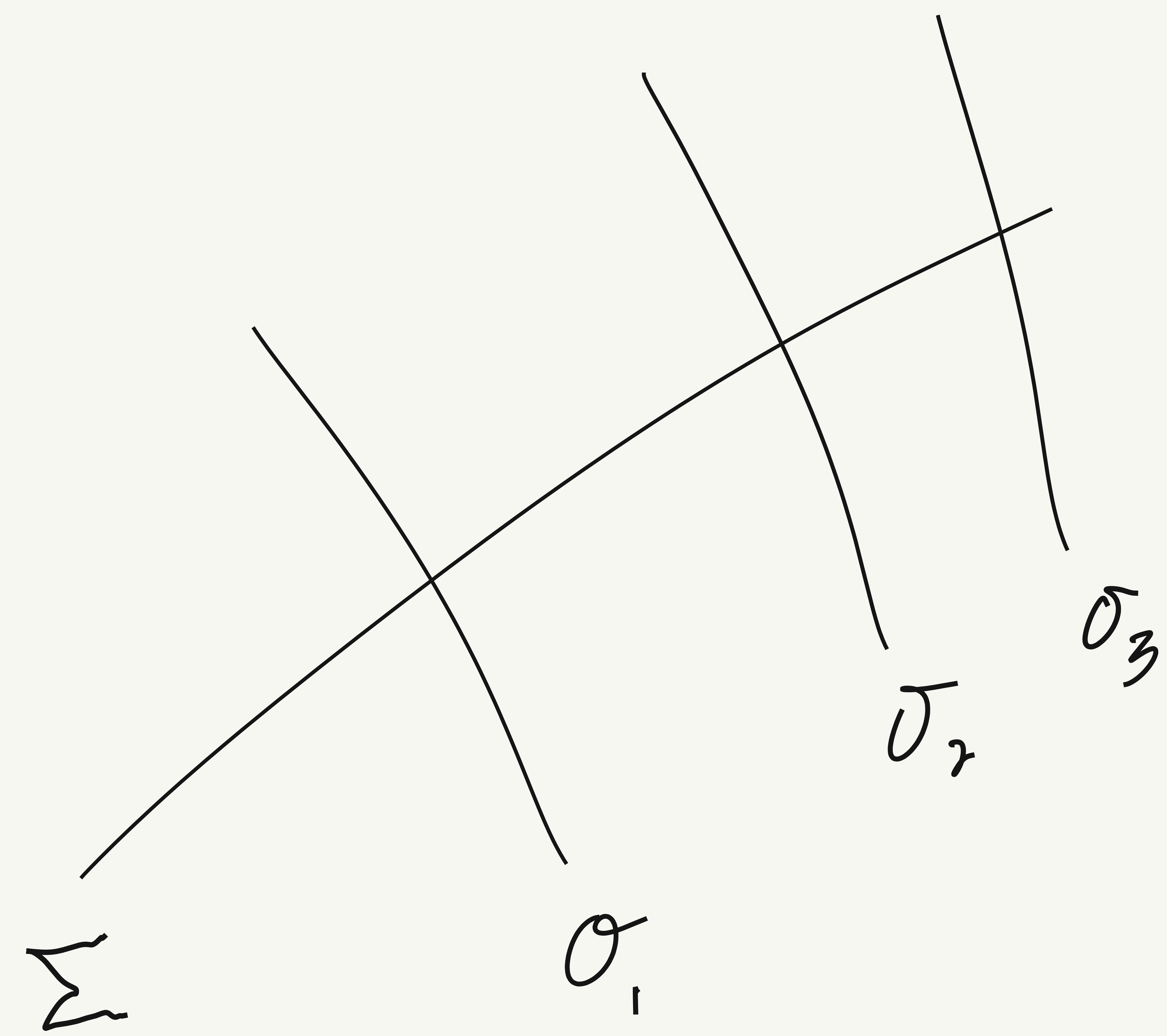
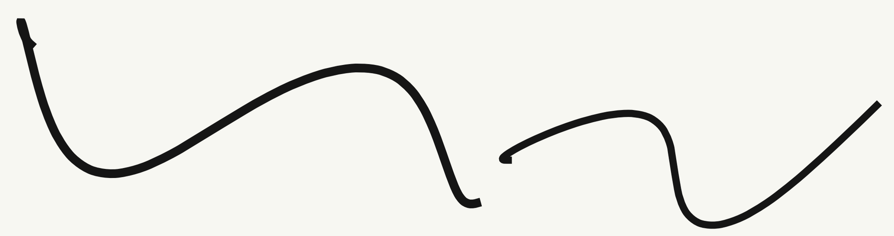
Solving Diffusion with Geometry and Max Cal

$$S[\gamma] = \int P(\gamma) \ln P(\gamma)$$



$$\dot{x} = f(x)$$



phase plane

(1) $\dot{x} = \vec{f}(x)$

$\gamma = \{ x(t) \dots x(t_f) \}$

$\forall x,\ \gamma(x) \nearrow f(x)$

$\nearrow$:= tangent to

(1) is solved by γ s.t. $\forall x,\ \gamma(x) \nearrow f(x)$

SDE

(2) $\dot{x} = f(x) + \eta(t)$ $\rightarrow$ $\int^t \eta \, ds$

$\equiv$

W_t s.t.

$\underbrace{W_s - W_t}$

γ now changes, given by a particular realisation of $\eta(t)$

noise term yields $P(\gamma)$

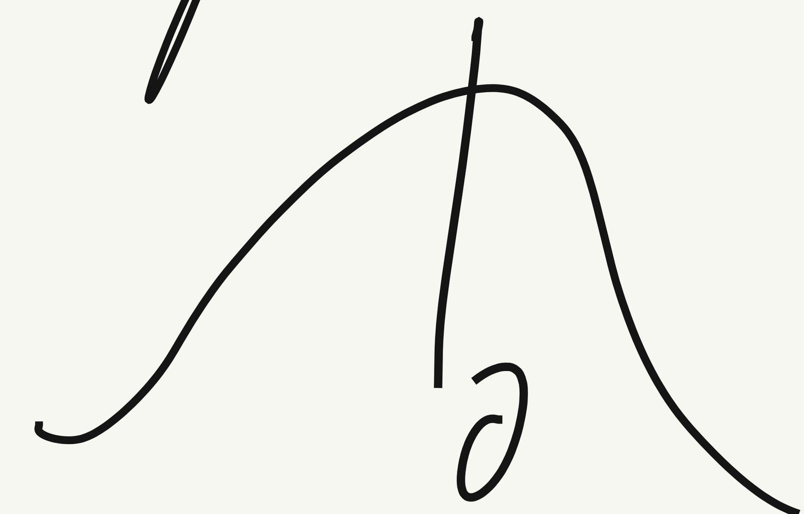
with each $\gamma \in \Gamma$.

$\Gamma :=$ set of possible paths

P ~~from~~ FP

$FP \Rightarrow Pr(\gamma)$

$\eta(t)$



Markovian expansion of Master equation Cf. Kramers-Moyal → FP =

Possibly non-linear

$$\frac{\partial \rho}{\partial t} = \alpha \frac{\partial}{\partial x} \rho(x,t) + \dots$$

$$\Rightarrow L^{\dagger} \rho = \dot{\rho}$$

with adjoint elliptic operator L^{+} encoding Parabolic partial derivatives.

Solve FP

$$\Rightarrow P(y)$$

$$\min \left(\int L \, dt \right)$$

S

$\delta S = 0$

Variational Euler - Lagrange from LAP

$$\delta S = 0 \Rightarrow \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0$$

$$m \ddot{x} = F$$

yields 'geodesic' curve that follows force acting on it

$$\gamma(t) = \{ x(t) \ldots x(t_f) \}$$

X

for some 'observable,' the 'output' X

R Feynman has a great lecture on this

$$\max\left(-\int P(x)\ln(P(x))\,dx\right)$$

Maximise entropy for model selection

$$y \Rightarrow \max\left(-\int P(y)\ln(P(y))\,dy\right)$$

$$\mathcal{J} \Rightarrow \max\left(-\int P(y\mid \mathcal{J})\ln(P(y\mid\mathcal{J}))\,dy\right)$$

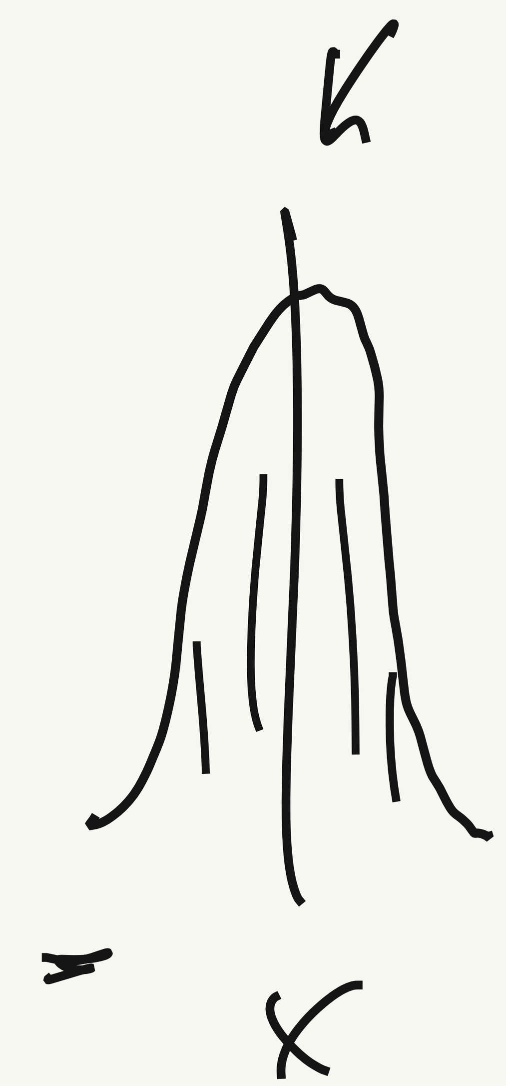
for $\mathcal{J}$ = constraints

$$-\int P(\gamma|J)\,\ln\left(\frac{P(\gamma|J)}{P(\gamma)}\right) = S[\gamma]$$

equivalent to

$\Rightarrow$ relative entropy between $P(\gamma)$ and $P(\gamma|J)$ $E[\gamma]$

$$\operatorname{argmax}\left(S[\gamma]\right) = \underline{\underline{P(\gamma)}}$$



Corresponds to $P(\gamma \in \Gamma)$ $\longleftarrow$ Inference is a density over possible / likely paths

$$S[y] = \int P(y|J) \ln\left(P(y|J)\right) dy$$

$$S'[y] = S[y] + \int [\lambda \cdot J] + \lambda_1 J_1 + \lambda_2 J_2 \; \ldots$$

Constraints become Lagrange multipliers

$$\lambda \cdot \int J(y) \, P(y) \, dy - J$$

take form of integral expectation
~ expected value

$$\frac{\partial S[y]}{\partial P(y)} = 0$$

$$\Rightarrow \ln(y) + \lambda \cdot J = 0$$

$$\Rightarrow y = \exp\{-\lambda \cdot J\}$$

Can plug this in to verify

Example – Boltzmann-Gibbs distribution is maximum entropy

$$\lambda \int J(y)\, P(y)\, dy$$

distribution given $J = \frac{E}{T}$ as a constraint.

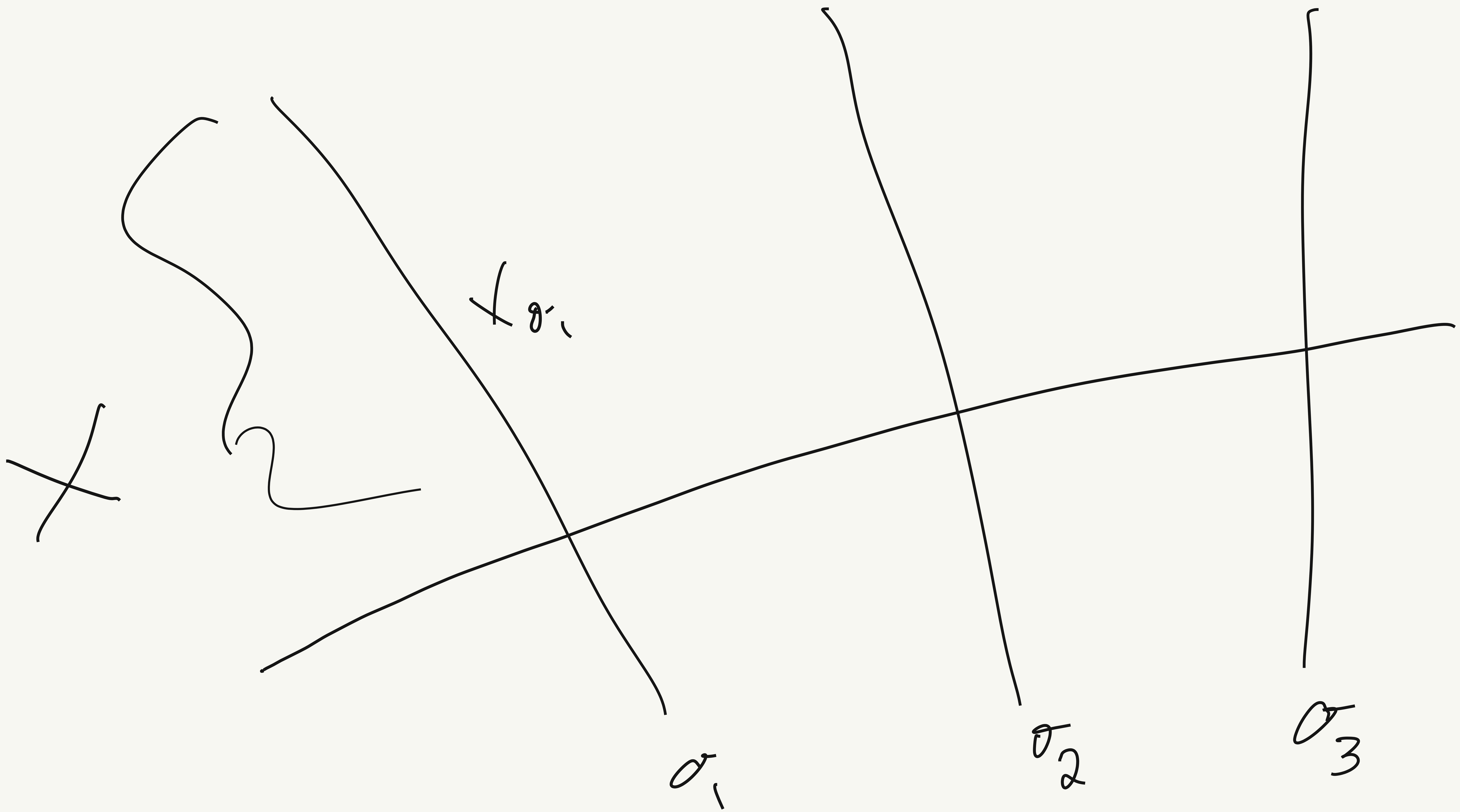
$$J = \frac{E}{T}$$

Boltzmann's constant

$$\lambda^{-1} = k_B$$

$$\ln(y) + \lambda J \implies y = \exp\{-\lambda J\}$$

$$\exp\left\{ -\frac{1}{k_B} \frac{E}{T} \right\}$$



A 'hairbrush' picture of internal field structure as a fibre bundle

Σ is $\mathbb{R}^2$

denote $\sigma \in \Sigma$ as base points

$X \; \bigsqcup_{\sigma} X_{\sigma_1}$

total space = bundle of fibres = disj. union of fibres

'Hair brush'

denote $x \in X$ points in fibres

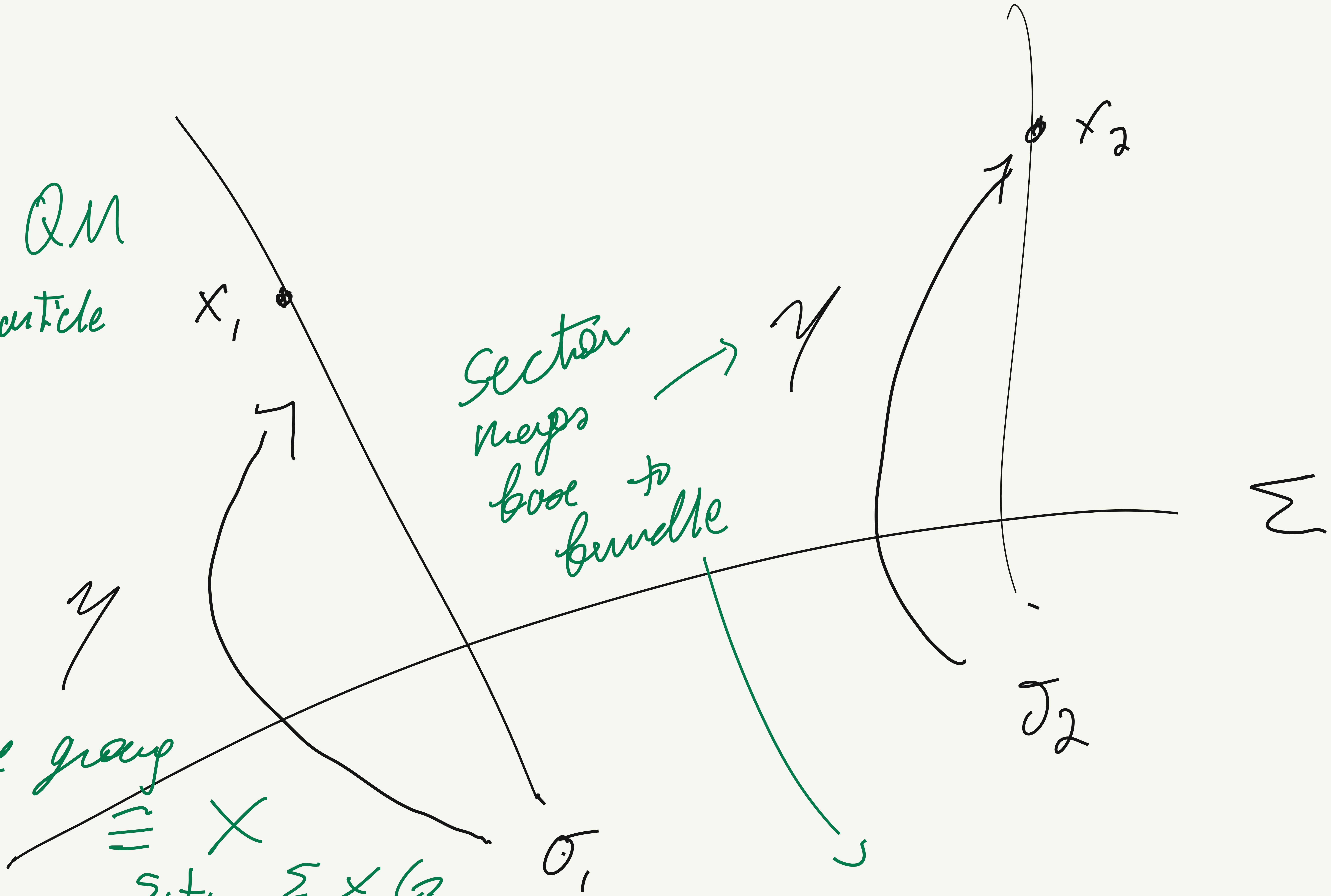
example: phase of a QM particle

$U(1) = S^1$

$e^{i\theta} \in U(1)$

$g \in G$, Lie group

γ



$\cong X$

s.t. $\Sigma \times G = X$

$\gamma : \Sigma \to X$

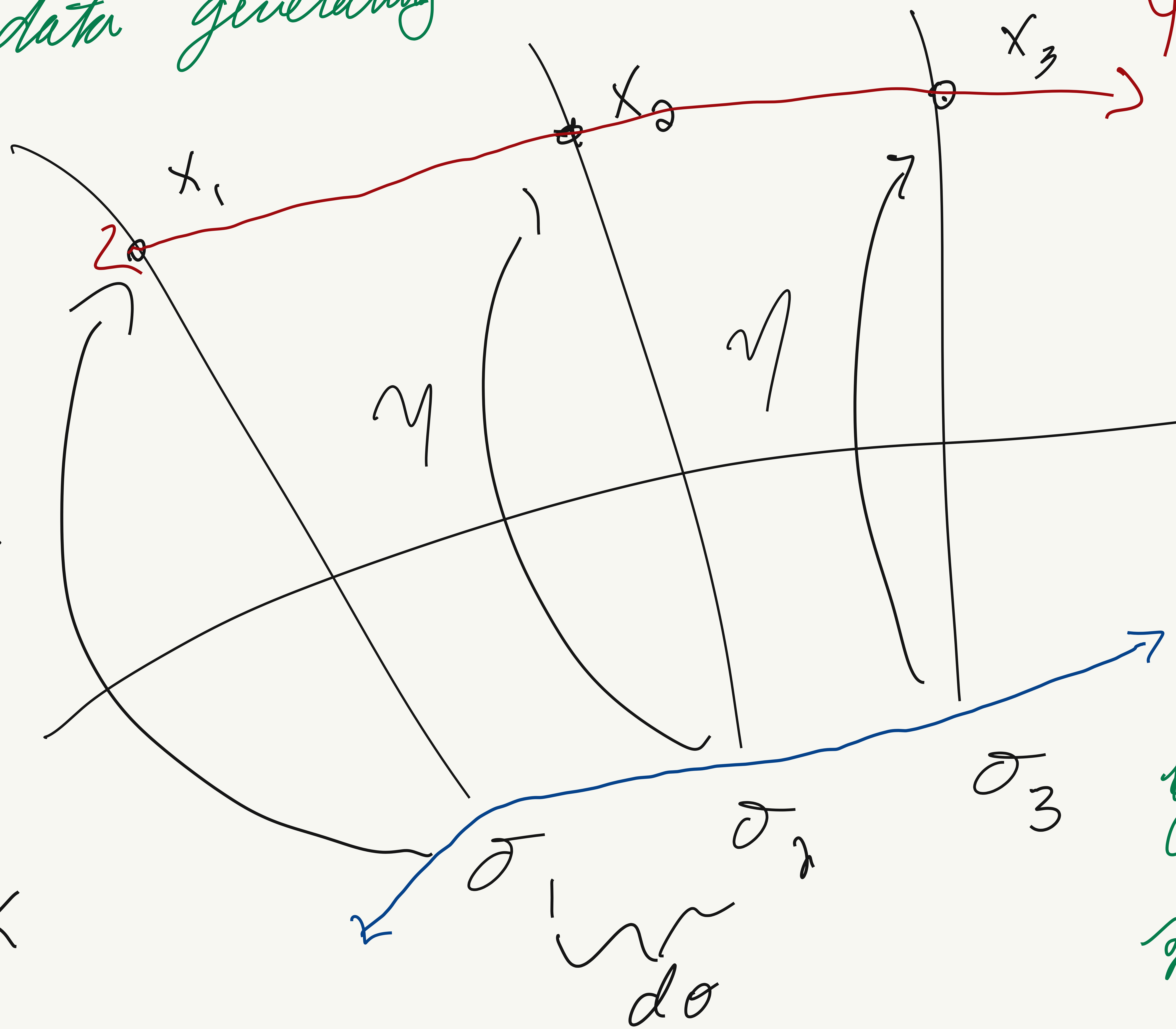
$\psi(x) =$ 

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

$\gamma : \Sigma \to X$

$\gamma(\sigma) = x$

From Propositions 1 and 2, η is a data generating process, which can be learned



Prop 1.

$\tilde{\varphi}$ ← Lifted curve, 'data'

$\eta : \Sigma \to X$

$\varphi : I \to \Sigma$

$[0,1]$

$\mathbb{R}^2$

$\eta \circ \varphi$

$\eta(\varphi)$

$I \to \Sigma$

$\Sigma \to X$

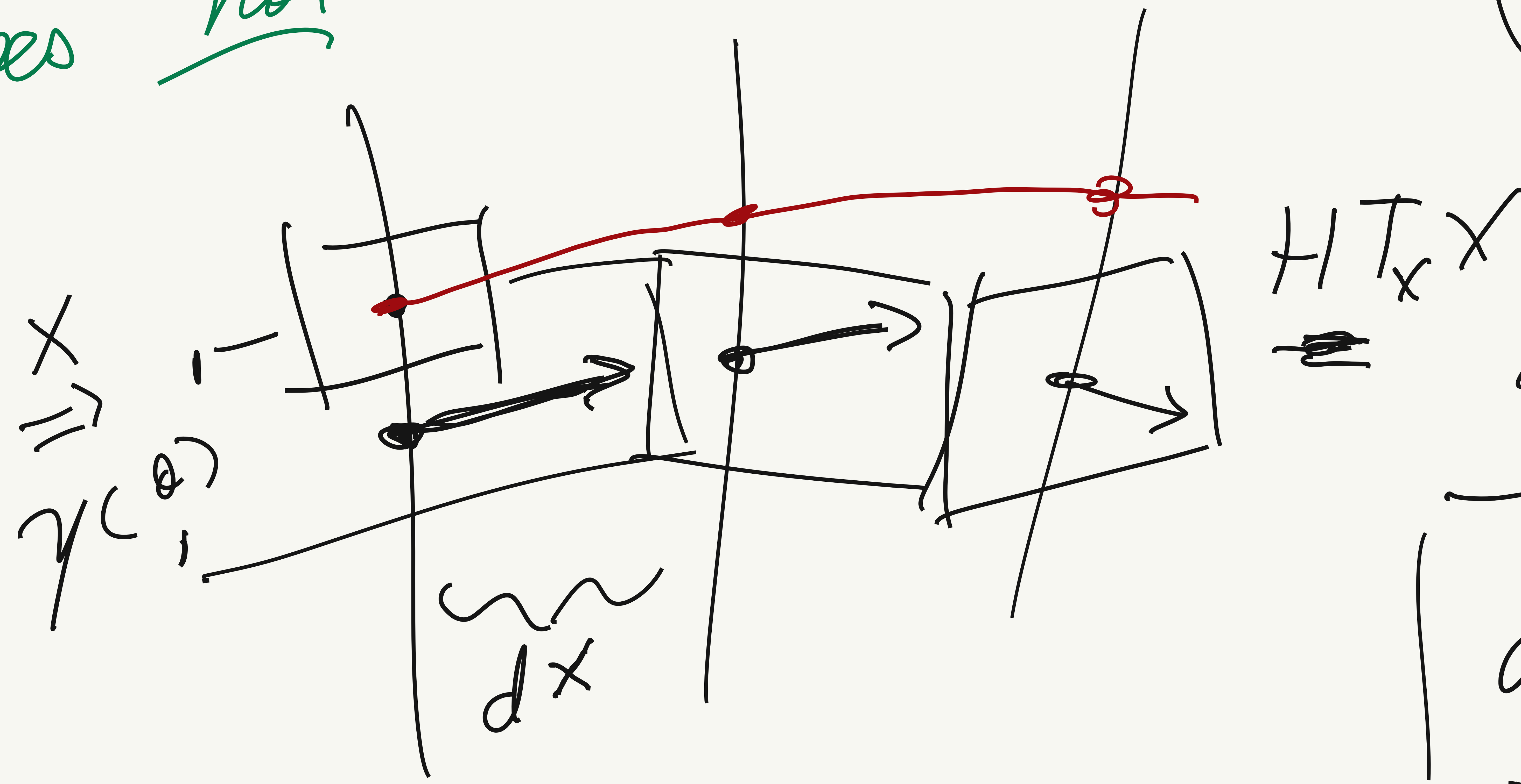
$\eta(\varphi) = \tilde{\varphi}$

φ ← base curve, spatial / temporal input

$\tilde{\varphi} = I \to X$

Prop 2.

Calculus does not work on bundles — necessitates connection



(3)

$$\gamma(\theta) = X$$

$$\gamma \circ \varphi = \tilde{\varphi}$$

$$\boxed{d\gamma : T\Sigma \longrightarrow T_{\gamma(\theta)}X}$$

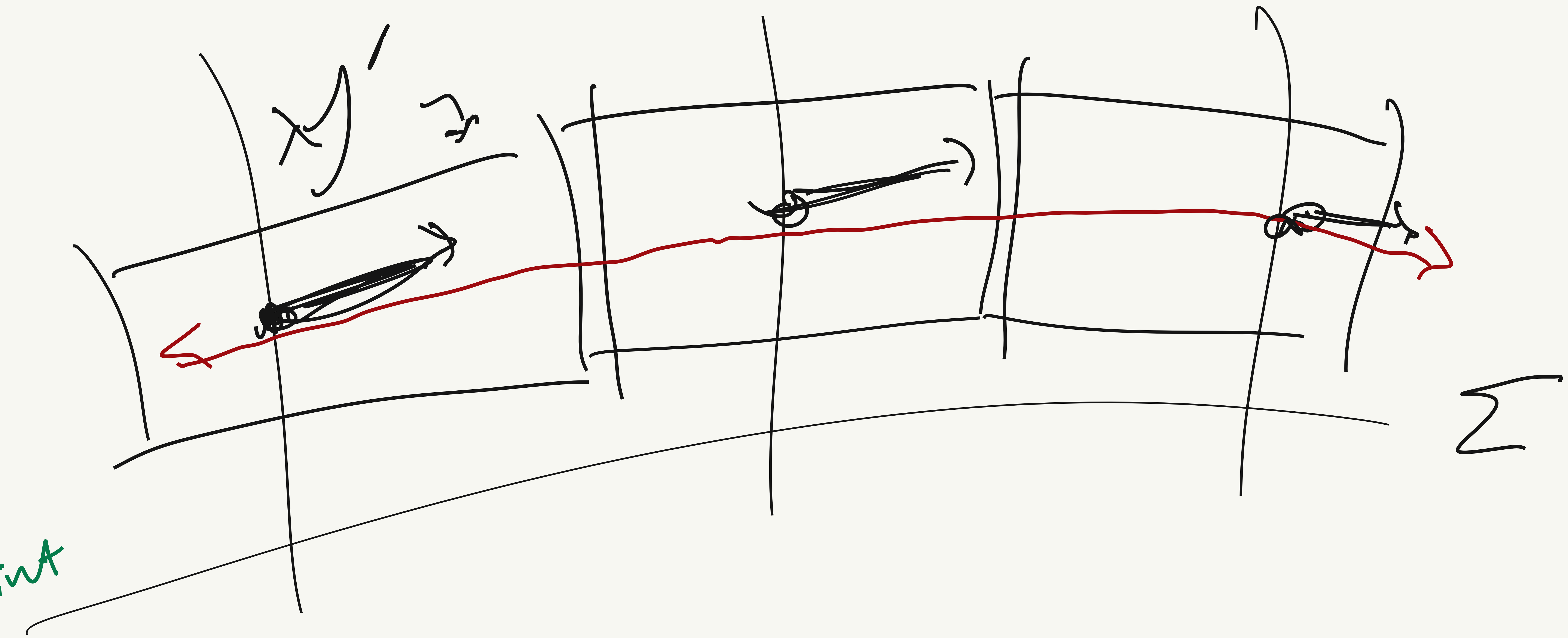
$$\tilde{\varphi}'(x)$$

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

$$dx$$

$$v \in T_\theta \Sigma \quad \forall \theta$$

Connection serves to constrain $\tilde{\varphi}'$ $\therefore$ Connection $\Leftrightarrow$ Constraint



$\nabla := HT_xX$

$d\gamma \mapsto T_xX$

$\tilde{\varphi}' = \gamma \circ \varphi$

= Observable

= unknown data generating process $\gamma \mid \gamma$

$\nabla := HT_xX$ is precisely a choice of tangent space

Theorem of Lag. Mult. $\longrightarrow$ $\tilde{\varphi}'(x) = -\lambda\gamma'(x)$ $\longleftarrow$ specifies possible tangent vectors *

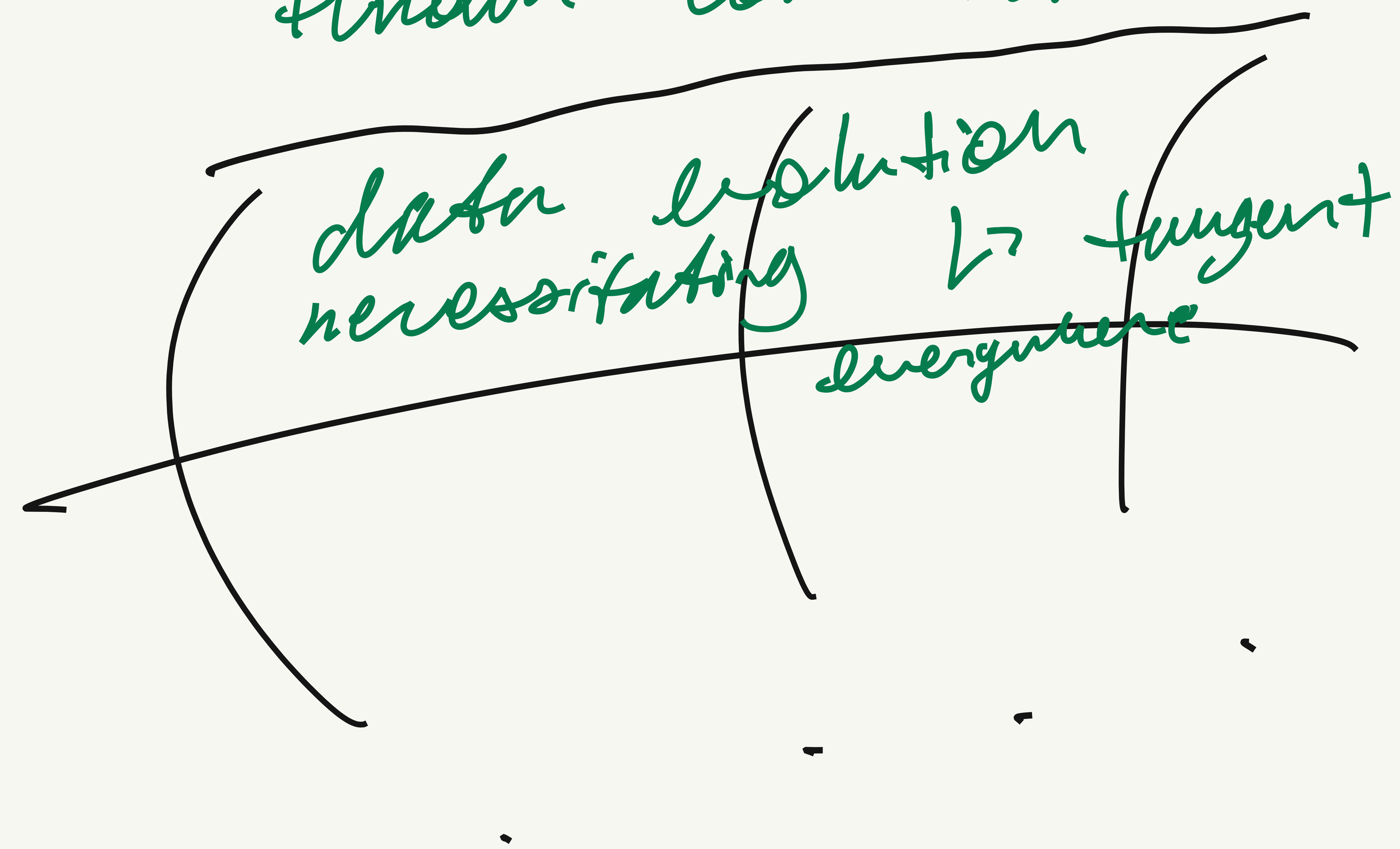
$f(x_c)' + \lambda g(x_c) = 0$

$x = \gamma = \{x(t) \dots x(t_f)\}$ * for all points x in the lift

$$-\lambda\gamma' = \tilde{\varphi}'\,\nabla\,\tilde{\varphi}$$

from $\tilde{\varphi}' + \lambda\gamma' = 0$

Constraints come from known constraints on



$$-\lambda\gamma' \Rightarrow D_{\varphi'}\gamma = 0$$

$$U_\alpha \Rightarrow \gamma_\alpha^{*}\nabla \Rightarrow A_\alpha \qquad A = \sum_\alpha A_\alpha\, du^\alpha$$

$u \in U_\alpha$

$$\nabla(\gamma)\,(\gamma : \Sigma \to X) \mapsto X \xrightarrow{\nabla} A$$

Here we define Parallel transportation in the bundle <u>AND</u> base tra(γ) $\Rightarrow$ an ODE in terms of γ

$$\gamma'_\alpha = \underbrace{-A_\alpha(\varphi)}_{\gamma^* \nabla} \gamma_\alpha$$

Solved by this $\rightarrow$

$$\gamma = \exp\Big\{ -\int_\Sigma \underbrace{A_\alpha(\varphi)}_{\gamma^* \nabla} \Big\}$$

or this for connection = constraint

$$\gamma = \exp\Big\{ -\int_\Sigma \lambda\, y' \Big\}$$

Implication

- If we suppose that a path obeys its geometry, identically, is determined by the potential acting on it,

$$\gamma = \exp\{-\lambda \gamma\}$$

$$\Longleftarrow \quad \mathrm{tra}(\gamma)$$

By some category (4) theory or cohomology, this is also the fibre-wise integration of the system

$$\int e^{S_{\mathrm{kin}}} \circ \mathrm{tra}(\gamma)\, d\gamma$$

$$\in \Omega(\mathrm{Hom}(\gamma))$$

$$d\gamma \wedge \mathrm{tra}(\gamma)$$

Constraints are (geometrically) potential which we obey when max(entropy) (this implies LAP, in turn, ∥-transport)...

A concrete example in reference to K A Dill's

$$J = -\int P(y)\ln(P(y))\,dy - \lambda\left(\int P(y)\,dy - 1\right) - \mu\left(\int C(y)P(y)\,dy - C\right) - \nu\left(\int W(y)P(y)\,dy - \Sigma\right) - \rho\left(\int H(y)P(y)\,dy + \Sigma\right)$$

$$J_h \geq F(y)_h$$

$E(J(y))$

'A minimal set of constraints for NESS systems'

$$0 = \frac{\partial S}{P(y)} = \ln(P(y)) - \lambda \cdot y$$

$$e^{(-\lambda \cdot y)}$$

$$\langle y \rangle = y$$

Constraints ought to be written as averages for computational reasons