How experimental requirements shape the mathematics of the laws of physics

Christine Aidala + Gabriele Carcassi

Physics Department
University of Michigan

ICASU Inaugural Conference
UIUC
May 19, 2022

Assumptions of Physics

https://icasu.illinois.edu/about

From the Director

In 2020, the College of Engineering at the University of Illinois Urbana-Champaign supported the creation of an intellectual home for interdisciplinary research aimed at addressing fundamental physics questions related to the universe. This research sits at the intersection of mathematics, computer science, astronomy and physics disciplines, such as nuclear physics, high energy physics, gravitation and cosmology. The ICASU is primarily an effort in theory and computation, but it is special in that all such work is driven by potential verification through experiment and observation. As such, the ICASU collaborates closely with experimental scientists in condensed matter and nuclear experiment, as well as with astrophysical observers in gravitational waves, particle astrophysics and traditional electromagnetic observations.

The project

Assumptions of Physics

https://assumptionsofphysics.org/

Gabriele Carcassi

Christine A. Aidala

University of Michigan

Introduction

- Past successes in fundamental physics have been reached through mathematical ideas (e.g. Einstein equations, Dirac equation, Yang-Mills theories, Higgs mechanism, …)
- This trend has been continuing (e.g. supersymmetry, supergravity, string theory, …) producing theories that are more esoteric and, sometimes, disconnected from experimental verification
- As a reaction, there is now an increased emphasis on developing theories that make verifiable predictions
- We believe we need to go further: the operational requirements of experimental verification are the actual foundational aspect of physics and they must play a central role in physical theories
    - We shouldn’t just look for theories that give verifiable predictions (not a side feature); they must incorporate the requirements of experimental verifiability at their core (the driving point)

Outline

- Show how we can "elevate" the discussion from mathematical constructs to physical principles, assumptions and requirements (**reverse physics**)

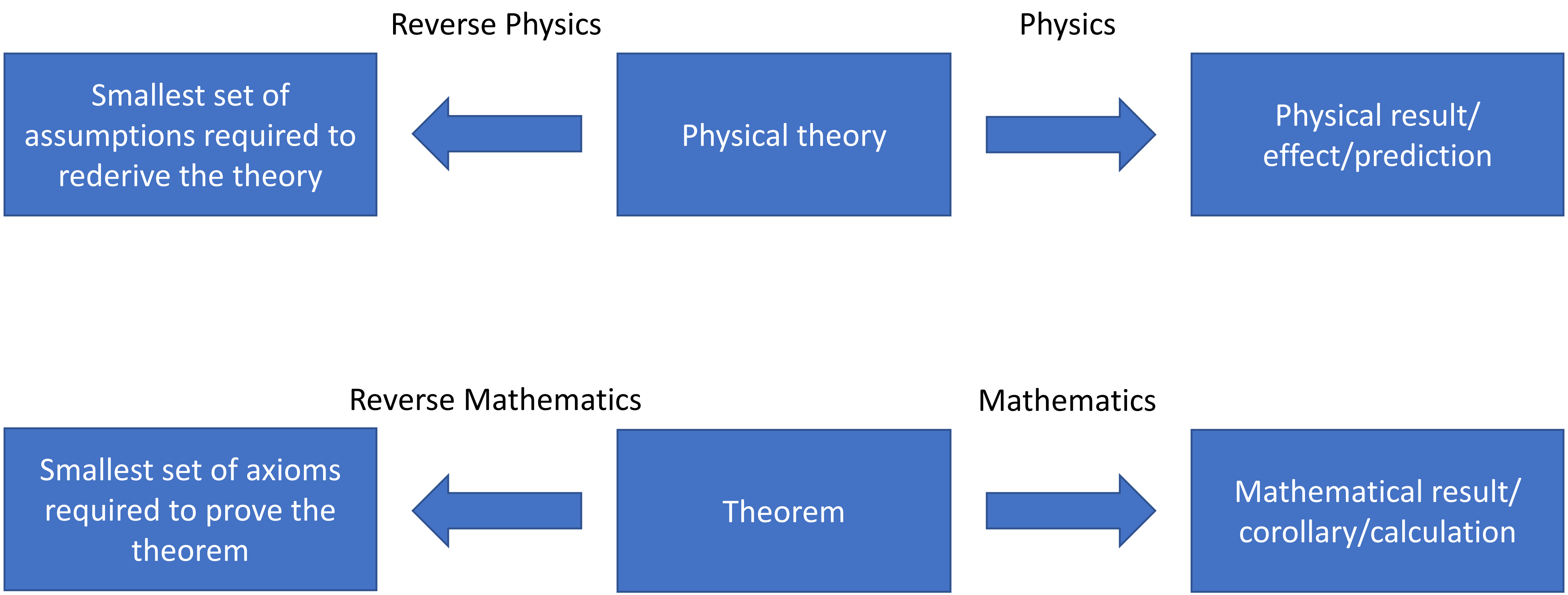
- Show that the current mathematical foundations are not quite what we need for physical theories (need for **physical mathematics**)

- See what new ideas can come out of this new outlook

Reverse Physics: from laws to physical assumptions

Reverse Physics: From Laws to Physical Assumptions

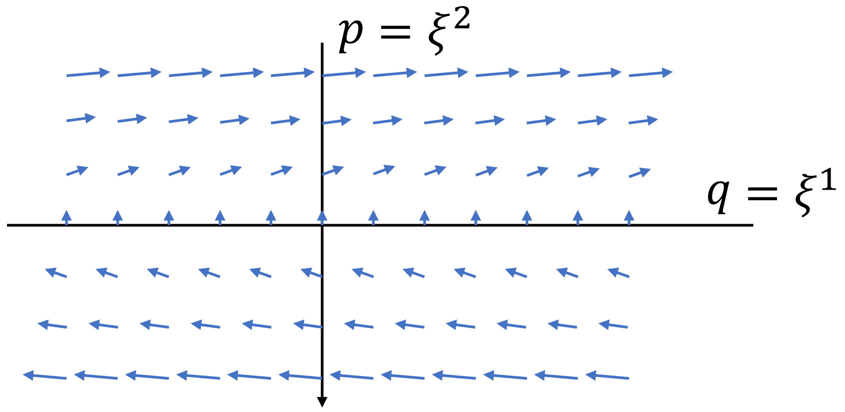
Gabriele Carcassi, Christine A. Aidala
Foundations of Physics (2022) 52:40


Reverse Physics
Physics
Smallest set of assumptions required to rederive the theory
Physical theory
Physical result/ effect/prediction
Reverse Mathematics
Mathematics
Smallest set of axioms required to prove the theorem
Theorem
Mathematical result/ corollary/calculation

(1) Hamilton's equations

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} = S^q$$

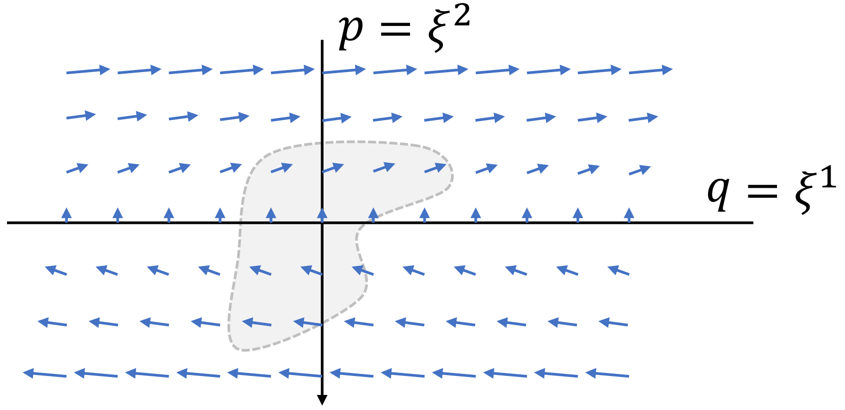
$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} = S^p$$



$$\xi^a = \{q, p\}$$

$$S^a = \frac{d\xi^a}{dt} = \left\{\frac{dq}{dt}, \frac{dp}{dt}\right\}$$

(2) Divergenceless displacement



Suppose S^a divergenceless

$$div(S^a) = \frac{\partial S^q}{\partial q} + \frac{\partial S^p}{\partial p} = 0$$

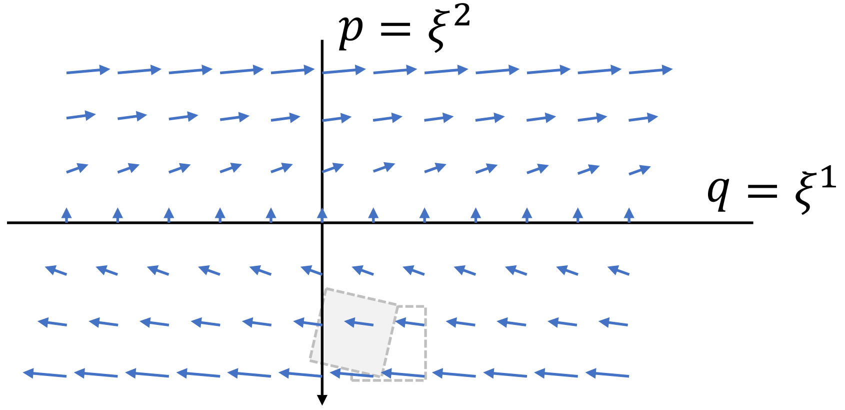
Then there exists a stream function H such that

$$\left\{\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q}\right\} = S^a = \frac{d\xi^a}{dt} = \left\{\frac{dq}{dt}, \frac{dp}{dt}\right\}$$

(3) Area conservation ($|J| = 1$)

Study how the area evolves

$$dQdP = |J|dqdp$$

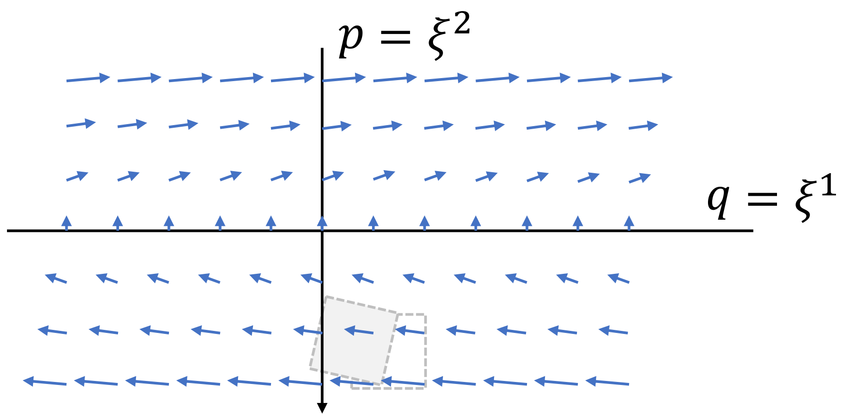


$$|J| = \begin{vmatrix} \frac{\partial Q}{\partial q} & \frac{\partial Q}{\partial p} \\ \frac{\partial P}{\partial q} & \frac{\partial P}{\partial p} \end{vmatrix} = \begin{vmatrix} 1 + \frac{\partial S^q}{\partial q}dt & \frac{\partial S^q}{\partial p}dt \\ \frac{\partial S^p}{\partial q}dt & 1 + \frac{\partial S^p}{\partial p}dt \end{vmatrix}$$

$div(S^a) = 0$

$$= 1 + \left(\frac{\partial S^q}{\partial q} + \frac{\partial S^p}{\partial p}\right)dt + O(dt^2)$$

(4) Deterministic and reversible evolution



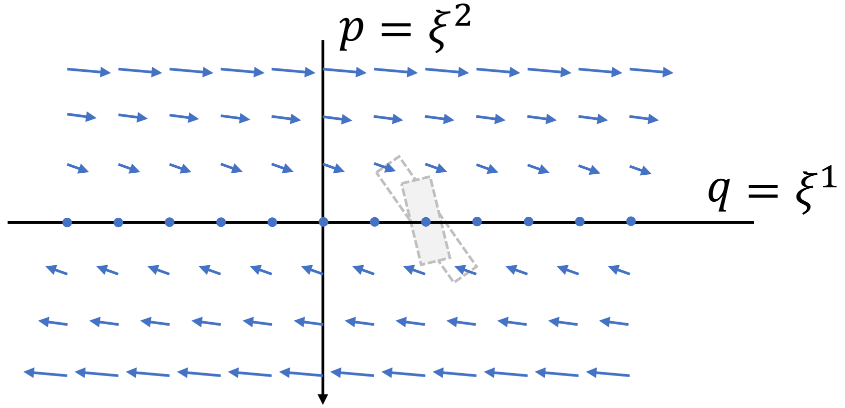
Statistical mechanics ⇒ use areas in phase space to count states

Area conservation ⟺ state count conservation

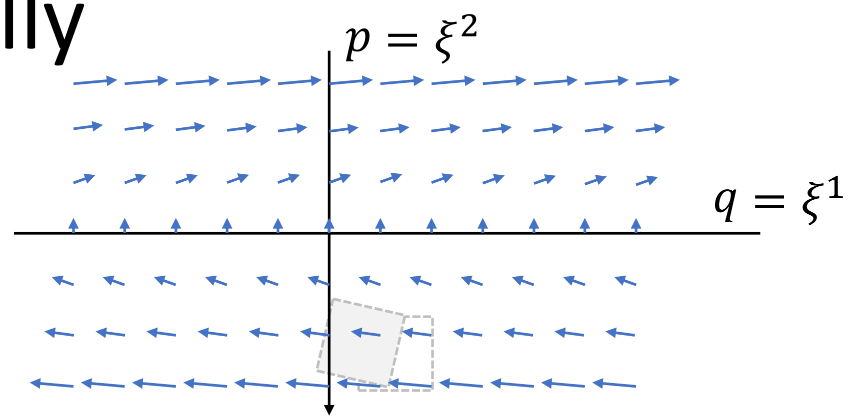
⟺ deterministic and reversible evolution

Key insight: det/rev is not just a bijection!
On continuous spaces, counting points is not enough!

A dissipative force maps points to points, but areas become smaller.



(5) Deterministic and thermodynamically reversible evolution



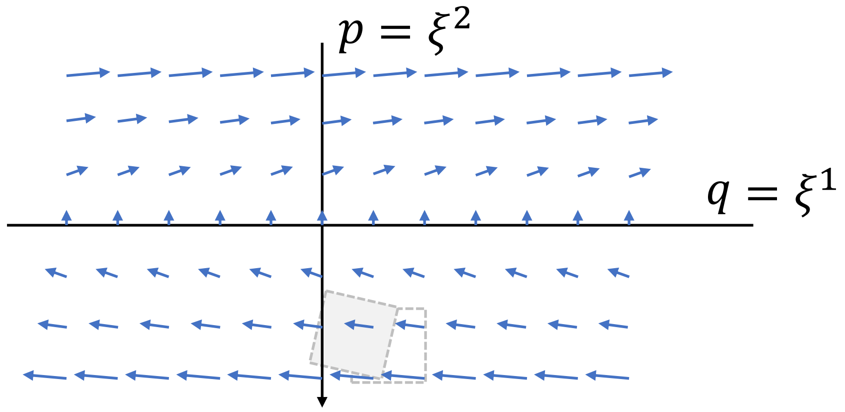
Link between statistical mechanics and thermodynamics

$$S = k_B \log W$$

Area conservation $\Longleftrightarrow$ entropy conservation

$\Longleftrightarrow$ thermodynamically reversible evolution

(6) Information conservation



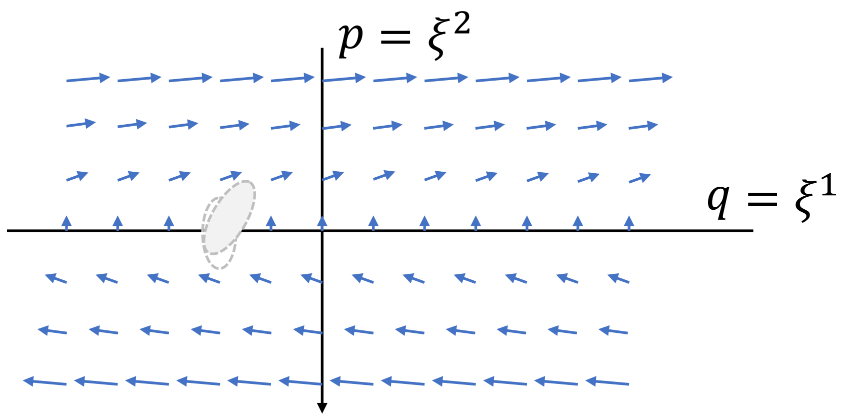
What about information entropy?

$$I[\rho(q,p)] = -\int \rho \log \rho \, dq dp$$

$$I[\rho(t+dt)] = I[\rho(t)] - \int \rho \log |J| \, dq dp$$

Area conservation $\Leftrightarrow$ information conservation

(7) Uncertainty conservation



What about uncertainty?

covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_q^2 & cov_{q,p} \\ cov_{p,q} & \sigma_p^2 \end{bmatrix}$$

Assuming a "very narrow" distribution

$$|\Sigma(t + dt)| = |J||\Sigma(t)||J|$$

Area conservation $\Leftrightarrow$ uncertainty conservation

(1) Hamilton's equations

All equivalent!

(2) Divergenceless displacement

(3) Area conservation ($|J| = 1$)

They should be thought as a "conceptual cluster": a series of ideas that "belong together"

(4) Deterministic and reversible evolution (i.e. isolation)

"Inverse" of Liouville's theorem

(5) Deterministic and thermodynamically reversible evolution

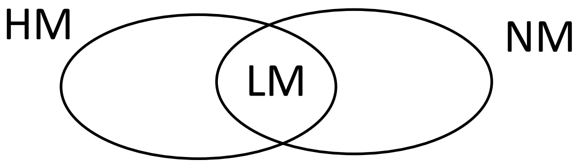
(6) Information conservation

(7) Uncertainty conservation

Three fundamental assumptions in Classical Mech

Infinitesimal Reducibility (IR) Determinism/Reversibility (D/R) Kinematic Equivalence (KE)

- IR ⇔ Classical phase space (symplectic manifolds ⇔ *unit independent* state count/densities/information entropy/thermodynamic entropy)
- IR+Directional degree of freedom ⇒ Space has three dimensions (2-sphere only symplectic manifold)
- IR+Directional degree of freedom ⇒ Classical analog for non-relativistic spin (open problem: relativistic analog)
- IR+D/R ⇔ Hamiltonian mechanics (Hamiltonian flow ⇔ conservation of state count/density/information entropy/thermodynamic entropy/dof independence)
- IR+D/R ⇒ energy-momentum co-vector, energy/Hamiltonian time component (pre-relativistic aspects w/o proper notion of space-time)
- IR+D/R ⇒ change of time variable changes the effective mass (similar to relativistic mass → rest mass scaled by time dilation)
- IR+D/R ⇒ classical antiparticles (w/o field theory, without quantum theory or full relativity/metric tensor)
- IR+D/R ⇒ classical uncertainty principle (uncertainty bound during evolution)
- IR+D/R ⇒ stationary action principle (with physical/geometrical interpretation, but w/o Lagrangian)
- IR+D/R+KE ⇒ Massive particles under scalar and vector potential forces
- IR+D/R+KE ⇒ $F^{\alpha\beta}$ is Poisson bracket between kinetic momenta; metric tensor as a geometrical feature of the tangent bundle ($dx^{\alpha} g_{\alpha\beta} du^{\beta}$); mass counts states per unit velocity; metric tensor locally flat (open problem: what about curvature?); speed of light converts count of possible time instants into number of possible spatial positions (i.e. ratio of measures, not speed).
- IR+D/R ⇔ Hamiltonian mechanics (HM); IR+KE ⇔ Newtonian mechanics (NM); IR+D/R+KE ⇔ Lagrangian mechanics (LM); LM = HM ∩ NM



Determinism and reversibility

$\Rightarrow$ existence and conservation of energy (Hamiltonian)

Why?

Stronger version of the first law of thermodynamics

Determinism and reversibility

$\Rightarrow$ past and future depend only on the state of the system

$\Rightarrow$ the evolution does not depend on anything else

$\Rightarrow$ the system is isolated

First law of thermodynamics!

$\Rightarrow$ the system conserves energy

Where does the bound on quantum uncertainty come from? Are there already other bounds in QM?

Let's look at the von Neumann entropy

$$I[\rho] = -tr(\rho \log \rho)$$

For a pure state $|\psi\rangle$

$$I[|\psi\rangle\langle\psi|] = 0$$

lowest possible entropy

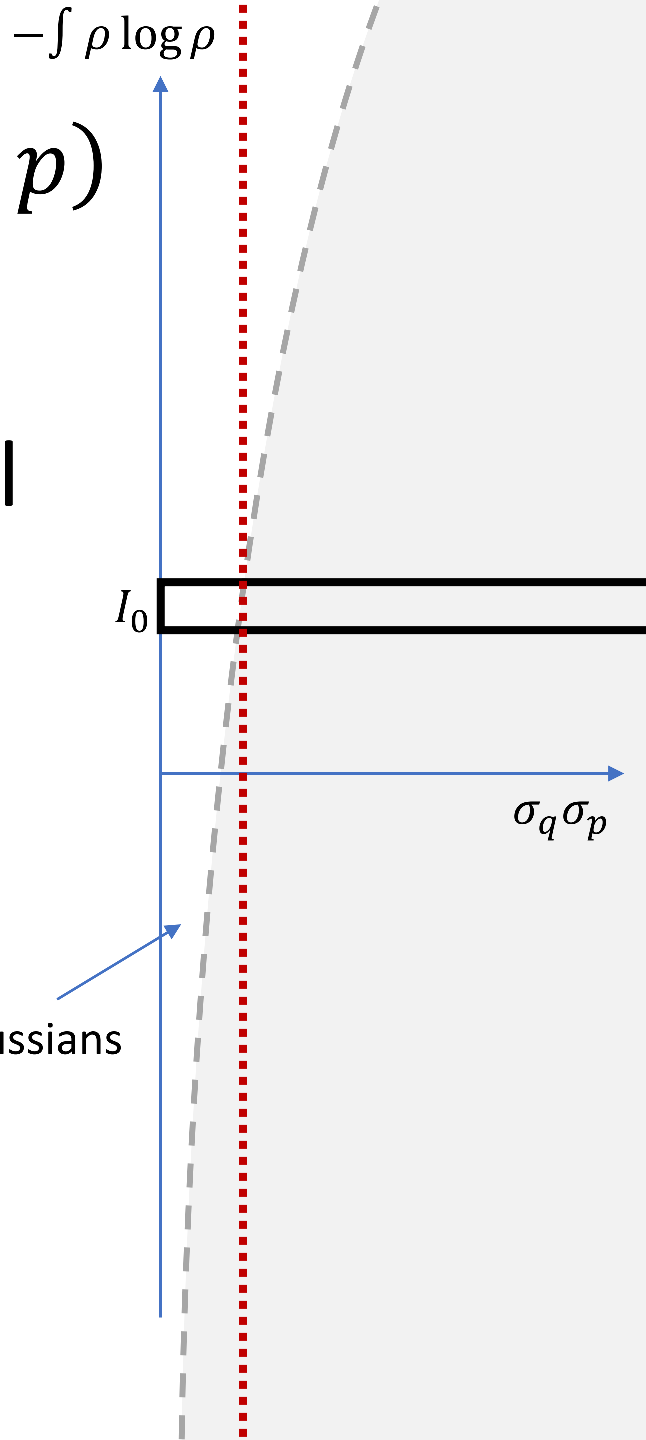
Could this bound, by itself, explain everything?

Take the space of all possible distributions $\rho(q,p)$ and order them by information/Gibbs entropy

Fix the entropy to a constant I_0 and consider all distributions with that entropy

They satisfy $\sigma_q\sigma_p \geq \frac{e^{I_0}}{2\pi e}$

equality for independent Gaussians



Lower bound on entropy ⇒ lower bound on uncertainty

Inverse does not work: lower bound on uncertainty does not give a lower bound on entropy

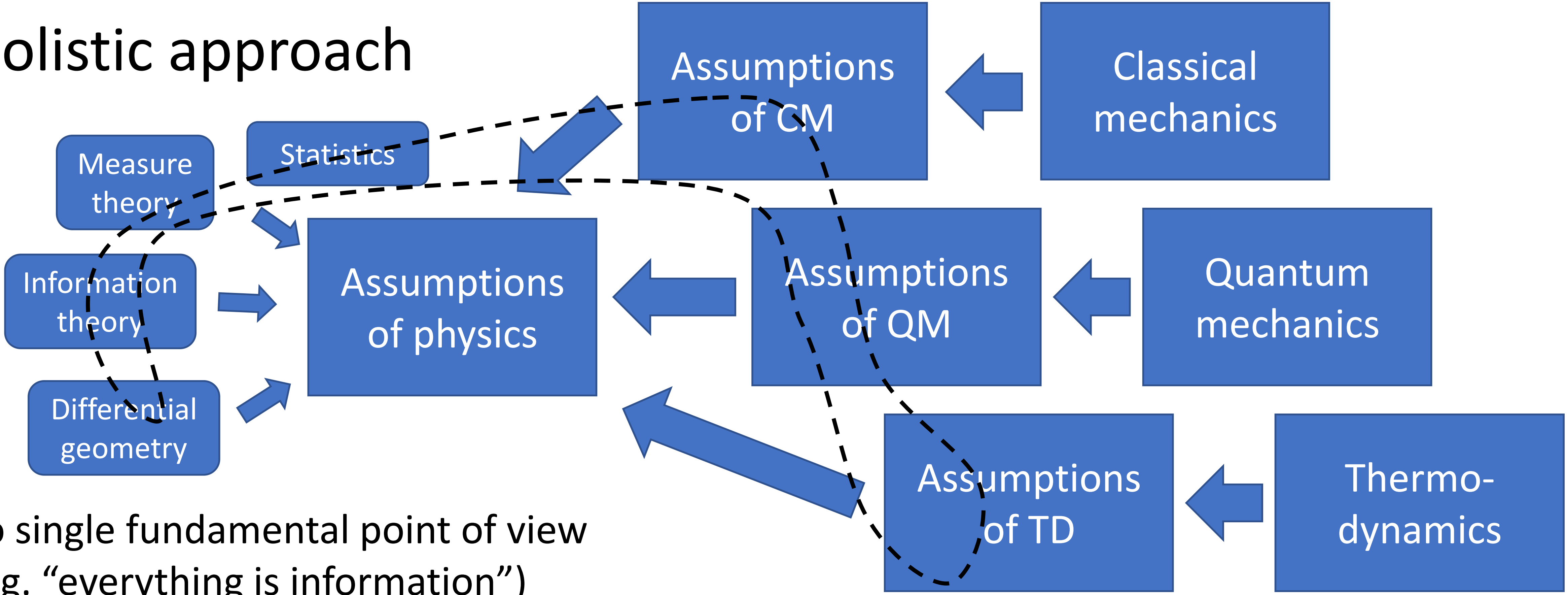
Lower bound for information entropy (Gibbs/von Neumann) ⇒ uncertainty principle (classical/quantum)

We don’t need the full quantum theory to derive the uncertainty principle: only the lower bound on entropy

The difference is that in classical mechanics we can prepare ensembles with arbitrarily low entropy…

… which is actually in contradiction with the third law of thermodynamics!!!

Holistic approach



No single fundamental point of view (e.g. “everything is information”)

Foundations of different theories are not disconnected

Find those “conceptual clusters” that span multiple areas of physics, math, …

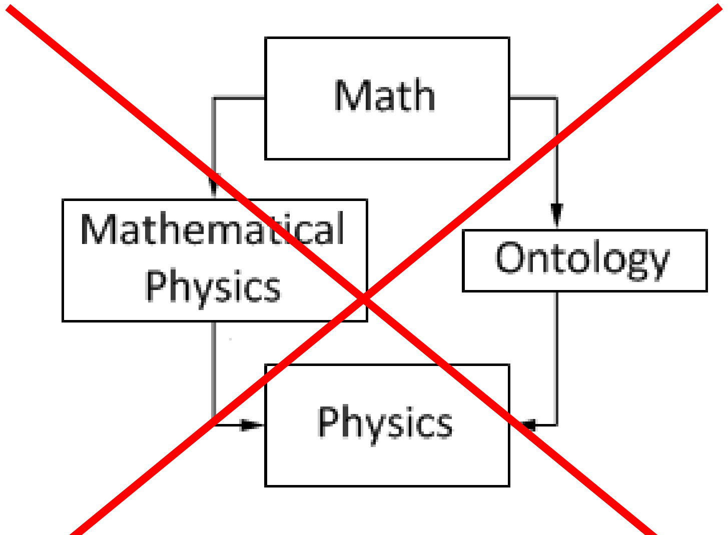
Reverse Physics

- "Reverse Physics" is an approach to the foundations of physics that starts from the physical laws and aims to "go back" to a suitable minimum number of physical assumptions
  - Reformulation – find better physical starting points.
  - Dependency analysis – find which part of a theory causes which effect.
  - Reconceptualization – substitute concepts with more general ones
- The goal is to fully map conceptual relationships and dependencies between different theories, different aspects of the theories, and to help foster higher level physical reasoning
- It is, by its nature, an interdisciplinary endeavor, and it can allow us to think more deeply about physical ideas and their relationships

Physical mathematics: from physical requirements to mathematical structures

In modern physics, mathematics is used as the foundation of our physical theories

From Hossenfelder's *Lost in Math*: "[…] finding a neat set of assumptions from which the whole theory can be derived, is often left to our colleagues in mathematical physics […]"



From Wikipedia "Mathematical Physics"

THIS CAN'T WORK!

But mathematics only deals with formal systems, without any connection to or concern about physical reality

David Hilbert: "Mathematics is a game played according to certain simple rules with meaningless marks on paper."

Bertrand Russell: "It is essential not to discuss whether the first proposition is really true, and not to mention what the anything is, of which it is supposed to be true."

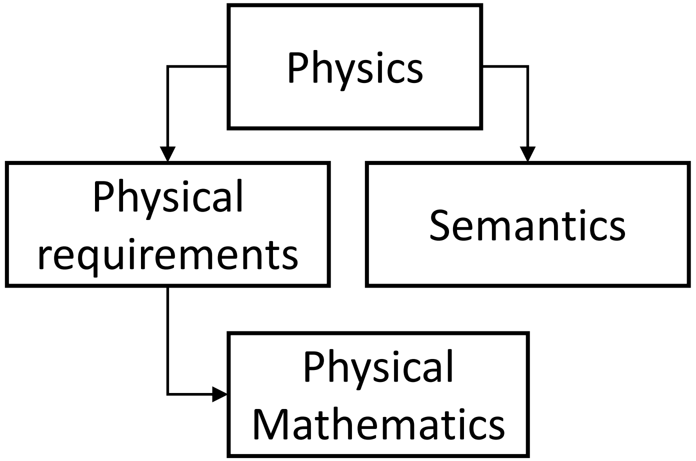
Formal definitions are neither necessary nor sufficient to do physics

Not useful in a lab

Moreover, there are choices at the foundations of mathematics

Do we accept the axiom of choice or the continuum hypothesis?

Physics is defined in terms of physical objects and operational definitions



Using assumptions and approximations, physical objects and their properties are idealized

The idealized model can then be expressed in the formal system

The idealization step is the most important part of this process, and it happens outside the formal system!

Starting with the math (i.e. the formal system) misses most of the physics

1^{st} basic requirement: experimental verifiability

Science deals with assertions whose truth can be defined/ascertained experimentally

⇒ **Verifiable statements**: assertions that can be experimentally verified in a finite time

The mass of the photon is less than 10^{-18} eV → Verifiable

The mass of the photon is exactly 0 eV → Not verifiable due to infinite precision, but falsifiable

T	SUCCESS (in finite time)
F	UNDEFINED
	FAILURE (in finite time)

Different logic of verifiable statements:

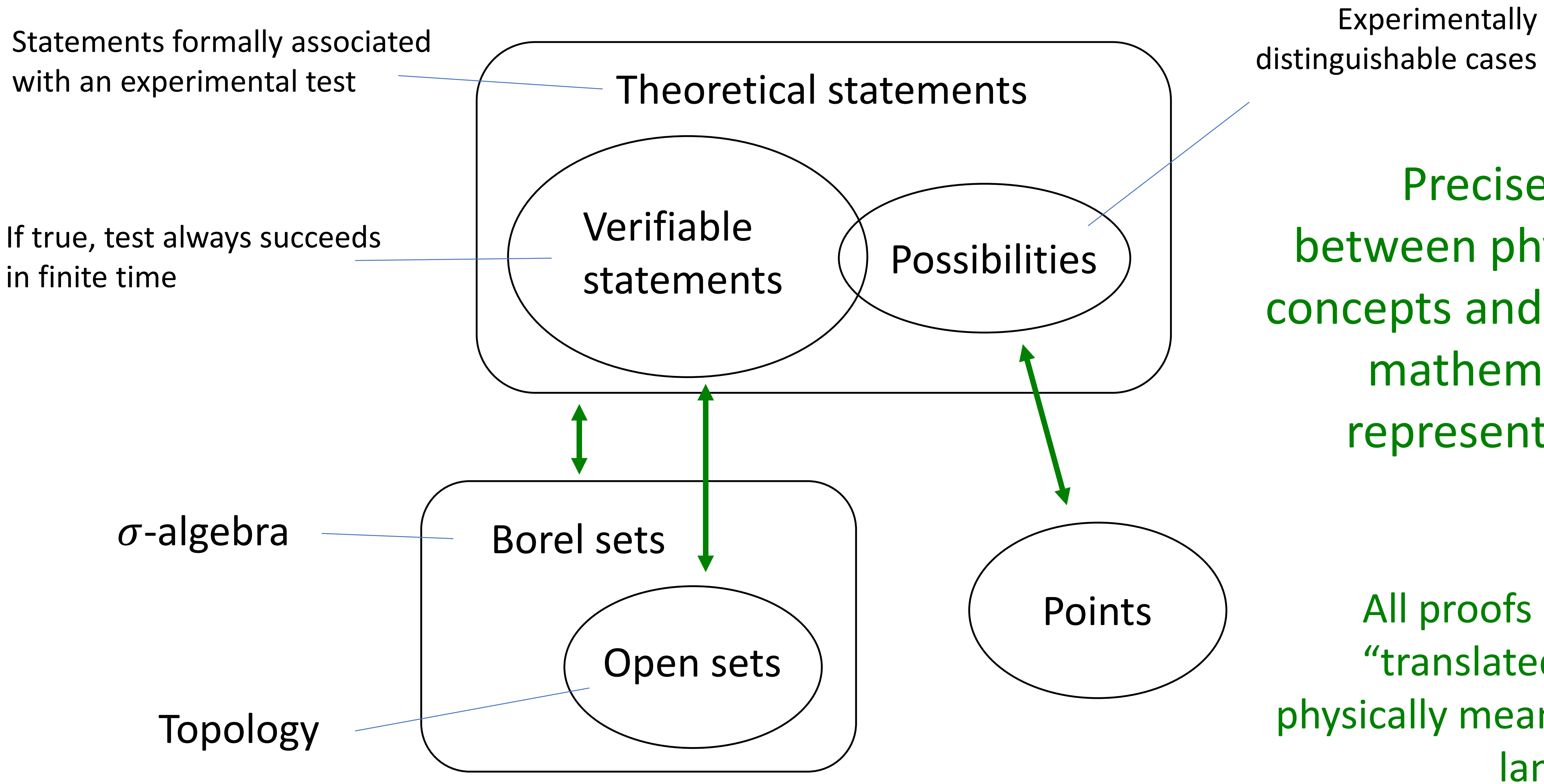
- Finite conjunction/logical AND (all tests must succeed in finite time)
- Countable disjunction/logical OR (once one test succeeds, we can stop)
- No negation/NOT (FALSE ≠ FAILURE)

Note: whether a specific statement is experimentally verifiable or even well defined may depend on context (e.g. premises, idealization, theory, etc…)

The mass of the electron is 511 ± 0.1 KeV

When measuring the mass, it is a verifiable hypothesis

When performing particle identification, it is assumed to be true



Precise map between physical concepts and their mathematical representation

All proofs can be "translated" into physically meaningful language

How do we define physical objects, e.g. time?

Formal definition (in math): some set with some properties (e.g. a variable that can be used as a parameter for the evolution of a system)

Ontological definition (in philosophy): some intrinsic feature of reality (e.g. continued sequence of existence and events)

These types of answers do not help us in a lab

An operational definition is necessary and sufficient: time is what is measured by a clock

What is a clock?

examples

The sun, the seasons, heart (pulse), a pendulum, ….

…, and anything else that can be synchronized to another clock.

operational feature of a general clock

⇒ If you can't synchronize clocks, you have a problem…

How do we formally model a clock?

What operational requirements lead to time as a real number?

A reference (i.e. a tick of a clock) is something that allows us to distinguish between a before and an after (mathematically, a triple of statements (b, o, a) with some properties)

A clock is a collection of references (synchronization ⇒ clocks have corresponding references)

To define an **ordered** sequence of "instants", the references must be (nec/suff conditions):
- Strict – an event is strictly before/on/after the reference (doesn't extend over the tick)
- Aligned – shared notion of before and after (logical relationship between statements)
- Refinable – overlaps can always be resolved

Additionally: between any two references we can have another reference ⇒ **real numbers**

These conditions are idealizations

How does this model of time break down?

The ticks of a clock have an extent and so do the events (references not strict)

If clocks have jitter, they cannot achieve perfect synchronization (references not aligned)

We cannot make clock ticks as narrow as we want (references not refinable)

No consistent **ordering**: no “objective” “before” and “after”

In relativity, different observers measure time differently, *but the order is the same*. We should expect this to fail at “small” scales.

A better understanding of space-time means creating a more realistic formal model that accounts for those failures

How do we define a system?

We can talk about a chair because we can manipulate it independently from the rest

- If the table moved every time the chair moved, we would talk about a chair-table system

We can talk about a chair here on the surface of the earth

- If we were on the surface of the sun, the chair would not exist for long

We can talk about a chair insofar that the internal dynamics is irrelevant

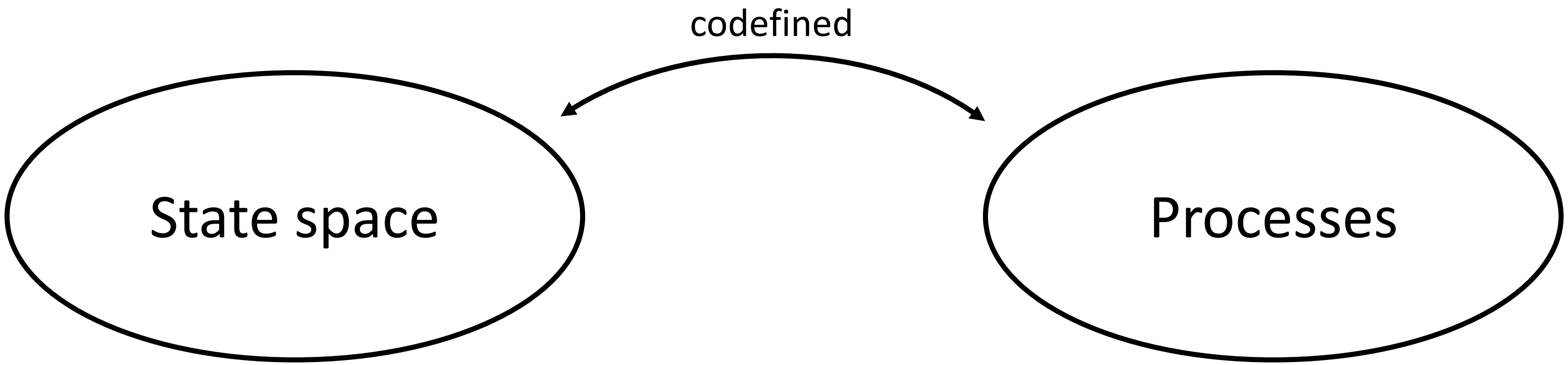
- If the detailed motion of each molecule were important, we would talk about a set of molecules

Defining a system is contingent upon the existence of processes that render it "independent" from the internal dynamics and the environment

Identifying conditions/processes for independent subsystems in QCD: factorization

- When we manipulate a proton as a whole, we are not generally sensitive to the internal dynamics of quarks and gluons. A proton is defined as a proton exactly when that internal dynamics is decoupled
- For inelastic proton collisions, sensitive to internal proton structure/dynamics, the QCD community has worked hard over the decades to identify (only a handful of) *factorizable* scattering processes in which it's a good approximation to think of independent subsystems—quarks and gluons—interacting
- But the vast majority of QCD processes cannot be treated as the interaction of simple independent subsystems
  - Are there identifiable subsystems within the proton at intermediate length/time scales?

Need a generalized theory of physical systems



State space must always be equipped with the processes under which the system is defined

Consistency requirements: state symmetries ↔ process symmetries; measurement processes ↔ open sets; system decoupling ↔ measure (and entropy) defined on states; …

The need for physical mathematics

- We can’t expect mathematicians to provide the formal structures we need for physics
  - they do not have enough understanding of the practical requirements of physics to create the appropriate abstractions
  - ⇒ the foundations of mathematics are not a good foundation for physics
- The proper foundation for physics is a conceptually consistent formal abstraction of **the practice of experimental science** (not “of the universe”)
  - We need to identify the formal structures that are appropriate to encode operational requirements and assumptions: **physically motivated mathematics**
- We can’t do this work without a deep understanding of how formal systems work, and how we can bridge the formal and informal parts
  - We need to understand which mathematical details to keep because they are physically relevant and which to “quotient out”
  - ⇒ we need a good understanding of the foundations of mathematics

New insights lead to new ideas

Measure theory plays a foundational role for theories of integration (e.g. geometrical sizes), probability and information theory: common physically motivated underpinning?

Consider the following statements:

"The position of the object is between 0 and 1 meters" and "The position of the object is between 2 and 3 kilometers"

"The fair die landed on 1" and "The fair die landed on 3 or 4"

"The first bit is 0 and the second bit is 1" and "The third bit is 0"

In all three cases, the first statement is "more precise", it is of a finer granularity (noted $\preccurlyeq$)

Constraining to a smaller volume gives finer description

Less likely events give more information

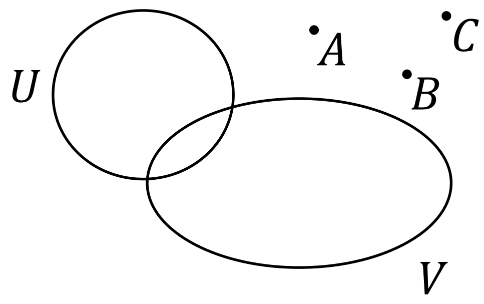
Statements with more information give a finer description

Comparing statements based on their granularity is another fundamental feature a physical theory must have

We need a generalized version of measure theory that covers all cases

Some statements are incomparable:

"The position of the object is between 0 and 1 meters" vs
"The velocity of the object is between 2 and 3 meters per seconds"

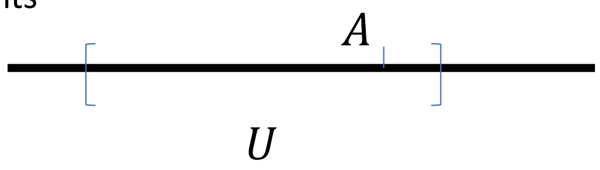


Comparability cannot be captured by a single measure:

$\{A\} \lesssim \{B, C\} \lesssim U \lesssim V$ while $\{A\} \not\gtrsim \{B, C\} \not\gtrsim U \not\gtrsim V$

Quantization breaks additivity:

Entropy in quantum mechanics is consistent with first two requirements



Single point is a single case (i.e. $\mu(\{A\}) = 1$)
Finite range carries finite information (i.e. $\mu(U) < \infty$)
Measure is additive for disjoint sets (i.e. $\mu(\cup U_i) = \sum \mu(U_i)$)

From what we understand, this is new mathematics

What could a generalized measure theory be useful for?

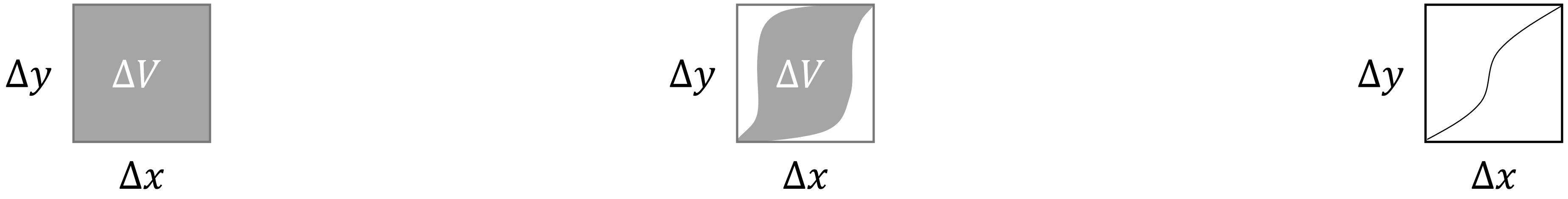
In a field theory, the value at each point is an independent d.o.f.

$\Rightarrow$ Measure of the volume "counts" the independent d.o.f.

Yet, in a singularity this can't be the case: value of the field at each point loses meaning;
Information encoded on the surface (holographic principle)

Flat space, zero curvature,
measure factorizes (i.e. $\Delta V = \Delta x \Delta y \Delta z$)

Singularity, infinite curvature,
"volume flattens"



What could a generalized measure theory be useful for?

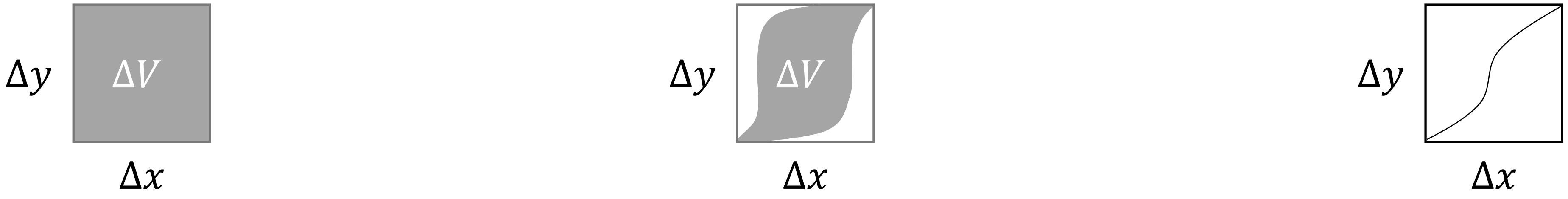
In a field theory, the value at each point is an independent d.o.f.

$\Rightarrow$ Measure of the volume "counts" the independent d.o.f.

Is the curvature an indicator for how independent the values of the fields are? Does "quantizing" space-time mean using a non-additive measure, so that the count of d.o.f. does not go to zero (but to a finite measure)?

Flat space, zero curvature,
measure factorizes (i.e. $\Delta V = \Delta x \Delta y \Delta z$)

Singularity, infinite curvature,
"volume flattens"

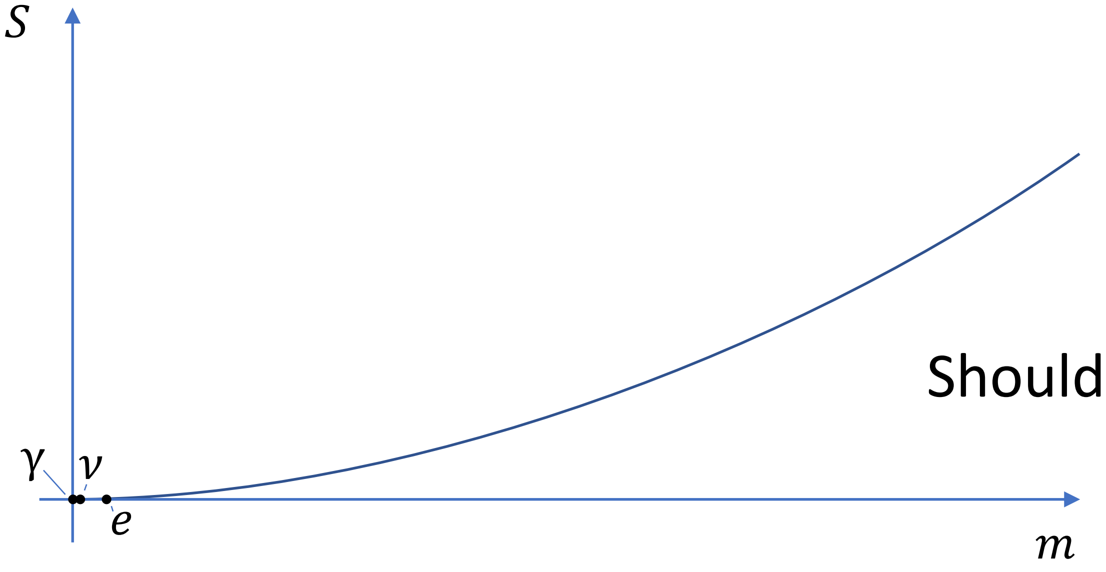


Fundamental particle:
object with no discernible substructure

BUT

Black hole:
has no discernible substructure

All pure states have zero entropy:
all fundamental particles
have zero entropy



$$S = \frac{k_B A}{4l^2} = 4\pi k_B \frac{M^2}{m_P^2}$$

Should same relationship apply?

Because relationship is quadratic,
zero entropy IS a good approximation
for small masses

Mass of $e \approx 10^{-30}$ Kg

Planck mass $\approx 10^{-8}$ Kg

BUT: macroscopic objects are made of MANY particles

Mass of Sun $\approx 10^{33}$ g

$\#e$ in 1 g of H $\approx 10^{23}$

$$S \approx \left(\frac{10^{-30}}{10^{-8}}\right)^2 10^{33} 10^{23} \approx 10^{12}$$

Other things to explore

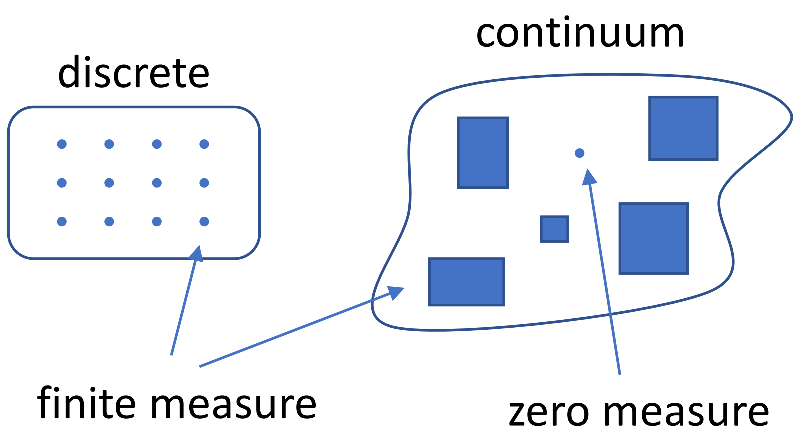
- In general relativity, the presence of mass/energy influences the definition of our units (though the metric tensor); what if it affects our clocks and rods in a deeper way? For example, the more energy is present, the harder it is to synchronize our clocks, the more time (and space) becomes fuzzy, to the point that it is no longer experimentally well defined (interior of a black hole)
- Consider the lattice of statements over a system at each time: a process will map one statement to another $\Rightarrow$ order theory as the most fundamental characterization of physical processes; fixed-point theorems
- …
- Any of these ideas start from a physical justification (not mathematical elegance or mathematical issues), and require us to develop new mathematics

Conclusion

- The solution to many open problems in the foundation of physics lies in a better understanding of the current mathematical tools, their physical meaning and the development of fundamentally new tools
- *Reverse physics* helps us reframe the current theories in terms of physical requirements and assumptions, shifting the attention away from math to physical ideas
- *Physical mathematics* helps us understand clearly how physical ideas are encoded into the formal systems, and find physically motivated generalizations
- In both cases, it is clear that what really drives the theoretical apparatus are the operational requirements, so that should be the focus of the foundations of physics

Supplemental

Quantifying discrete cases is fundamentally different than quantifying cases over the continuum



Why? Because fully identifying a discrete case requires finite information (finitely many experimental tests) while identifying a case from a continuum requires infinite information (an infinite sequence of increasingly precise tests)

This is something most physicists haven't yet fully digested

A single classical state in phase space (i.e. a microstate) ⇒ zero volume; minus infinite entropy; infinite information.

“Empty state” ⇒ one discrete case; zero entropy; finite information.

Quantum mechanics “fixes” this, by introducing a fixed lower bound on entropy.