

1. Motivation
2. Physics from Fisher Information?
3. Interpretive Issues

1. Physics From Fisher Information?

1. Motivation

- Recall: The *Fisher Information* $I_{\text{Fish}}(p; \theta)$ of a probability distribution $p(x; \theta)$ is given by:

$$I_{\text{Fish}}(p; \theta) = \int \left(\frac{\partial \log p}{\partial \theta} \right)^2 p dx$$

$$= - \int \left(\frac{\partial^2}{\partial \theta^2} \log p \right) p dx$$

- I_{Fish} is the average value of the sensitivity $-\frac{\partial^2}{\partial \theta^2} \log p(x; \theta)$ of a probability distribution $p(x; \theta)$ to θ .
- How spread out p is with respect to θ .
- The p that maximizes I_{Fish} has the greatest sensitivity/least spread.

- Now assume "shift invariance":

$$p(x) = p(x + \theta)$$

← - p remains unchanged under a shift transformation: $x \mapsto x' = x + \theta$

- So $\frac{\partial}{\partial \theta} = \frac{\partial}{\partial x} \frac{\partial x}{\partial \theta} = - \frac{\partial}{\partial x}$

- Then:

$$I_{\text{Fish}}(p) = \int \left(- \frac{\partial \log p}{\partial x} \right)^2 p dx = \int \frac{(\partial p / \partial x)^2}{p} dx$$

← $\left(\frac{\partial \log p}{\partial x} \right) = \frac{1}{p} \frac{\partial p}{\partial x}$

- Now assume: $p(x) = q^2(x)$, for some $q(x)$.

- Then:

$$I_{\text{Fish}}(q) = \int \frac{(\partial q^2 / \partial x)^2}{q^2} dx = \int \frac{(2q \partial q / \partial x)^2}{q^2} dx = 4 \int \left(\frac{\partial q}{\partial x} \right)^2 dx$$



I_{Fish} contains the square of a derivative term.

- What this entails:

The function $q(x)$ for which I_{Fish} takes a minimum or maximum (*i.e.*, "extremal") value is the solution to a 2nd-order partial differential equation.

- Why is this supposed to be important? The equations of motion of most theories in physics are 2nd-order partial differential equations.
- So (Friedan 1998): Perhaps the equations of motion (of most theories in physics) can be derived by an appropriate reconstrual of Fisher information.

Let's see how this is supposed to work...

2. Physics From Fisher Information?

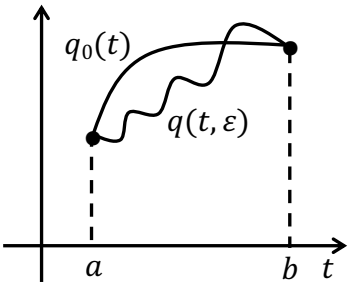
Task: Given a functional in the form of a definite integral

$$S(q) = \int_a^b L(t, q(t), q'(t))dt, \quad q'(t) \equiv \frac{\partial q(t)}{\partial t}$$

find the particular function $q_0(t)$ for which S take an *extremal value*.

- Consider small variations $q(t, \varepsilon)$ of $q_0(t)$ between the endpoints a, b :

$$q(t, \varepsilon) \equiv q_0(t) + \varepsilon \eta(t), \text{ where } \eta(a) = \eta(b) = 0$$
- $S(q(t, \varepsilon))$ takes an extremal value where it's slope is zero:



$$\begin{aligned}
 0 &= \frac{\partial S}{\partial \varepsilon} = \int_a^b \left(\frac{\partial L}{\partial q} \frac{\partial q}{\partial \varepsilon} + \frac{\partial L}{\partial q'} \frac{\partial q'}{\partial \varepsilon} \right) dt \quad \leftarrow \text{or: } \delta S \equiv \frac{\partial S}{\partial \varepsilon} d\varepsilon = 0 \\
 &= \int_a^b \left(\frac{\partial L}{\partial q} \frac{\partial q}{\partial \varepsilon} - \frac{d}{dt} \left(\frac{\partial L}{\partial q'} \right) \frac{\partial q}{\partial \varepsilon} \right) dt + \cancel{\frac{\partial L}{\partial q'} \frac{\partial q}{\partial \varepsilon}} \bigg|_b^a \\
 &= \int_a^b \left(\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial q'} \right) \right) \eta(t) dt
 \end{aligned}$$

Integration by parts:

$$\int_a^b u dv = uv \Big|_a^b - \int_a^b v du$$

$$u = \frac{\partial L}{\partial q'}, dv = \frac{\partial q'}{\partial \varepsilon} dt$$

- So:

$$\frac{\partial L}{\partial q} - \frac{d}{dt} \left(\frac{\partial L}{\partial q'} \right) = 0$$

Euler-Lagrange Equation

A solution $q(t)$ to the E-L equation is a function that extremizes S !

Task completed!

The function $L(t, q, q')$ is called the *Lagrangian*.

The functional $S(q) = \int_a^b L(t, q(t), q'(t)) dt$ is called the *action*.

Ex 1: What $q(t)$ extremizes $I_{\text{Fish}}(q) = 4 \int q'^2 dt$?

- Note: $L_{\text{Fish}}(t, q, q') = 4q'^2$

- So:
$$\frac{\partial L_{\text{Fish}}}{\partial q} - \frac{d}{dt} \left(\frac{\partial L_{\text{Fish}}}{\partial q'} \right) = -\frac{d}{dt} 2q' = -8 \frac{d^2 q(t)}{dt^2} = 0$$

↪ In the absence of interactions ($V = 0$), L_{Newt} is identical (up to a factor) to L_{Fish} .

Ex 2: What $q(t)$ extremizes $S_{\text{Newt}}(q) = \int \left[\frac{1}{2} m q'^2 - V(q) \right] dt$?

- Note: $L_{\text{Newt}}(t, q, q') = \frac{1}{2} m q'^2 - V(q)$

- So:
$$\frac{\partial L_{\text{Newt}}}{\partial q} - \frac{d}{dt} \left(\frac{\partial L_{\text{Newt}}}{\partial q'} \right) = -\frac{\partial V}{\partial q} - m \frac{d}{dt} q' = 0$$

- Or:
$$m \frac{d^2 q(t)}{dt^2} = -\frac{\partial V}{\partial q}$$

↪ Newton's 2nd order law of motion for the path $q(t)$ of a particle with mass m interacting with a potential $V(q)$!

Two Observations:

- (1) The fundamental equations of motion in physics can be derived by extremizing an action S with an appropriately formulated Lagrangian.
 - *Classical field theories: Newton's 2nd law, Maxwell's equations, Einstein equations.*
 - *Quantum mechanics: Schrödinger and Dirac equations.*
 - *Quantum field theories: equations of motion for QED, QCD, Standard Model.*
 - (2) In the absence of interactions, all of these Lagrangians are schematically identical to $L_{\text{Fish}} \sim q'^2$.
- Suggests: In the absence of interactions, the fundamental equations of motion in physics are the result of extremizing Fisher information; *i.e.*, requiring that $\delta I_{\text{Fish}} = 0$.

What could this mean??

But first: What **about** interactions?

What Frieden Has to Say (1998, pp. 70-72)

"Physical Information"

$$K(q) = I_{\text{Fish}}(q) - J(q)$$

- $I_{\text{Fish}}(q)$ = "Intrinsic" information of a physical system obtained during a measurement.
- $J(q)$ = "Bound" information of system prior to measurement.

"Axiom 1"

$$\delta I_{\text{Fish}} = \delta J$$

- "Conservation of (Fisher) information change". (During a measurement, small variations in I_{Fish} equal small variations in J .)

"Variational Principle"

$$\delta K = 0$$

- When changes in information are conserved during measurements (?), the physical information obtains an extremal value (pg. 71).

"Zero-condition"

$$I_{\text{Fish}} - \kappa J = 0, \quad \kappa \leq 1$$

- Suggested by the constraint $J \geq I_{\text{Fish}}$ ("...the information I_{Fish} in the data never exceeds the amount J in the phenomenon").

"Extreme Physical Information (EPI)"

"The variational principle and the zero-condition comprise the overall principle we will use to derive major physical laws. Since, by [the Variational Principle], physical information K is extremized, the principle is called 'extreme physical information' or EPI."

What EPI is supposed to do:

- (a) The *variational principle* $\delta K = \delta(I_{\text{Fish}} - J) = 0$ is supposed to generate the equations of motion.
 - *Variation of I_{Fish} provides the "kinetic" (interaction-free) terms.*
 - *Variation of J provides the interaction terms.*
- (b) The *zero-condition* is supposed to generate the particular form of J for a given phenomenon, with the help of "invariance principles".

Big Problem: Step (b) doesn't work! (Lavis & Streater 2002)

So: At most, EPI is a way to interpret the Lagrangian formulation of theories in physics in Fisher information-theoretic terms.

3. Interpretive Issues

How should the variational principle, $\delta K = \delta(I_{\text{Fish}} - J) = 0$, be interpreted?

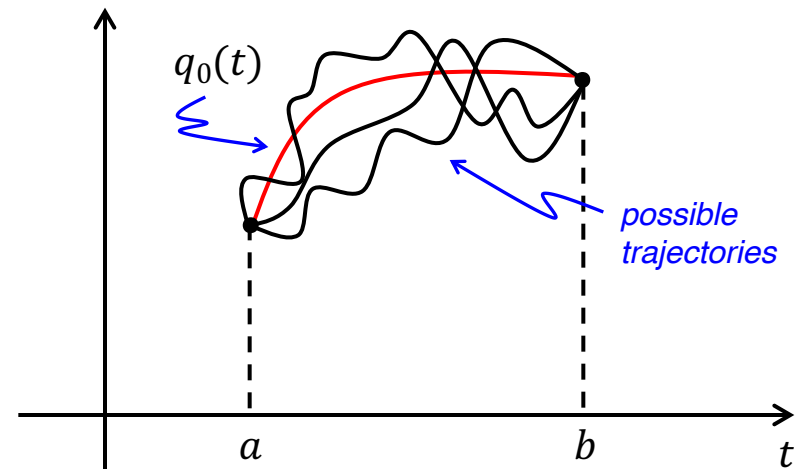
- Note: The Lagrangian formulation of physical theories is based on:

Principle of Least Action

- A physical system is described by an action functional $S(q)$, where q represents a possible quantity the system can possess.
- The *actual* quantity the system possesses is that q_0 that extremizes the action: $\delta S(q_0) = 0$.

Example: Massive Newtonian particle

- Described by the Newtonian action $S_{\text{Newt}}(q) = \int [\frac{1}{2}mq'^2 - V(q)]dt$, where $q(t)$ is a possible trajectory the particle can possess.
- The actual trajectory $q_0(t)$ is the one for which $\delta S_{\text{Newt}}(q_0) = 0$.



How should the Principle of Least Action be interpreted?

- (i) A fundamental physical principle?
- (ii) A mathematical way of encoding the equations of motion?

Problems with Option (i)

- (a) Why do actual physical quantities extremize the action?
 - *The action as a final cause of motion?*
 - *Contrast with local differential equations of motion as encoding efficient causes of motion.*
- (b) How is the form of the Lagrangian for a given system determined?
 - Standard approach: *Work backward from the equations of motion for a simple system to the form of its Lagrangian, and then extrapolate to more complicated systems.*
 - *But this seems to imply that the equations of motion are more fundamental.*

Does EPI offer a Fisher information-theoretic foundation for Option (i)?

Frieden comments:

- "Physics is... the science of measurement" (pg. 2).
- "EPI Theory proposes that all physical theory results from... imperfect observation" (pg. 3).
- A measurement is a game between observer and Nature in which the observer seeks to maximize info ($\delta K = 0$) and Nature seeks to maximize error (pp. 79-83).

- Recall: Key analogy

$$I_{\text{Fish}}(q(t)) = 4 \int q'^2 dt$$

The $q(t)$ that extremizes I_{Fish} is the most sensitive probability distribution (within a parameterized family) to changes in t (or the least spread in t).

$$S_{\text{Newt}}(q(t)) = \int \left[\frac{1}{2} m q'^2 - V(q) \right] dt$$

The $q(t)$ that extremizes S_{Newt} is the position function (path) for a Newtonian particle with mass m interacting with a potential $V(q)$.

Suggests: The $q(t)$ that extremizes S_{Newt} is the path that is most sensitive to changes in t . Given a collection of data sets generated by measurements of the position of a particle at different times, the Principle of Least Action says pick the curve that's the most sensitive to changes in t .

- Concern #1: What happened to the probabilities? ← *In the Newtonian case, $q(t)$ isn't a probability distribution.*
- Possible response: Fisher information is a measure of accuracy in fitting data to curves, and it doesn't matter what the curves represent. ← *They can be probability distributions (in which case we're constructing statistical models), or they can be physical quantities (in which case we're doing Lagrangian mechanics).*

- Recall: Key analogy

$$I_{\text{Fish}}(q(t)) = 4 \int q'^2 dt$$

The $q(t)$ that extremizes I_{Fish} is the most sensitive probability distribution (within a parameterized family) to changes in t (or the least spread in t).

$$S_{\text{Newt}}(q(t)) = \int \left[\frac{1}{2} m q'^2 - V(q) \right] dt$$

The $q(t)$ that extremizes S_{Newt} is the position function (path) for a Newtonian particle with mass m interacting with a potential $V(q)$.

Suggests: The $q(t)$ that extremizes S_{Newt} is the path that is most sensitive to changes in t . Given a collection of data sets generated by measurements of the position of a particle at different times, the Principle of Least Action says pick the curve that's the most sensitive to changes in t .

- Concern #2: Is the emphasis on curve-fitting an instrumentalist view of physics?

← *Theories in physics are "black boxes" that don't go beyond the input data.*

- Possible response: We use data sets to derive equations of motion, and the latter go beyond the data.

← *The data are used to find the optimal curve, which goes beyond the data.*

"This view regards reality as being perpetrated by requests for information" (Friedan 1998, pg. 108).

Charitable interpretation:

- Physics is based on collecting data sets.
- Physicists then attempt to fit these data sets to curves that represent physical quantities.
- And the way they do this is by determining which curve is the most sensitive to the data.

- Suggests: The extremization procedure isn't "out there" in Nature as a final cause; rather, it's a method physicists use to explain data.
- So: "Reality" is not information-theoretic, or observer-dependent; rather, our methods for describing reality are.