Assumptions of Physics Summer School 2025

Some results in Physical Mathematics

Gabriele Carcassi and Christine A. Aidala

Physics Department University of Michigan





Overview



Space of the well-posed scientific theories



3

Physics





Recover the reals





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

 \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



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







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: ${\sf b}$ and ${\sf a}$ are verifiable statements
- 2. the object can be on the reference: $o \not\equiv \bot$
- 3. if it's not before or after, it's on the reference: $\neg b \land \neg a \leq o$
- 4. if it's before and after, it's also on the reference: $b \land a \leq o$

A beginning reference has nothing before it. That is, $b \equiv \bot$. An ending reference has nothing after it. That is, $a \equiv \bot$. A terminal reference is either beginning or ending.

Before	On	After
Т	F	F
F	Т	F
F	F	т
т	Т	F
F	т	т
Т	Т	Т

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
Т	F	F
F	т	F
F	F	т



before

after

Multiple references

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

	b ₂	0 2	<i>a</i> ₂
b ₁	\checkmark	\checkmark	\checkmark
<i>o</i> ₁	\checkmark	\checkmark	\checkmark
<i>a</i> ₁	\checkmark	\checkmark	\checkmark



2. Aligned references

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



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



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

 a_2



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



 a_2

a

 b_2

 b_1

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





CHAPTER 7 PROPERTIES AND OUANTITIES

the possibilities themselves can be ordered, and how this ordering, in the end, is uniquely

the positivities the distributions the distribution of the ordered and the ordering, in the each is unique of distribution of the distribution of the distribution of the distribution 10^6 is narrower than "the quantity is less than 10^6 . 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 linear 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

are not linearly ordered: no direction is greater than the other. Therefore, no this context, as the order will an articularly but a quantity but an odderston of quantitative linear l has one natural satellite" is conjugent to the "the earth has more than zero natural se and fewer than two". Therefore we will define the order topology as the one generated by set of the type (a, ∞) and $(-\infty, b)$. A quantity, then, is an ordered property with the order topology.

Definition 3.4. A linear order on a set Q is a relationship $\leq Q \times Q \rightarrow B$ such that: mmetry) if $q_1 < q_2$ and $q_2 < q_2$ then $q_1 = q_2$

2. (transitivity) if $q_1 \leq q_2$ and $q_2 \leq q_1$ into $q_1 \leq q_2$ 3. (transitivity) if $q_1 \leq q_2$ and $q_2 \leq q_2$ then $q_1 \leq q_2$ 3. (total) at least $q_2 \leq q_2$ 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 topol

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

Definition 3.6. A mantitu for an emerimental domain Dy is a linearly ordered non Deminding uses $X \in X$ quantify for an experimental domain 2X is a trade-generating property formally, it is a tuple $(Q, \leq q)$ where (Q, q) is a property, $\leq Q \times Q \rightarrow B$ is a linear or and Q is a topological space with the order topology with respect to \leq .

As for properties, the quantity values are just symbols used to label the different cases set Q may correspond to the integers, real numbers or a set of words ordered alphabetic. The units are not captured by the numbers themselves: they are captured by the funct

⁹In other languages, there are two words to differentiate quantity as in "physical quantity" (e.g. grand cosee, grandeur) and as in "amount" (e.g. quantità, Menge, quantité). It is the second meaning of qua

at is explured here. ⁵The sentence: "the mass of the electron is 511±0.5 keV" could instead be referring to statistical unce-transformer in the statistical unce-The section: We make of the defirms a 511-63 keV could instead be referring to statistical inserts instead of an accury bound and would constitute a different ansatzing is atta-We will be treading these types of statistical statements have instances when the different manipular is be default bodies matematists this distribuilty bounds. ⁴Whom consulting the divisory, we use the fact that we can experimentally tell whether the word w looking for is budies or about the arounds) states.

ents of the original set and therefore reduces to countable conjunctions. Therefore, when forming D_k the only new elements will be the countable distunction Consider two countable sets $B_1, B_2 \subseteq B_b$. Their disjunctions $b_1 = \bigvee_{b \in B_1} b$ and $b_2 = \bigvee_{b \in B_2} b$.

esent the narrowest statement that is broader than all elements of the respec uppose that for each element of B_1 we can find a broader element in B_2 . Then b_2 , being proader than all elements of B_2 , will be broader than all elements of B_1 . But since by is where the main terms of B_2 , where B_1 is the interval of the main elements of B_1 . For since O_1 is the narrowest element that is broader than all elements in B_1 , we have $b_2 \ge b_1$. Conversely, uppose there is some element in B_1 for which there is no broader element in B_2 . Since he initial set is fully ordered, it means that that element of B_1 is broader than all the the initial set is fully broken, is means that common the common of D_1 is broader than the above that D_2 . This means that element is broader than D_2 and since D_1 is broader than all elements in B_1 we have $b_1 \ge b_2$. Therefore the domain D_b generated by B_b is linearly rdered by narrownes

Now we show that $(D_{\alpha, \geq})$ is linearly ordered. The basis B_{α} is linearly ordered by adness because the negation of its elements are part of B and are ordered by narrowness ote that broadness is the opposite order of narrowness and therefore a set linearly ordered w one is linearly ordered by the other. Therefore B_{τ} is also linearly ordered by narrowness by one is meanly ordered by the order. Therefore D_a is also meanly ordered by marked and so is D_a by the previous argument. Therefore D_a is ordered by broadness. To show that $D = D_b \cup \neg(D_a)$ is linearly ordered by narrowness, we only need to show

that the countable disjunctions of elements of B_b are either narrower or broade the contract of the negations of the negations of elements of D_0 in Let $D_1 \subset B_0$ and A_2 (disjunction $b_1 = \bigcup_{k=0}^{N} b$ represents the narrowest statement that is broader than all

proader than that one element in $\neg(A_2)$. But since $\neg a_2$ is narrower than all el broader than that one example. Conversely, suppose that for no element of $\neg(A_2)$, we have $\neg a_2 \le b_1$. Conversely, suppose that for no element $or (A_2)$ we broader statement in B_1 . As B is linearly ordered, it means that all elements in proader than all elements in B_1 . This means that all elements in $\neg(A_2)$ are bro b_1 and therefore $b_1 \leq \neg a_2$. Therefore D is linearly ordered by narr

Theorem 3.16 (Domain ordering theorem), An experimental domain D_X is rdered if and only if it is the combination of two experimental domains $D_X = D_a$

(i) D = D_b ∪ ¬(D_a) is linearly ordered by narrows (ii) all elements of D are part of a pair (s_b, −s_a) such that s_b ∈ D_b, s_a ∈ D_a ∈ either the immediate successor of s_{i} in D or $s_{i} \equiv \neg s_{i}$

(iii) if $s \in D$ has an immediate successor, then $s \in D_b$

Proof. Let D_Y be a naturally ordered experimental domain. Let B_b and B_a 1 Proof. Let D_X be a matrixed ordered experimental domain. Let D_2 and D_a (as in 3.12 which means $B = B_0 \cup B_a$ is the basis that generates the order topo D_b be the domain generated by B_b and D_a be the domain generated by B_a . T rated from D_b and D_a by finite conjunction and countable disjunction and $D_X = D_b \times D_a$.

3.2 OUANTITIES AND ORDERING

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 a quantity. For example, the domain for the mass of a system will be fully characterized by a rad number grater than or equal to zero. Each possibility will be indeduced 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 possibilitie themselves are ordered. That is, "the mass of the sustem is 1 Ka" proceedes "the mass of the system is 2 Ke^{θ}. This ordering of the possibilities will be linked to the natural topology a he mass of the system is less than # Ke", the distuction of all possibilities that come befor a units by the specific the mean a ray, the supported of an presentative take come between articular 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 assign a quantity to an experimental domain if it already has a natural ordering of the same type.

3.2. QUANTITIES AND ORDERING

Definition 3.12. Let D_X be a naturally ordered experimental domain and X its possilities. Define $B_b = \{ {}^{a}x < x_1{}^{\alpha} | x_1 \in X \}$, $B_a = \{ {}^{a}x > x_1{}^{\alpha} | x_1 \in X \}$ and $B = B_b \sqcup \neg (B_a)$.

Definition 3.13. Let (O, <) be an ordered set. Let $a_1 = c \in O$. Then a_2 is an immediate

control of A. Let (Q_{-5}) be an order to be a property for any property of the information scenessor of q_1 and q_1 is an immediate produces sors of q_2 if there is no element structure tween them in the ordering. That is, $q_1 < q_2$ and there is no $q \in Q$ such that $q_1 < q < w$ we 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 (B.

Proof. Let $f: X \to B_b$ be defined such that $f(x_1) = "x < x_1"$. As there is one is

aly one statement " $x < x_1$ " for each $x_1 \in X$, f is a bijection. Suppose $x_1 \leq x_2$

mp one assume $T \leq x_1$, for each $T \in X_1$. It is indexedue, suppose $z_1 \leq x_2$, $s_2 = f(z_2) = \frac{|x_1|^2}{|x_1|^2} + \frac{|x_1|^2$

we $g(x_1) \equiv \bigvee_{\{x \in X \mid x > x_1\}} x \equiv \left(\bigvee_{\{x \in X \mid x > x_1\}} x \right) \lor \left(\bigvee_{\{x \in X \mid x > x_2\}} x \right) \equiv g(x_1) \lor g(x_2)$ and therefore

 $\begin{array}{l} & (\kappa X(m_1) - \left\{ (\kappa X(m_1) - 1 \right\} (\kappa X(m_1) - 1 \right\} \\ & (\kappa X(m_1) - 1 \right\} (\kappa X(m_1) - 1) \\ & (\kappa X(m_1) - 1 \right) \\ & (\kappa X(m_1) - 1 \right$

n 3.15. Let B_b and B_a be two sets of verifiable statements such that kis linearly ordered by narrowness. Let D_b and D_a be the experimental doma wely generate and $D = D_b \cup \cdots (D_a)$. Then (D_b, π) , $(D_a, *)$ and (D, π) are linearly

irst we show that (\mathcal{D}, x) is linearly ordered. We have that \mathcal{B}_i is linearly ordered.

intwo how that $(\{E_n\}_n)$ is immary ordered. We have that E_n is intractly ordered so because it is a value of d which is immedy control of particle of a finite set of statements insarly ordered by matrowness. We constrain the disjunction of a finite set of statements insarly ordered by matrowness will extra be constable disjunctions, in the constraints disjunctions, and an element. The constrable disjunctions, in the constrable disjunction, instead, can we elements. But using these descutes again will so that introduce now ensure is of cosmable disjunctions is the constrable disjunctions in the first set of cosmable disjunctions is the first set of cosmable disjunctions in the first set of cosmable disjunctions is the first set of cosmable disjunctions in the first set of cosmable disjunctions is the first set of cosmable disjunctions in the first set of cosmable disjunctions is the first set of cosmable disjunctions in the first set of cosmable disjunctions is the first set of cosmable disjunction in the cosmic set of cosmable disjunctions is the first set of cosmable disjunction is the cosmic set of co

Note that determining whether the quantity is exactly equal to the reference is not as er

finite amount of time. That is, the reference itself can only be compared up to a finite leve

of precision. This may be a problem when constructing the references themselves: how do we

to precision: a not may be a process when donast uctual in a reasonates unanserves: now us we know that the marks on our rules are equally propared, or that the weights are equally propared, or that ticks of our clock are equally timed? It is a circular problem in the sense that, in a way, we need instruments of measurement to be able to create instruments of measurement.

Yet, even if our references can't be perfectly compared and are not perfectly equal, we can

(4) even is due detendent care be perfectly token and her not perfectly sequent, we can still say whether the value is well before our well after any of them. To make matters worse, the object we are measuring may itself have an extent. If we are measuring the position of a tiny ball, it may be clearly before or clearly after the nearest

mark, but it may also be partly before, partly on and partly after. One may try to sidestep the problem by measuring part of the object, say the position of the center of mass or of its

closest part. But this assumes we have a process to interact with only part of the object, and that part can only be before, on or after the reference. It may be a reasonable assumption in

that part can only be below, or or are the reference. It may be a reasonable assumption in many cases but we have to be mindful that we made that assumption: our general definition will have to be able to work in the less ideal cases.

discussion with the following definitions. A reference is represented by a set of three state-

uscussed with the bolowing dominations. A thereton is represented by a fee on time scale ments: they fell as whether the object is before, on or after a specific reference. To make sense, these have to satisfy the following minimal requirements. The before and the after statements must be verifiable, as otherwise they would not be usable as references. As the

reference must be somewhere, the on statement cannot be a contradiction. If the object is

not before and not after the reference, then it must be on the reference. If the object is before and after the reference, then it must also be on the reference. These requirements recognize

that, in general, a restructure and all extensions and so useds the UQEX relations in meritance. We can compare the extent of two references and say that one is finear than the other if the on statement is narrower than the other, and the before and after statements are wider. This corresponds to a finer tick of a rule or a finer pulse in our timing system. We say that

a reference is strict if the before, on and after statements are incompatible. That is, the three

Definition 3.17. A reference defines a before, an on and an after relationship between

iself and another object. Formally a reference $\mathbf{r} = (\mathbf{b}, \mathbf{o}, \mathbf{a})$ is a tuple of three statements

1. we can verify whether the object is before or after the reference: b and a are verifiable

A beginning reference has nothing before it. That is, $b \equiv \bot$. An ending reference has solving after it. That is, $a \equiv \bot$. A terminal reference is either beginning or ending.

that, in general, a reference has an extent and so does the object being measured

ses are distinct and can't be true at the same time

2. the object can be on the reference: $0 \neq 1$

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

uch that:

In our general mathematical theory of experimental science, we can capture the above

e mark on the ruler has a width, the balance has friction, the tick of our clock will last a

71

3.3. REFERENCES AND EXPERIMENTAL ORDERING

 (B_{i}, z) and (B, z) are linearly ordered sets. Moreover (B_{i}, z) , (B_{i}, z) are order isomo

 $x < x_1^n \land x \ge x_1^n \equiv \bigvee_{n=1}^{\infty} x \equiv \bot$

69

CHAPTER 3. PROPERTIES AND QUANTITIES

mass of the system is more than $q_1 K g^+$ 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 topologic and therefore spluty constitute a basis for the coperturbational domain. Now consider the statement $s_1 = the mass of the system is less than or equal to 1 K g^+$ with $s_2 = the mass of the system is less than 1 K g^+$. We have $s_2 < s_1$. In K_{11} , W are place with $s_2 = the mass of the system is <math>t_{12}$. Also, W is a single state of the system is the state of the system is the state of the system is the size the statement $s_1 = the mass of the system is the size of the size$ the value in s_2 with anything less than 1 Kg we'll still have $s_2 \in s_1$. Instead if we use a value greater than 1 Kg we'd have $s_1 \approx s_2$. In other words, if we call B the set that includes both e less-than-or-could and less-than statements this is also linearly ordered by narrowne the loss blue-or-equal and less-than statements this is also linearly ordered by narrowness. But the mass of the segment is its abase or equal to $t K^0$ is equivalent $t \to -K$ here mass of the spatra is generater than $t K^0_{\pi}$. In other words, $B = \delta_{\pi} \cup -\delta_{\pi}$) contain all the statements like M^0_{π} and $M^0_{$

is less than or equal to 1 Kg^{*} is the immediate successor of $s_2 =$ "the mass of the system is less than 1 Kg^{*}. That is, they are different and there can't be any other statement in B that ice non T, R_T . This is, how are universit nor there can t be any other statisfield in D mass is broken that g_T but narrows that g_T into the g_T fifter a single case. This will happen for any mass value. So B is composed of two exact copies of the ordering of X, where each element of one couple in immulating billowed by an element of the other case, G_T . So G_T we can a statement in B has an immulate successor, there must be only one case that separates the two. If we call μ the value of that to exe then the statement must be of the form 3^{-1} because of the system is less than $q_1 Kg^{*}$ while its immediate successor is of the form "the mass of the system is less than or equal to $q_1 Kg^{*}$: the successor is broader by just the possibility ated with q_1 . Therefore statements in B that have an immediate successor must be in B_k as well. The main result is that the above characterization of the basis of the domain is necessary

and sufficient to order the nossibilities. If an experimental domain has a basis composed of

3.2. OUANTITIES AND ORDERING

To prove (i), we have that \mathcal{B}_b and \mathcal{B}_a are linearly ordered by 3.14. We need to show that ie linear ordering holds across the sets. Let $x_1, x_2 \in X$ and consider the two statements $= D_h \cup \neg (D_n)$ is also linearly ordered.

 $D = D_{\delta} \supset \langle U_{\alpha} \rangle$ is also integrity ordered. To prove (ii), let $s_{\delta} \in D_{\delta}$. Take $s_{\alpha} \in D_{\alpha}$ such that $\neg s_{\alpha}$ is the narrowest statement in $\neg (D_{\alpha})$ that is broader than s_{α} . This exists because D_{α} is closed by infinite disjunction. As $s_1 \ge s_2$, let X_1 be the set of possibilities compatible with $\neg s_2$, but not compatible with s_2 . $s_a > s_a$, is A_1 be the set of possibility of a could find an element $x_1 < X_1$ such that $b_1 \leq u^* x \leq x_1^n \prec -s_a$. If X_1 contains one possibility, then $-s_a$ is the immediate successor. If 1 is empty then $s_b \equiv -s_a$. Similarly, we can start with $s_a \in D_a$ and find $s_b \in D_b$ such that s_b x_1 is enough the x_0 - x_0 - x_0 in x_0 , we can solve that x_0 . Let X_1 be the set of possibilities is the broadest statement in \mathcal{D}_b that is narrow than x_0 . Let X_1 be the set of possibilities compatible with $-x_0$ but not compatible with s_0 . If X_1 contains one possibility, then $-s_0$ is the immediate successor and if X_1 is empty then $s_1 = -x_0$. To prove (iii), let $s_1, s_2 \in D$ such that s_2 is the immediate successor of s_1 . This mean can write $s_2 \equiv s_1 \lor x_1$ for some $x_1 \in X$. This means $s_1 \equiv "x < x_1"$ while $s_2 \equiv "x \le x_1"$ and

therefore $s_1 \in \mathcal{B}_b$.

CHAPTER 3. PROPERTIES AND QUANTITIES

of the immediate successors. Let $(\cdot)^{++}: B_h \to B_n$ be the function such that $\neg(b^{++}) = \neg b^{++}$ is the immediate successor of b. Let b: $X \to B_b$ be the function such that $x = b(x) - b(x)^{++}$. On X define the ordering \leq such that $x_1 \leq x_2$ if and only if $b(x_1) \leq b(x_2)$. Since (B_b, \leq) s linearly ordered so is (X, \leq) . To show that the ordering is natural, suppose $x_1 < x_2$ hen $b(x_1) \leq -b(x_2)^{++} \leq b(x_2)$ and therefore $x_1 \leq b(x_2)$. We also have $-b(x_2)^{++} \leq b(x_2)$. and $o_{11}(x \to o_{12}) \to o_{12}(x)$ and therefore $x_1 \in o_{12}(x)$. We also have $\neg o_{11}(x) \to o_{12}(x) \to o_{12}(x) \to o_{12}(x)$. We also have $\neg o_{11}(x) \to o_{12}(x) \to o_{12}(x)$. We have $a_1(x) \to a_2(x) \to o_{12}(x)$. This means that given a possibilities $x_1 \in X$, all and only the possibilities lower than x_1 are compatible with $b(x_1)$ and therefore $b(x_1) \equiv "x < b_{12}(x)$. x_1^n , while all and only the possibilities greater than x_1 are compatible with $b(x_1)^{++}$ and herefore $b(x_1)^{*+} \equiv "x > x_1"$. The topology is the order topology and the domain has a

3.3 References and experimental ordering

In the previous section we have characterized what a quantity is and how it relates to an experimental domain. But as we saw in the first chapters, the possibilities of a domain are not objects that exist a priori: they are defined based on what can be verified experimentally. Therefore simply assigning an ordering to the possibilities of a domain does not answer the more fundamental question: how are quantities actually constructed? How do we, in practice create a system of references that allows us to measure a quantity at a given level of precision? What are the assumptions we make in that process?

In this section we construct ordering from the idea of a reference that physically defines a boundary between a before and an after. In general, a reference has an extent and may overlap with others. We define ordering in terms of references that are clearly before and overapy want Others, we denote the possibilities have as on references unat and covery neuron and after others. We see that the possibilities have a natural ordering only if they are generated from a set of references that is refinable (we can always find finer ones that do not overlap) and for which before/on/after are mutually exclusive coses. The possibilities, then, are the finest references possible.

We are by now so used of the ideas of real numbers, negative numbers and the number zero that it is difficult to realize that these are mental constructs that are, in the end, somewhat recent in the history of humankind. Yet geometry itself started four thousand years ago as an experimentally discovered collection of rules concerning lengths, areas and angles. That is, human beings were measuring quantities well before the real numbers were invented. So, how does one construct instruments that measure values?

To measure position, we can use a ruler, which is a series of equally spaced marks. We give a label to each mark (e.g. a number) and note which two marks are closest to the targe sition (e.g. between 1.2 and 1.3 cm). To measure weight, we can use a balance and a set of ually prepared reference weights. The balance can clearly tell us whether one side is heavier than the other, so we use it to compare the target with a number of reference weights and note the two closest (e.g. between 300 and 400 grams). A clock gives us a series of events to more new cool cosess (e.g., ostewen; sool and 'soo gland). A cook gives us a sense or evends to compare to (e.g. earth's rotation on its axis, the ticks of a clock). We can pour water from a reference container into another as many times as are needed to measure its volume. In all these cases what actually happens is similar: we have a reference (e.g. a mark on a ruler, a set of equally prepared weights, a number of ticks of a clock) and it is fairly easy to tel

CHAPTER 3. PROPERTIES AND QUANTITIES

Proof. By definition, we have $\neg b \land \neg a \leq o$ and by 1.23 $\neg (\neg b \land \neg a) \lor o \equiv T \equiv b \lor a \lor o$. 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 $b_1 \ge b_2$, $o_1 \le o_2$ and $a_1 \ge a_2$.

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. tisymmetry and transitivity properties and is therefore a partial order.

Definition 3.21. A reference is strict if its before, on and after statements are incom with a tible. Formally, r = (b, o, a) is such that $b \neq a$ and $o \equiv \neg b \land \neg a$. A reference is loose if it e not strict

Remark. In general, we can't turn a loose reference into a strict one. The on statement an be made strict by replacing it with $\neg b \land \neg a$. This is possible because o is not required to e 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 \mathbf{r}_1 is before reference \mathbf{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 s Mathematically, this type of orde

before and strictly after. It does n One may be tempted to define the roquires refining the references and refined references, not the original or

72

Definition 3.22. A reference is be he first it cannot be on or after the Proposition 3.23. Reference ord irreflexivity; not r < r transitivity: if r₁ < r₂ and r₂

and is therefore a strict partial of Proof. For irreflexivity, since th

nd therefore by o nova. Therefo reflexive.

Proposition 3.28. Let $\mathbf{r}_1 = (\mathbf{b}_1, \mathbf{o}_1, \mathbf{a}_1)$ and $\mathbf{r}_2 = (\mathbf{b}_2, \mathbf{o}_2, \mathbf{a}_2)$ be two strict references. Then *Proof.* Let $\mathbf{r}_1 < \mathbf{r}_2$. By 3.27, we have $-a_1 \leq b_2$. Conversely, let $-a_1 \leq b_2$. Then $-a_1 \neq -b_2$. because the references are strict, $\neg a_1 \equiv b_1 \lor o_1$ and $\neg b_2 \equiv o_2 \lor a_2$. Therefore $b_1 \lor o_1 \neq o_2 \lor a_2$ and $\mathbf{r}_1 < \mathbf{r}_2$ by definition.

Proof. We have $b_1 \lor o_1 \equiv (b_1 \lor o_1) \land \top \equiv (b_1 \lor o_1) \land (b_2 \lor o_2 \lor a_2) \equiv ((b_1 \lor o_1) \land b_2) \lor$

 $(b_1 \vee o_1) \land (o_2 \vee a_2) \equiv ((b_1 \vee o_1) \land b_2) \lor \perp \equiv (b_1 \vee o_1) \land b_2$. Therefore $b_1 \lor o_1 \preccurlyeq b_2$. And

Similarly, we have $o_2 \vee a_2 \equiv (o_2 \vee a_2) \wedge T \equiv (o_2 \vee a_2) \wedge (b_1 \vee o_1 \vee a_1) = ((o_2 \vee a_2) \wedge (b_1 \vee a_2))$

Since $b_1 \vee o_1 \vee a_1 \equiv T$, we have $\neg a_1 \preccurlyeq b_1 \vee o_1$. Similarly $\neg b_2 \preccurlyeq o_2 \vee a_2$. Since $b_1 \vee o_1 \neq o_2 \vee a_2$.

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

Since $b_1 \vee o_1 \neq o_2 \vee a_2,$ we have $b_1 \neq a_2$ which means $b_1 \preccurlyeq \neg a_2.$

 $(o_1 \lor a_2) \land a_1) \equiv \bot \lor ((o_2 \lor a_2) \land a_1) \equiv (o_2 \lor a_2) \land a_1$. Therefore $a_2 \lor o_2 \lessdot a_2) \land (o_1 \lor a_1) \equiv (o_2 \lor a_2) \land a_1$.

nce $b_1 \leq b_1 \vee o_1$, we have $b_1 \leq b_2$.

 $\leq o_2 \lor a_2$, we have $a_2 \leq a_1$.

 $a_1 \neq \neg b_2$ and therefore $\neg a_1 \neq b_2$.

 $1 < r_2$ if and only if $\neg a_1 \leq b_2$.

CHAPTER 3. PROPERTIES AND QUANTITIES

Definition 3.29. A reference is the immediate predecessor of another if nothing can be bound before the second and after the first. Formally, $r_1 < r_2$ and $a_1 * b_2$. Two references re consecutive if one is the immediate successor of the oth

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 nmediately before \mathbf{r}_2 then $\mathbf{b}_2 \equiv \neg \mathbf{a}_1$.

Proof. Let r_1 be immediately before r_2 . Then $a_1 \neq b_2$ which means $b_2 \preccurlyeq \neg a_1$. By 3.27 e also have $\neg a_1 \leq b_2$. 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 is immediately before \mathbf{r}_2 if and only if $\mathbf{b}_2 \equiv -\mathbf{a}_1$

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 $a_1 \neq \neg a_1$, therefore $a_1 \neq b_2$ and r_1 is immediately before r₂ by definition.

CHAPTER 3. PROPERTIES AND QUANTITIES

neans we can find $\mathbf{r}_3 = (\mathbf{b}, \neg \mathbf{b} \land \neg \mathbf{a}_2, \mathbf{a}_2)$ for some $\mathbf{b} \in \mathcal{D}_k$ such that $\mathbf{r}_3 < \mathbf{r}_2$ and therefore ¬a1 ≤ b ≤ ¬a2.

For the third, suppose $a_1 \in D_a$ and $b_2 \in D_b$ such that $\neg a_1 \prec b_2$. Then $r_1 = (\bot, \neg a_1, a_1)$ and $\mathbf{r}_2 = (\mathbf{b}_2, \neg \mathbf{b}_2, \bot)$ are strict references aligned with the domain such that $\mathbf{r}_1 < \mathbf{r}_2$ but \mathbf{r}_2 flate successor of \mathbf{r}_1 . This means we can find $\mathbf{r}_3 = (\mathbf{b}, \neg \mathbf{b} \land \neg \mathbf{a}, \mathbf{a})$ such that $\mathbf{r}_1 < \mathbf{r}_3 < \mathbf{r}_2$ and therefore $\neg a_1 \leq b < \neg a \leq \neg b_2$.

Proposition 3.37. Let D be an experimental domain generated by a set of refinable aligned references. Then all elements of D are part of a pair $(s_b, \neg s_a)$ such that $s_b \in D_b$, $s_a \in D_a$ 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 immediate successor, then $s \in D_b$.

Proof. Let $\mathcal D$ be an experimental domain generated by a set of refinable aligned strict eferences. Let $s_b \in D_b$. Let $A = \{a \in D_a | a \lor s_b \notin \top\}$. Let $s_a = \bigvee_{a \in A} a$. First we show that $s_b \leq \neg s_a$. We have $s_b \land \neg s_a \equiv s_b \land \neg \lor a \equiv s_b \land \land \neg a \equiv \land s_b \land \neg a$. For all $a \in A$ we have $a \lor s_b \notin T$, $\neg a \notin s_b$ which means $s_b \notin \neg a$ because of the total order of D. This means that

 $\wedge \neg a \equiv s_b$ 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 $s \in D$ is such that $s_h < s < -s_a$. Let $a \in D_a$ such that $s_{b} < -a$. By construction $a \in A$ and therefore $-s_{a} \leq -a$. Therefore we can't have $s_{b} < a < -s_{a}$. We also can't have $b \in D_{b}$ such that $s_{b} < b < -s_{a}$: by 3.36 we'd find $a \in D_{a}$ such that

 $s_b < a \le b < -s_0$ which was ruled out. So there are two cases. Either $s_b \neq -s_0$ then $s_b < -s_0$: s_0 is the immediate successor of b. Or $s_b \equiv \neg s_0$.

The same reasoning can be applied starting from $s_a \in D_a$ to find a $s_b \in D_b$ such that s_b is he 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 $s_1 \notin D_8$ and $s_2 \notin D_a$ we can always find another statement between the two. Then we must have that $s_1 \in D_k$ and $s_2 \in D_n$.

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

Proof. Suppose D_X is an experimental domain generated by a set of refinable aligned ences. Then by 3.34 and 3.37 the domain satisfies the requirement of theorem 16 and therefore is naturally ordered.

Now suppose D_X is naturally ordered. Define the set B_b , B_a and D as in 3.12. Let $R = \{(b, \neg b \land \neg a, a) | b \in B_b, a \in B_a, b \prec \neg a\}$ 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 $\neg b \land \neg a$ is not a contradiction since $b < \neg a$ means $b \neq a$ and $b \neq \neg a$. The on statement is broader than $\neg b \land \neg a$ as they are equivalent and it is broader than $b \land a$ as that is a contradiction since $b < \neg a$. Therefore R s 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.

3.3. REFERENCES AND EXPERIMENTAL ORDERING

For transitivity, if $\mathbf{r}_1 < \mathbf{r}_2$, we have $\mathbf{b}_1 \vee \mathbf{o}_1 \neq \mathbf{o}_2 \vee \mathbf{a}_2$ and therefore $\neg(\mathbf{b}_1 \vee \mathbf{o}_1) \geq \mathbf{o}_2 \vee \mathbf{a}_2$ by 1.23. Since $b_1 \vee o_1 \vee a_1 \equiv \tau$, we have $a_1 \ge \neg(b_1 \vee o_1)$. Similarly if $\mathbf{r}_2 < \mathbf{r}_3$ we'll have $a_2 \ge \neg (b_2 \lor o_2) \ge o_3 \lor a_3$. Putting it all together $\neg (b_1 \lor o_1) \ge o_2 \lor a_2 \ge a_2 \ge \neg (b_2 \lor o_2) \ge o_3 \lor a_3$. which means $b_1 \vee o_1 \neq o_3 \vee a_3$. 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 and

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 narrowe 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 notafter statements of the other. That is, alignment can also be defined in terms of narr 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 provides section, this was a necessary condition

3.3. REFERENCES AND EXPERIMENTAL ORDERING

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

Proposition 3.34. Let D be an experimental domain generated by a set of aligned stric for encoded and let $D = D_b \cup \neg (D_a)$. Then (D, \preccurlyeq) is linearly ordered.

Proof. By 3.26 we have that $B = B_b \cup \neg(B_a)$ is aligned by narrowness. By 3.15 the rdering 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 o not overlap, are linearly ordered and that cover the whole space. As we get to the fine 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 D be an experimental domain generated by a set of aligned references 2. The set of references is refinable if, given two strict references $r_1 = (b_1, o_1, a_1)$ and $\mathbf{p}_2 = (\mathbf{b}_2, \mathbf{o}_2, \mathbf{a}_2)$ 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 \mathbf{r}_1 , then we can find a strict reference \mathbf{r}_2 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 $o_1 < o_1$, 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

 $\mathbf{r}_1 < \mathbf{r}_2$ or $\mathbf{a}_2 \equiv \mathbf{a}_1$ and $\mathbf{r}_2 < \mathbf{r}_2$. Proposition 3.36. Let D be an experimental domain generated by a set of refinable aligned

77

3.4. DISCRETE QUANTITIES

Now we show that R consists of aligned strict references. We already saw that b * aand we also have $\neg b \land \neg a$ 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 inearly ordered by 3.14 which means the references are aligned.

To show R is refinable, note that each reference can be expressed as $(*x < x_1^n, *x_1 \le x \le x_2^n, *x > x_2^n)$ where $x_1, x_2 \in X$ and $*x_1 \le x \le x_2^n \equiv *x \ge x_1^n \land *x \le x_2^n$. That is, very reference is identified by two possibilities x_1, x_2 such that $x_1 \leq x_2$. Therefore take two references $\mathbf{r}_1, \mathbf{r}_2 \in R$ 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 $\mathbf{r}_1 < \mathbf{r}_2$ 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 " $x \le x_2$ " \le " $x < x_5$ " and " $x \le x_5$ " \le " $x < x_3$ ". Therefore the reference $\mathbf{r}_3 \in R$ dentified by (x_5, x_5) is between the two references. On the other hand, assume the second reference is finer than the first. Then $x_1 \le x_3$ and $x_4 \le x_2$ with either $x_1 \pm x_3$ or $x_4 \pm x_2$. Consider the references $\mathbf{r}_3, \mathbf{r}_4 \in R$ identified by (x_1, x_1) and (x_2, x_2) . Either $\mathbf{r}_3 < \mathbf{r}_2$ 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 we 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 ", "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

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



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

https://assumptionsofphysics.org/

Assumptions

Physics

Lack of order at small scales, (N.B. this is a toy model, each point should have order at large enough scale infinitely many neighbors) 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?

19

Takeaways

- Real numbers can be recovered from an idealized metrological model
 - Ordering of values comes from the ordering logical structure
 - $3 \le 5$ precisely because "there are less than 3 items" \le "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



Physics

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

Classical discrete

Classical continuum

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

Projective complex Hilbert space

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

State space

Phase space

Symplectic manifold

Ensembles

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

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

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

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 $+: [0,1] \times \mathcal{E} \times \mathcal{E} \to \mathcal{E}$ 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 $S : \mathcal{E} \to \mathbb{R}$, 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

26

https://assumptionsofphysics.org/

Assumptions Physics Strict concavity: S(pa + p̄b) ≥ pS(a) + p̄S(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(p₁, p₂) (i.e. the same for all ensemble spaces) such that S(pa + p̄b) ≤ I(p, p̄) + pS(a) + p̄S(b); if the equality holds, a and b are non-overlapping or orthogonal, noted a ⊥ b



Maximum increase when the ensembles are "completely different"

Increase is only a function of the mixing coefficient



Mixtures preserve orthogonality:^b a ⊥ b and a ⊥ c if and only if a ⊥ pb + pc for any p ∈ (0,1)



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



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(pa + p̄b) ≥ pS(a) + p̄S(b) with the equality holding if and only if a = b
- Upper variability bound: there exists a universal function I(p₁, p₂) (i.e. the same for all ensemble spaces) such that S(pa + p̄b) ≤ I(p, p̄) + pS(a) + p̄S(b); if the equality holds, a and b are non-overlapping or orthogonal, noted a ⊥ b
- Mixtures preserve orthogonality:^b a ⊥ b and a ⊥ c if and only if a ⊥ pb + pc for any p ∈ (0,1)

Ensemble variability ⇒ Entropy

No standard theory of entropic spaces



https://assumptionsofphysics.org/

Assumptions 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 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

$$S(\sum_i p_i a_i) = I(p_i) + \sum_i p_i S(a_i)$$
Uniform mixture of $n \times m$ elements

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

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

$$= I\left(\left\{\frac{1}{n}\right\}_{i=1}^n\right) + I\left(\left\{\frac{1}{m}\right\}_{i=1}^m\right) + \sum_{j=1}^n \sum_{k=1}^m \frac{1}{n} \frac{1}{m} S(a_{jk}).$$
Uniform mixture over n
uniform mixture over n
uniform mixtures of m elements

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

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

Assumptions Physics 33

Pick
$$a_i$$
 all orthogonal to each other

$$S(\sum_i p_i a_i) = I(p_i) + \sum_i p_i S(a_i)$$
Uniform mixture of m elements

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

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

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

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

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

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

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

https://assumptionsofphysics.org/

Assumptions Physics

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

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

$$= -\sum_{i=1}^n p_i \kappa \log \frac{m_i}{m} = -\kappa \sum_{i=1}^n p_i \log p_i.$$
From now on,
$$\log p \text{ is base 2}$$

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 Orthogonal ensembles

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

No "common component" $c \in \mathcal{E}$

such that

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

Coincide in classical Different

ensemble spaces



disjoint support





Simple but powerful definitions from basic axioms

Saturate upper entropy bound

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



Vector space constraints from entropy



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"

 $a \bullet a \bullet ra + \overline{r}b$

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 $I = \{i \in [1, n] \mid r_i \ge 0\}$ 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(b-a), is a triple formed by a real number $r \in \mathbb{R}$ and an ordered pair of ensembles $a, b \in \mathcal{E}$.

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 a, the **natural embedding** of \mathcal{E} into V_a is the map $\iota_a : \mathcal{E} \hookrightarrow V_a$, defined as $\iota_a(e) \to [(e-a)]$, that maps each ensemble to its difference from 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 $e_1, e_2, e_3 \in A$ 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 $f : A \to V$ 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 ⇒ green point between blue and purple curves



Assumptions

Physics

Geometric structures from entropy



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*: $MS(a, b) \ge 0$
- 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

$$\|\delta \mathbf{e}\|_{\mathbf{e}} = \sqrt{8MS(\mathbf{e}, \mathbf{e} + \delta \mathbf{e})}$$
$$g_{\mathbf{e}}(\delta \mathbf{e}_1, \delta \mathbf{e}_2) = \frac{1}{2} \left(\|\delta \mathbf{e}_1 + \delta \mathbf{e}_2\|_{\mathbf{e}}^2 - \|\delta \mathbf{e}_1\|_{\mathbf{e}}^2 - \|\delta \mathbf{e}_2\|_{\mathbf{e}}^2 \right)$$

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

Entropy strict concavity means the Hessian is negative definite

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



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

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

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

Another direct calculation from the definitions



Measure theoretic structures from entropy



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 $p = \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}$.



Proposition 1.153 (Exponential entropy subadditivity). Let $e_1, e_2 \in \mathcal{E}$. Let $S_1 = S(e_1)$ and $S_2 = S(e_2)$. Let $e = pe_1 + \bar{p}e_2$ for some $p \in [0,1]$ and S = S(e). Then $2^S \le 2^{S_1} + 2^{S_2}$, with the equality if and only if e_1 and e_2 are orthogonal and $p = \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}$.

Then

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

$$0 = \frac{d}{dp}S(pa + \bar{p}b) = \frac{d}{dp}(-p\log p - \bar{p}\log \bar{p} + pS_a + \bar{p}S_b)$$

= $-\log p - 1 + \log \bar{p} + 1 + S_a - S_b$
 $\log \frac{p}{\bar{p}} = \log 2^{S_a} - \log 2^{S_b}$
 $\log \frac{p}{1-p} = \log \frac{2^{S_a}}{2^{S_b}}$
 $p2^{S_b} = (1-p)2^{S_a}$
 $p(2^{S_a} + 2^{S_b}) = 2^{S_a}$
 $p = \frac{2^{S_a}}{2^{S_a} + 2^{S_b}}$

$$\begin{split} \bar{p} &= 1 - \frac{2^{S_a}}{2^{S_a} + 2^{S_b}} = \frac{2^{S_b}}{2^{S_a} + 2^{S_b}} \\ S(pa + \bar{p}b) &= -p \log p - \bar{p} \log \bar{p} + pS_a + \bar{p}S_b \\ &= -\frac{2^{S_a}}{2^{S_a} + 2^{S_b}} \log \frac{2^{S_a}}{2^{S_a} + 2^{S_b}} - \frac{2^{S_b}}{2^{S_a} + 2^{S_b}} \log \frac{2^{S_b}}{2^{S_a} + 2^{S_b}} \\ \text{Then calculate} &+ \frac{2^{S_a}}{2^{S_a} + 2^{S_b}} \log 2^{S_a} + \frac{2^{S_b}}{2^{S_a} + 2^{S_b}} \log 2^{S_b} \\ &= \frac{2^{S_a}}{2^{S_a} + 2^{S_b}} \log \left(2^{S_a} + 2^{S_b}\right) + \frac{2^{S_b}}{2^{S_a} + 2^{S_b}} \log \left(2^{S_a} + 2^{S_b}\right) \\ &= \frac{2^{S_a} + 2^{S_b}}{2^{S_a} + 2^{S_b}} \log \left(2^{S_a} + 2^{S_b}\right) \\ &= \frac{2^{S(pa+\bar{p}b)}}{2^{S(pa+\bar{p}b)}} = \log \left(2^{S_a} + 2^{S_b}\right) \\ &\geq 2^{S(pa+\bar{p}b)} = 2^{S_a} + 2^{S_b} \end{split}$$

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



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

of a non-additive measure

capacity also name

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

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

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?

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, $\operatorname{frac}_{e}(a) = \sup(\{p \in [0,1] | \exists b \in \mathcal{E} \text{ s.t. } e = pa + \overline{p}b\}).$

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

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

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

Fraction capacity is a non-additive probability measure

Assumptions Physics

https://assumptionsofphysics.org/

biggest p

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

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 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 Physics

