Newtonian Mechanics from the principle of Maximum Caliber

Diego González, Sergio Davis, Gonzalo Gutiérrez

Facultad de Ciencias, Universidad de Chile

MaxEnt 2014, September 26th, 2014
Outline

- Motivation
- MaxEnt and the divergence theorem
- The Maximum Caliber principle
- Newton’s 2nd law
- Inertia and potential energy
- Conclusions
Some motivation

- Why worry about interpretations?
- Ideas in the theme of “physics from inference”
- Several (most?) “laws of physics” are just of statistical nature
- We might find those “laws of physics” applied in completely unexpected contexts
For a system of continuous degrees of freedom $\vec{x}$ subjected to $M$ constraints ($j=1,\ldots,M$),

$$\langle f_j(\vec{x}) \rangle_I = F_j$$

(1)

the most unbiased probability distribution is

$$P(\vec{x}|I) = \frac{1}{Z} \exp \left( - \sum_{j=1}^{M} \lambda_j f_j(\vec{x}) \right)$$

(2)

Note that we are omitting for simplicity a constant “complete ignorance” distribution $P(\vec{x}|I_0)$. The Lagrange multipliers $\lambda_j$ can be obtained from

$$- \frac{\partial}{\partial \lambda_j} \ln Z = F_j.$$ 

(3)
Solving for $\lambda$: the divergence theorem

For an arbitrary (differentiable) distribution $P(\vec{x}|I)$, let us compute the expectation of a divergence (of an arbitrary vector field $\vec{\omega}$) and apply the divergence theorem. We get

$$\langle \nabla \cdot \vec{\omega}(\vec{x}) \rangle_I = -\langle \vec{\omega}(\vec{x}) \cdot \nabla \ln P(\vec{x}|I) \rangle_I.$$  \hfill (4)

For the special case of a MaxEnt distribution,

$$\langle \nabla \cdot \vec{\omega}(\vec{x}) \rangle_I = \sum_j \lambda_j \langle \vec{\omega}(\vec{x}) \cdot \nabla f_j(\vec{x}) \rangle_I.$$  \hfill (5)

This gives us an alternative route to solving for the Lagrange multipliers, a linear system of equations (by choosing different fields $\vec{\omega}$).

More details and implications of this relation in:
The Maximum Caliber principle

Suppose we are now making inferences about the trajectory $x(t)$ followed by a system in time. We constrain a functional $f[x(t); \tau]$ which is known for $\tau \in [t_i, t_f]$.

$$\langle f[x(t); \tau] \rangle_I = F(\tau),$$

We maximize the *caliber* or path entropy

$$S = -\int Dx(t) P[x(t)|I] \ln P[x(t)|I],$$

under the functional constraint, obtaining

$$P[x(t)|I] = \frac{1}{Z[\lambda(t)]} \exp \left( -\int_{t_i}^{t_f} d\tau \lambda(\tau) f[x(t); \tau] \right)$$

where $\lambda(\tau)$ is now a Lagrange multiplier function.

Jaynes called the quantity $S$ the *caliber* of the system, as it is analogous to the cross section of a barrel (the “tube” spawned by the possible paths).
The Maximum Caliber principle

\[ A[x(t)] = \alpha \int_{t_i}^{t_f} d\tau \lambda(\tau) f[x(t); \tau] \]

\[ P[x(t) | I] = \frac{1}{Z[\lambda(t) \lambda]} \exp \left(-\frac{1}{\alpha} A[x(t)] \right). \]

The action then is the “relevant quantity” for the trajectories, and \(1/\alpha\) its multiplier. It is just a constant with units of action introduced to make the exponent adimensional.

\[ \frac{\delta P}{\delta x(t)} \bigg|_{x=x_{cl}} = 0 \Rightarrow \frac{\delta A}{\delta x(t)} \bigg|_{x=x_{cl}} = 0. \] (6)

With \(\alpha > 0\), the most probable trajectory \(x_{cl}(t)\) (let us call it “classical”) is the one that minimizes the action \(A\).
Feynman’s path integral formulation:

\[ P[x(t)|I] = \frac{1}{Z} \exp \left( -\frac{1}{\alpha} A[x(t)] \right) \]

\[ P(x_1(t_1), x_2(t_2)|I) = \frac{1}{Z} \int \mathcal{D}x(t) \exp \left( -\frac{1}{\alpha} A[x(t)] \right) \]

\[ K(x_1, x_2; t_1, t_2) \propto \int \mathcal{D}x(t) \exp \left( \frac{i}{\hbar} S[x(t)] \right) \]
If $f[x(t); \tau] = f(x(\tau), \dot{x}(\tau); \tau)$ the action is the time integral of a Lagrangian,

$$A = \int_{t_1}^{t_2} d\tau L(x(\tau), \dot{x}(\tau); \tau)$$

$L(x, \dot{x}; t) = \alpha \lambda(t)f(x, \dot{x}; t)$

For the most probable trajectory $x_{cl}(t)$, the Euler-Lagrange equation holds,

$$\left. \left( \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) \right|_{x=x_{cl}} = 0.$$

(Euler-Lagrange equation)

We obtain an extremum principle for (classical) mechanics without invoking quantum mechanics as an underlying theory. Classical mechanics is just the most probable answer to our *inference over trajectories* problem.
Newton’s 2nd law

Of course we could from the start take the function \( f \) to be proportional to a known Lagrangian, but . . . Could we make it emerge for simpler constraints? Consider a coordinate which describes a random walk such that

1. The expected size of the displacement \((\Delta x)^2 = v(t)^2(\Delta t)^2\) is known at all times (i.e. we know the magnitude of the velocity)
2. The time-independent probability \( p(x) \) is known

It is direct to prove that

\[
\frac{d}{dt}(m(t)\dot{x}(t)) \bigg|_{x=x_{cl}} = \left( -\frac{d}{dx}\Phi(x) \right) \bigg|_{x=x_{cl}}. \tag{7}
\]

where \( m(t) \) and \( \Phi(x) \) are functions which impose the first and second constraint, respectively. Not only that, for the ensemble of trajectories under those general constraints,

\[
\left\langle \frac{d}{dt}(m(t)\dot{x}(t)) \right\rangle_I = -\left\langle \frac{d}{dx}\Phi(x) \right\rangle_I. \tag{8}
\]
Newton’s 2nd law

Our constraints are

\[ \langle \dot{x}(t)^2\rangle = v(t)^2 \]  
(9)

\[ \langle \delta(x(t) - X)\rangle = p(X). \]  
(10)

for all values of \( t \). Discretizing time,

\[ \langle (x_k - x_{k-1})^2\rangle = v_k^2 \Delta t^2 \]  
(11)

\[ \langle \delta(x_k - X)\rangle = p(X). \]  
(12)
Therefore the probability of the discrete-time trajectory $\vec{x} = (x_1, x_2, \ldots, x_N)$ is given by

$$P(\vec{x}|I) = \frac{1}{Z} \exp \left( -\frac{1}{\alpha} \sum_k \left[ \lambda_k (x_k - x_{k-1})^2 + \int dX \mu(X) \delta(x_k - X) \right] \right)$$

(13)

$$= \frac{1}{Z} \exp \left( -\frac{1}{\alpha} \sum_k \left[ \lambda_k (x_k - x_{k-1})^2 + \mu(x_k) \right] \right).$$

Continuous version:

$$P[x(t)|I] = \frac{1}{Z} \exp \left( -\frac{1}{\alpha} \int dt \left[ \lambda(t) \dot{x}(t)^2 + \mu(x(t)) \right] \right)$$

(14)
Newton’s 2nd law

\[ P[x(t)\mid I] = \frac{1}{Z} \exp \left( -\frac{1}{\alpha} \int dt \left[ \lambda(t)\dot{x}(t)^2 + \mu(x(t)) \right] \right) \]  

This corresponds to a system with Lagrangian

\[ L(x, \dot{x}, t) = \lambda(t)\dot{x}(t)^2 + \mu(x) = \frac{1}{2} m(t)\dot{x}(t)^2 - \Phi(x) \]  

under the identification \( m(t)/2 = \lambda(t) \) and \( \Phi(x) = -\mu(x) \). By the Euler-Lagrange equation,

\[ \frac{d}{dt} (m(t)\dot{x}(t))\bigg|_{x=x_{cl}} = \left( - \frac{d}{dx} \Phi(x) \right)\bigg|_{x=x_{cl}}. \]
How do we describe the “non-classical” trajectories?

\[ \langle B[x(t)] \rangle_I = \int Dx(t) P[x(t)|I] B[x(t)] \]

Generalizing the vector result,

\[ \langle \nabla \cdot \vec{\omega} (\vec{x}) \rangle_I = -\langle \vec{\omega} (\vec{x}) \cdot \nabla \ln P(\vec{x}|I) \rangle_I, \]

we obtain

\[ \langle \frac{\delta W[x(t)]}{\delta x(t)} \rangle_I = -\langle W[x(t)] \frac{\delta}{\delta x(t)} \ln P[x(t)|I] \rangle_I, \quad (18) \]

\[ \langle \frac{\delta W[x(t)]}{\delta x(t)} \rangle_I = \frac{1}{\alpha} \langle W[x(t)] \frac{\delta A}{\delta x(t)} \rangle_I = \frac{1}{\alpha} \langle W[x(t)] \cdot (\hat{E}_t L) \rangle_I. \quad (19) \]

where \( W[x(t)] \) is a test functional, and \( \hat{E}_t G = (\frac{\partial}{\partial x} - \frac{d}{dt} \frac{\partial}{\partial x}) G. \)
Newton’s 2nd law: ensemble

\[
\left\langle \frac{\delta W[x(t)]}{\delta x(t)} \right\rangle_I = \frac{1}{\alpha} \left\langle W[x(t)] \cdot (\hat{E}_t L) \right\rangle_I.
\] (20)

Choosing \( W[x(t)] \) to be a constant functional,

\[
\frac{1}{\alpha} \left\langle \hat{E}_t L \right\rangle_I = 0,
\]

Which is nothing but

\[
\left\langle \frac{d}{dt} (m(t) \dot{x}(t)) \right\rangle_I = -\left\langle \frac{d}{dx} \Phi(x) \right\rangle_I.
\] (21)

This derivation without the functional identity (just using the discretized version of the trajectory and the divergence theorem) is given in:
Inertia and potential energy as emergent

Notice the identification $m(t)/2 = \lambda(t)$. This gives a meaning for the mass parameter in the Lagrangian:

The more informative the constraint about the magnitude of the velocity is, the more massive the “particle” is.

Now notice that we also identified $\Phi(x) = -\mu(x)$. This means that

The more informative the constraint about the allowed positions of a particle, the more confined it is by a potential. Note that we assign a model for $p(x)$ and this leads us to a model for $\Phi(x)$. 
A slight generalization

\[ \langle \delta(\dot{x}(t) - V) \rangle_I = q(V; t) \]
\[ \langle \delta(x(t) - X) \rangle_I = p(X; t) \]

\[ P[x(t)|I] \propto \exp \left( - \int dt \left[ \int dV \delta(\dot{x}(t) - V)\lambda(V, t) + \int dX \delta(x(t) - X)\mu(X, t) \right] \right) \]
\[ = \frac{1}{Z} \exp \left( - \int dt [\lambda(\dot{x}(t), t) + \mu(x(t), t)] \right) \]

Now the Lagrange multiplier function \( \lambda(\dot{x}, t) \) is itself the kinetic energy! (again, a \textbf{model} for the probability of the velocity leads to a \textbf{model} for the kinetic energy)

\[ p = \frac{\partial L}{\partial \dot{x}} = \frac{\partial \lambda}{\partial \dot{x}} \Rightarrow m = \frac{\partial p}{\partial \dot{x}} = \frac{\partial^2 L}{\partial \dot{x}^2} \]

Taylor-expanding \( \lambda \) around \( \dot{x} = 0 \) up to second order will give you the classical (non-relativistic) kinetic energy.
A natural generalization of Maximum Entropy inference for dynamical systems is the Maximum Caliber principle.

It naturally provides us with a minimum action law, if we read the “relevant quantity” as an action.

Some properties can be computed in expectation over trajectories under the Maximum Caliber distribution.

It leads to Newton’s second law under relatively broad assumptions.
Thank you for your attention!

Acknowledgements: FONDECYT grant 1140514