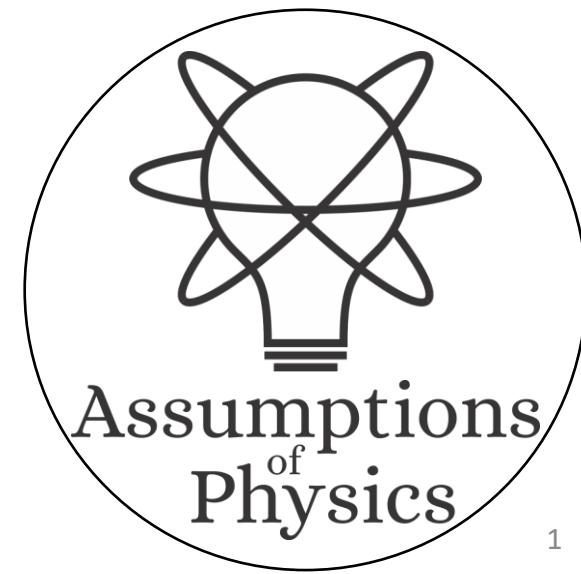


Assumptions of Physics
Summer School 2026

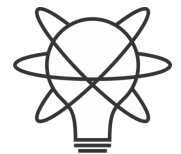
Some results in Physical Mathematics

Gabriele Carcassi, Christine A. Aidala, and **Tobias Thrien**

Physics Department
University of Michigan



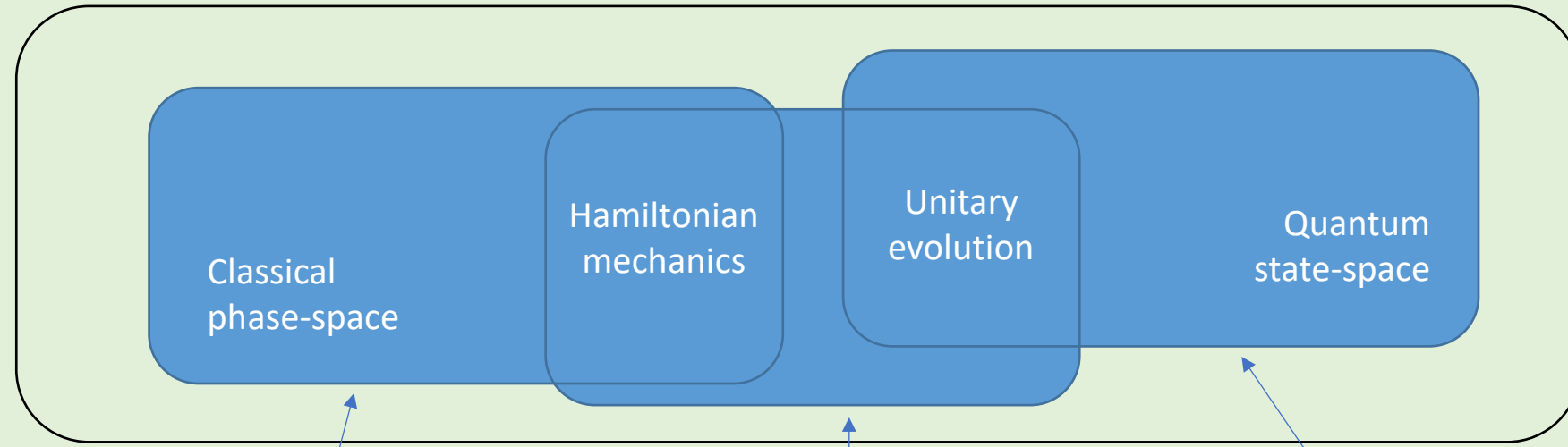
Overview



<https://assumptionsofphysics.org/>

Assumptions
of
Physics

Space of the well-posed scientific theories



Physical theories

Specializations of the general theory under the different assumptions

Assumptions

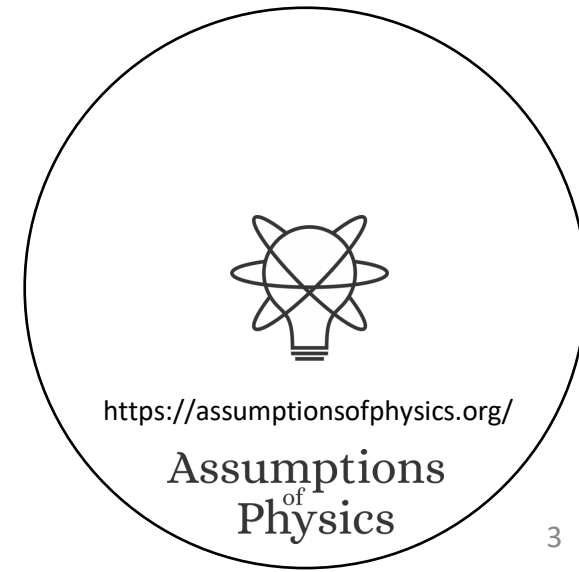
States and processes

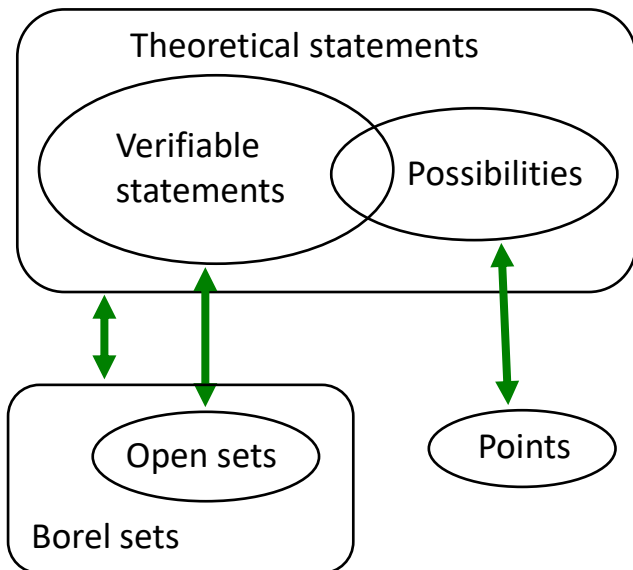
Information granularity

Experimental verifiability

General theory

Basic requirements and definitions valid in all theories





Experimental verifiability



Topology and σ -algebra

Quantities from references

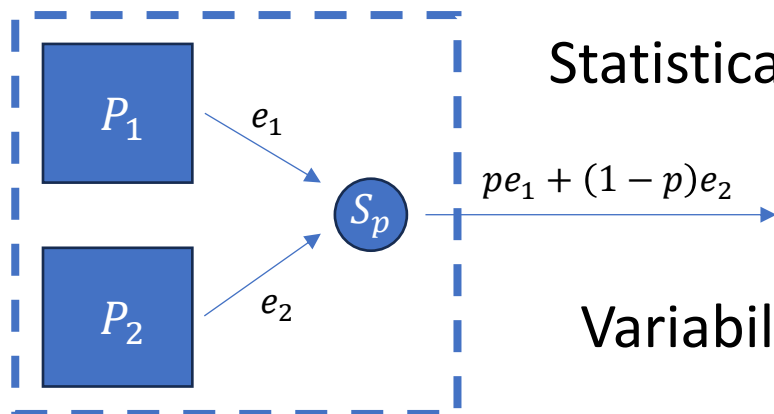


Refinable aligned
strict references



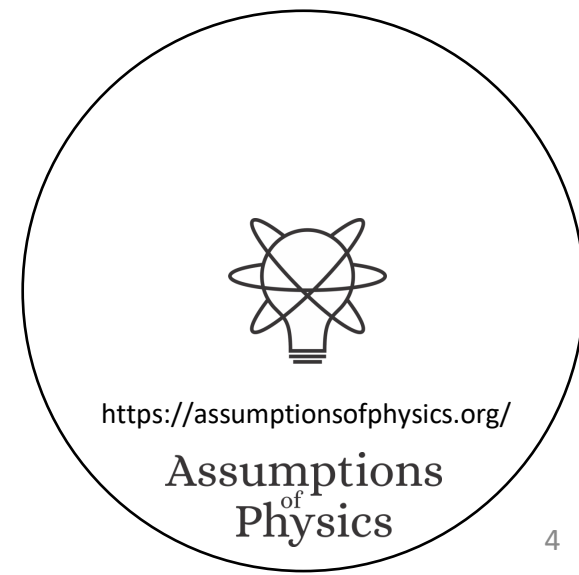
Real numbers

References act as
Dedekind cuts

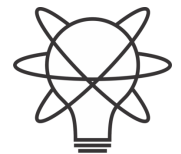


Statistical mixtures $\Rightarrow$ Convex (linear) structures

Variability/entropy $\Rightarrow$ Geometric structures



Recover the reals



<https://assumptionsofphysics.org/>

Assumptions
of
Physics

In physics, we use real numbers to model physical quantities

$\mathbb{R}$

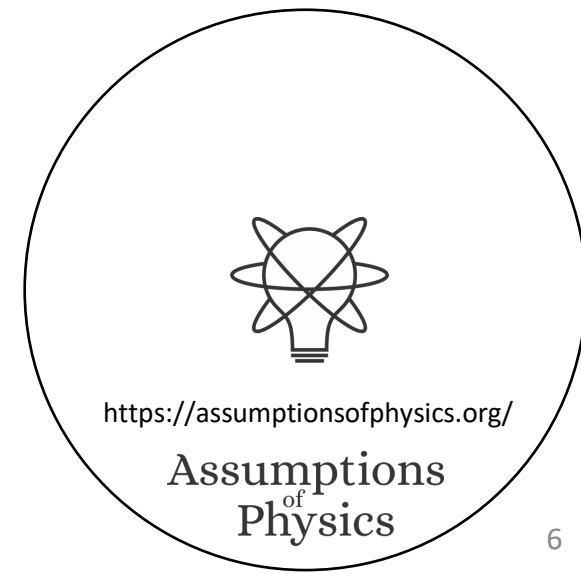
Needed for a lot of mathematical tools: exponential, gaussians, trigonometry, ...

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?

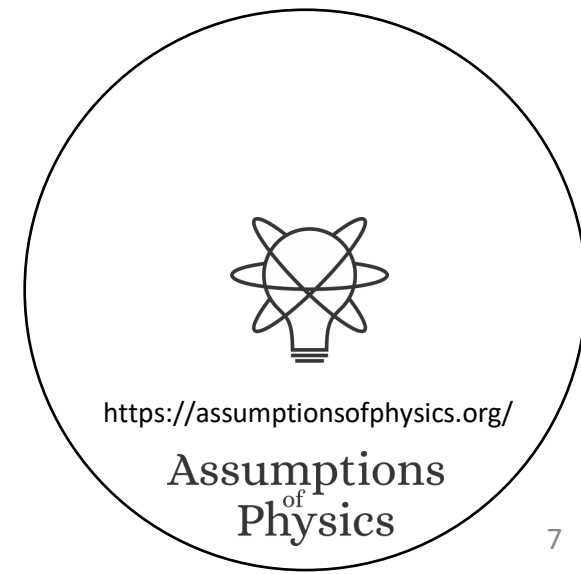


Order topology

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!!!



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

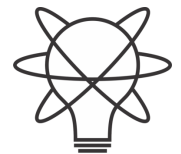
No greatest or smallest

2) The completion of a linear order is unique

Every subset has a supremum and an infimum

⇒ 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



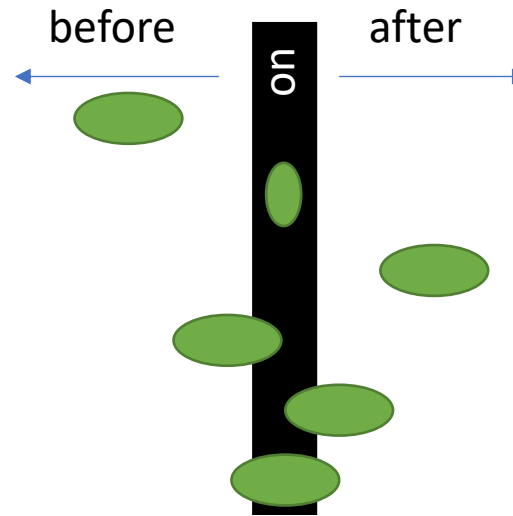
<https://assumptionsofphysics.org/>

Assumptions
of
Physics

How is order defined experimentally?

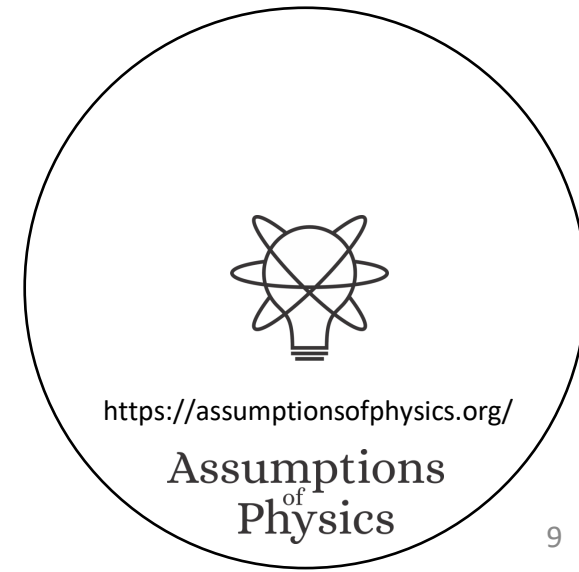
We have a way to compare two things, and decide whether one is bigger/smaller than the other. “Exactly the same” is a bit more difficult.

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

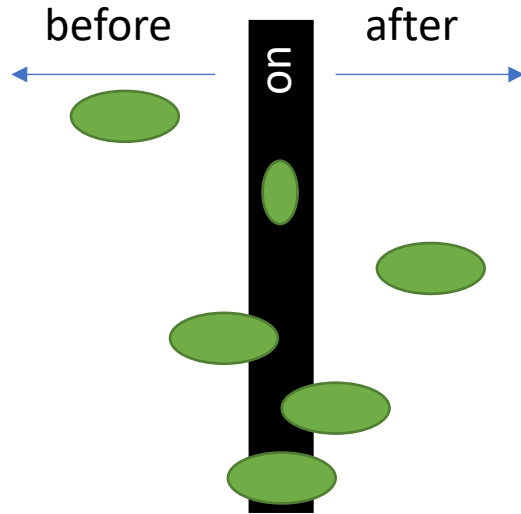


“Before/after a reference” are verifiable statements

wikipedia



Goal: make a mathematical model of a reference system



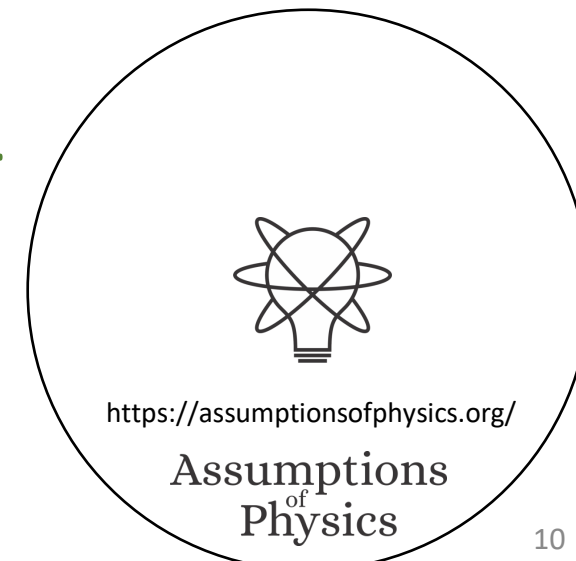
Definition 3.17. A *reference* defines a before, an on and an after relationship between itself and another object. Formally a reference $r = (b, o, a)$ is a tuple of three statements such that:

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: $o \neq \perp$
3. if it's not before or after, it's on the reference: $\neg b \wedge \neg a \leq o$
4. if it's before and after, it's also on the reference: $b \wedge a \leq o$

A *beginning reference* has nothing before it. That is, $b \equiv \perp$. An *ending reference* has nothing after it. That is, $a \equiv \perp$. 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

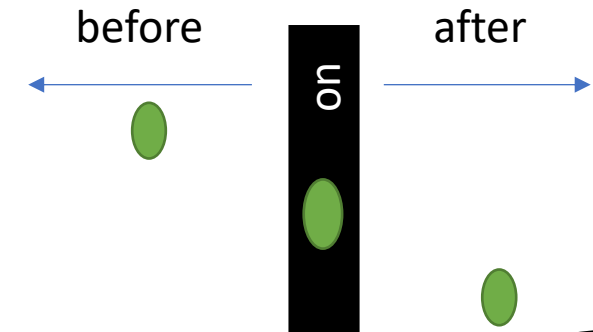
A reference system is a collection of references. The experimental domain for a quantity is the set of verifiable statements generated by all possible before/after statements.



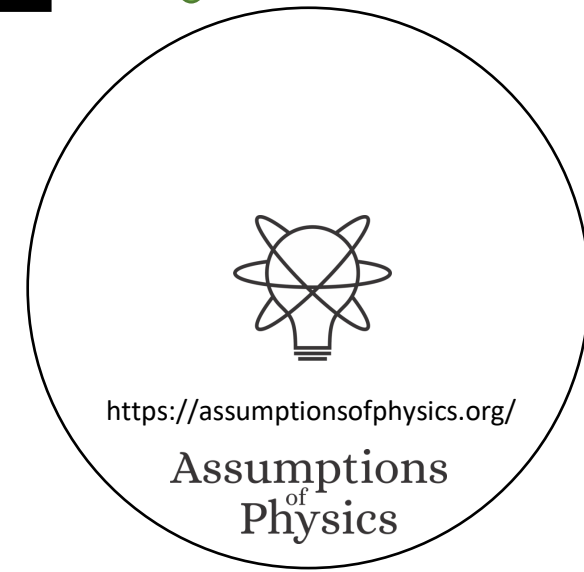
1. Strict references

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

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



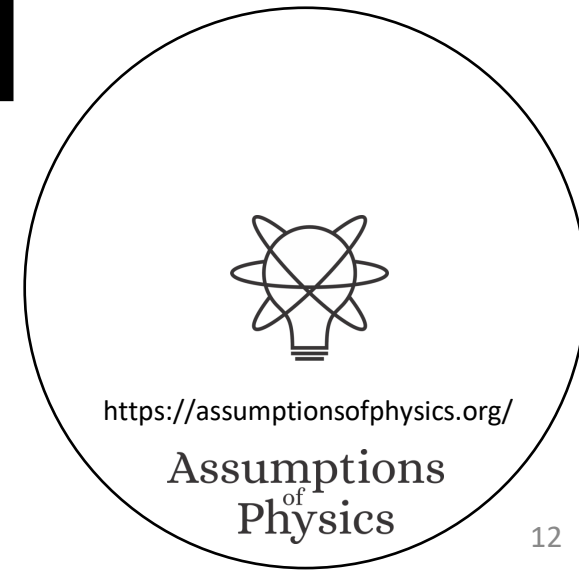
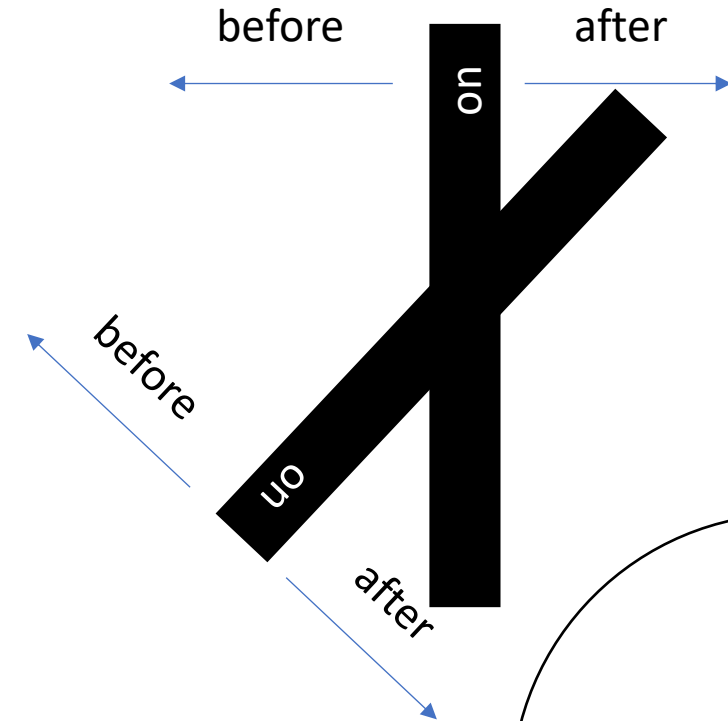
Physically, the extent of what we measure is assumed to be smaller than the extent of our reference



Multiple references

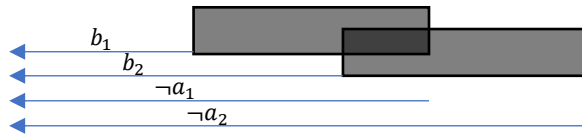
Without further constraints, references would not lead to a linear order

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



2. Aligned references

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

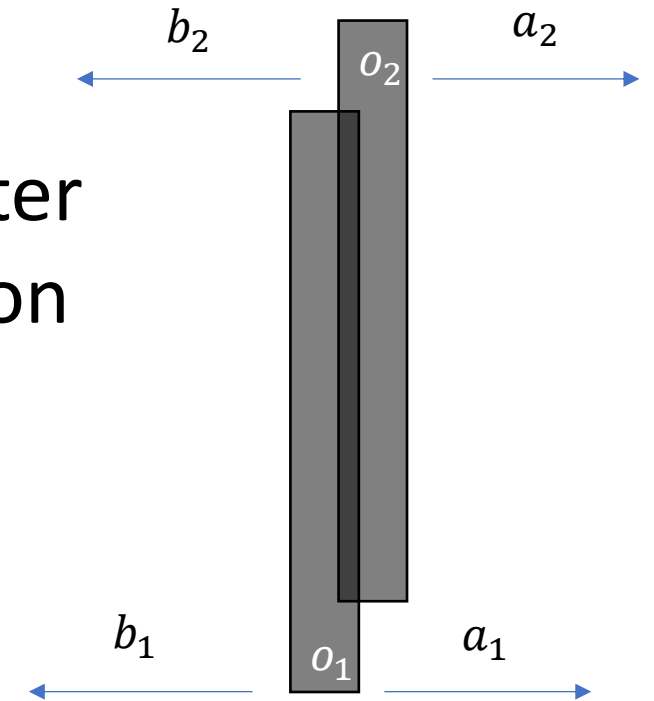


For example, $b_1 \preceq b_2 \preceq \neg a_1 \preceq \neg a_2$

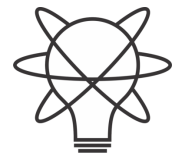
$\preceq$ Means that if the first statement is true

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

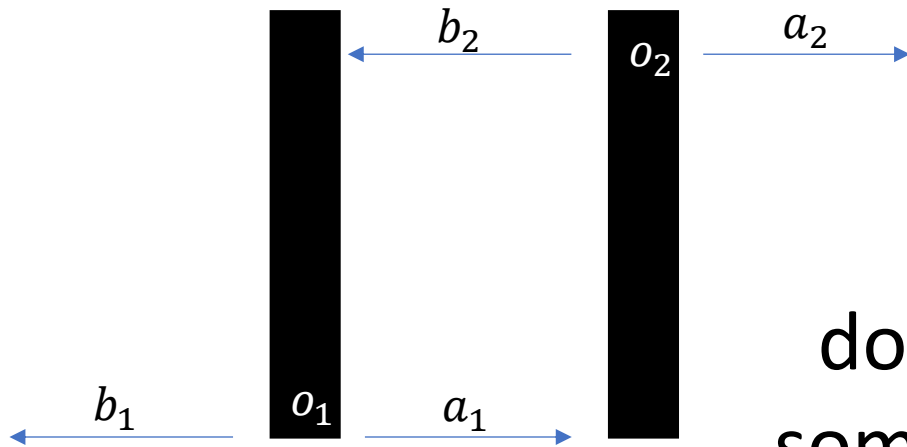


<https://assumptionsofphysics.org/>

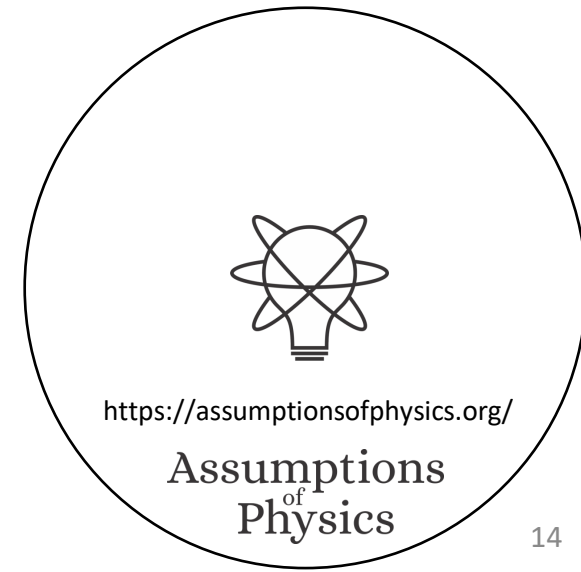
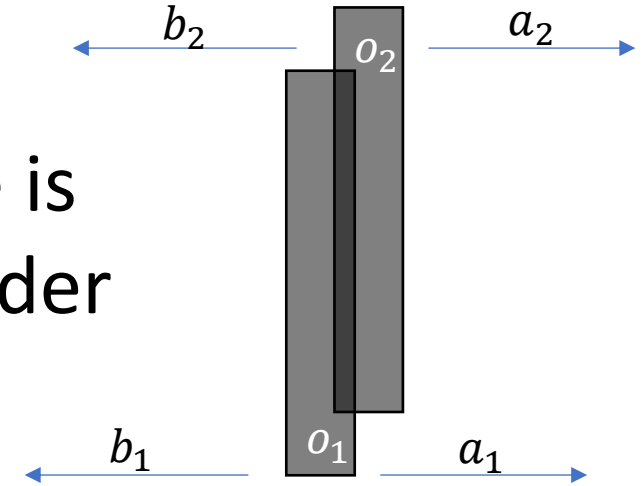
Assumptions
of
Physics

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



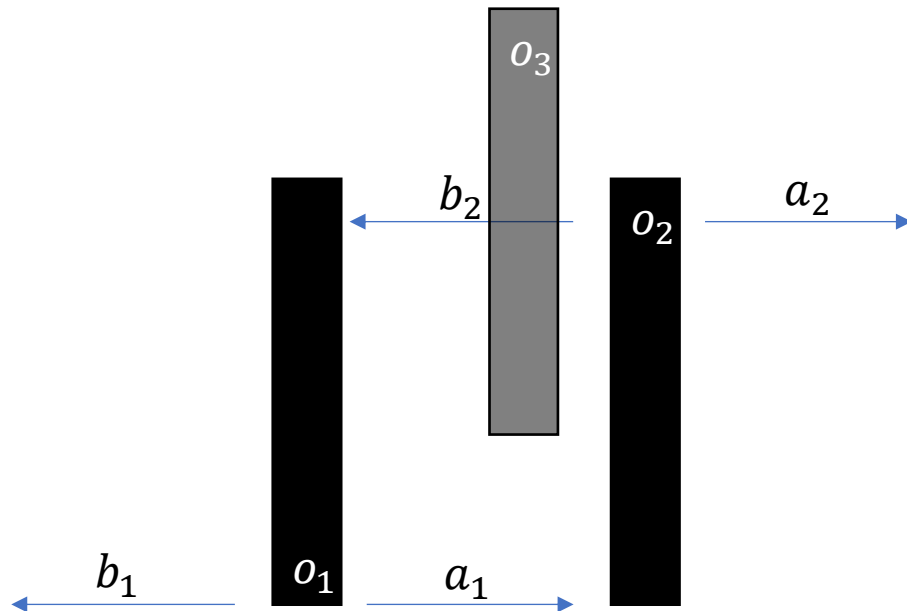
Moreover, if two references don't overlap and there can be something in between, we must be able to put a reference there



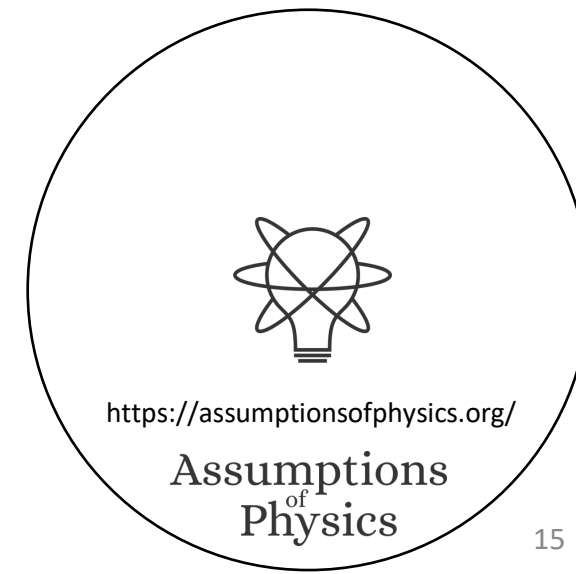
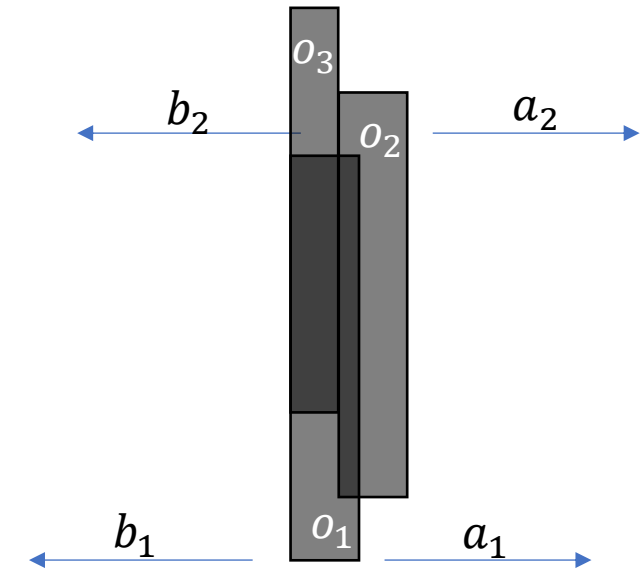
3. Refinable references

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



the possibility themselves can be ordered, and how this ordering, in the end, is uniquely characterized by statement narrowness: 10 is less than 42 because “the quantity is less than 10” is narrower than “the quantity is less than 42”.

As the defining characteristic for a quantity is the ability to compare its values, then the values must be ordered in some fashion from smaller to greater. Therefore, given two different values, one must be before the other. Mathematically, we call this order an order with such a characteristic as we can imagine the elements positioned along a line. Note that vectors are not linearly ordered: no direction is greater than the other. Therefore, in this context, a vector will not strictly be a quantity but a collection of quantities.

We also have to define how the order can be experimentally verified. The idea is that we should, at least, be able to verify that the value of a given quantity is before or after a set value. This allows us to construct bounds such as “the mass of the electron is 0.511 eV $\pm$ 4.5 keV” which we take to be equivalent to “the mass of the electron is more than 0.511 keV but less than 511.5 keV”. For integers, this also allows us to verify particular numbers “the earth has one natural satellite” is equivalent to “the earth has more than zero natural satellites and fewer than two”. Therefore we will define the order topology as the one generated by set of the type $(-\infty, a)$ and (a, ∞) .

A quantity, then, is an ordered property with the order topology.

Definition 3.4. A linear order on a set Q is a relationship $\leq$ $Q \times Q \rightarrow \mathbb{B}$ such that:

1. (antisymmetry) if $q_1 \leq q_2$ and $q_2 \leq q_1$ then $q_1 = q_2$.
2. (transitivity) if $q_1 \leq q_2$ and $q_2 \leq q_3$ then $q_1 \leq q_3$.
3. (total) at least $q_1 \leq q_2$ or $q_2 \leq q_1$.

A set together with a linear order is called a linearly ordered set.

Definition 3.5. Let $(Q, \leq)$ be a linearly ordered set. The order topology is the topology generated by the collection of sets of the form:

$$(a, \infty) = \{q \in Q \mid q < a\}, (-\infty, a) = \{q \in Q \mid q < b\}.$$

Definition 3.6. A quantity for an experimental domain D is a linearly ordered proper. Formally, it is a tuple $(Q, \leq, a)$ where $(Q, \leq)$ is a property, $Q \subseteq Q \times Q \rightarrow \mathbb{B}$ is a linear order and a is a topology space with the order topology with respect to $\leq$.

As for properties, the quantity values are too abstract to label the different cases. At most correspond to the letters of the alphabet and a few words ordered alphabetically. The units are not captured by the numbers themselves; they are captured by the function

“In other languages, there are two words to differentiate quantity in its “physical quantity” (e.g. grand (mass, gradient) and its “human” (e.g. quantita, massa, quantity). It is the second meaning that is required here.

The sentence: “The mass of the electron is 0.511 eV $\pm$ 4.5 keV” could instead be translated (using instead of an acronym) based on the words used. The sentence “the mass of the electron is 0.511 eV $\pm$ 4.5 keV” is not testing those types of statements where the mass of the electron is not the same as the defined bound statements that the mass of the electron is less than 0.511 keV but more than 511.5 keV. When examining the dictionary, we see the fact that we are experimentally told whether the word is looking for a before or after the one we are testing against.

that allows us to map statements to numbers and vice-versa.

As we want to understand quantities better, we concentrate on those experimental domains that are fully characterized by their order. For example, the domain for the mass of a system will be fully characterized by a real number greater than or equal to zero. Each possibility will be identified by a number which will correspond to the mass expressed in a particular unit, say in kg. As the values of the mass are ordered, we can also say that the possibilities themselves are ordered. That is, “the mass of the system is x kg” precedes “the mass of the system is y kg”. This ordering of the possibilities will be linked to the natural topology as “the mass of the system is less than x kg”, the depiction of all possibilities that come before a particular possibility, is a verifiable statement.

We call a natural order for the possibility a linear order on them such that the order topology is the natural topology. An experimental domain is fully characterized by a quantity if and only if it is naturally ordered and that quantity is ordered in the same way; it is order isomorphic. In other words, we can only sense a quantity to an experimental domain if it has the same order as a natural ordering of the same type.

mass of the system is more than x kg. K_R is also ordered by narrowness but with the reverse ordering of the possibilities: values. These are the very statements whose verifiable sets define the order topology and therefore jointly constitute a basis for the experimental domain.

Now consider the statement s_1 : “the mass of the system is less than or equal to x kg” with s_1 : “the mass of the system is less than x kg”. We have $s_1 \leq s_2$. In fact, if we replace the value in s_1 with anything less than x kg will still have $s_1 \leq s_2$. Instead if we use a value greater than x kg will have $s_1 \leq s_2$. In other words, if we call f the set that includes both the less-than-or-equal and less-than statements this is linearly ordered by narrowness. But “the mass of the system is less than or equal to x kg” is equivalent to “the mass of the system is greater than x kg”. In other words, $R : R(x) : R(x)$ contains all the statements like “the mass of the system is less than x kg” and “the mass of the system is more than x kg”. K_R and these are all linearly ordered by narrowness.

The ordering of R can be further characterized. Note that s_1 : “the mass of the system is less than or equal to x kg” is an immediate successor of s_2 : “the mass of the system is less than x kg”. That is, they are different and there can’t be any other statement in R that is between them s_1 but narrower than s_1 since they differ in the value of x . This will happen for any mass value. So R is composed of two exact copies of the ordering of X , where each element of one copy is immediately followed by an element of the other copy. Moreover, if a statement in R has an immediate predecessor in R , then it is also an immediate successor of the two. If we call g the value of that case, then the statement must be of the form “the mass of the system is less than g ” which its immediate successor is of the form “the mass of the system is less than or equal to g kg”. The statement is broader by just the possibility associated with g . Therefore statements in R that have an immediate successor must be in R_1 as well.

The main result is that the above characterization of the basis of the domain is necessary and sufficient to order the possibilities. If an experimental domain has a basis composed of

$$x^* < x_1^* \wedge x^* < x_2^* \Rightarrow x^* < x_1^* \vee x^* < x_2^*.$$

Definition 3.12. Let D_X be a naturally ordered experimental domain and X its poset. Define $R_X : (x^* < x_1^* \vee x^* < x_2^*) \rightarrow \mathbb{B}$ and $R : R(x) : R(x)$.

Definition 3.13. Let $(Q, \leq)$ be an ordered set. Let $q_1, q_2 \in Q$. Then q_1 is an immediate successor of q_2 and q_2 is an immediate predecessor of q_1 if there is no element in between them in the ordering. That is, $q_1, q_2 \in Q$ and there is no $q_3 \in Q$ such that $q_2 < q_3 < q_1$. Two elements are consecutive if one is the immediate successor of the other.

Proposition 3.14. Let D_X be a naturally ordered experimental domain. Then $(R_1, (R_2, \dots, R_n))$ are linearly ordered sets. Moreover $(R_1, \dots, R_n)$ are order isomorphic to $(X, \leq)$.

Proof. Let $f : X \rightarrow R_1$ be defined such that $f(x_1) = x^* < x_1^*$. As there is one only statement $x^* < x_1^*$ for each $x_1 \in X$, f is a bijection. Suppose $x_1 \leq x_2$, have $f(x_1) = (x_1^* \vee \dots \vee x_1^*)$ and $f(x_2) = (x_2^* \vee \dots \vee x_2^*)$ and third $f(x_1) \leq f(x_2)$. On the other hand if $f(x_1) \leq f(x_2)$ then we set $(-x_1, x_1) \subseteq (-x_2, x_2)$ with means $x_1 \leq x_2$. This means that g is an order isomorphism between $(R_1, \leq)$ and $(X, \leq)$. To show that R is linearly ordered, let $x_1, x_2 \in X$. If they both come from either R_1 or R_2 , then s_1 and s_2 are already ordered by narrowness. If not, consider the two statements $f^* < x_1^* \vee x^* < x_2^*$ and $x^* < x_1^* \vee x^* < x_2^*$. Since X is linearly ordered, either $(x \vee x_1) \leq x_2$ or $(x \vee x_2) \leq x_1$. Which means $f^* < x_1^* \vee x^* < x_2^*$ is linearly ordered by narrowness. Let D_X and R_X be two sets of verifiable statements such that D_X is linearly ordered by narrowness. Let D_X and R_X be the experimental domain fully generated and $D : D(x) : D(x)$. Then $(D_1, \dots, D_n)$ and $(D, \leq)$ are linearly

isot with respect to that $(D_1, \dots, D_n)$ is linearly ordered. We have that R_X is linearly ordered because it is the union of B which is linearly ordered by narrowness. Note junction of a finite set of statements linearly ordered by narrowness will return the narrowness and the distinction of a finite set of statements linearly ordered will return the broadest element. The countable distinction, instead, can be elements. But using these elements again will not introduce new one: the countable distinction will still be a countable distinction; the finite of countable distinction is the countable distinction of finite conjunctions.

3.15. Let R_1 and R_2 be two sets of verifiable statements such that D_X is linearly ordered by narrowness. Let D_X and R_X be the experimental domain fully generated and $D : D(x) : D(x)$. Then $(D_1, \dots, D_n)$ and $(D, \leq)$ are linearly

Proof. By definition, we have $-b \wedge -a < a$ and by 1.23 $-(b \wedge -a) \vee 0 \leq \neg b \vee a \vee 0$. $\square$

Definition 3.19. A reference $r_1 = (b_1, a_1, a_2)$ is finer than another reference $r_2 = (b_2, a_2, a_3)$ if $b_1 \leq b_2$, $a_1 < a_2$ and $a_1 \geq a_3$.

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. $\square$

Definition 3.21. A reference is strict if its before, and/or after statements are incompatible. Formally, $r = (b, a_1, a_2)$ such that $b \wedge a_1$ and $a_2 \wedge -b$ is not strict.

Remark. In general, we can turn a loose reference into a strict one. The on statement can be made strict by replacing it with $-b \wedge -a$. This is possible because it is not required to be verifiable. The before (and after) statements would need to be replaced with statements like $b \wedge -a$, 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 references weights, a clock will keep ticking. What does it mean that a reference comes after another in terms of the before/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 second. Mathematically, this type of order before and strictly after. It does so. One may be tempted to define the references refining the references and, if required, references, not the original $\square$

Definition 3.22. A reference is the first cannot be on or after the first.

Proposition 3.23. Reference order is:

- irreflexivity: not $r < r$
- transitivity: if $r_1 < r_2$ and $r_2 < r_3$

and is therefore a strict partial order.

Proof. For irreflexivity, since r and therefore $b \vee 0 \vee a$. Therefore irreflexive. $\square$

Definition 3.24. 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 $a_1 \leq b_2$. Two references are consecutive if one is the immediate successor of the other.

Proposition 3.30. Let $r_1 = (b_1, a_1, a_2)$ and $r_2 = (b_2, a_2, a_3)$ be two references. If r_1 is immediately before r_2 then $b_1 \leq b_2$.

Proof. Let r_1 be immediately before r_2 . Then $a_1 \leq b_2$ which means $b_2 \leq a_1$. By 3.27 we also have $-a_1 \leq b_2$. Therefore $b_2 \leq a_1$. $\square$

Proposition 3.31. Let $r_1 = (b_1, a_1, a_2)$ and $r_2 = (b_2, a_2, a_3)$ be two strict references. Then $r_1 < r_2$ if and only if $a_1 < a_2$.

Proof. Let r_1 be immediately before r_2 . Then $b_2 \leq a_1$ by 3.30. Conversely, let $b_2 \leq a_1$. Then $r_1 < r_2$ by 3.28. We also have $a_1 \leq a_2$, therefore $a_1 \leq b_2$ and r_1 is immediately before r_2 by definition. $\square$

means we can find $r_3 = (b, -b \wedge -a_2, a_2)$ for some $b \in D_X$ such that $r_1 < r_3$ and therefore $-a_1 \leq b \wedge -a_2$.

For the third, suppose $a_1 \leq b_2$ and $b_2 \leq b_3$ for some b_3 then $-a_1 < b_3$. Then $r_1 = (-a_1, a_1)$ and $r_2 = (-b_3, b_3)$ 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 $r_3 = (b, -b \wedge -a_2, a_2)$ such that $r_1 < r_3 < r_2$ and therefore $a_1 < b < a_2 \wedge -b < a_2$. $\square$

Proposition 3.37. Let D be an experimental domain generated by a set of verifiable aligned strict references. Then all elements of D are the part of a pair $(s, -s)$ such that $s \in D_X$, s is a reference, $-s$ is the immediate predecessor of s in D or s is a reference. Moreover if s has an immediate successor, then $s \in D_X$.

Proof. Let D be an experimental domain generated by a set of verifiable aligned strict references. Let $s \in D_X$. Let $s = (a \in D_X) \vee a \vee \neg 1$. Let $s = a$. First we show that $s \leq -a$. We have $s \leq a \leq s \wedge -a \vee s \wedge -a \leq a \wedge -a \leq -a \vee s \wedge -a \leq -a$. For all $a \in A$ we have $a \vee s \leq \neg a$, $-a \vee s$ which means $s \leq -a$ because of the total order of D . This means that $s \wedge -a \leq s$ for all $a \in A$, therefore $s \wedge -a \leq s$ and $s \leq -a$.

Next we show that no statement $s \in D$ is such that $s < s < -s$. Let $a \in D_X$ such that $s < a$. By construction $a \leq A$ and therefore $-a \leq -s$. Therefore we can’t have $s < a < -s$. We also can’t have $b \in D_X$ such that $s < b < -a$. By 3.36 we’d find $a \in D_X$ such that $s < a < -b$, which was ruled out. So there are two cases. Either $s \leq -a$, then $s \leq -a$; $-a \leq s$, the immediate successor of s is $-a$ or $s \leq -a$.

The same reasoning can be applied starting from $s \in D_X$ to find $a \in D_X$ such that s is the immediate predecessor of $-a$, 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 $s_1 \in D_X$ and $s_2 \in D$ such that s_2 is the immediate successor of s_1 . By 3.36, in all cases where $s_1 \in D_X$ and $s_2 \in D_X$ we can always find another statement between the two. Then we must have that $s_1 \in D_X$ and $s_2 \in D_X$. $\square$

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

Proof. Suppose D_X is a naturally ordered domain generated by a set of verifiable 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 D_X is naturally ordered. Define the set R_1, R_2 and D as in 3.12. Let $R = (b, -b \wedge -a) \vee (b \wedge a \in R_1, b < -a) \vee$ 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 $b \wedge -a$ is not a contradiction since $b \wedge -a$ means $b \wedge a$ and $b \wedge -a$. On the statement is broader than $-b \wedge -a$ as they are equivalent and it is broader than $b \wedge a$ as that is a contradiction since $b \wedge -a$. Therefore R is a set of references. Since the before and after statements of R coincide with the basis of the domain, D_X is generated by R .

For transitivity, if $r_1 < r_2$, we have $b_1 \vee 0 \vee a_2$ and therefore $-(b_1 \vee 0) \vee a_2 \vee b_2$ by 1.23. Since $b_1 \vee 0 \vee a_2 \leq 1$, we have $a_1 \geq -(b_1 \vee 0)$. Similarly if $r_2 < r_3$ we have $a_2 \geq -(b_2 \vee 0) \vee a_3$. Putting it all together $-(b_1 \vee 0) \vee a_2 \vee a_3 \geq -(b_2 \vee 0) \vee a_3$, which means $b_1 \vee 0 \vee a_3$.

Corollary 3.24. The relationship $r_1 \leq r_2$ 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. The 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 vertical vertical 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 narrow lines around the borders, and these will be ordered reference by reference: 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.

Definition 3.33. Let D be a domain generated by a set of references R . A reference $r = (b, a_1, a_2)$ is said to be aligned with D if D is D_X and $a \in D_X$.

Proposition 3.34. Let D be an experimental domain generated by a set of aligned strict references R and $D = D_X \cup \{D_X\}$. Then $(D, \leq)$ is linearly ordered.

Proof. By 3.26 we have that $D = B \cup D_X \cup \{D_X\}$ is aligned by narrowness. By 3.15 the ordering extends to D . $\square$

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 are consecutive, 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 show they are aligned, take two references. The before and not after statements of one reference, 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 third condition of the domain ordering theorem 3.16.

Definition 3.35. Let D be an experimental domain generated by a set of aligned references R . The set of references is refinable if, given two strict references $r_1 = (b_1, a_1, a_2)$ and $r_2 = (b_2, a_2, a_3)$ aligned with 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 D such that $r_1 < r_3 < r_2$.
- refine references if one is finer than the other; that is, if $q_1 \leq q_2$ and q_2 is not the immediate successor of q_1 , then we can find a strict reference r_3 aligned with D such that $q_1 < r_3$ and either $b_1 \leq b_2$ and $r_3 < r_2$ or $a_2 \leq a_3$ and $r_3 < r_2$.

Proposition 3.36. Let D be an experimental domain generated by a set of refinable aligned strict references R .

Now show that R consists of aligned strict references. We already saw that $b \wedge a$ and we also have $-b \wedge -a$ is incompatible with both b and a . The reference is strict, so 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 if R is refinable, note that each reference can be expressed as $(x^* < x_1^*, x_1^* \leq x_2^* \vee x_2^* < x_1^*)$ where $x_1, x_2 \in A$ and $x_1 \leq x_2 \vee x_2 \leq x_1$. Let $s_1 = (x_1^* \vee x_2^* < x_1^*)$. That is, every reference is identified by two possibilities x_1, x_2 such that $x_1 \leq x_2$. Therefore take two references $r_1 = (x_1, x_2)$ and $r_2 = (x_1, x_2)$ and (x_1, x_2) be the respective pair of possibilities we can use to express the references we have shown. Suppose $r_1 < r_2$ but they are not consecutive. Then $s_1 \leq x_2^* < x_2^* < x_1^*$. That is, we can find $x_3 \in A$ such that $x_2 \leq x_3$ which means $x_2 \leq x_3 \leq x_1^* \vee x_3 \leq x_2^* < x_1^*$. Therefore the reference $r_3 = (x_3, x_1)$ identified by (r_1, x_2) is between the two references. On the other hand, assume the second reference is finer than the first. Then $x_2 \leq x_3$ and $x_1 \leq x_2$ with either $x_1 \leq x_2$ or $x_1 \geq x_2$. Since $r_1 = (x_1, x_2)$ and $r_2 = (x_1, x_2)$ identified by (r_1, x_2) and (r_2, x_2) . If $x_1 < x_2$ and $r_1 < r_2$. Also note that the before statements of r_1 and r_2 are the same and the after statements of r_1 and r_2 are the same. Therefore we satisfy all the requirements and the set R is aligned by definition. $\square$

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 that are aligned with the domain. The before and not after statements of one reference 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. The before and not after statements must be aligned with the domain, 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_1 , we have the possibility “the value of the property is q_1 ” and “we have the reference r ” and “the value of the property is less than q_1 ”, “the value of the property is q_1 ”, “the value of the property is more than q_1 ”.

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 the integers. We will see that the integers are the only set of integers that 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

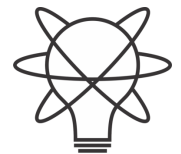
Additionally:

Dense order, which the closure on arbitrary union completes

Between any two references we can always have another reference $\Rightarrow$ **real numbers**

Only finitely many references between any two references $\Rightarrow$ **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)

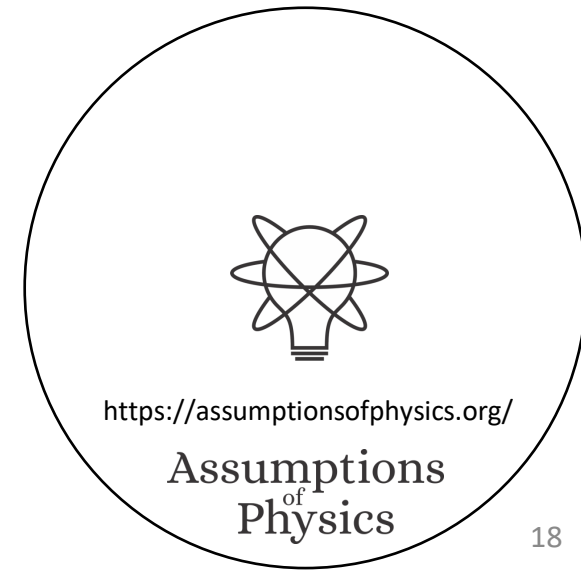
If clocks have jitter, they cannot achieve perfect synchronization (references not aligned)

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.

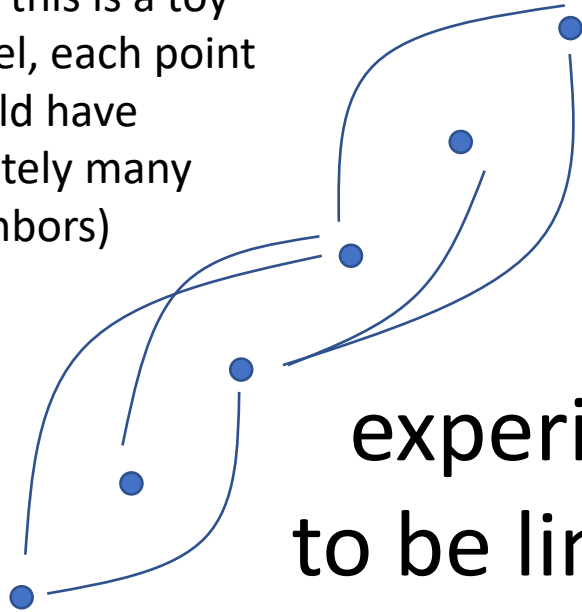
A better understanding of space-time means
creating a more realistic formal model that
accounts for those failures



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)

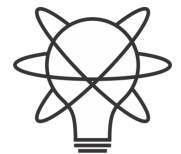


Lack of order at small scales,
order at large enough scale

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?

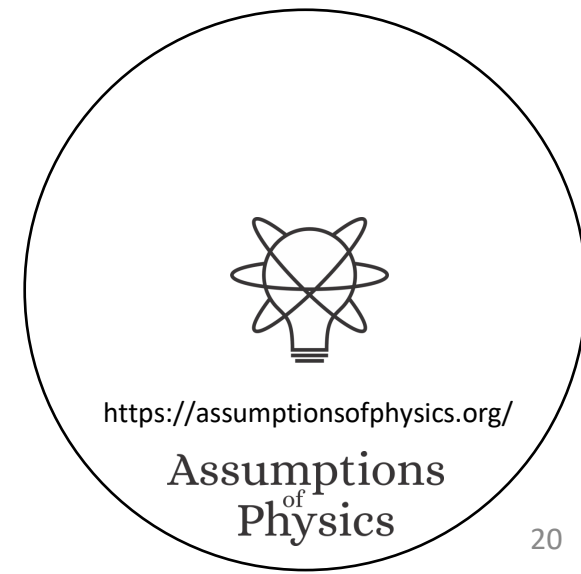


<https://assumptionsofphysics.org/>

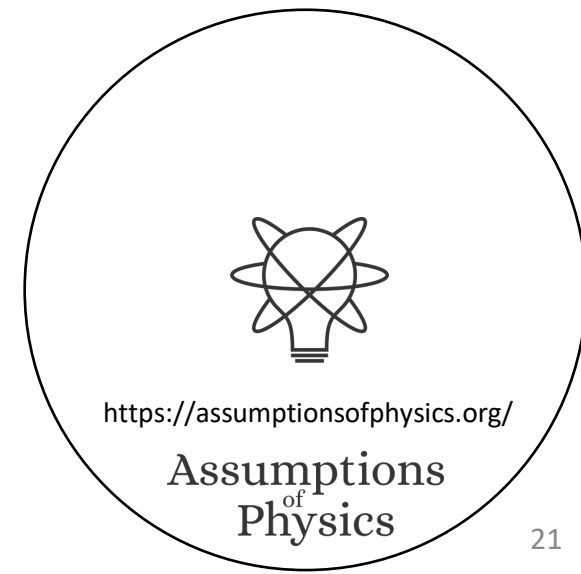
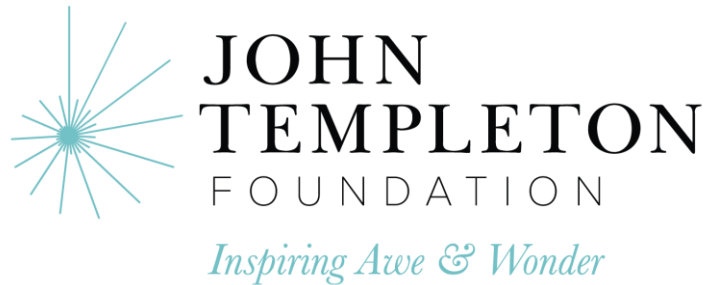
Assumptions
of
Physics

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” $\preceq$ “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



Ensemble spaces (generalized state spaces)



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

⇒ Scientific laws are relationships between ensembles

Preparation
procedure

Initial ensemble

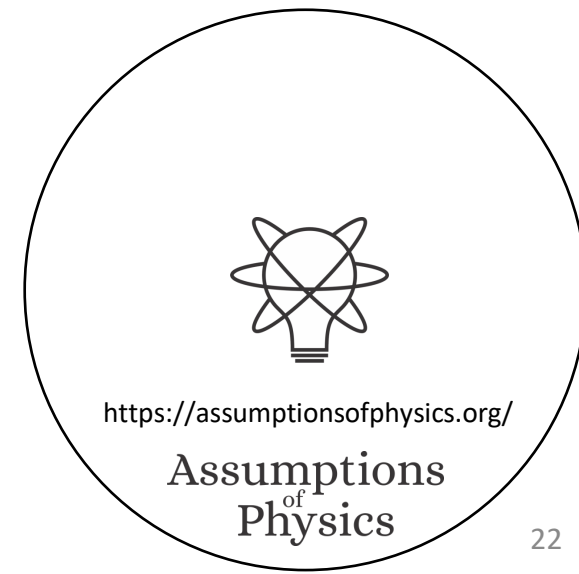
Physical
process

Final ensemble

Measurement
procedure

Statistical data

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



State space

Classical discrete $X = \{x_1, x_2, \dots\}$

Phase space
Symplectic manifold

Classical continuum $X = \{\mathbb{R}^{2n}, \omega\}$

Projective complex
Hilbert space

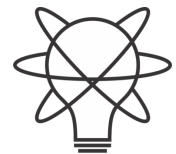
Quantum mechanics $X = P(\mathcal{H})$

Ensembles

$$\mathcal{E} = \{p_i \mid \sum_i p_i = 1\}$$

$$\mathcal{E} = \left\{ \rho \in L^1(\mathbb{R}^{2n}) \mid \int \rho dq^n dp_n = 1 \right\}$$

$$\mathcal{E} = \{ \text{positive semi-definite Hermitian with } \text{tr}(\rho) = 1 \}$$



<https://assumptionsofphysics.org/>

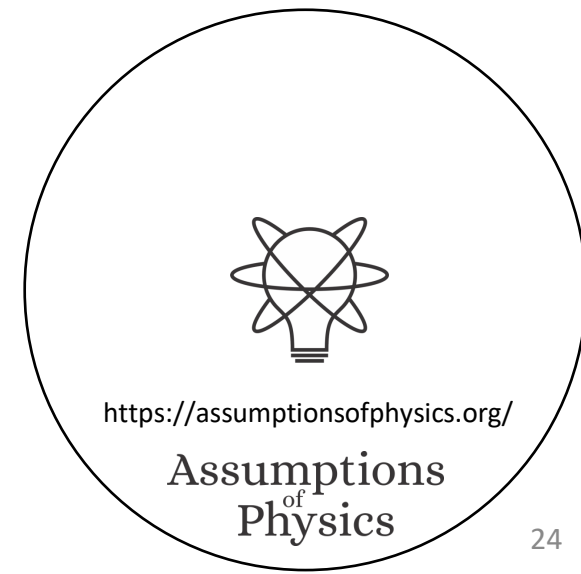
Assumptions
of
Physics

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 $\Rightarrow$ topological space

Topology is responsible for handling limits and infinite operations

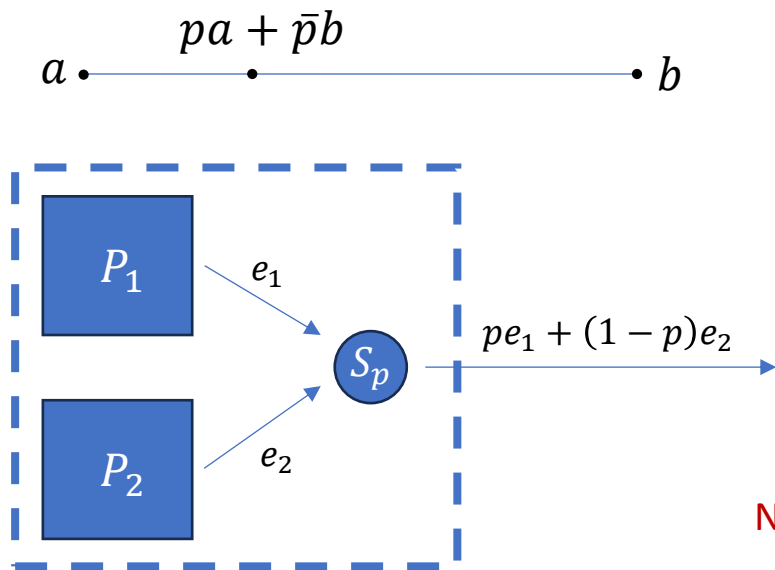
All other axioms are on finite elements



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 $+: [0, 1] \times \mathcal{E} \times \mathcal{E} \rightarrow \mathcal{E}$ called **mixing**, noted with the infix notation $pa + \bar{p}b$, with the following properties:*

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

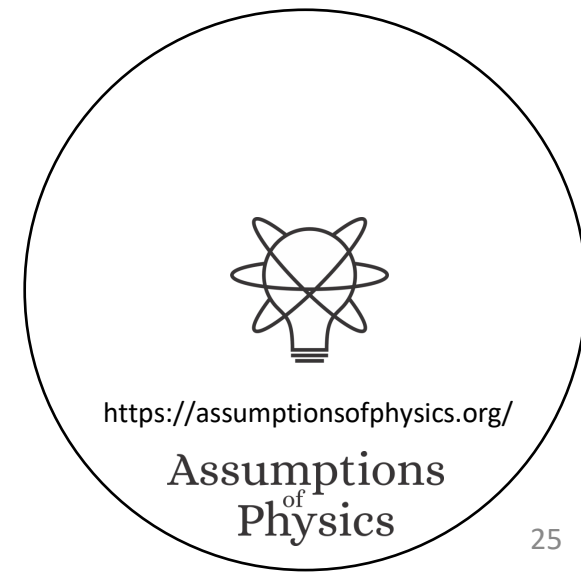
Statistical mixtures $\Rightarrow$ Convex structure



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

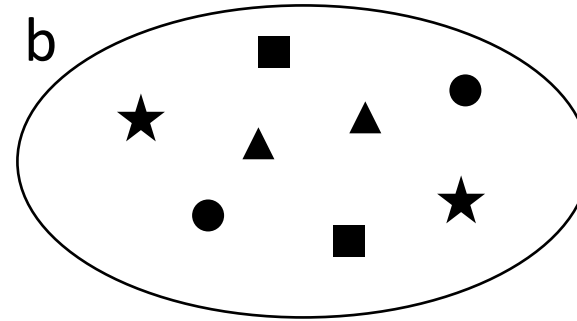
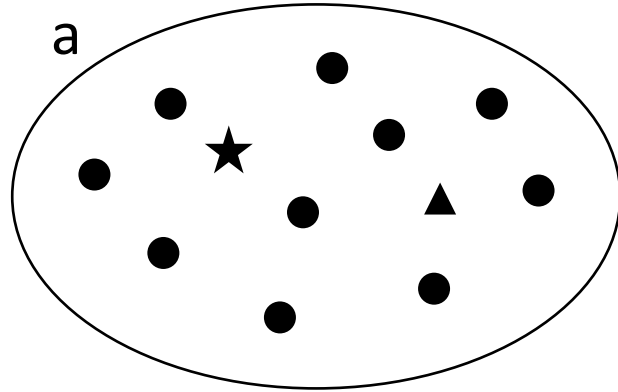
Topology tells us which
infinite mixtures $\sum_{i=1}^{\infty} p_i e_i$ converge
I.e. where experimental verifiability converges

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



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 $S : \mathcal{E} \rightarrow \mathbb{R}$, defined up to a positive multiplicative constant representing the unit numerical value. The entropy has the following properties:*

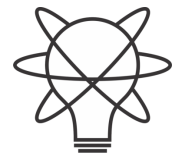
- *Continuity^a*



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

Variability: how similar are the preparations within an ensemble?

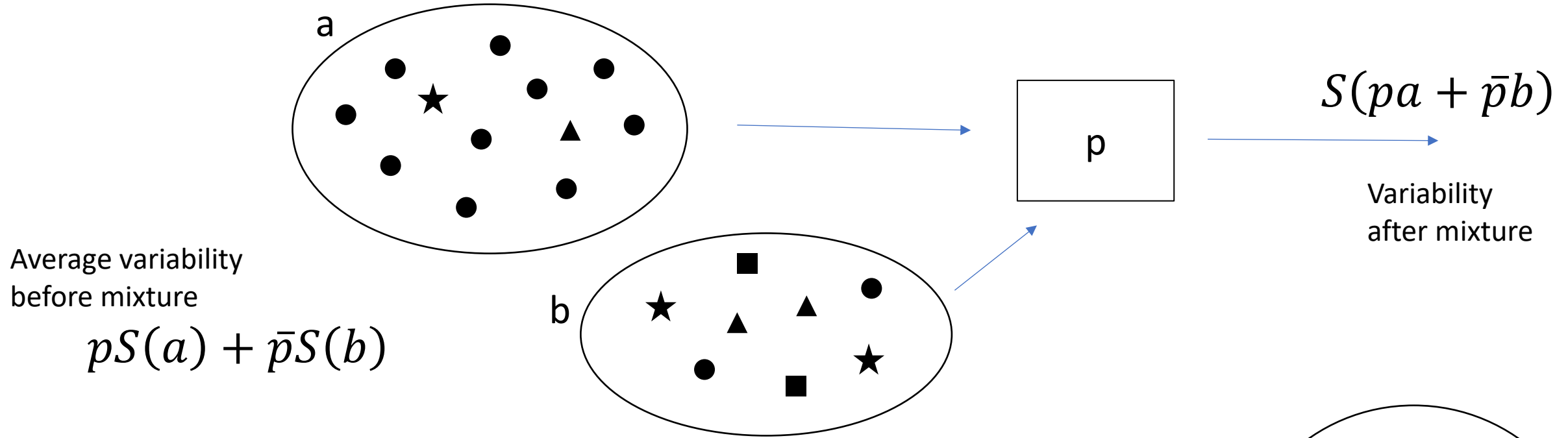
Entropy is a measure of variability



<https://assumptionsofphysics.org/>

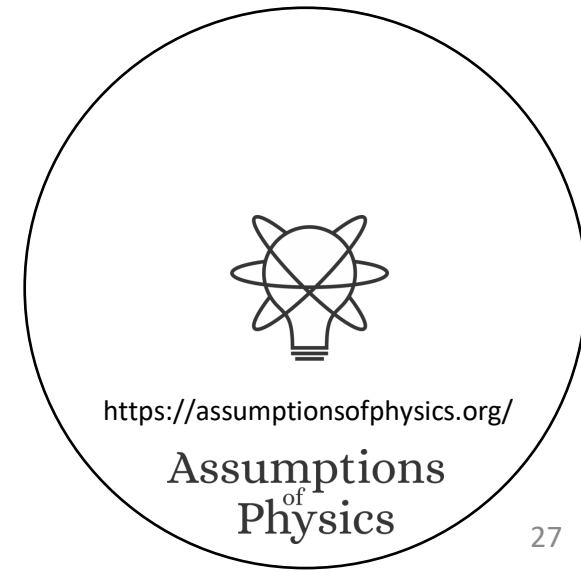
Assumptions
of
Physics

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

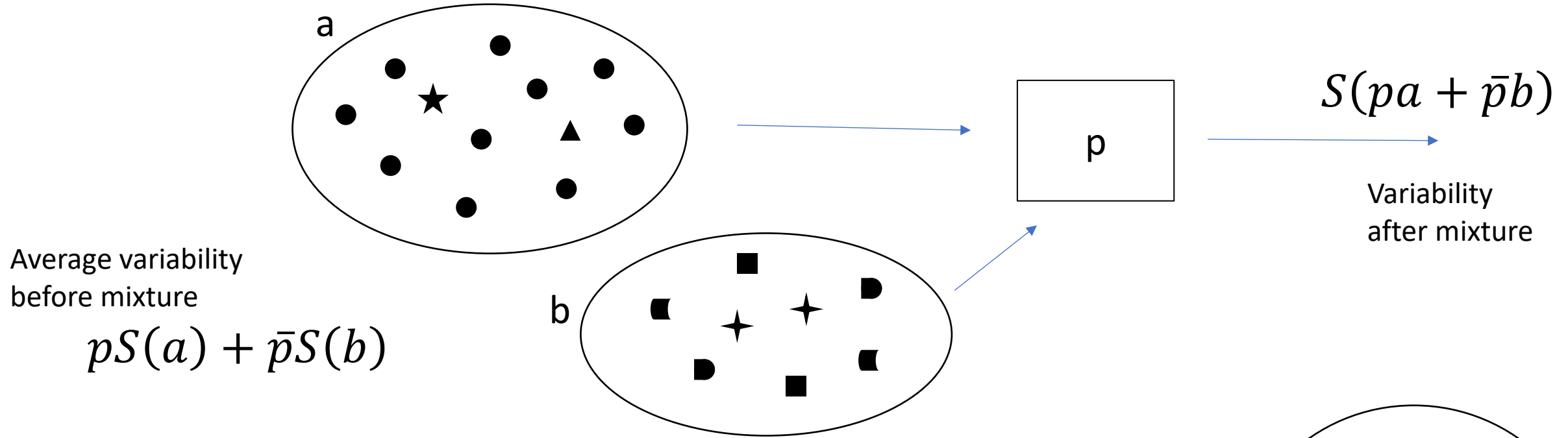


Final variability will be greater than average variability before mixture

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

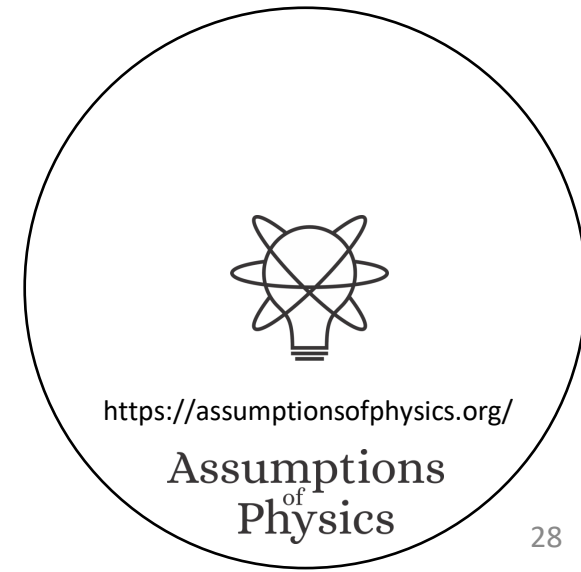


- **Upper variability bound:** there exists a universal function $I(p_1, p_2)$ (i.e. the same for all ensemble spaces) such that $S(pa + \bar{p}b) \leq I(p, \bar{p}) + pS(a) + \bar{p}S(b)$; if the equality holds, a and b are **non-overlapping or orthogonal**, noted $a \perp b$

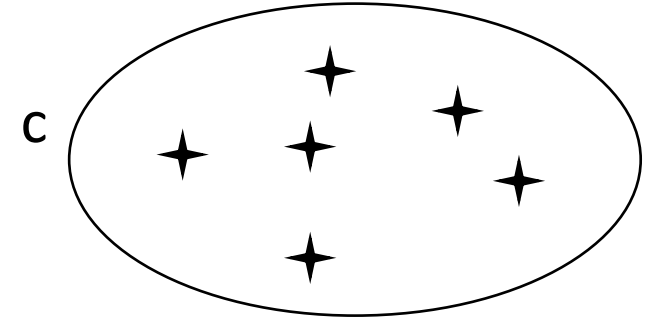
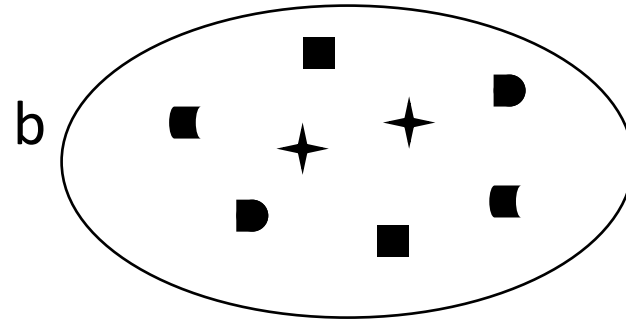
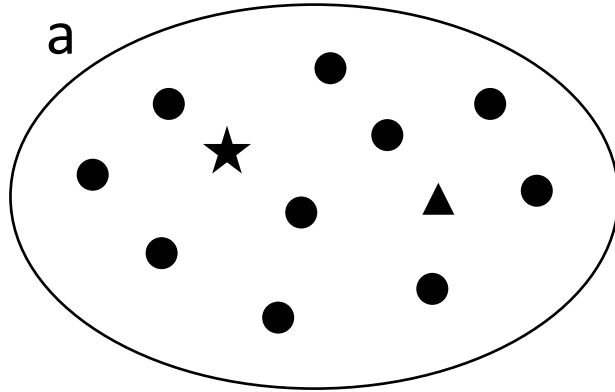


Maximum increase when the ensembles are
“completely different”

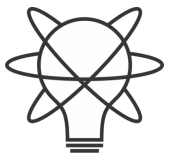
Increase is only a function of the mixing coefficient



- *Mixtures preserve orthogonality:*^b $a \perp b$ and $a \perp c$ if and only if $a \perp pb + \bar{p}c$ 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/>

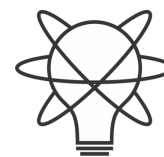
Assumptions
of
Physics

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 $S : \mathcal{E} \rightarrow \mathbb{R}$, defined up to a positive multiplicative constant representing the unit numerical value. The entropy has the following properties:*

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

Ensemble variability $\Rightarrow$ 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.*

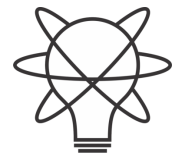
Axiom 1.21 (Axiom of entropy). *Every element of the ensemble is associated with an **entropy** which quantifies the variability of the preparations of the ensemble. Formally, an*

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

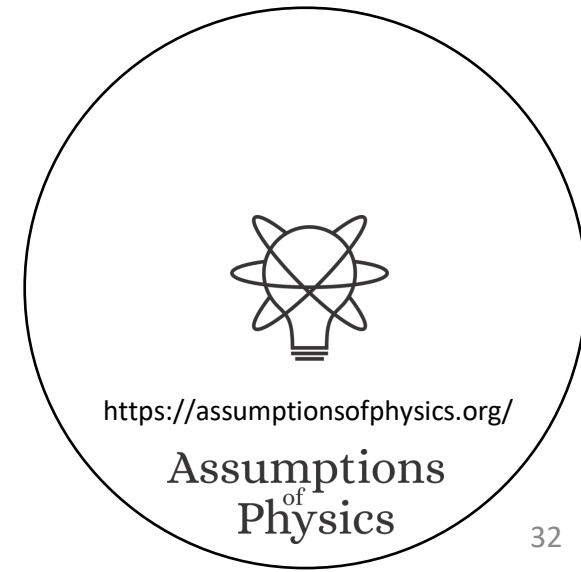
The entropy upper bound $I(p, \bar{p})$ is uniquely determined

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

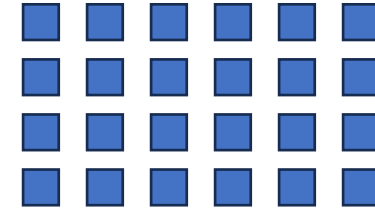


Pick a_i all orthogonal to each other

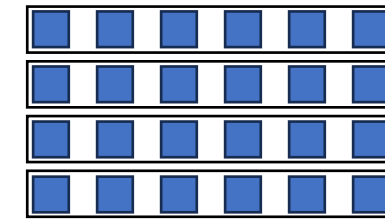
$$S(\sum_i p_i a_i) = I(p_i) + \sum_i p_i S(a_i)$$

$$\begin{aligned} S\left(\sum_{j=1}^n \sum_{k=1}^m \frac{1}{n} \frac{1}{m} a_{jk}\right) &= I\left(\left\{\frac{1}{nm}\right\}_{i=1}^{nm}\right) + \sum_{j=1}^n \sum_{k=1}^m \frac{1}{n} \frac{1}{m} S(a_{jk}) \\ &= S\left(\sum_{j=1}^n \frac{1}{n} \sum_{k=1}^m \frac{1}{m} a_{jk}\right) = I\left(\left\{\frac{1}{n}\right\}_{i=1}^n\right) + \sum_{j=1}^n \frac{1}{n} S\left(\sum_{k=1}^m \frac{1}{m} a_{jk}\right) \\ &= I\left(\left\{\frac{1}{n}\right\}_{i=1}^n\right) + \sum_{j=1}^n \frac{1}{n} \left(I\left(\left\{\frac{1}{m}\right\}_{i=1}^m\right) + \sum_{k=1}^m \frac{1}{m} S(a_{jk}) \right) \\ &= I\left(\left\{\frac{1}{n}\right\}_{i=1}^n\right) + I\left(\left\{\frac{1}{m}\right\}_{i=1}^m\right) + \sum_{j=1}^n \sum_{k=1}^m \frac{1}{n} \frac{1}{m} S(a_{jk}). \end{aligned}$$

Uniform mixture of $n \times m$ elements



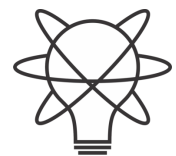
Must be the same



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

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

$$S\left(\sum_{i=1}^n \sum_{j=1}^{m_i} \frac{1}{m} a_{ij}\right) = I\left(\left\{\frac{1}{m}\right\}_{i=1}^m\right) + \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{1}{m} S(a_{ij})$$

$$= \kappa \log m + \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{1}{m} S(a_{ij})$$

$$= S\left(\sum_{i=1}^n \frac{m_i}{m} \sum_{j=1}^{m_i} \frac{1}{m_i} a_{ij}\right) = I\left(\left\{\frac{m_i}{m}\right\}_{i=1}^n\right) + \sum_{i=1}^n \frac{m_i}{m} S\left(\sum_{j=1}^{m_i} \frac{1}{m_i} a_{ij}\right)$$

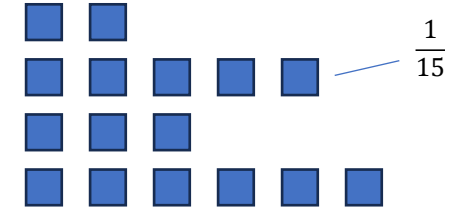
$$= I(\{p_i\}_{i=1}^n) + \sum_{i=1}^n \frac{m_i}{m} \left(I\left(\left\{\frac{1}{m_i}\right\}_{i=1}^{m_i}\right) + \sum_{j=1}^{m_i} \frac{1}{m_i} S(a_{ij}) \right)$$

$$= I(\{p_i\}_{i=1}^n) + \sum_{i=1}^n \frac{m_i}{m} I\left(\left\{\frac{1}{m_i}\right\}_{i=1}^{m_i}\right) + \sum_{i=1}^n \frac{m_i}{m} \sum_{j=1}^{m_i} \frac{1}{m_i} S(a_{ij})$$

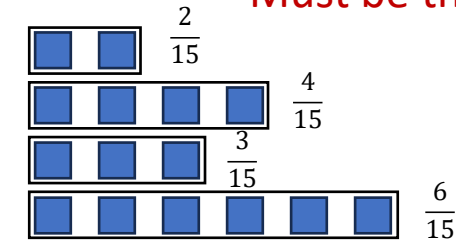
$$= I(\{p_i\}_{i=1}^n) + \sum_{i=1}^n p_i \kappa \log m_i + \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{1}{m} S(a_{ij}) .$$

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

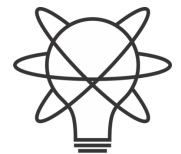
Uniform mixture of m elements



Must be the same



Rational distribution over uniform mixtures of grouped elements



<https://assumptionsofphysics.org/>

Assumptions
of
Physics

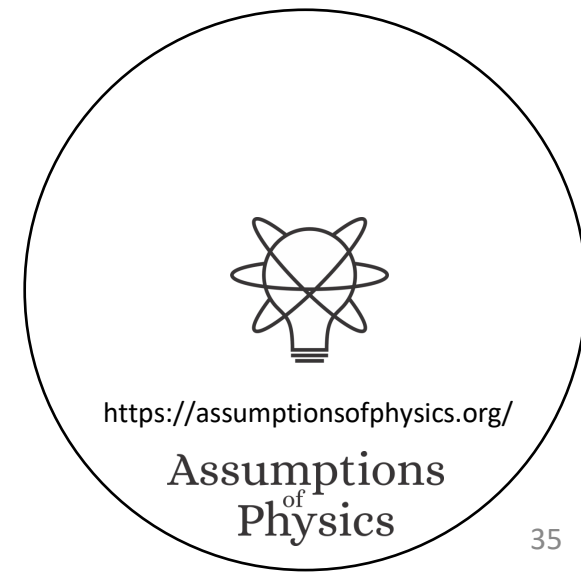
$$\begin{aligned}\kappa \log m &= I(\{p_i\}_{i=1}^n) + \sum_{i=1}^n p_i \kappa \log m_i \\ I(\{p_i\}_{i=1}^n) &= \kappa \log m - \sum_{i=1}^n p_i \kappa \log m_i = \sum_{i=1}^n p_i \kappa \log m - \sum_{i=1}^n p_i \kappa \log m_i \\ &= - \sum_{i=1}^n p_i \kappa \log \frac{m_i}{m} = -\kappa \sum_{i=1}^n p_i \log p_i.\end{aligned}$$

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

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

$$a \perp b$$

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

such that

$$a = p_1 c + \bar{p}_1 e_1$$

$$b = p_2 c + \bar{p}_2 e_2$$



Orthogonal ensembles

$$a \perp b$$

Saturate upper entropy bound

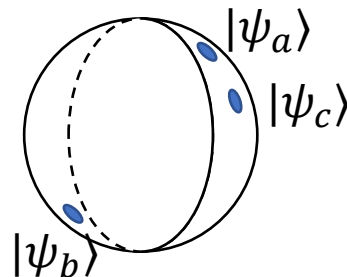
$$S(pa + \bar{p}b) = I(p, \bar{p}) + pS(a) + \bar{p}S(b)$$

Coincide in classical
ensemble spaces

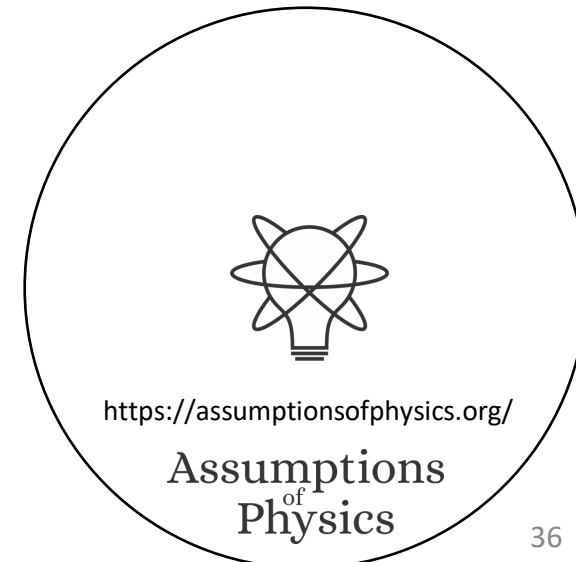


disjoint support

Different in quantum
ensemble spaces



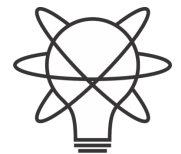
Simple but powerful
definitions from
basic axioms



<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 $pa + \bar{p}e = pb + \bar{p}e$ for some $p \in (0, 1)$ implies $a = b$.

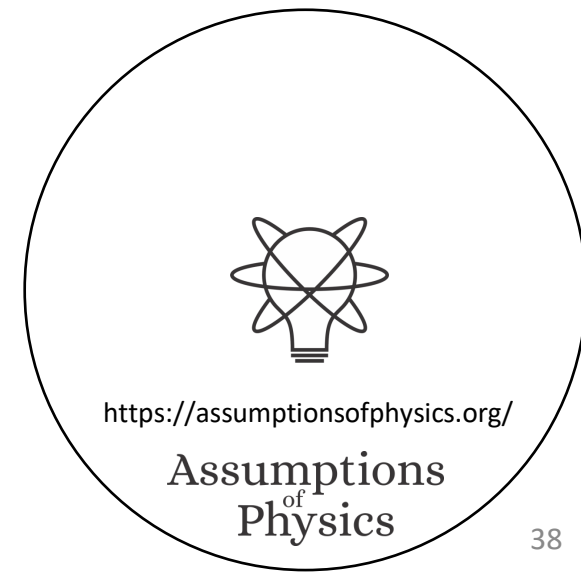
Theorem 1.43 (Ensemble spaces are cancellative). Let $\mathcal{E}$ be an ensemble space. Let $a, b, e \in \mathcal{E}$ such that $pa + \bar{p}e = pb + \bar{p}e$ for some $p \in (0, 1)$. Then $a = b$.

Entropy bounds force mixing to be “invertible”

$$a \bullet \text{---} \bullet b \text{---} \text{---} \bullet ra + \bar{r}b$$

Definition 1.53 (Affine combinations). Let $\{e_i\}_{i=1}^n \subseteq \mathcal{E}$ 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** $\sum_{i=1}^n r_i e_i$ is, if it exists, the ensemble $a \in \mathcal{E}$ such that $\sum_{i \in I} \frac{r_i}{r} e_i = \frac{1}{r} a + \sum_{i \notin I} \frac{-r_i}{r} e_i$ 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)

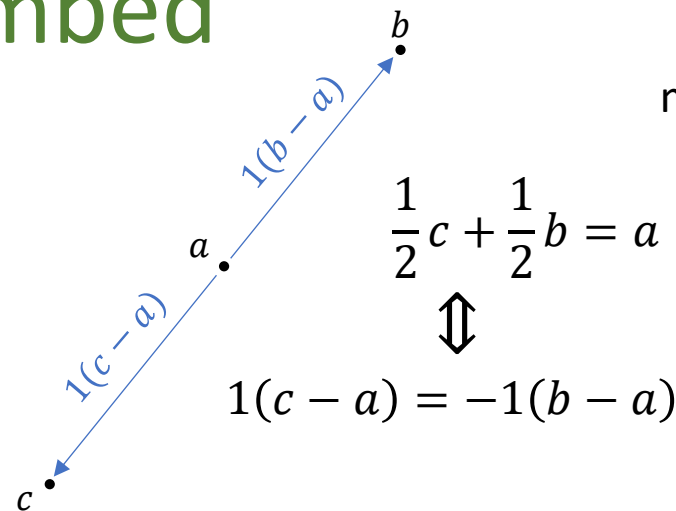


Definition 1.55 (Ensemble differences). Given an ensemble space, a difference between two ensemble represents the change required to transform one ensemble into another. Formally, an **ensemble difference**, noted $r(\mathbf{b} - \mathbf{a})$, is a triple formed by a real number $r \in \mathbb{R}$ and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

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

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

Ensemble spaces embed into vector spaces

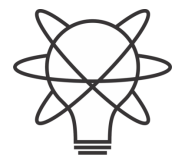


Vector space operations
derived from
mixing operations

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



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

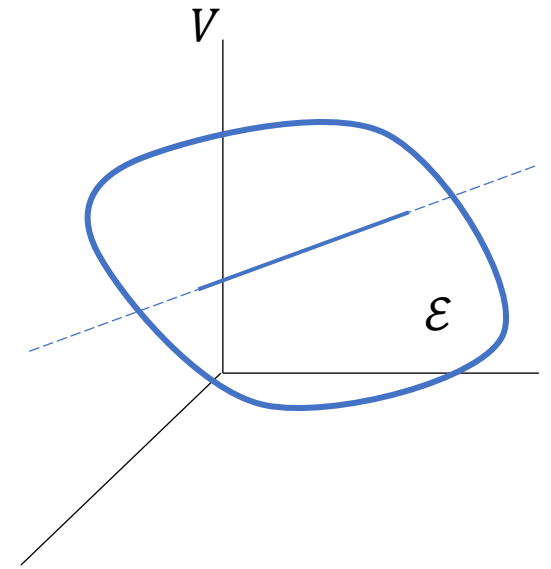


<https://assumptionsofphysics.org/>

Assumptions
of
Physics

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 $e_1, e_2, e_3 \in A$ there exists a permutation $\sigma : \{1, 2, 3\} \rightarrow \{1, 2, 3\}$ and $p \in [0, 1]$ such that $e_{\sigma(1)} = pe_{\sigma(2)} + \bar{p}e_{\sigma(3)}$.

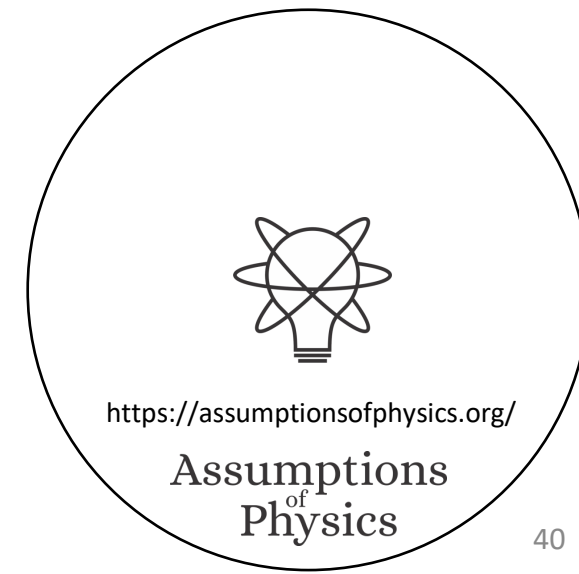
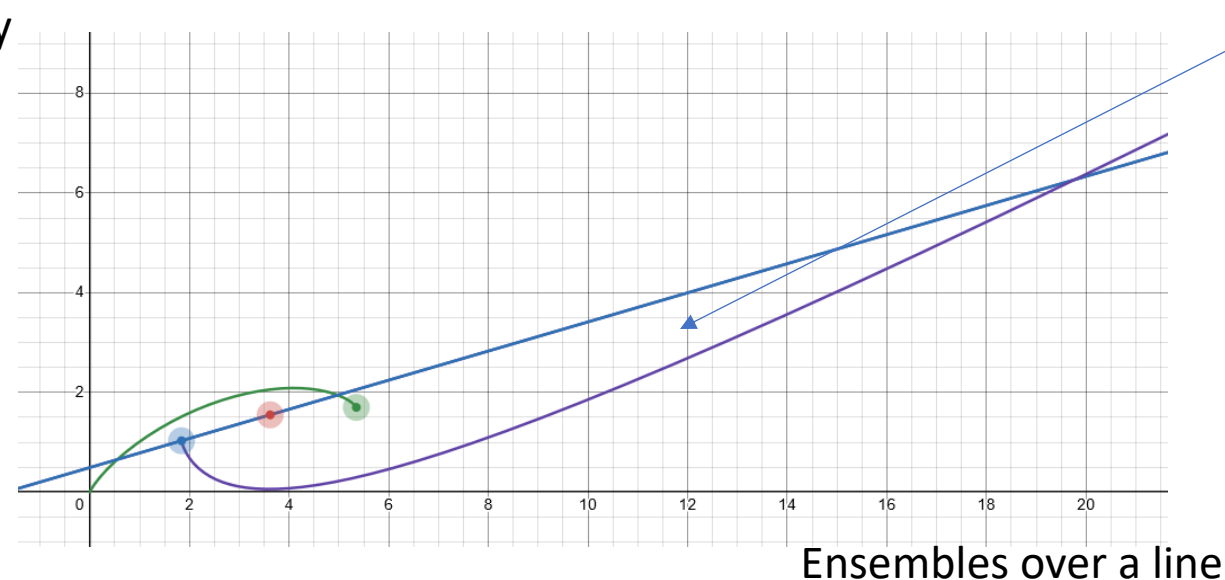
Theorem 1.52 (Lines are bounded). Let $A \subseteq \mathcal{E}$ be a line. Then we can find a bounded interval $V \subseteq \mathbb{R}$ and an invertible function $f : A \rightarrow V$ such that $f(pa + \bar{p}b) = pf(a) + \bar{p}f(b)$ for all $a, b \in A$.



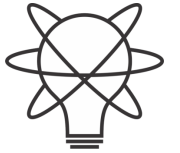
Ensemble spaces are bounded in all directions

Entropy bounds $\Rightarrow$ green point between blue and purple curves

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



Geometric structures from entropy



<https://assumptionsofphysics.org/>

Assumptions
of
Physics

How much does the entropy increase during mixture?

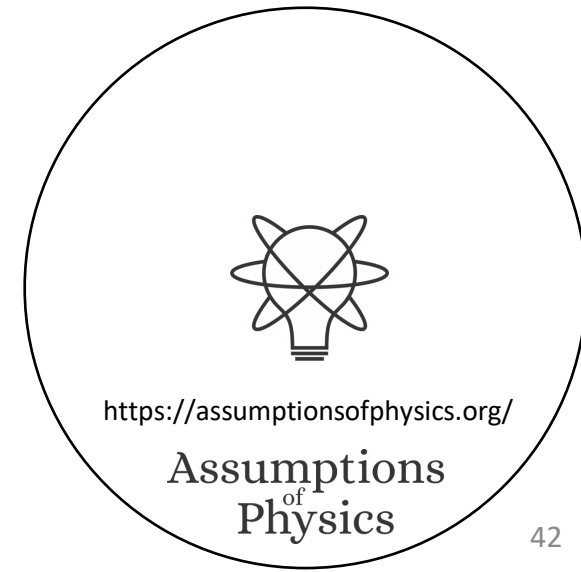
$$MS(a, b) = S\left(\frac{1}{2}a + \frac{1}{2}b\right) - \left(\frac{1}{2}S(a) + \frac{1}{2}S(b)\right)$$

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

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

Pseudo-distance defined from the entropy

(does not satisfy the triangle inequality)



Entropy imposes a metric on the ensemble space

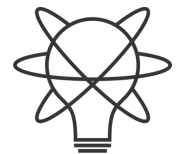
$$\|\delta \mathbf{e}\|_{\mathbf{e}} = \sqrt{8MS(\mathbf{e}, \mathbf{e} + \delta \mathbf{e})}$$

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

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

Entropy strict concavity means the Hessian is negative definite

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



<https://assumptionsofphysics.org/>

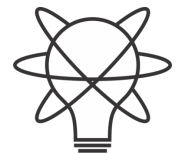
Assumptions
of
Physics

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

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

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

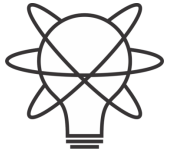
Another direct calculation
from the definitions



<https://assumptionsofphysics.org/>

Assumptions
of
Physics

Measure theoretic structures from entropy



<https://assumptionsofphysics.org/>

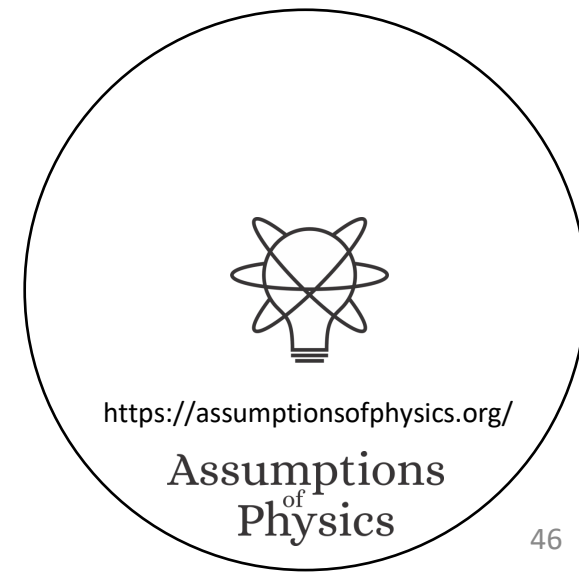
Assumptions
of
Physics

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

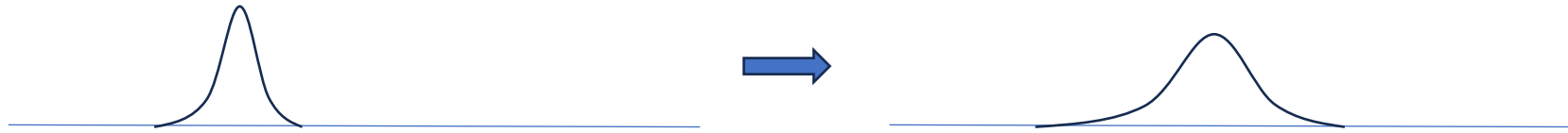
Count of states
(Phase-space volume)

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



We can think of an ensemble as spread over distinguishable cases



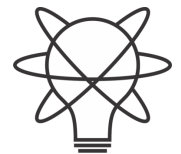
The number of distinguishable cases must increase with entropy



During mixing, the number of distinguishable cases at most sums

Note:

Proposition 1.153 (Exponential entropy subadditivity). *Let $\mathbf{e}_1, \mathbf{e}_2 \in \mathcal{E}$. Let $S_1 = S(\mathbf{e}_1)$ and $S_2 = S(\mathbf{e}_2)$. Let $\mathbf{e} = p\mathbf{e}_1 + \bar{p}\mathbf{e}_2$ for some $p \in [0, 1]$ and $S = S(\mathbf{e})$. Then $2^S \leq 2^{S_1} + 2^{S_2}$, with the equality if and only if $\mathbf{e}_1$ and $\mathbf{e}_2$ 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 $\mathbf{e}_1, \mathbf{e}_2 \in \mathcal{E}$. Let $S_1 = S(\mathbf{e}_1)$ and $S_2 = S(\mathbf{e}_2)$. Let $\mathbf{e} = p\mathbf{e}_1 + \bar{p}\mathbf{e}_2$ for some $p \in [0, 1]$ and $S = S(\mathbf{e})$. Then $2^S \leq 2^{S_1} + 2^{S_2}$, with the equality if and only if $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal and $p = \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}$.*

Maximum increase when ensembles are orthogonal.

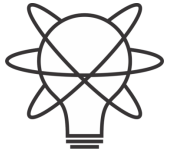
Find p that maximizes entropy:

$$\begin{aligned} 0 &= \frac{d}{dp} S(pa + \bar{p}b) = \frac{d}{dp} (-p \log p - \bar{p} \log \bar{p} + pS_a + \bar{p}S_b) \\ &= -\log p - 1 + \log \bar{p} + 1 + S_a - S_b \\ \log \frac{p}{\bar{p}} &= \log 2^{S_a} - \log 2^{S_b} \\ \log \frac{p}{1-p} &= \log \frac{2^{S_a}}{2^{S_b}} \\ p2^{S_b} &= (1-p)2^{S_a} \\ p(2^{S_a} + 2^{S_b}) &= 2^{S_a} \\ p &= \frac{2^{S_a}}{2^{S_a} + 2^{S_b}} \end{aligned}$$

Then calculate
final entropy:

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

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 $\Rightarrow$ 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 $\Rightarrow$ 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 $\text{scap}(A) = \sup(2^{S(\text{hull}(A))} \cup \{0\})$.

capacity also name
of a non-additive measure

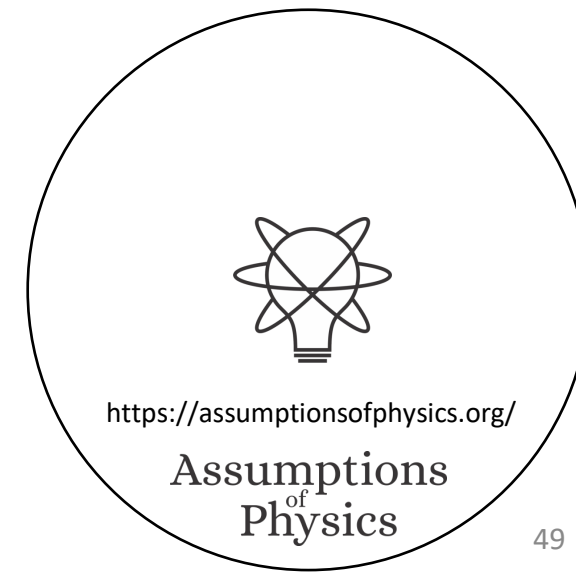
Proposition 4.134. The state capacity is a set function that is

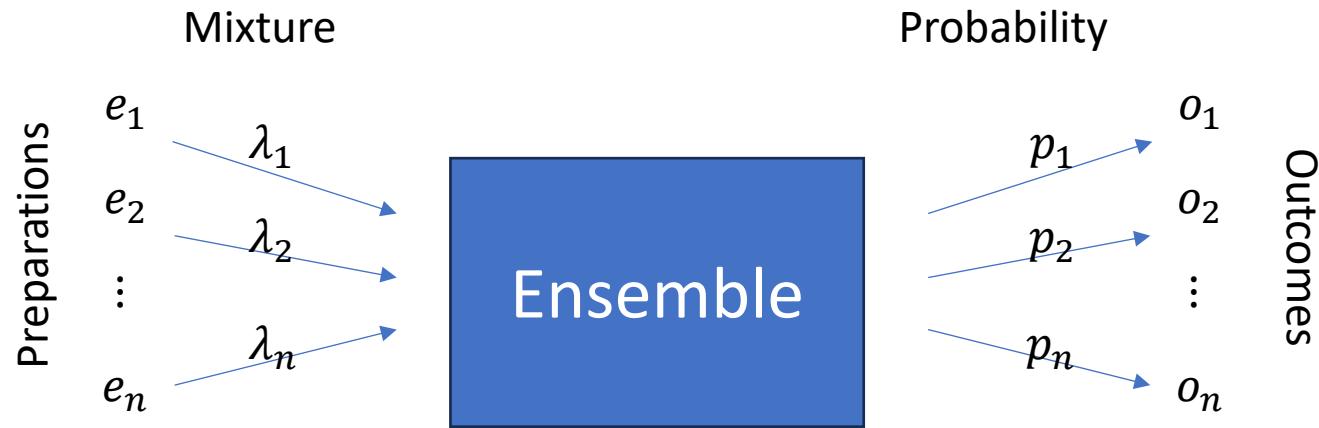
1. non-negative: $\text{scap}(A) \in [0, +\infty]$
2. monotone: $A \subseteq B \implies \text{scap}(A) \leq \text{scap}(B)$
3. subadditive: $\text{scap}(A \cup B) \leq \text{scap}(A) + \text{scap}(B)$
4. additive over orthogonal sets: $A \perp B \implies \text{scap}(A \cup B) = \text{scap}(A) + \text{scap}(B)$

fuzzy measure

State capacity is a non-additive measure

additive over orthogonal sets



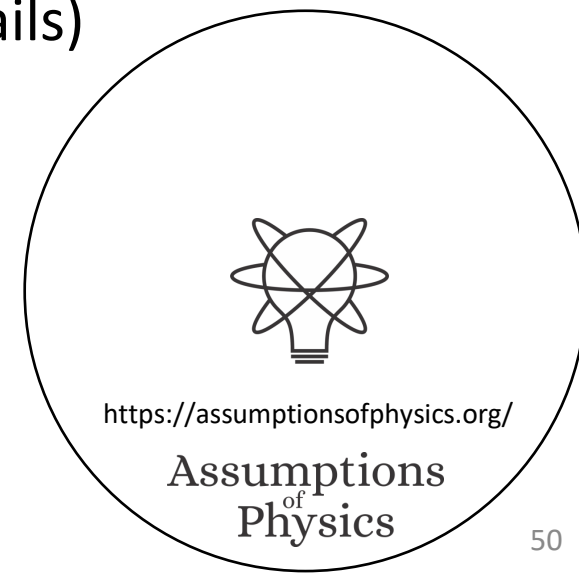


In classical mechanics, mixtures of preparations and probability of outcomes always coincide

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?



How much of e is a mixture of other ensembles?

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

Definition 1.83. Let $e, a \in \mathcal{E}$ 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, $\text{frac}_e(a) = \sup(\{p \in [0, 1] \mid \exists b \in \mathcal{E} \text{ s.t. } e = pa + \bar{p}b\})$.

Definition 1.85. Let $e \in \mathcal{E}$ 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, $\text{fcap}_e(A) = \sup(\text{frac}_e(\text{hull}(A)) \cup \{0\})$.

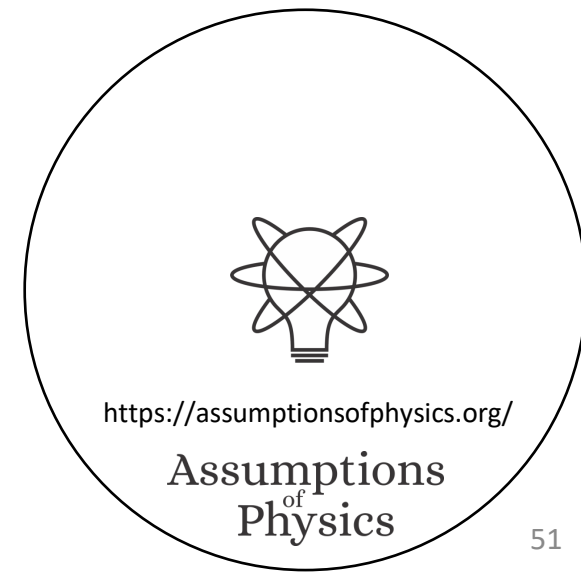
biggest p

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

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

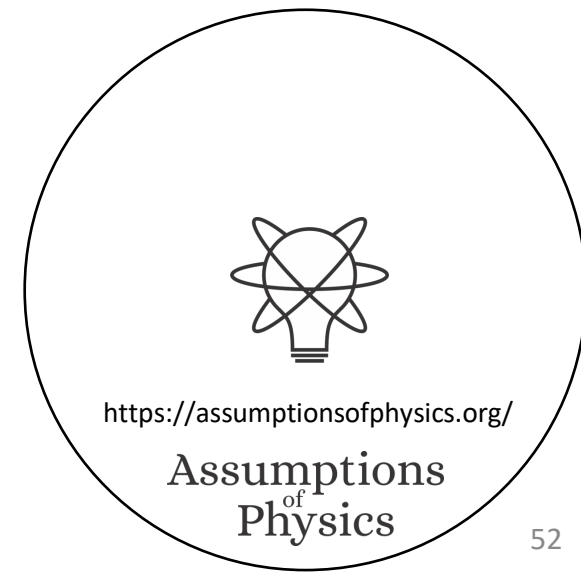
fuzzy measure

Fraction capacity is a non-additive probability measure



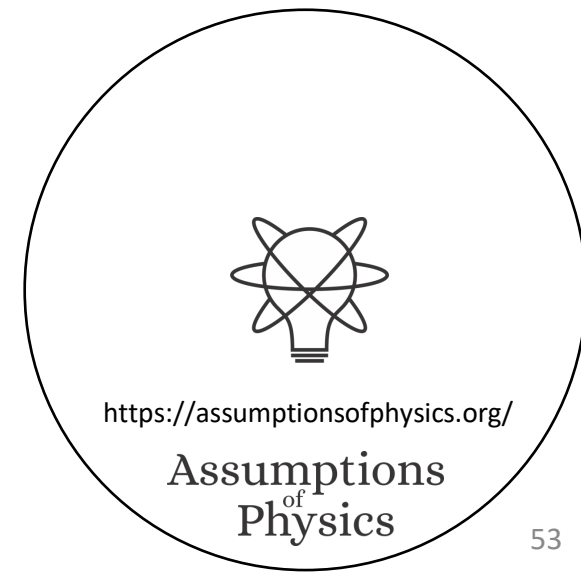
Takeaways

- 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



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



Assumptions of Physics is an open research program

Our main output is an open access book: <https://assumptionsofphysics.org/book/>

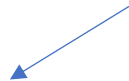
All our material is developed on GitHub: <https://github.com/assumptionsofphysics>

One YouTube channel dedicated to publicize results: <https://www.youtube.com/user/gcarcassi>

Another YouTube channel dedicated to research: <https://www.youtube.com/@AssumptionsofPhysicsResearch>

Activities coordinated through a Discord server (contact info@assumptionsofphysics.org for an invite)

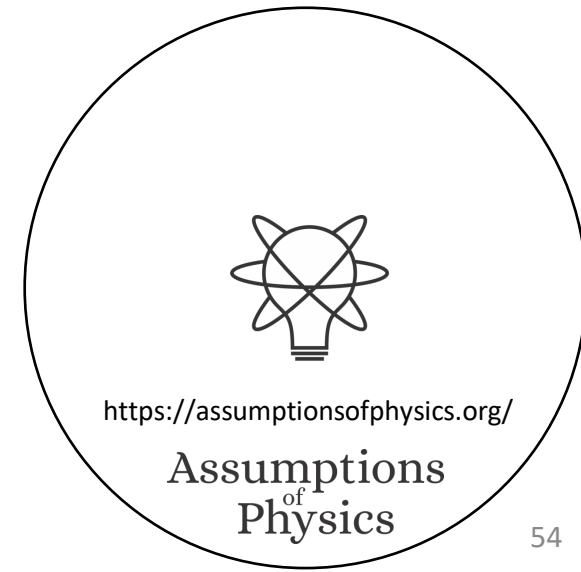
Livestream
discussions

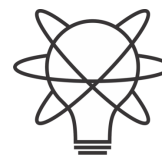


Always looking for experts to gain insights and/or help

Always looking for collaborations

Always looking for editors/journals/conferences that are sympathetic to the mission





<https://assumptionsofphysics.org/>

Assumptions
of
Physics