Assumptions of Physics
Summer School 2024

States and Processes

Gabriele Carcassi and Christine A. Aidala

Physics Department
University of Michigan

Assumptions of Physics

https://assumptionsofphysics.org

Assumptions of Physics

Main goal of the project

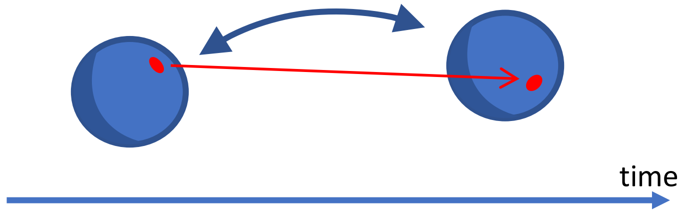
Identify a handful of physical starting points from which the basic laws can be rigorously derived

Assumptions of Physics

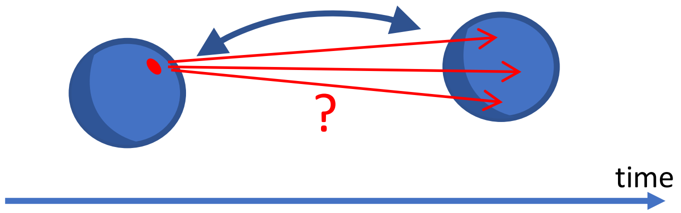
https://assumptionsofphysics.org

For example:

Infinitesimal reducibility ⇒ Classical state



Irreducibility ⇒ Quantum state



This also requires rederiving all mathematical structures from physical requirements

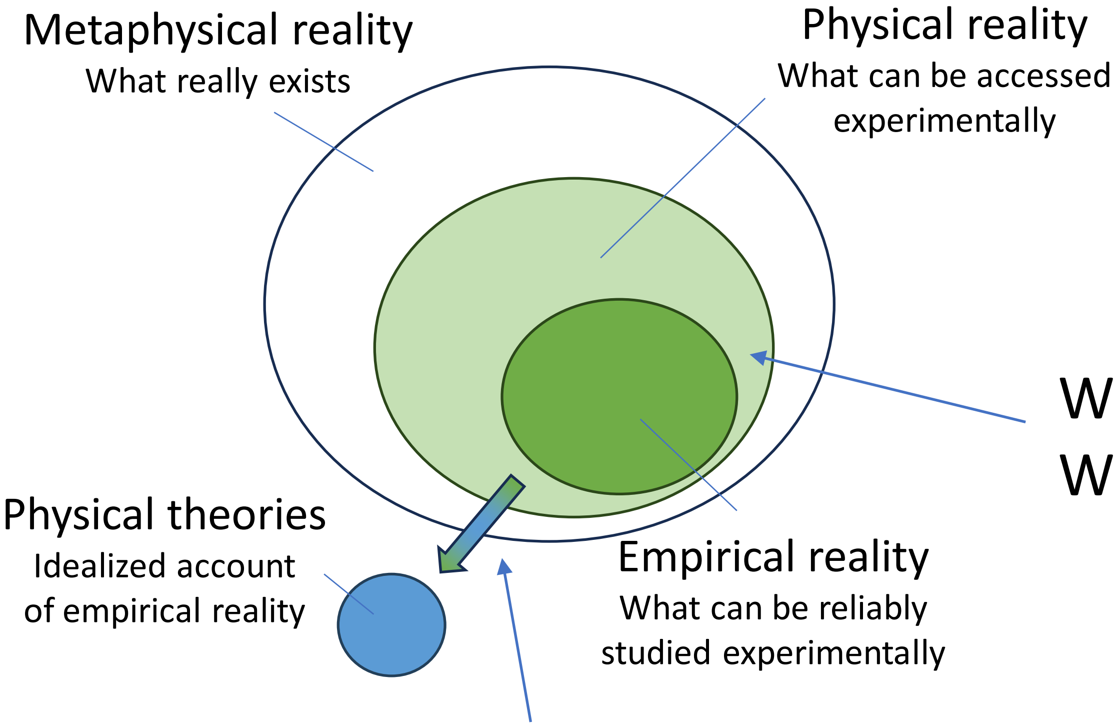
For example:

Science is evidence based ⇒ scientific theory must be characterized by experimentally verifiable statements ⇒ topology and σ-algebras

https://assumptionsofphysics.org/

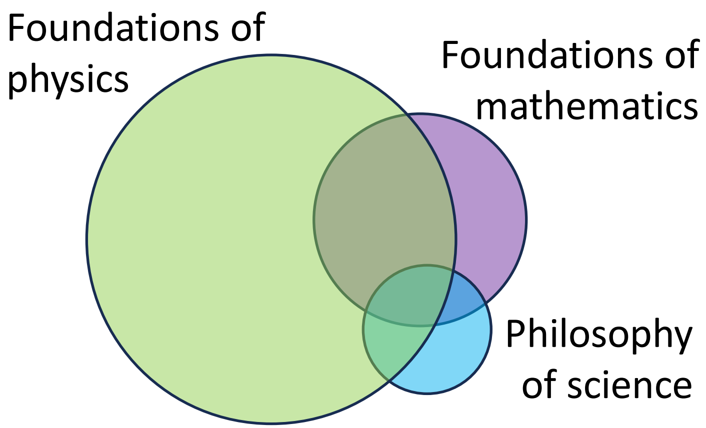
Assumptions of Physics

Underlying perspective



What is the boundary?
What are the requirements?

How exactly does the abstraction/idealization process work?



https://assumptionsofphysics.org/

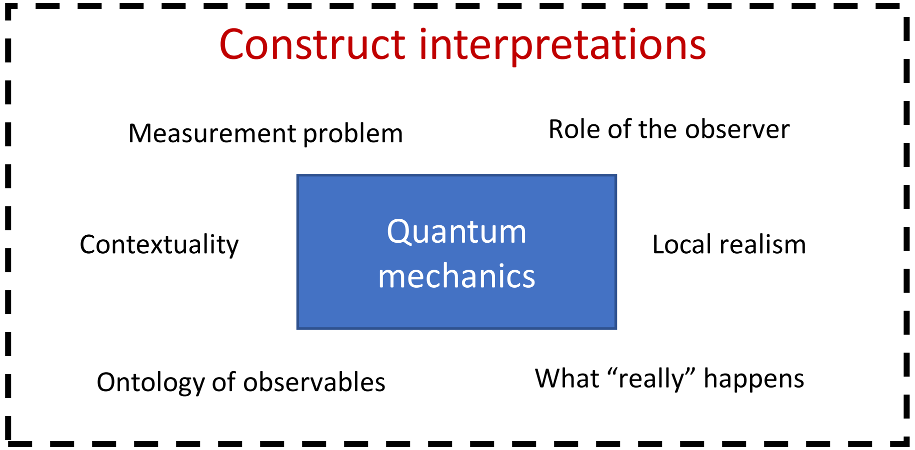
Assumptions of Physics

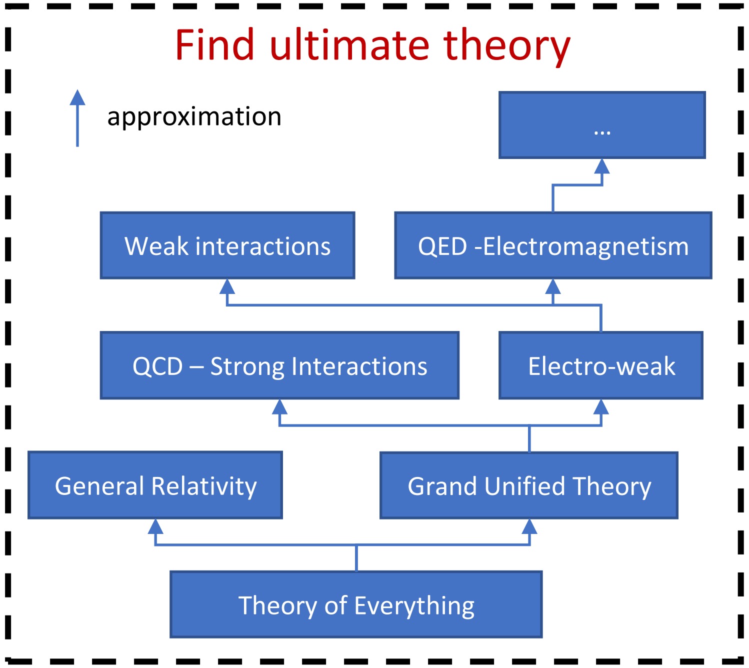
If physics is about creating models of empirical reality, the foundations of physics should be a theory of models of empirical reality

Requirements of experimental verification, assumptions of each theory, realm of validity of assumptions, …

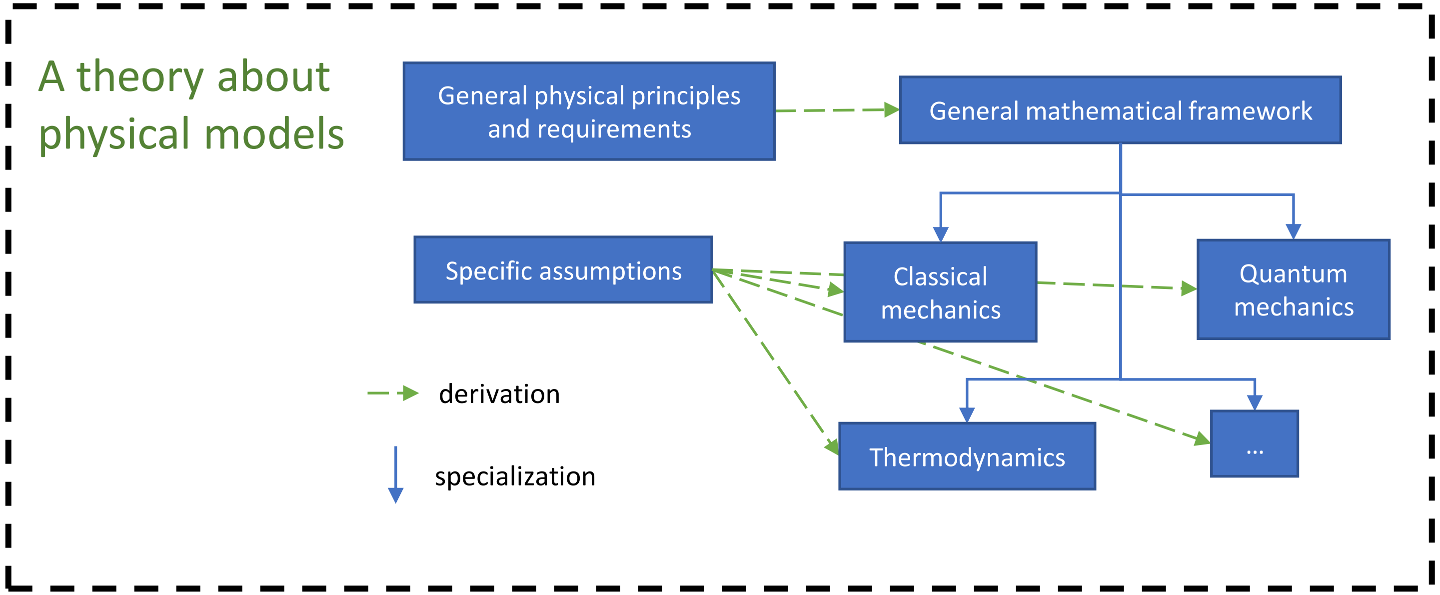
Different approach to the foundations of physics

Typical approaches





Our approach



https://assumptionsofphysics.org/

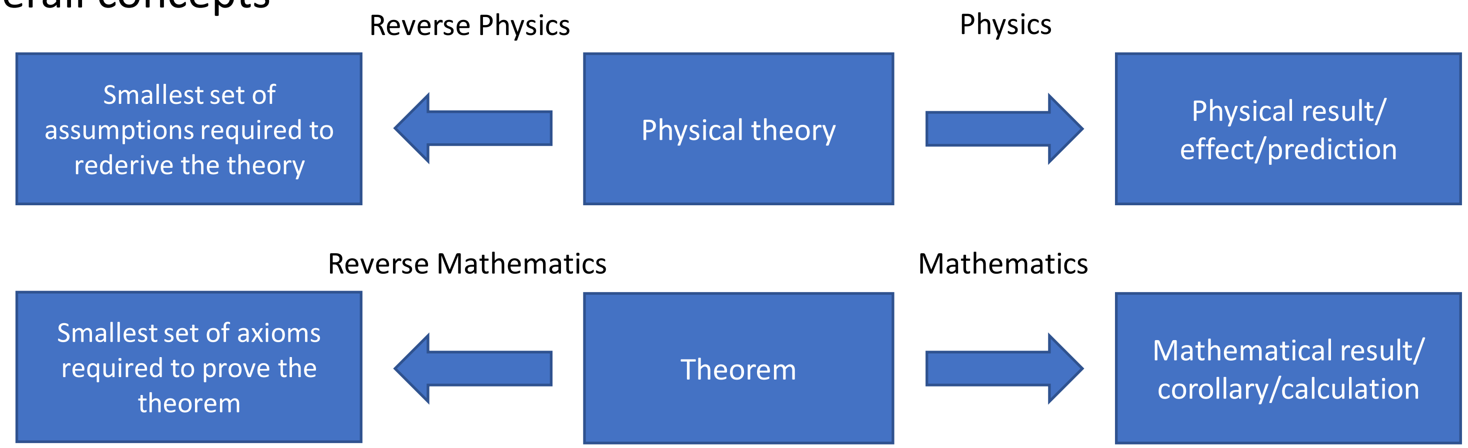
Assumptions of Physics

Find the right overall concepts

Reverse physics:

Start with the equations, reverse engineer physical assumptions/principles

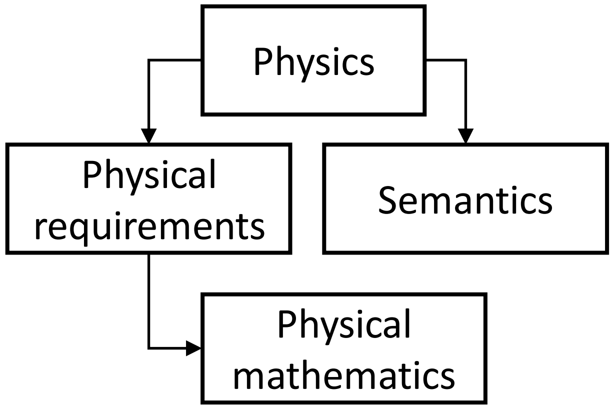
Found Phys **52**, 40 (2022)



Goal: find the right overall physical concepts, “elevate” the discussion from mathematical constructs to physical principles

Physical mathematics:

Start from scratch and rederive all mathematical structures from physical requirements



Goal: get the details right, perfect one-to-one map between mathematical and physical objects

https://assumptionsofphysics.org/

Assumptions of Physics

This session

Physical Mathematics: States and Processes

Assumptions of Physics*,*
Michigan Publishing (v2 2023)

https://assumptionsofphysics.org/

Assumptions of Physics

Space of ensembles $\mathcal{E}$

Ensembles can be mixed

$$e = pe_1 + (1-p)e_2$$

Convex structure

Assume unique complement

$$(e, p, e_1) \rightarrow e_2$$

Vector space structure

Assume differentiability

Sub-additive probability measure

$$p_e(U) = \sup\{\, p \mid pe_1 + (1-p)e_2 \,\}$$

Sub-additive "extent" measure

$$\mu(U) = \sup\{2^{\text{hull}(U)}\}$$

Ensembles must have an entropy well defined

$$S: \mathcal{E} \rightarrow \mathbb{R}$$

Strictly concave function on convex structure

Semi-metric

Jensen-Shannon divergence (JSD)

$$0 \leq S\left(\frac{1}{2}e_1 + \frac{1}{2}e_2\right) - \frac{1}{2}\left(S(e_1) + S(e_2)\right) \leq 1$$

Riemannian "manifold"

$$g_{ij} = -\partial_i \partial_j S$$

Fisher-Rao metric

Orthogonality

$e_1 \perp e_2$ if $\text{JSD}(e_1, e_2) = 1$

Subspaces

$$U = U^{\perp^{\perp}}$$

Contexts

$$U_1 \cap U_2 = 0 \Rightarrow U_1 \perp U_2$$

Additivity over contexts

Classical probability

https://assumptionsofphysics.org/

Assumptions of Physics

All names are placeholders

So feedback like “I wouldn’t call it that”, “the name is confusing”, … is not useful. We don’t even know what the right concepts are. Good naming is the final step.

Axioms of mixture and convex spaces

https://assumptionsofphysics.org/

Assumptions of Physics

Axiom 1.1 (Axiom of ensembles). *The state of a system is represented by an* ***ensemble****, which represents all possible preparations of equivalent systems prepared according to the same procedure. The set of all possible ensembles for a particular system is an* ***ensemble space****. Formally, an ensemble space is a T_0 second countable topological space where each element is called an ensemble.*

Can be defined experimentally

Physical laws are not about single instances: they are about reproducible relationships

$$F = ma$$

whenever this then that

Reproducibility already implies infinitely many copies
(you can always check one more time)

Also, preparations are never perfect
(can't prepare perfect initial conditions)

The "pure states"
(i.e. (q^i, p_i) and $P(\mathcal{H})$)
are idealized ensembles

⇒ Ensemble is the basic object for describing systems and states

https://assumptionsofphysics.org/

Assumptions of Physics

	States	Ensembles
Classical discrete	$X = \{x_1, x_2, \dots\}$	$\mathcal{E} = \{\, p_i \mid \sum_i p_i = 1\}$
Classical continuum	Phase space Symplectic manifold $X = \{\mathbb{R}^{2n}, \omega\}$	$\mathcal{E} = \{\, \rho \in C(\mathbb{R}^{2n}) \mid \int \rho \Pi_i dq^n dp_n = 1\}$
Quantum mechanics	Projective complex Hilbert space $X = P(\mathcal{H})$	$\mathcal{E} = \{$ positive semi-definite Hermitian, $\mathrm{tr}(\rho) = 1\,\}$

https://assumptionsofphysics.org/

Assumptions of Physics

Axiom 1.6 (Axiom of mixture). *An ensemble space $\mathcal{E}$ is equipped with an operation $+ : [0,1] \times \mathcal{E} \times \mathcal{E} \to \mathcal{E}$ called* ***mixing****, noted with the infix notation $p\mathsf{e}_1 + \bar{p}\mathsf{e}_2$, with the following properties:*

- ***Continuity****: the map $(p, \mathsf{e}_1, \mathsf{e}_2) \to p\mathsf{e}_1 + \bar{p}\mathsf{e}_2$ is continuous (the products $[0,1] \times \mathcal{E} \times \mathcal{E} \to \mathcal{E}$ have the product topology)*
- ***Identity****:* $1\mathsf{e}_1 + 0\mathsf{e}_2 = \mathsf{e}_1$
- ***Idempotence****:* $p\mathsf{e}_1 + \bar{p}\mathsf{e}_1 = \mathsf{e}_1$ *for all* $p \in [0,1]$
- ***Commutativity****:* $p\mathsf{e}_1 + \bar{p}\mathsf{e}_2 = \bar{p}\mathsf{e}_2 + p\mathsf{e}_1$ *for all* $p \in [0,1]$
- ***Associativity****:* $p_1\mathsf{e}_1 + \bar{p_1}\left(\overline{\left(\frac{p_3}{\bar{p_1}}\right)}\mathsf{e}_2 + \frac{p_3}{\bar{p_1}}\mathsf{e}_3\right) = \bar{p_3}\left(\frac{p_1}{\bar{p_3}}\mathsf{e}_1 + \overline{\left(\frac{p_1}{\bar{p_3}}\right)}\mathsf{e}_2\right) + p_3\mathsf{e}_3$ *for all* $p_1, p_3 \in [0,1]$

Ensembles can be mixed

⇒ Convex structure

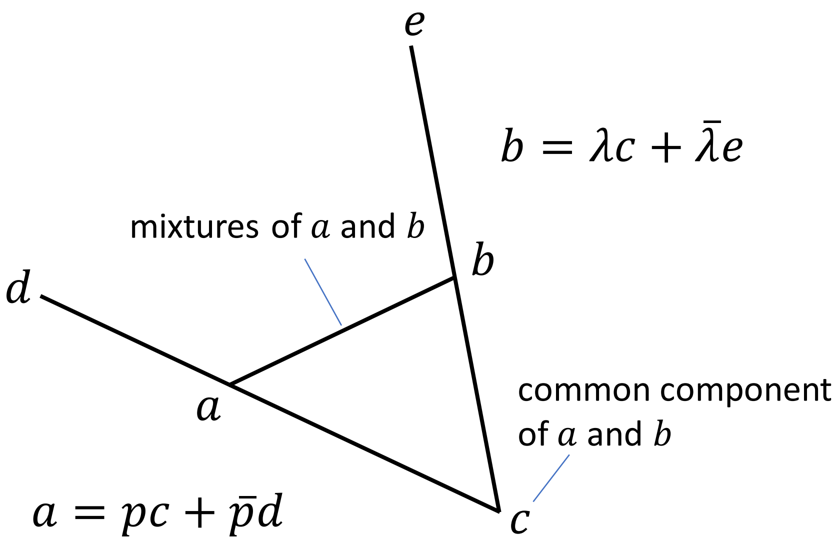
Classical probability distributions and quantum density operators have a convex structure

https://assumptionsofphysics.org/

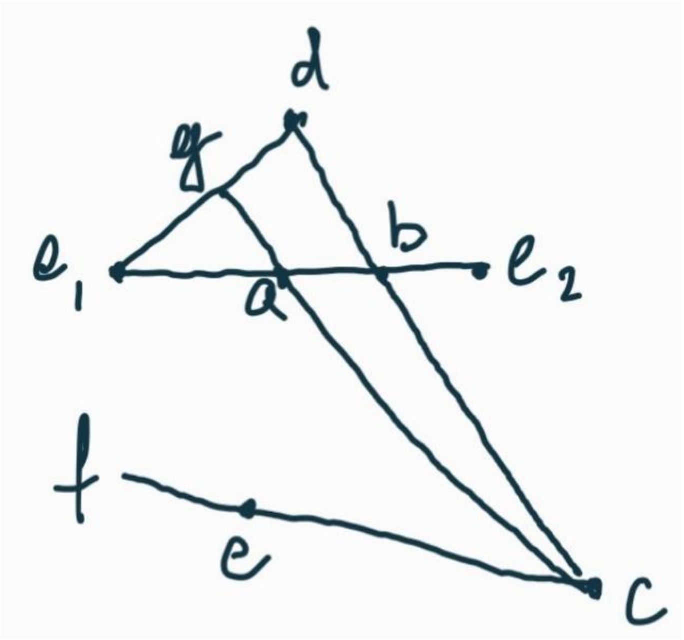
Assumptions of Physics

Definition 1.8. *Let $\mathcal{E}$ be an ensemble space. Let $\rho = \sum_i p_i \mathsf{e}_i$ where $\rho, \{\mathsf{e}_i\} \in \mathcal{E}$ and $p_i \in (0,1]$ such that $\sum p_i = 1$. We say that ρ is a* ***mixture*** *of $\{\mathsf{e}_i\}$ and each e_i is a* ***component*** *of ρ.*

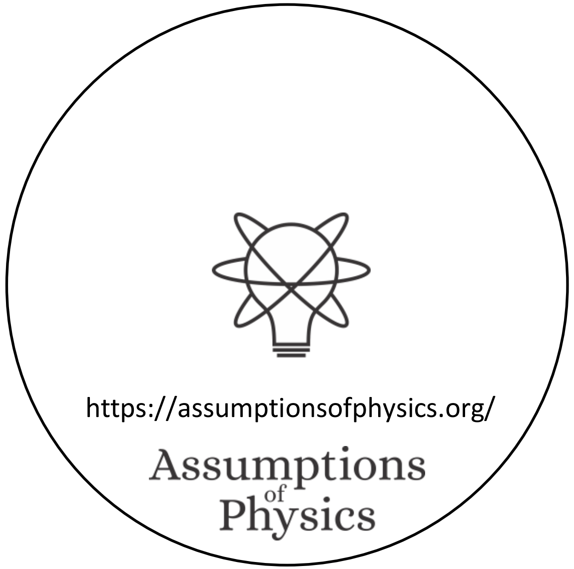
Definition 1.9. *Let $\mathcal{E}$ be an ensemble space and $\mathsf{e}_1, \mathsf{e}_2 \in \mathcal{E}$. We say that they* ***have a common component*** *if we can find $\mathsf{e}_3 \in \mathcal{E}$ such that $\mathsf{e}_1 = p_1\mathsf{e}_3 + \bar{p_1}\mathsf{e}_4$ and $\mathsf{e}_2 = p_2\mathsf{e}_3 + \bar{p_2}\mathsf{e}_5$ for some $\mathsf{e}_4, \mathsf{e}_5 \in \mathcal{E}$ and $p_1, p_2 \in (0,1)$. They are* ***distinct****, noted $\mathsf{e}_1 \perp \mathsf{e}_2$, otherwise.*



Proposition 1.11. *Let* $\mathsf{e}, \mathsf{e}_1, \mathsf{e}_2 \in \mathcal{E}$*. If* e *is distinct from a mixture of* e_1 *and* e_2 *then it is distinct from all mixtures of* e_1 *and* e_2 *and from both* e_1 *and* e_2*. That is, if* $\mathsf{e} \perp p\mathsf{e}_1 + \bar{p}\mathsf{e}_2$ *for some* $p \in (0,1)$ *then* $\mathsf{e} \perp p\mathsf{e}_1 + \bar{p}\mathsf{e}_2$ *for all* $p \in [0,1]$*. However, if* e *is distinct from* e_1 *and* e_2*, it is not necessarily true that* e *is distinct from a mixture of* e_1 *and* e_2*.*

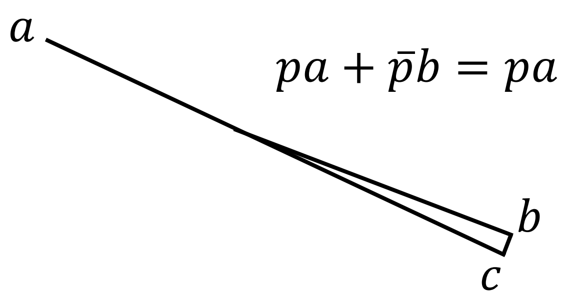


Suppose e distinct from a but not b



Definition 1.12. *Let $\rho = p\mathsf{e}_1 + \bar{p}\mathsf{e}_2$ with $p \in (0,1)$. Then we say that e_2 is a p-**complement** of e_1 towards ρ. An ensemble space is **complemented** if all p-complements are unique for all $p \in (0,1)$.*

Proposition 1.13. *An ensemble space is complemented if and only if it is a convex subset of a real vector space for which mixtures are linear combination.*



$$pa + \bar{p}b = pa + \bar{p}c$$

The mixing axioms allow us to get the same mixture just by changing one component

Requiring unique complement (i.e. cancellation axiom, unique inverse) recovers vector spaces

Both classical and quantum mechanics are complemented, but at this point it is not clear whether it is a necessary axiom

Takeaways

- Ensemble mixing provides a convex structure
- Invertibility of mixture recovers convex spaces
- TODOs:
  - Gather useful results for convex spaces
  - Understand how to recover topological vector spaces

Entropy

https://assumptionsofphysics.org/

Assumptions of Physics

Definition 1.16. *Given the coefficients* $\{p_i\} \in [0,1]$ *such that* $\sum p_i = 1$*, the* ***entropy of the coefficients*** *(also known as Shannon entropy) is defined as* $I(\{p_i\}) = -\sum p_i \log p_i$.

Axiom 1.17 (Axiom of entropy). *An ensemble space* $\mathcal{E}$ *is equipped a function* $S : \mathcal{E} \to \mathbb{R}$ *called* ***entropy*** *with the following properties*

- ***Continuity***
- ***Strict concavity:*** $S(p_1\mathsf{e}_1 + p_2\mathsf{e}_2) \geq p_1 S(\mathsf{e}_1) + p_2 S(\mathsf{e}_2)$ *with the equality holding if and only if* $\mathsf{e}_1 = \mathsf{e}_2$
- ***Upper variability bound:*** $S(p_1\mathsf{e}_1 + p_2\mathsf{e}_2) \leq I(p_1, p_2) + p_1 S(\mathsf{e}_1) + p_2 S(\mathsf{e}_2)$*; if the equality hold then,* e_1 *and* e_2 *are* ***disjunct****, noted* $e_1 \perp\!\!\!\perp e_2$

Entropy must be strictly concave: it cannot decrease during mixing, and it stays the same only when mixing an ensemble with itself

Maximum entropy increase is when ensembles are "completely different" (disjunct); in that case, the increase is only given by the choice of the ensemble

Assuming entropy is the variability within the ensemble

https://assumptionsofphysics.org/

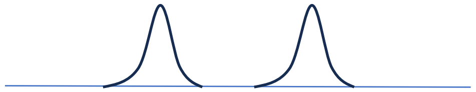
Assumptions of Physics

Definition 1.16. *Given the coefficients* $\{p_i\} \in [0,1]$ *such that* $\sum p_i = 1$, *the* ***entropy of the coefficients*** *(also known as Shannon entropy) is defined as* $I(\{p_i\}) = -\sum p_i \log p_i$.

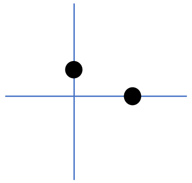
Axiom 1.17 (Axiom of entropy)**.** *An ensemble space* $\mathcal{E}$ *is equipped a function* $S : \mathcal{E} \to \mathbb{R}$ *called* ***entropy*** *with the following properties*

- ***Continuity***
- ***Strict concavity****:* $S(p_1\mathsf{e}_1 + p_2\mathsf{e}_2) \geq p_1 S(\mathsf{e}_1) + p_2 S(\mathsf{e}_2)$ *with the equality holding if and only if* $\mathsf{e}_1 = \mathsf{e}_2$
- ***Upper variability bound****:* $S(p_1\mathsf{e}_1 + p_2\mathsf{e}_2) \leq I(p_1, p_2) + p_1 S(\mathsf{e}_1) + p_2 S(\mathsf{e}_2)$*; if the equality hold then,* e_1 *and* e_2 *are* ***disjunct****, noted* $e_1 \perp\!\!\!\perp e_2$

In classical mechanics, disjunct ensembles correspond to probability distributions with disjoint support



In quantum mechanics, disjunct ensembles correspond to density operators in orthogonal subspaces



https://assumptionsofphysics.org/

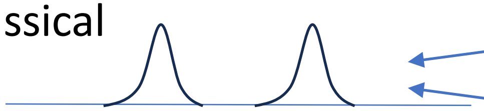
Assumptions of Physics

Definition 1.22. *An ensemble space $\mathcal{E}$ is* ***reducible*** *(or* ***classical****) if distinct ensembles are also disjunct. That is,* $\mathsf{e}_1 \perp \mathsf{e}_2$ *implies* $\mathsf{e}_1 \perp\!\!\!\perp \mathsf{e}_2$.

Proposition 1.23. *Continuous and discrete classical ensemble spaces are reducible.*

The difference between disjunctness and distinctness proves to be crucial

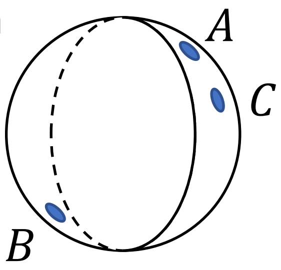
Classical



Distinct: no common subdistribution

Disjunct: equal mixture raises the entropy by 1

Quantum



All pure states are distinct (no common component)

Only orthogonal states are disjunct (equal mixture raises the entropy by 1)

The entropic structure (disjunctness) tells us how much the ensembles are similar or not
The convex structure (distinctness) tells us how much the common part can be separately studied

https://assumptionsofphysics.org/

Assumptions of Physics

Axiom 1.18 (Axiom of entropic disjunctness). *The entropy function of an ensemble space obeys the following properties*

- ***Disjunctness implies distinctness:*** $e_1 \perp\!\!\!\perp e_2$ *implies* $e_1 \perp e_2$
- ***Mixtures preserve disjunctness:*** *let* $e_1 = pe_2 + \bar{p}e_3$*, then* $e_4 \perp\!\!\!\perp e_1$ *if and only if* $e_4 \perp\!\!\!\perp e_2$ *and* $e_4 \perp\!\!\!\perp e_3$

Conceptually, disjunctness is a stronger property than distinctness

The above axioms can be justified from the physics

Can they be proved from the previous axioms or not?

Conjecture 1.19. *Let* $\mathsf{e}_1, \mathsf{e}_2 \in \mathcal{E}$ *be such that* $p\mathsf{e} + \bar{p}\mathsf{e}_1 = p\mathsf{e} + \bar{p}\mathsf{e}_2$ *for some* $p \in (0,1)$ *and* $\mathsf{e} \in \mathcal{E}$*. Then* e_1 *and* e_2 *are not disjunct.*

Remark. It seems very unlikely that differences in states that can be obscured by mixing would correspond to disjunct states. The more general question is whether the strict concavity of the entropy allows non-complemented spaces.

Is there any other interplay between convex structure and entropic structure?

https://assumptionsofphysics.org/

Assumptions of Physics

Takeaways

- Entropy provides an additional structure
- Disjunctness allows us to recognize orthogonal elements
- TODOs:
  - Better understand the interplay between convex and entropic structure
  - Entropy may be crucial to recover topological vector spaces

Dimensions of a subset of ensembles

https://assumptionsofphysics.org/

Classical statistical mechanics links count of states and entropy

$$S(\rho_U) = \log \mu(U)$$

Shannon/Gibbs entropy

Uniform distribution over U

Count of states

Fundamental postulate of statistical mechanics

Quantum statistical mechanics has a somewhat related expression

$$S(\rho_U) \leq \log(\dim_{\mathbb{C}}(\text{span}(U)))$$

Von Neumann entropy

Uniform distribution over U

Dimensionality of subspace

Equal for a uniform distribution over $\text{span}(U)$

Want a generalization of these relationships

https://assumptionsofphysics.org/

Assumptions of Physics

Proposition 1.24 (Exponential entropy subadditivity). *Let* $\mathsf{e}_1, \mathsf{e}_2 \in \mathcal{E}$. *Let* $S_1 = S(\mathsf{e}_1)$ *and* $S_2 = S(\mathsf{e}_2)$. *Let* $\mathsf{e} = p\mathsf{e}_1 + \bar{p}\mathsf{e}_2$ *for some* $p \in [0,1]$ *and* $S = S(\mathsf{e})$. *Then* $2^S \leq 2^{S_1} + 2^{S_2}$, *with the equality if and only if* e_1 *and* e_2 *are disjunct and* $p = \frac{2^{S_1}}{2^{S_1}+2^{S_2}}$.

Exponential of the entropy has key property

Proof. If p is fixed, the upper variability bound of entropy is saturated only if e_1 and e_2 are disjunct by definition. The entropy maximum for the mixed ensemble can only be achieved when the elements are disjunct, for some value of p.

$$
\begin{aligned}
0 = \frac{dS}{dp} &= \frac{d}{dp}S(\mathsf{e}) = \frac{d}{dp}\left(-p\log p - \bar{p}\log\bar{p} + pS_1 + \bar{p}S_2\right) \\
&= -\log p - 1 + \log\bar{p} + 1 + S_1 - S_2 \\
\log\frac{p}{\bar{p}} &= \log 2^{S_1} - \log 2^{S_2} \\
\log\frac{p}{1-p} &= \log\frac{2^{S_1}}{2^{S_2}} \\
p2^{S_2} &= (1-p)2^{S_1} \\
p(2^{S_1} + 2^{S_2}) &= 2^{S_1} \\
p &= \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}
\end{aligned}
$$

$$
\begin{aligned}
\bar{p} &= 1 - \frac{2^{S_1}}{2^{S_1} + 2^{S_2}} = \frac{2^{S_2}}{2^{S_1} + 2^{S_2}} \\
S &= S(\mathsf{e}) = -p\log p - \bar{p}\log\bar{p} + pS_1 + \bar{p}S_2 \\
&= -\frac{2^{S_1}}{2^{S_1} + 2^{S_2}}\log\frac{2^{S_1}}{2^{S_1} + 2^{S_2}} - \frac{2^{S_2}}{2^{S_1} + 2^{S_2}}\log\frac{2^{S_2}}{2^{S_1} + 2^{S_2}} \\
&\quad + \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}\log 2^{S_1} + \frac{2^{S_2}}{2^{S_1} + 2^{S_2}}\log 2^{S_2} \\
&= \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}\log\left(2^{S_1} + 2^{S_2}\right) + \frac{2^{S_2}}{2^{S_1} + 2^{S_2}}\log\left(2^{S_1} + 2^{S_2}\right) \\
&= \frac{2^{S_1} + 2^{S_2}}{2^{S_1} + 2^{S_2}}\log\left(2^{S_1} + 2^{S_2}\right) \\
\log 2^S &= \log\left(2^{S_1} + 2^{S_2}\right) \\
2^S &= 2^{S_1} + 2^{S_2}
\end{aligned}
$$

https://assumptionsofphysics.org/

Assumptions of Physics

Definition 1.27. *Let* $U \subseteq \mathcal{E}$ *be the subset of an ensemble space. The* **convex hull** *of* U*, noted* $\mathrm{hull}(U)$ *is the set of all possible mixtures that can be constructed with elements contained in* U*.*

Corollary 1.28. *The convex hull has the following properties*

1. $U \subseteq \mathrm{hull}(U)$
2. $U \subseteq V \implies \mathrm{hull}(U) \subseteq \mathrm{hull}(V)$
3. $\mathrm{hull}(\mathrm{hull}(U)) = \mathrm{hull}(U)$

and is therefore a closure operation

Definition 1.29. *Let $U \subseteq \mathcal{E}$ be the subset of an ensemble space. The* **dimension** *of U is defined as* $\dim(U) = \sup(2^{S(\mathrm{hull}(U))})$ *if* $U \neq \varnothing$ *and* $\dim(U) = 0$ *otherwise.*

Proposition 1.30. *The dimension is a set function that is*

1. *non negative:* $\dim(U) \in [0, +\infty]$
2. *monotone:* $U \subseteq V \implies \dim(U) \leq \dim(V)$
3. *subadditive:* $\dim(U \cup V) \leq \dim(U) + \dim(V)$
4. *additive over disjunct sets:* $U \perp\!\!\!\perp V \implies \dim(U \cup V) = \dim(U) + \dim(V)$

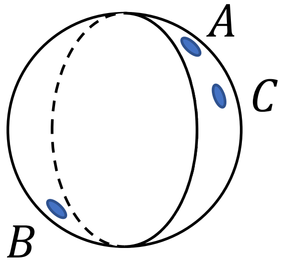
Dimension is the exponential of the highest entropy reachable through convex combinations

Recovers $S(\rho_U) \leq \log \dim(U)$

Highest entropy is reached
by uniform distribution
of disjunct elements

But it is not additive!

Need for non-additive measure



$\mu(\{A\}) = 2^0 = 1$

$\mu(\{A, B\}) = 2^1 = 2$

not additive

$\mu(\{A, C\}) < 2 = \mu(\{A\}) + \mu(\{C\})$

In quantum mechanics, literally $1 + 1 \leq 2$

	Single point		Finite continuous range	
	$\mu(U)$	$\log \mu(U)$	$\mu(U)$	$\log \mu(U)$
Counting measure $\mu(U) = \#U$ Number of points	1	0	$+\infty$	$+\infty$
Lebesgue measure $\mu([a, b]) = b - a$ Interval size	0	$-\infty$	$< \infty$	$< \infty$
"Quantized" measure $\mu(U) = \sup(2^{S(\mathrm{hull}(U))})$ Entropy over uniform distribution	1	0	$< \infty$	$< \infty$

Pick two!

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

Physically, we count states all else equal

Contextuality $\Longleftrightarrow$ non-additive measure

https://assumptionsofphysics.org/

Assumptions of Physics

Takeaways

- Upper entropy bound leads to a natural notion of "size" of a set which recovers statistical mechanics relationships in the general case
- This notion of size is, in general, not additive
- It is additive over disjunct sets
  - In classical mechanics, distinct = disjunct, so the measure is additive over pure states
  - In quantum mechanics, disjunct = orthogonal, so the measure is additive only over an orthogonal basis (i.e. measurement context)
- TODOs:
  - Better characterize the non-additivity

https://assumptionsofphysics.org/

Assumptions of Physics

Entropic geometry

https://assumptionsofphysics.org/

Assumptions of Physics

The space of classical statistical manifolds has a natural metric

The Fisher information metric then takes the form:[clarification needed]

$$g_{jk}(\theta) = -\int_R \frac{\partial^2 \log p(x,\theta)}{\partial \theta_j \, \partial \theta_k} p(x,\theta)\, dx.$$

There is a quantum analogue

The Bures metric may be defined as

$$[D_B(\rho, \rho + d\rho)]^2 = \frac{1}{2}\mathrm{tr}(d\rho G),$$

where G is the Hermitian 1-form operator implicitly given by

$$\rho G + G\rho = d\rho,$$

The Bures metric can be seen as the quantum equivalent of the Fisher information metric and can be rewritten in terms of the variation of coordinate parameters as

$$[D_B(\rho, \rho + d\rho)]^2 = \frac{1}{2}\mathrm{tr}\left(\frac{d\rho}{d\theta^\mu} L_\nu\right) d\theta^\mu d\theta^\nu,$$

Want a generalization of these objects

https://assumptionsofphysics.org/

Assumptions of Physics

Definition 1.31. *Given two ensembles* $\mathsf{e}_1, \mathsf{e}_2 \in \mathcal{E}$, *the* ***mixing entropy****, also called Jensen-Shannon divergence, is the increase in entropy associated to their mixture. That is:*

$$MS(\mathsf{e}_1, \mathsf{e}_2) = S\left(\frac{1}{2}\mathsf{e}_1 + \frac{1}{2}\mathsf{e}_2\right) - \left(\frac{1}{2}S(\mathsf{e}_1) + \frac{1}{2}S(\mathsf{e}_2)\right).$$

Corollary 1.32. *The mixing entropy obeys the following bounds*

$$0 \leq MS(\mathsf{e}_1, \mathsf{e}_2) \leq 1.$$

The lower bound is satisfied if and only if $\mathsf{e}_1 = \mathsf{e}_2$ *and the upper bound is satisfied if and only if* $\mathsf{e}_1 \perp\!\!\!\perp \mathsf{e}_2$.

It is a semi-metric (does not satisfy triangle inequality)

Proposition 1.33. *In discrete and continuous classical cases, the mixing entropy coincides with the Jensen-Shannon divergence. In quantum spaces it coincindes with the quantum Jensen-Shannon divergence.*

It generalizes the Jensen-Shannon divergence

Definition 1.34. *An ensemble space is **geometric** if it is complemented and has a twice differentiable entropy with respect to the mixing coefficients.*

Conjecture 1.35. *A geometric ensemble space is a convex subset of a real vector space. Moreover, every finite dimensional subspace is a smooth manifold.*

We need a vector space with a differentiable structure

https://assumptionsofphysics.org/

Assumptions of Physics

Definition 1.36. *Let $V \subseteq \mathcal{E}$ be a differentiable manifold embedded in the ensemble space. Let $\mathsf{e} \in \mathcal{E}$ be an ensemble and T_e the tangent space at that point. The **norm** of $\delta\mathsf{e} \in T_\mathsf{e}$ is given by*

$$\|\delta\mathsf{e}\|_\mathsf{e} = \sqrt{8MS(\mathsf{e}, \mathsf{e} + \delta\mathsf{e})}.$$

*The **metric tensor** (i.e. the inner product between $\delta\mathsf{e}_1, \delta\mathsf{e}_2 \in T_\mathsf{e}$) is given by*

$$g_\mathsf{e}(\delta\mathsf{e}_1, \delta\mathsf{e}_2) = \frac{1}{2}\left(\|\delta\mathsf{e}_1 + \delta\mathsf{e}_2\|_\mathsf{e}^2 - \|\delta\mathsf{e}_1\|_\mathsf{e}^2 - \|\delta\mathsf{e}_2\|_\mathsf{e}^2\right).$$

Recover a geometric structure in general

Theorem 1.37. *Let $\mathcal{E}$ be a geometric ensemble space and $V \subseteq \mathcal{E}$ a differentiable manifold embedded in the ensemble space. Then*

$$\|\delta\mathsf{e}\|_\mathsf{e}^2 = -\frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}, \delta\mathsf{e})$$

and

$$g_\mathsf{e}(\delta\mathsf{e}_1, \delta\mathsf{e}_2) = -\frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_1, \delta\mathsf{e}_2).$$

Metric tensor is just the Hessian of the entropy

Then V is a Riemannian manifold with g_e as the metric tensor and $\|\cdot\|_\mathsf{e}$ as the norm.

https://assumptionsofphysics.org/

Assumptions of Physics

Proofs are trivial

$$S(\mathsf{e}+\delta\mathsf{e}) = S(\mathsf{e}) + \frac{\partial S}{\partial \mathsf{e}}\delta\mathsf{e} + \frac{1}{2}\frac{\partial^2 S}{\partial \mathsf{e}^2}\delta\mathsf{e}\delta\mathsf{e} + O(\delta\mathsf{e}^3).$$

Expanding the definition of MS, we have

$$\begin{aligned}
MS(\mathsf{e},\mathsf{e}+\delta\mathsf{e}) &= S\left(\frac{1}{2}\mathsf{e} + \frac{1}{2}(\mathsf{e}+\delta\mathsf{e})\right) - \frac{1}{2}S(\mathsf{e}) - \frac{1}{2}S(\mathsf{e}+\delta\mathsf{e}) \\
&= S\left(\mathsf{e} + \frac{1}{2}\delta\mathsf{e}\right) - \frac{1}{2}S(\mathsf{e}) - \frac{1}{2}S(\mathsf{e}+\delta\mathsf{e}) \\
&= S(\mathsf{e}) + \frac{\partial S}{\partial \mathsf{e}}\frac{1}{2}\delta\mathsf{e} + \frac{1}{2}\frac{\partial^2 S}{\partial \mathsf{e}^2}\frac{1}{2}\delta\mathsf{e}\frac{1}{2}\delta\mathsf{e} + O(\delta\mathsf{e}^3) \\
&\quad - \frac{1}{2}S(\mathsf{e}) - \frac{1}{2}\left(S(\mathsf{e}) + \frac{\partial S}{\partial \mathsf{e}}\delta\mathsf{e} + \frac{1}{2}\frac{\partial^2 S}{\partial \mathsf{e}^2}\delta\mathsf{e}\delta\mathsf{e} + O(\delta\mathsf{e}^3)\right) \\
&= S(\mathsf{e}) + \frac{1}{2}\frac{\partial S}{\partial \mathsf{e}}\delta\mathsf{e} + \frac{1}{8}\frac{\partial^2 S}{\partial \mathsf{e}^2}\delta\mathsf{e}\delta\mathsf{e} \\
&\quad - S(\mathsf{e}) - \frac{1}{2}\frac{\partial S}{\partial \mathsf{e}}\delta\mathsf{e} - \frac{1}{4}\frac{\partial^2 S}{\partial \mathsf{e}^2} + O(\delta\mathsf{e}^3) \\
&= -\frac{1}{8}\frac{\partial^2 S}{\partial \mathsf{e}^2}\delta\mathsf{e}\delta\mathsf{e} + O(\delta\mathsf{e}^3).
\end{aligned}$$

Therefore

$$\|\delta\mathsf{e}\|^2 = 8MS(\mathsf{e},\mathsf{e}+\delta\mathsf{e}) = -\frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e},\delta\mathsf{e}).$$

$$\begin{aligned}
g_{\mathsf{e}}(\delta\mathsf{e}_1,\delta\mathsf{e}_2) &= \frac{1}{2}\left(\|\delta\mathsf{e}_1+\delta\mathsf{e}_2\|^2 - \|\delta\mathsf{e}_1\|^2 - \|\delta\mathsf{e}_2\|^2\right) \\
&= \frac{1}{2}\left(-\frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_1+\delta\mathsf{e}_2,\delta\mathsf{e}_1+\delta\mathsf{e}_2) + \frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_1,\delta\mathsf{e}_1) + \frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_2,\delta\mathsf{e}_2)\right) \\
&= -\frac{1}{2}\left(\frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_1,\delta\mathsf{e}_1) + \frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_1,\delta\mathsf{e}_2) + \frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_2,\delta\mathsf{e}_1) + \frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_2,\delta\mathsf{e}_2)\right. \\
&\quad \left. - \frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_1,\delta\mathsf{e}_1) - \frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_2,\delta\mathsf{e}_2)\right) \\
&= -\frac{\partial^2 S}{\partial \mathsf{e}^2}(\delta\mathsf{e}_1,\delta\mathsf{e}_2)
\end{aligned}$$

Proposition 1.41. *For a classical ensemble space, the metric corresponds to the Fisher-Rao metric.*

$$\log(x+dx) = \log(x) + d_x \log x \, dx + \frac{1}{2} d_x d_x \log x \, dx^2 + O(dx^3)$$
$$= \log(x) + \frac{1}{x} dx - \frac{1}{2}\frac{1}{x^2} dx^2 + O(dx^3)$$

$$S(\rho+\delta\rho) = -\int_X (\rho+\delta\rho)\log(\rho+\delta\rho)dx$$
$$= -\int_X (\rho+\delta\rho)\left[\log\rho + \frac{1}{\rho}\delta\rho - \frac{1}{2\rho^2}\delta\rho^2 + O(\delta\rho^3)\right]dx$$
$$= -\int_X \rho\log\rho dx - \int_x [\log\rho + 1]\, dx\delta\rho - \int_X \left[\frac{1}{\rho} - \frac{\rho}{2\rho^2}\right] dx\delta\rho^2 + \int_X dx O(\delta\rho^3)$$
$$= -\int_X \rho\log\rho dx - \int_x [\log\rho + 1]\, dx\delta\rho - \frac{1}{2}\int_X \frac{1}{\rho} dx\delta\rho^2 + \int_X dx O(\delta\rho^3)$$

Proofs are mere calculations

$$\frac{\partial^2 S}{\partial\rho^2}(\delta\rho,\delta\rho) = -\int_X \frac{1}{\rho}\delta\rho^2 dx = -\int_X \frac{1}{\rho}\delta\rho^2 dx + 0 = -\int_X \frac{1}{\rho}\delta\rho^2 dx + \delta^2(1)$$
$$= -\int_X \frac{1}{\rho}\delta\rho^2 dx + \delta^2 \int_X \rho dx = -\int_X \frac{1}{\rho}\delta\rho^2 dx + \int_X \delta^2 \rho dx$$
$$= \int_X \rho dx \left[-\frac{1}{\rho^2}\delta\rho^2 + \frac{1}{\rho}\delta^2\rho\right] = \int_X \rho dx\, \delta\left[\frac{1}{\rho}\delta\rho\right]$$
$$= \int_X \rho dx\, \delta^2 \log\rho$$

$$g_{\mathsf{e}}(d\theta^1, d\theta^2) = -\frac{\partial^2 S}{\partial\rho^2}\left(\frac{\partial\rho}{\partial\theta^1}d\theta^1, \frac{\partial\rho}{\partial\theta^2}d\theta^2\right) = -\int_X \rho dx \frac{\partial^2\log\rho}{\partial\theta^1\partial\theta^2} d\theta^1 d\theta^2$$

Takeaways

- The geometric structure of the ensemble space is due to the strict concavity of the entropy, and therefore it is fully general
- TODOs:
  - Recover the quantum case
  - In both the classical and quantum case, the JSD is the square of a distance function: can this be proven in general?

NOTE: differential geometry is limited to manifolds (i.e. finite dimensional spaces). Yet, what we find is more general.

Can we find a more general notion of differentiability, tangent spaces, etc… ?

https://assumptionsofphysics.org/

Assumptions of Physics

Differentiability

https://assumptionsofphysics.org/

Differentiability in math

Differentiable manifold	
Manifold	Differentiable structure

Mathematicians have developed several, increasingly abstract, definitions for differentials, derivatives, integrations, tangent vectors… are they suitable for physics?

Changes of coordinates are differentiable

Defined on top of Fréchet derivative

Vector defined as derivation of a scalar function

$$v: C^\infty(X, \mathbb{R}) \to C^\infty(X, \mathbb{R})$$
$$v(f) = v^i \partial_i f$$

vector basis

Does not make sense physically!

- velocity is not a derivation
- momentum is not a function of a derivation
- derivations ∂_i depend on units and can't be summed (e.g. $\partial_r + \partial_\theta$)
- Two mathematical notions of differentials (the new one and the one hidden in the Fréchet derivative)
- Infinitesimal objects are limits of finite objects, not the other way around

Differentials defined as linear functions of vectors

$$dx: V \to \mathbb{R}$$
$$dx(v) = dx\left(v^i \partial_i\right) = v^x$$

So are convectors, like momentum

Integrals defined on top of differential forms

$$\int_\gamma dx = \Delta x$$

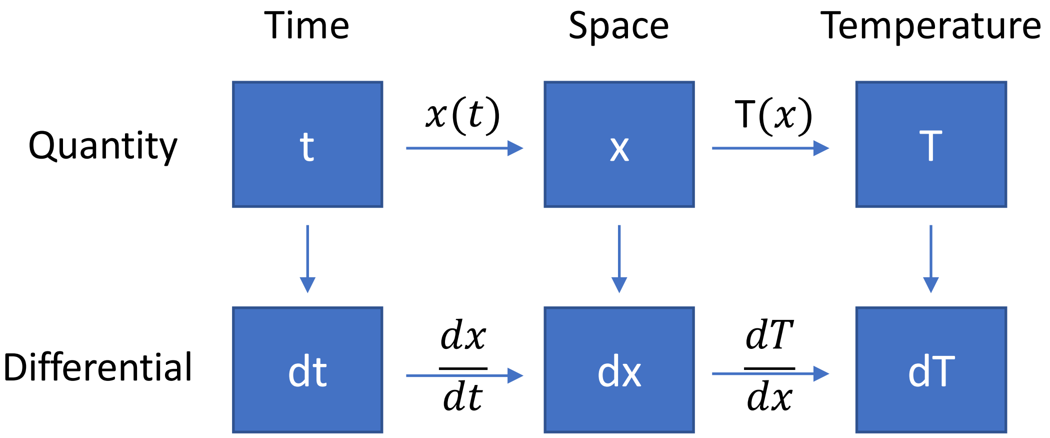
https://assumptionsofphysics.org/

Assumptions of Physics

Differentiability in physics

Infinitesimal reducibility ⇒ differentiability

General notion of differential as an infinitesimal change in ANY vector space



Differentiable function: infinitesimal changes map to infinitesimal changes

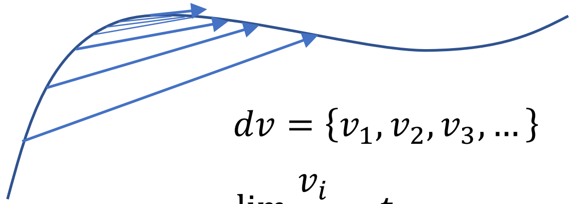
Differentiable space: infinitesimal changes are well-defined

Derivative: map between differentials

$$dx^i = \frac{dx^i}{dt} dt \qquad dT = \frac{\partial T}{\partial x^i} dx^i$$

velocity (vector)

gradient (covector)



$$dv = \{v_1, v_2, v_3, \dots\}$$

$$\lim_{i \to \infty} \frac{v_i}{a_i} = t$$

Tangent vector

Convergence at all points ⇒ differentiability of curve

Goal: one notion of derivative

$$dx^i e_i = dP$$

Manifold displacement (unit free)

Map between the two

Coordinate displacement (units of x^i)

https://assumptionsofphysics.org/

Assumptions of Physics

Definition 1.2 (Convergence envelope). *A **convergence envelope** $\{a_i\}_{i=1}^{\infty}$ is a sequence of non-zero elements of $\mathbb{R}$ that converges to 0.*

Defines how we go to zero

Definition 1.3. *Let V be a real vector space. A **differential** dv is a sequence of vectors $\{v_i\}_{i=1}^{\infty}$ such that there exists a vector $t \in V$ and a convergence envelope $\{a_i\}_{i=1}^{\infty}$ for which*

$$\lim_{i \to \infty} \frac{v_i}{a_i} = t.$$

*We call t the **tangent vector** of the differential and $\{a_i\}_{i=1}^{\infty}$ its **convergence envelope**. We note $dv[a_i\, t]$ the differential with its tangent vector and convergence envelope.*

Only requires the (topological) vector space structure

https://assumptionsofphysics.org/

Assumptions of Physics

Proposition 1.4. *Let dv be a differential. It can be expressed as*

$$v_i = a_i t_i = a_i(t + w_i)$$

where $\{t_i\}_{i=1}^{\infty}$ is a sequence of vectors that converges to t and w_i is a sequence of vectors that converges to 0.

Proposition 1.5. *Differentials respect the following property*

$$dv[a_i kt] = dv[ka_i t].$$

Some useful properties

https://assumptionsofphysics.org/

Assumptions of Physics

Proposition 1.6. *Differentials with the same convergence envelope form a vector space. That is*

$$b\,dv[a_i\,t] + c\,dv[a_i\,u] = dv[a_i\,(bt + cu)].$$

for any $t, u \in V$ *and* $b, c \in \mathbb{R}$.

Technically, we have a space of differentials for each a_i

Definition 1.7. *Let V and W be two vector spaces. Given a map $f : V \to W$, a sequence $\{v_i\}_{i=1}^{\infty}$ that converges to some $v \in V$ and a differential $dv[a_i t]$, we define the* ***image of the differential through the map*** *as $df(v_i, dv[a_i t]) = \{f(v_i + a_i t_i) - f(v_i)\}_{i=1}^{\infty}$. The map is* ***differentiable*** *at v if there exists a map $\left.\frac{df}{dV}\right|_v : V \to W$, called* ***derivative*** *such that $df(v_i, dv[a_i t]) = dw[a_i \left.\frac{df}{dV}\right|_v (t)]$ for all $\{v_i\}_{i=1}^{\infty}$ and for all differentials. That is, df maps differentials of V to differentials of W that have the same convergence envelope and a tangent vector that depends only on the original tangent vector.*

Differentiable maps are those that map differentials to differentials

$$v \mapsto f(v)$$

Only the tangent matters, not the convergence

$$a_i(t + \epsilon_i) \mapsto a_i \left(\frac{df}{dv}(t) + \zeta_i \right)$$

Same convergence

https://assumptionsofphysics.org/

Assumptions of Physics

Proposition 1.9. *The derivative must be a linear function.*

It is automatically linear!

Proof. Recall that $dv[a_i\, kt] = dv[ka_i\, t]$, therefore $dw[a_i \left.\frac{df}{dV}\right|_v (kt)] = df(v_i, dv[a_i\, kt]) = df(v_i, dv[ka_i\, t]) = dw[ka_i \left.\frac{df}{dV}\right|_v (t)] = dw[a_i\, k \left.\frac{df}{dV}\right|_v (t)]$. Therefore $\left.\frac{df}{dV}\right|_v (kt) = k \left.\frac{df}{dV}\right|_v (t)$.

We also have

$$\begin{aligned}
dw[a_i \left.\frac{df}{dV}\right|_v (t+u)] &= df(v_i, dv[a_i\, t+u]) = \{f(v_i + a_i(t_i + u_i)) - f(v_i)\}_{i=1}^{\infty} \\
&= \{f(v_i + a_i(t_i + u_i)) - f(v_i + a_i t_i) + f(v_i + a_i t_i) - f(v_i)\}_{i=1}^{\infty} \\
&= \{f((v_i + a_i t_i) + a_i u_i) - f(v_i + a_i t_i)\}_{i=1}^{\infty} + \{f(v_i + a_i t_i) - f(v_i)\}_{i=1}^{\infty} \\
&= df(v_i + a_i t_i, dv[a_i\, u]) + df(v_i, dv[a_i\, t]) \\
&= dw[a_i \left.\frac{df}{dV}\right|_v (u)] + dw[a_i \left.\frac{df}{dV}\right|_v (t)] \\
&= dw[a_i \left.\frac{df}{dV}\right|_v (u) + \left.\frac{df}{dV}\right|_v (t)]
\end{aligned}$$

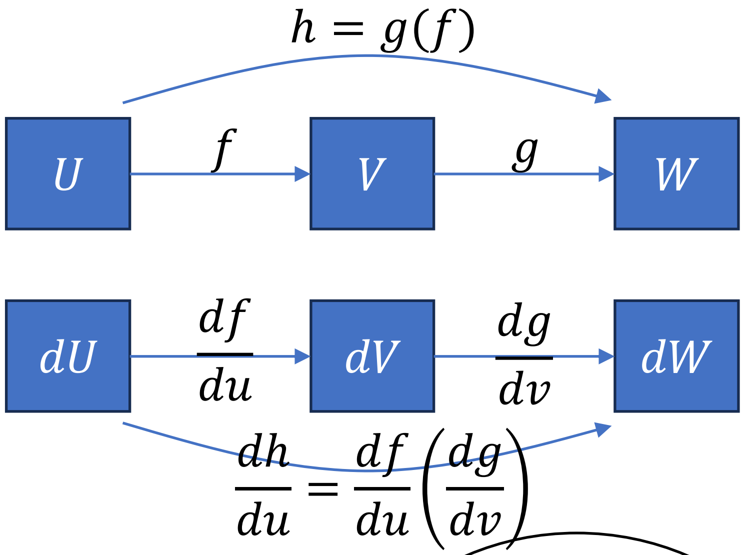
Generalizations of derivative (e.g. Fréchet) are DEFINED to be linear

https://assumptionsofphysics.org/

Assumptions of Physics

Proposition 1.14 (Chain rule). *Let U, V and W be three vector spaces. Let $f : U \to V$ and $g : V \to W$ be two differentiable maps and $h = g \circ f$ their composition. Then* $\frac{dh}{dU} = \frac{dg}{dV} \circ \frac{df}{dU}$.

Chain rule is simply function composition



Proof. Since f is differentiable, we have $df(u_i, du[a_i\, t]) = dv[a_i \left.\frac{df}{dU}\right|_u (t)]$ for all $u_i \to u$, convergence envelopes a_i and tangent vectors t. Since g is differentiable, we have $dg(v_i, dv[a_i\, t]) = dw[a_i \left.\frac{dg}{dV}\right|_v (t)]$ for all $v_i \to v$, convergence envelopes a_i and tangent vectors t. In particular, we have

$$\begin{aligned}
dw[a_i \left.\frac{dg}{dV}\right|_v \left(\left.\frac{df}{dU}\right|_u (t)\right)] &= dg(f(u_i), dv[a_i \left.\frac{df}{dU}\right|_u (t)]) \\
&= dg(f(u_i), df(u_i, du[a_i\, t])) \\
&= dg(f(u_i), df(u_i, \{a_i t_i\})) \\
&= dg(f(u_i), \{f(u_i + a_i t_i) - f(u_i)\})) \\
&= \{g(f(u_i) + f(u_i + a_i t_i) - f(u_i)) - g(f(u_i))\} \\
&= \{g(f(u_i + a_i t_i)) - g(f(u_i))\} \\
&= \{h(u_i + a_i t_i) - h(u_i)\} \\
&= dh(u_i, du[a_i\, t]).
\end{aligned} \tag{1.15}$$

Again, proof is half a page

Therefore the image of the differential through h is a differential with convergence envelope a_i and tangent vector $\left.\frac{dg}{dV}\right|_v \left(\left.\frac{df}{dU}\right|_u (t)\right)$. The derivative of h is

$$\frac{dh}{dU} = \frac{dg}{dV} \circ \frac{df}{dU} \tag{1.16}$$

Proposition 1.13. *Let $f : \mathbb{R} \to \mathbb{R}$. Then the standard analytical notions of differentiability and derivative coincide.*

Proposition 1.19. *Let V and W be two normed vector spaces. Then the notion of Fréchet derivative and the new derivative coincide.*

Coincides with standard notion of derivatives in specific cases

Differentiability: forms and linear functionals

Starting point: finite values defined on finite regions

k-functional, k-surface, k-form, k-vector

$$f_k(\sigma^k) = \int \theta_k(d\sigma^k)$$

Physically measurable quantities

zero-form

Differential forms: infinitesimal limit

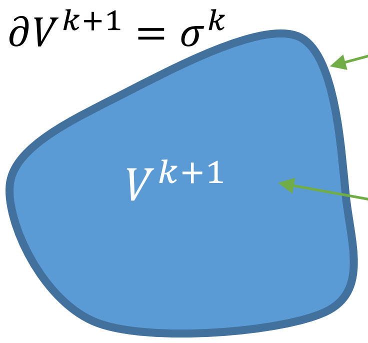
Temperature:	$T(P)$		zero-form
Work:	$W(\gamma) = \sum_i W(\gamma_i) = \int f(d\gamma)$	$f = dW/d\gamma$	one-form
Magnetic flux:	$\Phi(\sigma) = \sum_i \Phi(\sigma_i) = \iint B(d\sigma)$	$B = d\Phi/d\sigma$	two-form
Mass:	$m(V) = \sum_i m(V_i) = \iiint \rho(dV)$	$\rho = dm/dV$	three-form

Assume additivity over disjoint regions

Thinking in terms of relationships between finite objects leads to better physical intuition

The mathematics is contingent upon the assumption of infinitesimal reducibility (e.g. mass in volumes sums only if boundary effects can be neglected)

We can define functionals that act on boundaries



f^k Given a functional

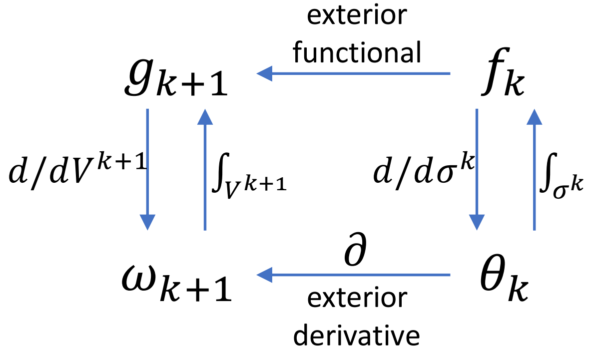
Define higher dimensional functional that acts on the boundary

$$g^{k+1}(V^{k+1}) \equiv f^k(\sigma^k)$$

Exterior functional

$$\partial\partial f^k(\sigma^{k+2}) = f^k(\partial\partial\sigma^{k+2}) = f^k(\emptyset) = 0$$

Boundary of a boundary is the empty set ⇒ exterior derivative of exterior derivative is zero



Reversing the exterior derivative is finding a (non-unique) potential

https://assumptionsofphysics.org/

Assumptions of Physics

Takeaways

- We can define a more general notion of derivative that is more in line with the old ideas of infinitesimals but still rigorous
- We can get better physical intuition because we can understand all infinitesimal objects as limits/decompositions on finite objects (on which the physics is actually defined)
- TODOs:
  - Complete the theory
  - Construct the analogue of differentiable manifold (i.e. topological space that is homeomorphic to a topological vector space at each point)

https://assumptionsofphysics.org/

Assumptions of Physics

Subspaces

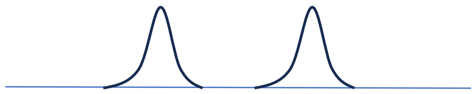
https://assumptionsofphysics.org/

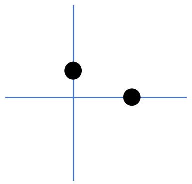
Assumptions of Physics

The basic definition do not tell us what are the spaces on which the ensembles are defined

We need to recover them… in a general way

Disjunctness allows us to detect disjoint support in classical mechanics and orthogonality in quantum mechanics





Use disjunctness to define subspaces

https://assumptionsofphysics.org/

Assumptions of Physics

Definition 1.49. *Let X be a set and $R \subseteq X \times X$ a symmetric relation. Given a subset $U \subseteq X$, we define the **R-complement** to be*

Will be disjunctness in the end

$$U^{RC} = \{a \in X \mid \forall b \in U, aRb\}.$$

All the elements that are orthogonal to all elements in U

Get a lot of properties out of very little

Proposition 1.50. *Let X be a set and $R \subseteq X \times X$ a symmetric relation. Then:*

1. $U \subseteq V \implies V^{RC} \subseteq U^{RC}$
2. $U \subseteq (U^{RC})^{RC}$
3. $U^{RC} = ((U^{RC})^{RC})^{RC}$
4. $U^{RC} = (V^{RC})^{RC} \iff (U^{RC})^{RC} = V^{RC}$
5. $(\bigcup_{i \in I} U_i)^{RC} = \bigcap_{i \in I} (U_i)^{RC}$
6. $\varnothing^{RC} = X$

Proposition 1.51. *Let X be a set and $R \subseteq X \times X$ a symmetric and irreflexive relation. Then*

1. $U \cap U^{RC} = \varnothing$
2. $X^{RC} = \varnothing$

Nothing is orthogonal to itself

Last few properties

https://assumptionsofphysics.org/

Assumptions of Physics

Definition 1.52. *Let X be a set and $R \subseteq X \times X$ a symmetric relation. Let $U \subseteq X$. The* ***R-subspace generated by** U is $\langle U \rangle_R = (U^{RC})^{RC}$. An* ***R-subspace** of X is a set $U \subseteq X$ such that $U = \langle U \rangle_R$. The* ***lattice of R-subspaces*** *is the set $\mathfrak{L} = \{U \subseteq X \mid U = \langle U \rangle_R\}$ ordered by inclusion.*

Corollary 1.53. *The lattice of R-subspaces $\mathfrak{L}$ is a topped $\bigcap$-structure on X and therefore is also a complete lattice.*

Subspaces defined as subsets that are the complement of their complement

This is usually a property for subspaces of vector spaces. We use it as the defining characteristic

Proposition 1.56. *Let X be an inner product space and $R = \{(a,b) \mid \langle a,b \rangle = 0\}$. Then:*

1. *R is an irreflexive and symmetric relation*
2. *$U^{RC} = U^{\perp}$*
3. *$\langle U \rangle_R = \mathrm{cl}_X(\mathrm{span}(U))$*

Recovers the notion of closed vector subspaces

Proposition 1.57. *Let $\mathcal{E}$ be an ensemble space, disjunctness $\perp\!\!\!\perp$ is an irreflexive symmetric relation.*

Definition 1.58. *Let $\mathcal{E}$ be an ensemble space and $X \subseteq \mathcal{E}$ be a subset. The* ***disjunct complement*** *$X^{\perp\!\!\!\perp} \subseteq \mathcal{E}$ is the set of all ensembles that are disjunct from all elements of X. An* ***ensemble subspace*** *is a subset $X \in \mathcal{E}$ such that $X = (X^{\perp\!\!\!\perp})^{\perp\!\!\!\perp}$.*

Now we apply the more general construction to our specific case

Note that this construction only uses the entropic structure

Valid even if there is not vector space structure (i.e. non-complemented space)

Recovers the correct notions

In the classical case, a subspace is the set of all distributions with support within a set U

In the quantum case, a subspace is the set of all density operators within a subspace of the Hilbert space

To recover the topology of the base space on top of which distributions are defined, the distributions must be continuous

Physically justifiable: experimental relationships are continuous functions

A continuous function is zero only on an open set

Hilbert spaces in quantum mechanics cannot work: $L^2(\mathbb{R}) \equiv L^\wedge 2(\mathbb{R}^n)$

Schwartz spaces do work: $S(\mathbb{R}) \not\equiv S(\mathbb{R}^n)$

Definition 1.62. *A **pure state** is a limit of an ensemble with the smallest entropy. Formally, a pure state $s \subset \mathfrak{L}$ is a non-empty collection of subspaces such that*

1. *if $\{X_i\}_{i \in I} \in s$ then $\bigcap_{i \in I} X_i \in s$*
2. *if $X \in s$ and $Y \in \mathfrak{L}$ such that $X \subseteq Y$, then $Y \in s$*
3. *if $X \in s$ and $Y \in \mathfrak{L}$ such that $Y \subseteq X$ and $Y \neq \varnothing$, then there exists $Z \in s$ such that $Z \subset X$.*

The set of all pure states for an ensemble space is noted $S(\mathcal{E})$.

Points can be recovered by looking for sequences of subspaces that become “smaller and smaller”

They are definitely not pure states…

Takeaways

- We can define a general notion of subspaces of ensembles based on disjunctness
- The geometry of the vector spaces (i.e. the inner product, orthogonality) is also uniquely defined by the entropy
- TODOs:
  - Complete the theory
  - Can we also define a notion of inner product from the entropy?

Probability and measure theory

https://assumptionsofphysics.org/

Assumptions of Physics

Classical probability

Possibilities (experimentally defined cases)

- the sample space Ω – an arbitrary non-empty set,
- the σ-algebra $\mathcal{F} \subseteq 2^{\Omega}$ (also called σ-field) – a set of subsets of Ω, called events, such that:
  - $\mathcal{F}$ contains the sample space: $\Omega \in \mathcal{F}$,
  - $\mathcal{F}$ is closed under complements: if $A \in \mathcal{F}$, then also $(\Omega \setminus A) \in \mathcal{F}$,
  - $\mathcal{F}$ is closed under countable unions: if $A_i \in \mathcal{F}$ for $i = 1, 2, \ldots$, then also $(\bigcup_{i=1}^{\infty} A_i) \in \mathcal{F}$
    - The corollary from the previous two properties and De Morgan's law is that $\mathcal{F}$ is also closed under countable intersections: if $A_i \in \mathcal{F}$ for $i = 1, 2, \ldots$, then also $(\bigcap_{i=1}^{\infty} A_i) \in \mathcal{F}$
- the probability measure $P : \mathcal{F} \to [0, 1]$ – a function on $\mathcal{F}$ such that:
  - P is countably additive (also called σ-additive): if $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{F}$ is a countable collection of pairwise disjoint sets, then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$,
  - the measure of the entire sample space is equal to one: $P(\Omega) = 1$.

Theoretical statements (statements with tests)

Probability each theoretical statement is true

Does not work for quantum mechanics!

Want a generalization for both

https://assumptionsofphysics.org/

Assumptions of Physics

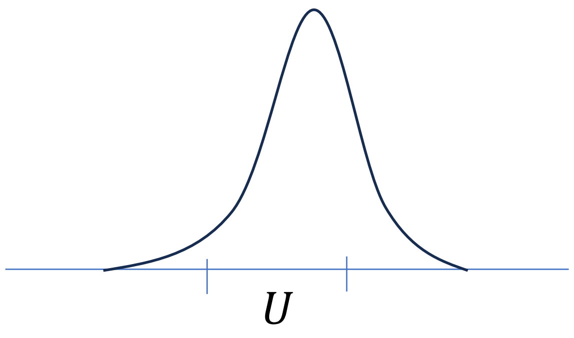
Definition 1.70. *Let* $\mathsf{e} \in \mathcal{E}$. *The* ***probability measure*** *for* e *is defined as* $p_{\mathsf{e}}(A) = \sup(\{p \in [0,1] \mid \exists\, \mathsf{e}_1 \in \text{hull}(A), \mathsf{e}_2 \in \mathcal{E} \text{ s.t. } \mathsf{e} = p\mathsf{e}_1 + \bar{p}\mathsf{e}_2\})$ *if* $A \neq \varnothing$ *and* $p_{\mathsf{e}}(A) = 0$ *otherwise.*

Corollary 1.71. *Let* $\mathsf{e}_1, \mathsf{e}_2 \in \mathcal{E}$. *Then* $p_{\mathsf{e}_1} \neq p_{\mathsf{e}_2}$.

Given a set of ensembles, the probability is the biggest mixing coefficient associated to the biggest component of the target ensemble

$$p(x) = p(x|U)p(U) + p(x|U^C)p(U^C)$$

$$e = e_1 p + e_2(1 - p)$$



Proposition 1.73. *The probability measure for an ensemble is*

1. *non negative and unit bounded:* $p_e(U) \in [0, 1]$
2. *monotone:* $U \subseteq V \implies p_e(U) \leq p_e(V)$
3. *subadditive:* $p_e(U \cup V) \leq p_e(U) + p_e(V)$

Non-additive in general. When do we recover additivity?

Non-additivity comes out when distinct ensembles are not disjunct.

Restrict to structure where distinct = disjunct.

Definition 1.74. *Let $\mathfrak{L}$ be the lattice of subspaces and $\wedge, \vee, (\cdot)^{\perp\!\!\!\perp}$ be respectively the join, meet and disjunct complement. A* ***context*** *is a lattice of subspaces $C \subseteq \mathfrak{L}$ such that*

1. *if $\{X_i\}_{i \in I} \in C$ then $\bigwedge_{i \in I} X_i \in C$, $\bigvee_{i \in I} X_i \in C$ and $X_i^{\perp\!\!\!\perp} \in C$*
2. *if $X, Y \in C$ and $X \cap Y = \emptyset$, then $X \perp\!\!\!\perp Y$.*

That is, the join, meet and complement operation in $\mathfrak{L}$ and C are the same (i.e. smallest subspace that contains all, biggest subspace contained by all, disjunct subspace). The additional property is that disjoint subspaces are disjunct.

The entire lattice of subspaces of CM is a context

Lattice of orthonormal subspaces in QM (i.e. a basis) is a context

Results from Quantum Logic can be reused at this point

Takeaways

- We can define a general of non-additive probability that is physically motivated and works for both classical and quantum mechanics (and beyond)
- The mixing coefficients are more fundamental and the probability comes out of decomposition of ensembles into parts
- TODOs:
  - Fully develop a non-additive measure theoretical parallel

https://assumptionsofphysics.org/

Assumptions of Physics

Other missing things

- Symplectic structure should be imposed on the space of ensembles to have coordinate invariance of entropy… How?
- Composite systems (i.e. product spaces) and independence of DOFs
- Quantities (linear maps from ensembles to real numbers… or other topological spaces?)
- Processes (linear ensemble maps)
  - Deterministic and reversible processes (entropy preserving processes)
  - Equilibration process (projections)
- Extension to field theory

Wrapping it up

- We should be able to construct a general theory of states and processes on minimal requirements
- Same concepts for all theories
- Lots of interesting mathematical work to do

https://assumptionsofphysics.org/

Assumptions of Physics

https://assumptionsofphysics.org/

Assumptions of Physics