Assumptions of Physics
Summer School 2025

Some results in Physical Mathematics

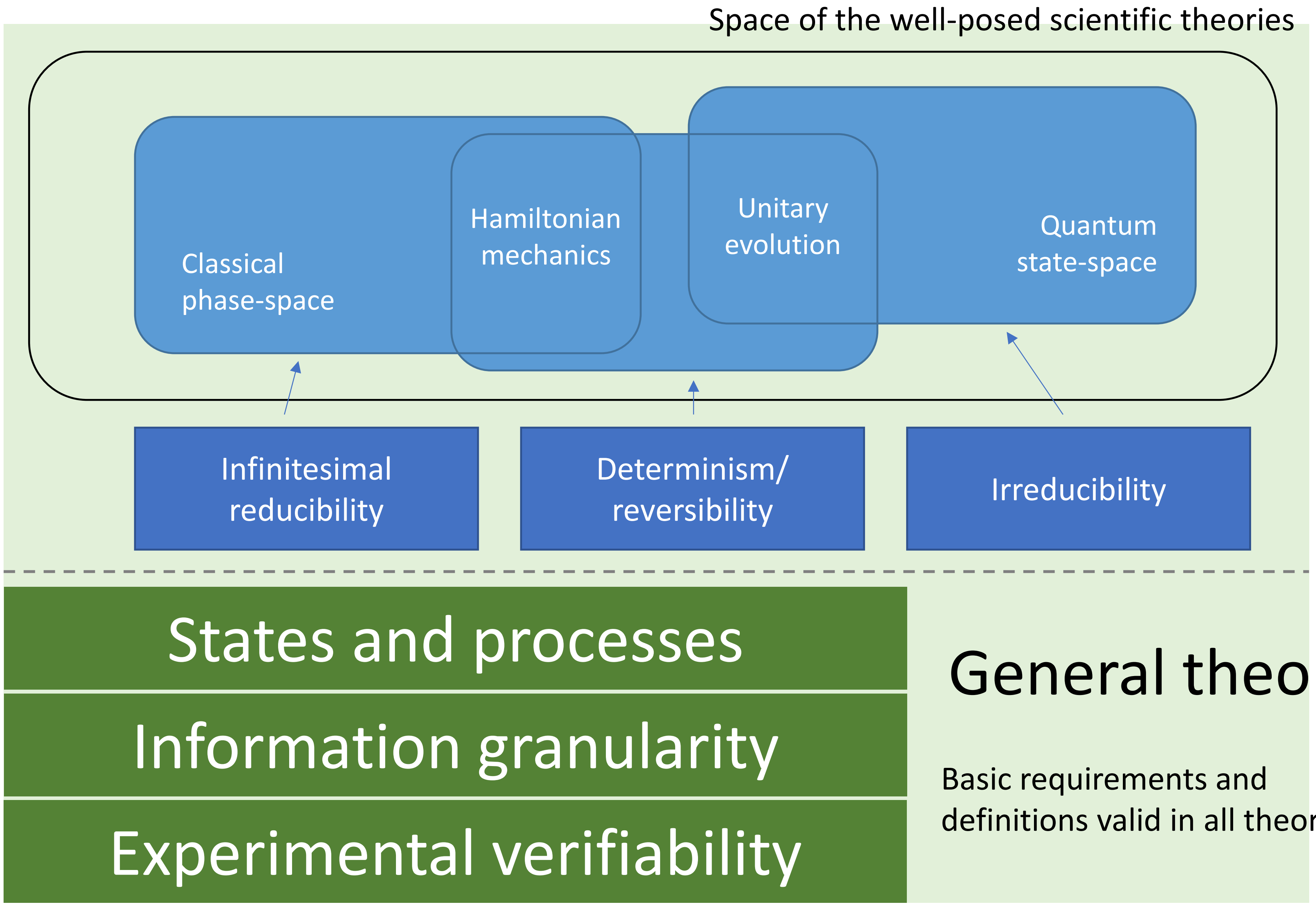
Gabriele Carcassi and Christine A. Aidala

Physics Department
University of Michigan

Overview

https://assumptionsofphysics.org/

Assumptions of Physics



Physical theories

Specializations of the general theory under the different assumptions

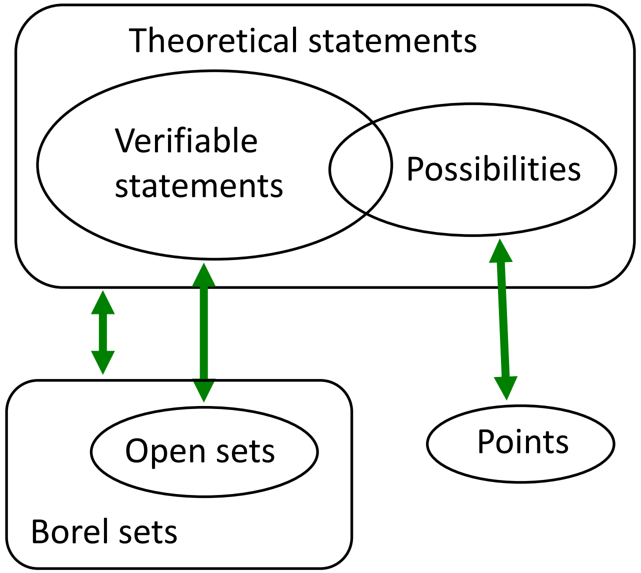
Assumptions

General theory

Basic requirements and definitions valid in all theories

https://assumptionsofphysics.org/

Assumptions of Physics

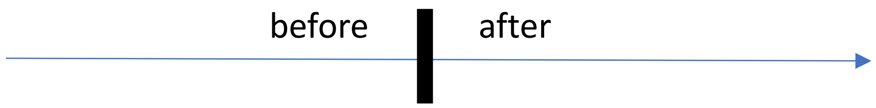


Experimental verifiability

$\Downarrow$

Topology and σ-algebra

Quantities from references

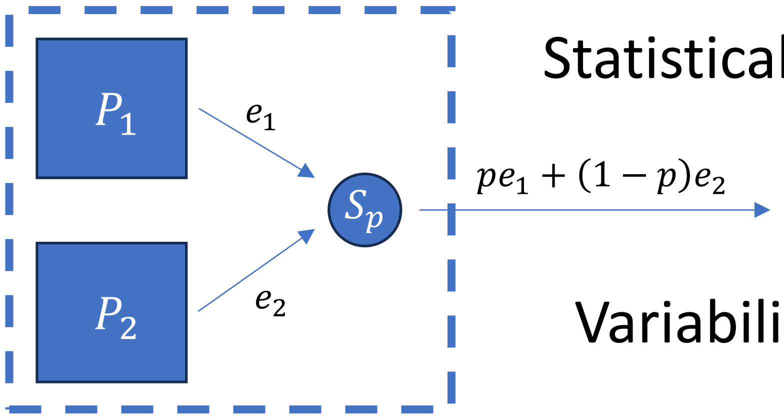


Refinable aligned strict references

$\Downarrow$

Real numbers

References act as Dedekind cuts



Statistical mixtures $\Rightarrow$ Convex (linear) structures

Variability/entropy $\Rightarrow$ Geometric structures

https://assumptionsofphysics.org/

Assumptions of Physics

Recover the reals

https://assumptionsofphysics.org/

Assumptions of Physics

$\mathbb{R}$

In physics, we use real numbers to model physical quantities

Needed for a lot of mathematical tools: exponential, gaussians, trigonometry, …

Yet, we know that infinite precision is an idealization

Many ways to mathematically construct the real numbers

As a field, as Cauchy sequence of rationals, as Dedekind cuts of rationals…

How can we define them in physical mathematics?
What do real numbers model?

https://assumptionsofphysics.org/

Assumptions of Physics

Order topology

Given $(Q, \leq)$, the order topology is the one generated by the test $\{q > a\}$ and $\{q < b\}$ for all $a, b \in Q$

This means that "the object is after a" and
"the object is before b" are verifiable statements

Note: topology only knows about order, not about the size of the intervals!!!

https://assumptionsofphysics.org/

Assumptions of Physics

Reals can be characterized as a type of linear order

Cantor's isomorphism theorem

1) Every countable dense linearly ordered set with no ends is order isomorphic to the rationals

Bijection with naturals

Always an element between two

One element is always before or after another

No greatest or smallest

2) The completion of a linear order is unique

Every subset has a supremum and an infimum

$\Rightarrow$ Every complete linearly ordered set with no ends that has a dense countable subset is order isomorphic to the reals

Since the reals are the completion of the rationals

https://assumptionsofphysics.org/

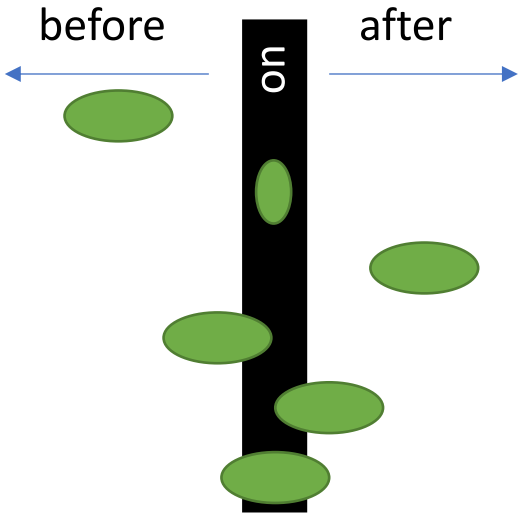
Assumptions of Physics

How is order defined experimentally?

wikipedia

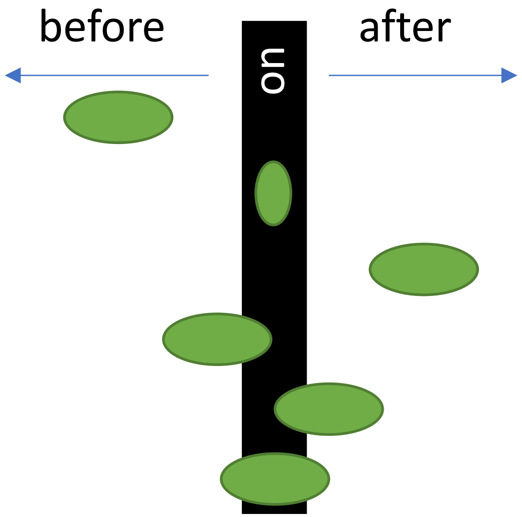
We have a way to compare two things, and decide whether one is bigger/smaller than the other. “Exactly the same” is a bit more difficult.

A **reference** (e.g. a tick of a clock, notch on a ruler, sample weight with a scale) is something that allows us to distinguish between a before and an after



“Before/after a reference” are verifiable statements

Goal: make a mathematical model of a reference system



Definition 3.17. *A* ***reference*** *defines a before, an on and an after relationship between itself and another object. Formally a reference* $\boldsymbol{r} = (\mathsf{b}, \mathsf{o}, \mathsf{a})$ *is a tuple of three statements such that:*

1. *we can verify whether the object is before or after the reference:* b *and* a *are verifiable statements*
2. *the object can be on the reference:* $\mathsf{o} \not\equiv \perp$
3. *if it's not before or after, it's on the reference:* $\neg\mathsf{b} \wedge \neg\mathsf{a} \preccurlyeq \mathsf{o}$
4. *if it's before and after, it's also on the reference:* $\mathsf{b} \wedge \mathsf{a} \preccurlyeq \mathsf{o}$

A ***beginning reference*** *has nothing before it. That is,* $\mathsf{b} \equiv \perp$*. An* ***ending reference*** *has nothing after it. That is,* $\mathsf{a} \equiv \perp$*. A* ***terminal reference*** *is either beginning or ending.*

Before	On	After
T	F	F
F	T	F
F	F	T
T	T	F
F	T	T
T	T	T

A reference system is a collection of references. The experimental domain for a quantity is the set of verifiable statements generated by all possible before/after statements.

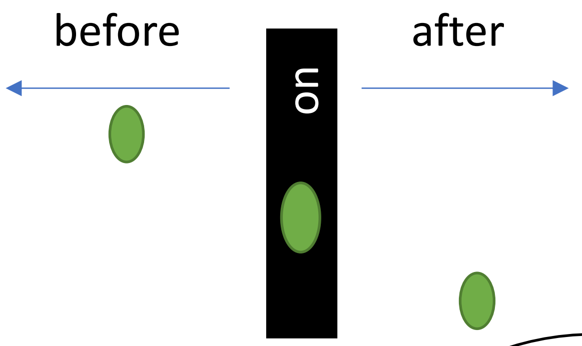
https://assumptionsofphysics.org/

Assumptions of Physics

1. Strict references

A reference is strict if before/on/after are mutually exclusive

Before	On	After
T	F	F
F	T	F
F	F	T



Physically, the extent of what we measure is assumed to be smaller than the extent of our reference

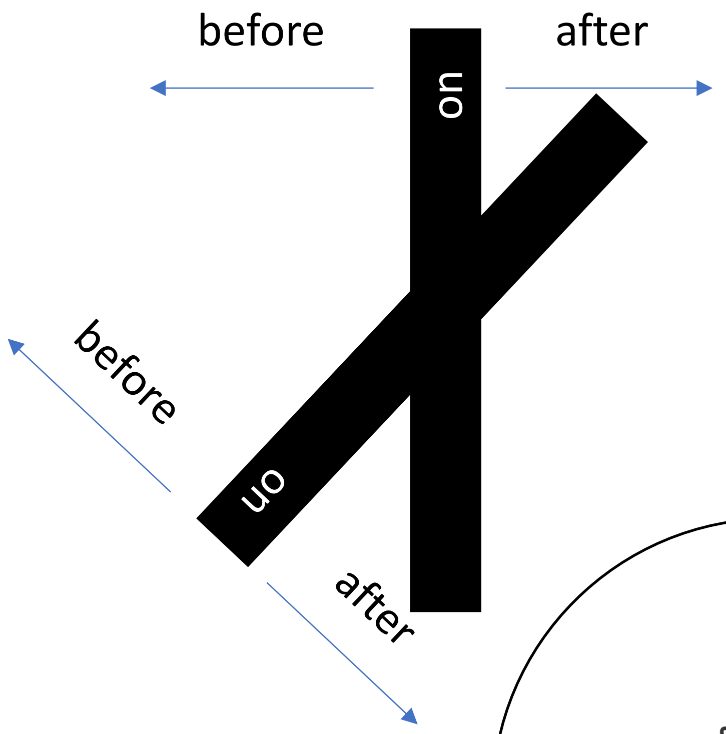
https://assumptionsofphysics.org/

Assumptions of Physics

Multiple references

Without further constraints, references would not lead to a linear order

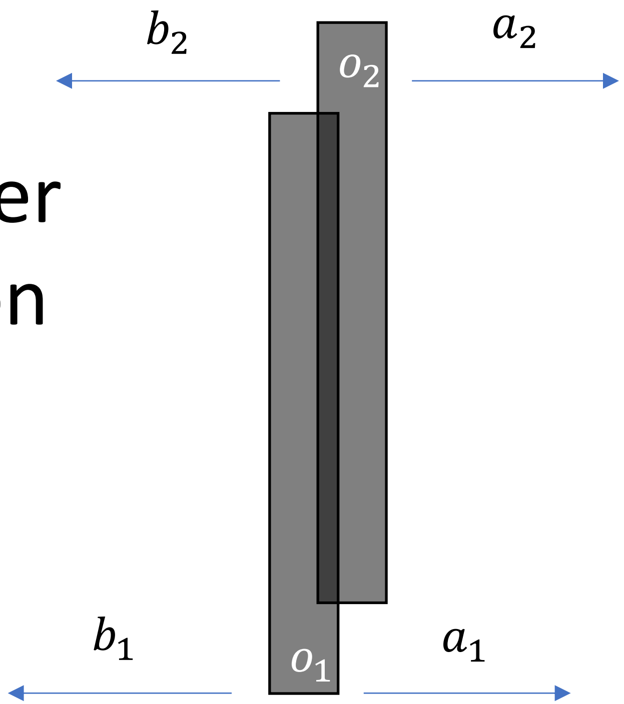
	b_2	o_2	a_2
b_1	✓	✓	✓
o_1	✓	✓	✓
a_1	✓	✓	✓



https://assumptionsofphysics.org/

Assumptions of Physics

2. Aligned references



Two references are aligned if the before and not-after statement can be ordered by narrowness/implication

For example, $b_1 \preccurlyeq b_2 \preccurlyeq \neg a_1 \preccurlyeq \neg a_2$

$\preccurlyeq$ Means that if the first statement is true then the second statement will be true as well

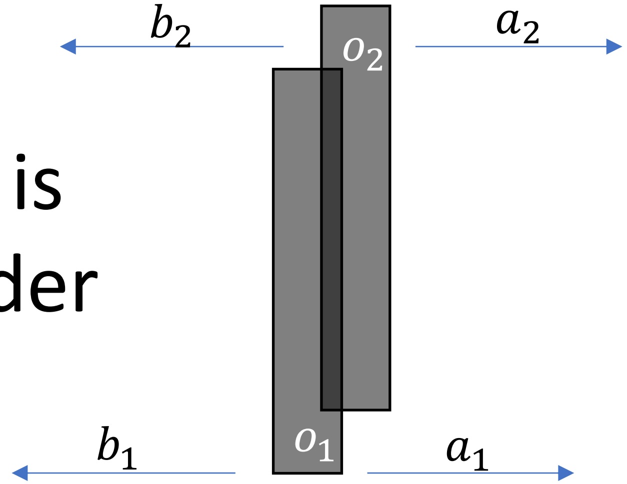
That is, the first statement is narrower, more specific

Basic insight: the ordering of the points corresponds to the order of statements by narrowness

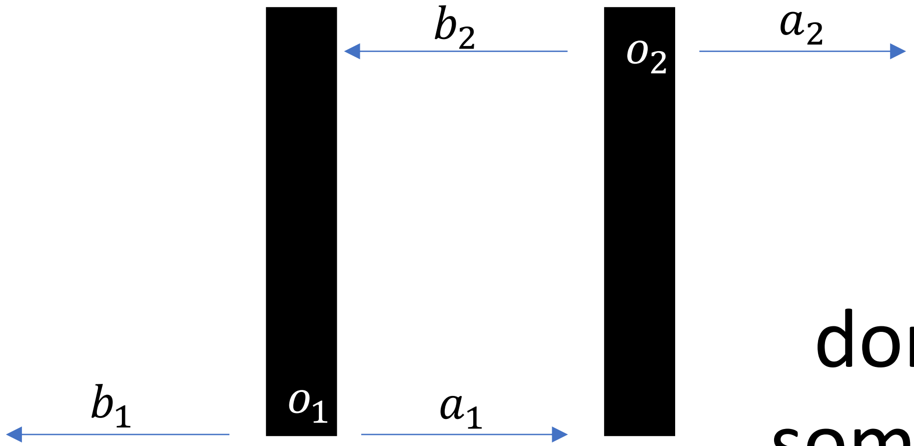
https://assumptionsofphysics.org/

Assumptions of Physics

Filling the whole region



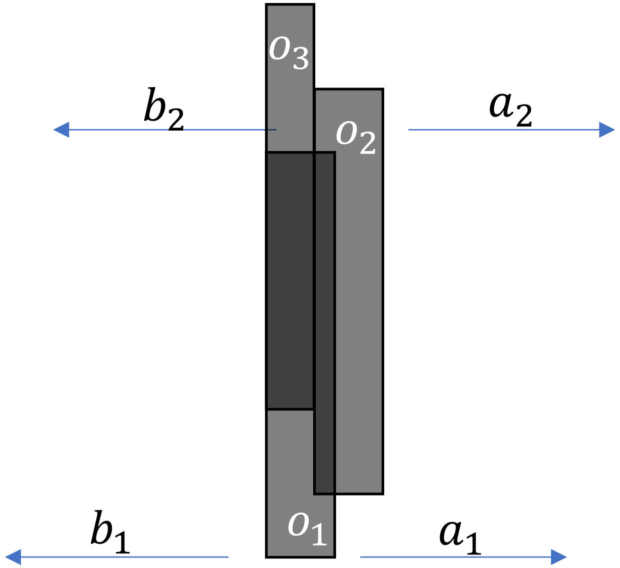
If two different references overlap, we can't say one is before the other: we can't fully resolve the linear order



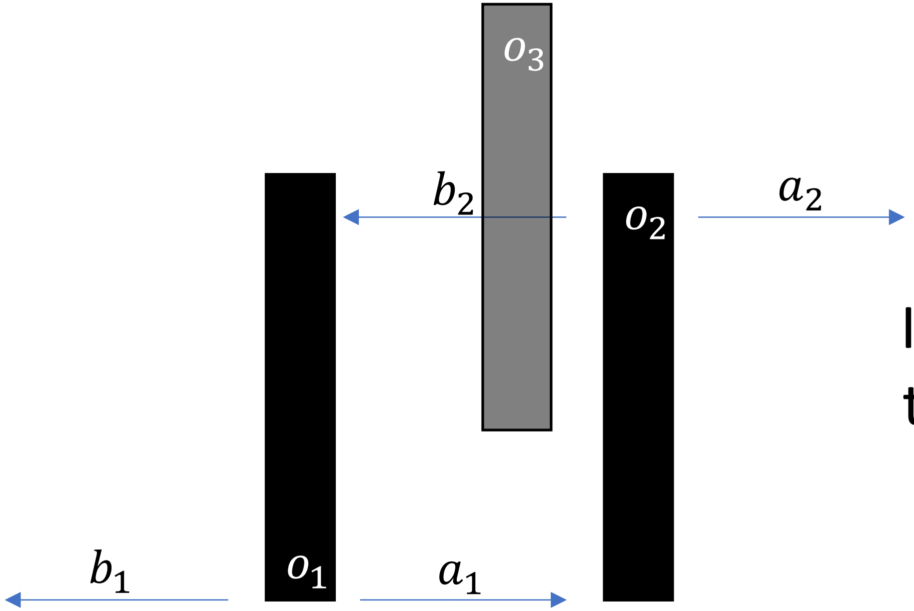
Moreover, if two references don't overlap and there can be something in between, we must be able to put a reference there

3. Refinable references

A set of references is refinable if we can address the previous two problems and resolve the full space



If two references overlap, we can find a reference that refines the overlap



If something can be found between two references, then there must be another reference in between

https://assumptionsofphysics.org/

Assumptions of Physics

Proof. By definition, we have $\neg b \wedge \neg a \preccurlyeq o$ and by 1.23 $\neg(\neg b \wedge \neg a) \vee o \equiv \top \equiv b \vee a \vee o$. □

Definition 3.19. *A reference $r_1 = (b_1, o_1, a_1)$ is **finer** than another reference $r_2 = (b_2, o_2, a_2)$ if $b_1 \succcurlyeq b_2$, $o_1 \preccurlyeq o_2$ and $a_1 \succcurlyeq a_2$.*

Corollary 3.20. *The finer relationship between references is a partial order.*

Proof. As the finer relationship is directly based on narrowness, it inherits its reflexivity, antisymmetry and transitivity properties and is therefore a partial order. □

Definition 3.21. *A reference is **strict** if its before, on and after statements are incompatible. Formally, $r = (b, o, a)$ is such that $b \not\equiv a$ and $o \equiv \neg b \wedge \neg a$. A reference is **loose** if it is not strict.*

Remark. In general, we can't turn a loose reference into a strict one. The on statement can be made strict by replacing it with $\neg b \wedge \neg a$. This is possible because o is not required to be verifiable. The before (and after) statements would need to be replaced with statements like $b \wedge \neg a$, which are not in general verifiable because of the negation.

To measure a quantity we will have many references one after the other: a ruler will have many marks, a scale will have many reference weights, a clock will keep ticking. What does it mean that a reference comes after another in terms of the before/on/after statements?

If reference r_1 is before reference r_2 we expect that if the value measured is before the first it will also be before the second, and if it is after the second it will also be after the first. Note that this is not enough, though, because as references have an extent they may overlap. And if they overlap one can't be after the other. To have an ordering properly defined we must have that the first reference is entirely before the second. That is, if the value measured is on the first it will be before the se[...]

Mathematically, this type of ord[...] before and strictly after. It does no[...] One may be tempted to define the [...] requires refining the references and, [...] refined references, not the original o[...]

Definition 3.22. *A reference is b[...] the first it cannot be on or after th[...]*

Proposition 3.23. *Reference ord[...]*

- *irreflexivity: not $r < r$*
- *transitivity: if $r_1 < r_2$ and r_2[...]*

*and is therefore a **strict partial** [...]*

Proof. For irreflexivity, since t[...] and therefore $b \vee o \not\equiv o \vee a$. Therefo[...] irreflexive.

For transitivity, if $r_1 < r_2$, we have $b_1 \vee o_1 \not\equiv o_2 \vee a_2$ and therefore $\neg(b_1 \vee o_1) \succcurlyeq o_2 \vee a_2$ by 1.23. Since $b_1 \vee o_1 \vee a_1 \equiv \top$, we have $a_1 \succcurlyeq \neg(b_1 \vee o_1)$. Similarly if $r_2 < r_3$ we'll have $a_2 \succcurlyeq \neg(b_2 \vee o_2) \succcurlyeq o_3 \vee a_3$. Putting it all together $\neg(b_1 \vee o_1) \succcurlyeq o_2 \vee a_2 \succcurlyeq a_2 \succcurlyeq \neg(b_2 \vee o_2) \succcurlyeq o_3 \vee a_3$, which means $b_1 \vee o_1 \not\equiv o_3 \vee a_3$. □

Corollary 3.24. *The relationship $r_1 \le r_2$, defined to be true if $r_1 < r_2$ or $r_1 = r_2$, is a partial order.*

As we saw, two references may overlap and therefore an ordering between them cannot be defined. But references can overlap in different ways.

Suppose we have a vertical line one millimeter thick and call the left side the part before the line and the right side the part after. We can have another vertical line of the same thickness overlapping but we can also have a horizontal line which will also, at some point, overlap. The case of the two vertical lines is something that, through finding finer references, can be given a linear order. The case of the vertical and horizontal line, instead, cannot. Intuitively, the vertical lines are aligned while the horizontal and vertical are not.

Conceptually, the overlapping vertical lines are aligned because we can imagine narrower lines around the borders, and those lines will be ordered references in the above sense: each line would be completely before or after, without intersection. This means that the before and not-after statements of one reference are either narrower or broader than the before and not-after statements of the other. That is, alignment can also be defined in terms of narrowness of statements.

Note that if a reference is strict, before and after statements are not compatible and therefore the before statement is narrower than the not-after statement. This means that, given a set of aligned strict references, the set of all before and not-after statements is linearly ordered by narrowness. As we saw in the previous section, this was a necessary condition

Proof. We have $b_1 \vee o_1 \equiv (b_1 \vee o_1) \wedge \top \equiv (b_1 \vee o_1) \wedge (b_2 \vee o_2 \vee a_2) \equiv ((b_1 \vee o_1) \wedge b_2) \vee ((b_1 \vee o_1) \wedge (o_2 \vee a_2)) \equiv ((b_1 \vee o_1) \wedge b_2) \vee \bot \equiv (b_1 \vee o_1) \wedge b_2$. Therefore $b_1 \vee o_1 \preccurlyeq b_2$. And since $b_1 \preccurlyeq b_1 \vee o_1$, we have $b_1 \preccurlyeq b_2$.

Similarly, we have $o_2 \vee a_2 \equiv (o_2 \vee a_2) \wedge \top \equiv (o_2 \vee a_2) \wedge (b_1 \vee o_1 \vee a_1) \equiv ((o_2 \vee a_2) \wedge (b_1 \vee o_1)) \vee ((o_2 \vee a_2) \wedge a_1) \equiv \bot \vee ((o_2 \vee a_2) \wedge a_1) \equiv (o_2 \vee a_2) \wedge a_1$. Therefore $a_2 \vee o_2 \preccurlyeq a_1$. And since $a_2 \preccurlyeq o_2 \vee a_2$, we have $a_2 \preccurlyeq a_1$.

Since $b_1 \vee o_1 \not\equiv o_2 \vee a_2$, we have $b_1 \not\equiv a_2$ which means $b_1 \preccurlyeq \neg a_2$.

Since $b_1 \vee o_1 \vee a_1 \equiv \top$, we have $\neg a_1 \preccurlyeq b_1 \vee o_1$. Similarly $\neg b_2 \preccurlyeq o_2 \vee a_2$. Since $b_1 \vee o_1 \not\equiv o_2 \vee a_2$, $\neg a_1 \not\equiv \neg b_2$ and therefore $\neg a_1 \preccurlyeq b_2$.

Since $b_1 \preccurlyeq b_2$, $a_2 \preccurlyeq a_1$, $b_1 \preccurlyeq \neg a_2$ and $\neg a_1 \preccurlyeq b_2$, the two references are aligned. □

Proposition 3.28. *Let $r_1 = (b_1, o_1, a_1)$ and $r_2 = (b_2, o_2, a_2)$ be two strict references. Then $r_1 < r_2$ if and only if $\neg a_1 \preccurlyeq b_2$.*

Proof. Let $r_1 < r_2$. By 3.27, we have $\neg a_1 \preccurlyeq b_2$. Conversely, let $\neg a_1 \preccurlyeq b_2$. Then $\neg a_1 \not\equiv \neg b_2$. Because the references are strict, $\neg a_1 \equiv b_1 \vee o_1$ and $\neg b_2 \equiv o_2 \vee a_2$. Therefore $b_1 \vee o_1 \not\equiv o_2 \vee a_2$ and $r_1 < r_2$ by definition. □

Definition 3.29. *A reference is the **immediate** predecessor of another if nothing can be found before the second and after the first. Formally, $r_1 < r_2$ and $a_1 \not\equiv b_2$. Two references are **consecutive** if one is the immediate successor of the other.*

Proposition 3.30. *Let $r_1 = (b_1, o_1, a_1)$ and $r_2 = (b_2, o_2, a_2)$ be two references. If r_1 is immediately before r_2 then $b_2 \equiv \neg a_1$.*

Proof. Let r_1 be immediately before r_2. Then $a_1 \not\equiv b_2$ which means $b_2 \preccurlyeq \neg a_1$. By 3.27 we also have $\neg a_1 \preccurlyeq b_2$. Therefore $b_2 \equiv \neg a_1$. □

Proposition 3.31. *Let $r_1 = (b_1, o_1, a_1)$ and $r_2 = (b_2, o_2, a_2)$ be two strict references. Then r_1 is immediately before r_2 if and only if $b_2 \equiv \neg a_1$.*

Proof. Let r_1 be immediately before r_2. Then $b_2 \equiv \neg a_1$ by 3.30. Conversely, let $b_2 \equiv \neg a_1$. Then $r_1 < r_2$ by 3.28. We also have $a_1 \not\equiv \neg a_1$, therefore $a_1 \not\equiv b_2$ and r_1 is immediately before r_2 by definition. □

Definition 3.33. *Let $\mathcal{D}$ be a domain generated by a set of references R. A reference $r = (b, o, a)$ is said to be aligned with $\mathcal{D}$ if $b \in \mathcal{D}_b$ and $a \in \mathcal{D}_a$.*

Proposition 3.34. *Let $\mathcal{D}$ be an experimental domain generated by a set of aligned strict references R and let $D = \mathcal{D}_b \cup \neg(\mathcal{D}_a)$. Then $(D, \preccurlyeq)$ is linearly ordered.*

Proof. By 3.26 we have that $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$ is aligned by narrowness. By 3.15 the ordering extends to D. □

Having a set of aligned references is not necessarily enough to cover the whole space at all levels of precision. To do that we need to make sure that, for example, between two references that are not consecutive we can at least put a reference in between. Or that if we have two references that overlap, we can break them apart into finer ones that do not overlap and one is after the other.

We call a set of references refinable if the domain they generate has the above mentioned properties. This allows us to break up the whole domain into a sequence of references that do not overlap, are linearly ordered and that cover the whole space. As we get to the finest references, their before statements will be immediately followed by the negation of their after statements, since there can't be any reference in between. Conceptually, this will give us the second and the third condition of the domain ordering theorem 3.16.

Definition 3.35. *Let $\mathcal{D}$ be an experimental domain generated by a set of aligned references R. The set of references is **refinable** if, given two strict references $r_1 = (b_1, o_1, a_1)$ and $r_2 = (b_2, o_2, a_2)$ aligned with $\mathcal{D}$, we can always:*

- *find an intermediate one if they are not consecutive; that is, if $r_1 < r_2$ but r_2 is not the immediate successor of r_1, then we can find a strict reference r_3 aligned with $\mathcal{D}$ such that $r_1 < r_3 < r_2$.*
- *refine overlapping references if one is finer than the other; that is, if $o_2 < o_1$, we can find a strict reference r_3 aligned with $\mathcal{D}$ such that $o_3 \preccurlyeq o_1$ and either $b_3 \equiv b_1$ and $r_3 < r_2$ or $a_3 \equiv a_1$ and $r_2 < r_3$.*

Proposition 3.36. *Let $\mathcal{D}$ be an experimental domain generated by a set of refinable aligned strict references R.*

means we can find $r_3 = (b, \neg b \wedge \neg a_2, a_2)$ for some $b \in \mathcal{D}_b$ such that $r_3 < r_2$ and therefore $\neg a_1 \preccurlyeq b \preccurlyeq \neg b_2$.

For the third, suppose $a_1 \in \mathcal{D}_a$ and $b_2 \in \mathcal{D}_b$ such that $\neg a_1 < b_2$. Then $r_1 = (\bot, \neg a_1, a_1)$ and $r_2 = (b_2, \neg b_2, \bot)$ are strict references aligned with the domain such that $r_1 < r_2$ but r_2 is not an immediate successor of r_1. This means we can find $r_3 = (b, \neg b \wedge \neg a, a)$ such that $r_1 < r_3 < r_2$ and therefore $\neg a_1 \preccurlyeq b < \neg a \preccurlyeq b_2$. □

Proposition 3.37. *Let $\mathcal{D}$ be an experimental domain generated by a set of refinable aligned strict references. Then all elements of D are part of a pair $(s_b, \neg s_a)$ such that $s_b \in \mathcal{D}_b$, $s_a \in \mathcal{D}_a$ and $\neg s_a$ is the immediate successor of s_b in D or $s_b \equiv \neg s_a$. Moreover if $s \in D$ has an immediate successor, then $s \in \mathcal{D}_b$.*

Proof. Let $\mathcal{D}$ be an experimental domain generated by a set of refinable aligned strict references. Let $s_b \in \mathcal{D}_b$. Let $A = \{a \in \mathcal{D}_a \mid a \vee s_b \not\equiv \top\}$. Let $s_a = \bigvee_{a \in A} a$. First we show that $s_b \preccurlyeq \neg s_a$. We have $s_b \wedge \neg s_a \equiv s_b \wedge \neg \bigvee_{a \in A} a \equiv \bigwedge_{a \in A} s_b \wedge \neg a$. For all $a \in A$ we have $a \vee s_b \not\equiv \top$, $\neg a \not\preccurlyeq s_b$ which means $s_b \preccurlyeq \neg a$ because of the total order of D. This means that $s_b \wedge \neg a \equiv s_b$ for all $a \in A$, therefore $s_b \wedge \neg s_a \equiv s_b$ and $s_b \preccurlyeq \neg s_a$.

Next we show that no statement $s \in D$ such that $s_b < s < \neg s_a$. Let $a \in \mathcal{D}_a$ such that $s_b < \neg a$. By construction $a \in A$ and therefore $\neg s_a \preccurlyeq \neg a$. Therefore we can't have $s_b < \neg a < \neg s_a$. We also can't have $b \in \mathcal{D}_b$ such that $s_b < b < \neg s_a$: by 3.36 we'd find $a \in \mathcal{D}_a$ such that $s_b < \neg a \preccurlyeq b < \neg s_a$ which was ruled out. So there are two cases. Either $s_b \not\equiv \neg s_a$ then $s_b < \neg s_a$: $\neg s_a$ is the immediate successor of s_b. Or $s_b \equiv \neg s_a$.

The same reasoning can be applied starting from $s_a \in \mathcal{D}_a$ to find a $s_b \in \mathcal{D}_b$ such that s_b is the immediate predecessor of $\neg s_a$ or an equivalent statement. This shows that all elements of D are paired.

To show that if a statement in D has a successor then it must be a before statement, let $s_1, s_2 \in D$ such that s_2 is the immediate successor of s_1. By 3.36, in all cases where $s_1 \notin \mathcal{D}_b$ and $s_2 \notin \mathcal{D}_a$ we can always find another statement between the two. Then we must have that $s_1 \in \mathcal{D}_b$ and $s_2 \in \mathcal{D}_a$. □

Theorem 3.38 (Reference ordering theorem). *An experimental domain is naturally ordered if and only if it can be generated by a set of refinable aligned strict references.*

Proof. Suppose $\mathcal{D}_X$ is an experimental domain generated by a set of refinable aligned strict references. Then by 3.34 and 3.37 the domain satisfies the requirement of theorem 3.16 and therefore is naturally ordered.

Now suppose $\mathcal{D}_X$ is naturally ordered. Define the set $\mathcal{B}_b$, $\mathcal{B}_a$ and D as in 3.12. Let $R = \{(b, \neg b \wedge \neg a, a) \mid b \in \mathcal{B}_b, a \in \mathcal{B}_a, b < \neg a\}$ be the set of all references constructed from the basis. First let us verify they are references. The before and after statements are verifiable since they are part of the basis. The on statement $\neg b \wedge \neg a$ is not a contradiction since $b < \neg a$ means $b \not\equiv a$ and $b \not\equiv \neg a$. The on statement is broader than $\neg b \wedge \neg a$ as they are equivalent and it is broader than $b \wedge a$ as that is a contradiction since $b < \neg a$. Therefore R is a set of references. Since the before and after statements of R coincide with the basis of the domain, $\mathcal{D}_X$ is generated by R.

Now we show that R consists of aligned strict references. We already saw that $b \not\equiv a$ and we also have $\neg b \wedge \neg a$ is incompatible with both b and a. The references are strict. To show they are aligned, take two references. The before and not after statements are linearly ordered by 3.14 which means the references are aligned.

To show R is refinable, note that all possibilities can be expressed as ("$x < x_1$", "$x_1 \le x \le x_2$", "$x > x_2$") where $x_1, x_2 \in X$ and "$x_1 \le x \le x_2$" ≡ "$x \ge x_1$" ∧ "$x \le x_2$". That is, every reference is identified by two possibilities x_1, x_2 such that $x_1 \le x_2$. Therefore take two references $r_1, r_2 \in R$ and let (x_1, x_2) and (x_3, x_4) be the respective pair of possibilities we can use to express the references as we have shown. Suppose $r_1 < r_2$ but they are not consecutive. Then "$x \le x_2$" < "$x < x_3$". That is, we can find $x_5 \in X$ such that $x_2 < x_5 < x_3$ which means "$x \le x_2$" < "$x < x_5$" and "$x \le x_5$" < "$x < x_3$". Therefore the reference $r_3 \in R$ identified by (x_5, x_5) is between the two references. On the other hand, suppose the second reference is finer than the first. Then $x_1 \le x_3$ and $x_4 \le x_2$ with either $x_1 \ne x_3$ or $x_4 \ne x_2$. Consider the references $r_3, r_4 \in R$ identified by (x_1, x_1) and (x_2, x_2). Either $r_3 < r_2$ or $r_2 < r_4$. Also note that the before statements of r_1 and r_3 are the same and the after statements of r_1 and r_4 are the same. Therefore they satisfy all the requirements and the set R is refinable by definition. □

To recap, experimentally we construct ordering by placing references and being able to tell whether the object measured is before or after. We can define a linear order on the possibilities, and therefore a quantity, only when the set of references meets special conditions. The references must be strict, meaning that before, on and after are mutually exclusive. They must be aligned, meaning that the before and not-after statement must be ordered by narrowness. They must be refinable, meaning when they overlap we can always find finer references with well defined before/after relationships. If all these conditions apply, we have a linear order. If any of these conditions fail, a linear order cannot be defined.

The possibilities, then, correspond to the finest references we can construct within the domain. That is, given a value q_0, we have the possibility *"the value of the property is q_0"* and we have the reference (*"the value of the property is less than q_0"*, *"the value of the property is q_0"*, *"the value of the property is more than q_0"*).

3.4 Discrete quantities

Now that we have seen the general conditions to have a naturally ordered experimental domain, we study common types of quantities and under what conditions they arise. We start with discrete ones: the number of chromosomes for a species, the number of inhabitants of a country or the atomic number for an element are all discrete quantities. These are quantities that are fully characterized by integers (positive or negative).

We will see that discrete quantities have a simple characterization: between two references there can only be a finite number of other references.

The first thing we want to do is characterize the ordering of the integers. That is, we want to find necessary and sufficient conditions for an ordered set of elements to be isomorphic to a subset of integers. First we note that between any two integers there are always finitely many elements. Let's call sparse an ordered set that has that property: that between two elements there are only finitely many. This is enough to say that the order is isomorphic to the integers.

Reference ordering theorem

To define an **ordered** sequence (e.g. of “instants”), the references must be (nec/suff conditions):

- Strict – an event is strictly before/on/after the reference (doesn’t extend over the “on”)
- Aligned – shared notion of before and after (logical relationship between statements)
- Refinable – overlaps can always be resolved

Gives you ordered points with the order topology

Additionally:

Dense order, which the closure on arbitrary union completes

Between any two references we can always have another reference ⇒ **real numbers**

Only finitely many references between any two references ⇒ **integers**

For time/space, these conditions are idealizations

https://assumptionsofphysics.org/

Assumptions of Physics

How does this model 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 “consistent” “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

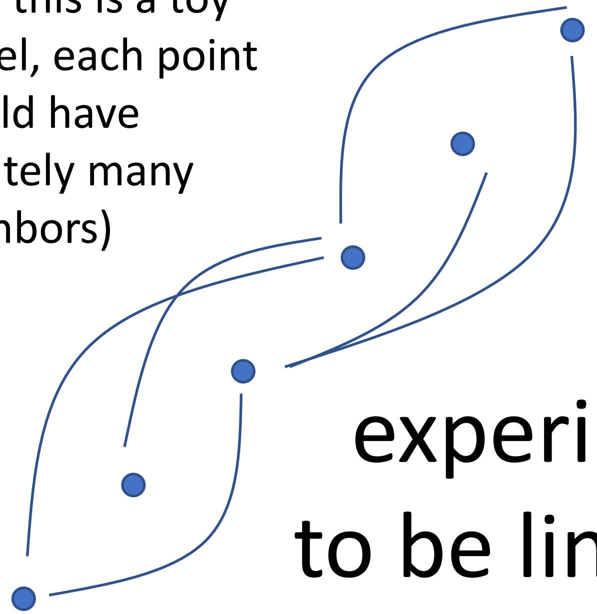
https://assumptionsofphysics.org/

Assumptions of Physics

What type of models should we use?

Hard to say, but we can argue from necessity

(N.B. this is a toy model, each point should have infinitely many neighbors)



Lack of order at small scales, order at large enough scale

What we can distinguish experimentally (i.e. topology) seems to be linked to how precisely we want to distinguish (i.e. geometry)

Current mathematical tools have a hard division between topology and geometry

Need new math?

https://assumptionsofphysics.org/

Assumptions of Physics

Takeaways

- Real numbers can be recovered from an idealized metrological model
  - Ordering of values comes from the ordering logical structure
  - $3 \leq 5$ precisely because “there are less than 3 items” $\preccurlyeq$ “there are less than 5 items”
- The hard part is the ordering, not the “continuity”
  - The difference between reals and integers is, as one intuitively expects, the ability to always find something in between two references
  - Completeness of the ordering is implied by the topology
- Failure of the idealization under real numbers would lead to a structure much richer (and more complicated) than a linear order

https://assumptionsofphysics.org/

Assumptions of Physics

Ensemble spaces
(generalized state spaces)

https://assumptionsofphysics.org/

Assumptions of Physics

Principle of scientific reproducibility. Scientific laws describe relationships that can always be experimentally reproduced.

⇒ Scientific laws are relationships between ensembles

Preparation procedure → Initial ensemble

Physical process → Final ensemble

Measurement procedure → Statistical data

A physical theory must AT LEAST describe which ensembles are possible within the theory

https://assumptionsofphysics.org/

Assumptions of Physics

	State space	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 $X = \{\mathbb{R}^{2n}, \omega\}$	$\mathcal{E} = \{\, \rho \in L^1(\mathbb{R}^{2n}) \mid \int \rho\, dq^n dp_n = 1\}$
Quantum mechanics	Projective complex Hilbert space $X = P(\mathcal{H})$	$\mathcal{E} = \{$ positive semi-definite Hermitian with $\mathrm{tr}(\rho) = 1 \}$

https://assumptionsofphysics.org/

Assumptions of Physics

Axiom 1.4 (Axiom of ensemble). *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.*

Experimental verifiability ⇒ topological space

Topology is responsible for handling limits and infinite operations

All other axioms are on finite elements

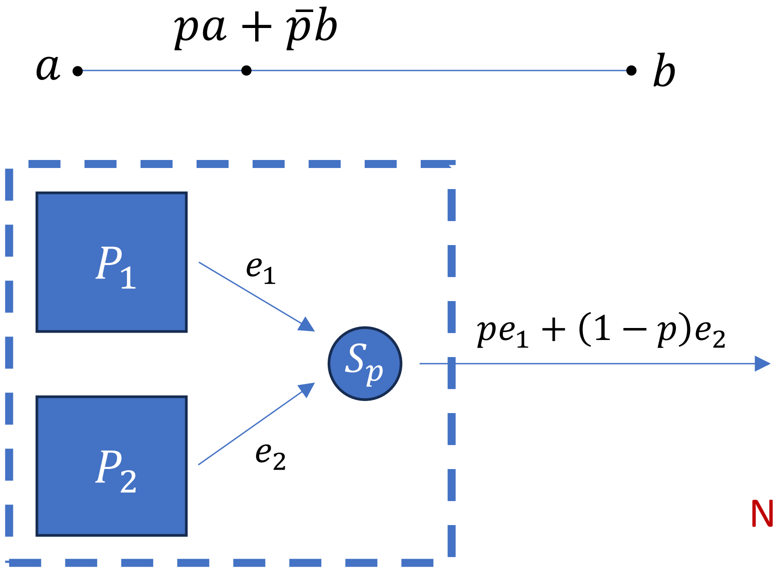
https://assumptionsofphysics.org/

Assumptions of Physics

Axiom 1.7 (Axiom of mixture). *The statistical mixture of two ensembles is an ensemble. Formally, 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 $\mathsf{p}\mathsf{a} + \bar{\mathsf{p}}\mathsf{b}$, with the following properties:*

- *Continuity: the map $+(p, \mathsf{a}, \mathsf{b}) \to \mathsf{p}\mathsf{a} + \bar{\mathsf{p}}\mathsf{b}$ is continuous (with respect to the product*

Statistical mixtures ⇒ Convex structure



Only finite mixtures $\sum_{i=1}^{n} p_i e_i$ are guaranteed

Topology tells us which
infinite mixtures $\sum_{i=1}^{\infty} p_i e_i$ converge

I.e. where experimental verifiability converges

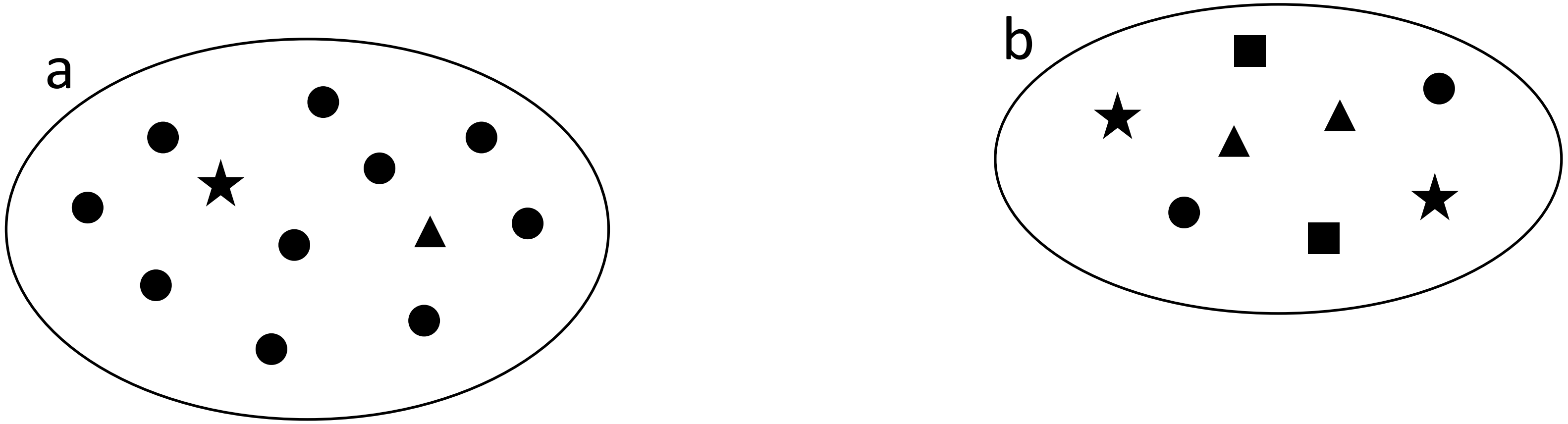
NB: the theory of topological vector spaces is well developed,
but not the theory of topological convex spaces

https://assumptionsofphysics.org/

Assumptions of Physics

Axiom 1.21 (Axiom of entropy). *Every element of the ensemble is associated with an **entropy** which quantifies the variability of the preparations of the ensemble. Formally, an ensemble space $\mathcal{E}$ is equipped with a function $S : \mathcal{E} \to \mathbb{R}$, defined up to a positive multiplicative constant representing the unit numerical value. The entropy has the following properties:*

- ***Continuity***[a]



Ensembles represent a collection of preparations which are, in general, not identical

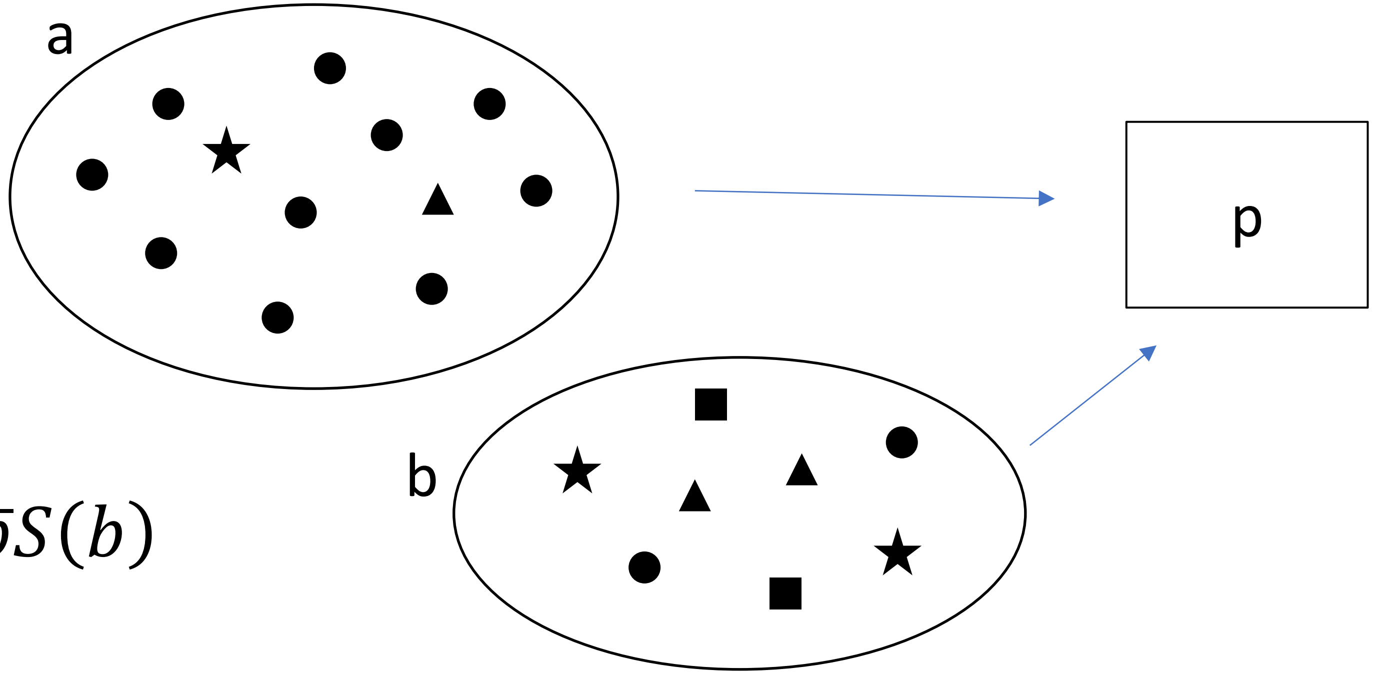
Variability: how similar are the preparations within an ensemble?

Entropy is a measure of variability

https://assumptionsofphysics.org/

Assumptions of Physics

- ***Strict concavity:*** $S(p\mathsf{a} + \bar{p}\mathsf{b}) \geq pS(\mathsf{a}) + \bar{p}S(\mathsf{b})$ *with the equality holding if and only if* $\mathsf{a} = \mathsf{b}$



$$S(pa + \bar{p}b)$$

Variability
after mixture

Average variability
before mixture

$$pS(a) + \bar{p}S(b)$$

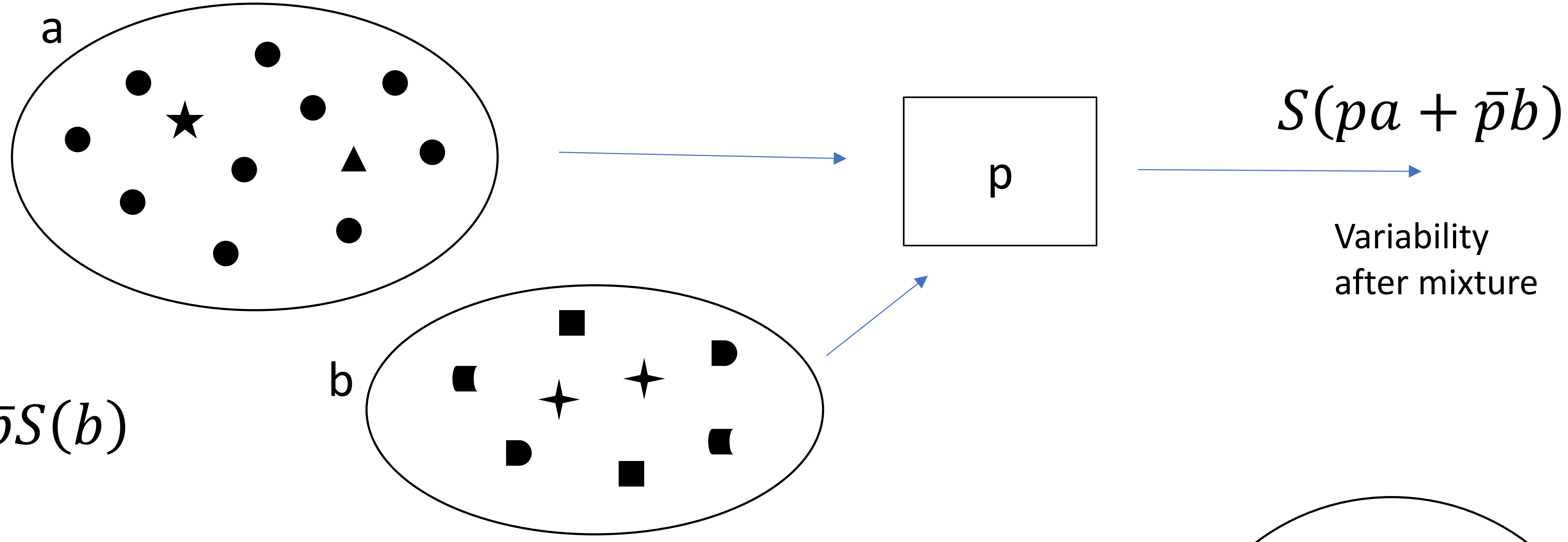
Final variability will be greater than average variability before mixture

If we mix an ensemble with itself, the variability is unchanged

https://assumptionsofphysics.org/

Assumptions of Physics

- ***Upper variability bound***: *there exists a universal function $I(p_1, p_2)$ (i.e. the same for all ensemble spaces) such that $S(\mathsf{pa} + \bar{p}\mathsf{b}) \leq I(p, \bar{p}) + pS(\mathsf{a}) + \bar{p}S(\mathsf{b})$; if the equality holds, a and b are* ***non-overlapping*** *or* ***orthogonal****, noted $\mathsf{a} \perp \mathsf{b}$*



Average variability
before mixture

$$pS(a) + \bar{p}S(b)$$

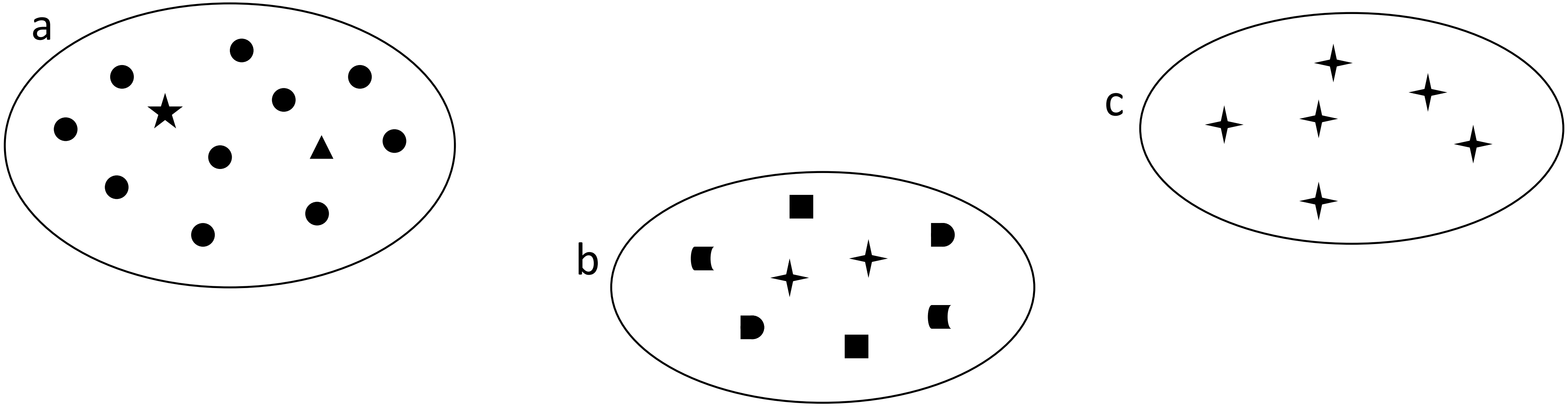
Maximum increase when the ensembles are “completely different”

Increase is only a function of the mixing coefficient

https://assumptionsofphysics.org/

Assumptions of Physics

- ***Mixtures preserve orthogonality:***[b] $\mathsf{a} \perp \mathsf{b}$ *and* $\mathsf{a} \perp \mathsf{c}$ *if and only if* $\mathsf{a} \perp p\mathsf{b} + \bar{p}\mathsf{c}$ *for any* $p \in (0, 1)$



If a has no elements in common with b or c, it has no elements in common with any mixture

https://assumptionsofphysics.org/

Assumptions of Physics

Axiom 1.21 (Axiom of entropy). *Every element of the ensemble is associated with an* ***entropy*** *which quantifies the variability of the preparations of the ensemble. Formally, an ensemble space $\mathcal{E}$ is equipped with a function $S : \mathcal{E} \to \mathbb{R}$, defined up to a positive multiplicative constant representing the unit numerical value. The entropy has the following properties:*

- ***Continuity***[a]
- ***Strict concavity:*** $S(p\mathsf{a} + \bar{p}\mathsf{b}) \geq pS(\mathsf{a}) + \bar{p}S(\mathsf{b})$ *with the equality holding if and only if* $\mathsf{a} = \mathsf{b}$
- ***Upper variability bound:*** *there exists a universal function $I(p_1, p_2)$ (i.e. the same for all ensemble spaces) such that $S(p\mathsf{a} + \bar{p}\mathsf{b}) \leq I(p, \bar{p}) + pS(\mathsf{a}) + \bar{p}S(\mathsf{b})$; if the equality holds,* a *and* b *are* ***non-overlapping*** *or* ***orthogonal****, noted* $\mathsf{a} \perp \mathsf{b}$
- ***Mixtures preserve orthogonality:***[b] $\mathsf{a} \perp \mathsf{b}$ *and* $\mathsf{a} \perp \mathsf{c}$ *if and only if* $\mathsf{a} \perp p\mathsf{b} + \bar{p}\mathsf{c}$ *for any* $p \in (0, 1)$

Ensemble variability ⇒ Entropy

No standard theory of entropic spaces

https://assumptionsofphysics.org/

Assumptions of Physics

Axiom 1.4 (Axiom of ensemble). *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*

Axiom 1.7 (Axiom of mixture). *The statistical mixture of two ensembles is an ensemble.*

Axiom 1.21 (Axiom of entropy). *Every element of the ensemble is associated with an* ***entropy*** *which quantifies the variability of the preparations of the ensemble. Formally, an*

These axioms specify very minimal **necessary** requirements on physical theories

How much can we derive just from these?

Physical mathematics starts with minimal requirements to force us to understand which requirements are truly independent and truly necessary

Here are a few things we are able to derive…

https://assumptionsofphysics.org/

Assumptions of Physics

The entropy upper bound $I(p, \bar{p})$ is uniquely determined

Theorem 1.25 (Uniqueness of entropy). *The entropy of the coefficients $I(p, \bar{p})$ is the Shannon entropy. That is, $I(p, \bar{p}) = -\kappa\,(p \log p + \bar{p} \log \bar{p})$ where $\kappa > 0$ is the arbitrary multiplicative constant for the entropy. For a mixture of arbitrarily many elements, $I(\{p_i\}) = -\kappa \sum_i p_i \log p_i$.*

Shannon entropy

Maximal increase in variability during mixing

Let’s see how it works

https://assumptionsofphysics.org/

Assumptions of Physics

Pick a_i all orthogonal to each other

$$S\left(\sum_i p_i a_i\right) = I(p_i) + \sum_i p_i S(a_i)$$

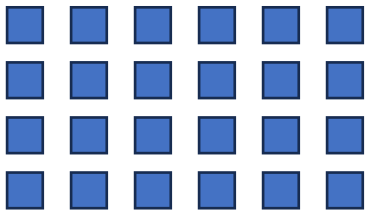
$$S\left(\sum_{j=1}^{n}\sum_{k=1}^{m}\frac{1}{n}\frac{1}{m}\mathsf{a}_{jk}\right) = I\left(\left\{\frac{1}{n}\frac{1}{m}\right\}_{i=1}^{nm}\right) + \sum_{j=1}^{n}\sum_{k=1}^{m}\frac{1}{n}\frac{1}{m}S\left(\mathsf{a}_{jk}\right)$$

$$= S\left(\sum_{j=1}^{n}\frac{1}{n}\sum_{k=1}^{m}\frac{1}{m}\mathsf{a}_{jk}\right) = I\left(\left\{\frac{1}{n}\right\}_{i=1}^{n}\right) + \sum_{j=1}^{n}\frac{1}{n}S\left(\sum_{k=1}^{m}\frac{1}{m}\mathsf{a}_{jk}\right)$$

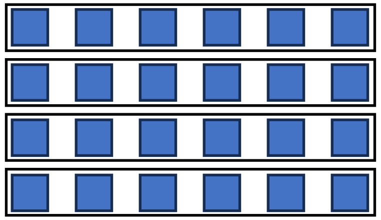
$$= I\left(\left\{\frac{1}{n}\right\}_{i=1}^{n}\right) + \sum_{j=1}^{n}\frac{1}{n}\left(I\left(\left\{\frac{1}{m}\right\}_{i=1}^{m}\right) + \sum_{k=1}^{m}\frac{1}{m}S\left(\mathsf{a}_{jk}\right)\right)$$

$$= I\left(\left\{\frac{1}{n}\right\}_{i=1}^{n}\right) + I\left(\left\{\frac{1}{m}\right\}_{i=1}^{m}\right) + \sum_{j=1}^{n}\sum_{k=1}^{m}\frac{1}{n}\frac{1}{m}S\left(\mathsf{a}_{jk}\right).$$

Uniform mixture of $n \times m$ elements



Must be the same



Uniform mixture over n uniform mixtures of m elements

$$I\left(\left\{\frac{1}{nm}\right\}_{i=1}^{nm}\right) = I\left(\left\{\frac{1}{n}\right\}_{i=1}^{n}\right) + I\left(\left\{\frac{1}{m}\right\}_{i=1}^{m}\right).$$

$$I\left(\left\{\frac{1}{n}\right\}_{i=1}^{n}\right) = \kappa \log n$$

https://assumptionsofphysics.org/

Assumptions of Physics

Pick a_i all orthogonal to each other

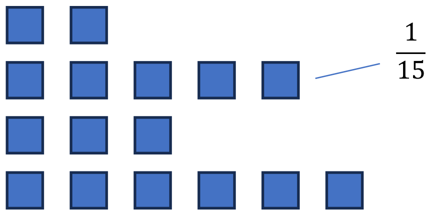
$$S(\textstyle\sum_i p_i a_i) = I(p_i) + \textstyle\sum_i p_i S(a_i)$$

$$\begin{aligned}
S\left(\sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{1}{m}\mathsf{a}_{ij}\right) &= I\left(\left\{\frac{1}{m}\right\}_{i=1}^{m}\right) + \sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{1}{m}S\left(\mathsf{a}_{ij}\right)\\
&= \kappa \log m + \sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{1}{m}S\left(\mathsf{a}_{ij}\right)
\end{aligned}$$

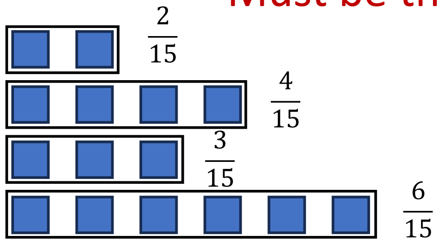
$$\begin{aligned}
= S\left(\sum_{i=1}^{n}\frac{m_i}{m}\sum_{j=1}^{m_i}\frac{1}{m_i}\mathsf{a}_{ij}\right) &= I\left(\left\{\frac{m_i}{m}\right\}_{i=1}^{n}\right) + \sum_{i=1}^{n}\frac{m_i}{m}S\left(\sum_{j=1}^{m_i}\frac{1}{m_i}\mathsf{a}_{ij}\right)\\
&= I\left(\left\{p_i\right\}_{i=1}^{n}\right) + \sum_{i=1}^{n}\frac{m_i}{m}\left(I\left(\left\{\frac{1}{m_i}\right\}_{i=1}^{m_i}\right) + \sum_{j=1}^{m_i}\frac{1}{m_i}S\left(\mathsf{a}_{ij}\right)\right)\\
&= I\left(\left\{p_i\right\}_{i=1}^{n}\right) + \sum_{i=1}^{n}\frac{m_i}{m}I\left(\left\{\frac{1}{m_i}\right\}_{i=1}^{m_i}\right) + \sum_{i=1}^{n}\frac{m_i}{m}\sum_{j=1}^{m_i}\frac{1}{m_i}S\left(\mathsf{a}_{ij}\right)\\
&= I\left(\left\{p_i\right\}_{i=1}^{n}\right) + \sum_{i=1}^{n}p_i\kappa\log m_i + \sum_{i=1}^{n}\sum_{j=1}^{m_i}\frac{1}{m}S\left(\mathsf{a}_{ij}\right).
\end{aligned}$$

$$\kappa \log m = I\left(\left\{p_i\right\}_{i=1}^{n}\right) + \sum_{i=1}^{n}p_i\kappa\log m_i$$

Uniform mixture of m elements



Must be the same



Rational distribution over
uniform mixtures of grouped elements

https://assumptionsofphysics.org/

Assumptions of Physics

$$\kappa \log m = I\left(\{p_i\}_{i=1}^{n}\right) + \sum_{i=1}^{n} p_i \kappa \log m_i$$

$$I\left(\{p_i\}_{i=1}^{n}\right) = \kappa \log m - \sum_{i=1}^{n} p_i \kappa \log m_i = \sum_{i=1}^{n} p_i \kappa \log m - \sum_{i=1}^{n} p_i \kappa \log m_i$$

$$= -\sum_{i=1}^{n} p_i \kappa \log \frac{m_i}{m} = -\kappa \sum_{i=1}^{n} p_i \log p_i .$$

$$I(p, \bar{p}) = -p \log p - \bar{p} \log \bar{p}$$

From now on,
$\log p$ is base 2

Proof “does not know” whether we are dealing with classical ensembles, quantum ensembles, or ensembles for a theory yet to be discovered

Proof is short (about two pages)

Two ways to define “exclusive” ensembles

Separate ensembles

$$\mathsf{a} \;⫪\; \mathsf{b}$$

No “common component” $\mathsf{c} \in \mathcal{E}$

such that
$$\mathsf{a} = p_1\mathsf{c} + \bar{p}_1\mathsf{e}_1$$
$$\mathsf{b} = p_2\mathsf{c} + \bar{p}_2\mathsf{e}_2$$

$$\Leftarrow$$

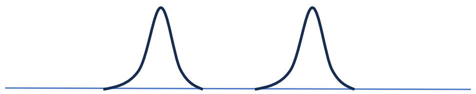
Orthogonal ensembles

$$\mathsf{a} \perp \mathsf{b}$$

Saturate upper entropy bound

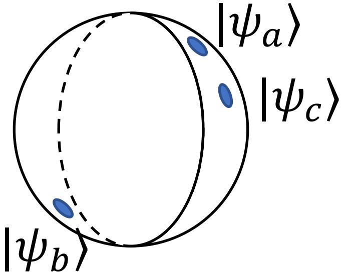
$$S(p\mathsf{a} + \bar{p}\mathsf{b}) = I(p, \bar{p}) + pS(\mathsf{a}) + \bar{p}S(\mathsf{b})$$

Coincide in classical ensemble spaces



disjoint support

Different in quantum ensemble spaces



Simple but powerful definitions from basic axioms

https://assumptionsofphysics.org/

Assumptions of Physics

Vector space constraints from entropy

https://assumptionsofphysics.org/

Assumptions of Physics

Definition 1.42. *A convex space X is* ***cancellative*** *if* $\mathsf{p}\mathsf{a} + \bar{p}\mathsf{e} = p\mathsf{b} + \bar{p}\mathsf{e}$ *for some* $p \in (0,1)$ *implies* $\mathsf{a} = \mathsf{b}$.

Theorem 1.43 (Ensemble spaces are cancellative)**.** *Let $\mathcal{E}$ be an ensemble space. Let* $\mathsf{a}, \mathsf{b}, \mathsf{e} \in \mathcal{E}$ *such that* $p\mathsf{a} + \bar{p}\mathsf{e} = p\mathsf{b} + \bar{p}\mathsf{e}$ *for some* $p \in (0,1)$*. Then* $\mathsf{a} = \mathsf{b}$.

Entropy bounds force mixing to be “invertible”

Definition 1.53 (Affine combinations)**.** *Let* $\{\mathsf{e}_i\}_{i=1}^n \subseteq \mathcal{E}$ *be a finite sequence of ensembles and* $\{r_i\}_{i=1}^n \subseteq \mathbb{R}$ *be a finite sequence of coefficients such that* $\sum_{i=1}^n r_i = 1$*. The* ***affine combination*** $\sum_{i=1}^n r_i \mathsf{e}_i$ *is, if it exists, the ensemble* $\mathsf{a} \in \mathcal{E}$ *such that* $\sum_{i \in I} \frac{r_i}{r} \mathsf{e}_i = \frac{1}{r}\mathsf{a} + \sum_{i \notin I} \frac{-r_i}{r} \mathsf{e}_i$ *where* $I = \{i \in [1,n] \mid r_i \geq 0\}$ *and* $r = \sum_{i \in I} r_i$.

Can define affine combinations (i.e. negative probabilities)

https://assumptionsofphysics.org/

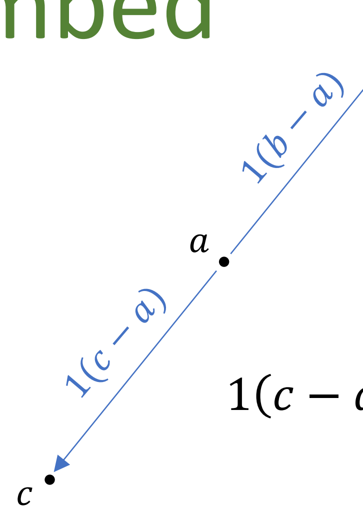
Assumptions of Physics

Definition 1.55 (Ensemble differences). *Given an ensemble space, a difference between two ensemble represents the change required to transform one ensemble into another. Formally, an* ***ensemble difference****, noted* $r(\mathsf{b}-\mathsf{a})$*, is a triple formed by a real number* $r \in \mathbb{R}$ *and an ordered pair of ensembles* $\mathsf{a}, \mathsf{b} \in \mathcal{E}$.

Theorem 1.65 (Differences from a vector space). *Let* $\mathsf{a} \in \mathcal{E}$ *be an interior point and let* $V = \{[r(\mathsf{b}-\mathsf{a})]\}$ *be the set of equivalence classes of ensemble differences from* a. *Then* V *is a vector space under the scalar multiplication and addition.*

Definition 1.69. *Given an internal point* a*, the* ***natural embedding*** *of* $\mathcal{E}$ *into* V_{a} *is the map* $\iota_{\mathsf{a}} : \mathcal{E} \hookrightarrow V_{\mathsf{a}}$*, defined as* $\iota_{\mathsf{a}}(\mathsf{e}) \to [(\mathsf{e}-\mathsf{a})]$*, that maps each ensemble to its difference from* a.

Ensemble spaces embed into vector spaces



Vector space operations derived from mixing operations

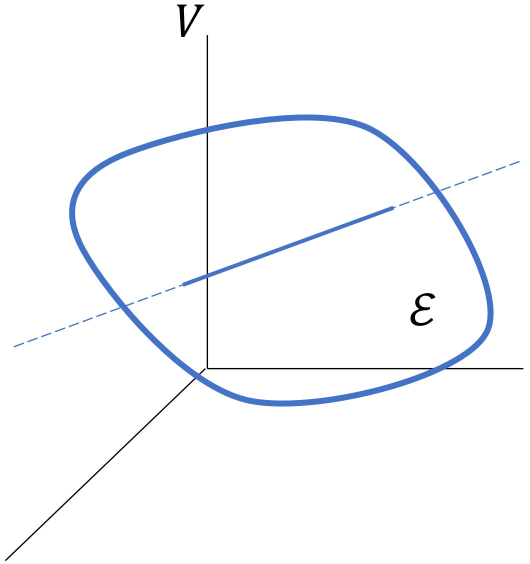
$$\frac{1}{2}c + \frac{1}{2}b = a$$

$$\Updownarrow$$

$$1(c-a) = -1(b-a)$$

Definition 1.50. *A **line** $A \subseteq \mathcal{E}$ is a convex subset such that for any three elements one can be expressed as a mixture of the other two. That is, for all $\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3 \in A$ there exists a permutation $\sigma : \{1,2,3\} \to \{1,2,3\}$ and $p \in [0,1]$ such that $\mathsf{e}_{\sigma(1)} = p\mathsf{e}_{\sigma(2)} + \bar{p}\mathsf{e}_{\sigma(3)}$.*

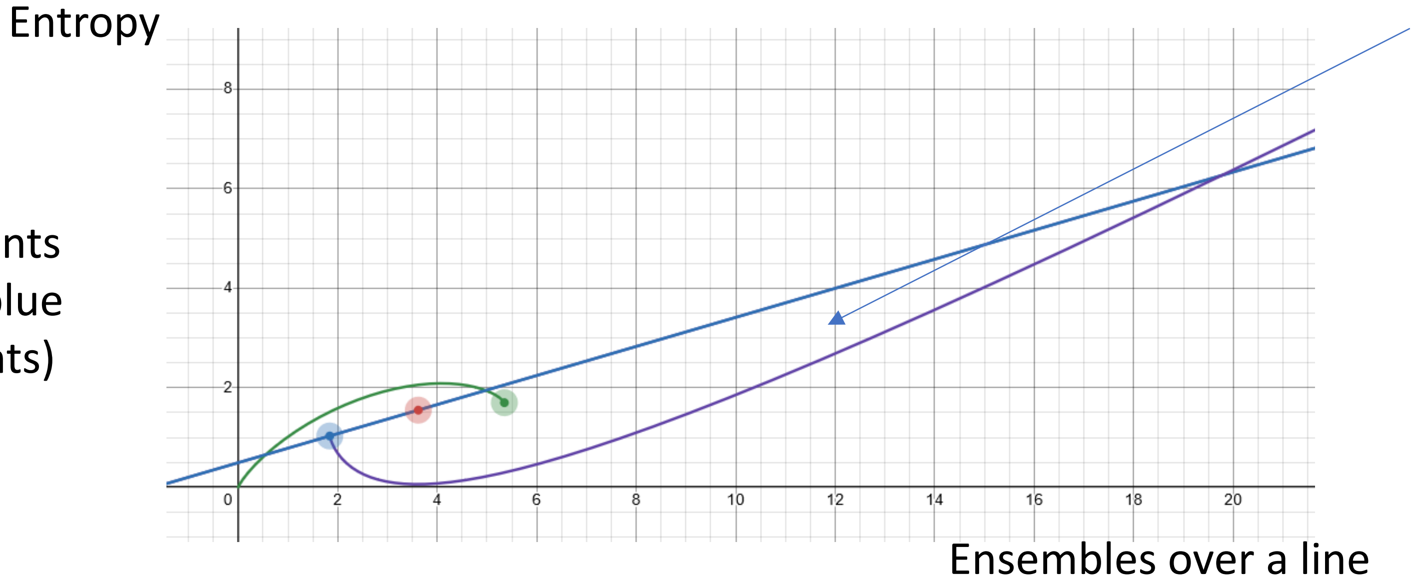
Theorem 1.52 (Lines are bounded). *Let $A \subseteq \mathcal{E}$ be a line. Then we can find a bounded interval $V \subseteq \mathbb{R}$ and an invertible function $f : A \to V$ such that $f(p\mathsf{a} + \bar{p}\mathsf{b}) = pf(\mathsf{a}) + \bar{p}f(\mathsf{b})$ for all $\mathsf{a}, \mathsf{b} \in A$.*



Ensemble spaces are bounded in all directions

Entropy bounds ⇒ green point between blue and purple curves

Fix three points (i.e. origin, blue and red points)



https://assumptionsofphysics.org/

Assumptions of Physics

Geometric structures from entropy

https://assumptionsofphysics.org/

Assumptions of Physics

How much does the entropy increase during mixture?

$$MS(\mathsf{a},\mathsf{b}) = S\left(\frac{1}{2}\mathsf{a} + \frac{1}{2}\mathsf{b}\right) - \left(\frac{1}{2}S(\mathsf{a}) + \frac{1}{2}S(\mathsf{b})\right)$$

1. *non-negativity:* $MS(\mathsf{a},\mathsf{b}) \geq 0$
2. *identity of indiscernibles:* $MS(\mathsf{a},\mathsf{b}) = 0 \iff \mathsf{a} = \mathsf{b}$
3. *unit boundedness:* $MS(\mathsf{a},\mathsf{b}) \leq 1$
4. *maximality of orthogonals:* $MS(\mathsf{a},\mathsf{b}) = 1 \iff \mathsf{a} \perp \mathsf{b}$
5. *symmetry:* $MS(\mathsf{a},\mathsf{b}) = MS(\mathsf{b},\mathsf{a})$

Recovers the Jensen-Shannon divergence (JSD)
(both classical and quantum)

Pseudo-distance defined from the entropy

(does not satisfy the triangle inequality)

https://assumptionsofphysics.org/

Assumptions of Physics

Entropy imposes a metric on the ensemble space

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

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

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

Entropy strict concavity means the Hessian is negative definite

Recovers Fisher-Rao information metric
(both classical and quantum)

https://assumptionsofphysics.org/

Assumptions of Physics

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

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

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

Another direct calculation from the definitions

https://assumptionsofphysics.org/

Assumptions of Physics

Measure theoretic structures from entropy

https://assumptionsofphysics.org/

Assumptions of Physics

In classical mechanics the entropy of a uniform distribution ρ_U over U is $S(\rho_U) = \log \mu(U)$

Count of states
(Phase-space volume)

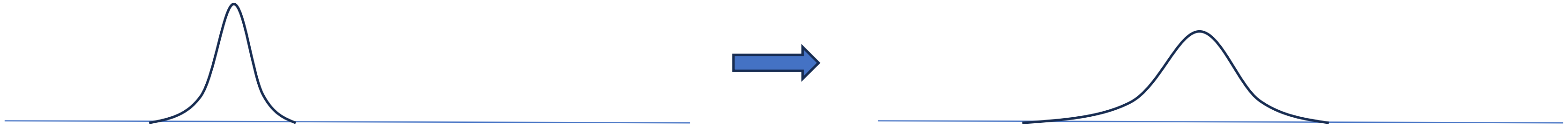
In ensemble spaces, entropy is the primitive notion. Can we define a notion of count of states that recovers the classical expression, but makes sense in the general theory, including quantum mechanics?

Note: given the set of all distributions with support U, the uniform distribution maximizes the entropy

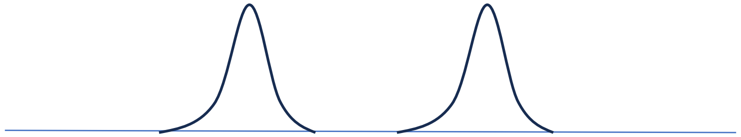
https://assumptionsofphysics.org/

Assumptions of Physics

We can think of an ensemble as spread over distinguishable cases



The number of distinguishable cases must increase with entropy



During mixing, the number of distinguishable cases at most sums

Note:

Proposition 1.153 (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 orthogonal and* $p = \frac{2^{S_1}}{2^{S_1}+2^{S_2}}$.

https://assumptionsofphysics.org/

Assumptions of Physics

Proposition 1.153 (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 orthogonal and* $p = \frac{2^{S_1}}{2^{S_1}+2^{S_2}}$.

Maximum increase when ensembles are orthogonal.
Find p that maximizes entropy:

$$\begin{aligned}
0 = \frac{d}{dp} S(p\mathsf{a} + \bar{p}\mathsf{b}) &= \frac{d}{dp}(-p\log p - \bar{p}\log\bar{p} + pS_a + \bar{p}S_b) \\
&= -\log p - 1 + \log \bar{p} + 1 + S_a - S_b \\
\log\frac{p}{\bar{p}} &= \log 2^{S_a} - \log 2^{S_b} \\
\log\frac{p}{1-p} &= \log\frac{2^{S_a}}{2^{S_b}} \\
p 2^{S_b} &= (1-p)2^{S_a} \\
p(2^{S_a} + 2^{S_b}) &= 2^{S_a} \\
p &= \frac{2^{S_a}}{2^{S_a}+2^{S_b}}
\end{aligned}$$

Then calculate final entropy:

$$\begin{aligned}
\bar{p} = 1 - \frac{2^{S_a}}{2^{S_a}+2^{S_b}} &= \frac{2^{S_b}}{2^{S_a}+2^{S_b}} \\
S(p\mathsf{a} + \bar{p}\mathsf{b}) &= -p\log p - \bar{p}\log\bar{p} + pS_a + \bar{p}S_b \\
&= -\frac{2^{S_a}}{2^{S_a}+2^{S_b}}\log\frac{2^{S_a}}{2^{S_a}+2^{S_b}} - \frac{2^{S_b}}{2^{S_a}+2^{S_b}}\log\frac{2^{S_b}}{2^{S_a}+2^{S_b}} \\
&\quad + \frac{2^{S_a}}{2^{S_a}+2^{S_b}}\log 2^{S_a} + \frac{2^{S_b}}{2^{S_a}+2^{S_b}}\log 2^{S_b} \\
&= \frac{2^{S_a}}{2^{S_a}+2^{S_b}}\log\left(2^{S_a}+2^{S_b}\right) + \frac{2^{S_b}}{2^{S_a}+2^{S_b}}\log\left(2^{S_a}+2^{S_b}\right) \\
&= \frac{2^{S_a}+2^{S_b}}{2^{S_a}+2^{S_b}}\log\left(2^{S_a}+2^{S_b}\right) \\
\log 2^{S(p\mathsf{a}+\bar{p}\mathsf{b})} &= \log\left(2^{S_a}+2^{S_b}\right) \\
2^{S(p\mathsf{a}+\bar{p}\mathsf{b})} &= 2^{S_a}+2^{S_b}
\end{aligned}$$

Again, simple calculation that does not depend on type of space

https://assumptionsofphysics.org/

Assumptions of Physics

Given a set of possible ensembles A, the count of configurations is the exponential of the maximum entropy reachable using mixtures. If A is the set of classical distributions over a particular support U, the maximum entropy is given by the uniform distribution ⇒ recovers the usual count of states! If A is the set of density matrices that has zero eigenvalues outside of a subspace H, the state capacity recovers the dimensionality of the space ⇒ count of distinguishable states!

Definition 4.133. *Let $A \subseteq \mathcal{E}$ be a subset of an ensemble space. The **state capacity** of A is defined as* $\mathrm{scap}(A) = \sup(2^{S(\mathrm{hull}(A))} \cup \{0\})$.

capacity also name
of a non-additive measure

Proposition 4.134. *The state capacity is a set function that is*

1. *non-negative:* $\mathrm{scap}(A) \in [0, +\infty]$
2. *monotone:* $A \subseteq B \implies \mathrm{scap}(A) \leq \mathrm{scap}(B)$
3. *subadditive:* $\mathrm{scap}(A \cup B) \leq \mathrm{scap}(A) + \mathrm{scap}(B)$
4. *additive over orthogonal sets:* $A \perp B \implies \mathrm{scap}(A \cup B) = \mathrm{scap}(A) + \mathrm{scap}(B)$

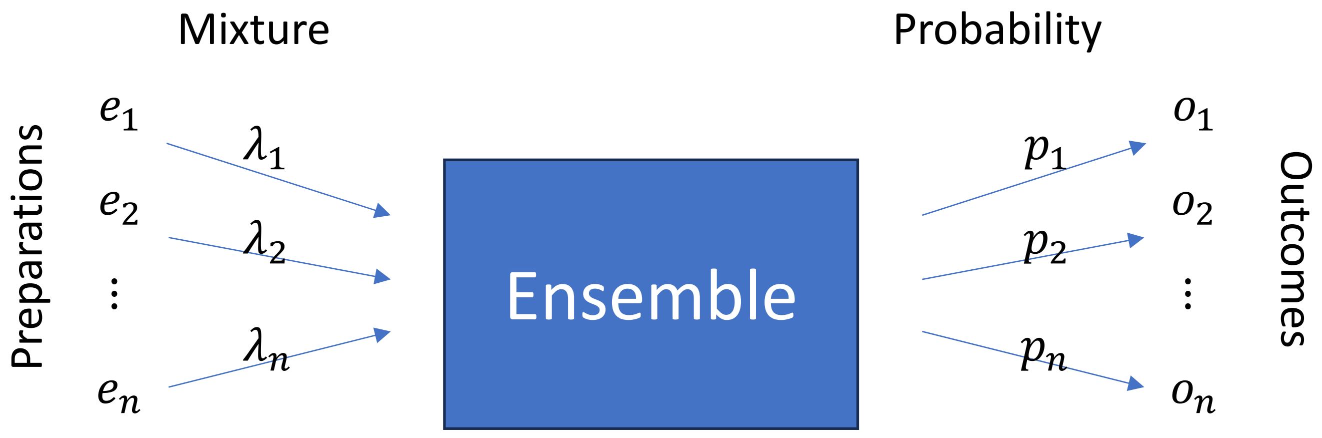
fuzzy measure

State capacity is a non-additive measure

additive over orthogonal sets

https://assumptionsofphysics.org/

Assumptions of Physics



In classical mechanics, mixtures of preparations and probability of outcomes always coincide

In quantum mechanics, they do not

⇒ quantum ensemble spaces not simplexes (i.e. classical probability fails)

Can we have common measure theoretic tools on the preparation side?

https://assumptionsofphysics.org/

Assumptions of Physics

How much of e is a mixture of other ensembles?

$$e = p\left(\sum_i \lambda_i a_i\right) + \bar{p}b$$

biggest p

Definition 1.83. *Let* $\mathsf{e}, \mathsf{a} \in \mathcal{E}$ *be two ensembles. The* ***fraction*** *of* a *in* e *is the greatest mixing coefficient for which* e *can be expressed as a mixture of* a. *That is,* $\mathrm{frac}_{\mathsf{e}}(\mathsf{a}) = \sup(\{p \in [0,1] \mid \exists\, \mathsf{b} \in \mathcal{E} \text{ s.t. } \mathsf{e} = p\mathsf{a} + \bar{p}\mathsf{b}\})$.

Definition 1.85. *Let* $\mathsf{e} \in \mathcal{E}$ *be an ensemble and* $A \subseteq \mathcal{E}$ *a Borel set. The* ***fraction capacity*** *of* A *for* e *is the biggest fraction achievable with convex combinations of* A. *That is,* $\mathrm{fcap}_{\mathsf{e}}(A) = \sup(\mathrm{frac}_{\mathsf{e}}(\mathrm{hull}(A)) \cup \{0\})$.

Proposition 1.87. *The fraction capacity for an ensemble is a set function that is*

1. *non-negative and unit bounded:* $\mathrm{fcap}_{\mathsf{e}}(A) \in [0,1]$
2. *monotone:* $A \subseteq B \implies \mathrm{fcap}_{\mathsf{e}}(A) \leq \mathrm{fcap}_{\mathsf{e}}(B)$
3. *subadditive:* $\mathrm{fcap}_{\mathsf{e}}(A \cup B) \leq \mathrm{fcap}_{\mathsf{e}}(A) + \mathrm{fcap}_{\mathsf{e}}(B)$
4. *continuous from below:* $\mathrm{fcap}_{\mathsf{e}}(\lim_{i\to\infty} A_i) = \lim_{i\to\infty} \mathrm{fcap}_{\mathsf{e}}(A_i)$ *for any increasing sequence* $\{A_i\}$
5. *continuous from above:* $\mathrm{fcap}_{\mathsf{e}}(\lim_{i\to\infty} A_i) = \lim_{i\to\infty} \mathrm{fcap}_{\mathsf{e}}(A_i)$ *for any decreasing sequence* $\{A_i\}$

fuzzy measure

Fraction capacity is a non-additive probability measure

https://assumptionsofphysics.org/

Assumptions of Physics

Takeaways

- The principle of scientific reproducibility requires the notion of ensembles
- We can develop a theory of states and processes using ensembles
  - Few physically justifiable starting points lead to a rich structure
  - Many definitions and proofs can be generalized in this setting
- These notions provide a core foundation that can link to various standard mathematical theories
  - Topological vector spaces are the foundation of modern functional analysis
  - Information geometry can be shown to link to symplectic geometry of classical mechanics and the inner product of quantum mechanics
  - Non-additive measures provide a generalization of classical measure theoretic structures
- Still a lot of work that needs to be done to complete the theory

https://assumptionsofphysics.org/

Assumptions of Physics

Wrapping it up

- Physical mathematics: derive the math required from physical requirements
  - In physics, mathematics is used to model physical systems, therefore we need mathematics that is designed specifically for that purpose
- Principle of scientific objectivity: science deals with evidence-based assertions
  - Requires notion of verifiable statements
  - Leads to topological spaces (open sets corresponds to verifiable statements) and σ-algebras (Borel sets correspond to statements associated with tests)
  - Real numbers can be derived by modeling an idealized reference system
- Principle of scientific reproducibility: scientific laws describe reproducible relationships
  - Requires notion of ensembles
  - Ensembles must be experimentally well-defined, allow statistical mixture and be associated with an entropy (that quantifies the variability of the instances of an ensemble)
  - Recovers notions of vector spaces, geometry and measure theory
- Hopefully this shows that we can build the required math from the ground up

https://assumptionsofphysics.org/

Assumptions of Physics

https://assumptionsofphysics.org/

Assumptions of Physics