. Suppose p has $\bar{r}$ as a suff. stat.

then $p = Z e^{r}$ and $\mathbb{E}_p[r] = \bar{r}$.

Since $p = p(r; \bar{r})$ and $\sigma(\bar{\mu})$

$= \bar{r}$ $\quad q = p$



$J = \eta$ and $\mathbb{E}_{z}[\eta] = \sigma(\bar{\eta})$

$\eta \mapsto \mu \implies \mathbb{E}\,\mu = \sigma^{-1}(\bar{\eta})$

max (S) Subject to

μ model η

on avg.

Introducing constraints decreases entropy

BUT does not minimise it.

$J = x^2 + y^2$

2D Gaussian

$p \propto -J$

J : constraint

p = prob.

Indeed $J = x^2 + y^2$ $\quad \|\vec{x}\|$

$\mu = [x, y]$ $\quad q(\mu) = \exp\{-\lambda J\}$

$$e^{-x^2 - y^2} = \mathcal{N}(0, \lambda)$$

$p(\text{state}) \sim -$ Constraints

Via max entropy / min FE

Constraints on states of a system are like preferences about what states are likely to be occupied

→ which are like potential functions for sampling dynamics.

So what?

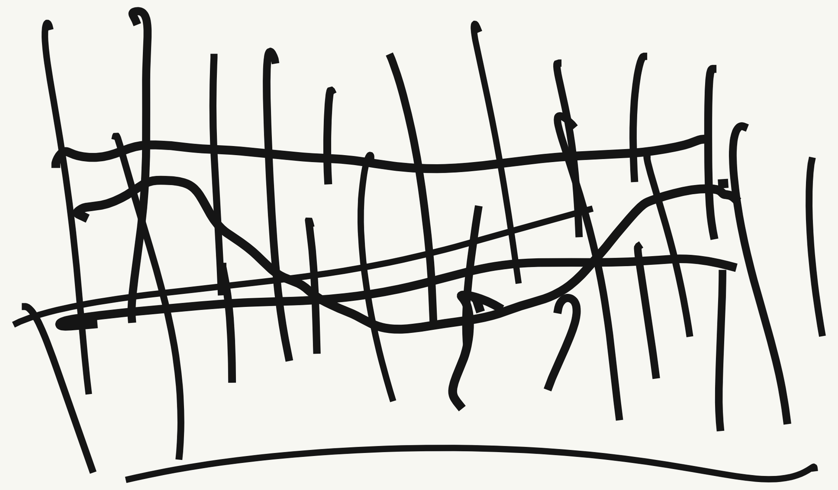
Potential functions are geometric features of a system

Example: general relativity

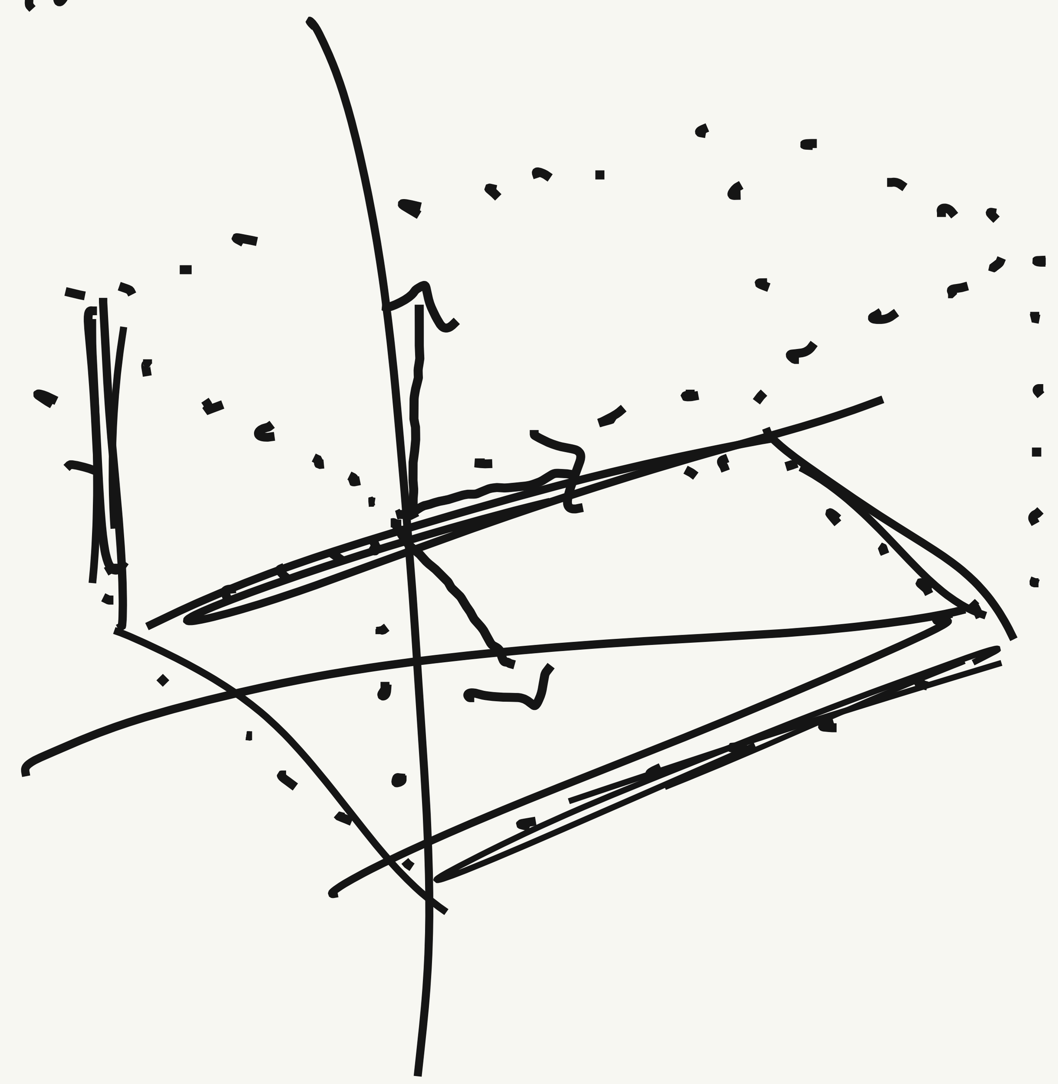
Some geometry

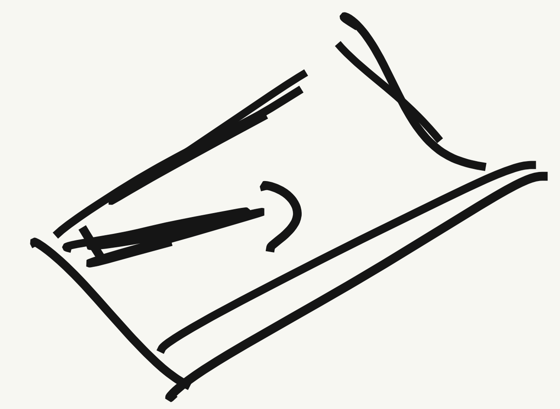
Equivalence principle

(Einstein)



$\mathbb{R}^2 + \mathbb{R} = \mathbb{R}^3$





$H_x T$

$V_x T$

dynamics of gauge

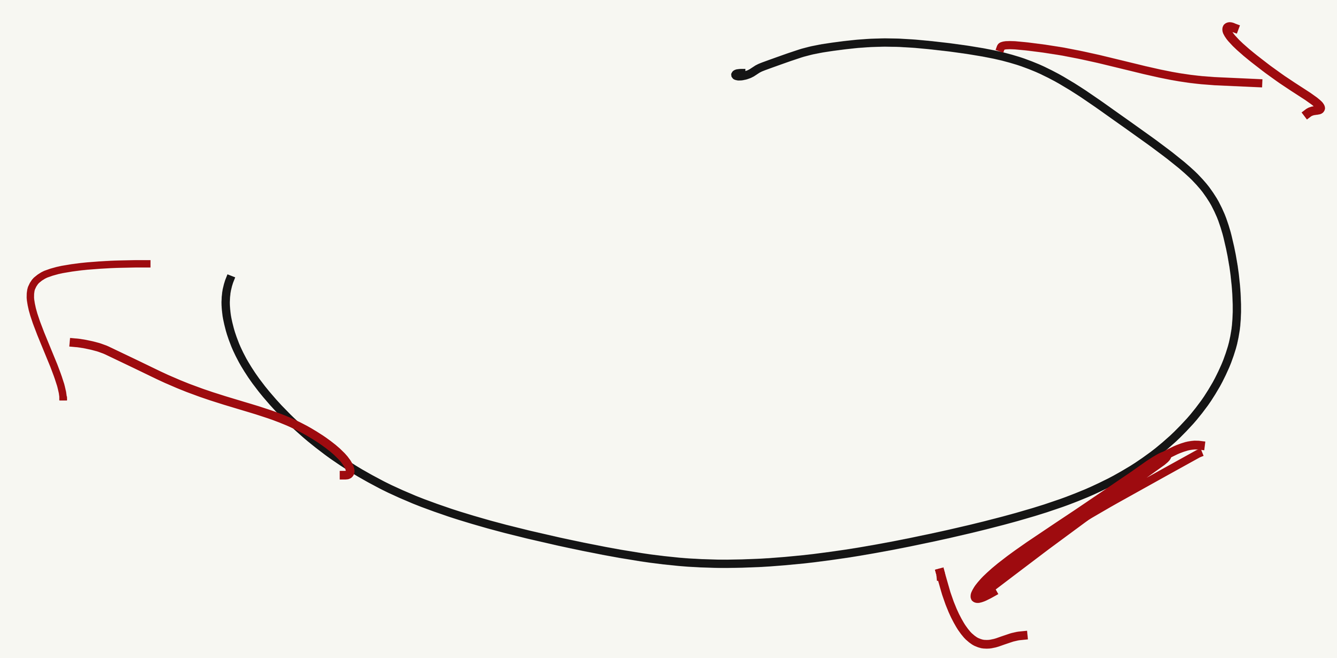
forced particles

split into particle dynamics and gauge force

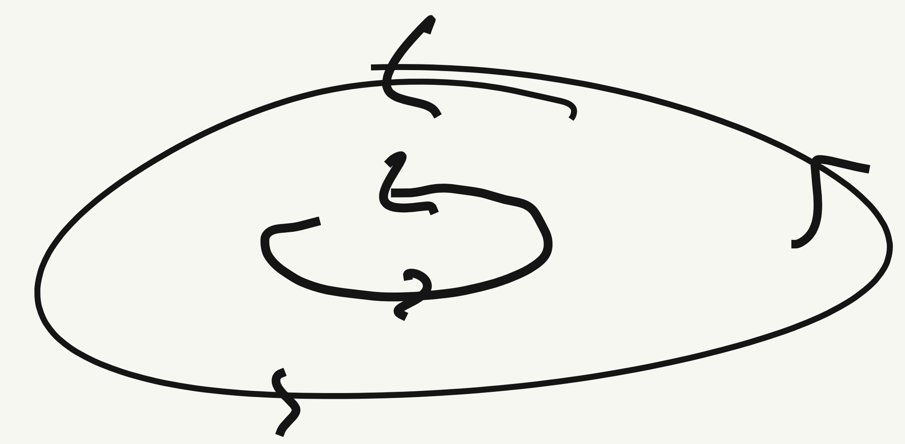
Spacetime

force (gravity)

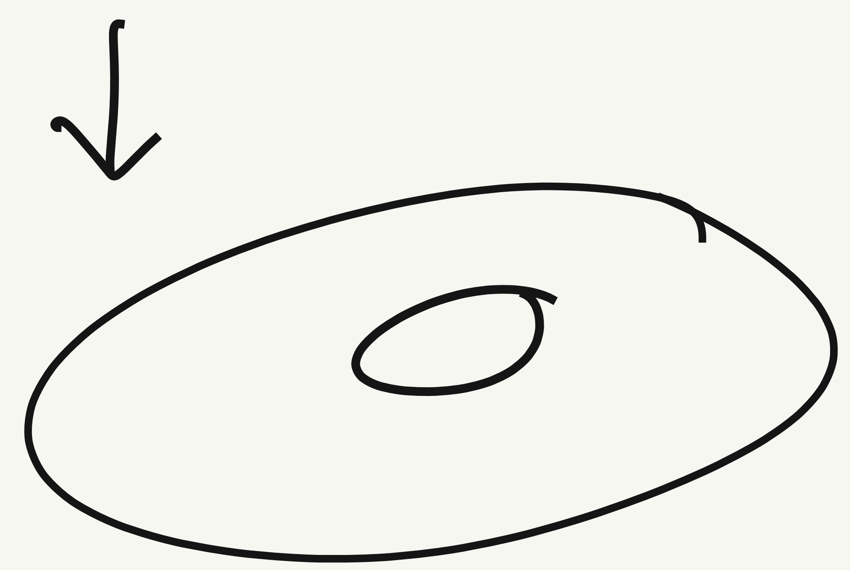
Example :

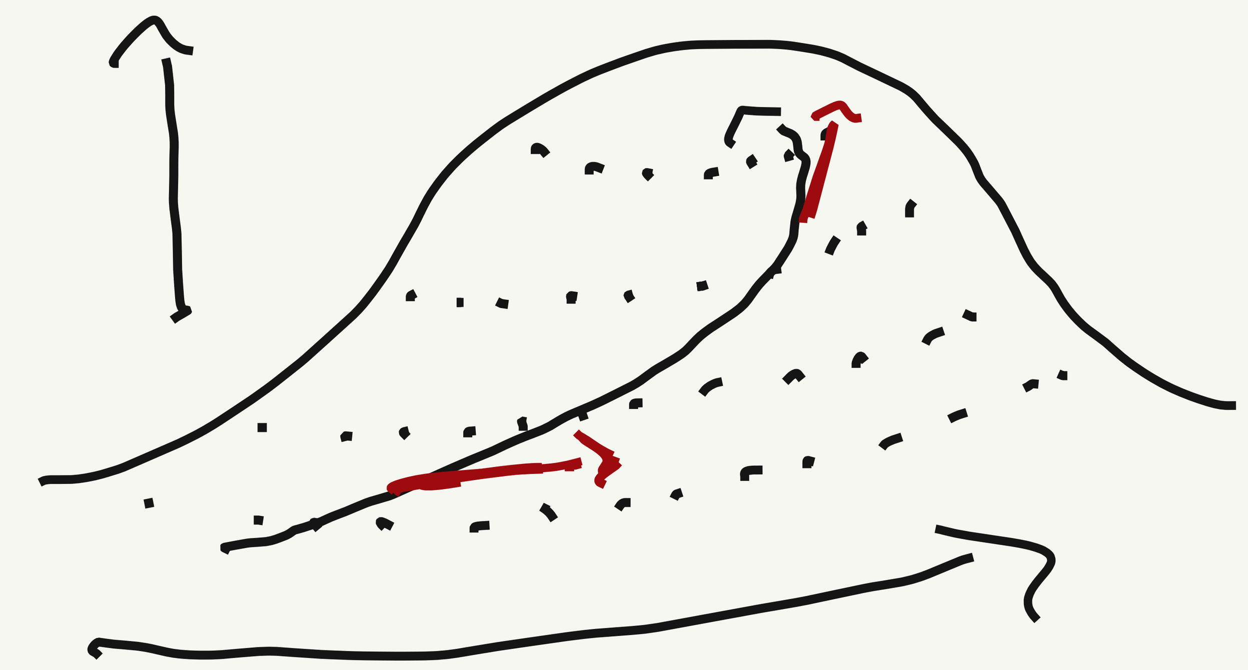


HP

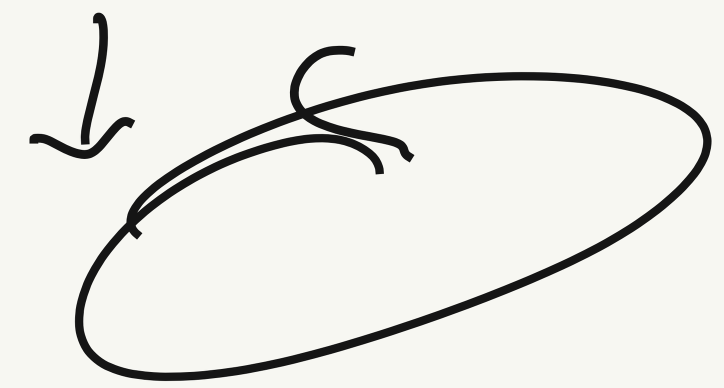


VP





$\mathbb{R}^3$



along isocont

$\Rightarrow$

max ent.

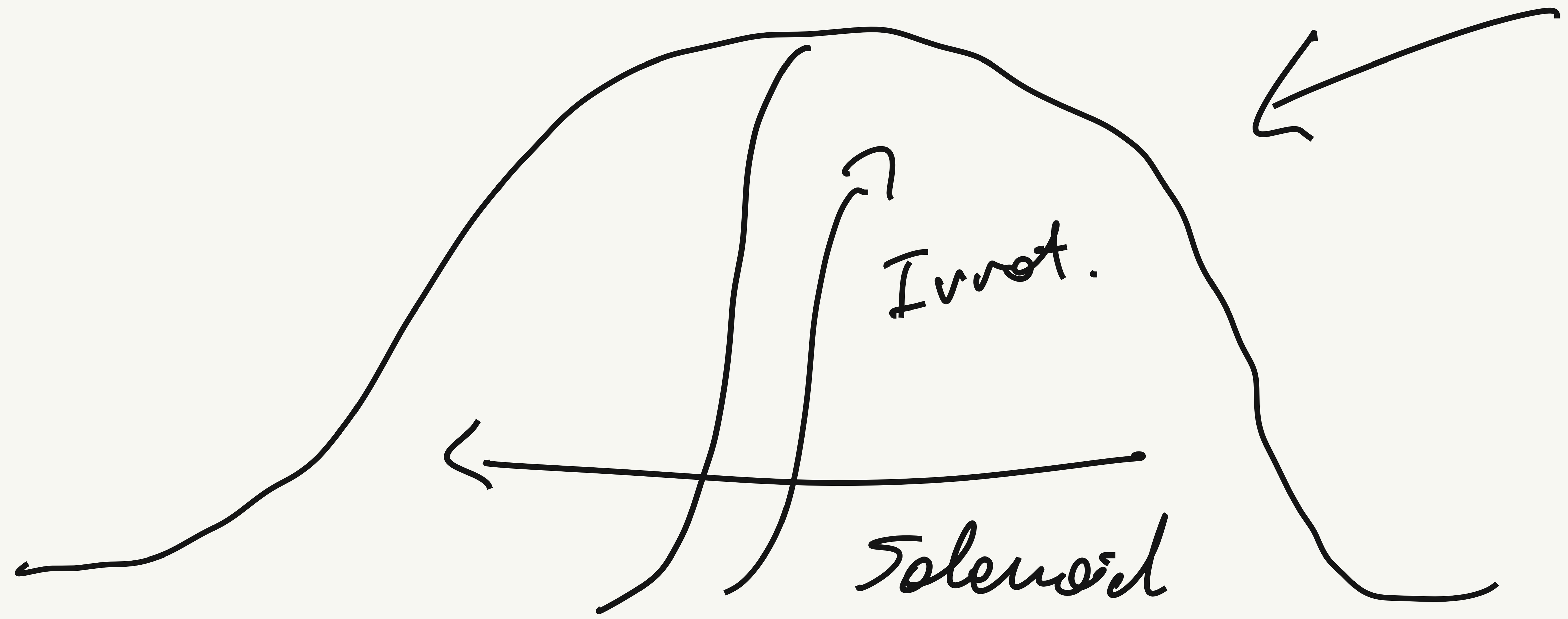
up

isocont.

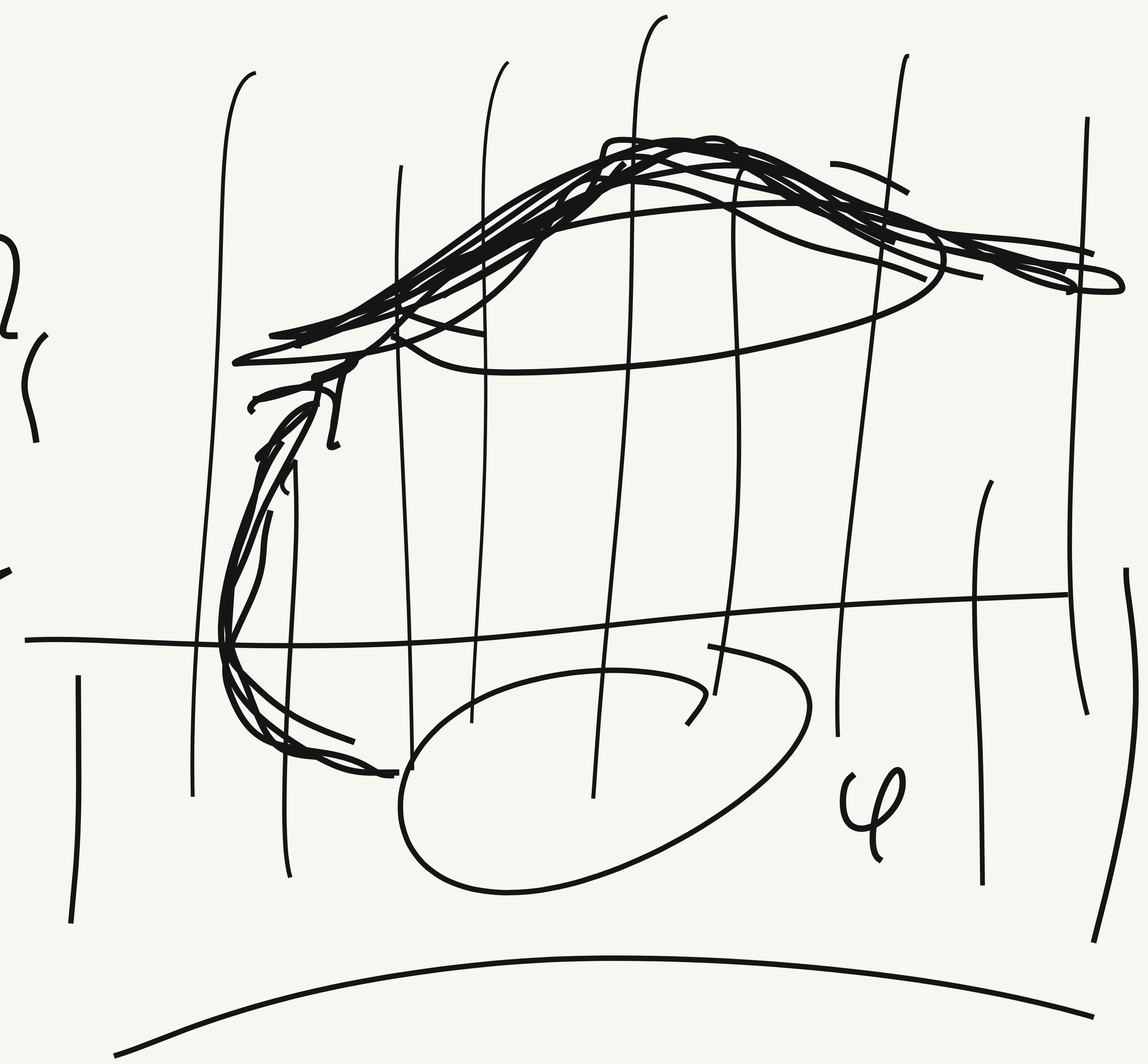
$\Rightarrow$

MAP

We have seen this before →



$\exp\{-\varphi\}$



$$\varphi = \int (x)$$

$$\exp\left\{ - \int (x) \right\}$$

$\ln p = -J$

$p = e^{-J}$

$-? = -J$

$\frac{\nabla p}{p} = \nabla \ln p$

Taking J as $-\ln\{p(\mu)\}$,

this flow is $-\nabla \ln\{p(\mu)\}$.

Proof. $\partial_{x_i} p(x_i) = -\partial_{x_i} J(x_i)\, p(x_i)$

$\nabla p(x_i) = -\nabla J(x_i)\, p(x_i)$

$\nabla \ln p = -\nabla J$

What we have shown:

- min FE $\iff$ max ent
- max ent $\iff$ potential on the state space

↳ Consequence:

flow towards a mode is a sort of gauge force.

Appendix: gauge symmetry

Frederic Schuller

Valery Rubakov

Leonardo A Lessa

"Notes on Gauge theories"

2019