When does a dynamical system implement a statistical inference?

Dalton A R Sakthivadivel

2 June 2026

THE GRADUATE CENTER
CITY UNIVERSITY OF NEW YORK

Overview

We will discuss the various notions of computation arising in (models of) dynamical systems and random processes in physics. None of these systems will be digital computers, but nevertheless can be viewed as undergoing state transitions from which an implicit, intrinsic computation can be extracted, resting on variational principles for motion

A brief account of 'anthropic computation' will lead to a specific discussion of open systems in statistical mechanics and the performance of Bayesian inference in variational descriptions of non-equilibrium steady state

Some more philosophical questions about the nature of representations of information, when teleological distinctions between goals and reality are justified and reified, and what forms of cognition and agency exist where in the natural world, will be raised

Further details

These slides and other lectures like it can be found later at darsakthi.github.io/talks. Some useful references will be contained on the final slides.

My viewpoint on this topic has been shaped by many insightful discussions with many friends and collaborators over the years. I am especially thankful to Scott Aaronson, Mel Andrews, Ken Dill, Karl Friston, Mike Levin, and Adam Safron.

This research is supported by the Einstein Chair at the Graduate Centre of the City University of New York and the Extropic Corporation.

Contents

From Hamilton to Jaynes

Variational principles in physics

How do we form globally-valid knowledge about the natural world from locally-made observations on innovations and outcomes?

We form general principles from which laws of physics can be derived

Historically this has had the general form of working backwards: given a known law implied by some understanding of the mechanics of a system...

... *e.g.* the energy of motion and the potential energy of a massive particle are functions of mass, velocity, and position...

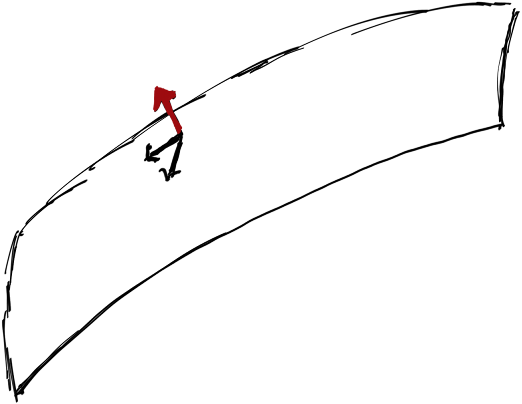
one finds a function of states or dynamics for which one can demonstrate that dynamics varying from the law are inconsistent with those mechanics.

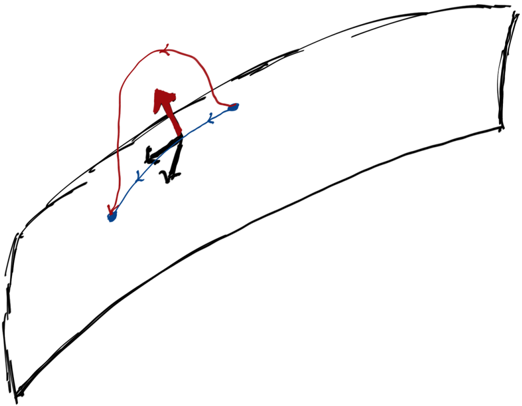
Variational principles in physics (continued)

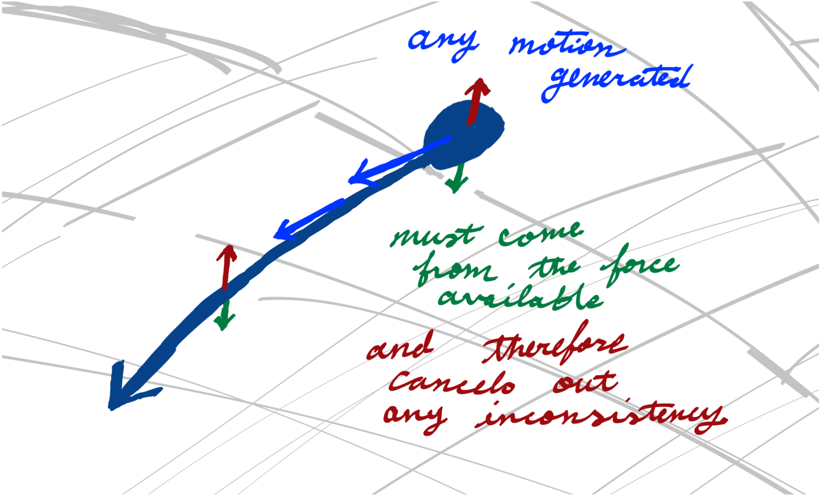
Consequences:

Physically-realisable dynamics are confined to a subspace of trajectories by the laws of motion constraining them

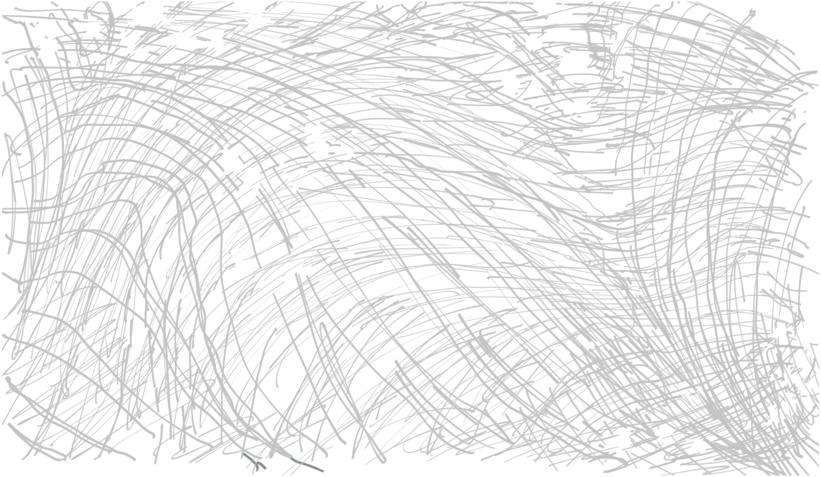
> It is a geometric fact that constraint satisfaction implies some function is made stationary, *i.e.*, its gradient vanishes.



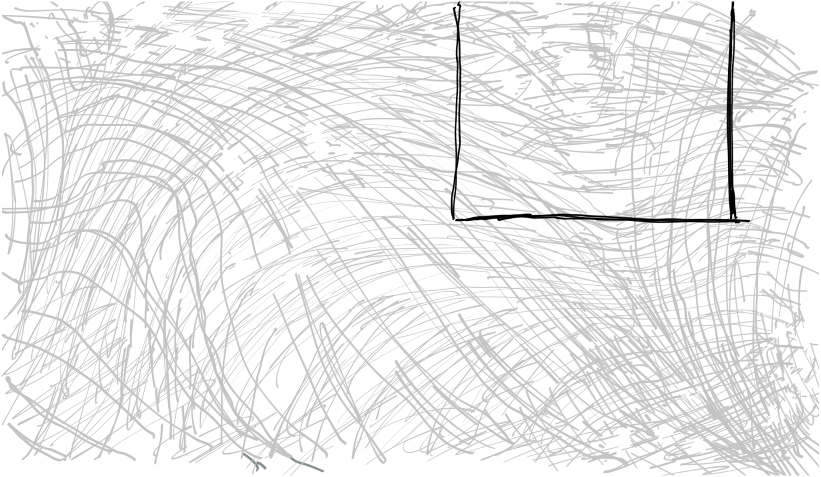



any motion generated
must come from the force available
and therefore cancels out any inconsistency

Newton's laws, Lagrange's equations, and Hamilton's principle



Newton's laws, Lagrange's equations, and Hamilton's principle

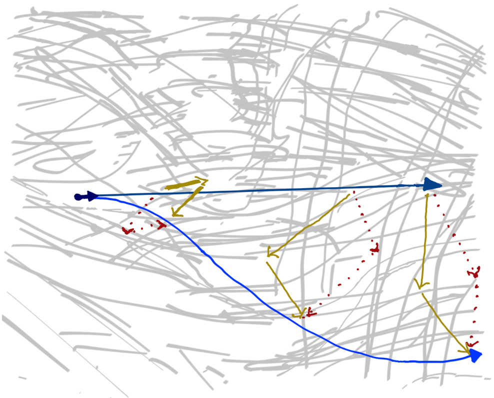


Newton's laws, Lagrange's equations, and Hamilton's principle

Newton's three laws (paraphrased):

1. Inertia is the tendency for momentum to resist change
2. The sum of forces acting on an object creates a change in momentum equal in magnitude and direction
3. Every action creates an action equal in magnitude and opposite in direction

Newton's laws, Lagrange's equations, and Hamilton's principle



Newton's three laws (paraphrased):

1. Inertia is the tendency for momentum to resis- change
2. The sum of forces acting on an object creates a change in momentum equal in magnitude and direction
3. Every action creates an action equal in magnitude and opposite in direction

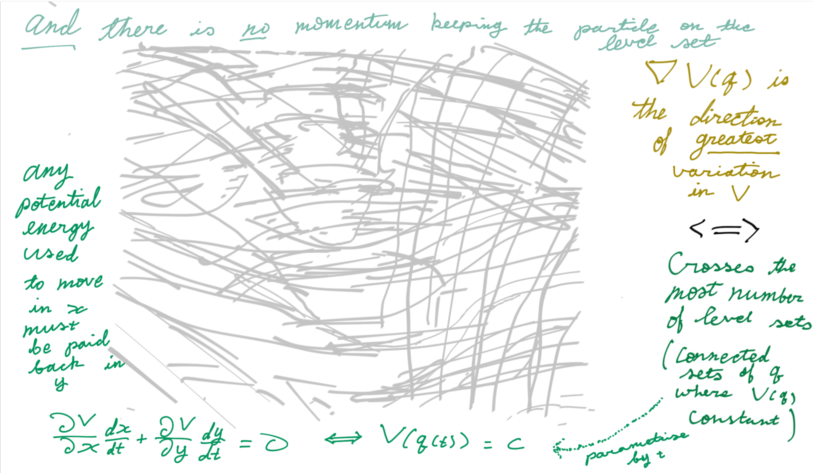
Newton's laws, Lagrange's equations, and Hamilton's principle

Now add the Conservation of energy:

1. In order to create a change in momentum a force must be applied by paying with potential energy ...

... and ^ then repaid.
eventually!

Newton's laws, Lagrange's equations, and Hamilton's principle



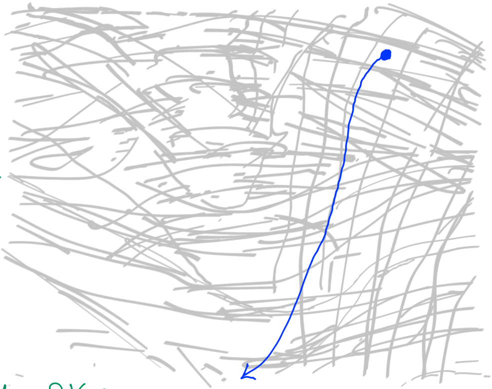
Newton's laws, Lagrange's equations, and Hamilton's principle

$\vec{v} = -\nabla V$

$\iff$

$\nabla V \cdot \dot{q}(t) = 0$

$\iff$



$$\frac{\partial V}{\partial x}\frac{dx}{dt} + \frac{\partial V}{\partial y}\frac{dy}{dt} = 0 \iff V(q(t)) = c$$

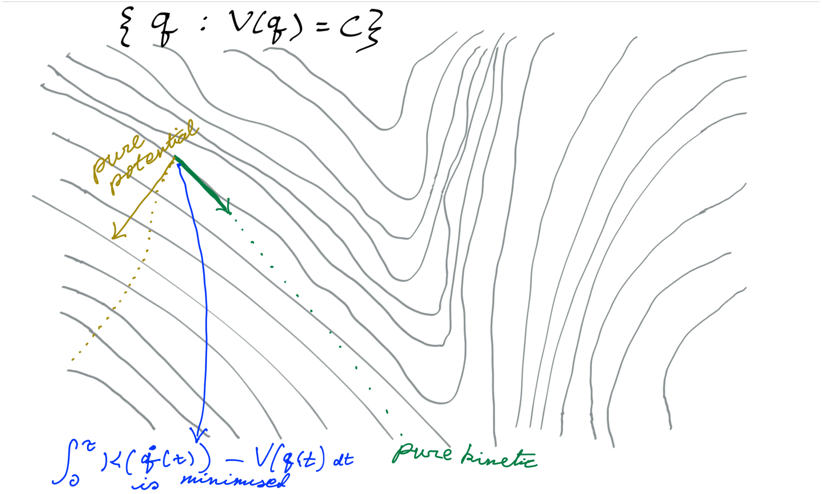
parametrise by t

$\nabla V(q)$ is the direction of greatest variation in V

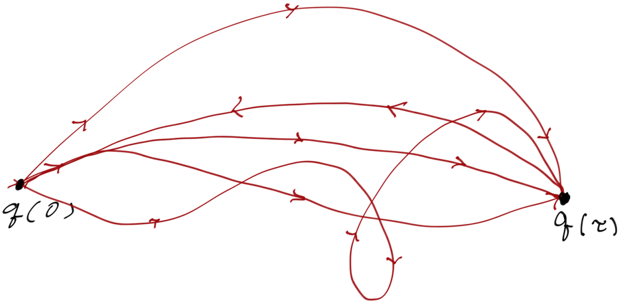
$\iff$

Crosses the most number of level sets (Connected sets of q where $V(q)$ constant)

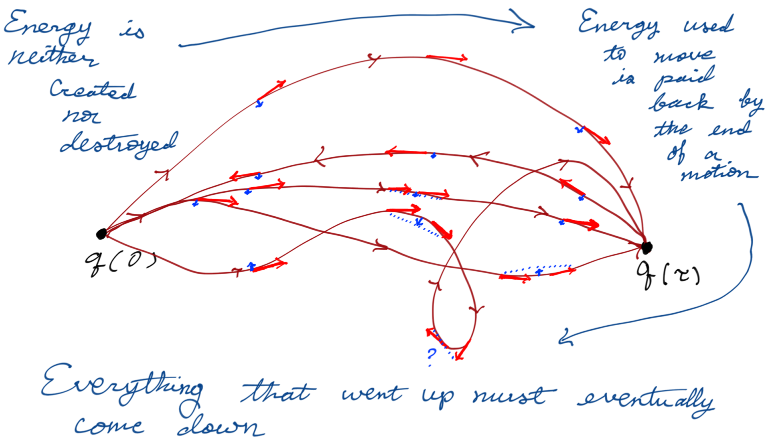
Newton's laws, Lagrange's equations, and Hamilton's principle



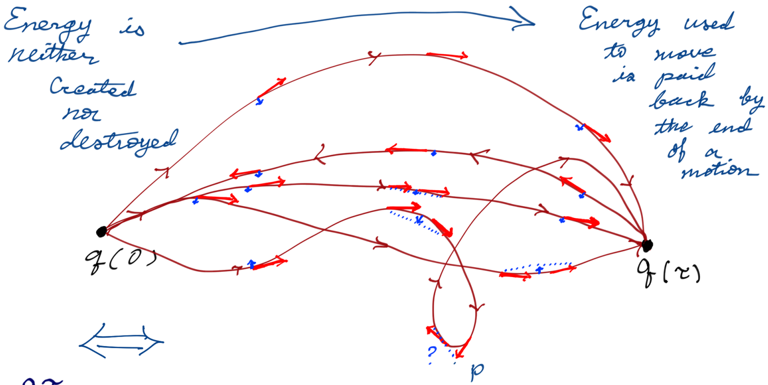
Newton's laws, Lagrange's equations, and Hamilton's principle



Newton's laws, Lagrange's equations, and Hamilton's principle

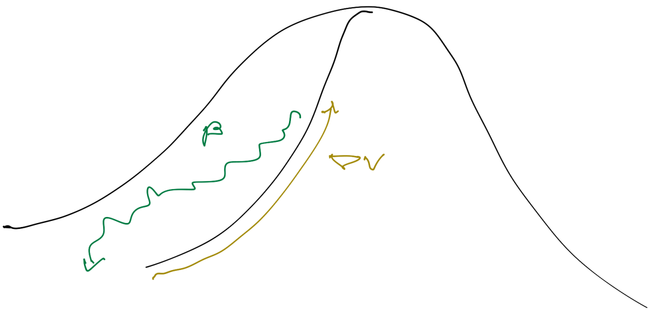


Newton's laws, Lagrange's equations, and Hamilton's principle



$$\int_0^\tau K(\dot{q}(t)) - V(q(t))\,dt = 0$$

Statistical mechanics



Statistical mechanics

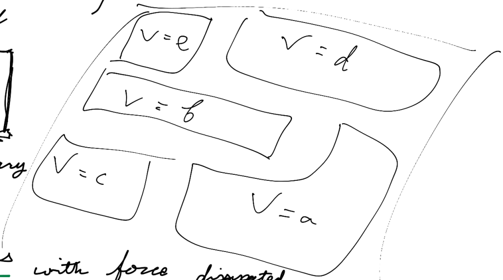
$$P\left(\underbrace{\text{system is in configuration } x}_{\text{microscopic inference}} \;\middle|\; \underbrace{\text{observed average of } V(x) = E}_{\text{macroscopic evidence}}\right)$$

is proportional to

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

- arises from a stationary action principle due to balance of kinetic fluctuations with force dissipated.

⌜At equilibrium⌟



Statistical mechanics

Namely: $\rho = \operatorname{argmin} \int \rho \log \rho - \left(\beta \int V\rho - E\right)$

Maximise number of possible configurations

Covers tendency of fluctuations to spread out due to high kinetic energy

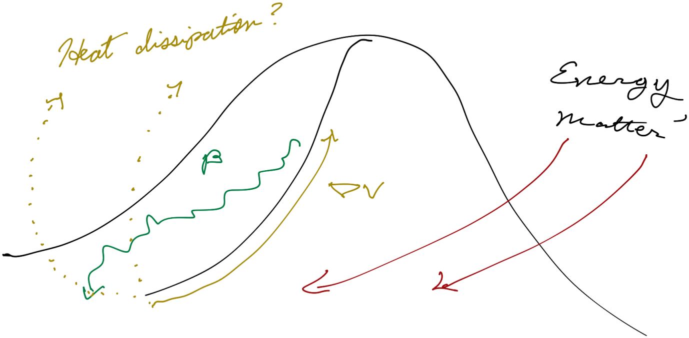
Constrained by average dissipative potential

Covers gradient climbing pulling fluctuations back by force

Statistical mechanics

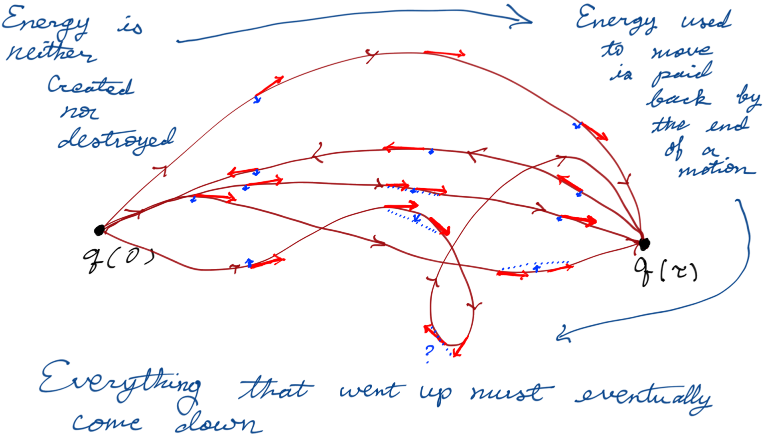
We owe these great insights perhaps foremostly to Edwin Jaynes, who synthesised the original work of Boltzmann, Gibbs, Landauer, and so on, with the work of Shannon, Wiener, and other early information theorists. Some comments are appended in the last slides.

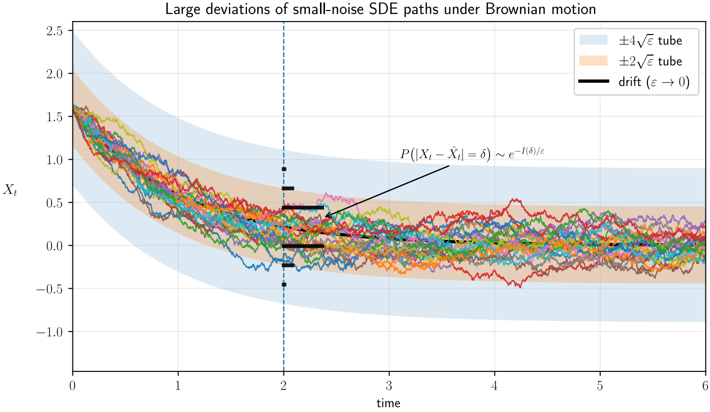
Open statistical mechanics?



And back again

(by way of Feynman)


Energy is neither Created nor destroyed
Energy used to move is paid back by the end of a motion
q(0)
q(τ)
?
Everything that went up must eventually come down


Large deviations of small-noise SDE paths under Brownian motion
$\pm 4\sqrt{\varepsilon}$ tube
$\pm 2\sqrt{\varepsilon}$ tube
drift ($\varepsilon \to 0$)
$P(|X_t - \hat{X}_t| = \delta) \sim e^{-I(\delta)/\varepsilon}$
X_t
time
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
0
1
2
3
4
5
6

Then ending at Bayes

(by an analysis of open statistical mechanics, and random paths of stationary action)

Modelling and metaphor; or: the modeller *vs.* the modelled

Question: does a particle really compute gradients as it moves through time and space?

A better question: does it matter?

A careful answer: not if we think correctly


Ben Nevis Inn
WEATHER FORECASTING STONE
CONDITION
FORECAST
Stone is Wet
Rain
Stone is Dry
Not Raining
Shadow on Ground
Sunny
White on Top
Snowing
Can't see Stone
Foggy
Swinging Stone
Windy
Stone Jumping Up & Down
Earthquake
Stone Gone
Tornado

We begin from a system of two coupled random variables evolving in time separated by a boundary,

$$X_t \xrightarrow{g} B_t \xrightarrow{h} Y_t$$

assumed to satisfy Itō SDEs

$$\begin{aligned} \mathrm{d}X_t &= f_1(X_t, B_t, t)\,\mathrm{d}t + D_1(X_t, B_t, t)\,\mathrm{d}W_t^1 \\ \mathrm{d}B_t &= f_2(X_t, B_t, Y_t, t)\,\mathrm{d}t + D_2(X_t, B_t, Y_t, t)\,\mathrm{d}W_t^2 \\ \mathrm{d}Y_t &= f_3(B_t, Y_t, t)\,\mathrm{d}t + D_3(B_t, Y_t, t)\,\mathrm{d}W_t^3 \end{aligned}$$

The precise coupling structure is specific to a given system and defines different classes of dynamics.

We begin from a system of two coupled random variables evolving in time separated by a boundary,

$$X_t \xrightarrow{g} B_t \xrightarrow{h} Y_t$$

assumed to satisfy Itō SDEs

$$\begin{aligned}
\mathrm{d}X_t &= f_1(X_t, B_t, t)\,\mathrm{d}t + D_1(X_t, B_t, t)\,\mathrm{d}W_t^1\\
\mathrm{d}B_t &= f_2(X_t, B_t, Y_t, t)\,\mathrm{d}t + D_2(X_t, B_t, Y_t, t)\,\mathrm{d}W_t^2\\
\mathrm{d}Y_t &= f_3(B_t, Y_t, t)\,\mathrm{d}t + D_3(B_t, Y_t, t)\,\mathrm{d}W_t^3
\end{aligned}$$

The precise coupling structure is specific to a given system and defines different classes of dynamics.

We begin from a system of two coupled random variables evolving in time separated by a boundary,

$$X_t \xrightarrow{g} B_t \xrightarrow{h} Y_t$$

assumed to satisfy Itō SDEs

$$\begin{aligned}
\mathrm{d}X_t &= f_1(X_t, B_t, t)\,\mathrm{d}t + D_1(X_t, B_t, t)\,\mathrm{d}W_t^1 \\
\mathrm{d}B_t &= f_2(X_t, B_t, Y_t, t)\,\mathrm{d}t + D_2(X_t, B_t, Y_t, t)\,\mathrm{d}W_t^2 \\
\mathrm{d}Y_t &= f_3(B_t, Y_t, t)\,\mathrm{d}t + D_3(B_t, Y_t, t)\,\mathrm{d}W_t^3
\end{aligned}$$

The precise coupling structure is specific to a given system and defines different classes of dynamics.

Suppose f_1 is not the gradient of a smooth function and X_t has a pullback attractor in the state space

Then under certain regularity assumptions there exists a non-equilibrium steady state density $p^*(x)$

In which case we have the normal form

$$\mathrm{d}X_t = -(Q - \Gamma)\nabla_x \log p^*(X_t)\,\mathrm{d}t + D\,\mathrm{d}W_t$$

with $Q^\top = -Q$, $\Gamma \geqslant 0$, and $2\Gamma = DD^\top$.

Suppose f_1 is not the gradient of a smooth function and X_t has a pullback attractor in the state space

Then under certain regularity assumptions there exists a non-equilibrium steady state density $p^*(x)$

In which case we have the normal form

$$\mathrm{d}X_t = -(Q - \Gamma)\nabla_x \log p^*(X_t)\,\mathrm{d}t + D\,\mathrm{d}W_t$$

with $Q^\top = -Q$, $\Gamma \geqslant 0$, and $2\Gamma = DD^\top$.

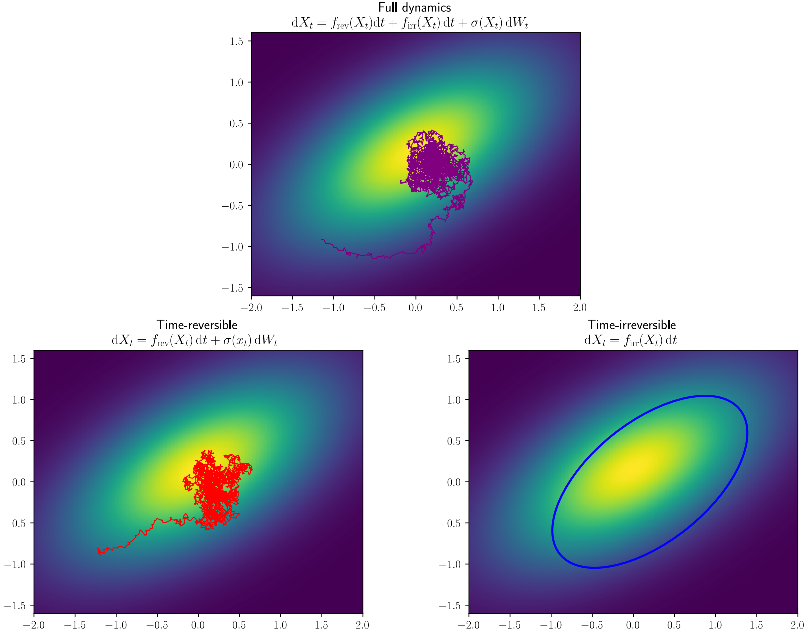
Suppose f_1 is not the gradient of a smooth function and X_t has a pullback attractor in the state space

Then under certain regularity assumptions there exists a non-equilibrium steady state density $p^*(x)$

In which case we have the normal form

$$\mathrm{d}X_t = -(Q - \Gamma)\nabla_x \log p^*(X_t)\,\mathrm{d}t + D\,\mathrm{d}W_t$$

with $Q^\top = -Q$, $\Gamma \geqslant 0$, and $2\Gamma = DD^\top$.


Full dynamics
$\mathrm{d}X_t = f_{\mathrm{rev}}(X_t)\mathrm{d}t + f_{\mathrm{irr}}(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t$
Time-reversible
$\mathrm{d}X_t = f_{\mathrm{rev}}(X_t)\,\mathrm{d}t + \sigma(x_t)\,\mathrm{d}W_t$
Time-irreversible
$\mathrm{d}X_t = f_{\mathrm{irr}}(X_t)\,\mathrm{d}t$

Now stipulate a variational posterior $q(y; x)$ and write

$$F(b, x) := \int q(y; x) \log q(y; x)\, \mathrm{d}y - \int q(y; x) \log p(y, b, x)\, \mathrm{d}y$$

Recall that $p(y, b, x) = p(y \mid b, x) p(b, x)$ and $\log ab = \log a + \log b$

Applying this and conditional independence we have

$$F(b, x) = \underset{q(y;x)}{\mathbf{E}}[\log q(y; x)] - \underset{q(y;x)}{\mathbf{E}}[\log p(y \mid b)] - \log p(b, x)$$

... so that

$$F(b, x) \geqslant -\log p(b, x).$$

Now stipulate a variational posterior $q(y; x)$ and write

$$F(b, x) := \int q(y; x) \log q(y; x) \, \mathrm{d}y - \int q(y; x) \log p(y, b, x) \, \mathrm{d}y$$

Recall that $p(y, b, x) = p(y \mid b, x) p(b, x)$ and $\log ab = \log a + \log b$

Applying this and conditional independence we have

$$F(b, x) = \underset{q(y;x)}{\mathbf{E}} [\log q(y; x)] - \underset{q(y;x)}{\mathbf{E}} [\log p(y \mid b)] - \log p(b, x)$$

... so that

$$F(b, x) \geqslant -\log p(b, x).$$

Now stipulate a variational posterior $q(y; x)$ and write

$$F(b, x) := \int q(y; x) \log q(y; x) \, \mathrm{d}y - \int q(y; x) \log p(y, b, x) \, \mathrm{d}y$$

Recall that $p(y, b, x) = p(y \mid b, x) p(b, x)$ and $\log ab = \log a + \log b$

Applying this and conditional independence we have

$$F(b, x) = \underset{q(y;x)}{\mathbf{E}}[\log q(y; x)] - \underset{q(y;x)}{\mathbf{E}}[\log p(y \mid b)] - \log p(b, x)$$

. . . so that

$$F(b, x) \geqslant -\log p(b, x).$$

Now stipulate a variational posterior $q(y;x)$ and write

$$F(b,x) := \int q(y;x) \log q(y;x)\, \mathrm{d}y - \int q(y;x) \log p(y,b,x)\, \mathrm{d}y$$

Recall that $p(y,b,x) = p(y \mid b,x)p(b,x)$ and $\log ab = \log a + \log b$

Applying this and conditional independence we have

$$F(b,x) = \underbrace{\boxed{\mathop{\mathbf{E}}_{q(y;x)}[\log q(y;x)] - \mathop{\mathbf{E}}_{q(y;x)}[\log p(y \mid b)]}}_{\geqslant 0 \ldots} - \log p(b,x)$$

... so that

$$F(b,x) \geqslant -\log p(b,x).$$

Now stipulate a variational posterior $q(y;x)$ and write

$$F(b,x) := \int q(y;x)\log q(y;x)\,\mathrm{d}y - \int q(y;x)\log p(y,b,x)\,\mathrm{d}y$$

Recall that $p(y,b,x) = p(y \mid b,x)p(b,x)$ and $\log ab = \log a + \log b$

Applying this and conditional independence we have

$$F(b,x) = \underbrace{\boxed{\mathop{\mathbf{E}}_{q(y;x)}[\log q(y;x)] - \mathop{\mathbf{E}}_{q(y;x)}[\log p(y \mid b)]}}_{\geqslant 0 \ldots} - \log p(b,x)$$

... so that

$$F(b,x) \geqslant -\log p(b,x).$$

Suppose there exists an x such that $q(y; x)$ equals $p(y \mid b)$ almost surely. Denote that x as x^*

Then $\mathbf{E}[\log q(y; x^*)] = \mathbf{E}[\log p(y \mid b)]$ and our SDE becomes

$$\mathrm{d}X_t = -(Q - \Gamma)\nabla_x F(b, x^*)\,\mathrm{d}t + D\,\mathrm{d}W_t$$

The difference of expectations is the *KL divergence* between a variational posterior and target distribution; the free energy is a tractable upper bound on model evidence

Implication:

All Markov-blanketted non-equilibrium processes on an attractor in their state space can be written as if implementing Bayesian inference over the likely causes of sensations

Suppose there exists an x such that $q(y; x)$ equals $p(y \mid b)$ almost surely. Denote that x as x^*

Then $\mathbf{E}[\log q(y; x^*)] = \mathbf{E}[\log p(y \mid b)]$ and our SDE becomes

$$\mathrm{d}X_t = -(Q - \Gamma)\nabla_x F(b, x^*)\,\mathrm{d}t + D\,\mathrm{d}W_t$$

The difference of expectations is the *KL divergence* between a variational posterior and target distribution; the free energy is a tractable upper bound on model evidence

Implication:

All Markov-blanketted non-equilibrium processes on an attractor in their state space can be written as if implementing Bayesian inference over the likely causes of sensations

Suppose there exists an x such that $q(y; x)$ equals $p(y \mid b)$ almost surely. Denote that x as x^*

Then $\mathbf{E}[\log q(y; x^*)] = \mathbf{E}[\log p(y \mid b)]$ and our SDE becomes

$$\mathrm{d}X_t = -(Q - \Gamma)\nabla_x F(b, x^*)\,\mathrm{d}t + D\,\mathrm{d}W_t$$

The difference of expectations is the *KL divergence* between a variational posterior and target distribution; the free energy is a tractable upper bound on model evidence

Implication:

All Markov-blanketted non-equilibrium processes on an attractor in their state space can be written as if implementing Bayesian inference over the likely causes of sensations

Suppose there exists an x such that $q(y; x)$ equals $p(y \mid b)$ almost surely. Denote that x as x^*

Then $\mathbf{E}[\log q(y; x^*)] = \mathbf{E}[\log p(y \mid b)]$ and our SDE becomes

$$\mathrm{d}X_t = -(Q - \Gamma)\nabla_x F(b, x^*)\,\mathrm{d}t + D\,\mathrm{d}W_t$$

The difference of expectations is the *KL divergence* between a variational posterior and target distribution; the free energy is a tractable upper bound on model evidence

Implication:

All Markov-blanketted non-equilibrium processes on an attractor in their state space can be written as if implementing Bayesian inference over the likely causes of sensations

Suppose there exists an x such that $q(y; x)$ equals $p(y \mid b)$ almost surely. Denote that x as x^*

Then $\mathbf{E}[\log q(y; x^*)] = \mathbf{E}[\log p(y \mid b)]$ and our SDE becomes

$$\mathrm{d}X_t = -(Q - \Gamma)\nabla_x F(b, x^*)\,\mathrm{d}t + D\,\mathrm{d}W_t$$

The difference of expectations is the *KL divergence* between a variational posterior and target distribution; the free energy is a tractable upper bound on model evidence

Implication:

All Markov-blanketted non-equilibrium processes on an attractor in their state space can be written as if implementing Bayesian inference over the likely causes of sensations

Recollection: a canonical minimal environmental estimator?

Interpretation (with respect to our model of a process as representing some extractable information):

If a system mainly does what one (a modeller) expects it to do, it can only be so surprising

For instance

- Observable stones must be concentrated on stone-like states
- Observable control systems must be concentrated on set points

Interpretation (with respect to our model of a process as representing some extractable information):

If a system mainly does what one (a modeller) expects it to do, it can only be so surprising

For instance

- Observable stones must be concentrated on stone-like states
- Observable control systems must be concentrated on set points

Interpretation (with respect to our model of a process as representing some extractable information):

If a system mainly does what one (a modeller) expects it to do, it can only be so surprising

For instance

- Observable stones must be concentrated on stone-like states
- Observable control systems must be concentrated on set points

So far all we have said is that, *via* the interactions across a shared boundary, coupled random dynamical systems estimate each others statistics

Ultimately: any 'thing' encodes a probability distribution over possible environmental states. . . because the environment must be conducive to it existing

Question: why bother?

Answer: complex systems are difficult to understand because of their interactions, so replacing couplings with the study of variational free energy is fruitful

So far all we have said is that, *via* the interactions across a shared boundary, coupled random dynamical systems estimate each others statistics

Ultimately: any 'thing' encodes a probability distribution over possible environmental states. . . because the environment must be conducive to it existing

Question: why bother?

Answer: complex systems are difficult to understand because of their interactions, so replacing couplings with the study of variational free energy is fruitful

So far all we have said is that, *via* the interactions across a shared boundary, coupled random dynamical systems estimate each others statistics

Ultimately: any 'thing' encodes a probability distribution over possible environmental states. . . because the environment must be conducive to it existing

Question: why bother?

Answer: complex systems are difficult to understand because of their interactions, so replacing couplings with the study of variational free energy is fruitful

So far all we have said is that, *via* the interactions across a shared boundary, coupled random dynamical systems estimate each others statistics

Ultimately: any 'thing' encodes a probability distribution over possible environmental states... because the environment must be conducive to it existing

Question: why bother?

Answer: complex systems are difficult to understand because of their interactions, so replacing couplings with the study of variational free energy is fruitful

References

- Aaronson, *NP-complete problems and physical reality*, 2005. On the arXiv and in press here.
- Andrews, *The Math & the Territory: On the Scientific Uses & Abuses of Machine Learning & Other Mathematical Modeling Strategies*, 2025. Viewable here.
  - (And many of their other works; see their Google scholar page and website.)
- Friston, Ramstead, Sakthivadivel, *A framework for the use of generative modelling in non-equilibrium statistical mechanics*, 2026. On the arXiv and in press here.
- Levin and Resnik, *Mind everywhere: A framework for conceptualising goal-directedness in biology and other domains*, 2026. Parts 1 and 2. Quite possibly found here and perhaps even here.
- Levin, *Technological approach to mind everywhere: An experimentally-grounded framework for understanding diverse bodies and minds*, 2022. In press here.

Further notes

Much more of Levin's work can be found on his website; see especially the theoretical work here.

In

- Pressé, Ghosh, Lee, Dill, *Reply to C. Tsallis' comment on our "Nonadditive entropies yield probability distributions with biases not warranted by the data"*, 2015

an excellent point is made about maximising entropy *vs.* modelling entropy maximisation (arXiv:1504.01822).

Some of my *worked example* paper may be insightful—though I no longer view certain turns of phrases and symbology in as favourable a light as I was compelled to do then—nonetheless the concepts are laid out in a useful way, especially with regard to synchrony and least action principles. See the expanded paper on the arXiv and the references therein.

Further further notes

Refer to volume 384 issue 2320 of Philosophical Transactions of the Royal Society A entitled *World models in natural and artificial intelligence*, 2026, collecting discussions of points related to what forms of cognition exist where. I was very pleased to assist Levin and Safron in the preparation of this volume and thank them for the opportunity. Find it at this link. (All publications in this issue are open access.)

Look out for the issue on agency being prepared, as well as our forthcoming joint work on adaptation and multiscale synchrony which further develops his theory of integrated world modelling.

Regarding maximum entropy

I highly recommend the 1957 paper of Jaynes in Physical Review, *Information theory and statistical mechanics*, and the later *Macroscopic prediction* manuscript. Copies of both are viewable here.

For a proof that maximising entropy is directly a least action principle under geometrised constraints (as above for Lagrangian mechanics) I cannot resist suggesting my results in *Entropy-maximising diffusions satisfy a parallel transport law*, 2022. I offer an account from estimation theory and asymptotics of probability measures of why basic statistical mechanics leads to a good (in fact, optimal, in a certain precise sense) rule for statistical inference in *The relation of bias with risk in empirically constrained inferences*, 2025. Find the former here and the latter here.