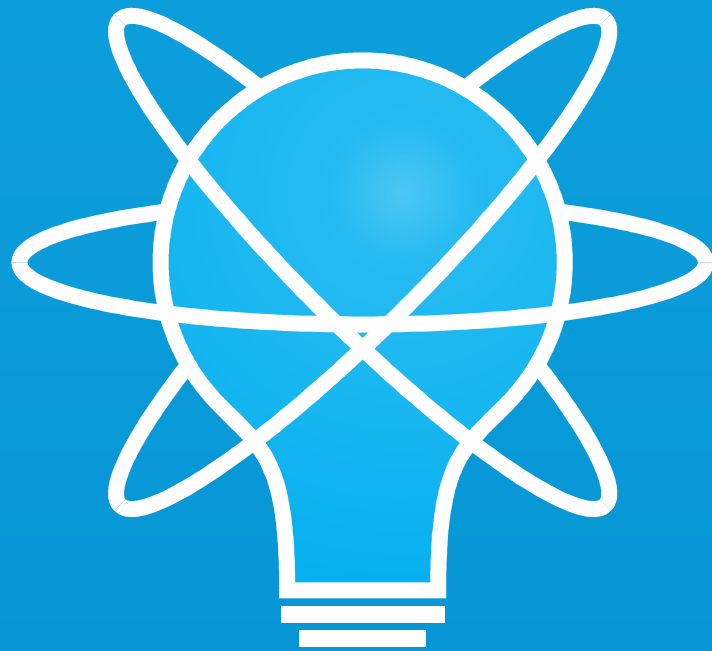


Assumptions of Physics



Gabriele Carcassi
Christine A. Aidala

Gabriele Carcassi, Christine A. Aidala

Assumptions of physics

Ver 2.0 - October 1, 2023

This edition was finalized on October 1, 2023. Older and newer versions can be found at <https://assumptionsofphysics.org/book>.

Copyright © 2018-23 Gabriele Carcassi, Christine A. Aidala

ASSUMPTIONSOFPHYSICS.ORG/BOOK

Licensed under the Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) (the “License”). You may not use this file except in compliance with the License. You may obtain a copy of the License at <https://creativecommons.org/licenses/by-nc-sa/4.0/>. This work is distributed under the License “as is”, without warranties or conditions of any kind, either express or implied.

Preface

This work is part of a larger project, Assumptions of Physics (<https://assumptionsofphysics.org>), that aims to identify a handful of physical principles from which the basic laws can be rigorously derived. The goal is to give physics (and science more in general) a renewed foundation that is mathematically precise, physically meaningful and philosophically consistent. Given the ambition and broad scope of the task, nothing would ever be written if one were to wait for the complete picture. Therefore this work contains only the parts of the project that are considered to be mature, and so it will be revised and expanded as progress is made. We give here a brief overview of the project, which can be useful to better understand the context of this work.

Overall goals of Assumptions of Physics

What do the basic laws of physics describe? Why is the state of a classical particle identified by position and momentum (i.e. a point on a symplectic manifold) while the state of a quantum system is identified by a wave function (i.e. a vector in a Hilbert space)? What assumptions are unique to each theory? What are, instead, the basic requirements that all physical theories must satisfy? Could we have had different laws? A lack of clear answers to these questions is, we believe, the biggest obstacle in the foundations of physics and prevents the resolution of outstanding problems in the field. **Our approach is to find a minimal set of physical assumptions that are necessary and sufficient to derive the different theories within a unified framework.** If we are able to do so, then we are guaranteed that all that the laws of physics say is encoded in those assumptions and we are able to answer those questions.

We found this approach to be very fruitful. It provides new insights into physics as a whole, the role of mathematics in physical theories and a gives a more solid conceptual foundation to both. It becomes clear why some mathematical structures are pervasive in science and what exactly they are meant to represent, while others will never play a role. The downside is that we have to touch many subjects in math (logic, topological spaces, measure theory, group theory, vector spaces, differential geometry, symplectic geometry, statistics, probability theory, ...), physics (Hamiltonian mechanics, Lagrangian mechanics, thermodynamics, quantum mechanics, electromagnetism, ...) and science in general (computer science theory, information theory, system theory, ...). In other words, **the only way to properly achieve the goal is to rebuild everything from the ground up:** formal rigor, physical significance and conceptual integrity are not something that can be added at the end, but they must be present from the beginning.

The main takeaway for us is that the foundations of science are one: no real progress can be made on the foundations of one subject without making progress on the foundations of

others. **What is needed is a general theory of experimental science: a theory that studies physical theories.** This provides a standard framework that defines basic concepts and requirements (e.g. experimental verifiability, granularity of descriptions, processes and states) that serve as a common basis for all theories. Each theory, then, is recovered by studying how these common objects become specialized under different assumptions. This book aims to build over time, piece by piece, this framework.

While the topic is necessarily inter-disciplinary, **this is still first and foremost a scientific book.** The material should be accessible to the mathematician and philosopher, but understand that it needs to resonate first and foremost with the experimental physicist and the engineer. The mathematical definitions and derivations are there to make the science precise, but they are not the main focus. In fact, the book is designed so that the mathematical definitions and proofs, highlighted with a green side bar, can be skipped altogether without loss of the big picture and the important details. Along the same line, the foundational discussions are there to articulate more precisely what it means to do science, so they will not indulge in other questions which may be of interest to the philosopher but not to the scientist.

A living work

For the project to be successful, we need to depart from some of the norms of academic research and academic publishing. For example, one typically develops his research program as a series of articles published in a peer reviewed journal that caters to a specific community. These articles are then typically collected as is or merged into a book. This does not work for this project. As journals are specialized into sometimes very narrow fields, this would create a set of disjoint articles that cater to different audiences, with no guarantee that they can fit into a unified vision. For us, the overall picture, if and how the different perspectives combine, is the most important feature. **In this sense, the book comes first and the articles are derivative works.** We need to pool expertise and ideas from a wide range of disciplines and make sure that the result makes sense from all angles.

As the goal is broad, the framework needs to evolve as new issues are solved and old ones are better understood. If one part changes, we have to make sure that everything is updated to keep conceptual consistency. **This book, therefore, is an ongoing project.** It will continue to grow organically, adding and revising chapters. As is standard practice in open source/free software communities, we need to “release early, release often” to gather feedback. Each new version supersedes all prior ones and will be superseded by future ones. There is therefore no “definitive version” in the near future as we don’t expect to “solve all of physics” in the near future. However, the framework will tend to converge as different parts become more settled.

The upshot is that one only needs to read the latest version of this work to be current. That is, one does not need to read a scattered set of papers, which require previous knowledge of a field, and follow how the ideas have changed. Just get the latest copy of the book, and if you do find areas you can help us expand or improve, let us know!

Project overview

Here we present a summary of the whole project and the status of each part as the layout will map to the structure of this work. We divide the work based on the two main techniques we use. The first, **reverse physics**, aims to identify the fundamental ideas and assumptions by reverse

engineering them from the current physical theories. The second, **physical mathematics**, aims to construct a rigorous mathematical framework from the ground up, based on the ideas and assumptions found by reverse physics.

Reverse physics

Reverse physics looks at the main physical theories, like classical mechanics, thermodynamics and quantum mechanics, to identify concepts that can be used to fully explain the common and different aspects of those theories. The idea is to find physically meaningful assumptions that can be shown to be equivalent to the physical laws.

The standard of rigor in this part is necessarily more relaxed as we do not have a guarantee that sufficiently mature mathematical tools exist to carry out the argument in a precise fashion. For example, we have found that the idea of a unit system is linked in a fundamental way to the notion of state spaces, yet we lack a fully developed mathematical framework to model units and their dependency. The goal is to test the ideas conceptually, find those that are broad enough and necessary enough to then justify investing further time in a more rigorous approach.

The following are examples of the type of assumptions we have found to be good starting points to rederive the different theories.

Determinism and reversibility: “The system undergoes deterministic and reversible evolution.” Mathematically, the physical properties of the system determine which category, in the mathematical sense, is used to describe the state space, and deterministic and reversible evolution will be an isomorphism in the category (i.e. a bijective map that preserves the physical properties of the system). Therefore the law of evolution is not just a bijective map, but is also a linear transformation, a differentiable map or an isometry depending on the context.

Infinitesimal reducibility: “Specifying the state of the whole system is equivalent to specifying the state of all its infinitesimal parts.” For example, we can study the motion of a ball, but we can also mark a spot in red and study the motion of the mark. Knowing the evolution of the whole ball means knowing the evolution of any arbitrary spot and vice-versa. Mathematically, the state of the whole will be a distribution over the state space of the parts. It will need to be a distribution whose value is invariant under coordinate transformations. The state space of the infinitesimal parts, then, comes equipped with an invariant two-form upon which we can define such a distribution. The state space is therefore a symplectic manifold, that is, the states of the infinitesimal parts are described by pairs of conjugate variables, which recovers phase space. If the previous assumption holds, deterministic and reversible evolution is a symplectomorphism, that is, deterministic and reversible evolution follows classical Hamiltonian mechanics. Proper handling of the time variable will give us a relativistic version of the framework without extra assumptions.

Irreducibility: “Specifying the state of the whole system tells us nothing about its infinitesimal parts.” For example, we can study the state of an electron by scattering photons off of it. But whenever a photon interacts with the electron, it interacts with the whole electron. There is no way to mark a part of an electron and study it independently from the rest. Mathematically, the state of the electron will be a distribution that evolves deterministically where the motion of each infinitesimal part cannot be further described.

Kinematic equivalence: “Specifying the motion of the system is equivalent to specifying its state and evolution.” This means that we will have to be able to re-express a distribution over kinematic variables (i.e. position and velocity) into a distribution over state variables (i.e. position and momentum) and vice-versa. Mathematically, the symplectic two-form will induce a symmetric tensor over the tangent space for position. This will give us a metric and will also allow us to reformulate the laws of motion according to Lagrangian mechanics. Because the transformation is linear, we are able to constrain the Hamiltonian to the one for massive particles under scalar and vector potential forces.

Physical mathematics

The first step in physical mathematics is the development of a general mathematical theory of experimental science, or simply a general theory, that lays down the basic axioms and definitions that are required for any physical theory. **We need a “theory of scientific theories” that allows us to define and study the set of all well-formed scientific theories.** The core tenet is the principle of scientific objectivity: “Science is universal, non-contradictory and evidence based.” This serves to define the subject and the basic formal requirements of a scientific theory. We divide the general theory into three main parts.

Experimental verifiability. This part defines a scientific theory as a set of statements. It provides the basic axioms which simply characterize the rules of logic as applied to experimentally verifiable statements. Mathematically, these requirements are captured by topologies and σ -algebras over the space of the possible cases that can be identified experimentally. This part is very well developed both conceptually and mathematically, and it is the only part currently included in this book.

Informational granularity. This part adds the ability to compare the level of description of different statements. Mathematically, this imposes a preorder on the algebra of statements which is intended to provide a common foundation to geometry, measure theory, probability theory and information theory. All those structures, in fact, provide ways to compare the size of sets of possible cases and the description provided in terms of those sets. This part is fairly developed conceptually but not yet well developed mathematically.

States and processes. This part defines the basic physical notions of processes, systems and states. It serves to bring to light a series of hidden assumptions that are needed, both physically and mathematically, to be able to talk about independent systems. These will lead to notions of entropy, state spaces and the geometrical structures that accompany them. This part is still being developed conceptually, though the progress is encouraging.

Current plan and status

In this version we have added a reverse physics part with a chapter on classical mechanics. We plan to follow that with a chapter on thermodynamics and statistical mechanics. For physical mathematics, the work is slowed down by the fact that general mathematical tools for the informational granularity appear to be missing. In particular, if we want state counting in quantum mechanics to be consistent with statistical mechanics and information theory, a more general notion of measure is needed.

Changelog

- 2023/10/01: Ver 2.0 - Divided the work into two main parts: Reverse Physics and Physical mathematics. Added chapter on reversing classical mechanics. Minor updates on the logic section.
- 2021/03/08: Ver 1.0 - Updated the first three chapters with minor changes: renamed tautology to certainty and contradiction to impossibility as they characterize better their role in the framework; made more formal justifications for the basic axioms and some of the basic definitions; causal relationships are now proved to be continuous instead of assumed to be continuous. Added Part II to include the results that are not yet fully formalized, to give a sense of the future scope of the work.
- 2019/07/07: Ver 0.3 - Reviewed first two chapters to clarify the idea of possible assignments and how contexts for function spaces are constructed.
- 2019/02/22: Ver 0.2 - Consolidated third chapter on properties, quantities and ordering.
- 2018/06/22: Ver 0.1 - Consolidated first two chapters that lay the foundation for the general theory.

Contents

Preface	v
Project overview	vi
I Reverse Physics	1
1 Classical mechanics	5
1.1 Formulations of classical mechanics	6
1.2 Inequivalence of formulations	8
1.3 Kinematics vs dynamics	12
1.4 Reversing Hamiltonian mechanics	19
1.5 Multiple degrees of freedom	28
1.6 Reversing differential topology	36
1.7 Reversing Lagrangian mechanics	44
1.8 Full kinematic equivalence and massive particles	53
1.9 Relativistic mechanics	58
1.10 Reversing phase space	69
1.11 Reversing Newtonian mechanics	80
1.12 Directional degree of freedom	82
1.13 Infinitesimal reducibility	87
1.14 Summary	94
II Physical Mathematics	97
1 Verifiable statements and experimental domains	101
1.1 Statements	101
1.2 Verifiable statements and experimental domains	114
1.3 Theoretical domains and possibilities	124
1.4 Topological spaces	133
1.5 Sigma-algebras	139
1.6 Decidable domains	141
1.7 Summary	143
1.8 Reference sheet	146
2 Domain combination and relationships	149

2.1	Dependence and equivalence between domains	149
2.2	Combining domains	156
2.3	Experimental domain for experimental relationships	162
2.4	Summary	168
3	Properties and quantities	169
3.1	Properties	169
3.2	Quantities and ordering	171
3.3	References and experimental ordering	178
3.4	Discrete quantities	185
3.5	Arbitrary precision and continuous quantities	188
3.6	When ordering breaks down	193
3.7	Summary	194
	IIIBlueprints for the work ahead	197
1	Reverse Physics	201
1.1	Classical mechanics	201
1.2	Thermodynamics	201
1.3	Quantum mechanics and irreducibility	203
2	Physical mathematics	209
2.1	Experimental verifiability	209
2.2	Informational granularity	209
2.3	States and processes	213
2.4	Open questions and possible extensions	215
3	Ensemble spaces	219
3.1	Ensembles	219
3.2	Mixtures and convex spaces	224
3.3	Quantities	228
3.4	Biggest component and probability	229
3.5	Entropy	231
3.6	Effective count	234
3.7	Entropic geometry	236
3.8	Ensemble subspaces	240
3.9	Contexts and probability measure	244
3.10	Lessons learned	247
	IV Appendix	251
A	Reference sheets for math and physics	253
A.1	Set theory	253
	Credits	255

Part I

Reverse Physics

Reverse physics is an approach to the foundations of physics that analyzes known theories to identify those physical principles and assumptions that can be taken as their conceptual foundation. This is the analogue of **reverse mathematics**, a program in mathematical logic that seeks to determine which axioms are required to prove mathematical theorems.

While some physical theories, such as Newtonian mechanics, thermodynamics and special relativity, are indeed founded on laws or principles, many theories, such as Hamiltonian and Lagrangian mechanics, quantum mechanics, and general relativity, are based on mathematical relationships that are simply postulated without a strict physical justification. Most modern theories are of the latter type, as theoretical physicists have increasingly focused on mathematical ideas rather than physical starting points. Therefore we do not know why the state of a quantum system can be represented by a ray in a Hilbert space, or what exact relationship is represented by the Einstein field equations. The goal of reverse physics, then, is to find suitable physical premises that can function as a more proper foundation for each theory. We stress that the premises have to be of a physical nature, such as “the system under study is isolated” or “the quantity is additive under system composition”. That is, they must be principles or assumptions that express some physical idea, not some abstract mathematical notion.

Another issue in modern physics is that it is currently a patchwork of different theories, and one is trained to match patterns and examples to decide which problems (or what aspects of a larger problem) should be treated with a particular theory. The goal here is to fully understand what physical situation each mathematical structure can suitably describe, what connections can be made across different theories and what exactly are the true limits of applicability of the different physical theories. One of the main results of reverse physics is that there are some core ideas that are shared by all physical theories, and that, in fact, physical theories are not as disconnected from each other as generally thought.

While this work explores connections between different mathematical and physical disciplines, it would be impractical to provide even a concise introduction to all. Therefore when encountering concepts like information entropy, Minkowski space-time, function spaces, quantum observables, Lagrange multipliers, intensive quantities, and so on, we can only provide the definition with very limited context, as most of the space must be dedicated to the connections between these concepts. It is up to the reader to decide whether to invest additional time on other sources, or to simply make a mental note and proceed further.

By nature, reverse physics does not aim to be mathematically precise, but rather conceptually precise. While the goal is to create a dictionary between physical concepts and their mathematical representation, this mapping cannot be absolutely perfect for a simple reason: we have no guarantee that the current mathematical structures are the correct ones to capture the physical ideas. Conversely, we cannot give a precise mathematical characterization of the physical ideas if these are not conceptually clear. Therefore conceptual clarity has to come before formal precision. Yet, it is true that to reach full conceptual clarity the ideas have to be sufficiently refined to allow a formal definition. This task is the purview of **physical mathematics**, which aims to find physically meaningful mathematical structures starting from definitions and axioms that can be fully supported by physical requirements and assumptions. Yet, it would be premature to engage in such activity without having thoroughly tested the conceptual framework beforehand.

Chapter 1

Classical mechanics

The standard view in physics is that classical mechanics is perfectly understood. It has three different but equivalent formulations, the oldest of which, Newtonian mechanics, is based on three laws. Classical mechanics is the theory of point particles that follow those laws. Unfortunately, this view is incorrect.

We will see that the three formulations are not equivalent, in the sense that there are physical systems that are Newtonian but not Hamiltonian and vice-versa. There are also a number of questions that have been left unanswered, such as the precise nature of the Hamiltonian or the Lagrangian, and what exactly the principle of stationary action represents physically. While shedding light on these issues, we will also find that classical mechanics already contains elements that are typically associated with other theories, such as quantum mechanics/field theories (uncertainty principle, anti-particles), thermodynamics/statistical mechanics (thermodynamic and information entropy conservation) or special relativity (energy as the time component of a four-vector). In other words, the common understanding of classical mechanics is quite shallow, and its foundations are, in fact, not separate from the ones of classical statistical mechanics or special relativity.

What reverse physics shows is that the central assumption underneath classical mechanics is that of **infinitesimal reducibility (IR)**: a classical system can be thought of as made of parts, which in turn are made of parts and so on; studying the whole system is equivalent to studying all its infinitesimal parts. This assumption, together with the assumption of **independence of degrees of freedom (IND)**, is what gives us the structure of classical phase space with conjugate variables. The additional assumption of **determinism and reversibility (DR)**, the fact that the description of the system at one time is enough to predict its future or reconstruct its past, leads us to Hamiltonian mechanics. On the other hand, assuming **kinematic equivalence (KE)**, the idea that trajectories in space are enough to reconstruct the state of the system and vice-versa, leads to Newtonian mechanics. The combination of all above assumptions, instead, leads to Lagrangian mechanics and, in particular, to massive particles under (scalar and vector) potential forces.

As a guide to the chapter, here is the list of main points in the order in which they will be presented, one for each section.

1. Review of classical formulations
2. Lagrangian mechanics is Hamiltonian mechanics and KE
3. Kinematics, in general, is not enough to reconstruct dynamics

4. Hamiltonian mechanics (one DOF) is equivalent to DR
5. Hamiltonian mechanics (multiple DOFs) is equivalent to DR plus IND
6. Differential calculus and its generalization, differential topology, study infinitesimally additive quantities that depend on geometric shapes (i.e. lines, surfaces, volumes)
7. The principle of least action is a consequence of DR, IND and KE
8. Massive particles under potential forces are a consequence of DR, IND and KE
9. Special relativity is a consequence of DR, IND and KE
10. Phase space is the only structure that makes distributions, state counting and entropy frame invariant
11. Newtonian mechanics is a consequence of KE
12. Three dimensional spaces are the only spaces for which distributions over directions are frame invariant
13. Classical particle states as points in phase space are equivalent to IR

1.1 Formulations of classical mechanics

In this section we will briefly review the three main formulations of classical mechanics. Our task is not to present them in detail, but rather to provide a brief summary of the equations so that we can proceed with the comparison. In particular, given that different conventions are used across formulations, within the same formulation and among different contexts (e.g. relativity, symplectic geometry), we will want to make the notation homogeneous to allow easier comparisons.

Newtonian mechanics

For all formulations, the system is modeled as a collection of point particles, though we will mostly focus on the single particle case. For a Newtonian system, the state of the system at a particular time t is described by the position x^i and velocity v^i of all its constituents. Each particle has its mass m , not necessarily constant in time, and, for each particle, we define kinetic momentum as $\Pi^i = mv^i$.¹

The evolution of our system is given by Newton's second law:²

$$F^i(x^j, v^k, t) = d_t \Pi^i. \quad (1.1)$$

Mathematically, if the forces F^i are locally Lipschitz³ continuous, then the solution $x^i(t)$ is

¹We will use the letter t for the time variable, x for position and v for velocity, which is a very common notation in Newtonian mechanics. However, we will keep using the same letters in Lagrangian mechanics as well, instead of q and \dot{q} , for consistency. Given that the distinction between kinetic and conjugate momentum is an important one, we will denote Π the former and p the latter. The Roman letters i, j, k, \dots will be used to span the spatial components (e.g. $i \in \{1, 2, 3\}$ for a particle in 3 dimensional space and $i \in \{1, 2, \dots, 3n\}$ for n particles), while we will use the Greek letters $\alpha, \beta, \gamma, \dots$ to span space-time components (e.g. $\alpha \in \{0, 1, 2, 3\}$ where the 0 value of the index is used for time). Unlike some texts, x^i do not represent Cartesian coordinates, and therefore they should be understood already as generalized coordinates.

²For derivatives, we will use the shorthand d_t for $\frac{d}{dt}$ and ∂_{x^i} for $\frac{\partial}{\partial x^i}$. For functions that depend on multiple arguments we use a free index to note that it depends on all elements; each argument will have a different index to highlight that there is no relationship between arguments.

³Lipschitz continuity means that the slope of the function is bounded. For example, \sqrt{x} in the neighborhood of 0 is not Lipschitz continuous as it has a vertical asymptote at that point. One can construct examples (e.g. Norton's dome) where the forces are not locally Lipschitz continuous, and therefore the initial position and

unique. That is, given position and velocity at a given time, we can predict the position and velocity at future times. We will assume a Newtonian system has this property.

An important aspect of Newtonian mechanics is that the equations are not invariant under coordinate transformation. To distinguish between apparent forces (i.e. those dependent on the choice of frame) and the real ones, we assume the existence of inertial frames. In an inertial frame there are no apparent forces, and therefore a free system (i.e. no forces) with constant mass proceeds in a linear uniform motion, or stays still.⁴

Lagrangian mechanics

The state for a Lagrangian system is also given by position x^i and velocity v^i . The dynamics is specified by a single function $L(x^i, v^j, t)$ called the Lagrangian. For each spatial trajectory $x^i(t)$ we define the action as $\mathcal{A}[x^i(t)] = \int_{t_0}^{t_1} L(x^i(t), d_t x^i(t), t) dt$. The trajectory taken by the system is the one that makes the action stationary:

$$\delta \mathcal{A}[x^i(t)] = \delta \int_{t_0}^{t_1} L(x^i(t), d_t x^i(t), t) dt = 0 \quad (1.2)$$

The evolution can equivalently be specified by the Euler-Lagrange equations:

$$\partial_{x^i} L = d_t \partial_{v^i} L. \quad (1.3)$$

Note that not all Lagrangians lead to a unique solution. For example, $L = 0$ will give the same action for all trajectories and therefore, strictly speaking, all trajectories are possible. The stationary action leads to a unique solution if and only if the Lagrangian is hyperregular, which means the Hessian matrix $\partial_{v^i} \partial_{v^j} L$ is invertible. Like in the Newtonian case, we will assume Lagrangian systems satisfy this property.

Unlike Newton's second law, both the Lagrangian and the Euler-Lagrange equations are invariant under coordinate transformations. This means that Lagrangian mechanics is particularly suited to study the symmetries of the system.

Hamiltonian mechanics

In Hamiltonian mechanics, the state of the system is given by position q^i and conjugate momentum p_i . The dynamics is specified by a single function $H(q^i, p_j, t)$ called the Hamiltonian.⁵ The evolution is given by Hamilton's equations:

$$\begin{aligned} d_t q^i &= \partial_{p_i} H \\ d_t p_i &= -\partial_{q^i} H \end{aligned} \quad (1.4)$$

velocity do not yield a unique solution (i.e. in Norton's dome, the body can stay on the top of the dome indefinitely, or it can fall down after an arbitrary amount of time). In this case, something else, outside the system, will necessarily determine what is the motion of the system, and therefore it is not true that the force and the state of the system fully determine the dynamics of the system.

⁴Recall that linear motion simply means that it describes a line in space, while uniform motion means that the speed is constant. Therefore we can have linear non-uniform motion (e.g. an object accelerated along the same direction) or a non-linear uniform motion (e.g. an object going around in a circle at constant speed).

⁵We use a different symbol for position in Hamiltonian mechanics because, while it is true that $q^i = x^i$, it is also true that $\partial_{q^i} \neq \partial_{x^i}$: the first derivative is taken at constant conjugate momentum while the second is taken at constant velocity. This creates absolute confusion when mixing and comparing Lagrangian and Hamiltonian concepts, which our notation avoids completely.

We will again want these equations to yield a unique solution, which means the Hamiltonian must be at least differentiable, and the derivatives must at least be Lipschitz continuous.

Hamilton's equations are invariant as well. The Hamiltonian itself is a scalar function which is often considered (mistakenly as we'll see later) invariant. This formulation is the most suitable for statistical mechanics as volumes of phase space correctly count the number of possible configurations.

1.2 Inequivalence of formulations

It is often stated in physics books that all three formulations of classical mechanics are equivalent. We will look at this claim in detail, and conclude that this is not the case: there are systems that can be described by one formulation and not another. More precisely, the set of Lagrangian systems is exactly the intersection of Newtonian and Hamiltonian systems.

Testing equivalence

We will consider two formalisms equivalent if they can be applied to exactly the same systems. That is, Newtonian and Lagrangian mechanics are equivalent if any system that can be described using Newtonian mechanics can also be described by Lagrangian mechanics and vice-versa. In general, in physics great emphasis is put on systems that can indeed be studied by all three, leaving the impression that this is always doable.⁶ However, just with a cursory glance, we realize that this can't possibly be the case.

The dynamics of a Newtonian system, in fact, is specified by three independently chosen functions of position and velocity, the forces applied to each degree of freedom (DOF). On the other hand, the dynamics of Lagrangian and Hamiltonian systems is specified by a single function of position and velocity/momentum, the Lagrangian/Hamiltonian. Intuitively, there are more choices in the dynamics for Newtonian systems than for Lagrangian and Hamiltonian.

Now, the reality is a bit trickier because the mathematical expression of the forces is not enough to fully characterize the physical system. We need to know in which frame we are, what coordinates are being used and the mass of the system, which is potentially a function of time. On the Lagrangian side, note that the Euler-Lagrange equations are homogeneous in L . This means that multiplying L by a constant leads to the same solutions, meaning that the same system can be described by more than one Lagrangian. The converse is also true: if one system is half as massive and is subjected to a force half as intense, the resulting Lagrangian is also simply rescaled by a constant factor. Therefore the map between Lagrangians and Lagrangian systems is not one-to-one: it is many-to-many. This is why we should never look simply at mathematical structures if we want to fully understand the physics they describe.

Regardless, our task is at the moment much simpler: we only need to show that there are Newtonian systems not expressible by Lagrangian or Hamiltonian mechanics. We can therefore limit ourselves to systems with a specific constant mass m in an inertial frame and write $a^i = F^i(x^j, v^k, t)/m$. Given that the force is arbitrary, the acceleration can be an arbitrary function of position, velocity and time. Similarly, we can write the acceleration of a

⁶If one asks the average physicist whether Newtonian and Hamiltonian mechanics are equivalent, the answer most of the time will be enthusiastically positive. If one then asks for the Hamiltonian for a damped harmonic oscillator, the typical reaction is annoyance due to the nonsensical question (damped harmonic oscillators do not conserve energy), followed by a realization and partial retraction of the previous claim. The moral of the story is to never take these claims at face value.

Lagrangian system as $a^i = F^i[L]/m$. That is, the acceleration is going to be some functional of the Lagrangian. Given the Euler-Lagrange equations 1.3, the map between the Lagrangian and the acceleration must be continuous in both direction: for small variation of the Lagrangian we must have a small variation of the equations of motion and therefore of the acceleration, and for small variations of the equations of motions we must have a small variation of the Lagrangian. But a continuous surjective map from the space of a single function (i.e. the Lagrangian) to the space of multiple functions (i.e. those that specify the acceleration in terms of position and velocity) does not exist,⁷ and therefore there must be at least one Newtonian system with constant mass expressed in an inertial frame that is not describable using Lagrangian mechanics. The same argument applies for Hamiltonian mechanics, since the dynamics in this case is also described by a single function in the same number of arguments. We therefore reach the following conclusion:

Insight 1.5. *Not all Newtonian systems are Lagrangian and/or Hamiltonian.*

Newtonian vs Lagrangian/Hamiltonian

We now want to understand whether all Lagrangian systems are Newtonian. Given what we discussed, we cannot expect to reconstruct the mass and force uniquely from the expression of the Lagrangian. We consider the mass and the frame fixed by the problem, together with the Lagrangian, and therefore we must only see whether we can indeed find a unique expression for the acceleration. From the Euler-Lagrange equations 1.3 we can write

$$\begin{aligned}\partial_{x^i} L &= d_t \partial_{v^i} L = \partial_{x^j} \partial_{v^i} L d_t x^j + \partial_{v^k} \partial_{v^i} L d_t v^k = \partial_{x^j} \partial_{v^i} L v^j + \partial_{v^k} \partial_{v^i} L a^k \\ \partial_{v^k} \partial_{v^i} L a^k &= \partial_{x^i} L - \partial_{x^j} \partial_{v^i} L v^j.\end{aligned}\tag{1.6}$$

To be able to write the acceleration explicitly, we must be able to invert the Hessian matrix $\partial_{v^k} \partial_{v^i} L$. As we noted before, this is exactly the condition for which the principle of stationary action leads to a unique solution, and we can better understand why. If it is not invertible at a point, the determinant is zero and therefore one eigenvalue is zero. The corresponding eigenvector corresponds to a direction for which the equation tells us nothing, and therefore a variation of the acceleration in that direction will not change the action. This is why the invertibility of the Hessian is required in order to obtain unique solutions.

What we find, then, is that for any Lagrangian system, which we assume to have a unique solution, we can explicitly write the acceleration as a function of position, velocity and time. Therefore

Insight 1.7. *All Lagrangian systems are Newtonian.*

Now we turn our attention to Hamiltonian mechanics and, similarly, we ask whether we can express the acceleration as a function of the state. We have

$$\begin{aligned}a^i &= d_t v^i = d_t d_t q^i = d_t \partial_{p_i} H = \partial_{q^j} \partial_{p_i} H d_t q^j + \partial_{p_k} \partial_{p_i} H d_t p_k \\ &= \partial_{q^j} \partial_{p_i} H \partial_{p_j} H - \partial_{p_k} \partial_{p_i} H \partial_{q^k} H.\end{aligned}\tag{1.8}$$

⁷Mathematically, the space of continuous functions $C(\mathbb{R}, \mathbb{R})$ and $C(\mathbb{R}^n, \mathbb{R})$ are not homeomorphic. Intuitively, the underlying reason is the same as to why a map from a volume to a line can't be continuous: in a volume you have infinitely many directions you can move away from a point, while on a line you only have two.

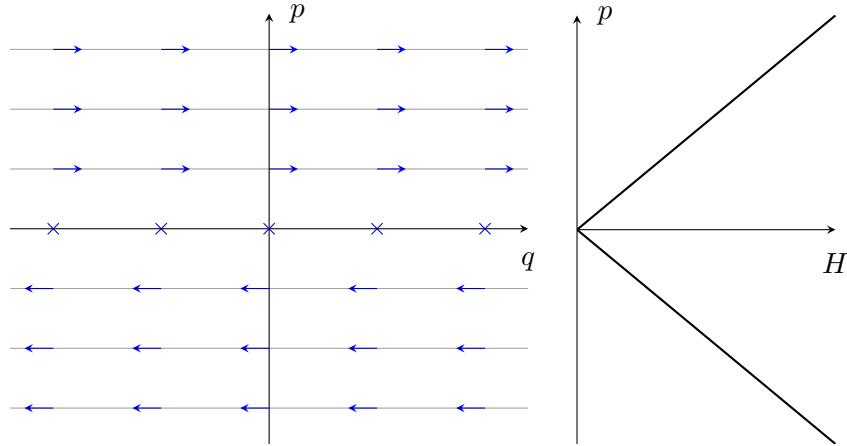


Figure 1.1: On the left, the phase-space diagram for a photon treated as a point particle. The Hamiltonian $H = c|p|$, on the right, is proportional to the modulus of p . Since H is not differentiable when $p = 0$, those states are excluded, consistent with the physics. The displacement field has only a q component, which is $+c$ above the horizontal axis and $-c$ below the horizontal axis.

This tells us that we can always write an explicit function for the acceleration. However, this is not enough. States in Newtonian mechanics are in terms of position and velocity, not position and momentum. For a Hamiltonian system to be equivalent to a Newtonian system we need to be able to write the momentum as a function of position and velocity and vice versa. Note that Hamilton's equations already give a way to express the velocity in terms of position and momentum. We just need that expression to be invertible, which means the Jacobian must be invertible. We must have:

$$|\partial_{p_i} v^j| = |\partial_{p_i} \partial_{p_j} H| \neq 0. \quad (1.9)$$

To be able to express momentum as a function of position and velocity, then, we need the Hessian of the Hamiltonian to be invertible (i.e. to have non-zero determinant).

Note that we had no such requirement for the Hamiltonian itself. For example, $H = 0$ leads to equations $d_t q^i = 0$ and $d_t p_i = 0$, which have unique solutions: both position $q^i(t) = k_{q^i}$ and momentum $p_i(t) = k_{p_i}$ are constants of motion. The Hessian, being the zero matrix, is not invertible, and in fact we cannot write momentum as a function of position and velocity: velocity $d_t q^i$ is always zero in all cases while conjugate momentum can be any value k_{p_i} . Though this case may not be physically interesting, it is a perfectly valid Hamiltonian system and shows that we should always check the trivial mathematical case. However, let us go through a more physically meaningful case.

Photon as a particle. If we want to treat the photon as a classical particle, we can write the Hamiltonian by expressing the energy as a function of momentum

$$H = \hbar|\omega| = c\hbar|k_i| = c|p_i|. \quad (1.10)$$

If we apply Hamilton's equations, we have

$$\begin{aligned} d_t q^i &= c \frac{p_i}{|p_i|} \\ d_t p_i &= 0. \end{aligned} \tag{1.11}$$

That is, the norm of the velocity is always c , the momentum decides its direction, and the momentum itself does not change in time, as shown in fig. 1.1. This is indeed the motion of a free photon. One can confirm, through tedious calculation, that the determinant of the Hessian is indeed zero, yet it is easier and more physically instructive to see that we cannot reconstruct the momentum from the velocity. Relativistically, all photons travel along the geodesics at the same speed, therefore two photons that differ only by the magnitude of the momentum will travel the same path.

Hamiltonian systems that are also Newtonian, then, need to satisfy this extra condition, so let us give it a name.

Assumption KE (Kinematic Equivalence). *The kinematics of the system is sufficient to reconstruct its dynamics and vice-versa. That is, specifying the motion of the system is equivalent to specifying its state and evolution.*

By kinematics we mean the motion in space and time and by dynamics we mean the state and its time evolution in phase space. We will need to analyze the difference between the two more in detail, but we should first finish our comparison between the different formulations.

Summing up, we find that

Insight 1.12. *Not all Hamiltonian systems are Newtonian: only those for which KE is valid.*

Lagrangian vs Hamiltonian

We now need to compare Lagrangian and Hamiltonian systems. The task is a lot easier because we already have a precise way to connect the two. If we are given a Lagrangian L , we define the conjugate momentum $p_i = \partial_{v^i} L$ and the Hamiltonian $H = p_i v^i - L$. If we are given a Hamiltonian H , we can define a Lagrangian $L = p_i v^i - H$ and a velocity $v^i = d_t q^i = \partial_{p_i} H$. The only detail that needs to be understood is whether this can be done for all Lagrangian and Hamiltonian systems.

While these expressions are always defined, we need to check whether we can change variables; whether we can write the Lagrangian in terms of position and velocity and the Hamiltonian in terms of position and momentum. Going from a Hamiltonian to a Lagrangian, it again means that we can write momentum as a function of position and velocity, and therefore assumption KE must hold. This makes sense: if all Lagrangian systems are Newtonian, and KE was required for a Hamiltonian system to be Newtonian, then it is also required for a Hamiltonian system to be Lagrangian. But the connection is stronger: KE is the *only* additional assumption we need to be able to write a Lagrangian given a Hamiltonian.

Going from a Lagrangian to a Hamiltonian, it means that we can write velocity as a function of position and momentum. Note that since we define conjugate momentum as the derivative of the Lagrangian, we can already express momentum as a function of position and velocity, which means we are simply asking that expression to be invertible. This is, again, assumption KE, just in the opposite direction. We must have

$$0 \neq |\partial_{v^i} p_j| = |\partial_{v^i} \partial_{v^j} L|. \tag{1.13}$$

This means that assumption **KE** is exactly the invertibility of the Hessian, the condition for unique solution of the Lagrangian. All Lagrangian systems that admit unique solutions, then, satisfy assumption **KE**. In fact, we can see that the Hessian determinants are related

$$|\partial_{v^i} \partial_{v^j} L| = |\partial_{v^i} p_j| = |\partial_{p_i} v^j|^{-1} = |\partial_{p_i} \partial_{p_j} H|^{-1}. \quad (1.14)$$

This means that every Lagrangian admits a Hamiltonian, but not every Hamiltonian admits a Lagrangian. Only the Hamiltonian systems for which **KE** is valid will also be Lagrangian systems, with a guaranteed unique solution given that **KE** is exactly the assumption needed for that as well. Therefore we conclude that

Insight 1.15. *Lagrangian systems are exactly those Hamiltonian systems for which **KE** is valid.*

Relationship between formulations

The relationship between the different formulations, then, can be summarized with the Venn diagram in fig. 1.2.

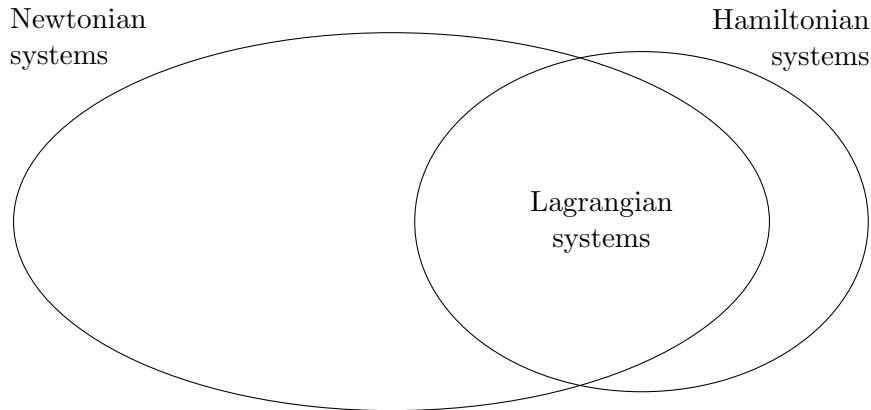


Figure 1.2: Not all Hamiltonian systems are Newtonian and not all Newtonian systems are Hamiltonian. All Lagrangian systems are both Newtonian and Hamiltonian.

We have found that **KE** is a constitutive assumption of Lagrangian mechanics, and that it clearly marks which Hamiltonian systems are Newtonian/Lagrangian. By constitutive assumption we mean an assumption that must be taken, either explicitly or implicitly, for a theory to be valid. But what makes a system Hamiltonian and what makes a system Newtonian? Can we find a full set of constitutive assumptions for classical mechanics?

1.3 Kinematics vs dynamics

We have seen the importance of the connection between kinematics and dynamics. In this section we will explore this link more deeply and come to the following conclusion: the kinematics of a system is not enough to reconstruct its dynamics.

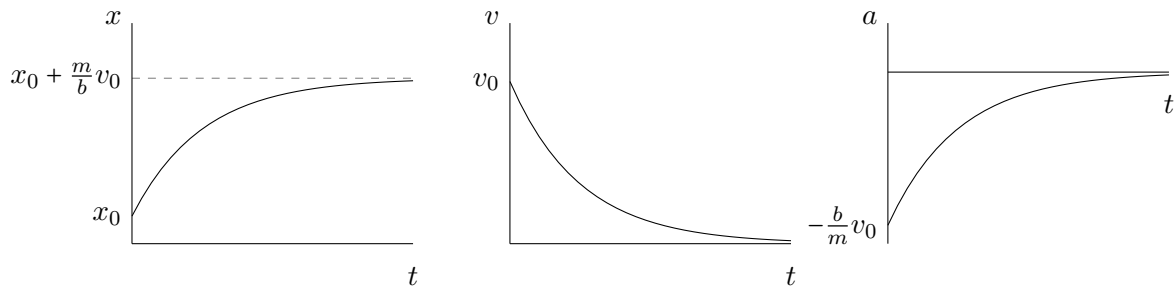


Figure 1.3: Evolution in time of position, velocity and acceleration for $ma = -bv$. Both acceleration and velocity will tend to zero as time increases. The position will tend to an equilibrium given by initial position and initial velocity.

Particle under linear drag

Let us first review exactly what the kinematics and dynamics are. Given a system, its kinematics is the description of its motion in space and time. Position, velocity, and acceleration are kinematic variables because they describe the motion. Kinematics is what Galileo studied and started to give a rigorous account of. The dynamics, instead, describes the cause of such motion. Force, mass, momentum, energy are dynamic quantities as they are used to describe why a body moves in a particular way. Dynamics is what Newton introduced and his second law, expressed as $F = ma$, clearly shows the link.

The link between the two concepts seems important given the constitutive role of **KE** in Lagrangian mechanics. Moreover, while both Newtonian and Hamiltonian mechanics are dynamical theories, in the sense that quantities like force and momentum are intrinsic parts of the respective theories, Lagrangian mechanics seems to be a purely kinematic theory, as it is described only by kinematic variables like position and velocity. Therefore it seems useful to characterize the kinematics-dynamics link as much as possible. Let's analyze a concrete example.

Suppose we are given the following equation:

$$ma = -bv. \quad (1.16)$$

The equation is in terms of kinematic variables and, given initial conditions x_0 and v_0 , it admits a unique solution, a unique trajectory. The solution, plotted in fig. 1.3, is

$$\begin{aligned} x(t) &= x_0 + v_0 \frac{m}{b} \left(1 - e^{-\frac{b}{m}t}\right) \\ v(t) &= v_0 e^{-\frac{b}{m}t} \\ a(t) &= -v_0 \frac{b}{m} e^{-\frac{b}{m}t} \end{aligned} \quad (1.17)$$

Can we reconstruct the forces acting on this system?

The obvious answer seems to be that the constant m represents the mass of the system and $F = -bv$ the force. This is the case of a particle under linear drag: the system is subjected to a frictional force that is proportional and opposite to the velocity. If we set the Lagrangian

$$L = \frac{1}{2}mv^2 e^{\frac{b}{m}t}. \quad (1.18)$$

and apply the Euler-Lagrange equation 1.3 we have

$$\begin{aligned}\partial_x L = 0 = d_t \partial_v L &= d_t \left(m v e^{\frac{b}{m}t} \right) = m a e^{\frac{b}{m}t} + \frac{b}{m} m v e^{\frac{b}{m}t} = e^{\frac{b}{m}t} (m a + b v) \\ m a &= -b v.\end{aligned}\tag{1.19}$$

Therefore we have a Lagrangian for the system. We can also find a Hamiltonian

$$\begin{aligned}p &= \partial_v L = m v e^{\frac{b}{m}t} \\ v &= \frac{p}{m} e^{-\frac{b}{m}t} \\ H = p v - L &= p \frac{p}{m} e^{-\frac{b}{m}t} - \frac{1}{2} m \left(\frac{p}{m} e^{-\frac{b}{m}t} \right)^2 e^{\frac{b}{m}t} = \frac{p^2}{m} e^{-\frac{b}{m}t} - \frac{1}{2} \frac{p^2}{m} e^{-\frac{b}{m}t} \\ &= \frac{1}{2} \frac{p^2}{m} e^{-\frac{b}{m}t}\end{aligned}\tag{1.20}$$

and apply Hamilton's equations 1.4

$$\begin{aligned}d_t q &= \partial_p H = \frac{p}{m} e^{-\frac{b}{m}t} \\ d_t p &= -\partial_q H = 0.\end{aligned}\tag{1.21}$$

The second equation tells us momentum is constant p_0 . Substituting the constant in the first equation, we have the velocity as a function of time, which we can integrate. We have

$$\begin{aligned}q(t) &= q_0 + \frac{p_0}{b} \left(1 - e^{-\frac{b}{m}t} \right) \\ p(t) &= p_0.\end{aligned}\tag{1.22}$$

The kinematics works perfectly, but the dynamics seems off, as shown in fig. 1.4. First of all, based on the physics, one would expect the momentum to be decreasing in time

$$p(t) = m v(t) = m v_0 e^{-\frac{b}{m}t}.\tag{1.23}$$

However, conjugate momentum is a constant of motion. For the energy, we would expect the Hamiltonian to match the kinetic energy

$$E(t) = \frac{1}{2} m v^2(t) = \frac{1}{2} m v_0^2 e^{-2\frac{b}{m}t}\tag{1.24}$$

but if we express the Hamiltonian in terms of velocity we have

$$H(t) = \frac{1}{2} \frac{p^2}{m} e^{-\frac{b}{m}t} = \frac{1}{2} \frac{1}{m} \left(m v(t) e^{\frac{b}{m}t} \right)^2 e^{-\frac{b}{m}t} = \frac{1}{2} m v^2(t) e^{\frac{b}{m}t} = \frac{1}{2} m v_0^2 e^{-\frac{b}{m}t}.\tag{1.25}$$

That is, the energy decreases more slowly than it should. This is not good.

Now, it is true that conjugate momentum is not the same as kinetic momentum. But the difference, as we will see much more clearly later, is caused by non-inertial non-Cartesian coordinate systems and/or the presence of vector potential forces.⁸ We are not at all in that case. Also, note that at time $t = 0$ the momentum and the energy do match our expectation, but not after. Therefore imagine a situation where friction is non-negligible only in a particular region. We would expect $p = m v$ to be valid before it enters, but not when it comes out. But wouldn't it come out in another region where we would expect $p = m v$ to work? This is strange. How should we proceed?

⁸The relationship is $p_i = m g_{ij} v^j + q A_i$. This reduces to $p_i = m v^i$ if and only if we are in an inertial frame with Cartesian coordinates (i.e. $g_{ij} = \delta_{ij}$) and no forces $A_i = 0$

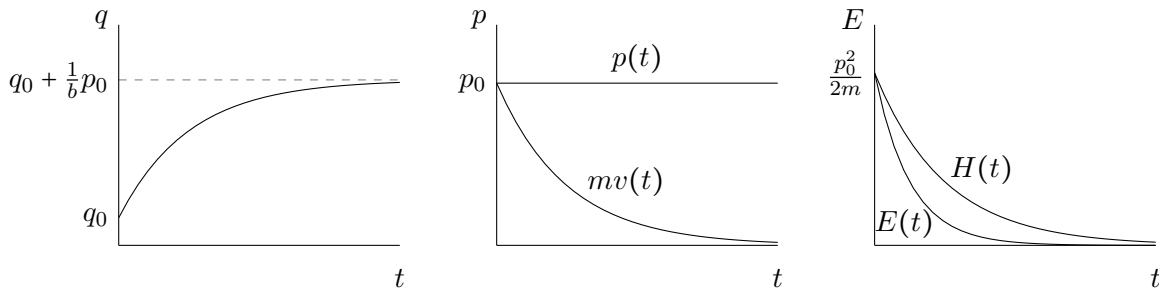


Figure 1.4: Trying to interpret $L = \frac{1}{2}mv^2e^{\frac{b}{m}t}$ and $H = \frac{1}{2}\frac{p^2}{m}e^{-\frac{b}{m}t}$ as respectively the Lagrangian and Hamiltonian of a particle under linear drag. While evolution of the position matches, note how the conjugate momentum is constant while the kinetic momentum decreases. Also, the Hamiltonian and the energy do not decrease at the same rate.

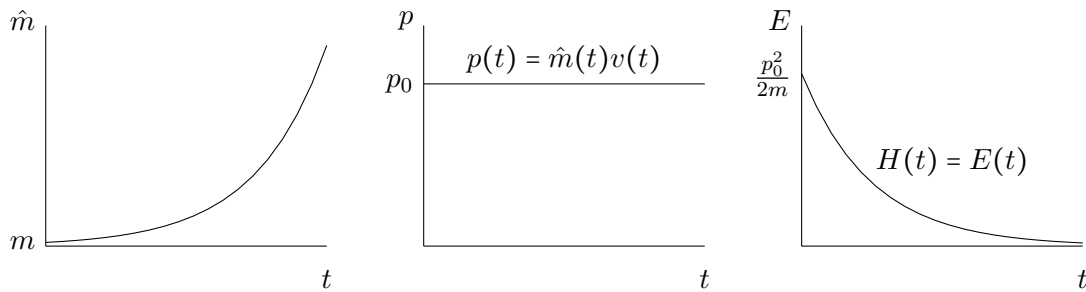


Figure 1.5: Showing how $L = \frac{1}{2}mv^2e^{\frac{b}{m}t}$ and $H = \frac{1}{2}\frac{p^2}{m}e^{-\frac{b}{m}t}$ can be interpreted as the Lagrangian and Hamiltonian of a variable mass system. The mass is increasing exponentially in time, while both conjugate and kinetic momentum remain constant. This means the velocity will need to decrease. The energy decreases at the same rate as the Hamiltonian.

Variable mass system

As it is typical in reverse physics, we will assume that things work in a reasonable way and that we simply have the wrong connection between physics and math. Recall that we started just with an equation, and we then interpreted m to be the mass of the system. Let's just assume that m is a constant with units of mass and define the actual mass of the system as the ratio between conjugate momentum and velocity. Looking back at 1.20, as shown in fig. 1.5, we have

$$\begin{aligned}
 \hat{m}(t) &= p(t)/v(t) = me^{\frac{b}{m}t} \\
 p(t) &= mv(t)e^{\frac{b}{m}t} = \hat{m}(t)v(t) \\
 H(t) &= \frac{1}{2}\frac{p^2(t)}{m}e^{-\frac{b}{m}t} = \frac{1}{2}\frac{p^2(t)}{\hat{m}} = \frac{1}{2}\hat{m}(t)v^2(t) = E(t)
 \end{aligned} \tag{1.26}$$

Now everything actually works perfectly: the relationship between velocity and conjugate momentum is respected, the Hamiltonian matches the kinetic energy. We just have a variable mass system. How and why does this work exactly?

Let us expand Newton's second law for a variable mass system.⁹ We have:

$$\begin{aligned} F^i &= d_t(\hat{m}v^i) = d_t\hat{m}v^i + \hat{m}a^i \\ \hat{m}a^i &= F^i - d_t\hat{m}v^i \end{aligned} \tag{1.27}$$

In particular, for our one dimensional case, let us set $F = 0$ and substitute \hat{m}

$$\begin{aligned} me^{\frac{b}{m}t}a &= 0 - d_tme^{\frac{b}{m}t}v = -\frac{b}{m}me^{\frac{b}{m}t}v \\ ma &= -bv. \end{aligned} \tag{1.28}$$

Therefore the same equation, the same kinematics, applies to a variable mass system that increases the mass over time. You can imagine, for example, a body that is absorbing mass from all directions, so that the balance of forces on the body is zero. The body, then, is not slowing down because of friction. It is slowing down because momentum is conserved, and if the mass is increasing, the velocity must be decreasing at the same rate. The energy, on the other hand, will decrease because the square of the velocity will decrease faster than the mass increases.

In Newtonian mechanics, we can readily distinguish these two cases because we have to be explicit about forces and masses. In Hamiltonian mechanics things are a bit more difficult because, as we will see later more precisely, conjugate momentum is not exactly kinetic momentum and the Hamiltonian is not exactly energy. Yet, conjugate momentum and the Hamiltonian are not kinematic quantities, they are dynamic quantities and therefore we can see that these would be different in different cases. In Lagrangian mechanics this is even more difficult to see because it looks like a purely kinematic theory, while it is not: the Lagrangian itself is not a purely kinematic entity. As we saw, Lagrangian mechanics implicitly assumes KE, which is a condition on the dynamics as well, and the Lagrangian itself is used to reconstruct conjugate momentum and the Hamiltonian. Moreover, if Lagrangian mechanics were a purely kinematic theory, and told us nothing about forces, energy or momentum, it would not be a complete formulation of classical mechanics.

So we have seen that the same kinematic equation can describe a constant mass dissipative system or a variable mass system. Is that it? Not quite. Recall that we mentioned that kinetic and conjugate momentum will differ in non-inertial frames. Note that we implicitly assumed that x and t represented the variables for an inertial observer, in the same way that we originally assumed m was the mass of the system. Could the same equation, then, be describing yet another system but in a non-inertial frame?

Non-inertial motion

Let's compare the motion of a particle traveling at constant velocity in an inertial frame, using t as the time variable, and the motion of a particle decelerating exponentially, using \hat{t} as the time variable

$$\begin{aligned} x(t) &= x_0 + v_0t \\ x(\hat{t}) &= x_0 + v_0\frac{m}{b}\left(1 - e^{-\frac{b}{m}\hat{t}}\right). \end{aligned} \tag{1.29}$$

⁹Note that, in general, the variable mass system should take into account the momentum gained or lost by the system when the mass is acquired or ejected. In our case, we are assuming that no momentum is lost, which means that either the mass is acquired/ejected uniformly from all directions or it is just an apparent change that depends on the change of coordinates.

Note the striking similarity: we can simply set

$$t = \frac{m}{b} \left(1 - e^{-\frac{b}{m}\hat{t}}\right) \quad (1.30)$$

which clearly takes us to a non-inertial frame since uniform motion is no longer uniform in the new frame.

Let's study how Newton's second law changes if we make a change of time variable while keeping the position variables unchanged

$$\begin{aligned} \hat{t} &= \hat{t}(t) \\ F^i &= d_t(m v^i) = d_t(m d_t x^i) = d_t \hat{t} d_{\hat{t}}(m d_{\hat{t}} \hat{t} d_{\hat{t}} x^i). \end{aligned} \quad (1.31)$$

If we set

$$\hat{m} = m d_t \hat{t} \quad (1.32)$$

we can express the previous equation in the following form

$$F^i = d_{\hat{t}} \hat{t} d_{\hat{t}}(\hat{m} d_{\hat{t}} x^i) = d_{\hat{t}} \hat{t} d_{\hat{t}}(\hat{m} \hat{v}^i) = d_{\hat{t}} \hat{t} \hat{F}^i. \quad (1.33)$$

This tells us that the second observer will see an effective mass rescaled exactly by the ratio between the time variables. Note that this is exactly what happens in special relativity: the clock for a boosted observer is dilated by a factor of γ which is exactly the factor used in the relativistic mass.¹⁰ If t is the time variable for an inertial frame and $t(\hat{t})$ is a non-linear function, the resulting frame will be non-inertial and the observer will see an effective variable mass system.

If we look at our problem this way, the rescaling of the mass, then, is not due to a truly variable mass, but a variable effective mass due to the slowing down of the clock. The body slows down because the non-inertial time is slowing down and the body appears to stop because the clock becomes infinitely slow. While this might sound like a contrived case,¹¹ these are exactly the type of situations a fully relativistic theory (i.e. one that works for all definitions of time and space variables) needs to take into account.

We can verify that this gives us the correct effective mass

$$\begin{aligned} d_{\hat{t}} \hat{t} &= d_{\hat{t}} \left(\frac{m}{b} (1 - e^{-\frac{b}{m}\hat{t}}) \right) = \frac{m}{b} d_{\hat{t}} (1 - e^{-\frac{b}{m}\hat{t}}) = -\frac{m}{b} d_{\hat{t}} e^{-\frac{b}{m}\hat{t}} = +\frac{m}{b} \frac{b}{m} e^{-\frac{b}{m}\hat{t}} = e^{-\frac{b}{m}\hat{t}} \\ \hat{m} &= m d_t \hat{t} = m (d_{\hat{t}} \hat{t})^{-1} = m e^{\frac{b}{m}\hat{t}}. \end{aligned} \quad (1.34)$$

And we can verify that we get the same equation by plugging in the time transformation in Newton's second law with a zero force

$$\begin{aligned} 0 &= d_t(m v) = d_t(m d_t x) = d_{\hat{t}} \hat{t} d_{\hat{t}}(m d_{\hat{t}} \hat{t} d_{\hat{t}} x) \\ &= e^{\frac{b}{m}\hat{t}} d_{\hat{t}}(m e^{\frac{b}{m}\hat{t}} \hat{v}) = e^{\frac{b}{m}\hat{t}} \left(m \frac{b}{m} e^{\frac{b}{m}\hat{t}} \hat{v} + m e^{\frac{b}{m}\hat{t}} \hat{a} \right) = e^{2\frac{b}{m}\hat{t}} (b\hat{v} + m\hat{a}) \end{aligned} \quad (1.35)$$

$$m\hat{a} = -b\hat{v}.$$

¹⁰It may be surprising to see a proto-relativistic effect showing up given that no assumption on space-time has been made. As we will see, these types of connections between different theories come up often in reverse physics.

¹¹On the surface, it sounds similar to what happens in general relativity with a black-hole. An observer that sees someone falling into a black hole will see him gradually slowing down as he approaches the event horizon and asymptotically stop there. The observer falling inside the black hole, instead, will perceive his time flowing uniformly and nothing special will happen as the event horizon is crossed.

Note that the expressions for momentum and energy will match the previous case because the system in the non-inertial frame looks like a variable mass system.

The relationship between kinematics and dynamics

This last case highlights a more subtle issue. In the two previous cases we were in the same inertial frame, we saw the same trajectory, the same kinematics, but we couldn't tell whether we were looking at a fixed mass system under linear drag or a variable mass system: we couldn't tell the dynamics. Now, we have the same system, a constant mass particle under no forces, described in two different frames, one inertial and one not. The motion of the system will naturally have different representations in the different frames, but this does not mean the motion or the causes of motion are different: it's the same object. Therefore we have the same kinematics even though we have different expressions for the same trajectory. The expression $x(t)$, then, is not enough to define the kinematics if we do not know exactly what x and t represent physically, if the frame is not given.

While typically one proceeds by defining the frame first and then the dynamics (i.e. the forces acting on the system), here we have followed a different approach: we first defined the dynamics (i.e. constant mass system under no forces) and then found the frame that matched the given kinematics (i.e. the trajectory or the relationship between velocity and acceleration). Given that Lagrangian and Hamiltonian mechanics are frame invariant, an intrinsic characterization of the system itself is exactly what we should be looking for. Saying, for example, that a system is subjected to no forces or to a linear drag is not frame invariant because forces are not frame invariant.

It is clear that the type of apparent variable mass due to non-inertial frames is unavoidable if we want to have a consistent theory with invariant laws. Therefore both Lagrangian and Hamiltonian mechanics must include these cases. However, it is not exactly clear what to do for true variable mass systems. From a cursory look, it would seem that everything is fine and there is no harm in including them. Yet again, from a cursory look we seemed to have a Lagrangian for a particle under linear drag. As we will see later, there are implicit connections between Lagrangian/Hamiltonian mechanics on one side and thermodynamics, statistical mechanics and special relativity on the other. Given that it is not clear to us whether these connections hold or not,¹² we will concentrate on the constant mass case from now on.

Let's recap what we learned. The biggest point is that we can't simply look at the kinematics and understand the causes of motion. The different formulations have different ways to relate the dynamics and the kinematics. Newtonian mechanics is the most clear about the dynamics as it makes us clearly spell out what is going on. This, however, comes at a cost: the equations are not covariant, meaning they have a different expression in different frames. The second law, in fact, is valid only for inertial frames with Cartesian coordinates. It is only in these frames, in fact, that a body will proceed in uniform motion if no forces are applied to it. If we are in polar coordinates, for example, the trajectory expressed in radius r and angle θ will not be linear. Even the notion of force is, if one looks closely, a bit ambiguous. In principle, we want to write both the second law $F = ma$ and the expression for work $dW = Fdx$. If dW is invariant under change of position variables, the force should be a covector and therefore $dW = F_i dx^i$. But since the acceleration a will change like a vector, we also have $F^i = ma^i$.

¹²For example, areas of phase space are connected to entropy. Does this connection hold with a variable mass system?

The notion of force in the second law and in the infinitesimal work are slightly different, and they coincide only if we are in an inertial frame and Cartesian coordinates.

On the other side, Hamiltonian and Lagrangian mechanics are coordinate independent: the laws remain the same if we change position variables. This makes them more useful in many contexts. Lagrangian mechanics is more useful when trying to study the symmetries of the system. Hamiltonian mechanics is more useful for statistical mechanics and to better separate degrees of freedom. However, this comes at a price. Hamiltonian and Lagrangian mechanics apply in fewer cases than Newtonian mechanics. As we saw, linear drag may look like it has a valid Hamiltonian/Lagrangian, but it doesn't. For quadratic drag or friction due to normal force, one cannot find a suitable trick, and is forced to use Rayleigh's dissipation functions which modify the Euler-Lagrange equations. This is not a coincidence: while Newtonian mechanics links kinematics and dynamics by choosing a particular frame, Hamiltonian and Lagrangian mechanics do so by fixing a type of system. It is the implicit knowledge of the type of system that allows us to reconstruct the dynamics just by looking at the kinematics in an unknown frame. What we need to understand, then, is what exactly is this restriction.

1.4 Reversing Hamiltonian mechanics

We now turn our attention to Hamiltonian mechanics and try to understand exactly what types of systems it focuses on. We will find twelve equivalent formulations of Hamiltonian mechanics that link ideas from vector calculus, differential geometry, statistical mechanics, thermodynamics, information theory and plain statistics. The overall result is that Hamiltonian mechanics focuses on systems that are assumed to be deterministic and reversible. We will see how the physical significance of that assumption differs from mathematically naive characterizations.

Mathematical characterizations

To simplify our discussion, we will first concentrate on a single degree of freedom. The first characterization of Hamiltonian mechanics is naturally in terms of the equations

$$\begin{aligned} d_t q &= \partial_p H \\ d_t p &= -\partial_q H. \end{aligned} \tag{HM-1D}$$

We will want to treat phase space as a generic two-dimensional space (i.e. manifold), like we would for a plane in physical space. We will reserve the term coordinate for the position variable q , while we will refer to the collection of position and momentum as state variables and will note them as $\xi^a = [q, p]$. We can now define the displacement field

$$S^a = d_t \xi^a = [d_t q, d_t p] \tag{1.36}$$

which is a vector field that defines the evolution of the system in time. Hamilton's equations, then, can be expressed as

$$\begin{aligned} S^q &= \partial_p H \\ S^p &= -\partial_q H. \end{aligned} \tag{1.37}$$

To bring out the geometric meaning of the equations, we introduce the matrix

$$\omega_{ab} = \begin{bmatrix} \omega_{qq} & \omega_{qp} \\ \omega_{pq} & \omega_{pp} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{SF-1D}$$

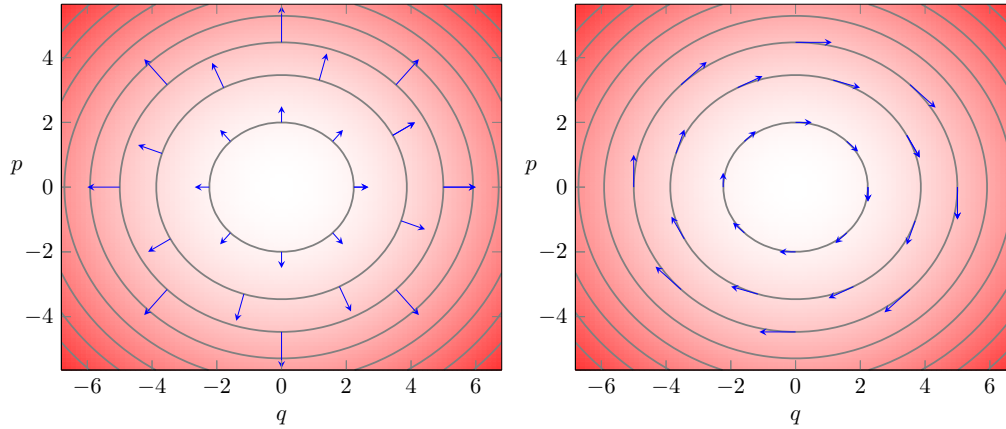


Figure 1.6: The surface plot shows the value of the Hamiltonian for a harmonic oscillator $H = \frac{p^2}{2m} + \frac{1}{2}kq^2$, red means higher value. The lines are the regions at constant energy H . On the left, the gradient of the Hamiltonian is shown. On the right, the displacement field is shown, which is the gradient rotated by a right angle. Note how the displacement is always parallel to the lines at constant energy.

which rotates a vector by a right angle.¹³ That is, if $v^a = [v^q, v^p]$, then $v_a = v^b \omega_{ba} = [-v^p, v^q]$.¹⁴ We can rewrite equation HM-1D as

$$S_a = S^b \omega_{ba} = \partial_a H \quad (\text{HM-G})$$

which tells us that the displacement field is the gradient of the Hamiltonian rotated by a right angle. Note that the gradient is perpendicular to the lines at constant energy. Therefore, as we can see in fig. 1.6, a right angle rotation gives us a vector field tangent to those lines, making it geometrically evident that the value of the Hamiltonian is a constant of motion. Condition HM-G is just a re-expression of HM-1D. Though it is already useful, we want to find different mathematical conditions which turn out to be equivalent to the equations.

We start by noting that the displacement field as expressed by 1.37 looks very similar to a curl of H , except that it is a two dimensional version. In vector calculus, a vector field is the curl of another field if and only if its divergence is zero.¹⁵ This holds here as well. First, we can verify that

$$\partial_a S^a = \partial_q S^q + \partial_p S^p = \partial_q \partial_p H - \partial_p \partial_q H = 0. \quad (1.38)$$

Geometrically, this means that the flow of S^a through a closed region is always zero, as shown in fig. 1.7. That is, $\oint (S^q dp - S^p dq) = 0$. Note that, since we are in a two dimensional space, a

¹³The notion of angle is technically ill-defined in phase space, but this slight imprecision makes it easier to get the point across.

¹⁴The notation is purposely similar to how indexes are raised and lowered in general relativity by the metric tensor $g_{\alpha\beta}$, since ω_{ab} plays a similar geometric role in phase space. One should be careful, however, that ω_{ab} is anti-symmetric (i.e. $\omega_{ab} = -\omega_{ba}$), so it matters which side is contracted. In terms of symplectic geometry, the rotated displacement field S_a corresponds to the interior product of the displacement field with the symplectic form, usually noted as $\iota_{S^a} \omega$ or $S^a \lrcorner \omega$.

¹⁵We will leave for now topological requirements as they would be a distraction from the overall point.

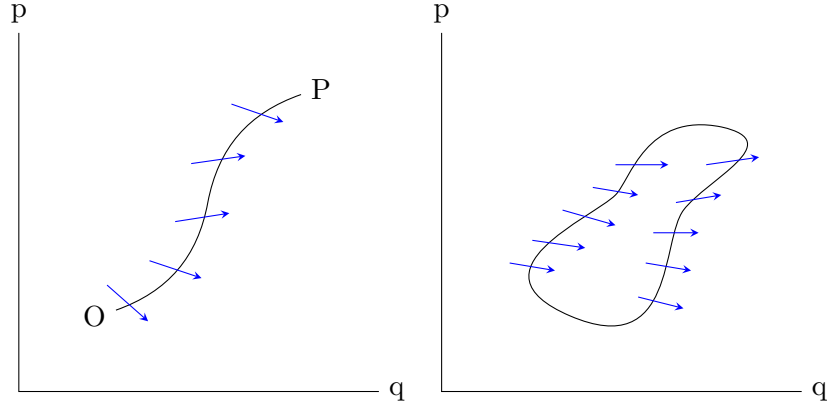


Figure 1.7: The flow of the displacement field S^a through a path, shown on the left, is equation to the difference of the Hamiltonian at the two points $\Delta H = \int_{OP} S^a \times d\xi^b$. The net flow of states through a region (i.e. the flow of the displacement field through the boundary) is zero, as shown on the right. This means that S^a is divergenceless and will admit a stream function, a potential, which corresponds to the Hamiltonian H .

hyper-surface has dimension $n - 1 = 2 - 1 = 1$ and therefore hyper-surfaces are lines. Therefore we have

$$\oint (S^q dp - S^p dq) = \oint (\partial_p H dp + \partial_q H dq) = \oint dH = 0. \quad (1.39)$$

That is, the flow of the displacement field is the line integral of the gradient of H , which is zero over a closed curve.

Conversely, we can see that each divergenceless field in two dimensions admits a stream function H that satisfies [HM-1D](#). Geometrically, we can construct H in the following way. Take a reference state O in phase space and assign $H(O) = 0$. For any other state P , consider the flow of S through any two lines that connect O and P . Given that the flow through the region contoured by those lines must be zero, the flow through each line must be equal. Therefore the flow through a line that connects O and P only depends on the states, it is path independent. We can assign $H(P) = \int_{OP} (S^q dp - S^p dq)$. If we expand the differential of H we have

$$dH = \partial_q H dq + \partial_p H dp = -S^p dq + S^q dp. \quad (1.40)$$

If we equate the components, we recover [HM-1D](#). Geometrically, at least for the one dimensional case, we can understand the difference of the Hamiltonian between two states as the flow of the displacement field between them.

We conclude that the following condition

$$\text{The displacement field is divergenceless: } \partial_a S^a = 0 \quad (\text{DR-DIV})$$

is equivalent to [HM-1D](#). Unlike [HM-G](#), this is a truly different mathematical condition.

Having looked at the flow through a region, we turn our attention to how regions themselves are transported by the evolution. Liouville's theorem states that volumes of phase space

are preserved during Hamiltonian evolution, which in our case will be areas over the q - p plane. To see this, let us review how variables transform, together with infinitesimal volumes:

$$\begin{aligned}
\hat{\xi}^a &= \hat{\xi}^a(\xi^b) \\
d\hat{\xi}^a &= \partial_b \hat{\xi}^a d\xi^b \\
d\hat{\xi}^1 \dots d\hat{\xi}^n &= |\partial_b \hat{\xi}^a| d\xi^1 \dots d\xi^n \\
d\hat{q}d\hat{p} &= \begin{vmatrix} \partial_q \hat{q} & \partial_p \hat{q} \\ \partial_q \hat{p} & \partial_p \hat{p} \end{vmatrix} dqdp
\end{aligned} \tag{1.41}$$

This tells us that, mathematically, a transformation is volume preserving if the determinant of the Jacobian $\partial_b \hat{\xi}^a$ is unitary. If \hat{q} and \hat{p} represent the evolution of q and p after an infinitesimal time step δt , we have

$$\begin{aligned}
\hat{q} &= q + S^q \delta t \\
\hat{p} &= p + S^p \delta t \\
\partial_b \hat{\xi}^a &= \begin{bmatrix} 1 + \partial_q S^q \delta t & \partial_p S^q \delta t \\ \partial_q S^p \delta t & 1 + \partial_p S^p \delta t \end{bmatrix} \\
|\partial_b \hat{\xi}^a| &= (1 + \partial_q S^q \delta t)(1 + \partial_p S^p \delta t) - \partial_p S^q \partial_q S^p \delta t^2 = 1 + (\partial_q S^q + \partial_p S^p) \delta t + O(\delta t^2).
\end{aligned} \tag{1.42}$$

Note that the first order term is proportional to the divergence of the displacement field, therefore the Jacobian determinant is equal to one if and only if the displacement is divergenceless. In other words, condition

$$\text{The Jacobian of time evolution is unitary: } |\partial_b \hat{\xi}^a| = 1 \tag{DR-JAC}$$

and condition

$$\text{Volumes are conserved through the evolution: } d\hat{\xi}^1 \dots d\hat{\xi}^n = d\xi^1 \dots d\xi^n \tag{DR-VOL}$$

are equivalent to [DR-DIV](#). We have found a third and a fourth way to characterize Hamiltonian evolution.

While condition [DR-VOL](#) is expressed in terms of areas, similar considerations will work for densities because a density is a quantity divided by an infinitesimal area. In fact densities

$$|\partial_b \hat{\xi}^a| \hat{\rho}(\hat{\xi}^a) = \rho(\xi^b). \tag{1.43}$$

transform in an equal and opposite way with respect to areas (i.e. the Jacobian determinant is on the other side of the equality). The unitarity of the Jacobian determinant, then, is equivalent to requiring that the density at an initial state is always equal to the density at the corresponding final state. Both areas and densities are transported unchanged by Hamiltonian evolution, as shown in [fig. 1.8](#). Therefore

$$\text{Densities are conserved through the evolution: } \hat{\rho}(\hat{\xi}^a) = \rho(\xi^b) \tag{DR-DEN}$$

is yet another equivalent characterization.

To get a yet different perspective, we can reframe these arguments in terms of ω_{ab} and S_a . Given two vectors v^a and w^a , the area of the parallelogram they form is $v^a w^p - v^p w^a$. This can

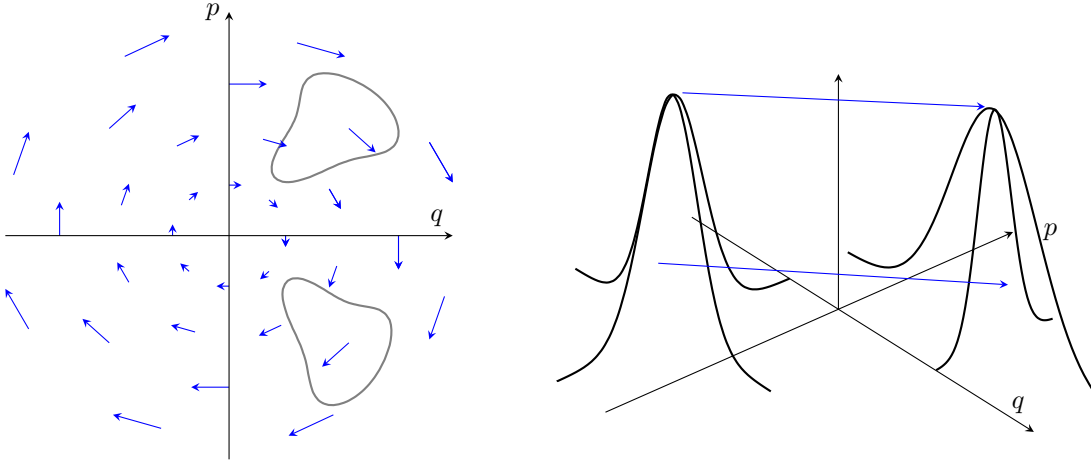


Figure 1.8: On the left side, we see how the displacement field S^a transports areas of phase space to equal areas of phase space. On the right, we see Hamiltonian evolution transports a probability distribution point by point. The value of the probability density remains the same as it moves over phase space.

be rewritten as $v^a \omega_{ab} w^b$, which means we can think of ω_{ab} as a tensor that, given two vectors, returns the area of the parallelogram they form.¹⁶ If we denote $\hat{v}^a = \partial_b \hat{\xi}^a v^b$ and $\hat{w}^a = \partial_b \hat{\xi}^a w^b$ the transformed vectors, the invariance of the area can be written as

$$v^a \omega_{ab} w^b = \hat{v}^c \omega_{cd} \hat{w}^d. \quad (1.44)$$

Since

$$\hat{v}^c \omega_{cd} \hat{w}^d = v^a \partial_a \hat{\xi}^c \omega_{cd} \partial_b \hat{\xi}^d w^b = v^a \hat{\omega}_{ab} w^b \quad (1.45)$$

the previous equivalence means that $\omega_{ab} = \hat{\omega}_{ab}$, that is ω_{ab} remains unchanged. In other words, preserving the area for all possible pairs of vectors is the same as preserving the tensor ω_{ab} that returns the areas. We now see that ω_{ab} plays such an important geometric role that

$$\text{The evolution leaves } \omega_{ab} \text{ invariant: } \hat{\omega}_{ab} = \omega_{ab} \quad (\text{DI-SYMP})$$

is yet another equivalent characterization of Hamiltonian mechanics.

It is useful to look more closely at the definition of the Poisson bracket

$$\{f, g\} = \partial_q f \partial_p g - \partial_p f \partial_q g = \begin{vmatrix} \partial_q f & \partial_p f \\ \partial_q g & \partial_p g \end{vmatrix}. \quad (1.46)$$

For a single degree of freedom, the Poisson bracket coincides with the Jacobian determinant, where f and g are the two new variables. It essentially tells us how the volume changes if we change state variables from $[q, p]$ to $[f, g]$. Canonical transformations, then, are those that

¹⁶More properly, ω_{ab} is a two-form.

do not change the units of area. The Poisson bracket can be expressed¹⁷ as

$$\{f, g\} = -\partial_a f \omega^{ab} \partial_b g = \partial_b g \omega^{ba} \partial_a f \quad (1.47)$$

where

$$\omega^{ab} = \begin{bmatrix} \omega^{qq} & \omega^{qp} \\ \omega^{pq} & \omega^{pp} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad (1.48)$$

is the inverse of ω_{ab} . The invariance of the Poisson brackets is equivalent to the invariance of the inverse of ω_{ab} , which is equivalent to [DI-SYMP](#). Therefore

$$\text{The evolution leaves the Poisson brackets invariant} \quad (\text{DI-POI})$$

is yet another equivalent characterization. So, again, we see how ω_{ab} plays a fundamental geometrical role.

We can also rewrite the flow of the displacement field

$$\int (S^q dp - S^p dq) = \int S^a \omega_{ab} d\xi^b = \int S_b d\xi^b \quad (1.49)$$

as the line integral of the rotated displacement field S_a . We can do that because in two dimensions the flow through a boundary is effectively a line integral along the boundary with the field rotated 90 degrees. This means that the following condition

$$\text{The rotated displacement field is curl free: } \partial_a S_b - \partial_b S_a = 0 \quad (\text{DI-CURL})$$

is equivalent to condition [DR-DIV](#).¹⁸ In fact, we can read equation [HM-G](#) as saying that the rotated displacement field is the gradient of the scalar potential H .

We can see that we have found plenty of alternative characterizations of Hamilton's equations [HM-1D](#) (or [HM-G](#)). Conditions [DR-DIV](#), [DR-JAC](#), [DR-VOL](#) and [DR-DEN](#) relate more directly to the displacement field S^a , while conditions [DI-SYMP](#), [DI-POI](#) and [DI-CURL](#) relate more directly to ω_{ab} and the rotated displacement field S_a . Nonetheless, they are all in terms of the mathematical description. While these are useful, the final goal of reverse physics is to find physical assumptions, not just equivalent mathematical definitions. So it is time to step back and try to understand what the math is really about.

Physical characterizations

Let us first reflect on what we just found out: the defining characteristic of Hamiltonian mechanics is not the transport of points, but the transport of areas and densities. If classical Hamiltonian mechanics were really about and only about point particles, there would be no reason for it to be characterized by [DR-DIV](#), [DR-VOL](#) or [DR-DEN](#). In fact, there would be

¹⁷To see how our definitions and notation map to that used in differential geometry, let us define $\partial^a H = \omega^{ab} \partial_b H$. Note that $\partial^a H$ corresponds to the Hamiltonian vector field of H usually noted X_H . The Poisson bracket is usually defined as $\omega(X_f, X_g)$. In our notation this becomes $\partial^a f \omega_{ab} \partial^b g = \omega^{ac} \partial_c f \omega_{ab} \omega^{bd} \partial_d g = \omega^{ac} \partial_c f \delta_a^d \partial_d g = \omega^{ac} \partial_c f \partial_a g$. One can see how the notation mimics the Einstein notation of general relativity and avoids the introduction of ad-hoc symbols.

¹⁸Those familiar with relativistic electromagnetism will recognize the expression $\partial_a S_b - \partial_b S_a$ as the generalization of the curl. More properly, it is the exterior derivative applied to a one-form.

no reason for the equations of motion [HM-1D](#) to be differentiable. Differentiable equations are exactly needed if we need to define the Jacobian, the transport of areas, or of densities defined on those areas. Classical point particles, then, are more aptly conceived not as points, but as infinitesimal regions of phase space, as distributions so peaked that only the mean value is important.

This, in retrospect, matches how classical mechanics is used in practice: planets, cannonballs, pendulums, beads on a wire, all the objects we study with classical mechanics are not point-like objects. They can be considered point-like if their size is negligible compared to the scale of the problem. If the distance between two celestial bodies is smaller than the sum of their radii, the point particle approximation clearly fails. This is also consistent with fluid dynamics and continuum mechanics, where we are literally studying the motion of infinitesimal parts of a material. It is interesting to see echoes of these considerations present in the mathematics.¹⁹

If we look at physics more broadly, we realize that in statistical mechanics we already have a physical interpretation for volumes of regions in phase space: they represent the number of states. Hamiltonian mechanics, then, maps regions while preserving the number of states. This means that, for each initial state there is one and only one final state, which leads to the following condition:

The evolution is deterministic and reversible. (DR-EV)

Note that by reversible here we mean that given the final state we can reconstruct the initial state. Given that areas measure the number of states, [DR-EV](#) is equivalent to [DR-VOL](#), which means this is another characterization of Hamiltonian mechanics. We can also see a connection to [DR-DEN](#). If we assign a density to an initial state, and we claim that all and only the elements that start in that initial state will end in a particular final state, we will expect the density of the corresponding final state to match. That is, if the evolution is deterministic and reversible, it may shuffle around a distribution, but it will never be able to spread it or concentrate it.

This makes us understand, at a conceptual level, why a dissipative system, like a particle under linear drag, is not a Hamiltonian system. A dissipative system will have an attractor: a point or a region to which the system will tend given enough time. This means that, in time, the area around the attractor must shrink, the density will concentrate over the attractor, but this is exactly what Hamiltonian systems cannot do. Therefore Hamiltonian systems cannot have attractors, they cannot be dissipative. By the same argument, they can't have unstable points or regions from which the system always goes away.

What may be confusing is that the motion of a particle under linear drag may seem reversible, in the sense that we are able to, given the final position and momentum, reconstruct the initial values. Mathematically, it maps points one-to-one and would seem to satisfy [DR-EV](#), even though it is not a Hamiltonian system. This is a perfect example of how focusing on just the points leads to the wrong physical intuition. Physically, we would say that a one meter range of position allows for more configurations than a one centimeter range, even though mathematically they have the same number of points. If we understand that states are infinitesimal areas of phase space, we can see that a dissipative system, though it does

¹⁹We will want to investigate this link in more detail later.

map the center points of infinitesimal areas one-to-one, it does not map the full infinitesimal area one-to-one. In this sense dissipative systems fail to be reversible.

Let that sink in: we found that, if the system is deterministic and reversible, it admits a Hamiltonian, a notion of energy, and that energy is conserved over time. This may seem like a surprising and unexpected result. In retrospect, we can make an argument for it based on familiar physics considerations. If a system is deterministic and reversible it means that its evolution only depends on the state of the system itself. This means that it does not depend on the state of anything else. A system whose evolution does not depend on anything else is an isolated system. Therefore a deterministic and reversible system is isolated, and from thermodynamics we know that an isolated system conserves energy. It should not be surprising, then, that a deterministic and reversible system conserves energy. However, we found that not only does it conserve energy, it defines it. Therefore this link between mechanics and thermodynamics is actually deeper than we may think at first, and we should explore it further.

The idea that a dissipative system is not reversible sounds true on thermodynamic grounds. But thermodynamic reversibility is not the ability to reconstruct the initial state, but rather the existence of a process that can undo the change. Alternatively, a process is thermodynamically reversible if it conserves thermodynamic entropy, which is a more precise characterization.²⁰ We should not, then, confuse the two notions of reversibility, but we can easily show their relationship. The fundamental postulate of statistical mechanics tells us that the thermodynamic entropy $S = k_B \log W$ is the logarithm of the count of states, which corresponds to volume in phase space. Since the logarithm is a bijective function, conservation of areas of phase space is equivalent to conservation of entropy. Therefore

The evolution is deterministic and thermodynamically reversible (DR-THER)

is yet another characterization of Hamiltonian mechanics.

There is another type of entropy that is also fundamental in both statistical mechanics and information theory: the Gibbs/Shannon entropy $I[\rho(\xi^a)] = -\int \rho \log \rho d\xi^1 \dots d\xi^n$ which is defined for each distribution $\rho(\xi^a)$. Recalling the transformation rules for both volumes 1.41 and densities 1.43, we have

$$\begin{aligned}
 I[\rho(\xi^a)] &= -\int \rho(\xi^a) \log \rho(\xi^a) d\xi^1 \dots d\xi^n \\
 &= -\int \hat{\rho}(\hat{\xi}^b) \left| \partial_a \hat{\xi}^b \right| \log \left(\hat{\rho}(\hat{\xi}^b) \left| \partial_a \hat{\xi}^b \right| \right) d\xi^1 \dots d\xi^n \\
 &= -\int \hat{\rho}(\hat{\xi}^b) \log \left(\hat{\rho}(\hat{\xi}^b) \left| \partial_a \hat{\xi}^b \right| \right) d\hat{\xi}^1 \dots d\hat{\xi}^n \\
 &= -\int \hat{\rho}(\hat{\xi}^b) \log \hat{\rho}(\hat{\xi}^b) d\hat{\xi}^1 \dots d\hat{\xi}^n - \int \hat{\rho}(\hat{\xi}^b) \log \left| \partial_a \hat{\xi}^b \right| d\hat{\xi}^1 \dots d\hat{\xi}^n \\
 &= I[\hat{\rho}(\hat{\xi}^b)] - \int \hat{\rho}(\hat{\xi}^b) \log \left| \partial_a \hat{\xi}^b \right| d\hat{\xi}^1 \dots d\hat{\xi}^n.
 \end{aligned} \tag{1.50}$$

Information entropy, then, remains constant if and only if the logarithm of the Jacobian determinant is zero, which means the Jacobian determinant is one. Therefore

The evolution conserves information entropy (DR-INFO)

²⁰The actual existence of a reverse process is not something that can always be guaranteed.

is equivalent to [DR-JAC](#) and is yet another characterization of Hamiltonian mechanics.

The fact that determinism and reversibility is equivalent to conservation of information entropy should not be, in retrospect, surprising. Given a distribution, its information entropy quantifies the average amount of information needed to specify a particular element chosen according to that distribution. If the evolution is deterministic and reversible, giving the initial state is equivalent to giving the final state and therefore the information to describe one or the other must be the same. Determinism and reversibility, then, can be understood as the informational equivalence between past and future descriptions.

Lastly, given that entropy is often associated with uncertainty, it may be useful to understand how Hamiltonian evolution affects uncertainty. Given a multivariate distribution, the uncertainty is characterized by the covariance matrix

$$\text{cov}(\xi^a, \xi^b) = \begin{bmatrix} \sigma_q^2 & \text{cov}_{q,p} \\ \text{cov}_{p,q} & \sigma_p^2 \end{bmatrix}. \quad (1.51)$$

The determinant of the covariance matrix gives us a coordinate independent quantity to characterize the uncertainty. If the distribution is narrow enough, we can use the linearized transformation to see how the uncertainty evolves after an infinitesimal time step δt . We have

$$|\text{cov}(\hat{\xi}^c, \hat{\xi}^d)| = |\partial_a \hat{\xi}^c \text{cov}(\xi^a, \xi^b) \partial_b \hat{\xi}^d| = |\partial_a \hat{\xi}^c| |\text{cov}(\xi^a, \xi^b)| |\partial_b \hat{\xi}^d|, \quad (1.52)$$

which means the uncertainty remains unchanged if and only if the Jacobian is unitary. So

$$\text{The evolution conserves the uncertainty of peaked distributions} \quad (\text{DR-UNC})$$

is equivalent to [DR-JAC](#) and is another characterization of Hamiltonian mechanics.

This connection gives us yet another insight on the nature of determinism and reversibility in physics. Given that all physically meaningful descriptions are finite precision, a system is deterministic and reversible in a physically meaningful sense if and only if the past/future descriptions can be reconstructed/predicted at the same level of precision. This gives us another perspective as to why areas and densities must be conserved.

Assumption of determinism and reversibility

We have found twelve equivalent characterizations that link Hamiltonian mechanics, vector calculus, differential geometry, statistical mechanics, thermodynamics, information theory and plain statistics. Though we only talked about the case of a single degree of freedom, it gives us a much better idea of what systems Hamiltonian mechanics is supposed to describe, those that satisfy the following

Assumption DR (Determinism and Reversibility). *The system undergoes deterministic and reversible evolution. That is, specifying the state of the system at a particular time is equivalent to specifying the state at a future (determinism) or past (reversibility) time.*

We can see how this concept is implemented mathematically: it is not simply a one-to-one map between points. Classical particles should be more properly thought of as infinitesimal regions of phase space. Conceptually, the count of states, the thermodynamic entropy and information entropy are all conserved, and are all equivalent characterizations of determinism

and reversibility. In terms of physical measurement, past and future states are given at the same level of uncertainty. But the most important lesson is that the foundations of classical mechanics are not disconnected from the foundations of all other disciplines we encountered. A full understanding of classical mechanics means understanding those connections as well.

1.5 Multiple degrees of freedom

We have seen how [DR](#) is a constitutive assumption for Hamiltonian mechanics, and in fact is equivalent to Hamiltonian mechanics for one degree of freedom. We now turn our attention to the general case, and we will find that [DR](#), by itself, is not enough to recover the equations. We will need an additional assumption, that of the independence of degrees of freedom.

First, let's take Hamilton's equations for multiple degrees of freedom

$$\begin{aligned} d_t q^i &= \partial_{p_i} H \\ d_t p_i &= -\partial_{q^i} H \end{aligned} \tag{HM-ND}$$

and re-express them in terms of generalized state variables. These will be noted as $\xi^a = [q^i, p_i]$ and will span a $2n$ -dimensional space (i.e. manifold). The displacement field will be

$$S^a = d_t \xi^a = [d_t q^i, d_t p_i] \tag{1.53}$$

which again is the vector field that defines the evolution of the system in time. Hamilton's equations, then, can be expressed as

$$\begin{aligned} S^{q^i} &= \partial_{p_i} H \\ S^{p_i} &= -\partial_{q^i} H. \end{aligned} \tag{1.54}$$

Similarly to the previous case, let's introduce the following matrix

$$\omega_{ab} = \begin{bmatrix} \omega_{q^i q^j} & \omega_{q^i p_j} \\ \omega_{p_i q^j} & \omega_{p_i p_j} \end{bmatrix} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes I_n \tag{SF-ND}$$

which performs a 90 degree rotation within each degree of freedom, switching the components between position and momentum. That is, if $v^a = [v^{q^i}, v^{p_i}]$, then $v_a = v^b \omega_{ba} = [-v^{p_i}, v^{q^i}]$.²¹ We can rewrite equation [HM-ND](#) as

$$S_a = S^b \omega_{ba} = \partial_a H \tag{1.55}$$

which notationally is the same as [HM-G](#). The insight that the displacement field is equal to the gradient of H rotated 90 degrees still applies, except there are now multiple ways, in principle, to do that rotation. It is only the one defined by ω_{ab} that works.

²¹For those versed in symplectic geometry, $v^a \omega_{ab}$ are the components of the one-form $\omega(v, \cdot)$. However, we are not going to call it a one-form as that assumes that the whole object is a map from a vector field to a scalar field, and we do not know whether that is the correct physical understanding. In other words, we want simply to understand what the quantities are doing without being tied, as much as possible, to a particular way to frame it. Full reverse engineering of differential geometry will be done in a later chapter, once the physics we need to describe is clear.

Conditions [DR-DIV](#), [DR-JAC](#), [DR-VOL](#) and [DR-DEN](#) are still satisfied and equivalent to each other. In fact, the divergence of the displacement field is zero

$$\partial_a S^a = \partial_{q^i} S^{q^i} + \partial_{p_i} S^{p_i} = \partial_{q^i} \partial_{p_i} H - \partial_{p_i} \partial_{q^i} H = 0 \quad (1.56)$$

and the Jacobian is unitary

$$\begin{aligned} \hat{q}^i &= q^i + S^{q^i} \delta t \\ \hat{p}_i &= p_i + S^{p_i} \delta t \\ \partial_b \hat{\xi}^a &= \begin{bmatrix} \delta_j^i + \partial_{q^j} S^{q^i} \delta t & \partial_{p_j} S^{q^i} \delta t \\ \partial_{q^j} S^{p_i} \delta t & \delta_i^j + \partial_{p_j} S^{p_i} \delta t \end{bmatrix} \\ |\partial_b \hat{\xi}^a| &= \left| \delta_j^i + \partial_{q^j} S^{q^i} \delta t \right| \left| \delta_i^j + \partial_{p_j} S^{p_i} \delta t \right| - \left| \partial_{p_j} S^{q^i} \delta t \right| \left| \partial_{q^j} S^{p_i} \delta t \right| \\ &= 1 + \left(\partial_{q^i} S^{q^i} + \partial_{p_i} S^{p_i} \right) \delta t + O(\delta t^2) \end{aligned} \quad (1.57)$$

since the first-order term is again the divergence. The Jacobian is still the multiplicative factor between past/future areas (and densities), and therefore they are conserved even in the case of multiple degrees of freedom.

However, these conditions are not equivalent to [HM-ND](#). The displacement field S^a has $2n$ components and is therefore specified by $2n$ functions. Conditions [DR-DIV](#), [DR-JAC](#), [DR-VOL](#) and [DR-DEN](#) specify the same single constraint, bringing down to $2n - 1$ the number of independent components. The choice of Hamiltonian provides another constraint, leaving $2n - 2$ choices undetermined. In the single degree of freedom case, $n = 1$, no choices are left, and therefore the displacement field is fully constrained. In the general case, however, this is not enough to fully characterize the evolution. Therefore [HM-ND](#) implies [DR-DIV](#), [DR-JAC](#), [DR-VOL](#) and [DR-DEN](#), but the converse is not true.

Let's see what happens to condition [DI-SYMP](#), the invariance of ω in the general case. We have

$$\begin{aligned} \hat{\omega}_{ab} &= \partial_a \hat{\xi}^c \omega_{cd} \partial_b \hat{\xi}^d \\ &= (\delta_a^c + \partial_a S^c \delta t) \omega_{cd} (\delta_b^d + \partial_b S^d \delta t) \\ &= \omega_{ab} + (\partial_a S^c \omega_{cb} + \omega_{ad} \partial_b S^d) \delta t + O(\delta t^2) \\ &= \omega_{ab} + (\partial_a (S^c \omega_{cb}) + \partial_b (S^d \omega_{ad})) \delta t + O(\delta t^2) \\ &= \omega_{ab} + (\partial_a (S^c \omega_{cb}) - \partial_b (S^d \omega_{da})) \delta t + O(\delta t^2). \end{aligned} \quad (1.58)$$

Therefore, the invariance of ω_{ab} is equivalent to

$$\partial_a (S^c \omega_{cb}) - \partial_b (S^c \omega_{ca}) = 0. \quad (1.59)$$

In terms of the rotated displacement field S_a we have the more compact form

$$\partial_a S_b - \partial_b S_a = 0. \quad (1.60)$$

This tells us that the rotated displacement field S_a is curl free, which is the same condition as [DI-CURL](#), therefore [DI-CURL](#) and [DI-SYMP](#) are equivalent conditions also in the general case.

Note that Hamilton's equations state that the rotated displacement field is the gradient of the Hamiltonian, and therefore

$$\partial_a S_b - \partial_b S_a = \partial_a(S^c \omega_{cb}) - \partial_b(S^c \omega_{ca}) = \partial_a \partial_b H - \partial_b \partial_a H = 0, \quad (1.61)$$

which simply verifies that the curl of the gradient is zero. Conversely, if S_a is curl-free, then it admits a scalar potential H such that

$$S_a = S^b \omega_{ba} = \partial_a H \quad (1.62)$$

which recovers Hamilton's equations. Therefore **HM-ND**, **DI-CURL** and **DI-SYMP** are equivalent.

The relationship between Poisson brackets and ω^{ab} is the same in the general case, therefore **DI-POI** and **DI-SYMP** are equivalent as well.

To sum up, in the general case **HM-ND**, **HM-G**, **DI-SYMP**, **DI-POI** and **DI-CURL** are all equivalent and therefore full characterizations of Hamiltonian mechanics in the general case. These imply **DR-DIV**, **DR-JAC**, **DR-VOL** and **DR-DEN**, which are all equivalent to one another, but weaker conditions that cannot recover Hamiltonian mechanics in full. For the second set of conditions, we already have an intuitive geometrical picture: the net flow of the displacement within a region of phase space is zero, volumes are preserved and so are densities. We need to build a stronger geometrical intuition for the first set, which is actually the more fundamental one.

Condition **DI-SYMP** tells us that $v^a \omega_{ab} w^b$ is a conserved quantity, no matter what vectors v^a and w^b we choose. In the case of a single degree of freedom, this represented the area of the parallelogram formed by the two vectors, which was also the volume of the region. In the general case, we still have two vectors, but the situation is a bit more complicated.

We can gain an understanding by looking at the outer product decomposition for ω_{ab} we saw in **SF-ND**. This tells us that what happens within a degree of freedom is different from what happens across degrees of freedom. If we pick a single degree of freedom $1 \leq x \leq n$ and two vectors $v = v^q e_{q^x} + v^p e_{p_x}$ and $w = w^q e_{q^x} + w^p e_{p_x}$ that stretch along that degree of freedom, then we have

$$v^a \omega_{ab} w^b = v^q w^p - v^p w^q. \quad (1.63)$$

That is, within each degree of freedom, ω_{ab} computes the area of the parallelogram. Since ω_{ab} is conserved, parallelograms within any degree of freedom will be mapped to parallelograms of the same size.

If we pick two different DOFs x and y and two corresponding vectors $v = v^q e_{q^x} + v^p e_{p_x}$ and $w = w^q e_{q^y} + w^p e_{p_y}$, then we have

$$v^a \omega_{ab} w^b = 0. \quad (1.64)$$

This defines a notion of orthogonality between different degrees of freedom. Since ω_{ab} is conserved, this notion of orthogonality is preserved during the evolution: orthogonal degrees of freedom are mapped to orthogonal degrees of freedom.

Those familiar with general relativity and/or Riemannian geometry may gain more insight by the following analogy. In those cases, the metric tensor g_{ij} defines the geometry by defining the scalar product between vectors. That is, given two vectors v^i and w^j ,

$v^i g_{ij} w^j = |v||w| \cos \theta_{vw}$. Therefore the metric tensor defines the length and angles for vectors. In Cartesian coordinates, the metric tensor is a unitary matrix of the same dimension of the space. The form ω_{ab} does something in some sense similar and in some sense different. It defines areas within degrees of freedom and angles between them. Hamiltonian evolution preserves these areas and angles.

If areas and orthogonality are preserved, then volumes are preserved as well. The volume of a parallelepiped formed by parallelograms on orthogonal degrees of freedom will simply be the product of the areas of the parallelograms. Therefore we can understand why Hamiltonian mechanics satisfies **DR-VOL**. We can also understand why **DR-VOL** is not enough to recover Hamiltonian mechanics. An evolution could stretch one degree of freedom while shrinking another by the same amount. The total volume would remain the same, even though the area in each degree of freedom wouldn't. For example, take the system of equations:

$$\begin{aligned} d_t q^1 &= S^{q^1} = \frac{p_1}{m} \\ d_t p_1 &= S^{p_1} = -b p_1 \\ d_t q^2 &= S^{q^2} = \frac{p_2}{m} \\ d_t p_2 &= S^{p_2} = b p_2 \end{aligned} \tag{1.65}$$

The first degree of freedom is a particle under linear drag, while the second is a particle accelerated (not decelerated) proportionally to its momentum by the same coefficient. We can verify that

$$\partial_a S^a = \partial_{q^1} \frac{p_1}{m} + \partial_{p_1} (-b p_1) + \partial_{q^2} \frac{p_2}{m} + \partial_{p_2} (b p_2) = -b + b = 0 \tag{1.66}$$

the divergence is zero and therefore **DR-DIV** and **DR-VOL** are satisfied. However

$$\partial_{q^1} S_{p_1} - \partial_{p_1} S_{q^1} = \partial_{q^1} S^{q^1} \omega_{q^1 p_1} - \partial_{p_1} S^{p_1} \omega_{p_1 q^1} = \partial_{q^1} \frac{p_1}{m} (1) - \partial_{p_1} (-b p_1) (-1) = -b. \tag{1.67}$$

The curl of S_a , then, is not zero, **DI-CURL** is not satisfied, nor are **DI-SYMP** and **HM-ND**. The system is not Hamiltonian precisely because we are not preserving the areas within each independent DOF: the first is shrunk and the second stretched.

Now that we have a more precise understanding of the mathematics and the geometry, we should turn to the physics. Note that all the previous physical conditions **DR-EV**, **DR-THER**, **DR-INFO** and **DR-UNC** are equivalent to **DR-VOL** and **DR-JAC**. Therefore determinism and reversibility is clearly a constitutive assumption of Hamiltonian mechanics in the general case, but it cannot be the only one. Ideally, we would like to find a condition that is independent of **DR**. However, we saw that **DI-SYMP** implies **DR-VOL**, therefore the mathematics does not already give us two independent conditions we can map to the physics.

This is an important aspect to understand for reverse physics: the mapping between physical and mathematical conditions need not necessarily be one to one. A single mathematical condition can map to multiple physical ones, or the same physical condition can map to multiple mathematical ones. We saw before that determinism and reversibility forces the evolution map to be both bijective and volume preserving. Mathematically, these are two independent conditions. We can have a bijection that is not volume preserving (e.g. a linear transformation that stretches one side) or a volume preserving map that is not bijective (e.g. a map from

\mathbb{R} to \mathbb{R} that maps all rationals to 0 while leaving all the irrationals the same). Yet, a physically meaningful deterministic and reversible map must do both. Here we have the opposite: Hamiltonian mechanics implies determinism and reversibility, but is also implying at least another physical condition, and we need to understand which and whether it is physically independent.

Let's start from what we have already established: the phase space volume quantifies the number of states in the region. It stands to reason that the area on each degree of freedom identifies the number of configurations for that degree of freedom. Therefore, given two vectors, $v = v^q e_{q^x} + v^p e_{p_x}$ and $w = w^q e_{q^x} + w^p e_{p_x}$, constrained on a single degree of freedom x , the area of the parallelogram they identify, $v^q w^p - v^p w^q$, quantifies the number of configurations. Therefore ω_{ab} returns the number of configurations within each degree of freedom. What about between degrees of freedom?

As we saw, the conserved volume, which means the total number of states, is the product of those degrees of freedom that are orthogonal. In terms of ω_{ab} , x and y are orthogonal if $\omega_{q^x q^y} = \omega_{q^x p_y} = \omega_{p_x q^y} = \omega_{p_x p_y} = 0$. In this case, the volume is simply the product of the phase-space areas on the x and y degrees of freedom. This means that the total number of states is the product of the configurations of each degree of freedom. Physically, it means that a configuration choice for x does not constrain the configurations for y . This means that the degrees of freedom are independent.

Given that there are different notions of independence, let us go through an example. Suppose we have a rabbit farm and describe its state with the number of males and females. These two variables are independent: if we say there are 231 females it doesn't, in principle, tell us anything about the number of males. Now, we may expect the population of both sexes to be about equal, and we may even find that in most rabbit farms that is the case, but this does not describe something about the nature of the variables themselves: it describes the nature of rabbit farms. Chicken farms, for example, would be predominantly females, as those are the ones that lay eggs.

Now we could choose to describe the rabbit farm with another set of variables: the number of females and the total number of rabbits. In this case, the variables are not independent. If we find that there are 231 females, it tells us, in principle, that there must be at least 231 rabbits. Conversely, if we find that there are 231 rabbits, there can only be up to 231 females. This dependence is not a feature of the rabbit farms. It does not just happen to be that there are no farms where the number of female rabbits exceeds the total number of rabbits. There can't be one.

This type of independence is very different from the notion of independence in terms of statistics and probability. The latter is in terms of whether the probability distribution factorizes. That is, if $P(f, m)$ is the probability that a particular rabbit farm has f females and m males, the distributions of males $P(m)$ and females $P(f)$ are independent if $P(f, m) = P(f)P(m)$.

We therefore have two notions of independence. One is on the variables themselves and whether they can allow (i.e. whether we can measure) different combinations. One is on the probability distribution we may have in a particular case and whether it factorizes. The orthogonal directions in phase space, then, are independent in the first, stronger, sense. The degrees of freedom themselves are independent, regardless of what probability distribution one may put on top.

One may ask whether there is a link between the two, and in fact there is. Going back to

our rabbits, we can easily see that, given any distribution $P(f)$ on the females, we can choose any distribution $P(m)$ on the males and set $P(f, m) = P(f)P(m)$. However, this does not happen for the total number. Suppose we chose $P(f)$ and we wanted to find a $P(f + m)$ such that $P(f, f + m) = P(f)P(f + m)$. The probability of the total number of rabbits could not change based on the number of females. If the probability of having 231 females is non-zero, then the probability of having less than 231 total rabbits in that case must be zero. But since we want the probability of having less than 231 rabbits independent of the number of females, then it must be zero for all cases. That is, the probability $P(f + m)$ must be zero for all numbers smaller than the greatest value of f such that $P(f) \neq 0$. If there is no such greatest value of f , for example f follows a geometric distribution, no $P(f + m)$ can exist that is independent of $P(f)$.

The conclusion is that only independent variables can support independent distributions.²² It should be clear that this observation is something that goes beyond the physical underpinning of Hamiltonian mechanics: it is something that applies to any variable to which we want to assign a probability distribution. As such, we do not want to expand the scope too much at this point, though clearly we will need to explore this more in full.

Here we limit ourselves to concluding that the following four conditions

The system is decomposable into independent DOFs	(IND-DOF)
The system allows statistically independent distributions over each DOF	(IND-STAT)
The system allows informationally independent distributions over each DOF	(IND-INFO)
The system allows peaked distributions where the uncertainty is the product of the uncertainty on each DOF	(IND-UNC)

are equivalent. The first means that the count of states factorizes, the second that probability distributions can factorize, the third that the information entropy can sum, and the fourth that the determinant of the covariance matrix can factorize. Since only independent variables can support statistically independent distributions, [IND-DOF](#) is equivalent to [IND-STAT](#). Statistical independence of random variables coincides with independence of information entropy, therefore [IND-STAT](#) is equivalent to [IND-INFO](#). The uncertainty for peaked distributions factorizes if and only if the joint distribution is the product of independent distributions, therefore [IND-STAT](#) is equivalent to [IND-UNC](#).

Clearly these conditions are independent from [DR](#). We can imagine a deterministic and reversible system that cannot be broken into separate independent degrees of freedom, and we can imagine a system that can be broken into separate independent degrees of freedom that does not evolve deterministically or reversibly. The question is whether assuming independent degrees of freedom and deterministic and reversible evolution is enough to recover Hamiltonian mechanics.

The first thing to check, then, is whether we have enough constraints to recover ω_{ab} . Assuming that the system can be broken up into independent degrees of freedom, we must be

²²Formally, let (Ω, \mathcal{F}, P) be a probability space, let $X : \Omega \rightarrow E_X$ and $Y : \Omega \rightarrow E_Y$ be two random variables and $Z : \Omega \rightarrow E_X \times E_Y$ be their joint random variable (i.e. $Z(\omega) = (X(\omega), Y(\omega))$). Then X and Y are independent in this stronger sense if $Z(\Omega) = X(\Omega) \times Y(\Omega)$ and are statistically independent if the cumulative distribution function $F_Z(x, y) = F_X(x)F_Y(y)$ factorizes. Alternatively, they are independent if the σ -algebra generated by the joint distribution Z is the product of the σ -algebras generated by X and Y , and are statistically independent if the σ -algebras generated by X and Y are independent in the standard probability sense.

able to define the count of configurations for each degree of freedom, and independent degrees of freedom must be orthogonal. The fact that ω_{ab} will return zero if it acts on directions belonging to independent degrees of freedom is really telling us that ω_{ab} counts not just configurations, but independent configurations. This, in retrospect, makes sense. But, so far this seems to restrict ω_{ab} to be

$$\omega_{ab} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix}. \quad (1.68)$$

The fact that ω_{ab} must return the area in each degree of freedom constrains the left matrix of the outer product to the one above. The fact that ω_{ab} needs to return zero across independent degrees of freedom constrains the right matrix of the outer product to be diagonal. However, there is nothing, at this point, that seems to constrain the area of each DOF to map to the same count of configurations. Naturally, we could simply rescale conjugate momentum in each DOF to homogenize the count, but this would be an arbitrary freedom. Is there something forcing all the a_i to be the same?

Note that q^i and p_i are not the only variables that form independent degrees of freedom. If we take two independent DOFs x and y , $x+y$ and $x-y$ will also form independent degrees of freedom. That is, ω_{ab} defines orthogonality for all DOFs, not just those of a particular basis. Changing x and y to $x+y$ and $x-y$ will effectively apply a rotation on the diagonal matrix, which will remain diagonal only if the coefficients on the diagonal are the same.

This tells us that if we want ω_{ab} to properly capture the independence of linear combinations of independent DOFs, the diagonal matrix must have the same coefficient. This coefficient represents the freedom we have in choosing the units of omega with respect to the units of everything else. In SI units, by convention, the product between q^i and p_i is in $J \cdot s$ (i.e. the same units of h and angular momentum) and we set the coefficient to 1. Therefore expressing the number of configurations with the same units for all DOFs is not an extra constraint, but it is necessary to keep track of the dependency relationship for all degrees of freedom.

Therefore the other constitutive assumption of Hamiltonian mechanics is

Assumption IND (Independent DOFs). *The system is decomposable into independent degrees of freedom. That is, the variables that describe the state can be divided into groups that have independent definition, units and count of states.*

This assumption leads to conditions [IND-DOF](#), [IND-STAT](#), [IND-INFO](#) and [IND-UNC](#), which we saw implies the existence of a form ω_{ab} that defines the independence of DOFs together with the count of independent configurations for each DOF. Conversely, assuming [HM-ND](#) means defining an ω_{ab} such that [IND](#) is satisfied. The last question is whether [DR](#) and [IND](#) are enough to recover [HM-ND](#).

As we saw, [IND](#) by itself means the existence of the counting form ω_{ab} . Meanwhile, [DR](#) by itself means the conservation of the total number of states, the volume. These two mathematical conditions, by themselves, do not lead to Hamiltonian mechanics. We need that the counting form itself is conserved, meaning that DOF independence and the configuration count is preserved. This boils down to the following question: does it make sense, on physical

grounds, to have a deterministic and reversible evolution that takes a system decomposable into independent degrees of freedom and turns it into a system that is no longer decomposable? More specifically, can deterministic and reversible evolution take two independent degrees of freedom and break their independence?

We should remind ourselves that independence here is not statistical independence. Clearly, Hamiltonian evolution can add correlations, and therefore evolve the product of two independent distributions into something that is no longer factorizable. This is not the question at hand. The notion of independence on the table is the one that tells us that all the combinations of configurations are at least possible. Recall the example of the rabbits: number of females and total rabbits are not independent because we cannot have more females than rabbits. This is the type of independence we are interested in. Can deterministic and reversible evolution lose this type of independence?

The answer is, as you may expect, no. This is easily seen in the finite case. Suppose we have two integer variables, $1 \leq x \leq 3$ and $1 \leq y \leq 3$. If they are independent, we have a total of 9 distinct cases. If the evolution is deterministic and reversible, we will still have 9 distinct cases, which means the variables must remain independent. Note that we may introduce a correlation between x and y . But still, we need 9 total cases.

The issue here is that the case of independent variables is maximal as it posits that all combinations of configurations are possible. Therefore, the only way to make variables no longer independent is to decrease the number of distinct cases, which cannot happen during deterministic and reversible evolution. The same must happen in the infinite or the continuous case, because finite ranges must still be comparable with finite sizes which must hold the same property. Therefore when we take both physical assumptions **IND** and **DR**, the physics tells us that independence of degrees of freedom must be preserved, which means preserving ω_{ab} as well. In other words:

Insight 1.69. *Hamiltonian mechanics is exactly the deterministic and reversible evolution of a system decomposable into a finite collection of independent degrees of freedom.*

This conclusion is yet another example of why just looking at the math is not enough. Two physical conditions taken separately may each impose one mathematical condition, but it is not necessarily true that imposing them together will only impose the conjunction of the two mathematical conditions. More than a problem with math in general, we believe it is an indication that the math we currently use is not the “physically correct” one as it does not seem to be capturing the entirety of the physical conditions.

Note that, in principle, we could ask the evolution to preserve the independence of DOFs without requiring **DR**. As we saw, fixing all independent DOFs fixes ω_{ab} up to a scalar factor, which could change during the evolution. The volume would stretch or shrink depending on the factor, stretching or shrinking each DOF by the same amount. An example of this would be a particle under linear drag in three dimensions. Since a faster moving object will be subjected to a greater frictional force, a spread in momentum Δp will become smaller in time, and will tend to zero as time increases. Given that the friction coefficient is the same for all directions, all degrees of freedom will shrink at the same rate. If we understand the volume as entropy, this tells us that the only way we can add or remove entropy to/from a system while preserving the independence of the degrees of freedom is by dividing that entropy contribution equally among each DOF. In other words, preservation of DOF independence

gives us a sort of equipartition of entropy change. It is again striking to find these connections between disciplines at such a basic level.

To conclude, we now have a mathematically and physically precise way to characterize Hamiltonian evolution. We had found that Hamiltonian mechanics did not apply to all cases, and now we know exactly to which cases it applies: systems described by finitely many independent degrees of freedom undergoing deterministic and reversible evolution.

1.6 Reversing differential topology

In the previous sections, we saw that elements of differential topology and differential geometry started appearing: we employed a generalization of the curl and we used the form ω_{ab} to lower indexes, much like one does in general relativity with the metric tensor $g_{\alpha\beta}$. Since we will use these tools more and more, we should take a detour and understand the physical significance of the tools themselves. We will end up with a generalized notion of integral and of differential operations that work with an arbitrary number of dimensions. We will also conclude that to reach the full potential of reverse physics, we need to apply the same techniques not just to the equations themselves, but to the mathematical tools we use to formulate them.

The main reason we are forced to abandon vector calculus in favor of differential topology and geometry is that vector calculus works in three dimensions but does not generalize to an arbitrary dimensional space. Special and general relativity, for example, live on a four-dimensional space-time; Hamiltonian and Lagrangian mechanics live in phase space, which can have an arbitrarily large number of degrees of freedom. Similarly to what we have done with the equations, we will start from the expressions of vector calculus, see their limitations, and construct generalized ones. We warn the reader who is already familiar with differential topology that this process will give us notation and concepts that are slightly different from what is used by mathematicians. We will discuss these differences at the end of the section.

The first tool we need to generalize is that of line, surface and volume integrals. For example, the mass m within a region V can be understood as the sum of the contributions of a mass density ρ within each infinitesimal volume dV :

$$m(V) = \iiint_V \rho dV. \quad (1.70)$$

Similarly, the magnetic flux Φ through a surface Σ can be understood as the sum of the contributions of the magnetic field \vec{B} over each infinitesimal $d\vec{\Sigma}$:

$$\Phi(\Sigma) = \iint_{\Sigma} \vec{B} \cdot d\vec{\Sigma}. \quad (1.71)$$

Lastly, the work W over a path γ can be understood as the sum of the contributions of a force field \vec{f} over each infinitesimal segment $d\vec{\gamma}$:

$$W(\gamma) = \int_{\gamma} \vec{f} \cdot d\vec{\gamma}. \quad (1.72)$$

Note the pattern: the functionals W , Φ and m all take a region of space while \vec{f} , \vec{B} and ρ act on infinitesimal regions. However the pattern is not completely consistent. In the line integral case, we have the product between a vector representing the force and a vector representing the displacement along the line. In the surface integral case we have the product between a

pseudo-vector representing the magnetic force and a pseudo-vector representing the normal to the surface element. In the volume integral case we have the product between a pseudo-scalar representing the density and a pseudo-scalar representing the volume of the infinitesimal region. Each operation is slightly ad-hoc. Moreover, a surface has a single perpendicular direction only in three dimensions. In four dimensions, for example, there are multiple different perpendiculars to the same plane. Lastly, they all require a notion of product between vectors, an inner product, which in differential geometry is only defined on Riemannian spaces, those that define a metric tensor. That is, those spaces in which angles and distances are well defined. In physics we are so used to working with spaces that have an inner product that it may seem that all spaces provide one, but that is not the case. In phase space, there is no way to compare differences in position with differences in momentum, therefore we do not have a metric to take a scalar product; there is no overall notion of distance and angle. In general, if we imagine the space that represents all possible outcomes of a blood test, this will form a manifold as each test can be fully described by a finite set of continuous quantities. In this space, there is no natural notion of distance and angles between directions, there is no notion of geometry. As we will see, the notion of integral, the idea of a quantity that can be understood as the sum of infinitely many infinitesimally small contributions, does not require either a particular number of dimensions or a notion of distance and angle.

Suppose we understood \vec{f} , \vec{B} and ρ not as above, but as maps that for each infinitesimal region return an infinitesimal contribution dW , $d\Phi$ and dm . We could simply write

$$\begin{aligned} W(\gamma) &= \int_{\gamma} dW = \int_{\gamma} f(d\gamma) \\ \Phi(\Sigma) &= \iint_{\Sigma} d\Phi = \iint_{\Sigma} B(d\Sigma) \\ m(V) &= \iiint_V dm = \iiint_V \rho(dV). \end{aligned}$$

This pattern is straightforward and more easily generalized. We call these functions of infinitesimal regions k -forms, where k is the dimensionality of the infinitesimal region they take as an argument. The force, in this notation, is a one-form (or covector) as it takes one dimensional infinitesimal regions (i.e. vectors); the magnetic field is a two-form; the density is a three-form. We can also say that a scalar field, like the temperature, is a zero-form, as it takes points, zero dimensional objects.

Since k -forms act over infinitesimal regions, they will have some key properties. First, note that each infinitesimal region can be understood as a parallelepiped, and a parallelepiped is fully identified by its sides. Therefore, a k -form can be understood as acting on a set of infinitesimal displacements, the sides of the parallelepiped, whose number matches the dimensionality of the form. A one-form will take one displacement, a two-form two displacements and so on. Second, as they are linear functions of the infinitesimal regions, they will also be linear functions of the vectors that define these infinitesimal regions. Lastly, all forms must be anti-symmetric because switching the order of the sides does not change the parallelepiped, but it changes its orientation.

We can write displacements and forms in terms of components and basis elements

$$\begin{aligned}
dP &= dx^i e_i \\
f &= f_i e^i \\
B &= B_{ij} e^i \otimes e^j \\
\rho &= \rho_{ijk} e^i \otimes e^j \otimes e^k.
\end{aligned} \tag{1.73}$$

The anti-symmetry means that switching the indexes introduces a minus sign. For example, $B_{ij} = -B_{ji}$ and $\rho_{ijk} = -\rho_{jik}$. Therefore, our integrals can be written as

$$\begin{aligned}
W(\gamma) &= \int_{\gamma} f_i dx^i \\
\Phi(\Sigma) &= \iint_{\Sigma} B_{ij} dx_1^i dx_2^j \\
m(V) &= \iiint_V \rho_{ijk} dx_1^i dx_2^j dx_3^k.
\end{aligned} \tag{1.74}$$

Each k -form, then, is a fully anti-symmetric covariant tensor, with one index for each dimension of the form. In this view, the B^x component of the magnetic field, then, becomes the B_{yz} component. That is, instead of being the component that gives us the magnetic flux along the x direction, it is the component that gives us the flux through the yz plane. Similarly, ρ_{xyz} is better understood not as the value of the mass density at a point, but rather the value of the density over the xyz volume. The indexes make unit dependency apparent as well: B_{yz} must be in units of magnetic flux divided by the units of y and z ; ρ_{xyz} must be in units of mass divided by the units of x , y and z ; in spherical coordinates, $B_{r\theta}$ must be in units of magnetic flux divided by units of r and θ . In fact, one can argue that these tools are here exactly to keep track of the units of the components.

We can go further, and rewrite the integrals in terms of parametrizations of the surfaces and the components of the forms along said parametrization. We have

$$\begin{aligned}
W(\gamma) &= \int_{\gamma} f_i dx^i = \int_{a_u}^{b_u} f_i \partial_u x^i du = \int_{a_u}^{b_u} f_u du \\
\Phi(\Sigma) &= \iint_{\Sigma} B_{ij} dx_1^i dx_2^j = \int_{a_u}^{b_u} \int_{a_v}^{b_v} B_{ij} \partial_u x_1^i du \partial_v x_2^j dv = \int_{a_u}^{b_u} \int_{a_v}^{b_v} B_{uv} dudv \\
m(V) &= \iiint_V \rho_{ijk} dx_1^i dx_2^j dx_3^k = \int_{a_u}^{b_u} \int_{a_v}^{b_v} \int_{a_w}^{b_w} \rho_{ijk} \partial_u x_1^i du \partial_v x_2^j dv \partial_w x_3^k dw \\
&= \int_{a_u}^{b_u} \int_{a_v}^{b_v} \int_{a_w}^{b_w} \rho_{uvw} dudvdw.
\end{aligned} \tag{1.75}$$

These expressions are useful to write the integrals in terms of an integral of k variables. This is particularly useful if the surfaces possess different symmetries than the forms, which allows one to use the parametrization appropriately.

We now have expressions that are easier to generalize. If we denote S^k the space of all the k -dimensional subregions of an n -dimensional manifold, a k -functional $F_k : S^k \rightarrow \mathbb{R}$ is an additive functional that takes a k -surface σ^k and returns a number. This can be expressed as

$$\begin{aligned}
F_k(\sigma^k) &= \int_{\sigma^k} \omega_k(d\sigma^k) = \int_{\sigma^k} \omega_{i_1 i_2 \dots i_k} dx_1^{i_1} dx_2^{i_2} \dots dx_k^{i_k} \\
&= \int_{a_{u_1}}^{b_{u_1}} \int_{a_{u_2}}^{b_{u_2}} \dots \int_{a_{u_k}}^{b_{u_k}} \omega_{u_1 u_2 \dots u_k} du_1 du_2 \dots du_k,
\end{aligned} \tag{1.76}$$

which is the integral of a k -form $\omega_k : V^k \rightarrow \mathbb{R}$. Here V is the space of vectors, of infinitesimal displacements, and an infinitesimal k -surface $d\sigma^k$ is defined by an ordered set of k vectors, therefore is an element from the Cartesian product $V^k = V \times V \times \dots \times V$. The k -form takes an infinitesimal k -surface $d\sigma^k$ and returns a number. We now have a way to express local objects and integrals in a generalized way.

In the spirit of reverse physics, we ask: what are the physical assumptions that are required to use these mathematical objects? Note that we do not measure mass density ρ directly: we measure mass m within a finite region V and the size of the finite region V , and we calculate the mass density by dividing the first by the second. The mass density ρ is the limit for which the region V shrinks to a point.²³ The same happens with the other quantities. We do not measure a force f directly, but rather the work that a force performs, for example, by deforming a spring. In other words, we measure the finite value of the functional F over a finite region σ , and the form ω returns the limit of F for infinitesimal regions. Note that it is crucial for F to be additive, that is, if σ_1 and σ_2 are two disjoint regions, $F(\sigma_1 \cup \sigma_2) = F(\sigma_1) + F(\sigma_2)$. Physically, it means that the contribution of one sub-region is independent from the other, because if this isn't the case, we cannot assign a unique value to each region. While this may seem a relatively harmless assumption, it may not hold in general. For example, suppose we have a system distributed in space. Given the mass-energy equivalence, the mass is an additive functional only if the interaction between the parts can be neglected. In fact, if the interaction energy at the boundary is so large that it is of the same scale of the mass within the regions, then it is not true that the total mass is simply the mass within each region. Therefore, if we write the mass density ρ , we have implicitly assumed that the interaction energy between parts can be neglected. While this is going to be true in a large number of cases, it is something we have to keep in mind. The functional F , in fact, is a more general physical object as it may exist even though infinitesimal additivity fails.

If the functional F respects the limits, in the sense that small variations of the region result in small variations of the value of the functional, then we can express the functional F as the integral of a form ω .²⁴ Note that the standard mathematical definitions are inverted with respect to what makes physical sense. Mathematically, we first define the form and then its integration, and we may worry whether the integral exists or diverges; physically, we first define the functional of finite regions and worry about the existence of an infinitesimal limit, the form, under additional physical assumptions. This is important because, as we said before, it tells us that it is the functional that survives when the assumption fails, not the form. It also means that, if we want to get a robust physical intuition, we should concentrate on the functionals, the finite objects, rather than the forms, the infinitesimal objects.

We therefore reach the following:

Insight 1.77. *Differential k -forms represent the infinitesimal contributions of an infinitesimally additive quantity F_k that depends on a k -dimensional surface.*

The issue, however, is that while we have a sense that, *mutatis mutandi*, an infinitesimally additive functional corresponds to a differential form, we would need to show that differential

²³This is essentially the definition of the Radon-Nikodym derivative used in measure theory.

²⁴Mathematically, one needs to be precise as to what these small variations are, and what the space of regions is. Effectively, we need to define what it means for k -surfaces to be differentiable. This is more in the scope of physical mathematics than reverse physics.

forms cannot represent anything else. We currently do not have such an argument, though we suspect it should exist. One may first note that no measurement can really be conducted at a point, and therefore it has to extend, at least in principle, over a k -dimensional region. Therefore we could argue that all measurements are functionals, additive or not. This means that we would need a map between the value of the components of the form at all points and the value of the functional over all regions. To reach our goal, we would need to show that this map can be established only if the functional is infinitesimally additive, which may not be a strong enough condition by itself. While we are already implicitly assuming the map to be differentiable, we could impose further requirements on the measurement functional. We could impose locality, in the sense that changes of the form within a region must change the value of the functional for said region; it would also mean that changes outside of the region do not affect measurements in the region. While we do not know whether these requirements are sufficient, these are the types of questions we would need to answer in order to have a fully physically meaningful treatment of differential forms.

Having generalized the idea of integration, let us now turn to differential operators such as gradient, curl and divergence. Suppose that T is a scalar field, like the temperature. Then we have the following relationship

$$\int_{\gamma} \vec{\nabla} T \cdot d\vec{\gamma} = T(B) - T(A). \quad (1.78)$$

That is, the integral of the gradient of T along γ gives us the difference of T evaluated at the endpoints, the boundary of the line. If \vec{f} is a vector field, like the force, we have

$$\iint_{\Sigma} \vec{\nabla} \times \vec{f} \cdot d\vec{\Sigma} = \int_{\gamma=\partial\Sigma} \vec{f} \cdot d\vec{\gamma} = W(\gamma) \quad (1.79)$$

That is, the integral of the curl of \vec{f} over a surface Σ is equal to the line integral of \vec{f} over the boundary. If \vec{B} is a pseudo-vector field, like the magnetic field, we have

$$\iiint_V \vec{\nabla} \cdot \vec{B} dV = \iint_{\Sigma=\partial V} \vec{B} \cdot d\vec{\Sigma} = \Phi(\Sigma). \quad (1.80)$$

That is, the integral of the divergence of \vec{B} over a region V is equal to the surface integral of \vec{B} over the boundary. Note the pattern: the integral of the differential operator applied to the bulk becomes the integral of the original object over the boundary.

Let us understand how this works in terms of functionals. Given the temperature $T(P)$, we can construct a line functional that for each line returns the difference in temperature at the endpoints. Given the work line functional $W(\gamma)$, we can construct a surface functional that for each surface returns the work needed to go around the contour. Given the magnetic flux surface functional $\Phi(\Sigma)$, we can construct a volume functional that for each volume returns the magnetic flux over the boundary. In general, suppose we have a k -functional $F_k : S^k \rightarrow \mathbb{R}$. To each k -surface σ^k we can associate a quantity $F_k(\sigma^k)$. Now, suppose we are given a $(k+1)$ -surface σ^{k+1} . While F_k cannot act on σ^{k+1} , the boundary $\partial\sigma^{k+1}$ is a k -surface, therefore we can evaluate $F_k(\partial\sigma^{k+1})$. Therefore, we can define the $(k+1)$ -functional ∂F_k such that $\partial F_k(\sigma^{k+1}) \mapsto F_k(\partial\sigma^{k+1})$, which we call the exterior functional of F_k .²⁵ Since the

²⁵Mathematically, we would have to prove that ∂F_k is a functional. Again, these mathematical details are left for the physical mathematics section. For now, we are more interested in the conceptual understanding.

exterior functional is an additive functional, it will have a corresponding form that acts on the infinitesimal region. Conceptually, the gradient of T is the one-form that corresponds to the boundary functional of T ; the curl of \vec{f} is the two-form that corresponds to the boundary functional of the work W , the line integral of \vec{f} ; the divergence of \vec{B} is the three-form that corresponds to the boundary functional of the magnetic flux Φ , the surface integral of \vec{B} . The problem is, again, that the gradient, curl and divergence are not expressed in a way that is easy to generalize.

Given that all three operators are in terms of ∇ , which in components is written ∂_i , we would like to write something along the lines of

$$\begin{aligned}\partial F_k(\sigma^{k+1}) &= \int_{\sigma^{k+1}} \partial_{i_0} \wedge \omega_{i_1 i_2 \dots i_k} dx^{i_0} dx^{i_1} dx^{i_2} \dots dx^{i_k} \\ &= \int_{\partial\sigma^{k+1}} \omega_{i_1 i_2 \dots i_k} dx^{i_1} dx^{i_2} \dots dx^{i_k} = F_k(\partial\sigma^{k+1}).\end{aligned}\tag{1.81}$$

The operation \wedge , which we call exterior product, must be such that we recover something consistent to the previous operations. That is, we need

$$\begin{aligned}\partial_i \wedge T &= \partial_i T \\ \partial_i \wedge F_j &= \partial_i F_j - \partial_j F_i \\ \partial_i \wedge B_{jk} &= \partial_i B_{jk} + \partial_j B_{ki} + \partial_k B_{ij}\end{aligned}\tag{1.82}$$

These would be the expressions we need to recover the gradient, curl and divergence respectively. Let us study them to see the pattern. First of all, each expression takes a k -form and returns a $(k+1)$ -form by adding a derivation along each index. Given that we have a derivation for each index, the number of terms matches the number of indexes of the final form, which is $k+1$. Each term changes index by taking a cyclic permutation. Recall that the forms are anti-symmetric, therefore each permutation of two indexes introduces a minus sign. A cyclic permutation of $k+1$ elements corresponds to k pair swaps. If $k+1$ is odd, then, each cyclic permutation will correspond to an even number of sign switches, which cancel out. The pattern, then, generalizes in the following way

$$\begin{aligned}\partial_{i_0} \wedge \omega_{i_1 i_2 \dots i_k} &= \partial_{i_0} \omega_{i_1 i_2 \dots i_k} + (-1)^k \partial_{i_1} \omega_{i_2 \dots i_k i_0} + (-1)^{2k} \partial_{i_2} \omega_{i_3 \dots i_k i_0 i_1} + \dots \\ &\quad + (-1)^{k \cdot k} \partial_{i_k} \omega_{i_0 i_1 \dots i_{k-1}} \\ &= \sum_{j=0}^k (-1)^{j \cdot k} \partial_{i_{j \bmod k+1}} \omega_{i_{j+1 \bmod k+1} i_{j+2 \bmod k+1} \dots i_{j+k \bmod k+1}}.\end{aligned}\tag{1.83}$$

The use of $\bmod k+1$ in the generalized expression makes sure that the index jumps from i_k to i_0 . This gives us a fully anti-symmetric tensor which matches the gradient, curl and divergence in the simple cases.²⁶ This operation is called the exterior derivative.

While we have the expression for the exterior derivative, we would like to understand why and how the expression works. Geometrically, we can imagine integrating along a parallelepiped, which becomes

$$\int_{\sigma^{k+1}} (\partial \wedge \omega_k)(d\sigma^{k+1}) = \int_{\sigma^{k+1}} \partial_{i_0} \wedge \omega_{i_1 i_2 \dots i_k} dx_0^{i_0} dx_1^{i_1} \dots dx_k^{i_k}$$

²⁶In principle, we could write the expression in different equivalent ways. Here we used cyclic permutations as it gives a more elegant expression.

$$\begin{aligned}
&= \int_{a_{u_0}}^{b_{u_0}} \int_{a_{u_1}}^{b_{u_1}} \cdots \int_{a_{u_k}}^{b_{u_k}} (\partial_{u_0} \omega_{u_1 u_2 \cdots u_k} + (-1)^k \partial_{u_1} \omega_{u_2 \cdots u_k u_0} \\
&+ \cdots + (-1)^{k \cdot k} \partial_{u_k} \omega_{u_0 u_1 \cdots u_{k-1}}) du_0 du_1 du_2 \cdots du_k \\
&= \int_{a_{u_1}}^{b_{u_1}} \cdots \int_{a_{u_k}}^{b_{u_k}} \int_{a_{u_0}}^{b_{u_0}} du_0 \partial_{u_0} \omega_{u_1 u_2 \cdots u_k} du_1 du_2 \cdots du_k \\
&+ (-1)^k \int_{a_{u_0}}^{b_{u_0}} \int_{a_{u_2}}^{b_{u_2}} \cdots \int_{a_{u_k}}^{b_{u_k}} \int_{a_{u_1}}^{b_{u_1}} du_1 \partial_{u_1} \omega_{u_2 \cdots u_k u_0} du_0 du_2 \cdots du_k \\
&+ \cdots + (-1)^{k \cdot k} \int_{a_{u_0}}^{b_{u_0}} \int_{a_{u_1}}^{b_{u_1}} \cdots \int_{a_{u_k}}^{b_{u_k}} du_k \partial_{u_k} \omega_{u_0 u_1 \cdots u_{k-1}} du_0 du_1 \cdots du_{k-1} \\
&= \left[\int_{a_{u_1}}^{b_{u_1}} \cdots \int_{a_{u_k}}^{b_{u_k}} \omega_{u_1 u_2 \cdots u_k} du_1 du_2 \cdots du_k \right]_{a_{u_0}}^{b_{u_0}} \\
&+ (-1)^k \left[\int_{a_{u_0}}^{b_{u_0}} \int_{a_{u_2}}^{b_{u_2}} \cdots \int_{a_{u_k}}^{b_{u_k}} \omega_{u_2 \cdots u_k u_0} du_0 du_2 \cdots du_k \right]_{a_{u_1}}^{b_{u_1}} \\
&+ \cdots + (-1)^{k \cdot k} \left[\int_{a_{u_0}}^{b_{u_0}} \int_{a_{u_1}}^{b_{u_1}} \cdots \int_{a_{u_{k-1}}}^{b_{u_{k-1}}} \omega_{u_0 u_1 \cdots u_{k-1}} du_0 du_1 \cdots du_{k-1} \right]_{a_{u_k}}^{b_{u_k}} \\
&= \int_{\partial \sigma^{k+1}} \omega_k (d\partial \sigma^{k+1}).
\end{aligned}$$

What happens is that each direction of integration will match a derivative in the same direction, and therefore will reduce to the integration of ω on opposing sides of the parallelepiped. This happens for each direction, and therefore the whole integral will reduce to the integration of ω on the surface of the parallelepiped. We have verified that equation 1.81, which is known as the generalized Stokes theorem, indeed works and it includes, as particular cases, the gradient theorem 1.78, the curl theorem 1.79 and the divergence theorem 1.80.

Another aspect of vector calculus is given by the following identities

$$\begin{aligned}
\vec{\nabla} \times \vec{\nabla} T &= 0 \\
\vec{\nabla} \cdot \vec{\nabla} \times \vec{f} &= 0.
\end{aligned} \tag{1.84}$$

To generalize them, we note that the exterior product \wedge is an anti-commutative and associative operation. Therefore we have

$$\partial_i \wedge \partial_j \wedge \omega_{l_1 l_2 \cdots l_k} = (\partial_i \partial_j - \partial_j \partial_i) \wedge \omega_{l_1 l_2 \cdots l_k} = 0 \wedge \omega_{l_1 l_2 \cdots l_k} = 0. \tag{1.85}$$

In other words, the exterior derivative applied twice returns zero, no matter on what form it is applied. Given that the curl is the exterior derivative applied to one forms and the gradient is the exterior derivative applied to zero forms, the fact that the curl of the gradient is zero is simply an application of the more general property. The same applies for the divergence of the curl. As we saw, mathematically the property is easy enough to verify, but we get no insight into its meaning. To understand the geometrical significance of the generalized relationship, recall that the exterior derivative of the form is associated with the exterior functional. We should, then, look at what happens when we construct the exterior functional of an exterior functional. We have

$$\partial \partial F_k(\sigma^{k+2}) = \partial F_k(\partial \sigma^{k+2}) = F_k(\partial \partial \sigma^{k+2}) = F_k(\emptyset) = 0. \tag{1.86}$$

In words, the exterior of the exterior functional of F_k equals F_k applied to the boundary of the boundary. However, the boundary of a boundary is always the empty set, and any functional applied to the empty set must be zero. In fact, since functionals are additive, we must have

$$F_k(\sigma^k) = F_k(\sigma^k \cup \emptyset) = F_k(\sigma^k) + F_k(\emptyset). \quad (1.87)$$

Therefore, we can see that the identity $\partial \wedge \partial \wedge \omega = 0$ for every form ω is ultimately a direct consequence that $\partial \partial U = \emptyset$ for every set U .²⁷ This shows how studying differential relationships in terms of finite functionals can give a more geometrically meaningful picture.

The last element of vector calculus we need to generalize is the idea of potentials. That is,

$$\begin{aligned} \vec{\nabla} \times \vec{f} = 0 &\Rightarrow \vec{f} = \vec{\nabla} V \\ \vec{\nabla} \cdot \vec{B} = 0 &\Rightarrow \vec{B} = \vec{\nabla} \times \vec{A}. \end{aligned} \quad (1.88)$$

These are generalized by the following formula

$$\partial_i \wedge \omega_{l_1 l_2 \dots l_k} = 0 \Rightarrow \omega_{l_1 l_2 \dots l_k} = \partial_{l_1} \wedge \theta_{l_2 \dots l_k}. \quad (1.89)$$

That is, if the exterior derivative of a k -form is zero, then there exists a $(k-1)$ -form whose exterior derivative is the original k -form.²⁸

To sum up, we have generalized the idea of integration over k -dimensional submanifolds of an n -dimensional space, which leads to the idea of k -functionals over finite regions and k -forms over the infinitesimal ones; we have seen that the finite functionals are physically more fundamental than the infinitesimal forms. We have seen how k -functionals induce $(k+1)$ -functionals by acting on the boundaries of the $k+1$ dimensional regions. We have seen that the exterior derivative gives us the form associated to the exterior functionals, and that this operation generalizes the notion of gradient, curl and divergence. While this does not exhaust all that can be done in differential topology and differential geometry, this is enough for what we need to use in the following sections.

Those already familiar with differential topology will have noticed that our notation and definitions do not quite match the ones typically used in math textbooks. The issue is that said notation and definitions do not match what we need to physically capture, and therefore it would be a mistake to employ them. Let us briefly see why. First of all, in the context of differential topology vectors are defined as directional derivatives. That is, a vector $v = v^i \partial_i$ is an operator that acts on scalar functions. A velocity, which in physics we think of as a vector, is not a directional derivative. In differential topology, a covector $\theta = \theta_i dx^i$ is defined

²⁷Note that we are talking about the boundary of a manifold, not the boundary of a set in the topological sense.

²⁸Technically, closed forms (i.e. those whose exterior derivative is zero) are not necessarily exact forms (i.e. those that are the exterior derivative of another form). This is true only on contractible regions (i.e. those regions that can be continuously shrunk to a point, that do not have holes). While this is a subtle mathematical point, we can understand it by looking at the corresponding functionals. An exact form corresponds to a functional that returns zero for any closed surface. A closed form, however, is guaranteed to return zero only on closed surfaces that are contractible, that can be continuously shrunk to a point. For example, the functional associated with a closed form may return non-zero over a closed surface that encloses a hole, something the functional associated with an exact form cannot do. Therefore, all exact forms are closed, but not all closed forms are exact. However, if we restrict ourselves to a contractible region, the two definitions are the same.

as a map from a vector to a scalar number. Conjugate momentum, whose components change as a covector, is not a map. In differential topology, a differential dx^i is a covector, and it is the exterior derivative of the coordinate x^i . This means that differentials are maps such that $dx^i(\partial_{x^j}) = \delta_j^i$. In physics, this is not how we think of differentials and integration. In fact, consider the expression $\int f_i dx^i$. To write it in terms of invariant objects, we would have $f = f_i e^i$ and $dx = dx^i e_i$, where e^i and e_i are the co-basis and basis respectively, and therefore $e^i(e_j) = \delta_j^i$. So we obtain $\int f(dx) = \int f_i e^i(dx^j e_j) = \int f_i dx^i = \int df$. Therefore, in physics, the differential dx is more properly thought of as a vector, where the dx^i are the contravariant components, while f as a map from a vector dx to the differential df , which is what is integrated. Therefore the differential is the vector, while the force is the covector. This is in contrast to the use in differential topology. Moreover, in differential topology there is a notion of a single tangent space where all vectors live, which is not compatible with the idea of units. Consider the basis ∂_i . Given that coordinates are expressed in different units, we cannot simply sum derivatives along different directions. For example, in polar coordinates ∂_r may have units of inverse meters while ∂_θ of inverse radians. Worst of all, a directional derivative is taken with respect to a parameter, which will also have units and physical dimension. For example, if v represents a velocity, the components v^i would also depend on the units of space and time. This would mean that units of vectors, the tangent space of a manifold, depend not only on the physical dimensions of the space, but also on the physical dimensions of all possible parameters along which we may want to define a directional derivative. This would mean that the tangent space is not definable only in terms of the units of the manifold itself, and therefore is not defined just in terms of the manifold itself.

The takeaway message here is the following: the mathematical tools we inherit from mathematics are not necessarily designed to capture the physical relationships we need to capture. Mathematicians only care about formal definitions, regardless of what, or if, they represent physically. In physics we do not have this luxury. If we want to have meaningful physical theories, which is ultimately the goal of reverse physics, we need to revisit the mathematical tools we use to formulate them.

1.7 Reversing Lagrangian mechanics

Now that we have a good geometric and physical feel for Hamiltonian mechanics, and that we have a general understanding of what the tools of differential topology describe physically, we will analyze Lagrangian mechanics more in detail. Conceptually, we already know that Lagrangian mechanics is Hamiltonian mechanics plus assumption [KE](#). We will see that the flow of states in phase space admits a vector potential and the Lagrangian is the scalar product between that potential and the displacement along the path. The principle of least action is geometrically equivalent to asking for paths that are always tangent to the displacement field. Moreover, the principle of least action is better understood as a property of Hamiltonian evolution, and assumption [KE](#) is only required to express the product between potential and displacement in terms of velocity instead of momentum.

Kinematic assumption revisited

As we saw in section [1.2](#), Lagrangian mechanics is the subset of Hamiltonian mechanics for which assumption [KE](#) is valid, which means Lagrangian mechanics is equivalent to assuming

DR, **IND** and **KE**. For the first two assumptions, we found a host of equivalent mathematical, geometric and physical formulations. Let's see what can we find for **KE**.

First of all, let us summarize the conditions we have already found. Hamilton's equations always impose that the velocity is a function of the state variables:

$$v^i = d_t q^i = \partial_{p_i} H. \quad (1.90)$$

At fixed position, the Jacobian of the transformation is therefore the Hessian of the Hamiltonian

$$\partial_{p_i} v^j = \partial_{p_i} \partial_{p_j} H. \quad (1.91)$$

Therefore condition

At every position, the relationship between momentum and velocity is invertible and differentiable (WKE-INV)

and condition

At every point, the Hessian of the Hamiltonian is non-singular (hyperregularity of H): $|\partial_{p_i} \partial_{p_j} H| \neq 0$ (WKE-HYP)

are equivalent to each other. The non-singularity of the Hessian can also be understood as strict monotonicity of the derivative along momentum. Therefore

The Hamiltonian is twice differentiable and concave (or convex) in momentum (WKE-CONC)

is another equivalent condition.

Note that the relationship between velocity and momentum is not just invertible, but differentiable as well. This may seem like an additional condition from **KE**, but recall that the dynamics is not in terms of points, but rather cells of phase space and density distributions. In order to express those geometric elements in terms of the kinematic variables, we must make sure that the Jacobian determinant of the transformation from state variables to kinematic variables is well-defined and non-zero. We have

$$|J| = \begin{vmatrix} \partial_{q^j} x^i & \partial_{p_j} x^i \\ \partial_{q^j} v^i & \partial_{p_j} v^i \end{vmatrix} = \begin{vmatrix} \delta_j^i & 0 \\ \partial_{q^j} v^i & \partial_{p_j} v^i \end{vmatrix} = |\delta_j^i| |\partial_{p_j} v^i| - |0| |\partial_{q^j} v^i| = |\partial_{p_j} v^i|. \quad (1.92)$$

The Jacobian determinant of the change of variables, then, coincides with the Jacobian determinant of the relationship between velocity and momentum at constant position. Therefore the following conditions are all equivalent to each other and to condition **WKE-INV**.

The Jacobian of the transformation between state variables and kinematic variables is non-singular. (WKE-NSIN)

Densities over phase space can be expressed in terms of position and velocity: $\rho(x^i, v^j) |J| = \rho(q^i, p_j)$. (WKE-DEN)

Areas and volumes in phase space can be expressed in kinematic variables: $dx^1 \dots dx^n dv^1 \dots dv^n = |J| dq^1 \dots dq^n dp_1 \dots dp_n$. (WKE-VOL)

The symplectic form ω_{ab} can be expressed in kinematic variables. (WKE-SYMP)

The displacement field S^a can be expressed in kinematic variables. (WKE-DISP)

The insight is that differentiability between state variables and kinematic variables is required to be able to express the objects we used to characterize the deterministic and reversible dynamics. The only physical requirement, then, is to be able to express the dynamics in terms of the kinematics, which is exactly what assumption **KE** already requires.

On the physical meaning of the Lagrangian

A key problem is to understand what the Lagrangian represents physically. It is often introduced as the difference between kinetic and potential energy. However, this does not work in general. Take the Lagrangian for a particle with charge q under an electromagnetic field with electric potential V and magnetic potential A_i .

$$L = \frac{1}{2}m|v^i|^2 + qv^i A_i - qV. \quad (1.93)$$

The first term is clearly kinetic energy, the last clearly potential energy, but what about the middle term? It depends on velocity, so it would appear to be a kinetic term, but it also depends on the potential. It's both and neither. The characterization of Lagrangian as difference between kinetic and potential energy, then, works for some systems but not in general and it is best abandoned.

Another problem in understanding what the Lagrangian represents is that it is not unique. Now, a similar problem exists for the Hamiltonian, in the sense that we can sum an arbitrary constant to any Hamiltonian without changing the equations of motion. However, the degeneracy for a Lagrangian is far worse. For example, let $f(x^i, t)$ be an arbitrary function of position and time. We can set

$$L' = L + \partial_{x^i} f(x^i, t)v^i + \partial_t f(x^i, t). \quad (1.94)$$

We can see how the Euler-Lagrange equations 1.3 are affected

$$\begin{aligned} 0 &= \partial_{x^i} L' - d_t \partial_{v^i} L' = \partial_{x^i} L + \partial_{x^i} \partial_{x^j} f v^j + \partial_{x^i} \partial_t f - d_t (\partial_{v^i} L + \partial_{x^i} f) \\ &= \partial_{x^i} L - d_t \partial_{v^i} L + \partial_{x^i} \partial_{x^j} f v^j + \partial_{x^i} \partial_t f - \partial_{x^j} \partial_{x^i} f d_t x^j - \partial_t \partial_{x^i} f d_t t \\ &= \partial_{x^i} L - d_t \partial_{v^i} L + \partial_{x^i} \partial_{x^j} f v^j + \partial_{x^i} \partial_t f - \partial_{x^j} \partial_{x^i} f v^j - \partial_t \partial_{x^i} f \\ &= \partial_{x^i} L - d_t \partial_{v^i} L. \end{aligned} \quad (1.95)$$

That is, the equations of motion given by L' are the same as the ones given by L . Therefore the actual value, or the difference of values, of the Lagrangian is physically meaningless. This makes the question even more puzzling: what is the Lagrangian?

The extended phase space

Given that we have a good understanding of Hamiltonian mechanics, let's work on the equations that link the two. We have

$$\begin{aligned} L &= p_i v^i - H \\ &= p_i d_t q^i - H d_t t \\ &= \begin{bmatrix} p_i & 0 & -H \end{bmatrix} \begin{bmatrix} d_t q^i \\ d_t p_i \\ d_t t \end{bmatrix}. \end{aligned} \quad (1.96)$$

The Lagrangian, then, can be understood as the scalar product of two vectors. The second one is the displacement along the path, however it is the displacement not just in position and momentum, but over time as well. The correct setting to understand Lagrangian mechanics,

then, is phase space extended by the time variable. If we redefine $\xi^a = [q^i \ p_i \ t]$ to include the time variable, we have

$$S^a = d_t \xi^a = [d_t q^i \ d_t p_i \ d_t t]. \quad (1.97)$$

We can then define the new vector θ_a in terms of its components

$$\theta_a = [p_i \ 0 \ -H]. \quad (1.98)$$

We also need to generalize ω_{ab} to the extended phase space. Looking at equation [HM-G](#), the idea is to put the gradient of the Hamiltonian in the time component. If we set

$$\omega_{ab} = \begin{bmatrix} \omega_{q^i q^j} & \omega_{q^i p_j} & \omega_{q^i t} \\ \omega_{p_i q^j} & \omega_{p_i p_j} & \omega_{p_i t} \\ \omega_{t q^j} & \omega_{t p_j} & \omega_{tt} \end{bmatrix} = \begin{bmatrix} 0 & \delta_j^i & \partial_{q^i} H \\ -\delta_i^j & 0 & \partial_{p_i} H \\ -\partial_{q^j} H & -\partial_{p_j} H & 0 \end{bmatrix}, \quad (\text{SF-EPS})$$

we have

$$\begin{aligned} S^a \omega_{a q^j} &= S^{q^i} \omega_{q^i q^j} + S^{p_i} \omega_{p_i q^j} + S^t \omega_{t q^j} \\ &= -S^{p_j} - S^t \partial_{q^j} H = -S^{p_j} - \partial_{q^j} H = 0 \\ S^a \omega_{a p_j} &= S^{q^i} \omega_{q^i p_j} + S^{p_i} \omega_{p_i p_j} + S^t \omega_{t p_j} \\ &= S^{q^j} - S^t \partial_{p_j} H = S^{q^j} - \partial_{p_j} H = 0 \\ S^a \omega_{a t} &= S^{q^i} \omega_{q^i t} + S^{p_i} \omega_{p_i t} + S^t \omega_{tt} \\ &= S^{q^i} \partial_{q^i} H + S^{p_i} \partial_{p_i} H \\ &= \partial_{p_i} H \partial_{q^i} H - \partial_{q^i} H \partial_{p_i} H = 0. \end{aligned} \quad (1.99)$$

This means that, on the phase space extended by time, Hamilton's equations become

$$S_a = S^b \omega_{ba} = 0. \quad (\text{HM-EPS})$$

Note that while the position and momentum components of S_a still perform a rotation, the time component does not. So we can't understand S_a as a rotated displacement. Recall that $v^a \omega_{ab} w^b = v_b w^b$ quantified the number of states in the parallelepiped formed by v^a and w^b . If $v_b w^b = 0$, then the parallelepiped does not identify states on an independent DOF. For each vector v^a , the covector v_a identifies the direction that forms an independent DOF with v^a . If we only have position and momentum, we can see that you get the direction rotated by ninety degrees along each DOF. If we extend phase space with time, time does not add a new independent degree of freedom. In fact, the direction given by the displacement field S^a should give us no new independent states: there should be no direction v^b in phase space such that $S^a \omega_{ab} v^b \neq 0$. In other words, $S_b = S^a \omega_{ab}$ must be zero and this is what equation [HM-EPS](#) says.

Potential of the flow - 1 DOF case

We wrote the Lagrangian $L = \theta_a d_t \xi^a$ in terms of θ_a and $d_t \xi^a$, but we have yet to understand what θ_a is. If we compare it to ω_{ab} , we note that the first has the Hamiltonian as a component,

while the second has its derivative. Given that ω_{ab} is anti-symmetric, it is a two-form, we may want to calculate the anti-symmetrized derivative of θ_a , the exterior derivative. We have

$$\begin{aligned}
\partial_a \theta_b - \partial_b \theta_a &= \begin{bmatrix} \partial_{q^i} \theta_{q^j} - \partial_{q^j} \theta_{q^i} & \partial_{q^i} \theta_{p_j} - \partial_{p_j} \theta_{q^i} & \partial_{q^i} \theta_t - \partial_t \theta_{q^i} \\ \partial_{p_i} \theta_{q^j} - \partial_{q^j} \theta_{p_i} & \partial_{p_i} \theta_{p_j} - \partial_{p_j} \theta_{p_i} & \partial_{p_i} \theta_t - \partial_t \theta_{p_i} \\ \partial_t \theta_{q^j} - \partial_{q^j} \theta_t & \partial_t \theta_{p_j} - \partial_{p_j} \theta_t & \partial_t \theta_t - \partial_t \theta_t \end{bmatrix} \\
&= \begin{bmatrix} \partial_{q^i} p_j - \partial_{q^j} p_i & \partial_{q^i} 0 - \partial_{p_j} p_i & \partial_{q^i} (-H) - \partial_t p_i \\ \partial_{p_i} p_j - \partial_{q^j} 0 & \partial_{p_i} 0 - \partial_{p_j} 0 & \partial_{p_i} (-H) - \partial_t 0 \\ \partial_t p_j - \partial_{q^j} (-H) & \partial_t 0 - \partial_{p_j} (-H) & \partial_t (-H) - \partial_t (-H) \end{bmatrix} \\
&= \begin{bmatrix} 0 - 0 & 0 - \delta_i^j & -\partial_{q^i} H - 0 \\ \delta_j^i - 0 & 0 - 0 & -\partial_{p_i} H - 0 \\ 0 + \partial_{q^j} H & 0 + \partial_{p_j} H & -\partial_t H + \partial_t H \end{bmatrix} \\
&= \begin{bmatrix} 0 & -\delta_i^j & -\partial_{q^i} H \\ \delta_j^i & 0 & -\partial_{p_i} H \\ \partial_{q^j} H & \partial_{p_j} H & 0 \end{bmatrix}.
\end{aligned} \tag{1.100}$$

This means that the form ω_{ab} is minus the exterior derivative of θ_a :

$$\begin{aligned}
\omega_{ab} &= -(\partial_a \theta_b - \partial_b \theta_a) = -\partial_a \wedge \theta_b \\
\theta_a &= [p_i \quad 0 \quad -H].
\end{aligned} \tag{1.101}$$

In other words, ω_{ab} has a null exterior derivative and θ_a is its potential.

Given that the extended phase space for a single degree of freedom is three dimensional, we can use standard vector calculus to gain more understanding. In this case, we have

$$\begin{aligned}
\theta_a &= [p \quad 0 \quad -H] \\
\omega_{ab} &= \begin{bmatrix} 0 & 1 & \partial_q H \\ -1 & 0 & \partial_p H \\ -\partial_q H & -\partial_p H & 0 \end{bmatrix} = \begin{bmatrix} 0 & S^t & -S^p \\ -S^t & 0 & S^q \\ S^p & -S^q & 0 \end{bmatrix} = \epsilon_{abc} S^c
\end{aligned} \tag{1.102}$$

where ϵ_{abc} is the fully anti-symmetric Levi-Civita symbol which returns the sign of the permutation of the variables (i.e. $\epsilon_{qpt} = \epsilon_{ptq} = \epsilon_{tqp} = 1$ while $\epsilon_{tpq} = \epsilon_{pqt} = \epsilon_{qtp} = -1$). We then have the following relationship

$$\begin{aligned}
\epsilon_{abc} S^c &= \omega_{ab} = -\partial_a \wedge \theta_b = -(\partial_a \theta_b - \partial_b \theta_a) \\
\vec{S} &= -\vec{\nabla} \times \vec{\theta}_b
\end{aligned} \tag{1.103}$$

Apart from the minus, this is the same relationship we would have between the magnetic field B^i and its vector potential A_i . Therefore the same concepts from vector calculus apply: S^a is divergenceless, admits a vector potential, and the flow of the displacement over a closed surface is zero. We can understand this as assumption DR implemented in the extended phase space. Moreover, we have the following relationship:

$$\int_{\Sigma} \omega_{ab} dx_1^a dx_2^b = \int_{\Sigma} \epsilon_{abc} S^c dx_1^a dx_2^b \tag{1.104}$$

That is, the integral of the form ω_{ab} corresponds to the surface integral of the displacement field S^a .²⁹ In phase space extended by time, then, ω_{ab} quantifies both the number of states over a surface and the flow of S^a through the surface. Under assumption DR this works because we have one and only one evolution for each state, therefore quantifying the number of evolutions that intersect a given surface is the same as quantifying the number of states that flow through them.

One striking feature of the vector potential θ_a is that it is fully characterized by a single arbitrary component, the time component $-H$. The magnetic vector potential A_i , instead, is characterized by two arbitrary components. What are the exact physical conditions that allow that to happen?

Suppose that S^a is a divergenceless field, then we can write it as minus the curl of a potential

$$\theta_a = [\theta_q \quad \theta_p \quad \theta_t]. \quad (1.105)$$

Vector potentials are defined up to the gradient of an arbitrary function (i.e. up to a gauge), since $\nabla \times (\theta + \nabla f) = \nabla \times \theta$. We can choose f such that $\partial_p f = -\theta_p$, and therefore we can set, without loss of generality, the momentum component to zero

$$\theta_a = [\theta_q \quad 0 \quad \theta_t]. \quad (1.106)$$

So far, this procedure can be applied to any potential of a divergenceless field. However, the displacement field S^a is particular in that its time component $S^t = d_t t = 1$ is unitary. That is, states flow at a uniform rate in time. Therefore we must have

$$S^t = -(\partial_q \theta_p - \partial_p \theta_q) = \partial_p \theta_q = 1. \quad (1.107)$$

Integrating the relationship we find that $\theta_q = p + c(q, t)$, where $c(q, t)$ is an arbitrary function. We can choose $c(q, t) = 0$ without loss of generality since that arbitrariness corresponds to a choice of gauge. In fact, we can choose $f(q, t)$ such that $\partial_q f = -c(q, t)$. Given that $\partial_p f(q, t) = 0$, this will not impact the previous arbitrary choice of $\theta_p = 0$. We have

$$\theta_a = [p \quad 0 \quad \theta_t]. \quad (1.108)$$

At this point, we simply rename the last component to $-H$ and have

$$\theta_a = [p \quad 0 \quad -H]. \quad (1.109)$$

The form of the vector potential, then, is set by the fact that the flow of S^a is both divergenceless and uniform along the time direction.

This discussion tells us exactly what θ_a is in the case of a single degree of freedom: it is the potential of the displacement field S^a or, equivalently, of the form ω_{ab} . Since we are in the single degree of freedom case, assumption DR is enough to recover Hamiltonian mechanics, and this is the same in the extended-phase-space formulation.

²⁹Note that, strictly speaking, while ω_{ab} is a tensor, ϵ_{abc} and S^a are not. Relationship 1.103, then, is valid only if we chose position, momentum and time as variables. Yet, as we will see much later, we can change the position variable and time variable and still have the same expression by redefining momentum and energy appropriately.

We also learned another thing: the Hamiltonian is the time component of the potential, therefore it is not a scalar. That is, if we made a change in the time variable $\hat{t} = \hat{t}(t)$, we would have

$$\hat{H} = -\theta_{\hat{t}} = -d_{\hat{t}}t \theta_t = d_{\hat{t}}t H. \quad (1.110)$$

It transforms like a scalar only under purely spatial change of variables.

Lagrangian and action

Now that we have seen what θ_a is, we can go back to the Lagrangian and the action. Suppose that we have a path γ , not necessarily an actual evolution, that proceeds along the time variable. That is, we can write

$$\begin{aligned} \gamma &= [q^i(t), p_i(t), t] \\ d\gamma &= d\xi^a = d_t \xi^a dt. \end{aligned} \quad (1.111)$$

where we used the time variable as its affine parameter. Note that, in this case, the displacement is along the generic path γ , which is not necessarily an actual evolution. Therefore $d_t \xi^a = S^a$ if and only if γ is an actual evolution of the system. For a generic path we can write:

$$\begin{aligned} L &= p_i v^i - H = \theta_a d_t \xi^a \\ \mathcal{A}[\gamma] &= \int_{\gamma} L dt = \int_{\gamma} \theta_a d_t \xi^a dt = \int_{\gamma} \theta_a d\xi^a = \int_{\gamma} \theta d\gamma \end{aligned} \quad (1.112)$$

This tells us that

$$\begin{aligned} &\text{The action along a path } \gamma \text{ is the line integral of the vector potential } \theta_a \text{ of} \\ &\text{the form } \omega_{ab}. \end{aligned} \quad (1.113)$$

We now have a precise geometric characterization of the action and of the Lagrangian. As for the physics, it tells us conclusively that both the Lagrangian and the action, by themselves, are unphysical. That is, the value of the Lagrangian at a given position, velocity and time, or the value of the action for a given path is not a physically meaningful value. The vector potential is not a physical quantity: it depends on a choice of gauge, which does not correspond to a physically well-defined object.

We are now in the position to understand why, as we have seen before, the Lagrangian for a given system is not unique: it depends on the vector potential, and the vector potential is not uniquely defined. We can see how the change of Lagrangian 1.94 corresponds to redefining the vector potential as

$$\theta'_a = \theta_a + \partial_a f(q^i, t). \quad (1.114)$$

We have

$$\begin{aligned} L &= \theta_a d_t \xi^a \\ L' &= \theta'_a d_t \xi^a = \theta_a d_t \xi^a + \partial_a f d_t \xi^a = L + \partial_{q^i} f d_t q^i + \partial_t f d_t t \\ &= L + \partial_{x^i} f(x^i, t) v^i + \partial_t f(x^i, t). \end{aligned} \quad (1.115)$$

which is the same expression we had in 1.94.

But if the action and the Lagrangian are unphysical, how can we use them to derive the laws of evolution, which are clearly physical? Note that the laws are not expressed in terms of the action, but in terms of the variation of the action. Let γ' , then, be a small variation of the path γ with the same endpoints. Note that γ and γ' form a closed loop. Take a surface Σ that is enclosed by that boundary, we can use Stokes' theorem:

$$\begin{aligned}\delta\mathcal{A}[\gamma] &= \delta \int_{\gamma} L dt = \delta \int_{\gamma} \theta_a d\xi^a = \int_{\gamma} \theta_a d\xi^a - \int_{\gamma'} \theta_a d\xi^a = \oint_{\partial\Sigma} \theta_a d\xi^a = \int_{\Sigma} \partial_a \wedge \theta_b d\xi^a d\eta^b \\ &= - \int_{\Sigma} \omega_{ab} d\xi^a d\eta^b,\end{aligned}\tag{1.116}$$

where $d\eta^b$ is the displacement of a point from γ to the corresponding point on the variation γ' . The variation of the action, then, corresponds to the surface integral of ω_{ab} , which is physical. Geometrically, it corresponds to the flow of the evolutions through the surface enclosed by the path and its variation. Note that, because the flow is divergenceless, it does not matter which surface is chosen, since all surfaces that share the same boundaries will correspond to the same flow. This is a striking observation: while the action is unphysical, its variation is physical.

We now are in a perfect position to fully understand the principle of stationary action. This states that actual evolutions are given by paths for which the variation is zero. That is, $\int_{\Sigma} \omega_{ab} d\xi^a d\eta^b$ has to be zero for all possible $d\eta^b$. This means we must have $\omega_{ab} d\xi^a = 0$. Since S^a is the only degenerate direction for ω_{ab} , we must have $d\xi^a = S^a dt$. In other words, the paths that make the action stationary are exactly the paths whose tangent vector applied to ω_{ab} return zero.

Geometrically, if a path γ is an evolution, it will always be tangent to the displacement field S^a . If we make a small variation γ' , the two paths will enclose an infinitesimal strip Σ which will be tangent to S^a as well. Therefore the flow of S^a through Σ will be zero. Given that the flow through Σ is minus the variation of the action, the variation of the action for all evolutions will be zero. Conversely, if a path is not an evolution, at some point its tangent will be different from the displacement field. Therefore we will be able to find a variation γ' for which the surface Σ will “catch” some flow. That is, we will have a variation of the path for which the variation of the action is non-zero. In other words, the action principle is a geometrically roundabout way to ask for those paths that are always tangent to the displacement field S^a , which is divergenceless and has constant flow in time.

Note that the geometric and physical interpretation we have given to the Lagrangian and the action lives in the extended phase space, which is part of the Hamiltonian formulation. In fact, the only place where we need assumption [KE](#) is when we want to write $p_i d_t q^i - H$ as a function of velocity instead of momentum. The integral of the vector potential and its variation, in fact, can still be written even in terms of momentum, and therefore the geometric and physical characterization of the principle of stationary action applies for all Hamiltonian systems, not just the Lagrangian subset. In other words, the principle of stationary action is really more a property of Hamiltonian mechanics than Lagrangian mechanics. The only difference is that in Lagrangian mechanics, since assumption [KE](#) applies, it can be expressed purely in terms of kinematic variables.

Multiple DOFs

The above discussion captures all the important elements of Lagrangian mechanics and the action principle even if it is limited to the single DOF case. Still, we should generalize to the

multiple DOFs case. Note that the equations we wrote for the Lagrangian and the action are already in generalized form. The only thing that we need to do is derive the expression for the vector potential θ_a in the general case.

If we compare **SF-EPS** with **SF-ND**, we find that the non-temporal components of the form are identical. Therefore, we can still understand it as quantifying the number of configurations within each DOF and the degree of independence across DOFs, while adding the idea that the displacement field does not contribute new configurations. Mathematically, the displacement field identifies the only direction in which the form ω_{ab} is degenerate. Let us see what the non-temporal components tell us in terms of the potential. We have:

$$\begin{aligned}\omega(e^{q^i}, e^{p_j}) &= (-\partial\theta)_{q^i p_j} = -(\partial_{q^i}\theta_{p_j} - \partial_{p_j}\theta_{q^i}) = \delta_j^i \\ \omega(e^{q^i}, e^{q^j}) &= (-\partial\theta)_{q^i q^j} = -(\partial_{q^i}\theta_{q^j} - \partial_{q^j}\theta_{q^i}) = 0 \\ \omega(e^{p_i}, e^{p_j}) &= (-\partial\theta)_{p_i p_j} = -(\partial_{p_i}\theta_{p_j} - \partial_{p_j}\theta_{p_i}) = 0\end{aligned}\tag{1.117}$$

We can use our gauge freedom to set $\theta_{p_1} = 0$, much in the same way we did before. We now have $\partial_{q^1}\theta_{p_1} = 0$ and, by the first condition, $\partial_{p_1}\theta_{q^1} = 1$. Integrating, we have $\theta_{q^1} = p_1 + g(q^i, p_2, p_3, \dots, t)$ where g is an arbitrary function which we can set to zero, since it corresponds to a choice of gauge. Therefore we have:

$$\theta_a = [p_1 \quad \theta_{q^2} \quad \dots \quad \theta_{q^n} \quad 0 \quad \theta_{p_2} \quad \dots \quad \theta_{p_n} \quad \theta_t].\tag{1.118}$$

Note that the components for the first degree of freedom do not depend on the other degrees of freedom. That is, for all $i > 1$, $\partial_{q^i}\theta_{q^1} = \partial_{p_i}\theta_{q^1} = \partial_{q^i}\theta_{p_1} = \partial_{p_i}\theta_{p_1} = 0$. But by using conditions 1.117, we find that the converse is true as well: the components of all other degrees of freedom do not depend on the first. That is, for all $i > 1$, $\partial_{q^1}\theta_{q^i} = \partial_{p_1}\theta_{q^i} = \partial_{q^1}\theta_{p_i} = \partial_{p_1}\theta_{p_i} = 0$.

We can then use, again, our gauge freedom with a function that does not depend on the first two variables to set $\theta_{p_2} = 0$. And, with the same reasoning, we will be able to set $\theta_{q^2} = p_2$. And then, again, find that the first two degrees of freedom do not depend on the others, etc. This will exhaust all DOFs and we can set $\theta_t = -H$ as before. This will find 1.101.

To recap, we have found that the expression of **SF-EPS** are equivalent to assuming **DR** and **IND** and is also equivalent to 1.101. Note that the expression of ω_{ab} and θ_a are in terms of specific coordinates, while the fact that ω_{ab} is the exterior derivative of θ_a , instead, is coordinate independent. However, there is a characterization of ω_{ab} that is fully coordinate independent.

Since ω_{ab} admits a potential, its exterior derivative is zero. We also have that the displacement field identifies the only degenerate direction. That is, if $v^a\omega_{ab} = 0$ then $v^a = fS^a$ for some scalar function f . Physically, this corresponds to saying that temporal displacement is the only direction that does not provide independent configurations: as we said before, time does not provide new possible configurations for the system. It turns out that these are enough conditions to find canonical coordinates $[q^i, p_i, t]$ such that we can express ω_{ab} as **SF-EPS**.³⁰ Therefore

The two-form ω_{ab} that quantifies independent configurations is closed
(i.e. has zero exterior derivative) and its only direction of degeneracy (DI-SYME)
is identified by the displacement field S^a

³⁰This result is known as Darboux's theorem. Unfortunately we haven't been able to find a proof that is short and/or physically significant, though we suspect such a proof should exist.

is an equivalent characterization of Hamiltonian mechanics in the extended phase space.

The above condition must imply both **DR** and **IND**. We can break down the two contributions in the following way. Suppose we have a system with n independent degrees of freedom, meaning the extended phase space is of dimension $N = 2n + 1$. Assumption **DR** tells us we have a way to measure the flow of the evolutions over a hyper-surface, which corresponds to an $(N - 1)$ -form $\Omega_{a_1 \dots a_{2n}}$ such that:

$$S^{a_1} \Omega_{a_1 \dots a_{2n}} = 0. \quad (1.119)$$

This is just stating that, given that $\Omega_{a_1 \dots a_{2n}}$ measures the flow through an infinitesimal $2n$ -dimensional parallelepiped, if one of the sides is along the direction of flow S^a then the flow will be parallel to the parallelepiped. If the space is charted by position, momentum and time, we can write:

$$\begin{aligned} \Omega_{a_1 \dots a_{2n}} &= \epsilon_{a_1 \dots a_{2n} a_{2n+1}} S^{a_{2n+1}} \\ \int \epsilon_{a_1 \dots a_{2n} a_{2n+1}} S^{a_{2n+1}} d\xi^{a_1} \dots d\xi^{a_{2n}} &= \int \Omega_{a_1 \dots a_{2n}} d\xi^{a_1} \dots d\xi^{a_{2n}}. \end{aligned} \quad (1.120)$$

On the right side, we see that we are really integrating the flow S^a through the hypersurface defined by the differentials $d\xi^{a_i}$, which corresponds to the integral of $\Omega_{a_1 \dots a_{2n}}$ through the hypersurface. Note that, under **DR**, we have one and only one evolution for each state, therefore measuring the flow of evolutions is equivalent to measuring the number of states that travel through the surface.

As we said, this setup was motivated by assumption **DR** alone. If we add **IND**, then we can write:

$$\Omega_{a_1 \dots a_{2n}} = \omega_{a_1 a_2} \wedge \omega_{a_3 a_4} \wedge \dots \wedge \omega_{a_{2n-1} a_{2n}}. \quad (1.121)$$

This tells us that the total state count becomes the product of the configurations along each degree of freedom.

The geometric and physical understanding of the action and the Lagrangian are the same as in the single DOF case: the action is the line integral of the vector potential θ_a and the variation of the action is minus the surface integral of ω_{ab} between the path and its variation. The only difference is that this corresponds not to the flow of total states, but to the flow of configurations for a single degree of freedom. Still, the flow is zero only if the path is everywhere tangent to the displacement field S^a , and therefore if the path is an actual evolution of the system.

1.8 Full kinematic equivalence and massive particles

In the previous section we recovered Lagrangian mechanics by formulating the principle of stationary action in Hamiltonian form, and then used assumption **KE** in form **WKE-INV** simply to express it in terms of the kinematic variables. This is the weakest use of the assumption, as it only assumes we have an invertible map between velocity and momentum. However, we will find that this is not enough to define a meaningful count of configurations over kinematic variables, which requires states to be uniformly distributed over velocity. This leads to a stronger version of assumption **KE**, which requires the map between velocity and conjugate momentum to be linear, which in turn fixes the dynamics to the one of massive particles under scalar and vector potential forces.

Linear map between linear spaces

Condition **WKE-INV** imposes, at every position, an invertible relationship $p_i = p_i(q^j, v^j)$ between conjugate momentum and velocity which can be arbitrary. Note that both velocity and momentum are linear objects, and therefore it would seem natural to require a linear relationship between the two. Though we do not have, at this point, a clear physical reason to impose this restriction, let's see what it would entail.

If the relationship between momentum and velocity at every point is linear, we have that the Jacobian is only a function of position. Moreover, since it is a linear map between a vector and a covector, the map must be a tensor. We have

$$\partial_{v^i} p_j = m g_{ij} \quad (1.122)$$

where m is a constant that transforms units of velocity to units of momentum while g_{ij} is the actual linear map in terms of spatial coordinates. Since $p_j = \partial_{v^j} L$, we have

$$m g_{ij} = \partial_{v^i} p_j = \partial_{v^i} \partial_{v^j} L = \partial_{v^j} \partial_{v^i} L = \partial_{v^j} p_i = m g_{ji} \quad (1.123)$$

and find that the tensor g_{ij} is symmetric.

If we integrate equation 1.122 we have

$$p_i = m g_{ij} v^j + q A_i \quad (1.124)$$

where A_i are arbitrary functions of position and q is an arbitrary constant. Since g_{ij} and A_i are functions of position only, we can see that, if we fix position, the relationship is a line where $m g_{ij}$ is the slope and A_i the value of momentum for zero velocity. Note that

$$v^i = d_t q^i = \partial_{p_i} H = \frac{1}{m} g^{ij} (p_j - q A_j). \quad (1.125)$$

If we integrate again we find

$$H = \frac{1}{2m} (p_i - q A_i) g^{ij} (p_j - q A_j) + V \quad (1.126)$$

where V is yet another arbitrary function. This is exactly the Hamiltonian for a massive particle under scalar and vector potential forces: m is the mass, g_{ij} is the metric tensor, q is the charge, A_i is a vector potential and V is the scalar potential.³¹ Therefore, we saw that imposing a linear relationship between velocity and conjugate momentum gives us massive particles under potential forces.

Conversely, if we start by imposing the above Hamiltonian for massive particles under potential forces, we find a linear relationship between conjugate momentum and velocity. Therefore condition

$$\text{There is a linear relationship between conjugate momentum and velocity} \quad (\text{FKE-LIN})$$

and condition

$$\text{The system under study is a massive particle under scalar and vector potential forces} \quad (\text{FKE-POT})$$

³¹Note that we have not separated the charge from the potentials.

are equivalent. At this point, having seen enough of reverse physics, it should be clear that this can't simply be a coincidence and it warrants more exploration.

Recall, that in 1.92 we saw that the Jacobian determinant of the transformation from state variables to kinematic variables was equal to the Jacobian determinant of the relationship between velocity and momentum. Since mg_{ij} is the Jacobian between velocity and momentum, we find

$$\begin{aligned}\partial_{p_i} v^j &= \frac{1}{m} g^{ij} \\ |J| = |\partial_{p_i} v^j| &= \frac{|g^{ij}|}{m} = \frac{|g_{ij}|^{-1}}{m} = \frac{1}{m |g_{ij}|}.\end{aligned}\tag{1.127}$$

Recall that the existence of a non-singular Jacobian is what allowed us to express all differential objects (e.g. densities, phase-space areas/volumes, the symplectic form) in terms of kinematic variables. This forces those expressions to be in terms of position only. We therefore have that the following conditions are equivalent to each other and equivalent to [FKE-LIN](#) and [FKE-POT](#).

The Jacobian of the transformation between state variables and kinematic variables is a non-singular function of position only. (FKE-NSIN)

Densities over phase space can be expressed in terms of position and velocity by rescaling the value at each point: $\rho(x^i, v^j) |J(x^i)| = \rho(q^i, p_j)$. (FKE-DEN)

Areas and volumes in phase space can be expressed in kinematic variables, and the transformation depends on position only: $dx^1 \cdots dx^n dv^1 \cdots dv^n = |J(x^i)| dq^1 \cdots dq^n dp_1 \cdots dp_n$. (FKE-VOL)

The symplectic form ω_{ab} can be expressed in kinematic variables, and its components are a linear function of velocity. (FKE-SYMP)

To understand better these conditions, consider a density ρ_{qp} over phase space and its expression ρ_{xv} over kinematic variables. Since there is a factor of a Jacobian determinant between the two, if the density takes the same value between two states as expressed in position and momentum, the density as defined over position and velocity will not necessarily have the same value. Position and momentum have a special property in that the density of states is uniform in terms of those variables, and therefore comparing densities in those variables is simply a matter of comparing the value of the density. Canonical coordinates are exactly those state variables for which this property holds. While kinematic variables are, in general, not canonical, the linearity between velocity and momentum imposes a lesser version of this property: it makes it so that, at least at the same position, we can compare areas and densities. Therefore if density for two different values of velocity at the same point matches, $\rho_{xv}(x^i, v_1^j) = \rho_{xv}(x^i, v_2^j)$, then we really have the same density of states. Therefore

Density expressed in velocity at the same position is proportional to the density over states (FKE-PROP)

is another equivalent condition.

In this formulation, one starts to wonder whether the lack of this property would make physical sense. We can understand that non-linear transformations of position, since they stretch and shrink space differently at different points, would make it more complicated to

understand whether two densities at two different points are the same. On the other hand, velocity is defined locally over infinitesimal changes of position, therefore non-linear changes in velocity do not arise when changing units and coordinates.

Mass and inertial frames

Another way of looking at it is the following. Suppose we have a density $\rho(q^i, p_j)$ defined over a finite region of phase space that is constant along momentum. That is, $\rho(q^i, p_j^1) = \rho(q^i, p_j^2)$ for all p_j^1 and p_j^2 . Assuming we have a linear relationship between momentum and velocity we have

$$\begin{aligned} \rho_{xv}(x^i, v_1^j) &= m|g_{ij}|\rho_{qp}(q^i, mg_{ij}v_1^j + \mathbf{q}A_i) = m|g_{ij}|\rho_{qp}(q^i, mg_{ij}v_2^j + \mathbf{q}A_i) \\ &= \rho_{xv}(x^i, v_2^j). \end{aligned} \tag{1.128}$$

Therefore if we have a linear map between velocity and momentum, uniform distributions along momentum correspond to uniform distributions along velocity. The converse is also true: if uniform distributions along momentum correspond to uniform distributions along velocity, then, the Jacobian of the transformation cannot depend on either, and it must depend on position only. Therefore

Uniform distributions along momentum correspond to uniform distributions along velocity (FKE-UNIF)

is another equivalent characterization of [FKE-LIN](#). Failure of this condition, then, would mean that physical states are not uniformly distributed over velocity at the same position, and some velocities correspond to higher state densities than others. This does not sound like something that would mesh well with the principle of relativity.

We can go one step further. Suppose we can find a coordinate system for which g_{ij} over a finite region is the identity matrix δ_{ij} . Physically, this corresponds to a Cartesian coordinate system in an inertial frame. In this frame, $|g_{ij}| = 1$ and therefore densities over kinematic variables differ from densities over phase space by a constant of proportionality: the mass. We have the following characterization of mass

Inertial mass tells us how many states there are per unit of area of position-velocity in an inertial (Cartesian) frame. (IM)

This characterization of mass works even for massless particles, while the standard one does not. In fact, if we define mass to be the resistance of a body to acceleration, we would expect a zero mass body to be extremely easy to accelerate, while this is not the case: a massless particle travels at the same constant speed. On the other hand, as we saw before, massless particles do not satisfy [KE](#), and therefore the state cannot be reconstructed from position and velocity: areas of position-velocity correspond to zero states. So why are more massive bodies harder to accelerate? The more massive the body, the more states per unit of velocity, and therefore reaching the same velocity means going through more states. If we understand force as a change of state per unit time, then we have an intuitive explanation that matches the new characterization.

Note that we focused on Cartesian coordinates, not just inertial frames. That is not a coincidence: the motion for a free particles appears linear and uniform only in these coordinates. That is, trajectories obey the linear expression $x^i = v^i t + x_0^i$ only in Cartesian coordinates. The

first law of Newton, the fact that in an inertial frame a body travels in linear and uniform motion, implicitly implies the ability to use Cartesian coordinates. But, in light of what we have seen before, it also implies the existence of mass. In fact, we can, in true reverse physics spirit, turn this on its head. Suppose that the motion of free particles is linear and uniform. The velocity will remain the same along an evolution; the distance between two particles moving at the same velocity along the same line will also remain the same. Therefore consider the plane charted by position and velocity along a particular direction. Take a parallelogram where two sides are at constant velocity: its area would be the length of those sides Δx times the height, which will be given by the difference in velocity Δv . As time evolves, the sides at constant velocity will move but will remain at the same velocity. The parallelogram will remain a parallelogram with the same height and base: the area is conserved. Now, clearly this is a deterministic and reversible system. Position and velocity identify all states and already provide an area that is conserved over time. It must be, then, that the count of states is proportional to the area in position-velocity, and therefore we have the relationship $p_i = mv^i$. In other words, once we assumed the existence of inertial frames, we already implicitly introduced the idea of inertial mass as characterized above.

The above discussion links the linearity between velocity and momentum to the existence of inertial frames. Therefore

At each position, there exists a local inertial frame (FKE-INER)

is equivalent to [FKE-LIN](#). Note that here we are claiming that only local inertial frames are needed. To see that, note that a change of position coordinates fully determines the change of both position and velocity. Therefore a change of coordinate can neither introduce nor remove a velocity dependency from the expression $\partial_{v^i} p_j$. If $\partial_{v^i} p_j = mg_{ij}$ has no velocity dependency, we can always find local spatial coordinates such that $g_{ij} = \delta_{ij}$. Therefore, $\partial_{v^i} p_j$ has no velocity dependency if and only if we can find a local inertial (Cartesian) frame.

So we found that [FKE-LIN](#) is linked to inertia and the existence of inertial frames, but in what sense is it linked to assumption [KE](#)? Suppose one gave us all the kinematics of a system, which are fully defined by the reference frame and all possible trajectories. Velocity, as a quantity, is fully defined in terms of position and time. Now, assumption [KE](#) tells us we must be able to recover the full dynamics. We must be able, for example, to convert a distribution over kinematic variables to one over state variables. This is essentially a change of units of the density ρ from position-velocity to position-momentum. But if velocity is a derived quantity, fully defined by position, then it should be the case that the transformation rule of ρ is only a function of position. Therefore condition

The position fully defines the units of all state variables, therefore (FKE-UNIT)
an invertible transformation between momentum and velocity

is an equivalent characterization of [FKE-LIN](#). Ultimately, this must be the case if [KE](#) holds: trajectories in space-time are fully specified by units of position, and since we must be able to reconstruct states from them, so the units of momentum must depend on them.

While this argument works on physics grounds, we cannot carry it out in a mathematically precise way. The issue is that current mathematical structures are devoid of the concept of units and of the type of definitional dependence we used. What does it mean, in a precise way, that velocity is a derived unit from position and time? What does it mean that temperature is

a derived unit from energy and entropy? Covariance and contravariance capture some hint in terms of change of units (i.e. how do units of velocity (or momentum) change if I change units of position?), but this is sufficient only when some initial relationship is assumed. Capturing this initial relationship requires conceptual and mathematical tools that are not currently available.

1.9 Relativistic mechanics

In this section we analyze what happens if we consider change of coordinates that mix space and time. We will see that relativistic elements appear even if a notion of metric tensor is not introduced, and the particle/anti-particle duality emerges. Moreover, under the full kinematic equivalence special relativity is recovered.

Hamiltonian mechanics on the extended phase space

The Hamiltonian formalism is invariant under generic change of spatial variables, but if we want to introduce generic changes of variables that mix space and time we have a problem. In the standard formulation, time is the parameter of the evolution $q^i(t)$, meaning that we can mix the spatial variables q^i while leaving t alone. In the formulation extended by time, time has a double role of parameter and variable. That is, we write $q^i(t)$ and $t(t)$. Therefore mixing q^i with t will affect the evolution parameter as well. What we need to do is to separate the time variable from the parameter of the evolution. We can group the space-time variables as

$$q^\alpha = [t, q^i]. \quad (1.129)$$

We introduce an affine parameter s , and therefore a trajectory in space-time will be noted as $q^\alpha(s)$. Under a change of coordinate, the variables q^α will mix but the affine parameter s will not change.

Now we have understood how to deal with time, we have to understand how to deal with energy. As we saw before, the Hamiltonian is no longer invariant under transformations that affect time. Moreover, even if we start with a time-independent Hamiltonian, it will not remain so under a generic coordinate transformation that mixes time and space. This means that, during the evolution, energy will not be constant but will need to increase and decrease. Another, more physical, way to look at it is that energy, like momentum, is not an absolute quantity but rather a relative quantity with respect to an observer. If we imagine an observer that is accelerating and decelerating, in the same way that he will see momentum change, he will see the energy change as well. In the same way that momentum is sensitive to changes of units along the corresponding spatial direction, energy is sensitive to change of time units. To be able to characterize these relationships better, we introduce an energy variable E which we group with the momentum variables and, since $-H$ was the time component of the potential θ_a , we write

$$p_\alpha = [-E, p_i] \quad (1.130)$$

so that p_α is a covector.

What we did is add a temporal degree of freedom, and with that addition, the equations look a lot like standard Hamiltonian mechanics for multiple degrees of freedom. We have

$$\begin{aligned}\xi^a &= [q^\alpha \ p_\alpha] \\ S^a &= d_s \xi^a = [d_s q^\alpha \ d_s p_\alpha]\end{aligned}\tag{1.131}$$

We define the potential θ_a ³² and the form ω_{ab}

$$\begin{aligned}\theta_a &= [p_\alpha \ 0] \\ \omega_{ab} &= \partial_a \wedge \theta_b = \begin{bmatrix} \omega_{q^\alpha q^\beta} & \omega_{q^\alpha p_\beta} \\ \omega_{p_\alpha q^\beta} & \omega_{p_\alpha p_\beta} \end{bmatrix} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix},\end{aligned}\tag{SF-FEPS}$$

and the equations of motion will be

$$S^a \omega_{ab} = \partial_b \mathcal{H}.\tag{1.132}$$

The function \mathcal{H} is called the Hamiltonian constraint. Typically, the constraint is chosen such that

$$\mathcal{H} = 0.\tag{1.133}$$

Relativistic free particle

Given that this formulation is probably unfamiliar to most, let us see how it works for a free particle in a Cartesian (and inertial) reference frame. In this case, it is more convenient to use $q^0 = ct$ as the time variable, which, given 1.110, means using $p_0 = -E/c$ for the energy. The Hamiltonian constraint is

$$\mathcal{H} = \frac{1}{2m} (p_\alpha \eta^{\alpha\beta} p_\beta + m^2 c^2) = \frac{1}{2m} (p_i \delta^{ij} p_j - (E/c)^2 + m^2 c^2),\tag{1.134}$$

where $\eta^{\alpha\beta}$ is the Minkowski metric. Hamilton's equations give

$$\begin{aligned}d_s q^0 &= d_s ct = \partial_{p_0} \mathcal{H} = \frac{1}{2m} \partial_{-E/c} (-(-E/c)^2) = \frac{1}{2m} - 2(-E/c) = \frac{E/c}{m} \\ d_s t &= \frac{E}{mc^2} \\ d_s q^i &= \partial_{p_i} \mathcal{H} = \frac{1}{2m} \partial_{p_i} (p_i \delta^{ij} p_j) = \frac{p^i}{m} \\ d_s p_0 &= d_s (-E/c) = \partial_{q^0} \mathcal{H} = 0 \\ d_s p_i &= \partial_{q^i} \mathcal{H} = 0.\end{aligned}\tag{1.135}$$

Therefore we have

$$p^\alpha = m d_s q^\alpha = m u^\alpha.\tag{1.136}$$

³²Recall that when phase space was extended with just time, we had $\theta_a = [\theta_{q^i} \ \theta_{p_i} \ \theta_t] = [p_i \ 0 \ -H]$. Now that it is also extended with energy, we have $\theta_a = [\theta_t \ \theta_{q^i} \ \theta_{-E} \ \theta_{p_i}] = [-E \ p_i \ 0 \ 0] = [\theta_{q^\alpha} \ \theta_{p_\alpha}] = [p_\alpha \ 0]$.

We recognize u^α as the four-velocity, and the affine parameter s as proper time. With this in mind, we can rewrite the Hamiltonian constraint as

$$\mathcal{H} = \frac{1}{2m} (mu^\alpha \eta_{\alpha\beta} mu^\beta + m^2 c^2) = \frac{1}{2} m |u|^2 + \frac{1}{2} m c^2, \quad (1.137)$$

which looks like a kinetic energy term constructed from the four-velocity. Setting \mathcal{H} to zero, then, means setting the norm squared of the four-velocity to $-c^2$, which is consistent with special relativity. It also means that

$$\begin{aligned} (-E/c)^2 &= m^2 c^2 + p_i p^i \\ E &= \pm \sqrt{c^2 |p_i|^2 + (m c^2)^2}. \end{aligned} \quad (1.138)$$

That is, the Hamiltonian constraint sets the relationship between energy and momentum.

The geometry of extended phase space

Note that while the mathematical structure of the extended phase space looks the same at first glance as the one of standard phase space, it has important differences. Let's concentrate on the Hamiltonian constraint. In standard Hamiltonian mechanics, the Hamiltonian H gives a value of energy that is conserved by the system during the evolution, but the system can be in different states with different energy. The Hamiltonian constraint \mathcal{H} , instead, is a quantity that is not only conserved by the system during the evolution, it is always the same for the given system. In the example above, the Hamiltonian constraint essentially is a constraint on the mass of the system. Therefore the space of states is not really the full $2n + 2$ dimensional manifold charted by time, position, energy and momentum, but it is the $2n + 1$ dimensional sub-manifold that is given by the constraint $\mathcal{H} = 0$. The constraint lowers the dimensionality by imposing a relationship between energy and the other state variables. The Hamiltonian constraint, then, plays a double role: as a generator of the evolution over the affine parameter s and as an equation of state of the system.³³

Since over valid states we will have both $\mathcal{H} = 0$ and $E = H$, we can write

$$\mathcal{H} = \lambda(H - E). \quad (1.139)$$

where λ is a function of the extended phase space. This expression provides a link between the Hamiltonian constraint and the standard Hamiltonian. To understand what λ is, note that

$$d_s t = \partial_{-E} \mathcal{H} = \lambda. \quad (1.140)$$

Therefore λ is the rate of change between time and the affine parameter. Additionally, given that time must still be a possible parameter for the evolution, $t(s)$ must be an invertible strictly monotonic function. Therefore we must always have $d_s t \neq 0$ at least over the region

³³Note that to apply the constraint to a distribution ρ , one can simply write $\mathcal{H}\rho = 0$. This means that ρ can be different from zero only over the region where \mathcal{H} is zero, and therefore ρ is non-zero only for those points that satisfy the Hamiltonian constraint. Moreover, note that in the case of a free particle in a Cartesian frame, the Hamiltonian constraint $\mathcal{H}\rho = 0$ is the classical analogue of the Klein-Gordon equation $\left(\frac{\hbar^2}{c^2} \partial_t^2 - \hbar^2 \nabla^2 + m^2 c^2\right) \psi = 0$.

where $E = H$. We can now verify that the new formulation recovers the old.

$$\begin{aligned}
\mathcal{H} &= \lambda(H - E) = 0 \\
E &= H \\
d_t t &= 1 = d_t s d_s t = d_t s \partial_{-E} \mathcal{H} = d_t s \lambda \\
d_t s &= \frac{1}{\lambda} \\
d_t q^i &= d_t s d_s q^i = d_t s \partial_{p_i} \mathcal{H} = \frac{1}{\lambda} (\partial_{p_i} \lambda(H - E) + \lambda \partial_{p_i} H) = \partial_{p_i} H \\
d_t p_i &= d_t s d_s p_i = -d_t s \partial_{q^i} \mathcal{H} = -\frac{1}{\lambda} (\partial_{q^i} \lambda(H - E) + \lambda \partial_{q^i} H) = -\partial_{q^i} H \\
d_t E &= d_t s d_s E = d_t s \partial_t \mathcal{H} = \frac{1}{\lambda} (\partial_t \lambda(H - E) + \lambda \partial_t H) = \partial_t H
\end{aligned} \tag{1.141}$$

Therefore if we set $\mathcal{H} = H - E$, the affine parameter will have to be equal to time up to an arbitrary additive constant, and therefore the formulations are equivalent.

The expression above seems to imply that the Hamiltonian constraint always has to correspond to one Hamiltonian. This is not the case. Suppose we have

$$\mathcal{H} = (H_1 - E)(H_2 - E) \tag{1.142}$$

so that $H_1(\xi^a) \neq H_2(\xi^a)$ at all points. This means that there are two regions of the extended phase space that satisfy the Hamiltonian constraints, one for H_1 and one for H_2 . These two regions are disconnected, therefore states from different regions cannot be connected by Hamiltonian evolution. They essentially represent two different types of system encoded into the same equation. In principle, the same idea can be extended to have more than two regions.

While this may seem just a mathematical artifact, note that this is exactly what happens for the Hamiltonian constraint of a free particle. In fact we can write:

$$\mathcal{H} = \frac{1}{2mc^2} (\sqrt{c^2|p_i|^2 + (mc^2)^2} + E) (\sqrt{c^2|p_i|^2 + (mc^2)^2} - E) \tag{1.143}$$

Given that the solutions for energy have opposite sign, one finds

$$\lambda = \frac{1}{2mc^2} (E + E) = \frac{E}{mc^2} \tag{1.144}$$

in agreement with what we found before.

Note that for a free particle, similarly to what one finds in quantum field theory, we have both positive and negative energy solutions. Given that $d_s t = E/mc^2$, what happens is the affine parameter is anti-aligned with respect to time. For the negative energy solutions, then, s will be minus proper time instead of proper time.³⁴

In general, $d_s t$ may be positive or negative, which corresponds to the affine parameter being aligned or anti-aligned with respect to time. Note that since $d_s t$ cannot be equal to zero, therefore states for which $d_s t > 0$ can never be connected with states for which $d_s t < 0$.

³⁴This is a much more precise version of the claim that anti-particles “travel backwards in time”. They do not. What happens is that the function $t(s)$ is parameterized in the opposite direction. However, the affine parameter s is not physically significant.

Therefore these two types of states must always exist in different regions of the extended phase space. In analogy with what happens in quantum field theory, we call particle states those for which $d_s t > 0$ and anti-particle states those for which $d_s t < 0$. We therefore find the following

Insight 1.145. *A frame-invariant notion of determinism and reversibility (i.e. allowing generalized coordinate transformations in Hamiltonian mechanics) gives us the notion of anti-particles, even in classical particle mechanics.*

Now that we got a better feel for what the Hamiltonian constraint is and how it works, let's go back to explore the geometry and physical significance of the extended phase space. Looking at ω_{ab} it may seem that adding the temporal degree of freedom means just adding another independent DOF, but that is not the case. If we start with standard phase space and add time, we are not adding new independent configurations: if we have deterministic and reversible motion, given the state at one time, we know the state at all times. So, while adding time allows us to talk about all states at all times, we are not really adding new states because states are defined and counted at equal time. In phase space extended by time only, this was captured by the fact that the form ω_{ab} applied to the displacement field always gives zero. If we now add energy, the conjugate of time, things are even more constrained. The energy, in fact, is a function of the state of the system at a particular time. The addition of energy, then, adds no configurations at all.

Now that it is clear that $(t, -E)$ do not constitute an independent degree of freedom, we should understand the significance of the minus sign. As we saw, it was needed so that p_α is a covector, but that does not tell us anything in terms of the geometry. What we are really saying is that an infinitesimal rectangle of size dt and dE has a negative contribution to the count of states. Shouldn't it have no contribution at all?

We saw that orthogonality in phase space means DOF independence. If we have two independent DOFs we can write:

$$\omega(dq^1 + dq^2, dp_1 + dp_2) = dq^1 dp_1 + dq^2 dp_2. \quad (1.146)$$

The above expression, then, can be understood as an areal version of Pythagoras's theorem: instead of summing the square distances, we are summing areas directly. That is, given the orthogonality of the independent DOFs, we have a right triangle-like structure where the two sides are the areas in each DOF and the hypotenuse is the area determined by ω . Now, because spatial and temporal degrees of freedom are not independent, i.e. they are not orthogonal, the form of the above expression cannot be the correct one for the temporal DOF.

The naive consideration would simply be to disregard the temporal degree of freedom, since time does not contribute new states, and set

$$\omega(dq + dt, dp + dE) = dq dp. \quad (1.147)$$

However, this does not work either. This would say that the temporal DOF does not identify states at all, which is not the case. Suppose we study a free particle with one degree of freedom. Let's assume that for $t = 0$, $q = 0$. We have

$$\begin{aligned} q &= \frac{p}{m} t \\ E &= \frac{p^2}{2m}. \end{aligned} \quad (1.148)$$

If we look at the region where momentum is positive, the relationship is bijective. Therefore time and energy can be used to identify, and therefore count, states. The issue is that those states are not new states, they are the same ones that are identified, and counted, by position and momentum.

If we look at the expressions derived before, we have

$$\omega(dq + dt, dp + dE) = dq dp + dt d(-E) = dq dp - dt dE. \quad (1.149)$$

We can rearrange this expression as

$$dq dp = \omega(dq + dt, dp + dE) + dt dE. \quad (1.150)$$

Now, compare 1.146 with 1.150. In the second expression, the area in the spatial DOF is the hypotenuse, while the area on the temporal DOF and the count of states given by ω are the sides. That is, while the temporal DOF and the spatial DOF are not orthogonal, the temporal DOF is orthogonal to the surface where states are counted. If we have a region defined at equal time, for example, the count of states reduces to the count of spatial configurations, which makes sense. In the general case, the differentials dt and dE are defined over surfaces at constant (q, p) , while the differentials dq and dp are defined over surfaces at constant (t, E) . These, as we said, are not orthogonal. This means that the area identified by dq and dp may have a non-zero projection over the temporal degree of freedom. Given that the state count should be defined at equal time, the area given by $dq dp$ does not always properly count states. If we want the count of states to be defined at equal time, this needs to be done on the surface that is orthogonal to the surface where dt and dE are defined, which therefore forms a right triangle-like structure with the spatial and temporal degrees of freedom.

The minus sign for the energy, then, has both a geometrical and physical meaning. If we used E instead of $-E$ as the variable, the minus sign would show up in the form ω_{ab} , which would probably be better as it would make the geometrical feature more clear. However, $p_\alpha = [E p_i]$ would not match the definition used in relativity and it would not form a covector, so it would be confusing in a different way. If there is a better overall notation and grouping, we haven't yet found it. At any rate, we have found that

Insight 1.151. *A frame-invariant notion of determinism and reversibility (i.e. allowing generalized coordinate transformations in Hamiltonian mechanics) gives elements of special relativity (i.e. energy-momentum four-vector), even without the notion of a metric tensor.*

What is interesting is that we didn't need to add any further assumption: DR and IND are sufficient. The only thing we needed to add was the ability to make generalized space-time transformations.

Relativistic kinematics

It is time to add assumption KE. In this generalized setting, the trajectory, and therefore the velocity, will be in terms of the affine parameter s . Therefore we set

$$d_s q^\alpha = d_s t d_t q^\alpha = d_s t [d_t q^0, d_t q^i] = d_s t [d_t q^0, v^i]. \quad (1.152)$$

Given that the Hamiltonian constraint on the extended phase space plays the same role as the Hamiltonian on the standard phase space, we will have an equivalent principle of stationary

action once KE is taken. The action can be written as

$$A[\gamma] = \int_{\gamma} L ds = \int_{\gamma} (p_{\alpha} u^{\alpha} - \mathcal{H}) ds \quad (1.153)$$

Under the full kinematic assumption, we have

$$\begin{aligned} \partial_{u^{\alpha}} p_{\beta} &= m g_{\alpha\beta} \\ p_{\alpha} &= m g_{\alpha\beta} u^{\beta} + \mathfrak{q} A_{\alpha} \\ u^{\alpha} &= d_s q^{\alpha} = \partial_{p_{\alpha}} \mathcal{H} = \frac{1}{m} g^{\alpha\beta} (p_{\beta} - \mathfrak{q} A_{\beta}) \\ \mathcal{H} &= \frac{1}{2m} (p_{\alpha} - \mathfrak{q} A_{\alpha}) g^{\alpha\beta} (p_{\beta} - \mathfrak{q} A_{\beta}) + U \end{aligned} \quad (1.154)$$

If we set $U = \frac{1}{2} m c^2$, this is the Hamiltonian constraint for a massive particle under potential forces.³⁵

As we saw, the best way to understand the geometry of phase space is to use the state variables q^{α} and p_{α} . However, understanding the physics is tricky as conjugate momentum is a gauge dependent quantity. One thing we can do is actually work on phase space with kinematic quantities, and express elements like θ_a and ω_{ab} in those variables. We have:

$$\begin{aligned} q^{\alpha} &= x^{\alpha} \\ p_{\alpha} &= m g_{\alpha\beta} u^{\beta} + \mathfrak{q} A_{\alpha} \\ x^{\alpha} &= q^{\alpha} \\ u^{\beta} &= \frac{1}{m} g^{\beta\alpha} (p_{\alpha} - \mathfrak{q} A_{\alpha}) \\ \mathcal{H} &= \frac{1}{2} m u^{\alpha} g_{\alpha\beta} u^{\beta} + \frac{1}{2} m c^2 \end{aligned} \quad (1.155)$$

Note how the Hamiltonian constraint, in kinematic variables, is always the same, regardless of the forces acting on the particle.

To calculate the expressions of θ_a and ω_{ab} in kinematic variables, we first see how the covector basis transforms. We have

$$\begin{aligned} e^{q^{\alpha}} &= \partial_{x^{\beta}} q^{\alpha} e^{x^{\beta}} + \partial_{u^{\gamma}} q^{\alpha} e^{u^{\gamma}} = \delta_{\beta}^{\alpha} e^{x^{\beta}} + 0 e^{u^{\gamma}} = e^{x^{\alpha}} \\ e^{p_{\alpha}} &= \partial_{x^{\beta}} p_{\alpha} e^{x^{\beta}} + \partial_{u^{\gamma}} p_{\alpha} e^{u^{\gamma}} \\ &= (m \partial_{x^{\beta}} g_{\alpha\gamma} u^{\gamma} + \mathfrak{q} \partial_{x^{\beta}} A_{\alpha}) e^{x^{\beta}} + m g_{\alpha\gamma} e^{u^{\gamma}} \end{aligned} \quad (1.156)$$

Now we can express the forms in terms of state variables and perform variable substitution.

$$\begin{aligned} \theta &= \theta_a e^a = p_{\alpha} e^{q^{\alpha}} = (m g_{\alpha\beta} u^{\beta} + \mathfrak{q} A_{\alpha}) e^{x^{\alpha}} \\ \omega &= \omega_{ab} e^a \otimes e^b = \omega_{q^{\alpha} p_{\beta}} e^{q^{\alpha}} \otimes e^{p_{\beta}} + \omega_{p_{\alpha} q^{\beta}} e^{p_{\alpha}} \otimes e^{q^{\beta}} = e^{q^{\alpha}} \otimes e^{p_{\alpha}} - e^{p_{\alpha}} \otimes e^{q^{\alpha}} \\ &= e^{x^{\alpha}} \otimes \left[(m \partial_{x^{\beta}} g_{\alpha\gamma} u^{\gamma} + \mathfrak{q} \partial_{x^{\beta}} A_{\alpha}) e^{x^{\beta}} + m g_{\alpha\gamma} e^{u^{\gamma}} \right] - \left[(m \partial_{x^{\beta}} g_{\alpha\gamma} u^{\gamma} + \mathfrak{q} \partial_{x^{\beta}} A_{\alpha}) e^{x^{\beta}} + m g_{\alpha\gamma} e^{u^{\gamma}} \right] \otimes e^{x^{\alpha}} \\ &= (m \partial_{x^{\beta}} g_{\alpha\gamma} u^{\gamma} + \mathfrak{q} \partial_{x^{\beta}} A_{\alpha}) e^{x^{\alpha}} \otimes e^{x^{\beta}} + m g_{\alpha\beta} e^{x^{\alpha}} \otimes e^{u^{\beta}} - (m \partial_{x^{\beta}} g_{\alpha\gamma} u^{\gamma} + \mathfrak{q} \partial_{x^{\beta}} A_{\alpha}) e^{x^{\beta}} \otimes e^{x^{\alpha}} - m g_{\alpha\beta} e^{u^{\beta}} \otimes e^{x^{\alpha}} \\ &= (m u^{\gamma} (\partial_{x^{\beta}} g_{\alpha\gamma} - \partial_{x^{\alpha}} g_{\beta\gamma}) + \mathfrak{q} (\partial_{x^{\beta}} A_{\alpha} - \partial_{x^{\alpha}} A_{\beta})) e^{x^{\alpha}} \otimes e^{x^{\beta}} + m g_{\alpha\gamma} e^{x^{\alpha}} \otimes e^{u^{\gamma}} - m g_{\alpha\gamma} e^{u^{\gamma}} \otimes e^{x^{\alpha}} \end{aligned}$$

³⁵An open problem is understanding whether U is constrained to be a constant or not.

$$(1.157)$$

We introduce the following two expressions

$$\begin{aligned} F_{\alpha\beta} &= \partial_{x^\alpha} A_\beta - \partial_{x^\beta} A_\alpha \\ G_{\alpha\beta\gamma} &= \partial_{x^\alpha} g_{\beta\gamma} - \partial_{x^\beta} g_{\alpha\gamma} \end{aligned} \quad (1.158)$$

We have

$$\begin{aligned} \theta_a &= [mg_{\alpha\beta}u^\beta + \mathbf{q}A_\alpha \ 0] \\ \omega_{ab} &= \begin{bmatrix} -mG_{\alpha\beta\gamma}u^\gamma - \mathbf{q}F_{\alpha\beta} & mg_{\alpha\beta} \\ -mg_{\alpha\beta} & 0 \end{bmatrix} \end{aligned} \quad (1.159)$$

We recognize $F_{\alpha\beta}$ as the electromagnetic field tensor. However, it is still an open problem what $G_{\alpha\beta\gamma}$ represents. It has a direct relationship with the Christoffel symbols $\Gamma_{\alpha\beta\gamma}$ ³⁶

$$\begin{aligned} G_{\alpha\beta\gamma} &= \partial_\alpha g_{\beta\gamma} - \partial_\beta g_{\alpha\gamma} = \frac{1}{2}(\partial_\alpha g_{\beta\gamma} - \partial_\beta g_{\alpha\gamma}) - \frac{1}{2}(\partial_\beta g_{\alpha\gamma} - \partial_\alpha g_{\beta\gamma}) \\ &= \frac{1}{2}(\partial_\gamma g_{\alpha\beta} + \partial_\alpha g_{\beta\gamma} - \partial_\beta g_{\alpha\gamma}) - \frac{1}{2}(\partial_\gamma g_{\alpha\beta} + \partial_\beta g_{\alpha\gamma} - \partial_\alpha g_{\beta\gamma}) \\ &= \Gamma_{\beta\alpha\gamma} - \Gamma_{\alpha\beta\gamma}. \end{aligned} \quad (1.160)$$

If we express the two tensors in covariant components, we find³⁷

$$\begin{aligned} G^{\alpha\beta\gamma} &= g^{\alpha\delta} g^{\beta\epsilon} g^{\gamma\zeta} G_{\delta\epsilon\zeta} = g^{\alpha\delta} g^{\beta\epsilon} g^{\gamma\zeta} \partial_\delta g_{\epsilon\zeta} - g^{\alpha\delta} g^{\beta\epsilon} g^{\gamma\zeta} \partial_\epsilon g_{\delta\zeta} \\ &= -g^{\alpha\delta} \partial_\delta g^{\beta\gamma} + g^{\beta\epsilon} \partial_\epsilon g^{\alpha\gamma} \\ &= \partial^\beta g^{\alpha\gamma} - \partial^\alpha g^{\beta\gamma} \end{aligned} \quad (1.161)$$

$$\begin{aligned} F^{\alpha\beta} &= g^{\alpha\gamma} g^{\beta\delta} F_{\gamma\delta} = g^{\alpha\gamma} g^{\beta\delta} \partial_\gamma A_\delta - g^{\alpha\gamma} g^{\beta\delta} \partial_\delta A_\gamma \\ &= g^{\alpha\gamma} \partial_\gamma (g^{\beta\delta} A_\delta) - g^{\beta\delta} \partial_\delta (g^{\alpha\gamma} A_\gamma) - g^{\alpha\gamma} \partial_\gamma g^{\beta\delta} A_\delta + g^{\beta\delta} \partial_\delta g^{\alpha\gamma} A_\gamma \\ &= \partial^\alpha A^\beta - \partial^\beta A^\alpha + \partial^\beta g^{\alpha\gamma} A_\gamma - \partial^\alpha g^{\beta\delta} A_\delta \\ &= \partial^\alpha A^\beta - \partial^\beta A^\alpha + G^{\alpha\beta\gamma} A_\gamma \end{aligned} \quad (1.162)$$

The link between between $F^{\alpha\beta}$ and $G^{\alpha\beta\gamma}$ is not specific to the electromagnetic field. In fact

$$\begin{aligned} \nabla^\alpha A^\beta - \nabla^\beta A^\alpha &= g^{\alpha\gamma} \nabla_\gamma A^\beta - g^{\beta\gamma} \nabla_\gamma A^\alpha \\ &= g^{\alpha\gamma} (\partial_\gamma A^\beta + \Gamma_{\gamma\delta}^\beta A^\delta) - g^{\beta\gamma} (\partial_\gamma A^\alpha + \Gamma_{\gamma\delta}^\alpha A^\delta) \\ &= \partial^\alpha A^\beta - \partial^\beta A^\alpha + g^{\alpha\delta} g^{\beta\epsilon} A^\gamma (\Gamma_{\epsilon\delta\gamma} - \Gamma_{\delta\epsilon\gamma}) \\ &= \partial^\alpha A^\beta - \partial^\beta A^\alpha + G^{\alpha\beta\gamma} A_\gamma \end{aligned} \quad (1.163)$$

³⁶It is not the torsion as it anti-symmetrizes the first two indexes of the Christoffel symbols, while the torsion uses the second two. In fact, we used the expression for a connection with no torsion to derive the expression.

³⁷The derivation uses the following relationship

$$\partial_\alpha g^{\beta\gamma} = -g^{\beta\delta} g^{\gamma\epsilon} \partial_\alpha g_{\delta\epsilon}$$

which can be derived from the following

$$0 = \partial_\alpha \delta_\epsilon^\beta = \partial_\alpha (g^{\beta\delta} g_{\delta\epsilon}) = g_{\delta\epsilon} \partial_\alpha g^{\beta\delta} + g^{\beta\delta} \partial_\alpha g_{\delta\epsilon} = g^{\gamma\epsilon} g_{\delta\epsilon} \partial_\alpha g^{\beta\delta} + g^{\gamma\epsilon} g^{\beta\delta} \partial_\alpha g_{\delta\epsilon} = \partial_\alpha g^{\beta\gamma} + g^{\gamma\epsilon} g^{\beta\delta} \partial_\alpha g_{\delta\epsilon}$$

The expression $G_{\alpha\beta\gamma}$ is linked to a sort of covariant exterior derivative for vectors. Unfortunately, we do not truly understand its geometrical significance.

To find the equations for motions, we first calculate the Poisson brackets between the kinematic variables.

$$\{x^\alpha, x^\beta\} = \{q^\alpha, q^\beta\} = 0 \quad (1.164)$$

$$\begin{aligned} \{x^\alpha, u^\beta\} &= \{q^\alpha, \frac{1}{m}g^{\beta\gamma}(p_\gamma - \mathbf{q}A_\gamma)\} = \{q^\alpha, \frac{1}{m}g^{\beta\gamma}p_\gamma\} - \{q^\alpha, \frac{\mathbf{q}}{m}g^{\beta\gamma}A_\gamma\} \\ &= \frac{1}{m}g^{\alpha\beta} \end{aligned} \quad (1.165)$$

$$\begin{aligned} \{u^\alpha, u^\beta\} &= \frac{1}{m^2}\{g^{\alpha\gamma}p_\gamma - \mathbf{q}A^\alpha, g^{\beta\delta}p_\delta - \mathbf{q}A^\beta\} \\ &= \frac{1}{m^2}[\{g^{\alpha\gamma}p_\gamma, g^{\beta\delta}p_\delta\} - \{g^{\alpha\gamma}p_\gamma, \mathbf{q}A^\beta\} - \{\mathbf{q}A^\alpha, g^{\beta\delta}p_\delta\} + \{\mathbf{q}A^\alpha, \mathbf{q}A^\beta\}] \\ &= \frac{1}{m^2}[g^{\alpha\gamma}\{p_\gamma, p_\delta\}g^{\beta\delta} + p_\gamma\{g^{\alpha\gamma}, p_\delta\}g^{\beta\delta} + g^{\alpha\gamma}\{p_\gamma, g^{\beta\delta}\}p_\delta + p_\gamma\{g^{\alpha\gamma}, g^{\beta\delta}\}p_\delta \\ &\quad - g^{\alpha\gamma}\{p_\gamma, \mathbf{q}A^\beta\} - p_\gamma\{g^{\alpha\gamma}, \mathbf{q}A^\beta\} - \{\mathbf{q}A^\alpha, p_\delta\}g^{\beta\delta} - \{\mathbf{q}A^\alpha, g^{\beta\delta}\}p_\delta] \\ &= \frac{1}{m^2}[G^{\alpha\beta\gamma}p_\gamma + \mathbf{q}(\partial^\alpha A^\beta - \partial^\beta A^\alpha)] \\ &= \frac{1}{m^2}[G^{\alpha\beta\gamma}m(g_{\gamma\delta}u^\delta + \mathbf{q}A_\gamma) + \mathbf{q}(\partial^\alpha A^\beta - \partial^\beta A^\alpha)] \\ &= \frac{1}{m^2}[G^{\alpha\beta\gamma}mg_{\gamma\delta}u^\delta + \mathbf{q}(G^{\alpha\beta\gamma}A_\gamma + \partial^\alpha A^\beta - \partial^\beta A^\alpha)] \\ &= \frac{1}{m^2}[G^{\alpha\beta\gamma}mg_{\gamma\delta}u^\delta + \mathbf{q}F^{\alpha\beta}] \end{aligned} \quad (1.166)$$

We find the evolution of the kinematic variables by calculating the Poisson bracket with the Hamiltonian constraint.

$$\begin{aligned} d_s x^\alpha &= \{x^\alpha, \mathcal{H}\} = \frac{1}{2}m\{x^\alpha, u^\beta g_{\beta\gamma}u^\gamma\} = mu^\beta g_{\beta\gamma}\{x^\alpha, u^\gamma\} \\ &= mu^\beta g_{\beta\gamma} \frac{1}{m}g^{\alpha\gamma} = u^\alpha \end{aligned} \quad (1.167)$$

$$\begin{aligned} d_s u^\alpha &= \{u^\alpha, \mathcal{H}\} = \frac{1}{2}m\{u^\alpha, u^\beta g_{\beta\gamma}u^\gamma\} \\ &= mu^\beta g_{\beta\gamma}\{u^\alpha, u^\gamma\} + \frac{1}{2}mu^\beta u^\gamma\{u^\alpha, g_{\beta\gamma}\} \\ &= mu^\beta g_{\beta\gamma} \frac{1}{m^2}(G^{\alpha\gamma\delta}mg_{\delta\epsilon}u^\epsilon + \mathbf{q}F^{\alpha\gamma}) - \frac{1}{2}mu^\beta u^\gamma \frac{1}{m}g^{\alpha\delta}\partial_\delta g_{\beta\gamma} \\ &= u^\beta u^\epsilon g_{\beta\gamma}g_{\epsilon\delta}G^{\alpha\gamma\delta} - \frac{1}{2}u^\beta u^\gamma g^{\alpha\delta}\partial_\delta g_{\beta\gamma} + \frac{\mathbf{q}}{m}F^{\alpha\gamma}g_{\gamma\beta}u^\beta \\ &= u^\beta u^\epsilon g^{\alpha\gamma}G_{\gamma\beta\epsilon} - \frac{1}{2}u^\beta u^\gamma g^{\alpha\delta}(\partial_\beta g_{\gamma\delta} - \partial_\beta g_{\gamma\delta} + \partial_\delta g_{\beta\gamma}) + \frac{\mathbf{q}}{m}F^{\alpha\gamma}g_{\gamma\beta}u^\beta \end{aligned}$$

$$\begin{aligned}
&= u^\beta u^\gamma g^{\alpha\delta} G_{\delta\beta\gamma} - u^\beta u^\gamma g^{\alpha\delta} \frac{1}{2} (\partial_\gamma g_{\delta\beta} + \partial_\delta g_{\beta\gamma} - \partial_\beta g_{\gamma\delta}) + \frac{q}{m} F^{\alpha\gamma} g_{\gamma\beta} u^\beta \\
&= u^\beta u^\gamma g^{\alpha\delta} (\Gamma_{\beta\delta\gamma} - \Gamma_{\delta\beta\gamma} - \Gamma_{\beta\delta\gamma}) + \frac{q}{m} F^{\alpha\gamma} g_{\gamma\beta} u^\beta \\
&= -u^\beta u^\gamma g^{\alpha\delta} \Gamma_{\delta\beta\gamma} + \frac{q}{m} F^{\alpha\gamma} g_{\gamma\beta} u^\beta
\end{aligned}$$

$$D_s u^\alpha = d_s u^\alpha + u^\beta d_s x^\gamma \Gamma_{\beta\gamma}^\alpha = \frac{q}{m} F^{\alpha\gamma} g_{\gamma\beta} u^\beta \quad (1.168)$$

These are geodesic equations modified by the electromagnetic force, which are consistent with general relativity.

We have found that the addition of assumption **KE** in the generalized case gives us relativistic mechanics and only relativistic mechanics. There was no choice at any point, therefore, in this sense, relativistic mechanics is the only options that works.

Insight 1.169. *Relativistic mechanics is a consequence of **DR**, **IND** and **KE**.*

Let's see how the Minkowski metric and the speed of light emerges from what we have already discussed. The strong kinematic assumptions imposes a linear relationship between velocity which, in absence of forces, and be written as

$$p_\alpha = m g_{\alpha\beta} u^\alpha. \quad (1.170)$$

Given that $g_{\alpha\beta}$ is a symmetric tensor that depends only on space-time, it can be diagonalized at a point P with a suitable coordinate choice. While the velocity is in terms of an affine parameter s , we can express it in terms of time since $d_s x^\alpha = d_{st} d_t x^\alpha = \lambda d_t x^\alpha$. Since we can set $x^0 = t$, the above equation becomes

$$\begin{aligned}
-E &= m g_{00} \lambda d_t t = \lambda m g_{00} \\
p_i &= m g_{ii} \lambda d_t x^i = \lambda m g_{ii} v^i.
\end{aligned} \quad (1.171)$$

Dimensional analysis on the spatial components tells us that already work in the sense that g_{ii} must be pure numbers, under the assumption that s has dimensions of time. For the time component, instead, dimensional analysis tells us that g_{00} must be the square of a velocity. As we saw before, solution with positive energy are those for which s and λ are aligned, therefore g_{00} must be a negative quantity. Therefore we set $g_{00} = -c^2$, and we recognize c as the speed of light. Note that, technically, the constant does not play at all a role of speed in this discussion, just as a conversion factor from the temporal component of the four-velocity to the energy.

We can change units so that $g_{00} = -1$ and $g_{ii} = 1$. For the spatial components, it is just a matter of rescaling the units. For time, units need to be changed as well by setting $p_0 = -E/c$ and $x^0 = ct$. Note that the product of the two remains of the same unit, and the two are still conjugate as space variables. Therefore we have shown that, at every point, we can find a set of space-time coordinates such as $g_{\alpha\beta} = \eta_{\alpha\beta}$ where $\eta_{\alpha\beta}$ is the Minkowki metric. If we use proper time as the affine parameter s , then λ becomes $\gamma = \frac{1}{\sqrt{1 + \frac{v^2}{c^2}}}$

The failure of Galilean relativity

Given that our assumptions led directly to relativistic mechanics, non-relativistic mechanics must fail to satisfy the strong kinematic assumption. We have not identified the exact reason of the failure, though we have identified an interesting issue.

Galilean space-time transformations correspond to

$$\begin{aligned}\hat{t} &= t \\ \hat{x} &= x + v_0 t \\ \hat{v} &= v + v_0\end{aligned}\tag{1.172}$$

The canonical transformation induced by that space-time variable change is the following

$$\begin{aligned}t &= \hat{t} \\ q &= \hat{q} - v_0 \hat{t} \\ \hat{p} &= \partial_{\hat{q}} q p + \partial_{\hat{t}} t (-E) = p \\ -\hat{E} &= \partial_{\hat{t}} q p + \partial_{\hat{t}} t (-E) = -v_0 p - E\end{aligned}\tag{1.173}$$

which gives us different rules for how momentum and energy transform. Note, for example, that the momentum remains unchanged. The expected rules would be

$$\begin{aligned}\hat{p} &= m\hat{v} = mv + mv_0 = p + mv_0 \\ \hat{E} &= \frac{1}{2}m\hat{v}^2 = \frac{1}{2}m(v + v_0)^2 = \frac{1}{2}mv^2 + mvv_0 + \frac{1}{2}m(v_0)^2 \\ &= E + v_0 p + \frac{1}{2}m(v_0)^2.\end{aligned}\tag{1.174}$$

The difference between the two expressions is a constant, therefore Galilean change of variables together with non-relativistic expressions for momentum and energy are still canonical transformations. However, accommodating that constant requires a change of gauge. That is, if we start in an inertial frame where

$$p_\alpha = m\lambda g_{\alpha\beta} u^\beta\tag{1.175}$$

we end up in another inertial frame where

$$\begin{aligned}p_\alpha &= m\lambda g_{\alpha\beta} u^\beta + A_\alpha \\ A_\alpha &= \left[\frac{1}{2}m|v_0|^2 \quad mv_0^i \right].\end{aligned}\tag{1.176}$$

The new vector potential is still a constant field, therefore the forces do not change. Still, the spatial components have a different value than the time ones, therefore the direction of momentum is corrected.

What we find, then, is that Galilean transformations do not preserve the relationship between kinetic momentum and conjugate momentum. To recover the original relationship, one has to perform a gauge transformation. But this is contrary to the expectation that the laws are the same for all inertial frames. There are likely other deeper issues at play, which we hope to uncover in the future.

Metric tensor revisited

As we saw, the metric tensor appears not as defining the distances in space-time but rather as the linear relationship between velocity and conjugate momentum. It does, however, end up playing an important geometric role: in 1.159 we see that the metric tensor times the mass is the off diagonal component of the form ω_{ab} . Therefore, if we have a range of positions dx^α and a range of velocities dv^β , the number of configurations they identify is given by $dx^\alpha m g_{\alpha\beta} dv^\beta$. In other words,

Insight 1.177. *the metric tensor allows us to count states in terms of the kinematic variables.*

This corroborates what we saw in 1.127: the determinant of the metric tensor is the Jacobian determinant that allows us to express densities over phase space as densities over kinematic variables. In relativity, the square root of the determinant of the metric tensor appears to create an invariant volume element. That is, $\int_U \sqrt{|g_{\alpha\beta}|} dx^0 dx^1 dx^2 dx^3$ defines the volume of space-time region U . However, we can also define an invariant volume element over the space of kinematic variables. That is, $\int_V |g_{\alpha\beta}| dx^0 \dots dx^3 du^0 \dots du^3$ defines the volume of the position-velocity space. The determinant of the metric tensor, then, is more directly giving us the size of a volume taken in position-velocity space instead of space-time. This also connects the ability to specify distributions in terms of kinematic variables since, as we saw in FKE-DEN, the Jacobian determinant of the transformation between position and velocity is proportional to $|g_{\alpha\beta}|$.

Why is it, then, that defining volumes over kinematic quantities fixes the geometry of space-time? Note that both velocities and differentials of space transform like vectors since $du^\alpha = dx^\alpha/d\tau$. Therefore if $dx^\alpha g_{\alpha\beta} du^\beta$ is invariant, then $dx^\alpha g_{\alpha\beta} dx^\beta$ will also be invariant. The geometry of space-time, then, is set by the geometry of phase space.

There are other interesting open questions to explore to better understand this relationship. For example, in relativity, one does put spatial and time variables on the same plane. However, we saw that the primary role of the metric tensor is so that we deal with distributions and integration, and these work differently in space and time. Consider, in fact, a distribution $\rho(x^i, v^j, t)$. This would represent the evolution over time of a distribution over kinematic variables. At each time, the distribution has to integrate to one, assuming it is normalized, which means the units of ρ are $\left[\frac{1}{[x]^3 [v]^3} \right]$, the inverse of the cube of distance times velocity. Note that every observer has to see the same thing: a normalized distribution at each moment in time. Additionally, this has to happen no matter what ρ is. An open question in reverse physics, then, is whether imposing this requirement is already enough to recover parts of relativity.

1.10 Reversing phase space

In this section we will find the assumptions required to rederive the structure of phase space. The principle of relativity will play a key role as the structure of classical phase space is the only structure that allows us to define densities and entropy in a way that is coordinate invariant.

Properties of phase space

We have seen that DR and IND are the constitutive assumptions of Hamiltonian mechanics, in the sense that they fully characterize Hamiltonian evolution. The addition of KE in its

full version recovers both Lagrangian mechanics and massive particles under potential forces. The relativistic version of the theories comes out without additional assumptions simply by properly dividing the role of time as a variable from time as a parameter. But in all this discussion we assumed, without questioning, that states are identified by pairs of variables: position-velocity or position-momentum. Is this a coincidence or is there an underlying reason for it? Can we find assumptions from which the structure of phase space itself can be recovered?

Since we are interested in the structure of phase space itself, let's go back to its original version, without the extension to time or the temporal DOF. Throughout the whole discussion, conjugate momentum p_i has been written with the index down, tacitly assuming it to be a covector. This means it obeys the following transformation rules under coordinate changes:

$$\begin{aligned}\hat{q}^i &= \hat{q}^i(q^j) \\ \hat{p}_i &= \partial_{\hat{q}^i} q^j p_j\end{aligned}\tag{1.178}$$

When a metric tensor is defined, that is when assumption [KE](#) is valid, the difference between vector and covector blurs, as we can always transform one into the other, but this is not the general case. We have, therefore, this condition

$$\text{Conjugate momentum } p_i \text{ changes like a covector under changes of coordinates } q^i \tag{PS-COV}$$

which characterizes the relationship between q^i and p_i . We want to stress that these changes of coordinates do not mix space and time variables. Therefore we are only talking different choices of spatial coordinates at equal time.

Given that the form ω_{ab} plays a fundamental role, as it defines the geometry of phase space in terms of state count, let's see how it transforms during a coordinate change. We have:

$$\begin{aligned}\hat{\omega}_{ab} &= \begin{bmatrix} \omega_{\hat{q}^i \hat{q}^j} & \omega_{\hat{q}^i \hat{p}_j} \\ \omega_{\hat{p}_i \hat{q}^j} & \omega_{\hat{p}_i \hat{p}_j} \end{bmatrix} = \begin{bmatrix} \partial_{q^k} \hat{q}^i & \partial_{p_k} \hat{q}^i \\ \partial_{q^k} \hat{p}_i & \partial_{p_k} \hat{p}_i \end{bmatrix} \begin{bmatrix} \omega_{q^k q^l} & \omega_{q^k p_l} \\ \omega_{p_k q^l} & \omega_{p_k p_l} \end{bmatrix} \begin{bmatrix} \partial_{q^l} \hat{q}^j & \partial_{q^l} \hat{p}_j \\ \partial_{p_l} \hat{q}^j & \partial_{p_l} \hat{p}_j \end{bmatrix} \\ &= \begin{bmatrix} \partial_{q^k} \hat{q}^i & 0 \\ \partial_{q^k} \hat{p}_i & \partial_{\hat{q}^i} q^k \end{bmatrix} \begin{bmatrix} 0 & \delta_{kl} \\ -\delta_{kl} & 0 \end{bmatrix} \begin{bmatrix} \partial_{q^l} \hat{q}^j & \partial_{q^l} \hat{p}_j \\ 0 & \partial_{\hat{q}^j} q^l \end{bmatrix} \\ &= \begin{bmatrix} 0 & \partial_{q^l} \hat{q}^i \\ -\partial_{\hat{q}^i} q^l & \partial_{q^l} \hat{p}_i \end{bmatrix} \begin{bmatrix} \partial_{q^l} \hat{q}^j & \partial_{q^l} \hat{p}_j \\ 0 & \partial_{\hat{q}^j} q^l \end{bmatrix} \\ &= \begin{bmatrix} 0 & \partial_{q^l} \hat{q}^i \partial_{\hat{q}^j} q^l \\ -\partial_{\hat{q}^i} q^l \partial_{q^l} \hat{q}^j & -\partial_{\hat{q}^i} q^l \partial_{q^l} \hat{p}_j + \partial_{q^l} \hat{p}_i \partial_{\hat{q}^j} q^l \end{bmatrix} \\ &= \begin{bmatrix} 0 & \partial_{\hat{q}^j} \hat{q}^i \\ -\partial_{\hat{q}^i} \hat{q}^j & -\partial_{\hat{q}^i} \hat{p}_j + \partial_{\hat{q}^j} \hat{p}_i \end{bmatrix} = \begin{bmatrix} 0 & \delta_{ij} \\ -\delta_{ij} & 0 \end{bmatrix}\end{aligned}\tag{1.179}$$

For the last step, we are taking partial derivatives of the new variables with respect to the new variables. The derivative is equal to one if we are taking the partial derivative with respect to the same variable and zero otherwise. We find that condition

$$\text{The form } \omega_{ab} \text{ is invariant under changes of coordinates } q^i \tag{PS-SYMP}$$

is implied by [PS-COV](#).

As usual, we ask whether the converse is true. That is, suppose we perform a change of coordinates $\hat{q}^i = \hat{q}^i(q^j)$ for which ω_{ab} is invariant. Does this pose restrictions on how conjugate momentum changes? We have

$$\begin{aligned}
\hat{\omega}_{ab} &= \begin{bmatrix} \omega_{\hat{q}^i \hat{q}^j} & \omega_{\hat{q}^i \hat{p}_j} \\ \omega_{\hat{p}_i \hat{q}^j} & \omega_{\hat{p}_i \hat{p}_j} \end{bmatrix} = \begin{bmatrix} \partial_{q^k} \hat{q}^i & \partial_{p^k} \hat{q}^i \\ \partial_{q^k} \hat{p}_i & \partial_{p^k} \hat{p}_i \end{bmatrix} \begin{bmatrix} \omega_{q^k q^l} & \omega_{q^k p_l} \\ \omega_{p_k q^l} & \omega_{p_k p_l} \end{bmatrix} \begin{bmatrix} \partial_{q^l} \hat{q}^j & \partial_{q^l} \hat{p}_j \\ \partial_{p_l} \hat{q}^j & \partial_{p_l} \hat{p}_j \end{bmatrix} \\
&= \begin{bmatrix} \partial_{q^k} \hat{q}^i & 0 \\ \partial_{q^k} \hat{p}_i & \partial_{p^k} \hat{p}_i \end{bmatrix} \begin{bmatrix} 0 & \delta_{kl} \\ -\delta_{kl} & 0 \end{bmatrix} \begin{bmatrix} \partial_{q^l} \hat{q}^j & \partial_{q^l} \hat{p}_j \\ 0 & \partial_{p_l} \hat{p}_j \end{bmatrix} \\
&= \begin{bmatrix} 0 & \partial_{q^l} \hat{q}^i \\ -\partial_{p_l} \hat{p}_i & \partial_{q^l} \hat{p}_i \end{bmatrix} \begin{bmatrix} \partial_{q^l} \hat{q}^j & \partial_{q^l} \hat{p}_j \\ 0 & \partial_{p_l} \hat{p}_j \end{bmatrix} \\
&= \begin{bmatrix} 0 & \partial_{q^l} \hat{q}^i \partial_{p_l} \hat{p}_j \\ -\partial_{p_l} \hat{p}_i \partial_{q^l} \hat{q}^j & -\partial_{p_l} \hat{p}_i \partial_{q^l} \hat{p}_j + \partial_{q^l} \hat{p}_i \partial_{p_l} \hat{p}_j \end{bmatrix} = \omega_{ab} = \begin{bmatrix} 0 & \delta_{ij} \\ -\delta_{ij} & 0 \end{bmatrix}
\end{aligned} \tag{1.180}$$

The two off diagonal terms impose the same constraint

$$\partial_{q^l} \hat{q}^i \partial_{p_l} \hat{p}_j = \delta_{ij}. \tag{1.181}$$

In matrix terms, we are taking the product of two matrices and equating it to the identity matrix. Therefore the two matrices are the inverse of each other:

$$\partial_{p_l} \hat{p}_j = (\partial_{q^l} \hat{q}^j)^{-1} = \partial_{\hat{q}^j} q^l. \tag{1.182}$$

This means that the change of position variables induces a change of momentum

$$\hat{p}_j = \partial_{\hat{q}^j} q^i p_i + A_j(q^k), \tag{1.183}$$

where A_j are arbitrary functions, which we can set to zero without loss of generality.³⁸ This makes momentum a covector.

The second constraint comes from the bottom diagonal term. Using the newly found transformation rules, we have

$$-\partial_{p_l} \hat{p}_i \partial_{q^l} \hat{p}_j + \partial_{q^l} \hat{p}_i \partial_{p_l} \hat{p}_j = -\partial_{\hat{q}^i} q^l \partial_{q^l} \hat{p}_j + \partial_{q^l} \hat{p}_i \partial_{\hat{q}^j} q^l = -\partial_{\hat{q}^i} \hat{p}_j + \partial_{\hat{q}^j} \hat{p}_i = 0. \tag{1.184}$$

This constraint is satisfied simply because position and momentum are different state variables, and partial derivatives along one variable are taken keeping the others constant. By imposing the preservation of the form ω_{ab} under an arbitrary change of coordinates, we recovered the transformation law of momentum as a covector. Therefore conditions [PS-COV](#) and [PS-SYMP](#) are equivalent.

The invariance of the form ω_{ab} under coordinate changes means that all those properties that were invariant under Hamiltonian evolution are also invariant under equal-time coordinate changes. Therefore condition

$$\text{The Poisson brackets are invariant under equal-time coordinate changes} \tag{PS-POI}$$

³⁸The arbitrary functions change the value of zero momentum, potentially at every point in a different way. While this is mathematically possible, it would make no physical sense that a coordinate change would induce a change in the zero reference for momentum. We will see later how this is related to gauge transformations.

is equivalent to [PS-SYMP](#).

The following conditions are equivalent to each other and are implied by [PS-SYMP](#), but do not imply it.

The system allows statistically independent distributions over each DOF under any choice of coordinates q^i	(PSI-DEN)
The system allows informationally independent distributions over each DOF under any choice of coordinates q^i	(PSI-INFO)
The system allows peaked distributions where the uncertainty is the product of the uncertainty on each DOF under any choice of coordinates q^i	(PSI-UNC)

Physically, these correspond to the independence of DOFs, which is assumption [IND](#). The only difference is that the assumption must be valid for all equal-time coordinate changes, which is just an application of the principle of relativity.

Lastly, the following conditions are all equivalent but independent of the PSI conditions and are implied by [PS-SYMP](#), but do not imply it.

Phase space volumes are invariant under equal-time changes of coordinates q^i	(PSV-VOL)
The Jacobian for the transformation induced by equal-time changes of coordinates q^i is unitary	(PSV-JAC)
Densities over phase space are invariant under equal-time changes of coordinates q^i	(PSV-DEN)
Thermodynamic entropy is invariant under equal-time changes of coordinates q^i	(PSV-THER)
Information entropy is invariant under equal-time changes of coordinates q^i	(PSV-INFO)
Uncertainty of peaked distributions is invariant under equal-time changes of coordinates q^i	(PSV-UNC)

Physically, these correspond to requiring that the count of states is the same under all choices of coordinates at a given time. In the same way that conservation of ω_{ab} in time was equivalent to assumptions [DR](#) and [IND](#) combined, any PSV condition together with any PSI condition will be equivalent to any PS condition. That is, the phase space structure corresponds to invariance of state count plus independence of DOFs.

It should be evident that all these properties are necessary if we want to have a physically meaningful state space. If state count, densities, entropy, independence of DOFs were not properties that all equal-time observers could agree on, there would be no notion of an objective state to begin with, and it would be pointless to even talk about isolation, determinism, thermodynamics and so on. All these properties, then, are constitutive assumptions for any state space, as without them there are no well defined states. Having established that phase space has all these properties, do these properties define phase space? That is, suppose the states of our system are identified by a finite number of continuous quantities, meaning that the state space is a manifold; does the invariance of those properties constrain that space to be phase space?

Invariance over the continuum

To simplify the matter, let us consider the case of a single DOF. One way to define the problem is to ask that if a distribution is uniform for one observer, it should be uniform for all equal-time observers. Uniform distributions are important in statistical mechanics since the macrostate of an isolated system is assumed to be a uniform distribution over all possible microstates that satisfy a few constraints, such as the value of the energy, the number of particles and so on. This is very similar to the “principle of indifference” in classical probability, which assigns equal probabilities to outcomes for which there is no justifiable preference.

For example, if we have a fair die, meaning that the die itself and the mechanism for throwing it do not have a preference for any side, the principle tells us to assign equal probability to all sides.³⁹ If the die has 6 sides, we assign probability 1/6 for each side. The part of probability theory that is linked to combinatorics works like this. Note that what number we use to label each side of the die, or whether we use something else (e.g. for poker dice, images of playing cards) is irrelevant for the probability assignment. Therefore, when applied to discrete variables, the principle of indifference is invariant under relabeling.

If we try to apply the same idea on the continuum, however, things are not as simple. Suppose we have a factory that produces boxes at random, with side from $1m$ to $3m$. Furthermore, let’s assume that the manufacturing procedure does not have a preference for the size of the boxes. Using the principle of indifference, we assume a uniform distribution for the side from $1m$ to $3m$. However, we could have equally said that the factory does not have a preference for the total volume, and assign a uniform distribution for the volume from $1m^3$ to $27m^3$. The problem is that $x \rightarrow x^3$ is a non-linear transformation, and uniform distributions do not remain uniform under non-linear transformations. Another way to frame it, on the continuum we do not have a probability, but a probability density, which has units of probability over units of the variable. In the first case the units were probability over meters while in the second probability over meters cubed. The density, then, is unit dependent, it depends on the variable.

This is essentially our problem:

Insight 1.185. *Densities over continuum variables depend on the choice of variable and therefore do not, in general, satisfy the principle of relativity.*

If we have a uniform distribution over a variable, it will only remain uniform under linear changes of variable. This means that density, thermodynamic entropy, information entropy, all the properties we looked at before, are not the same under variable changes. Mathematically, the issue is that only transformations with unitary Jacobian preserve these properties, which is not the general case.

How does phase space solve the problem? Phase space is defined by two variables, q and p . A change of coordinates only changes q arbitrarily. Once the q change is determined, p changes as a covector. Around each point, we have:

$$\begin{aligned} d\hat{q} &= d_q \hat{q} dq \\ d\hat{p} &= d_{\hat{q}} q dp \end{aligned} \tag{1.186}$$

³⁹The principle of indifference is often stated in terms of degree of belief: it is the agent that has no reason to prefer one outcome over the other. We state it in more objective terms: the system and preparation procedure do not have a preference. If an agent believes a die to be fair, while in fact the die is loaded, the principle of indifference gives wrong empirical results, and therefore, for us, it does not apply.

which means

$$d\hat{q}d\hat{p} = d_q\hat{q}dq d_{\hat{q}}q dp = d_qq dq dp = dq dp. \quad (1.187)$$

The area is conserved under a generic change of q precisely because p changes in the opposite way.⁴⁰

While the math is clear, what is the physics behind this? Why does conjugate momentum magically change when we change position? This is better understood if we think in kinematic variables: position and velocity. Why would velocity change when we change position? Because units of velocity are units of position over time. If we redefine units of position, we will redefine units of velocity as well. As we said before, the mathematical tools we currently use do not capture all the physical elements, and unit dependence is something that mathematics currently fails to capture. The defining role of coordinate variables q^i , then, is not that they describe position, but rather the following:

Insight 1.188. *The coordinate variables q^i define the unit system.*

What are the units of conjugate momentum, then? Given that the product $dqdp$ is invariant and it measures the count of configurations for one DOF, p_i must have units of configurations divided by units of the corresponding q^i . By convention, we measure phase-space areas in units of angular momentum, in units of action h , but why do we do that? Suppose that we track position by an angle in radians q^θ . Conjugate momentum p_θ will be expressed in units of area of phase space, and therefore angular momentum. This is the connection between angular momentum and phase-space areas: the conjugate of a dimensionless quantity must have the same physical dimensions of areas of phase space. However, if we change the angle from radians to degrees, conjugate momentum will be units of h over degrees. That is, it is improper to say the count of states is expressed in units of angular momentum. Moreover, while two quantities that represent the same physical quantity must have the same physical dimension, the converse is not true.⁴¹ Ultimately, areas in phase space are measured with those units because, in an inertial Cartesian coordinate frame, linear kinetic momentum times distance can be used to count the configurations of the system, as we saw when discussing the full kinematic equivalence.

To recap, we saw that we cannot define a coordinate invariant density over a single continuous variable, but we can do that over two variables, if one has inverse units with respect

⁴⁰In statistical mechanics, it is often said that Liouville's theorem justifies the use of phase space volume as state count. Liouville's theorem states that under Hamiltonian evolution the density, or equivalently the phase space volume is conserved, which corresponds to assumption DR. As we mentioned before, the conservation of volume in time would not be physically meaningful if the volume were not first an objective quantity. Therefore Liouville's theorem does *not* justify the use of phase space volume as state count: it is the invariance under equal-time coordinate changes that does.

⁴¹In general, units and physical dimensions keep track of some aspects of physical quantities, but not everything. For example, one can show that, dimensionally, pressure is equal to energy over volume. This relationship is actually useful when describing fluids. On the other hand, energy density is another useful quantity to, for example, measure the capacity of batteries. There is no relationship between the energy density of a battery and its pressure. Another puzzle: if one multiplies radians and meters, what is the resulting unit? Consider the area of the side of a cylinder section θrh where θ is the angular size of the section, r is the radius and h is the height of the cylinder. If we multiply θ and r , we get the length of the arc, which is a distance: we get meters. However, if we multiply θ and h , we do not get a length, we have an angle times length: we get meters times radians. The issue is that an angle is really a ratio between the length of the arc and the length of the radius. Only a multiplication by the correct radius will give back a distance.

to the first. Is this the only case? What happens if we use three or more variables? Suppose that q is the only variable that defines the unit system for a set of states, meaning that all other variables have derived units. Then under unit change $\hat{q} = \hat{q}(q)$ the units of all other variables must be uniquely defined. Suppose that we add a single additional constraint, that the measure of states is invariant. How many total variables can the state space have? We saw there must be more than one variable, or we cannot change the unit arbitrarily. Now suppose that we have three or more variables. Given that we have only two constraints, the choice of unit change and the area conservation, these are not enough to fully determine the transformation of all variables, and therefore the units of all variables. Therefore it would not be true the units of q fully determine the units of all variables. This tells us that, to have densities that are invariant under equal-time coordinate changes, a state space for a system that is fully characterized by a single unit must have exactly one additional quantity, conjugate momentum, that changes covariantly under unit change. We reach the following

Insight 1.189. *Degrees of freedom are two dimensional because only these allow coordinate invariant densities and count of configurations.*

Invariance over multiple degrees of freedom

We saw how things work for a single DOF, let's see how they generalize for multiple independent DOFs. Suppose states are identified by m variables ξ^a , but the unit system is fully identified by a subset q^i of n variables. This would be the case, for example, if we are studying particle trajectories as the spatial variables define the unit system. Moreover, suppose we assume **IND**, that the system is decomposable into independent DOFs. In this case, independence of the variables ties in with the independence of the DOFs. That is, if the unit of one variable depends on the unit of another, they cannot belong to independent DOFs; on the other hand, if two q^i define independent units, they must belong to independent DOFs by definition.

For example, suppose we have three degrees of freedom. Then we must have three variables q^x , q^y and q^z that define the units of each degree. Suppose we change q^x to \hat{q}^x while leaving the others unchanged. Given the premise and what we discussed above, there must be another variable p_x with inverse units. This must be an additional variable given that q^y and q^z must have different units. Given that the DOFs are independent, we should also be able to fix the value of both q^x and p_x and obtain a subspace identified by all the remaining variables. Now suppose that we change q^y by itself. By the same logic, there must be another variable p_y . This must have units of inverse q^y , and therefore cannot be any of the variables we already have. We repeat the logic and we find another variable p_z . At this point, there cannot be other variables: we would have to introduce units that do not depend on q^x , q^y and q^z , but this contradicts the assumption that q^i define the unit system. In general, then, if we have n degrees of freedom, we must have $2n$ state variables: the q^i that define the units and the conjugate quantities p_i expressed in inverse units.

We now want to introduce a form ω_{ab} that returns the count of independent configurations for an infinitesimal parallelogram. Here we simply use the same arguments that we gave discussing assumption **IND**. The number of independent configurations identified by a parallelogram within the same degree of freedom will be given by the area of the parallelogram. On the other hand, a parallelogram across degrees of freedom, for example formed by Δq^x and Δp_y , will not properly identify independent configurations. Note, in fact, that their product

is not unit invariant. This means ω_{ab} must match the familiar form. In short, assumption **IND** plus requiring that the count of states is the same for all equal-time observers gives us the structure of phase space. That is, condition

The space allows coordinate invariant distributions over equal-time
independent DOFs (PS-INV)

is equivalent to **PS-SYMP** and therefore to **PS-COV** and **PS-POI**.

Phase space and its structure, then, are neither a coincidence nor a choice. Phase space is the only space that allows us to define key concepts, like uniform distributions or reversibility, in an invariant way, which would otherwise be ill-defined over the continuum. In other words:

Insight 1.190. *The structure of phase space is exactly the structure needed to define state densities, thermodynamic entropy, information entropy and statistical uncertainty over continuous quantities in a way that satisfies the principle of relativity for equal-time observers.*

The structure of phase space, then, comes out from simply keeping track of unit dependence between variables. This should sound familiar, as unit dependence was used in condition **FKE-UNIT** to motivate why assumption **KE** should be implemented in its full form **FKE-LIN**. The existence of a metric tensor, of inertial mass, and more could be seen as stemming from assuming that the change of variables from dynamical to kinematic variables depended only on position, on q^i . Now we see that the very structure of phase space rests precisely on the fact that coordinates are those variables that define the units. This leads to the following observation: while taking the weak form **WKE-INV** of assumption **KE** leads to no mathematical inconsistencies, it would be physically inconsistent. On one side, when defining phase space, we are saying that the units of the coordinates q^i determine both the units of velocity and of momentum; on the other side, when introducing **WKE**, we claim that the coordinates q^i do not determine by themselves the unit transformation between velocity and momentum. The fact that we can write a mathematically consistent model that is physically inconsistent is evidence that we need a better mathematical specification of our physical theories. Namely, we need a way to encode unit relationship between variables.

We are also now in a position to answer another question: why are the laws of physics second order? That is, why is it that external forces specify the acceleration of the system, and not the velocity or the jerk (i.e. the derivative of the acceleration)? Note that, once the structure of phase space is derived, we know that, under **KE**, position and velocity fully determine the state, and, during the evolution, the change of state is fully determined by the acceleration. In other words, the laws of physics are second order exactly because the state is given by variable pairs. Since we have an explanation as to why the state space must obey that structure, that same explanation tells us why the laws are second order. Therefore, we have

Insight 1.191. *The principle of relativity ultimately requires the laws of motion to be second order.*

The principle of relativity, then, is responsible for much more than the invariance of the laws, it is also responsible for the structure of the space and the nature of the laws themselves.

Invariance under relative motion

So far, we have applied the principle of relativity only to transformations that leave time unchanged. For completion, let us apply it to the more usual case, when space and time coordinates are mixed. Consider phase space extended by time only. If we do not mix space and time variables, we keep the equal-time surfaces the same. Therefore we are only requiring that areas, density and entropy over each time slice remain the same. Moreover, we have no requirement to relate what happens on two different surfaces at equal time. That is, we will have the structure of phase space at each time, but there is no connection between these structures at different times. However, if we mix space and time variables, equal-time surfaces for one observer are not necessarily equal-time surfaces for another.

Now suppose we have a distribution that is uniform for one observer. The only way that it is uniform for all other observers is if it remains constant in time as well. In other words, the principle of relativity will imply the invariance of the form ω_{ab} even under space-time transformations, which means, as we saw, Hamiltonian mechanics. The only way that we can satisfy the principle of relativity, then, seems to assume deterministic and reversible motion. That is, condition

The space allows coordinate invariant distributions over independent
and temporal DOFs (DI-INV)

is equivalent to [DI-SYMP](#), and is therefore equivalent to Hamiltonian mechanics.

This may be surprising at first, but in retrospect it makes sense. The invariance among equal-time observers imposed that the density at each state was the same for everybody. In general, however, the equal-time surface for two generic observers will not be the same, but only intersect, in a region. Clearly, on that region the density must be the same for both, but what happens in the regions that are different? Well, if they still need to see the same density distribution, with the same entropy, then we must be able to map one region to the other in a bijective way. That is, the evolution must map each state of one surface to a state of the other surface while preserving the density. But this is exactly what deterministic and reversible evolution does. If we don't have this, if densities spread or if states are not mapped one-to-one, the two observers will see a different state.

As another way to see this, suppose we have a box of gas at equilibrium. Now suppose that at time t_0 we start heating it very slowly, such that we can assume the gas is at equilibrium at each time. The system, then, will increase its entropy. Suppose we stop the process at time t_1 , so that the temperature of the gas remains constant after that. Now, the equal-time surface of an observer boosted with respect to the gas will cross the original frame at different times, and therefore will see the gas at different temperatures in different places at the same time. The moving observer will see a system out of equilibrium. So, again, processes that are completely isolated, for which entropy is conserved, are the only processes that are truly relativistic.

This tells us that if we want to study a non-deterministic and/or non-reversible process, in a way that is frame independent, we need to be able to compare it to other processes that can be assumed to be deterministic and reversible. It is only by comparing the two that we are going to be able to ascertain precisely by how much the first one fails to be deterministic or reversible. Suppose, in fact, that we want to construct a coordinate system to study our non-deterministic and non-reversible process. Operationally, this means devising a system of rods and clocks so that we can correlate the quantities measured with our spatial and time references. Therefore we need some guarantee that, as time evolves, the spatial and temporal

references will have set spatial and temporal relationships. For example, if we leave a mark at a particular position, we may expect the mark to remain there; if the clock ticks uniformly in time, we may expect it to do so going forward. But this implicitly requires the existence of at least some deterministic and reversible process that can be used to distribute our references in space and time. If we do not have access to any deterministic process, no reliable coordinate system can be constructed. If the processes are not reversible, references will drift and the precision of our reference system will degrade since DR-UNC is not satisfied.

We close, then, with the following.

Insight 1.192. *Hamiltonian evolution over phase space is exactly the structure needed to define state densities, thermodynamic entropy, information entropy and statistical uncertainty over continuous quantities in a way that satisfies the principle of relativity.*

Gauge transformations

Throughout all this work we have seen that position q^i and momentum p_i are the true state variables: they allow us to identify states uniquely in all circumstances, properly count them and define densities that are invariant under coordinate changes and are defined at a particular instant in time. On the other hand, kinematic variables x^i and v^i identify states only under KE, they properly count states and define densities only in Cartesian inertial frames for systems subject to no forces, and velocity is technically defined over two infinitesimally close instants in time. Yet, under assumption KE, the information content of the two sets of variables are the same. Moreover, position and velocity are the variables that are more directly linked to experimentation: while we can measure velocity and kinetic momentum directly, conjugate momentum is not a direct observable.

We now pose the question: if velocity is the variable we actually measure, is conjugate momentum even uniquely defined? That is, is it possible to change momentum in a way that leaves velocity, and therefore the trajectories, unchanged? Since the position does not change, the new state variables can be written as (q, \hat{p}) where the new conjugate momentum $\hat{p}_i = \hat{p}_i(q^j, p_k)$ is a function of the old variables. We want the new variables to be conjugate, therefore the following relationships in terms of the Poisson brackets must hold:

$$\begin{aligned} \{q^i, q^j\} &= 0 \\ \{q^i, \hat{p}_j\} &= \delta_j^i \\ \{\hat{p}_i, \hat{p}_j\} &= 0. \end{aligned} \tag{1.193}$$

Using the second equation we find

$$\begin{aligned} \delta_j^i = \{q^i, \hat{p}_j\} &= \partial_{q^k} q^i \partial_{p_k} \hat{p}_j - \partial_{p_k} q^i \partial_{q^k} \hat{p}_j \\ &= \delta_k^i \partial_{p_k} \hat{p}_j - 0 \partial_{q^k} \hat{p}_j \\ \partial_{p_i} \hat{p}_j &= \delta_j^i. \end{aligned} \tag{1.194}$$

Integrating this last equation yields

$$\hat{p}_j = p_j + G_j(q^i) \tag{1.195}$$

where G_j are arbitrary functions of position. Using the third Poisson bracket, we have:

$$\begin{aligned}
0 &= \{\hat{p}_i, \hat{p}_j\} = \{p_i + G_i, p_j + G_j\} \\
&= \{p_i, p_j\} + \{p_i, G_j\} + \{G_i, p_j\} + \{G_i, G_j\} \\
&= 0 - \{G_j, p_i\} + \{G_i, p_j\} + 0 \\
&= -\partial_{q^i} G_j + \partial_{q^j} G_i \\
&= \text{curl}(G_j).
\end{aligned} \tag{1.196}$$

This G_j is a curl free field, it admits a scalar potential $f(q^i)$. Therefore the general transformation is of the form

$$\hat{p}_j = p_j + \partial_j f(q^i). \tag{1.197}$$

In other words, conjugate momentum is defined up to a gauge, in the same way that the electromagnetic potential is defined up to a gauge. In fact, these are essentially the same gauges.

Recall the expression

$$u^\alpha = \frac{1}{m} g^{\alpha\beta} (p_\beta - \mathbf{q} A_\beta). \tag{1.198}$$

If the four-velocity u^α and metric tensor $g^{\alpha\beta}$ are gauge independent, and the electromagnetic four-potential A_β is gauge dependent, then it must be that conjugate momentum p_β must be gauge dependent in the exact opposite way such that their difference is gauge independent. As another way to see this, suppose we have a charged particle in a field free region. In an inertial frame, we would write

$$p_i = mv^i. \tag{1.199}$$

But this not only requires us to use Cartesian coordinates, it also requires us to have chosen the gauge for the magnetic field such that $A_i = 0$. Choosing a gauge for momentum, then, is the same as choosing a gauge for the magnetic field.

If we look again at equation 1.195, note that what the gauge does is redefine the zero momentum at each point. The G_j field represents the new value of the old zero. The new zeros must form a curl free field because the new conjugate momenta still need to form independent DOFs. Recall that when recovering phase space, in equation 1.183, we saw that change of units of position left unspecified a potential change for zero momentum. The gauge freedom is exactly that change.

This tells us, in a more direct way, why assumption KE requires gauge theories. Given that the true observables are the kinematic variables, conjugate momentum is defined only up to a curl-free field that represents the arbitrary definition of zero momentum. Given that the vector potential of the interaction field tells us how the zero momentum is mapped to zero velocity, this inherits the same arbitrariness. The same relationship will be there in quantum mechanics, where the arbitrariness of the zero momentum will be implemented in the arbitrariness of absolute phases.

Entropy and dissipative evolution

Having seen the link between the principle of relativity and deterministic and reversible evolution, let's look a bit more closely at what happens during non-reversible evolution. We saw that a damped harmonic oscillator is deterministic but not reversible, as it does not preserve the count of states. The evolution, in fact, has an equilibrium, and regions around it get smaller and smaller. This presents a problem: if areas get smaller, the entropy goes down. But this is a dissipative system: shouldn't the entropy go up?

To understand the problem, suppose Alice takes a pendulum at rest, displaces the weight and lets it go. Suppose Alice told Bob, "In an hour, the position of the pendulum will be within 1 mm of the equilibrium." This would not give Bob much information as, after an hour, all the energy will have likely dissipated, no matter how it was initialized. Suppose Alice told Bob, "I let the weight go within 1 mm of the equilibrium." This gives Bob a lot more information, as the range of possible trajectories of the pendulum has greatly decreased. Because the motion is irreversible, statements at the same level of precision give more information in the past than in the future.

In terms of information, then, the difference is whether we are asking about how much we know about the position and momentum at a specific time, or how much we know about the system in terms of its overall evolution. If the system satisfies [DR](#), the two are the same. If it doesn't, we have a problem. Intuitively, we would want the state of the system to always provide the same amount of information about the system, but under non-deterministic or non-reversible evolution, this does not work. The amount of information changes in time and, for a boosted observer, it will also change in space and momentum. This brings to light a fundamental problem when defining states and systems: it would seem that such definition can only be given if, at least in some cases, the system can be studied under deterministic and reversible motion.

While the notion of state starts being problematic, the notion of evolution, however, is automatically invariant. A damped harmonic oscillator may not satisfy [DR](#), but it does satisfy [KE](#), which means that each state at a particular time will correspond to a particular trajectory and, given a region of phase space at a particular time, we will be able to quantify the flow of evolutions through that region. As time goes on, the evolutions become more concentrated, and therefore it is more difficult to tell them apart with measurements of the same precision. This is the indication that the system is not reversible. The proper indicator of irreversibility, then, is not the number of states within the region, but the number of evolutions over the region of phase space. This will increase for dissipative systems around the equilibria, and will remain unchanged over reversible evolution.

Note that the flow of evolutions through a surface was exactly the geometric interpretation we found for the principle of stationary action. In the extended phase space, the symplectic form could be understood as quantifying the flow of the displacement field instead of the count of states. Again, this seems to be hinting that the true nature of entropy is not the count of states, but the count of evolutions. We will explore these ideas when applying reverse physics to thermodynamics and statistical mechanics.

1.11 Reversing Newtonian mechanics

In this section we return to Newtonian mechanics, and see that assumption [KE](#) is the only assumption that characterizes Newtonian mechanics.

Inertia and forces

We have seen that all systems that satisfy assumptions **DR**, **IND** and **KE** are Lagrangian systems and, therefore, Newtonian given that the acceleration is a function of position and velocity. We also saw that some dissipative systems, like a particle under friction, do not satisfy **DR** though are Newtonian systems. So now the question is whether **IND** and **KE**, by themselves, are enough to characterize Newtonian mechanics.

The first task is to verify that all Newtonian systems satisfy the assumptions. The second law $F^i = ma^i$ in inertial frames makes a link between the dynamics, the force, and the kinematics, the acceleration. Moreover, the force is specified through kinematic variables, which makes it clear that the kinematics is enough to reconstruct the dynamics. Therefore assumption **KE** is satisfied somewhat trivially because Newtonian systems specify both kinematics and dynamics. Translating the dynamics from forces and masses into energy and conjugate momentum is complicated by the fact that there is no single way to decompose the total force into a conservative and non-conservative part. However, we can always set conjugate momentum equal to kinetic momentum in a Cartesian inertial frame.

As for **IND**, note that each new variable introduces its own force term, its own velocity and acceleration. The difficulty here is that, since the dynamics does not satisfy **DR**, this independence may not be kept over time. For example, the forces may be dissipative in such a way that the region with equilibria has lower dimensionality, introducing correlations that are effectively variable constraints. In other words, we do have a notion of independence at each time, but it is not the same notion throughout the evolution. Mathematically, if we are given the map from velocity to conjugate momentum at each time, we would be able to write a form ω_{ab} at each time. Given that the evolution is still differentiable, the Jacobian exists allowing us to map the count of configurations, and the form, back and forth in time. Since the evolution does not, in general, satisfy **DR**, the form is not conserved over time.

Having seen that Newtonian systems satisfy **IND** and **KE**, we have to show the converse, that all systems that satisfy those assumptions are Newtonian systems. We will need all the insights we gained in the previous section. We saw that the principle of relativity mixed with assumption **IND** recovers the structure of phase space. Assumption **KE** requires that the kinematics is specified by the dynamics, therefore the state at each time is enough to identify the trajectory of the system. This means that position and velocity must be invertible functions of position and momentum. This also means that the count of configurations can be calculated in terms of regions of position and velocity, which therefore must be well defined over time. Given that finite regions are mapped to finite regions, the evolution must be differentiable, an acceleration must be well defined and must be a function of position and velocity.

Given that position determines the unit system, condition **FKE-UNIT** is satisfied which is equivalent to **FKE-INNER**. Therefore we find the existence of locally inertial systems, in which a system not subjected to forces will travel in uniform linear motion. This recovers the first law. We already saw that acceleration was fully specified by position and velocity. The mass is the coefficient that recovers the correct count of states, which will also fix the units for the force. This recovers the second law.

Note that, in principle, a more complete account of Newtonian mechanics could be given. This has not been pursued given how Lagrangian and Hamiltonian mechanics play a much bigger role in field theories and in quantum mechanics.

1.12 Directional degree of freedom

When recovering the structure of phase space, we only relied on the premise that q^i define the units. We are now going to use that premise to define a directional degree of freedom, that is a degree of freedom that identifies a direction in space. We will see that, since a DOF must be composed of two variables, the only directional DOFs that are possible are those in three dimensional space.

Magnetic dipole

Spatial and temporal DOFs are of fundamental importance given that every object must be located in space and time. We now want to focus our attention on a different type of DOF, one that captures a direction in space. In general, objects are not completely spherically symmetric and can be distinguished by their orientation. This is also true for fundamental particles, as their intrinsic angular momentum, their spin, gives them an orientation that can be detected and manipulated through magnetic forces.

To study this case, let's consider a magnetic dipole. The magnetic moment μ_i can be written as

$$\mu_i = \frac{q}{2m} L_i, \quad (1.200)$$

where q is the electric charge, m is the mass and L_i the angular momentum. The Hamiltonian for a magnetic dipole subjected to a magnetic field B^i is given by

$$H = -\mu_i B^i = -\frac{q}{2m} L_i B^i. \quad (1.201)$$

To write the equations of motion, we need to know what the conjugate variables of L_i are or, equivalently, the form ω_{ab} or the Poisson brackets. Typically, one is given the Poisson brackets, therefore we will start from those. They are

$$\{L_i, L_j\} = L_k \epsilon_{ijk}. \quad (1.202)$$

In Hamiltonian mechanics, we can write the evolution of any quantity $f(q^i, p_i)$ as

$$d_t f = \partial_{q^i} f d_t q^i + \partial_{p_i} f d_t p_i = \partial_{q^i} f \partial_{p_i} H + \partial_{p_i} f (-\partial_{q^i} H) = \{f, H\}. \quad (1.203)$$

Therefore we have

$$\begin{aligned} d_t L_i &= \{L_i, H\} = \left\{ L_i, -\frac{q}{2m} L_j B^j \right\} = -\frac{q}{2m} B^j \{L_i, L_j\} = -\frac{q}{2m} B^j L_k \epsilon_{ijk} \\ d_t \vec{L} &= -\frac{q}{2m} \vec{B} \times \vec{L}. \end{aligned} \quad (1.204)$$

Without loss of generality, we can assume that the magnetic field is oriented along the z axis. We have

$$\begin{aligned} d_t L_x &= -\frac{q}{2m} B^z L_y \epsilon_{xzy} = \frac{q}{2m} B^z L_y \\ d_t L_y &= -\frac{q}{2m} B^z L_x \epsilon_{yzx} = -\frac{q}{2m} B^z L_x \\ d_t L_z &= -\frac{q}{2m} B^z L_i \epsilon_{zzi} = 0. \end{aligned} \quad (1.205)$$

If we integrate, assuming $L_i = L_i^0$ at time $t = 0$, we have

$$\begin{aligned}\omega &= -\frac{\mathfrak{q}}{2m}B^z \\ L_x(t) &= L_x^0 \cos \omega t - L_y^0 \sin \omega t \\ L_y(t) &= L_x^0 \sin \omega t + L_y^0 \cos \omega t \\ L_z(t) &= L_z^0\end{aligned}\tag{1.206}$$

which gives us the Larmor precession of the magnetic moment. This is clearly a Hamiltonian system, it obeys Hamilton's equations, but it seems to have three variables, the components of L_i , instead of an even number. Shouldn't we have conjugate pairs?

Note that the norm L of the vector is a constant of motion, only the angle changes. We can then write $L_i = Ln_i$ where n_i is a unit vector. The state space can be understood as the space of all possible vectors with the same magnitude, which is the surface of a 2-sphere. The size of an area A over the surface of a sphere of radius r can be measured using the solid angle

$$\Omega = \frac{A}{r^2},\tag{1.207}$$

which is already a unit independent quantity. The infinitesimal solid angle can be written in terms of the polar angle φ and the azimuthal angle θ

$$d\Omega = \sin \varphi d\varphi d\theta.\tag{1.208}$$

The two variables are not conjugate, as the area is not simply the product of the differentials. However, we can write

$$d\Omega = d(-\cos \varphi)d\theta.\tag{1.209}$$

The $\cos \varphi$ is simply the component along the z direction, while θ is the angle on the x - y plane. Changing the orientation of the solid angle, and relabeling for clarity $\theta = \theta^{xy}$ we have

$$d\Omega = dn_z d\theta^{xy},\tag{1.210}$$

which means that the angle on a plane and the component orthogonal to the plane are conjugate. If we pick the variable θ^{xy} as the coordinate q and L_z as the conjugate p , we have two variables with the right geometric relationship and with the right units. Therefore

$$\begin{aligned}\omega_{\theta^{xy} L_z} &= -\omega_{L_z \theta^{xy}} = 1 \\ \{\theta^{xy}, L_z\} &= -\{L_z, \theta^{xy}\} = 1\end{aligned}\tag{1.211}$$

We can recover L_i with the following expressions

$$\begin{aligned}L_x &= \cos \theta^{xy} \sqrt{L^2 - L_z^2} \\ L_y &= \sin \theta^{xy} \sqrt{L^2 - L_z^2} \\ L_z &= L_z.\end{aligned}\tag{1.212}$$

We can also recover the Poisson brackets for the components

$$\begin{aligned}
\{L_x, L_y\} &= \partial_{\theta^{xy}} L_x \partial_{L_z} L_y - \partial_{L_z} L_x \partial_{\theta^{xy}} L_y \\
&= -\sin \theta^{xy} \sqrt{L^2 - L_z^2} \sin \theta^{xy} \left(\frac{-L_z}{\sqrt{L^2 - L_z^2}} \right) \\
&\quad - \cos \theta^{xy} \left(\frac{-L_z}{\sqrt{L^2 - L_z^2}} \right) \cos \theta^{xy} \sqrt{L^2 - L_z^2} \\
&= \sin^2 \theta^{xy} L_z + \cos^2 \theta^{xy} L_z = L_z
\end{aligned} \tag{1.213}$$

$$\begin{aligned}
\{L_y, L_z\} &= \partial_{\theta^{xy}} L_y \partial_{L_z} L_z - \partial_{L_z} L_y \partial_{\theta^{xy}} L_z \\
&= \cos \theta^{xy} \sqrt{L^2 - L_z^2} - 0 = L_x
\end{aligned} \tag{1.214}$$

$$\begin{aligned}
\{L_z, L_x\} &= \partial_{\theta^{xy}} L_z \partial_{L_z} L_x - \partial_{L_z} L_z \partial_{\theta^{xy}} L_x \\
&= 0 - (-\sin \theta^{xy} \sqrt{L^2 - L_z^2}) = L_y
\end{aligned} \tag{1.215}$$

We can verify that Hamilton's equations work as before. If the magnetic field is aligned along the z direction, we have

$$\begin{aligned}
H &= -\frac{\mathbf{q}}{2m} L_z B^z \\
d_t \theta^{xy} &= \partial_{L_z} H = -\frac{\mathbf{q}}{2m} B^z \\
d_t L_z &= -\partial_{\theta^{xy}} H = 0.
\end{aligned} \tag{1.216}$$

Given initial conditions θ_0^{xy} and L_z^0 , we have

$$\begin{aligned}
L_z(t) &= L_z^0 \\
\theta^{xy}(t) &= \omega t + \theta_0^{xy} \\
L_x^0 &= \cos(\theta_0^{xy}) \sqrt{L^2 + (L_z^0)^2} \\
L_y^0 &= \sin(\theta_0^{xy}) \sqrt{L^2 + (L_z^0)^2} \\
L_x(t) &= \cos \theta^{xy}(t) \sqrt{L^2 - L_z^2(t)} = \cos(\omega t + \theta_0^{xy}) \sqrt{L^2 - (L_z^0)^2} \\
&= \cos(\omega t) \cos(\theta_0^{xy}) \sqrt{L^2 - (L_z^0)^2} - \sin(\omega t) \sin(\theta_0^{xy}) \sqrt{L^2 - (L_z^0)^2} \\
&= L_x^0 \cos \omega t - L_y^0 \sin \omega t \\
L_y(t) &= \sin \theta^{xy}(t) \sqrt{L^2 - L_z^2(t)} = \sin(\omega t + \theta_0^{xy}) \sqrt{L^2 - (L_z^0)^2} \\
&= \sin(\omega t) \cos(\theta_0^{xy}) \sqrt{L^2 - (L_z^0)^2} + \cos(\omega t) \sin(\theta_0^{xy}) \sqrt{L^2 - (L_z^0)^2} \\
&= L_x^0 \sin \omega t + L_y^0 \cos \omega t
\end{aligned} \tag{1.217}$$

which recovers the precession.

Let us sum up what we learned. The magnetic dipole is described by a single DOF, where the conjugate variables are the angle on a plane and the component of the angular momentum perpendicular to the plane. A constant force does not correspond to a constant angular acceleration, but to a constant angular velocity. Now the question is how and to what extent this can all be generalized.

Generalizing to directional quantities

The above treatment works without modification regardless of whether the magnetic dipole describes a small solenoid, a small rotating charge distribution, or the spin of a single particle. Spin was originally thought to be due to a rotational motion of a particle, hence the name. However, this would require, for example, an electron to be spinning at velocities that would exceed the speed of light, therefore this view is no longer favored. But if spin is not due to a rotation, why does it have the same properties as angular momentum? Is it a coincidence?

Let us call directional quantity a quantity l_i defined only by a direction in space, meaning that the magnitude of the quantity $|l_i| = l$ is fixed, an intrinsic feature of the object, while the direction can change. A magnetic dipole, generated by either spin or rotating charge, can be considered a directional quantity. The question now is whether the directional nature is enough to recover its properties.

If we want to keep track of a direction in space, the angle θ^{xy} of the component of the direction on a plane is a natural variable to use. Therefore θ^{xy} is a natural choice for a unit variable. Because of the structure of phase space, we will have a conjugate quantity L_z , that we will call directional momentum. Given that a directional quantity is fully determined by a direction in space, the count of possible configurations is proportional to a solid angle. As we saw before in the context of a magnetic dipole, this means that the conjugate quantity L_z must be proportional to the component of the directional quantity perpendicular to the plane where θ^{xy} is defined. Therefore $l_z = kL_z$ for some constant k .

Recall now that the product of conjugate quantities must always be in units of phase space. We already noted, in fact, that if a coordinate is dimensionless, then its conjugate must have dimensions of angular momentum. Therefore directional momentum L_z must have dimension of an angular momentum, regardless of what the directional quantity l_i is about. Thus, even if the directional quantity is not “an amount of rotation” defined by an angular momentum, we can still write $l_i = kL_i$, where L_i is expressed in units of angular momentum. That is, a directional momentum L_i arises naturally every time we have a directional quantity. In the case of a magnetic dipole $k = \frac{\mu_i}{L_i} = \frac{q}{2m}$ is the gyromagnetic ratio.

We want to stress that for a directional quantity both variables, the directional momentum L_z along a given direction z and its angle θ^{xy} on the perpendicular plane, are effectively tracking a single quantity. This is different from what happens in a spatial DOF, where the coordinate q determines position while the conjugate p determines conjugate momentum, and therefore the velocity, which are different properties of the system. This fact tells us that a force acts differently in the two cases. A force for a spatial DOF will, in general, affect momentum, and therefore velocity, which means it will impart an acceleration on position. However, for a directional quantity the only thing that can be changed is a direction, therefore a force can only impart a change in direction, a directional velocity. To sum up, a force over position will impart a constant spatial acceleration while a force over a direction will impart a constant directional velocity.

For a rotating rigid body, the situation may seem different. We do have an angle and an angular velocity, and the angular momentum is the angular velocity times the moment of inertia, much like kinetic momentum is mass times velocity. Given that angular momentum is conserved, angular velocity is conserved and a force needs to be used to change the angular velocity. Therefore, at first glance, the situation seems different from the magnetic dipole. On the other hand, we are dealing with an angular momentum in both cases, therefore the

situation can't be that different.

The issue here is that there are two directions: one is the direction of the axis of rotation, the other is the orientation of the object on the plane of rotation. While the second one constantly changes, the first may be fixed. Note that the ability to discern the second assumes the ability to discern parts of the rigid body. If the object is too small with respect to our resolution, the orientation angle cannot be defined. On the other hand, even if we study the overall object, we would still be able to discern the direction and magnitude of the angular momentum by performing a rotation. In fact, if the rotation is performed along a direction different from the direction of the angular momentum, a torque will need to be applied to perform the rotation. The greater the angular momentum, the greater the torque. We are now in a case similar to the magnetic dipole, where a constant force is needed to impart a constant velocity.

We have seen that if we want to describe a directional quantity, and only a directional quantity, with state variables, there is only one way to do it, and it will lead to a notion of directional momentum. Spin and angular momentum are, in our definition, two types of directional momentum. Is this the only way to do it? That is, why does a directional DOF only include the directional quantity, without the directional velocity?

Suppose we want to describe a particle, in the sense of an infinitesimally small part of an object, which is not only characterized by a position, but also by a directional quantity. If we pick a specific time t , the position and direction at that particular time are independent variables, in all senses: we can choose units independently, we can choose from all possible configurations of each variable and a choice for one variable does not constrain the choice for another variable. It is true that we will need to relate the direction identified by the directional variable to the reference frame used by the position coordinates. Yet, we are not forced to choose a particular plane for the directional angle θ .

The situation between velocity and the directional velocity, however, is different. Contrast the following two cases. In the first, we are in an inertial frame with a particle at rest, whose directional quantity changes with a constant directional velocity. In the second, we are in a rotating frame with a particle at the center, whose directional quantity changes with a constant directional velocity. If we look only at the directional quantity, these two cases are indistinguishable: we need to know which frame we are in. But knowing the frame, ultimately, means knowing the definition of our spatial variables. That is, there is a definitional dependency between the spatial variables and the directional variables, which are, in the end, defined in space. We are free to choose the relationship between the variables at a specific time. But to relate the directional quantity at different times, we must know how the spatial coordinates change in time. Therefore, the directional quantity is an independent quantity with respect to the spatial DOFs, but the directional velocity is not. We reach the following:

Insight 1.218. *An independent directional DOF must be restricted to a directional quantity.*

We concluded that an independent directional DOF only includes a directional quantity, so the structure we have for the magnetic dipole is an instance of a general structure. However, note how the directional quantity fits exactly in one DOF. Is this necessary? That is, could we have a directional quantity that fits in more than one pair of conjugate quantities? In other words, is it possible to have directional quantities that break up into multiple independent DOFs?

As we said before, identifying a direction in three dimensions is equivalent to identifying a point on a 2-sphere. If we had an n -dimensional space, a direction would correspond to a point on an $(n - 1)$ -sphere. That sphere would have to allow a description in terms of conjugate variables with a suitable ω_{ab} . It is a result of symplectic geometry that the 2-sphere is the only sphere to allow that structure. Therefore

Insight 1.219. *An independent directional DOF can only exist in a three dimensional space.*

In other words, three dimensional space is special as it is the only space that allows us to talk about directions as forming an independent degree of freedom. This is yet another example of how simple unassuming premises have significant consequences.

1.13 Infinitesimal reducibility

We have seen that the structure of phase space (i.e. the symplectic structure) is exactly the structure needed to describe density, count of states and entropy in a coordinate invariant way. This implicitly assumes that states are points on a differentiable manifold, that is they are fully identified by real variables that only allow differentiable changes of variables. We will see that this corresponds to the assumption of infinitesimal reducibility.

Particles as infinitesimal parts

At this point, we have ample evidence that the correct fundamental object that describes classical systems is not a point on phase space. All physical objects have finite spatial extent and assuming otherwise leads to problems. A finite mass concentrated at a point is incompatible with general relativity, as it would puncture space-time and not travel along geodesics,⁴² and with quantum mechanics, as it would imply uniform spread in momentum thus requiring infinite kinetic energy.⁴³ The correct physical object, instead, is a distribution over phase space. It explains the very structure of phase space (i.e. the distribution has to be frame invariant), it explains why the laws of motion are differentiable (i.e. they transport a density, not just points) and why they follow Hamiltonian mechanics for isolated, i.e. deterministic and reversible, systems. We can therefore conclude that

The state of a classical system is given by a distribution over phase
space. (IR-DIST)

As usual in reverse physics, we want to find equivalent physical assumptions for the same statement.

While point particles should not be considered fundamental classical objects, the whole of classical mechanics assumes them and uses them. Therefore we should understand exactly what they represent.

One way to understand point particles is as an approximation. When we calculate the trajectory of the earth around the sun, the motion of a cannonball or the average displacement of a molecule under Brownian motion, we know that the object is not actually point-like. But

⁴²Mathematically, it would mean that all mass-energy would be concentrated in points of infinite curvature, leaving the rest of space-time flat with no mass.

⁴³Mathematically, the wave-function would be a δ -function, which is not even an element of the Hilbert space L^2 .

since the body is rigid, and assuming that the effect of external forces on the different parts of the body can be neglected, we can get away with studying the motion of the center of mass. Clearly, we need to check that the assumption holds during the whole motion: if the distance between a comet and the sun becomes smaller than the radius of the sun, the sun cannot be assumed to be point-like anymore. The approximation will also fail if the force exerted on different parts of a body is greater than the force that keeps it together.

The point particle approximation is not unique to classical mechanics. Quantum mechanics, through the Ehrenfest theorem, also allows for a point particle approximation, which will fail if the wave-function is spread over a region where the potential changes are non-negligible. Since the approximation is not unique to classical physics, this characterization of a point particle does not help us understand the fundamental constitutive assumptions of classical mechanics.

When discussing the nature of Hamiltonian evolution, however, we saw that states were better understood as infinitesimal regions of phase space. This means that classical particles, at least this more fundamental version of them, can't be understood as a standalone physical object: they should really be thought of as the limit of some recursive subdivision. That is, classical particles should really be thought of as an infinitesimal part of something bigger, and the 'part' in 'particle' should be understood literally.

The actual classical object, then, is not the infinitesimal part, but the whole object that can be, in principle, divided into infinitesimal parts. In the same way that, when discussing differential topology, we had quantities over finite regions that could be broken into the infinitesimal contributions, the whole object, together with its notion of state, can be understood as made up of infinitesimal parts.

A classical system can be thought of as being made of infinitesimal parts, called particles. (IR-INF)

All classical systems follow the assumption, not just mechanical systems. Continuum mechanics and fluid dynamics assume that the materials are a continuum of infinitesimally small parts; classical electromagnetism assumes that the electromagnetic radiation can be decomposed into arbitrarily small signals at all frequencies. Moreover, the assumption clearly fails in quantum mechanics. In that case, we cannot talk about the state of a part of an electron; materials are not a continuum of infinitesimally small parts; the intensity of electromagnetic radiation cannot be made arbitrarily small.

Divisible vs reducible

Before going forward, we need to be clear as to what "being made of" actually means. For example, if we say that a table is made of a horizontal top and four legs, we may mean that we can take the table apart and study its components independently, or that we can describe the whole table by describing the top and the legs. While often one can do both, these are actually different properties. That is, divisibility, the ability to "cut" an object into independent parts, is not the same as reducibility, the ability to describe an object in terms of parts. Let us go through some examples.

Suppose we have a planarian worm and we divide it in half. After some time, the tail will regrow a head, and the head will regrow a tail. We will be left with two worms. A planarian worm is divisible into two worms, in the sense that we have a process by which a worm is

divided into two worms, but it is not reducible to two worms, in the sense that describing one worm is not the same as describing two worms.

Now, suppose we have a magnet. We can describe it by describing its north and south pole. Now suppose we divide it in half. Each half will be a new magnet with its own north and south pole. A magnet is reducible to a north and a south pole, meaning that describing a magnet is the same as describing its poles, but it is not divisible into a north and a south pole, meaning that we do not have a process that can separate the two.

We can also find differences between divisibility and reducibility with fundamental particles. Suppose we have a muon. If we wait some time, it will decay into an electron, a neutrino and an antineutrino. A muon divides itself into the three particles, but it is not reducible to the three particles: the state of a muon is not equivalent to the state of an electron and two neutrinos.

Suppose we have a proton. This is not a fundamental particle and it is described in terms of quarks and gluons. However, if we try to separate one of its quarks, the interaction energy will create new quark-antiquark pairs and the proton will divide into multiple hadrons. A proton is reducible to quarks and gluons, but it is not divisible into them.

Divisibility here means that we have a physical process that starts with the overall system and ends with parts that can now be independently manipulated. Mathematically, if S_C is the state space of the overall system and S_A and S_B are the state spaces of the parts, we have a time evolution map $U_t : S_C \rightarrow S_A \times S_B$ that takes an initial state of the composite and returns a state for each part. This is not what we want.

Reducibility means that describing the whole is the same as describing the parts. Mathematically, the state space of the whole system $S_C = S_A \times S_B$ is exactly the Cartesian product of the parts. In other words, what we are taking apart is not really the object itself, but its state. This is the type of partitioning we are going to use.

To make reducibility more concrete and intuitive, suppose we have a ball. We can throw the ball, study the motion of the whole ball. We can also take a red marker, make a red dot on the ball, and study the motion of the red dot. We say the ball is reducible because studying the motion of the whole ball is equivalent to studying the motion of all possible red dots. It is infinitesimally reducible if we assume that we can make the red dot arbitrarily small. This is the property we are interested in. Conversely, suppose we have an electron. We can study the motion of the electron, but we cannot make a red dot on the electron. We have no process at our disposal that can tag part of an electron so that we interact with and study only that part. Whenever we interact, we interact with the whole electron. This means that the assumption does not hold in that case.

Condition [IR-INF](#), then, is a property of classical systems and only classical systems, and it is clear that [IR-DIST](#) implies [IR-INF](#). But is the converse true? That is, if something can be thought of as being made of infinitesimal parts, does it follow that the state of those infinitesimal parts should be represented by points in phase space? In other words, is condition [IR-INF](#) enough to characterize classical systems?

Classical systems and infinitesimal reducibility

Suppose we have a system that satisfies condition [IR-INF](#). Under this premise, we have two state spaces: S_C for the full system and S_P for the infinitesimal parts, the particles. Each state for the full system must tell us exactly how much of the system is in each region of S_P . That

is, for each state $s \in S_C$ we have an associated real valued set function⁴⁴ $f(U) \in [0, 1]$ that for every region $U \in S_P$ returns the fraction of the system in that region of particle state space.⁴⁵ Moreover, reducibility implies f to be additive, meaning that if U is the disjoint union of two regions U_1 and U_2 , then $f(U) = f(U_1) + f(U_2)$. Because we are assuming infinitesimal reducibility, this must hold not just in the finite case, but also in the countable case. Mathematically, f is a bounded measure. Since the state of the whole system is fully identified by the state of the parts, for each state $s \in S_C$ there is one and only one f . We can understand S_C as the space of such functions.

Depending on whether we are considering a specific instance of a particular system or an ensemble of similarly prepared systems, the value of the function f will have a different physical meaning. It could be understood as the fraction of a single system that has a particular property (i.e. half of the ball is to the right of the line), or it could be understood as the probability that a particular instance of an ensemble has a particular property (i.e. half the time the ball is to the right of the line). From now on, we will assume we are talking about an actual system, but all that we say will apply to the ensemble case as well.

The fact that $f(U)$ is real valued is also implied by condition **IR-INF** since the parts are infinitesimal and therefore there can't be a smallest increment.⁴⁶ It also implies that the fraction $f(\{x\})$ associated to a single particle state $x \in S_P$ must be zero. Otherwise we would be associating a finite non-zero fraction to an infinitesimal part x , which would make it not infinitesimal. The particle state space S_P of an infinitesimally reducible system, then, must be charted by continuous variables.⁴⁷

So far, we have found that, under condition **IR-INF**, the state space S_C of a system is the space of all possible functions $f(U) \in [0, 1]$ where $U \subseteq S_P$ is a subset of all the possible particle states. Furthermore, the particle state space S_P must be charted by continuous variables, it must be a manifold. We now need to recover the notion of differentiability and of invariant density.

To be able to fully characterize the state of the whole system by the state of the parts, we need to be able to quantify the fraction of the system for each particle state. We saw that the fraction for each particle state is zero, which makes sense because particle states are limits. Therefore the fraction associated to a particle state should be a limit as well: it should be a fraction density, similar to mass or charge density at a point. A fraction density will quantify the fraction present over a region of particle states, meaning that is expressed as a fraction over count of particle states. We therefore need to be able to quantify how many states there are in each region of S_P ; we must be able to say which regions have equal number of states. Over a discrete space this would be trivial, we would just count the number of points, but over the continuum finite ranges have infinitely many points and counting points

⁴⁴A set function is a function that takes sets as an argument.

⁴⁵Technically, the region U must be a Borel set, as these are the regions that are associated with an experimental procedure. These details are established in the physical mathematics part of the book.

⁴⁶Technically, this only limits the output of $f(U)$ to be a bounded dense linear order. However, the function f must be closed under arbitrary countable addition, bringing in all the limits, meaning that the ordering must be complete. Lastly, the whole system can always be divided into countably many pieces, which tells us that the order must have a countable dense subset. Since the order is bounded, dense, complete and has a countable dense subset, it is order isomorphic to $[0, 1] \subset \mathbb{R}$.

⁴⁷Again, technically we only found that the space must be, in some sense, dense. To show that it is "dense like the reals", we need the link between topology and experimental verifiability that we establish in physical mathematics.

doesn't work. We must have another set function $\mu(U)$, which will also be countably additive as disjoint regions will identify entirely different states, making μ another measure. However, μ will be bounded only from below (i.e. no set can have fewer than zero states) as S_P can have potentially infinitely many possible states.

Note that, because of the infinitesimal reducibility assumption, we will be able to find smaller and smaller subsets of S_P . The value of μ will keep decreasing and, in the limit of infinitesimal subdivision, we will have $\mu(\{x\}) = 0$. That is, a single particle state counts as zero in terms of number of states. In a way, the math is already telling us that particle states do not exist literally and that the assumption of infinitesimal reducibility shouldn't be taken literally. It is a simplifying assumption and should be understood as such. In the same vein, as we can define distributions over phase space that have arbitrarily narrow spread, the entropy of those distributions can be an arbitrarily low number, which will tend to $-\infty$ in the limit of a δ -distribution. This is clearly a problem, since thermodynamics requires the entropy to be non-negative. These two problems are in fact the same problem, since the entropy for a uniform distribution is the logarithm of the count of states. In other words, we shouldn't be surprised that classical mechanics fails for small objects (i.e. narrow distributions), or for ensembles with low entropy: it is exactly when condition [IR-INF](#) fails.

Continuing our derivation, we have seen that both the measure f for the fraction of the system and the measure μ for the count of particle states are zero for sets with single points. More in general, we must have that whenever $\mu(U) = 0$, then $f(U) = 0$. That is, we cannot assign a finite fraction to a set of states that has no finite state count. That would, again, mean that we have a finite fraction associated to an infinitesimal subdivision, which is physically untenable. Mathematically, this means that f is absolutely continuous with respect to μ , and therefore we can define a density $\rho = d_\mu f$ such that $f(U) = \int_U \rho d\mu$.⁴⁸ That is, the state of the whole system can be described by a density over the particle states.

As we said, S_P is a manifold. The measure μ is a feature of that manifold, in the sense that it depends only on the properties of the particle states, and it will be the same for all states of the whole system. Of all possible state variables that we can use to chart the manifold, then, it is convenient to choose those that can express μ as a density of states over units of the state variables. Mathematically, we are asking that the Lebesgue measure induced by the variable $\xi : S_P \rightarrow \mathbb{R}$ is absolutely continuous with respect to μ . This will also require that the variables are differentiable with respect to each other: it will lead to the differentiable structure. That is

Insight 1.220. *The differentiable structure of a state space is exactly the ability to express state count and fractions as densities over said states.*

Lastly, both f and μ must not depend on the choice of state variables. Therefore the proper expression of $\rho = d_\mu f$ must also not depend on the choice of state variables. This, plus assumption [IND](#), as we have seen, requires the existence of the form ω , and recovers phase space, the symplectic manifold. Therefore we have recovered [IR-DIST](#).

Condition [IR-DIST](#) is therefore equivalent to condition [IR-INF](#). Classical mechanics describes exactly those systems that follow

Assumption IR (Infinitesimal Reducibility). *The state of the system is reducible to the state of its infinitesimal parts. That is, specifying the state of the whole system is equivalent*

⁴⁸Mathematically, ρ is the Radon-Nikodym derivative.

to specifying the state of its parts, which in turn is equivalent to specifying the state of its subparts and so on.

Assumption **IR** is, therefore, the constitutive assumption of classical mechanics. Assumptions **IR**, **IND** and **DR** are the constitutive assumptions of the Hamiltonian formulation. Assumptions **IR**, **IND** and **KE** are the constitutive assumptions of Newtonian mechanics. Assumptions **IR**, **IND**, **DR** and **KE** are the constitutive assumptions of Lagrangian mechanics.

We have found all the constitutive assumptions of all formulations of classical mechanics. We now are guaranteed that every result of classical mechanics is, one way or the other, explained by those assumptions. These are the only physical ideas that are strictly required to understand all the general aspects of classical mechanics. There is nothing else. Therefore the goal of reverse physics for classical mechanics is reached.

Classical uncertainty principle

Before concluding, let us turn again to the problem of zero measure on individual particle states. As we said, a distribution over phase space can be made arbitrarily narrow and, in the limit of a δ -function, the support will be a single point. Since the entropy is the logarithm of the phase-space volume, which is zero for a single state, the entropy of a δ -function is minus infinity.

Negative entropy is not compatible with the third law of thermodynamics, which states that the entropy must be non-negative. Therefore, on thermodynamics grounds, the spread of a distribution cannot be made infinitesimal. This means that any distribution that is physically meaningful must have a finite spread in position and momentum. Conceptually, this sounds very close to the uncertainty principle of quantum mechanics, so it's worth looking at it more closely.

First of all, since we saw that units have an important role, consider the expression for thermodynamic entropy $S = k_B \log W$ where k_B is the Boltzmann constant and W is a volume of phase space. Given that a logarithm must take pure numbers, the expression is dimensionally incorrect. To correct it, we should take

$$S = k_B \log \frac{W}{W_0} \quad (1.221)$$

where W_0 can be understood as the volume for which the thermodynamic entropy is zero. As we saw, units of momentum are units of configurations, of areas in phase space, over units of position. Therefore we can choose units of momentum such that W_0 is, numerically, one, but not dimensionally. To keep track of the value and units of W_0 separately, then, let's write

$$W_0 = w_0 \mathbf{1}^{qp} \quad (1.222)$$

where w_0 is the pure numeric value and $\mathbf{1}^{qp}$ is the product of the units of q and p . If we substitute in the above expression, we find

$$S = k_B \log \frac{W}{w_0 \mathbf{1}^{qp}} = k_B \log \frac{W}{\mathbf{1}^{qp}} - k_B \log w_0 = k_B \log W - k_B \log w_0 \quad (1.223)$$

where $\log W$ is the logarithm of the numeric value of W . This clearly show that changing w_0 , the numeric value of W_0 , changes the zero for the entropy.

To make the Gibbs-Shannon entropy consistent with the previous definition, we must set

$$I[\rho] = -k_B \int \rho \log(W_0 \rho) dqdp. \quad (1.224)$$

If we take a uniform distribution ρ_U over a region U with volume W , we find

$$\begin{aligned} I[\rho] &= -k_B \int \rho_U \log(W_0 \rho_U) dqdp = -k_B \int_U \frac{1}{W} \log \frac{W_0}{W} dqdp \\ &= k_B \log \frac{W}{W_0} \frac{\int_U dqdp}{W} = k_B \log \frac{W}{W_0} \end{aligned} \quad (1.225)$$

Similarly to before, we find

$$\begin{aligned} I[\rho] &= -k_B \int \rho \log(W_0 \rho) dqdp = -k_B \int \rho \log(\rho \mathbf{1}^{qp}) dqdp - k_B \int \rho \log w_0 dqdp \\ &= -k_B \int \rho \log(\rho \mathbf{1}^{qp}) dqdp - k_B \log w_0 \\ &= -k_B \int \rho \log \rho dqdp - k_B \log w_0 \end{aligned} \quad (1.226)$$

which again shows that w_0 changes the zero for entropy.

We are now ready to study the relationship between spreads in classical phase space and entropy. We want to find the distribution ρ with zero entropy that minimizes the uncertainty. To do that, we set up a minimization problem using Lagrange multipliers. We want to minimize the product of the variances $\sigma_q^2 \sigma_p^2 \equiv \int (q - \mu_q)^2 \rho dqdp \int (p - \mu_p)^2 \rho dqdp$, where μ_q and μ_p are the mean position and momentum, under two constraints: ρ integrates to 1 and its entropy is zero. We have:

$$\begin{aligned} L &= \int (q - \mu_q)^2 \rho dqdp \int (p - \mu_p)^2 \rho dqdp \\ &\quad + \lambda_1 \left(\int \rho dqdp - 1 \right) + \lambda_2' \left(-k_B \int \rho \ln(W_0 \rho) dqdp - 0 \right) \\ &= \int (q - \mu_q)^2 \rho dqdp \int (p - \mu_p)^2 \rho dqdp \\ &\quad + \lambda_1 \left(\int \rho dqdp - 1 \right) + \lambda_2 \left(- \int \rho \ln \rho dqdp - \ln w_0 \right) \\ \delta L &= \int \delta \rho \left[(q - \mu_q)^2 \sigma_p^2 + \sigma_q^2 (p - \mu_p)^2 + \lambda_1 - \lambda_2 \ln \rho - \lambda_2 \right] dqdp = 0 \\ \lambda_2 \ln \rho &= \lambda_1 - \lambda_2 + (q - \mu_q)^2 \sigma_p^2 + \sigma_q^2 (p - \mu_p)^2 \\ \rho &= e^{\frac{\lambda_1 - \lambda_2}{\lambda_2}} e^{-\frac{(q - \mu_q)^2 \sigma_p^2}{\lambda_2}} e^{-\frac{\sigma_q^2 (p - \mu_p)^2}{\lambda_2}} \end{aligned}$$

We recognize the distribution as the product of two independent Gaussians. We can therefore use the standard expression and calculate the entropy, which must be zero. We find:

$$\begin{aligned} \rho &= \frac{1}{2\pi\sigma_q\sigma_p} e^{-\frac{(q-\mu_q)^2}{2\sigma_q^2}} e^{-\frac{(p-\mu_p)^2}{2\sigma_p^2}} \\ I[\rho] &= k_B \ln \left(2\pi e \frac{\sigma_q \sigma_p}{W_0} \right) = 0 = k_B \ln 1 \\ \sigma_q \sigma_p &= \frac{W_0}{2\pi e} \end{aligned}$$

We have found that any distribution $\rho(q, p)$ with non-negative entropy (i.e. that satisfies the third law of thermodynamics) satisfies the inequality

$$\sigma_q \sigma_p \geq \frac{W_0}{2\pi e}. \quad (1.227)$$

Moreover, the inequality is saturated (i.e. becomes an equality) for Gaussian distributions.

Compare this with the Heisenberg uncertainty principle of quantum mechanics, which tells us that the every state satisfies the inequality

$$\sigma_q \sigma_p \geq \frac{\hbar}{2}, \quad (1.228)$$

which is saturated by Gaussian states. Given the parallel, we call equation 1.227 the classical uncertainty principle. Note that the classical uncertainty principle has a clear physical explanation: the third law of thermodynamics. Given that all quantum states have non-negative entropy, the same explanation can be taken for quantum mechanics as well.

This result also tells us the problem with Maxwell's demon. In this setup, a gas is separated into two chambers with a door. A demon that sees the exact position and momentum of all particles can open the door letting a particle go only from one specific side to the other, decreasing the entropy of the system. But knowing the exact position of all particles corresponds to having access to a state of minus infinite entropy, which is forbidden by the third law. Naturally, if we have access to a source of minus infinite entropy, we can always decrease the entropy of any system by a finite amount without violating the second law: given that minus infinity minus a finite amount is equal to minus infinity, the entropy is technically not decreasing.

This final result should really make it clear that the foundations of statistical mechanics and thermodynamics are not separate from the foundations of classical mechanics and quantum mechanics. In any theory, states are better understood as ensembles, and pure states should be understood as the most precise ensemble that can be prepared in the theory.

1.14 Summary

Let's step back and sum up all that we learned by applying the reverse physics approach to classical mechanics. If we start from scratch, the first assumption that we need to set is Infinitesimal Reducibility [IR](#). This tells us that the domain of classical physics is those objects that can be thought of as being made of arbitrarily small parts. For these objects, specifying the state of the whole system is equivalent to specifying the state of all the infinitesimal parts. In this context, a particle is an infinitesimal part, the limit of recursive reduction of parts into smaller parts. Mathematically, the assumption tells us that the state space of the particles is a differentiable manifold, together with a volume measure that defines the count of states for each region. The state space of a classical object is a distribution over such manifold.

We then add assumption [IND](#) that tells us that the system is decomposable into independent degrees of freedom. This means that not only are we able to count states over regions of state space, but we must be able to count configurations along each DOF. Moreover, the count of configurations over multiple DOFs must be the product of the count over each DOF. The only way to obtain this structure, so that the count is defined independently of the units used to describe states, is for the manifold to be even dimensional, and for any state variable

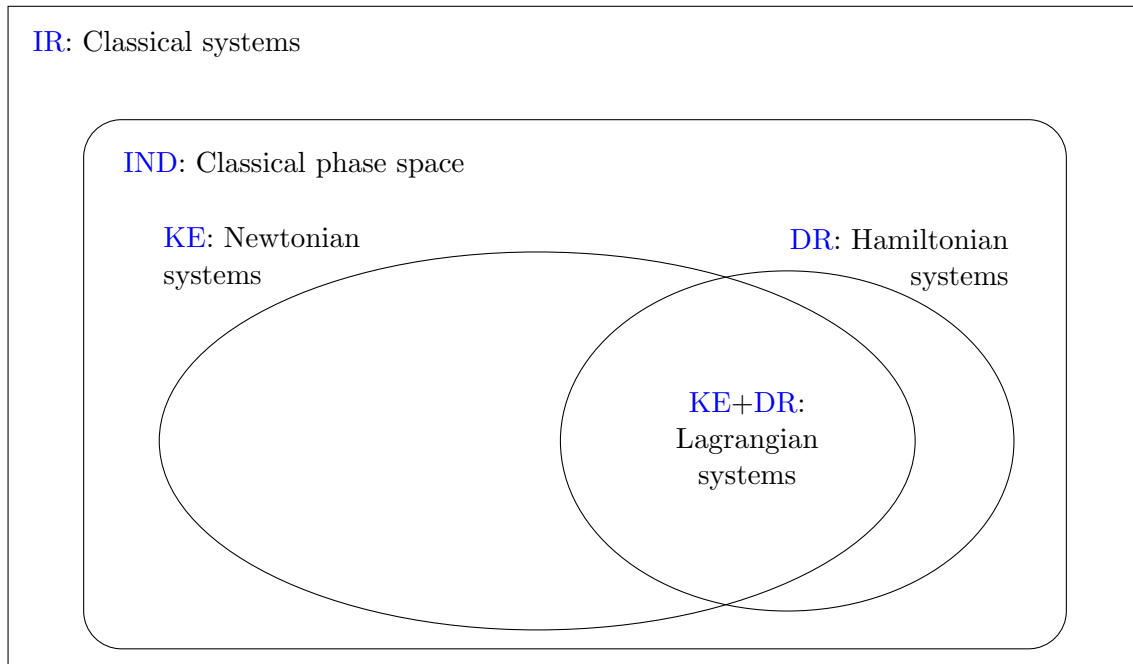


Figure 1.9: Relationship between the constitutive assumptions and the different formulations of classical mechanics. All classical systems satisfy infinitesimal reducibility. The independence of DOF recovers classical phase space. An ideal gas, with state variables $[PVT]$ is an example of a classical system that does not satisfy IND. Another example is a systems with non-holonomic constraints. Newtonian system are fully characterized by spatial trajectories. Hamiltonian systems are characterized by deterministic and reversible dynamics. Lagrangian systems are characterized by both.

q^i that defines an independent unit, there is a conjugate state variable p_i whose units are count of states over units of the corresponding q^i . This recovers the structure of classical phase space. Mathematically, this gives us the structure of a symplectic manifold, where the symplectic form ω_{ab} counts the configurations over the infinitesimal parallelogram defined by two vectors.

If we add assumption **DR** that tells us that the evolution is deterministic and reversible, we recover Hamiltonian mechanics. For a system to be deterministic and reversible, in fact, we must map not only each initial state to one and only one final state, but we must map initial to final regions while preserving the state count. Moreover, the notion of independence must be preserved, and therefore the way configurations are counted over independent DOFs must also be preserved. Under these conditions, the displacement vector field, which tells us how states move in phase space, allows a potential H , which corresponds to the Hamiltonian. This can be generalized to the time-dependent case, which recovers relativistic features even without the notion of a metric tensor. Mathematically, the form ω_{ab} must be preserved, meaning deterministic and reversible evolution is a symplectomorphism. Each symplectomorphism can be characterized by a function H , the Hamiltonian.

If, instead, we add assumption **KE** that tells us that the dynamics is recoverable from the kinematics, we recover Newtonian mechanics. If the dynamics is recoverable from the kinemat-

ics, the momentum must be a function of position and velocity. The units of position are the only independent units, and therefore define the units of momentum and also the transformation between densities over dynamic and kinematic variables. This forces the transformation between position and velocity to be linear. If no forces are present, we can find coordinates such that the linear transformation is simply a multiplication by a constant. That is, $p_i = mv^i$. These are the inertial frames in Cartesian coordinates and m is the inertial mass. In this case, assumption **DR** applies as well, and, since momentum is constant, the velocity is constant. If a force is present, this can be expressed in terms of position and velocity, recovering $F = ma$.

If we take both assumption **DR** and **KE**, Lagrangian mechanics is recovered. The action is the line integral over the potential θ_a of the form $\omega_{ab} = -\partial_a \wedge \theta_b$, and since the vector potential is unphysical, so is the action. The variation of the action is physical as, by Stokes' theorem, it will give the surface integral of ω_{ab} , which is zero if and only if the path is always tangent to the direction of motion, which happens only if the path is an actual evolution of the system.

All the elements of classical mechanics, both physically and mathematically, can be understood just in terms of these assumptions and the concepts that they require.

Part II

Physical Mathematics

Physical mathematics is an approach to the mathematical foundations of physics that seeks to construct mathematical structures strictly from axioms and definitions that can be rigorously justified from physical requirements. This is in contrast to current approaches typically followed in **mathematical physics**, which take tools developed within mathematics and apply them to physics or physics-inspired problems.

If our goal is to fully rederive physical theories from physical assumptions, we need to have a precise mapping between physical objects and mathematical ones. Understanding the axioms and definitions of the mathematical tools used in a physical theory, then, is not just “mathematical detail” of no concern to the physicist, but rather the precise stipulation of properties that certain physical objects must have under suitable, possibly simplifying, assumptions. In this sense, there is no “correct” structure in a mathematical sense, because the correct structure is the one suited to the physical problem at hand.

It should be clear that mathematicians are generally ill-equipped to determine whether mathematical structures are physically significant. As David Hilbert stated, “Mathematics is a game played according to certain simple rules with meaningless marks on paper.” Regarding mathematical axioms, Bertrand Russell claimed, “It is essential not to discuss whether the first proposition is really true, and not to mention what the anything is, of which it is supposed to be true.” Mathematics knows the rules of everything but the meaning of nothing. It is therefore unreasonable to expect that the foundations of mathematics, by themselves, can provide any foundation for physics.

In the same way that elaborate correct mathematical theories stem from minimal correct mathematical theories (not elaborate incorrect mathematical theories); that large living creatures grow from small living creatures (not large dead creatures); sophisticated physically meaningful theories come from simple physically meaningful theories (and not from sophisticated meaningless ones). Meaningfulness, like correctness or aliveness, is not something that can be imposed after the fact. Therefore the only way to develop physically meaningful mathematical structures is to develop them from scratch: we cannot simply take higher level mathematical objects and “sprinkle meaning”, an interpretation, on top.

The goal of physical mathematics, then, is to find how to turn physical assumptions into precise mathematical requirements, such that we are guaranteed to know what exactly each mathematical object represents and under which physical conditions.

A new standard for scientific rigor

From the above discussion, it follows that the standard of rigor mathematicians have developed for their field is not sufficient for the purpose of physical mathematics. Mathematics only deals with formal systems, whose starting points are a set of definitions and rules that are taken as is. At that point, correctness of the premise cannot be established, only self-consistency. Therefore mathematics fails to deal with the most delicate and interesting parts of the foundations of physics: the physical assumptions and how they are encoded into the formal framework. We therefore need rules and standards for rigorously handling the informal parts of the framework and, since there are no guidelines for this, we set our own standard.

We call an **axiom** a proposition that brings new objects or new properties of established objects within the formal framework. A **definition**, instead, is a proposition that further characterizes objects and properties already present in the formalism. **An axiom or a definition is well posed only when it is clear what the objects represent physically and what aspects are captured mathematically.** Therefore each axiom and definition

is composed of two parts. The first characterizes the objects and properties within the informal system, tells us what they represent physically. The second part, typically preceded by “Formally”, characterizes the part that is captured by the formal system. Axioms and definitions are followed by a **justification** when it is necessary to explain why the elements in the informal system must be mapped into the formal system in the way proposed. Some definitions are purely formal and as such do not require justifications. As this argument spans both the formal and informal systems, this cannot be a mathematical proof in the modern sense. In particular, the justification for an axiom must argue why those objects must exist.

The above standard makes sure we have a perfect identification between formal and informal objects. All mathematical symbols correspond to physical objects and all the relevant physical concepts are captured by the math. All subsequent propositions and proofs, then, can be carried out in the formal system, where it is easier to check for consistency and correctness. However, all the proofs can, if needed, be translated into the informal language and given physical meaning.

Chapter 1

Verifiable statements and experimental domains

In this chapter we lay the foundations for our general mathematical theory of experimental science: a formalism that is broad enough to be applied to any area of scientific investigation. It is based on the idea of **verifiable statements**, assertions that can be experimentally shown to be true. Whether it is physics or chemistry, economics or psychology, medicine or biology, the goal is always to find some objective truth about the natural world that is supported by experimental data.

We group verifiable statements into **experimental domains** which represent the list of all possible verifiable answers to a particular scientific question. From those, we define **theoretical domains** which add those statements that, though not directly verifiable, are associated to an experimental test with no guarantee of termination. Within each theoretical domain, we find those particular statements that, if true, imply the truthfulness or falsehood of all other statements. We call these the **possibilities** of the domain as they identify the complete descriptions that are admissible. To answer a scientific question, then, is to find which possibility is consistent with experimental data: the one that correctly predicts the result of all experimental tests.

We'll see how the above organization always exists on any given set of verifiable statements. That is, it is a fundamental structure for all sciences. We'll also see that these concepts are deeply intertwined with fundamental mathematical tools: experimental domains map to topologies while theoretical domains map to σ -algebras. These two core mathematical structures provide the foundation for differential geometry, Lie algebras, measure theory, probability theory and other mathematical branches that are heavily used in physics and other sciences.

As a consequence of this connection, we can build a more precise, intuitive and insightful understanding of what these mathematical structures are meant to represent in the scientific world. It also reveals why these mathematical tools are so pervasive and successful in science.

1.1 Statements

Statements, like *“the mass of the electron is 511 ± 0.5 keV”* or *“that animal is a cat”*, will be the cornerstone of our general mathematical theory of experimental science. In this section we will outline the basic definitions that allow us to combine statements into other statements

(e.g. “*that animal is black*” and “*that animal a cat*” gives “*that animal is a black cat*”) and to compare their content (e.g. “*the mass of the electron is 511 ± 0.5 keV*” and “*the mass of the electron is 0.511 ± 0.0005 MeV*” are equivalent).

We will start from a somewhat different starting point than what is customary in order to address a few issues. The first is that we need to develop a formal framework to handle the relationship of statements that are not themselves formally defined.¹ The second issue is that the truth values in our context are in general found experimentally, not through deduction. The role of the logic framework is to keep track of the consistency between the possible hypothetical cases. That is, we need relationships that capture the idea of causal relationship, that one is true because the other is true, and not merely by coincidence.² We remind the reader that the mathematical sections, highlighted with a green bar on the side, can be skipped without loss of conceptual understanding in case one is not interested in all the details.

As a starting point, we need to define what science is: the systematic study of the physical world through observation and experimentation. We therefore introduce the principle of scientific objectivity that will guide us throughout this work. This states that science is universal, non-contradictory and evidence based.

Principle of scientific objectivity. Science is universal, non-contradictory and evidence based.

Consider assertions like “*jazz is marvelous*” or “*green and red go well together*”. These are not objective: there is no agreed upon definition or procedure for what constitutes marvelous music or good color combination. Because of their nature, they can’t be the subject of scientific inquiry. This does not mean that marvelous music or good color combinations do not exist or are not worth studying.³ What the principle tells us is simply that if we choose to do science, we are limiting ourselves to those assertions that are either true or false (i.e. non-contradictory) for everybody (i.e. universal): assertions that have a single truth value. We call these assertions statements: they are the basic building blocks of a scientific description as only these can be studied scientifically.

Logical consistency, though, is not just a property of individual statements. Consider the following two sentences:

“*the next sentence is true*”
 “*the previous sentence is false*”

Each assertion, by itself, would be fine but their combination makes it impossible to assign truth values to both in a way that is logically consistent. For this reason we group statements into logical contexts, sets of statements for which it is possible to assign truth values to all in a way that is logically consistent.

Definition 1.1. The *Boolean domain* is the set $\mathbb{B} = \{\text{FALSE}, \text{TRUE}\}$ of all possible truth values.

¹This is the main reason we cannot simply use the tools of mathematical logic.

²In contrast, in standard logic “the moon is a satellite of earth” implies “5 is prime”.

³In fact, one can argue that most of the things that make life worth living (e.g. love, friendship, arts, purpose and so on) defy objective characterization and, therefore, that science gives us certain truth about trifling matters.

Axiom 1.2 (Axiom of context). A **statement** s is an assertion that is either true or false. A **logical context** \mathcal{S} is the collection of all those statements that share a common semantics and have, therefore, well defined logical relationships. Formally, a logical context \mathcal{S} is a collection of elements called statements upon which is defined a function $\text{truth} : \mathcal{S} \rightarrow \mathbb{B}$.

Justification. As science is universal and non-contradictory, it must deal with assertions that have clear meaning, well-defined logical relationships and are associated with a unique truth value. A priori, we only assume these objects exist, simply because we cannot proceed if we do not. A posteriori, we see that a particular set of statements works in practice, which shows that science could indeed be done with those statements. We are therefore justified to assume the existence of sets of assertions that have the aforementioned properties. We call a logical context such a group of assertions and we represent it formally as a set \mathcal{S} . We call statement each assertion within a logical context and represent it formally with an element s . We say $s \in \mathcal{S}$ if the statement s belongs to the logical context \mathcal{S} .

Given a context, we are also justified to assume the existence of a function $\text{truth} : \mathcal{S} \rightarrow \mathbb{B}$ such that $\text{truth}(s) = \text{TRUE}$ for all $s \in \mathcal{S}$ that are true for everybody and $\text{truth}(s) = \text{FALSE}$ for all $s \in \mathcal{S}$ that are false for everybody. In fact, note that no $s \in \mathcal{S}$ can be both true and false for everybody as it would be contradictory. Note that no $s \in \mathcal{S}$ can be both true for some and false for others as it would not be universal. Note that no $s \in \mathcal{S}$ can be neither true nor false for everybody. Suppose it were. Then its truth can never be, even potentially, settled with experimental evidence. Therefore either $\text{truth}(s) = \text{TRUE}$ or $\text{truth}(s) = \text{FALSE}$ for all $s \in \mathcal{S}$.

Note that the statement is the concept asserted, not the words used to express the concept. A sentence in a particular language is neither necessary nor sufficient to define a statement. Consider the English language and take the sentence “*Snoopy is a dog*”. The truth will depend on whether or not a fictional character qualifies as a dog. The sentence, by itself, is not enough to determine the truth value. Conversely, take the idea that a particular animal is a dog. We can express the concept in Italian as “*quell’animale è un cane*” without using an English sentence. Therefore English sentences are neither necessary nor sufficient to express a statement. This will be true of any other language. Therefore, the basic notion of statement is considered prime, independent of its expression. Moreover, all statements are considered equally prime.

Also note that logical consistency is not a property of an individual statement but rather of a set of statements. Consider the two statements: “*the next sentence is true*” and “*the previous sentence is false*”. The two statements together are not logically consistent as assuming one to be true leads to a proof for it to be false and vice-versa. If we changed the latter statement to “*the previous sentence is true*” then both statements would be true and logically consistent. Therefore logical consistency is defined on the logical context, and each statement has to be defined as belonging to a context.

Finally, note that the existence of a truth function on the context imposes logical consistency on the logical context without worrying about the details of how this is achieved. □

The idea of statements has their origin in the philosophical tradition of classical logic, “*Socrates is a man*” being a classic example. Any language can be used to form them, formal or natural, as indeed any language is used in practice. This means we are not going to care what

particular syntax (i.e. symbols and grammar rules) is used.⁴ In fact, even a grammatically incorrect statement is fine as long as the intent is clear. On the other hand, we are going to care about the semantics of the statements (i.e. their content and meaning). Therefore we will consider “*Socrates is a man*” and “*Socrate è un uomo*” to be the same statement because they provide the same assertion but in different languages.

Moreover, when we say “*Socrates is a man*” it has to be clear who is Socrates and what a man is. If it weren’t, we would have no idea what to experimentally test and how. This is also important because the mere content of a set of statements puts constraints on what can be found to be true or false. Consider the statement “*that cat is a swan*”. There is nothing to experimentally test here: based on the definitions of cat and swan we know the statement can never be true, no matter what particular cat we are considering. The statement provides no new information. Consider, instead, the statements “*that animal is a mammal*” and “*that animal is a bird*”. Based on the content, each of the statements can be found true or false separately, but they can’t be found true together simply by how mammals and birds are defined.

We want each logical context to keep track not only of the truth value of each statement, but also of which truth combinations are possible merely because of the content. Given a logical context \mathcal{S} we will call an assignment for \mathcal{S} a map $a : \mathcal{S} \rightarrow \mathbb{B}$ that for each statement gives a truth value. For example, consider the following table:

“ <i>that animal is a cat</i> ”	“ <i>that animal is a mammal</i> ”	“ <i>that animal is a bird</i> ”	...
T	T	T	...
T	T	F	...
F	F	T	...
F	T	F	...

Each line represents an assignment for all statements, each represented by a column. The set of all assignments is the set of all functions from \mathcal{S} to \mathbb{B} , which in set theory notation is $\mathbb{B}^{\mathcal{S}}$, and corresponds to all possible permutations. Some assignments, like the one in the first line, are not consistent with the content of the statements. We will call possible the ones that are allowed and impossible the ones that aren’t. Naturally, the truth must be one of the possible assignments. As the context itself needs to tell us which assignments are possible, we can imagine it comes equipped with the set $\mathcal{A}_{\mathcal{S}} \subseteq \mathbb{B}^{\mathcal{S}}$ of all possible assignments for that context.⁵

Some statements may be allowed to either be true or false while others may only be allowed one option. We will call certainty a statement that can only be true, like “*that cat is an animal*”. We will call impossibility a statement that can only be false, like “*that cat is a*

⁴In general, statements in this context are not necessarily well formed formulas, predicates or similar concepts in the context of mathematical or propositional logic. Scientific investigation in the broad sense of learning from experimentation predates math and formal languages: information about agriculture, astronomy, metallurgy, botany and the like were collected and used even before the written word. Moreover, cognitive scientists have shown that children start using deliberate experimentation at a very young age to understand the world around them, even before their speech is fully developed. Ultimately, that knowledge is encoded in the language of electrical and biochemical signals. Formal languages are indeed extremely helpful in that they allow us to be more precise and to better keep track of possible inconsistencies, but ultimately one always needs natural language to give meaning and context to the mathematical symbols.

⁵In the same way that we do not try to capture how statements are constructed, we are not going to capture how logical consistency is established. We just assume a mechanism is available and therefore one can check whether an assignment is possible or not.

swan”. We will call a contingent statement one that can be either true or false as its truth is contingent on the assignment, like “that animal is a cat”.

Note that the semantics, the meaning of the statements, plays an important role in defining the possible assignments but, in general, does not define the truth values.⁶ For example, even if the meaning of “the next race is going to be won by Secretariat” is clear, and so is its logical relationship to “the next race is going to be lost by Secretariat”, we may be none the wiser about its truthfulness.

Definition 1.3. Given a set of statements, an **assignment** associates a truth value with each statement. Formally, an assignment for a logical context \mathcal{S} is a map $a : \mathcal{S} \rightarrow \mathbb{B}$, an element of $\mathbb{B}^{\mathcal{S}}$. An assignment for a set of statements $S \subseteq \mathcal{S}$ is a map $a : S \rightarrow \mathbb{B}$ while an assignment for a statement $s \in \mathcal{S}$ is a truth value $t \in \mathbb{B}$.

Justification. Given that an assignment for a set $S \subseteq \mathcal{S}$ associates a truth value to each statement, it is identified by a map $a : S \rightarrow \mathbb{B}$. The assignment for a single statement is given by a single truth value $t \in \mathbb{B}$. \square

Axiom 1.4 (Axiom of possibility). A **possible assignment** for a logical context \mathcal{S} is a map $a : \mathcal{S} \rightarrow \mathbb{B}$ that assigns a truth value to each statement in a way consistent with the content of the statements. Formally, each logical context comes equipped with a set $\mathcal{A}_{\mathcal{S}} \subseteq \mathbb{B}^{\mathcal{S}}$ such that $\text{truth} \in \mathcal{A}_{\mathcal{S}}$. A map $a : \mathcal{S} \rightarrow \mathbb{B}$ is a possible assignment for \mathcal{S} if $a \in \mathcal{A}_{\mathcal{S}}$.

Justification. The meaning associated to each statement in the context may prevent some assignments from being logically consistent. Consider, for example, the statements “that animal is a cat” and “that animal is a dog”. Given that an animal cannot be both a cat and a dog, an assignment that associated true to both statements would not be logically consistent. Given that the context must give clear meaning to all statements, it must be able to clarify whether an assignment $a \in \mathbb{B}^{\mathcal{S}}$ is consistent with the meaning of all statements. We call these possible assignments.

For each logical context, then, we are justified to assume the existence of a set $\mathcal{A}_{\mathcal{S}} \subseteq \mathbb{B}^{\mathcal{S}}$ such that $a \in \mathcal{A}_{\mathcal{S}}$ if and only if a is a possible assignment. \square

Definition 1.5. Let $S \subseteq \mathcal{S}$ be a set of statements. Then $a : S \rightarrow \mathbb{B}$ is a possible assignment for S if there exists $\bar{a} \in \mathcal{A}_{\mathcal{S}}$ such that $\bar{a}(s) = a(s)$ for all $s \in S$. Let $s \in \mathcal{S}$ be a statement. Then $t \in \mathbb{B}$ is a possible assignment for s if there exists $\bar{a} \in \mathcal{A}_{\mathcal{S}}$ such that $\bar{a}(s) = t$.

Justification. Let $S \subseteq \mathcal{S}$ be a set of statements. If $a : S \rightarrow \mathbb{B}$ is a possible assignment then there must be a way to assign the remaining statements in the domain in a way that is logically consistent. Therefore there must be an $\bar{a} \in \mathcal{A}_{\mathcal{S}}$ such that $\bar{a}(s) = a(s)$ for all $s \in S$. Similarly, $t \in \mathbb{B}$ is a possible assignment for s if there exists $\bar{a} \in \mathcal{A}_{\mathcal{S}}$ such that $\bar{a}(s) = t$. This justifies the definition. \square

Definition 1.6. Statements are categorized based on their possible assignments.

- A certain statement, or **certainty**, is a statement \top that must be true simply because of its content. Formally, $a(\top) = \text{TRUE}$ for all possible assignments $a \in \mathcal{A}_{\mathcal{S}}$.

⁶In other formalisms the semantics is said to define the truth values, not in ours.

- An impossible statement, or **impossibility**, is a statement \perp that must be false simply because of its content. Formally, $a(\perp) = \text{FALSE}$ for all possible assignments $a \in \mathcal{A}_S$.
- A statement is **contingent** if it is neither certain nor impossible.

Justification. Since a certainty must be true, no possible assignment can assign it to be false. Therefore $a(\top) = \text{TRUE}$ for all $a \in \mathcal{A}_S$. Similarly, an impossibility must be false and no possible assignment can assign it to be true. Therefore $a(\perp) = \text{FALSE}$ for all $a \in \mathcal{A}_S$. This justifies the definitions. \square

Corollary 1.7. A statement $s \in \mathcal{S}$ can only be exactly one of the following: impossible, contingent, certain.

Proof. Let $s \in \mathcal{S}$ be a statement. If it is contingent, by definition, it is neither certain nor impossible. If it is not contingent, it is either certain or impossible. If s is certain, then $a(s) = \text{TRUE}$ for all possible assignments $a \in \mathcal{A}_S$. This means $a(s) \neq \text{FALSE}$ for all possible assignments $a \in \mathcal{A}_S$ and therefore s is not impossible. If s is impossible, then $a(s) = \text{FALSE}$ for all possible assignments $a \in \mathcal{A}_S$. This means $a(s) \neq \text{TRUE}$ for all possible assignments $a \in \mathcal{A}_S$ and therefore s is not certain. \square

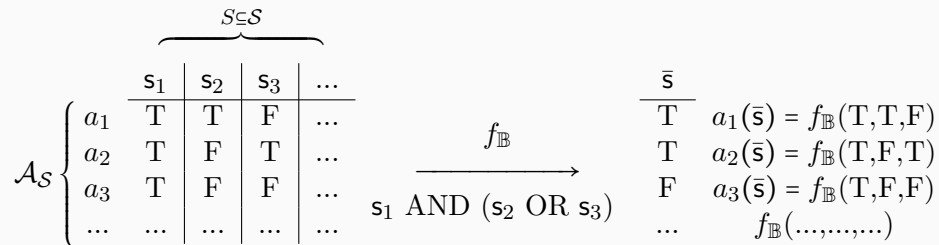
Next we want to keep track of statements whose truth depends on the truth of other statements. Consider “that animal is a cat” and “that animal is not a cat”: if the first one is true then the second is false and vice-versa. In this sense, the second statement depends on the first. Therefore, in general, a statement depends on other statements if its truth is determined by the truth values of the other statements in every possible assignment.

Since statements are intangible, there are no limits to the number of arguments one statement may depend on. For example, consider the statement “the mass of the electron is 511 ± 0.5 keV” and the set of all the statements of the form “the mass of the electron is exactly x keV” with $510.5 < x < 511.5$. If any of the latter is true then the original statement is true as well. Given that x is a real number, that is uncountably many statements so the original statement can be seen as a function of uncountably many statements. Therefore we will assume we can always create a statement that arbitrarily depends on an arbitrary set of statements.

Definition 1.8. Let $\bar{s} \in \mathcal{S}$ be a statement and $S \subseteq \mathcal{S}$ be a set of statements. Then \bar{s} **depends on** S (or it is a function of S) if we can find an $f_{\mathbb{B}} : \mathbb{B}^S \rightarrow \mathbb{B}$ such that

$$a(\bar{s}) = f_{\mathbb{B}}(\{a(s)\}_{s \in S})$$

for every possible assignment $a \in \mathcal{A}_S$. We say \bar{s} depends on S **through** $f_{\mathbb{B}}$. The relationship is illustrated by the following diagram:



Axiom 1.9 (Axiom of closure). *We can always find a statement whose truth value arbitrarily depends on an arbitrary set of statements. Formally, let $S \subseteq \mathcal{S}$ be a set of statements and $f_{\mathbb{B}} : \mathbb{B}^S \rightarrow \mathbb{B}$ an arbitrary function from an assignment of S to a truth value. Then we can always find a statement $\bar{s} \in \mathcal{S}$ that depends on S through $f_{\mathbb{B}}$.*

Justification. Let $S \subseteq \mathcal{S}$ be a set of statements and let $f_{\mathbb{B}} : \mathbb{B}^S \rightarrow \mathbb{B}$ be an arbitrary function. Consider the statement $\bar{s} =$ “this statement is true if and only if the function $f_{\mathbb{B}}$ applied to the truth values of S is true”. The meaning of \bar{s} is determined by the meaning of the statements in S , and, for every possible assignment $a \in \mathcal{A}_{\mathcal{S}}$, \bar{s} is assigned the truth value $f_{\mathbb{B}}(\{a(s)\}_{s \in S})$ that is the same for everybody. Therefore \bar{s} shares a common semantics with the statements in S , and has a well-defined logical relationship with all other statements in \mathcal{S} . Given that a context, by definition, contains all such statements, we are justified to assume $\bar{s} \in \mathcal{S}$. \square

Remark. If it can be proven that the closure can always be performed, that proof, together with the definition of closure, would replace the axiom. This would be preferable but, given that the closure would be on possibly infinite elements, this is a delicate problem that we have not addressed yet.

Corollary 1.10. *Functions on truth values induce functions on statements. Formally, let I be an index set and $f_{\mathbb{B}} : \mathbb{B}^I \rightarrow \mathbb{B}$ be a function. There exists a function $f : \mathcal{S}^I \rightarrow \mathcal{S}$ such that*

$$a(f(\{s_i\}_{i \in I})) = f_{\mathbb{B}}(\{a(s_i)\}_{i \in I})$$

for every indexed set $\{s_i\}_{i \in I} \subseteq \mathcal{S}$ and possible assignment $a \in \mathcal{A}_{\mathcal{S}}$.

Proof. Given I and $f_{\mathbb{B}} : \mathbb{B}^I \rightarrow \mathbb{B}$ we can construct $f : \mathcal{S}^I \rightarrow \mathcal{S}$ as follows. Given an indexed set $\{s_i\}_{i \in I} \in \mathcal{S}^I$, let $S \subseteq \mathcal{S}$ be the set of elements in the indexed set and define $\bar{f}_{\mathbb{B}} : \mathbb{B}^S \rightarrow \mathbb{B}$ such that $\bar{f}_{\mathbb{B}}(\{a(s)\}_{s \in S}) = f_{\mathbb{B}}(\{a(s_i)\}_{i \in I})$. Then by axiom 1.9 we can find a statement \bar{s} that depends on S through $\bar{f}_{\mathbb{B}}$ and we can set $f(\{s_i\}_{i \in I}) = \bar{s}$. We have $a(f(\{s_i\}_{i \in I})) = f_{\mathbb{B}}(\{a(s_i)\}_{i \in I})$ for all indexed sets and for all possible assignments. \square

To better characterize truth functions, we borrow ideas and definitions from Boolean algebra which is the branch of algebra that operates on truth values. Boolean algebra is fundamental in logic and computer science, since every digital circuit ultimately is implemented on two-state systems (e.g. high/low voltage, up/down magnetization). The most fundamental elements in that algebra are the following three simple operations: negation (i.e. logical NOT), conjunction (i.e. logical AND) and disjunction (i.e. logical OR).

Suppose $s_1 =$ “the sauce is sweet” and $s_2 =$ “the sauce is sour”. We can apply the three operations to make this table:

Operator	Gate	Symbol	Example
Negation	NOT	$\neg s_1$	“the sauce is not sweet”
Conjunction	AND	$s_1 \wedge s_2$	“the sauce is sweet and sour”
Disjunction	OR	$s_1 \vee s_2$	“the sauce is at least sweet or sour”

Table 1.1: Boolean operations on statements.

Most languages, natural or symbolic, typically already provide similar operations, as

the examples show. Technically, though, we should consider the ones defined here as meta-operations that are defined outside the language of the statements. For example, “ x is the position of a ball” \wedge “ $\frac{d^2x}{dt^2} = -g$ ” stitches together an English statement with a calculus statement into a new statement that is neither. This kind of mix should be allowed as it does happen in practice.

Definition 1.11. The **negation or logical NOT** is the function $\neg : \mathbb{B} \rightarrow \mathbb{B}$ that takes a truth value and returns its opposite. That is: $\neg \text{TRUE} = \text{FALSE}$ and $\neg \text{FALSE} = \text{TRUE}$. We also call negation $\neg : \mathcal{S} \rightarrow \mathcal{S}$ the related function on statements.

Definition 1.12. The **conjunction or logical AND** is the function $\wedge : \mathbb{B} \times \mathbb{B} \rightarrow \mathbb{B}$ that returns TRUE only if all the arguments are TRUE. That is: $\text{TRUE} \wedge \text{TRUE} = \text{TRUE}$ and $\text{TRUE} \wedge \text{FALSE} = \text{FALSE} \wedge \text{TRUE} = \text{FALSE} \wedge \text{FALSE} = \text{FALSE}$. We also call conjunction $\wedge : \mathcal{S} \times \mathcal{S} \rightarrow \mathcal{S}$ the related function on statements.

Definition 1.13. The **disjunction or logical OR** is the function $\vee : \mathbb{B} \times \mathbb{B} \rightarrow \mathbb{B}$ that returns FALSE only if all the arguments are FALSE. That is: $\text{FALSE} \vee \text{FALSE} = \text{FALSE}$ and $\text{TRUE} \vee \text{FALSE} = \text{FALSE} \vee \text{TRUE} = \text{TRUE} \vee \text{TRUE} = \text{TRUE}$. We also call disjunction $\vee : \mathcal{S} \times \mathcal{S} \rightarrow \mathcal{S}$ the related function on statements.

Proposition 1.14. A logical context \mathcal{S} is closed under negation, arbitrary conjunction and arbitrary disjunction.

Proof. Negation, arbitrary conjunction and arbitrary disjunction are particular truth functions. Their output always exists by axiom (1.9). \square

Now we have all the elements to define when two statements have the same logical content. Consider the two statements “that animal is a bird” and “that animal has feathers”: since all birds and only birds have feathers they give us the same information. Consider “the mass of the electron is $511 \pm 0.5 \text{ keV}$ ” and “the mass of the electron is $0.511 \pm 0.0005 \text{ MeV}$ ”: they represent the same measurement but in different units. So, how can we express the fact that two statements s_1 and s_2 give us the same information? The idea is that they can never be assigned opposite truth values. If we assigned true to the first, then the second must be true as well. If we assigned false to the first, then the second must be false.⁷

Definition 1.15. Two statements s_1 and s_2 are **equivalent** $s_1 \equiv s_2$ if they must be equally true or false simply because of their content. Formally, $s_1 \equiv s_2$ if and only if $a(s_1) = a(s_2)$ for all possible assignments $a \in \mathcal{A}_{\mathcal{S}}$.

Justification. If two statements must be equally true or false simply because of their content, then their value must be the same in all possible assignments, which justifies the definition. \square

Again, we want to stress that this notion of equivalence is not based on the truth value

⁷This technique allows us to do something analogous to model theory. Two statements are equivalent if their truth is equal for all consistent truth assignments, which in our framework play the part of the models of model theory. But in our context the assignments are only hypothetical: there isn’t a model in which “this is a cat” and another in which “this is a dog”. There is only one truth value, the one we find experimentally.

(i.e. whether the statements happen to be both true or false) or on whether they are the same statement (i.e. whether they assert the same thing): it is based on the possible assignment (i.e. whether there exists a possible assignment in which the two statements have a different truth value) and therefore on the content of the statements. We sum up the difference in the following table.

Relationship	Symbol	First statement	Second statement	Also known as
Statement equality	$s_1 = s_2$	“Swans are birds”	“I cigni sono uccelli”	Semantic equivalence
Statement equivalence	$s_1 \equiv s_2$	“Swans are birds”	“Swans have feathers”	Logical equivalence
Truth equality	$\text{truth}(s_1) = \text{truth}(s_2)$	“Swans are birds”	“The earth is round”	Material equivalence

Table 1.2: Different types of statement relationships.

Note that the relationships in the table are ordered by strength: if two statements are equal, they are also equivalent; if two statements are equivalent, they have equal truth. The reverse is not true in general: two statements with equal truth may not be equivalent; two equivalent statements may not be equal.

Intuitively, statement equivalence answers the question: do these two statements carry the same information? Is experimentally testing the first the same as experimentally testing the second? If that’s the case, they are essentially equivalent to us. So much so, that from now on we will implicitly assume two different statements to be inequivalent.⁸

There are a number of useful properties that statement equivalence satisfies.

Corollary 1.16. *All certainties are equivalent. All impossibilities are equivalent.*

Proof. Let $\top_1, \top_2 \in \mathcal{S}$ be two certainties. Then for every possible assignment $a \in \mathcal{A}_{\mathcal{S}}$ we have $a(\top_1) = \text{TRUE} = a(\top_2)$ and therefore $\top_1 \equiv \top_2$ by definition.

Let $\perp_1, \perp_2 \in \mathcal{S}$ be two impossibilities. Then for every possible assignment $a \in \mathcal{A}_{\mathcal{S}}$ we have $a(\perp_1) = \text{FALSE} = a(\perp_2)$ and therefore $\perp_1 \equiv \perp_2$ by definition. \square

Corollary 1.17. *Two statements s_1 and s_2 are equivalent if and only if $(s_1 \wedge s_2) \vee (\neg s_1 \wedge \neg s_2) \equiv \top$.*

Proof. Let $a \in \mathcal{A}_{\mathcal{S}}$ and $s = (s_1 \wedge s_2) \vee (\neg s_1 \wedge \neg s_2)$. We have $a(s) = (a(s_1) \wedge a(s_2)) \vee (\neg a(s_1) \wedge \neg a(s_2))$. We have the following truth table

⁸Technically, when we’ll say that s is a statement we actually mean s is an equivalence class of statements. We are not going to be explicit about the distinction, though, as we feel it simply distracts without adding greater clarity. We’ll let the context determine what is meant.

s_1	s_2	s
T	T	T
T	F	F
F	T	F
F	F	T

Note that the assignments for which $a(s_1) = a(s_2)$ are exactly the assignments for which $a(s) = \text{TRUE}$. Therefore $s_1 \equiv s_2$ if and only if s is a certainty. \square

Corollary 1.18. *Statement equivalence satisfies the following properties:*

- *reflexivity:* $s \equiv s$
- *symmetry:* if $s_1 \equiv s_2$ then $s_2 \equiv s_1$
- *transitivity:* if $s_1 \equiv s_2$ and $s_2 \equiv s_3$ then $s_1 \equiv s_3$

and is therefore an **equivalence relationship**.

Proof. Statement equivalence is defined in terms of truth equality within all possible assignments and will inherit reflexivity, symmetry and transitivity from it. \square

Corollary 1.19. *A logical context \mathcal{S} satisfies the following properties:*

- *associativity:* $a \vee (b \vee c) \equiv (a \vee b) \vee c$, $a \wedge (b \wedge c) \equiv (a \wedge b) \wedge c$
- *commutativity:* $a \vee b \equiv b \vee a$, $a \wedge b \equiv b \wedge a$
- *absorption:* $a \vee (a \wedge b) \equiv a$, $a \wedge (a \vee b) \equiv a$
- *identity:* $a \vee \perp \equiv a$, $a \wedge \top \equiv a$
- *distributivity:* $a \vee (b \wedge c) \equiv (a \vee b) \wedge (a \vee c)$, $a \wedge (b \vee c) \equiv (a \wedge b) \vee (a \wedge c)$
- *complements:* $a \vee \neg a \equiv \top$, $a \wedge \neg a \equiv \perp$
- *De Morgan:* $\neg a \vee \neg b \equiv \neg(a \wedge b)$, $\neg a \wedge \neg b \equiv \neg(a \vee b)$

Therefore \mathcal{S} is a **Boolean algebra** by definition.

Proof. The left and right expressions for each equivalence correspond to the same truth function applied to the same statements. Therefore, the left side is true if and only if the right side is true and they are therefore equivalent by definition. \square

These operations and properties define the **algebra of statements**. While we started from a slightly different premise, the relationships we found are the logical identities of classical logic. These are exactly what we need to make sure the truth values of all our statements are consistent.

Equivalence is not the only semantic relationship that we want to capture. Consider the contents of the following:

- $s_1 = \text{“that animal is a cat”}$
- $s_2 = \text{“that animal is a mammal”}$
- $s_3 = \text{“that animal is a dog”}$
- $s_4 = \text{“that animal is black”}$

The second will be true whenever the first is true. In this case we say the first statement is narrower than the second ($s_1 \preceq s_2$) because “that animal is a cat” is more specific than “that

animal is a mammal”. The third will be false whenever the first is true. In this case we say that they are incompatible ($s_1 \not\approx s_3$) because “*that animal is a dog*” and “*that animal is a cat*” can never be true at the same time. The fourth will be true or false regardless of whether the first is true. In this case we say that they are independent ($s_1 \perp s_4$) because knowing whether “*that animal is a cat*” tells us nothing about whether “*that animal is black*”. As for equivalence, we can define these relationships upon the previous definitions.

Definition 1.20. *Given two statements s_1 and s_2 , we say that:*

- s_1 **is narrower than** s_2 (noted $s_1 \leq s_2$) if s_2 is true whenever s_1 is true simply because of their content. That is, for all $a \in \mathcal{A}_S$ if $a(s_1) = \text{TRUE}$ then $a(s_2) = \text{TRUE}$.
- s_1 **is broader than** s_2 (noted $s_1 \geq s_2$) if $s_2 \leq s_1$.
- s_1 **is compatible to** s_2 (noted $s_1 \approx s_2$) if their content allows them to be true at the same time. That is, there exists $a \in \mathcal{A}_S$ such that $a(s_1) = a(s_2) = \text{TRUE}$.

The negation of these properties will be noted by $\not\leq$, $\not\geq$, $\not\approx$ respectively.

Definition 1.21. *The elements of a set of statements $S \subseteq \mathcal{S}$ are said to be **independent** (noted $s_1 \perp s_2$ for a set of two) if the assignment of any subset of statements does not depend on the assignment of the others. That is, a set of statements $S \subseteq \mathcal{S}$ is independent if given a family $\{t_s\}_{s \in S}$ such that each $t_s \in \mathbb{B}$ is a possible assignment for the respective s we can always find $a \in \mathcal{A}_S$ such that $a(s) = t_s$ for all $s \in S$.*

Proposition 1.22. *The above operations obey the following relationships:*

- (i) $s_1 \leq s_2$ if and only if $s_1 \wedge \neg s_2 \equiv \perp$
- (ii) $s_1 \leq s_2$ if and only if $s_1 \wedge s_2 \equiv s_1$
- (iii) $s_1 \leq s_2$ if and only if $s_1 \vee s_2 \equiv s_2$
- (iv) $s_1 \approx s_2$ if and only if $s_1 \wedge s_2 \neq \perp$
- (v) $s_1 \not\approx s_2$ if and only if $s_1 \wedge \neg s_2 \equiv s_1$
- (vi) $s_1 \leq s_1 \vee s_2$
- (vii) $s_1 \wedge s_2 \leq s_1$
- (viii) $s_1 \leq s_2$ if and only if $\neg s_1 \vee s_2 \equiv \top$

Proof. For (i), consider the following truth table

s_1	s_2	$s_1 \wedge \neg s_2$
T	T	F
T	F	T
F	T	F
F	F	F

The assignment for which $a(s_1) = \text{TRUE}$ and $a(s_2) = \text{FALSE}$ is exactly the assignment for which $a(s_1 \wedge \neg s_2) = \text{TRUE}$. Therefore $s_1 \leq s_2$ if and only if $s_1 \wedge \neg s_2$ is impossible.

For (ii), consider $s_1 \wedge s_2 \equiv s_1 \wedge s_2 \vee \perp$. Since $s_1 \leq s_2$, $s_1 \wedge \neg s_2 \equiv \perp$. We have $s_1 \wedge s_2 \vee \perp \equiv (s_1 \wedge s_2) \vee (s_1 \wedge \neg s_2) \equiv s_1 \wedge (s_2 \vee \neg s_2) \equiv s_1 \wedge \top \equiv s_1$. Therefore $s_1 \wedge s_2 \equiv s_1$. The same logic can be applied in reverse.

For (iii), consider $s_1 \vee s_2 \equiv s_1 \vee s_2 \wedge \top \equiv (s_1 \vee s_2) \wedge (\neg s_2 \vee s_2) \equiv (s_1 \wedge \neg s_2) \vee s_2$. Since $s_1 \leq s_2$, $s_1 \wedge \neg s_2 \equiv \perp$. We have $(s_1 \wedge \neg s_2) \vee s_2 \equiv \perp \vee s_2 \equiv s_2$. Therefore $s_1 \wedge s_2 \equiv s_2$. The same logic can be applied in reverse.

For (iv), consider the following truth table

s_1	s_2	$s_1 \wedge s_2$
T	T	T
T	F	F
F	T	F
F	F	F

The assignment for which $a(s_1) = a(s_2) = \text{TRUE}$ is exactly the assignment for which $a(s_1 \wedge s_2) = \text{TRUE}$. Therefore $s_1 \approx s_2$ if and only if $s_1 \wedge s_2$ is not impossible.

For (v), consider $s_1 \wedge \neg s_2 \equiv s_1 \wedge \neg s_2 \vee \perp$. Since $s_1 \not\approx s_2$, $s_1 \wedge s_2 \equiv \perp$. We have $s_1 \wedge \neg s_2 \vee \perp \equiv (s_1 \wedge \neg s_2) \vee (s_1 \wedge s_2) \equiv s_1 \wedge (\neg s_2 \vee s_2) \equiv s_1 \wedge \top \equiv s_1$. Therefore $s_1 \wedge \neg s_2 \equiv s_1$. The same logic can be applied in reverse.

For (vi), we have $s_1 \wedge \neg(s_1 \vee s_2) \equiv s_1 \wedge \neg s_1 \wedge \neg s_2 \equiv \perp \wedge \neg s_2 \equiv \perp$. Therefore $s_1 \leq s_1 \vee s_2$.

For (vii), we have $s_1 \wedge s_2 \wedge \neg s_1 \equiv s_1 \wedge \neg s_1 \wedge s_2 \equiv \perp \wedge s_2 \equiv \perp$. Therefore $s_1 \wedge s_2 \leq s_1$.

For (viii), suppose $s_1 \leq s_2$. We have $\neg s_1 \vee s_2 \equiv \neg s_1 \vee (s_1 \vee s_2) \equiv (\neg s_1 \vee s_1) \vee s_2 \equiv \top \vee s_2 \equiv \top$. Conversely, suppose $\neg s_1 \vee s_2 \equiv \top$. We have $s_1 \vee s_2 \equiv s_1 \vee s_2 \vee \top \equiv s_1 \vee s_2 \vee (\neg s_1 \vee s_2) \equiv (s_1 \vee \neg s_1) \vee s_2 \equiv \top \vee s_2 \equiv s_2$ and therefore $s_1 \leq s_2$. \square

Proposition 1.23. *Statement narrowness satisfies the following properties:*

- *reflexivity:* $s \leq s$
- *antisymmetry:* if $s_1 \leq s_2$ and $s_2 \leq s_1$ then $s_1 \equiv s_2$
- *transitivity:* if $s_1 \leq s_2$ and $s_2 \leq s_3$ then $s_1 \leq s_3$

and is therefore a **partial order**.

Proof. For reflexivity, $s \wedge \neg s \equiv \perp$ and therefore $s \leq s$ by 1.22.

For antisymmetry, $s_1 \wedge s_2 \equiv s_1$ since $s_1 \leq s_2$ and $s_1 \wedge s_2 \equiv s_2$ since $s_2 \leq s_1$. Therefore $s_1 \equiv s_2$. Conversely, suppose $s_1 \equiv s_2$. Then $s_1 \wedge s_2 \equiv s_1 \equiv s_2$. Therefore $s_1 \leq s_2$ and $s_1 \geq s_2$.

For transitivity, we have $s_1 \equiv s_1 \wedge s_2 \equiv s_1 \wedge s_2 \wedge s_3 \equiv s_1 \wedge s_3$ and therefore $s_1 \leq s_3$. \square

Proposition 1.24. *Every subset $S \subseteq \mathcal{S}$ of statements has a supremum. That is, there exists an element $\bar{s} \in \mathcal{S}$ such that $s \leq \bar{s}$ for all $s \in S$. This, by definition, means the \mathcal{S} is a **complete** Boolean algebra and, as a consequence, that the distributivity and De Morgan laws in 1.19 hold in the infinite case.*

Proof. Let $S \subseteq \mathcal{S}$ be an arbitrary set of statements. Consider $\bar{s} = \bigvee_{e \in S} e$. This statement exists by 1.9. Let $s \in S$. Using the properties in 1.19 we have $s \vee \bar{s} \equiv s \vee \left(\bigvee_{e \in S} e \right) \equiv s \vee s \vee \left(\bigvee_{e \in (S \setminus \{s\})} e \right) \equiv s \vee \left(\bigvee_{e \in (S \setminus \{s\})} e \right) \equiv \bigvee_{e \in S} e = \bar{s}$. By 1.22 we have $s \leq \bar{s}$ for any $s \in S$. Therefore any $S \subseteq \mathcal{S}$ admits a supremum \bar{s} .

A complete Boolean algebra is one such that every subset admits a supremum, therefore the algebra of statements is complete by definition. For a complete Boolean algebra infinite

distributivity and De Morgan laws hold, therefore they will hold in the algebra of statements as well. \square

It should be noted that statement narrowness captures more than just the idea of one statement being more specific than another. Consider “*this harp seal is white*” and “*this harp seal is less than one year old*”. Since harp seals have white fur only for their first month, the first one can never be true while the second is not. Therefore “*this harp seal is white*” \leq “*this harp seal is less than one year old*”. By the same token, we also have “*I lighted the fuse*” \leq “*the bomb will go off*”. That is, narrowness can also capture causal relationships, which is essential if we want to develop a basic theory of scientific investigation.⁹ Intuitively, a statement is narrower than another if it provides at least as much or more information when true. If we experimentally verified the narrower statement, then we already know that the broader one is also verified.

It should also be noted that independence is not transitive and pair-wise independence is not sufficient. Consider the following statements for an ideal gas:

1. “*the pressure is 101 ± 1 kPa*”
2. “*the volume is 1 ± 0.1 m³*”
3. “*the temperature is 293 ± 1 Kelvin*”

Since the three quantities are linked by the equation of state $PV = nRT$, any two statements are independent but the three together aren’t. This notion of independence is similar, and in fact related, to statistical independence and linear independence.

We finish this section by showing how every logical dependence among statements can be naturally expressed in terms of negation, conjunction and disjunction. Consider the statement $s =$ “*the sauce is sweet and sour or it is neither*”. This depends on the two statements $s_1 =$ “*the sauce is sweet*” and $s_2 =$ “*the sauce is sour*” defined before: if we know whether s_1 and s_2 are true, we can tell whether s is true as well. The idea is that we can express the dependence as all possible assignments for $\{s_1, s_2\}$ that make the result true. For example, s will be true if the sauce is sweet and sour or if it is not sweet and not sour. That is: $(s_1 \wedge s_2) \vee (\neg s_1 \wedge \neg s_2)$. Similarly, the statement “*the sauce is not sweet-and-sour*” can be expressed as $(s_1 \wedge \neg s_2) \vee (\neg s_1 \wedge s_2) \vee (\neg s_1 \wedge \neg s_2)$ since it is going to be true in all cases except the one where the sauce is sweet and sour.

Each of the cases is the conjunction of all independent statements where each one appears only once, negated or not. We call these expressions minterms. A function can always be expressed as the disjunction of all the minterms for which the function is true. This is called its disjunctive normal form because it is a canonical way to express the function in terms of disjunctions.

Definition 1.25. Let $S \subseteq \mathcal{S}$ be a set of statements. A *minterm* of S is a conjunction where each element appears once and only once, either negated or not. That is, it is a statement $m \in \mathcal{S}$ that can be written as $m \equiv \bigwedge_{s \in S} (\neg)^{a(s)} s$ where $a : S \rightarrow \mathbb{B}$, $\neg^{\text{TRUE}} s = s$ and

⁹We considered using the term implication directly, but it seems that it leads to confusion. Implication in classical logic is something different: it is simply another truth function. Moreover, saying that an impossibility is narrower than all other statements sounds better than saying that an impossibility implies all other statements.

$\neg^{\text{FALSE}} \mathbf{s} = \neg \mathbf{s}$. In this notation, in a given $a_0 \in \mathbb{B}^S$ we have $a_0(\mathbf{m}) = \text{TRUE}$ if and only if $a_0(\mathbf{s}) = a(\mathbf{s})$ for all $\mathbf{s} \in S$.

Proposition 1.26. *Let $\bar{\mathbf{s}} \in \mathcal{S}$ be a statement that depends on a set of statements $S \subseteq \mathcal{S}$ through $f_{\mathbb{B}} : \mathbb{B}^S \rightarrow \mathbb{B}$. Then we can express $\bar{\mathbf{s}}$ in its **disjunctive normal form** as a disjunction of minterms of S , that is $\bar{\mathbf{s}} \equiv \bigvee_{a \in A} \left(\bigwedge_{\mathbf{s} \in S} (\neg)^{a(\mathbf{s})} \mathbf{s} \right)$ where $A \subseteq \mathbb{B}^S$ is a set of assignments for S . In this notation, in a given $a_0 \in \mathbb{B}^S$ we have $a_0(\bar{\mathbf{s}}) = \text{TRUE}$ if and only if there is an $a \in A$ such that $a_0(\mathbf{s}) = a(\mathbf{s})$ for all $\mathbf{s} \in S$.*

Proof. We first show that this can be done for a function $f_{\mathbb{B}}$ that returns TRUE for only one assignment. Let $a : S \rightarrow \mathbb{B}$ be an assignment for S and suppose $f_{\mathbb{B}} : \mathbb{B}^S \rightarrow \mathbb{B}$ is such that $f_{\mathbb{B}}(\bar{a}) = \text{TRUE}$ if and only if $\bar{a} = a$. Now consider the minterm $\mathbf{m}_a = \bigwedge_{\mathbf{s} \in S} (\neg)^{a(\mathbf{s})} \mathbf{s}$ and an assignment $\bar{a} : \mathcal{S} \rightarrow \mathbb{B}$ for the whole context. We have that $\bar{a}(\mathbf{m}_a) = \text{TRUE}$ if and only if $\bar{a}(\mathbf{s}) = a(\mathbf{s})$ for all $\mathbf{s} \in S$. Then $\bar{a}(\bar{\mathbf{s}}) = \bar{a}(\mathbf{m}_a)$ for all possible assignments $\bar{a} \in \mathcal{A}_{\mathcal{S}}$ and therefore $\bar{\mathbf{s}} \equiv \bigwedge_{\mathbf{s} \in S} (\neg)^{a(\mathbf{s})} \mathbf{s}$.

Now we generalize the result for arbitrary functions. Let $f_{\mathbb{B}} : \mathbb{B}^S \rightarrow \mathbb{B}$ be a generic function. Let $A = \{a \in \mathbb{B}^S \mid f_{\mathbb{B}}(a) = \text{TRUE}\}$. Let $\mathbf{m}_a = \bigwedge_{\mathbf{s} \in S} (\neg)^{a(\mathbf{s})} \mathbf{s}$ be the minterm associated with an assignment $a \in A$. Consider $\bar{\mathbf{m}} = \bigvee_{a \in A} \mathbf{m}_a$ and an assignment $\bar{a} : \mathcal{S} \rightarrow \mathbb{B}$ for the whole context. We have that $\bar{a}(\bar{\mathbf{m}}) = \text{TRUE}$ if and only if $\bar{a}(\mathbf{s}) = a(\mathbf{s})$ for some $a \in A$ and for all $\mathbf{s} \in S$. Then $\bar{a}(\bar{\mathbf{s}}) = \bar{a}(\bar{\mathbf{m}})$ for all possible assignments $\bar{a} \in \mathcal{A}_{\mathcal{S}}$ and therefore $\bar{\mathbf{s}} \equiv \bigvee_{a \in A} \left(\bigwedge_{\mathbf{s} \in S} (\neg)^{a(\mathbf{s})} \mathbf{s} \right)$ \square

With these tools in place we are in a position to formulate models that are universal and non-contradictory. These models will be a collection of statements with a well defined content and well defined possible cases. Each statement's truth value will be discovered experimentally.

1.2 Verifiable statements and experimental domains

We now focus on those statements that are verifiable: we have a way to experimentally confirm that the statement is true. The main result of this section is that not all functions of verifiable statements are verifiable statements. For example, since a test has to finish in a finite amount of time we are not going to be able to verify a statement that is the conjunction (i.e. logical AND) of infinitely many statements. We are also going to group verifiable statements into experimental domains which represent all the experimental evidence about a scientific subject that can be acquired in an indefinite amount of time.

The previous section took care of universality and non-contradiction, but the principle of scientific objectivity requires science to be evidence based. Consider the statements “23 is a prime number”, “it is immoral to kill a person exclusively for monetary gain” or “God is eternal”. They deal with abstract concepts that cannot be defined operationally and therefore cannot be experimentally verified conclusively. Again, this does not mean these concepts are of less significance, just that they cannot be the subject of scientific inquiry.¹⁰

¹⁰In fact, one may be more interested in them precisely because of their abstract, and therefore less transient, nature.

Limiting the scope of our discussion to objects and properties that are well defined physically is also not enough. For example, “*the electron is green*” or “*1 meter is equal to 5 Kelvin*” are still not suitable scientific statements as the relationships established are not physically meaningful. Even when the relationship is meaningful, we may still not be able to validate it experimentally. For example, “*there is no extra-terrestrial life*” or “*the mass of the electron is exactly 9.109×10^{-31} kg*” are not statements that can be verified in practice. In the first case, we would need to check every corner of the universe and find none, with the closest galaxy like ours, Andromeda, being 2.5 million light-years away; in the second case, we will always have an uncertainty associated with the measurement, however small.

So we have to narrow the scope to those and only those statements that can be verified experimentally. That is, first we have to provide an experimental test: a repeatable experimental procedure (i.e. evidence based) that anyone (i.e. universal) can in principle execute and obtain consistent results (i.e. non-contradictory). Second, we must guarantee that the test always terminates successfully if and only if the statement is true. This is both the power and the limit of scientific inquiry: it gives us a way to construct a coherent description of the physical world but it is limited to those aspects that can be reliably studied experimentally.

Note that certainty and impossibility are trivially verifiable since we know a priori that they are true and false respectively. Also note that if we have two statements that are equivalent, having a test that verifies one means we have a test that verifies the other as well. For example, since “*that animal is a bird*” is equivalent to “*that animal has feathers*”, checking whether the animal has feathers is equivalent to checking whether the animal is a bird. The subtlety here is that the evidence may be indirect and it is only the relationship between statements (i.e. the theoretical model) that guarantees the validity of the verification. This should not worry us, though, because in this strict sense most experimental data is indirect. It comes from a chain of inductions (e.g. the pulse of light is produced, bounces off a moving target and changes frequency due to the Doppler effect, the light signal is transduced to an electronic signal which finally is displayed on the device) and therefore needs a theoretical framework to be properly understood.

Axiom 1.27 (Axiom of verifiability). *A **verifiable statement** is a statement that, if true, can be shown to be so experimentally. Formally, each logical context \mathcal{S} contains a set of statements $\mathcal{S}_v \subseteq \mathcal{S}$ whose elements are said to be verifiable. Moreover, we have the following properties:*

- every certainty $\top \in \mathcal{S}$ is verifiable
- every impossibility $\perp \in \mathcal{S}$ is verifiable
- a statement equivalent to a verifiable statement is verifiable

Justification. To give a better justification for this and later axioms, we introduce the following pseudo-mathematical concepts. As science is evidence based, for each logical context \mathcal{S} we must have at our disposal a set of **experimental tests**, which we indicate with \mathcal{E} . Each element $e \in \mathcal{E}$ is a repeatable procedure (i.e. it can be restarted and stopped at any time) that anybody can execute and will always terminate successfully, terminate unsuccessfully or never terminate. As the tests must provide evidence, the output of the test must depend on the truth values of the statements; therefore we can assume we have

a function $\text{result} : \mathcal{E} \times \mathcal{A}_{\mathcal{S}} \rightarrow \{\text{SUCCESS}, \text{FAILURE}, \text{UNDEFINED}\}$. For a statement s to be verifiable, there must be an experimental test e that succeeds if and only if the statement is true. That is, for all $a \in \mathcal{A}_{\mathcal{S}}$, $\text{result}(e, a) = \text{SUCCESS}$ if and only if $a(s) = \text{TRUE}$.

Certainties and impossibilities can be associated with trivial tests that always terminate successfully or unsuccessfully respectively. This justifies assuming them to be verifiable. As for equivalence and verifiability, let $s_1, s_2 \in \mathcal{S}$ and suppose s_1 is verifiable. This means there is a test $e \in \mathcal{E}$ such that $\text{result}(e, a) = \text{SUCCESS}$ if and only if $a(s_1) = \text{TRUE}$. Since the statements are equivalent, $a(s_1) = \text{TRUE}$ if and only if $a(s_2) = \text{TRUE}$. Therefore $\text{result}(e, a) = \text{SUCCESS}$ if and only if $a(s_2) = \text{TRUE}$. We are therefore justified to assume s_2 is verifiable.

Precisely defining experimental tests as procedures will present the same type of problems in defining statements or logical consistency, therefore we leave them as primary concepts. But as primary concepts they only complicate the formal framework without adding insights, therefore we use them only as part of the justifications for the axioms. \square

Experimental tests are the second and last building block of our general mathematical theory of experimental science. As with statements, any language (e.g. natural, formal, engineering drawings, computer programs, ...) can in principle be used to describe the procedure, which can be arbitrarily complicated. It may require building detectors, gathering large amounts of data and performing complicated computations. We are not going to care how these procedures are described, just that it is done in a way that allows us to execute the test.¹¹

As an example, consider a procedure along the lines of:

1. find a swan
2. if it's black terminate successfully
3. go to step 1

If a black swan exists, at some point we'll find it and the test will be successful. If a black swan does not exist, then the procedure will never terminate and the result is undefined. This is something anybody can do and will eventually always provide the same result: it is an experimental test. It also terminates successfully if and only if a black swan exists, so the statement "*black swans exist*" is verifiable.

Note that, in principle, science can also study statements that can be refuted experimentally. But the negation of those is a statement that can be verified experimentally. Therefore we lose nothing by only focusing on verification.¹²

In the previous section we saw that we can combine statements into new statements. How about verifiable statements? Can we always combine verifiable statements into other verifiable statements? Since all truth functions can be constructed from the three basic Boolean operations, the question becomes: can we construct experimental tests for the negation, conjunction and disjunction of verifiable statements?

The first important result is that the negation of an experimental test, an experimental test that is successful when the first is not successful, does not necessarily exist. Consider our black swan example, an experimental test for the negation would be a procedure that terminates

¹¹Trying to formalize a universal language for experimental tests is not only impractical but also conceptually problematic. To know what we can test experimentally is to know what is physically possible, which is equivalent to knowing the laws of physics, which is what we are trying to construct a framework for.

¹²Mathematically, the spaces of verifiable statements and refutable statements are dual to each other.

successfully if black swans do not exist. But the given procedure never finishes in that case, so it is not just a matter of switching success with failure. Because of non-termination, not-successful does not necessarily mean failure.¹³ Moreover, it is a result of computability theory that some problems are undecidable: they do not allow the construction of an algorithm that always terminates with a correct yes-or-no answer. So we know that in some cases this is not actually possible.

In the same vein we are able to confirm experimentally that “*the mass of this particle is not zero*” but not that “*the mass of this particle is exactly zero*” since we always have uncertainty in our measurements of mass. Even if we could continue shrinking the uncertainty arbitrarily, we would ideally need infinite time to shrink it to zero. What this means is that not all answers to the same question can be equally verified. Is the mass of the photon exactly zero? We can either give a precise “no” or an imprecise “it’s within this range.” Is there extra-terrestrial life? We can either give a precise “yes” or an imprecise “we haven’t found it so far.”¹⁴

Remark. The **negation or logical NOT** of a verifiable statement is not necessarily a verifiable statement.

Justification. A statement $s \in \mathcal{S}$ is verifiable if we can find $e \in \mathcal{E}$ such that $\text{result}(e, a) = \text{SUCCESS}$ if and only if $a(s) = \text{TRUE}$ for all $a \in \mathcal{A}_{\mathcal{S}}$. This means that for some $a \in \mathcal{A}_{\mathcal{S}}$ we may have $a(s) = \text{FALSE}$ and $\text{result}(e, a) = \text{UNDEFINED}$. That is, the test is not guaranteed to terminate if the statement is false. Therefore we cannot, in general, use e to construct a statement that terminates whenever s is false. Therefore we are not justified, in general, to assume that the negation of a verifiable statement is verifiable. \square

While this is true in general, we can still test the negation of many verifiable statements. Consider the statement “*this swan is black*”. It allows the following experimental test:

1. look at the swan
2. if it’s black terminate successfully
3. terminate unsuccessfully

Note that, since the test always terminates, we can switch failure to success and vice-versa. In this case we can test the negation and we say that the statement is decidable: we can decide experimentally whether it is true or false. It is precisely when and only when the test is guaranteed to terminate, that we can test the negation.

Definition 1.28. A **falsifiable statement** is a statement that, if false, can be shown to be so experimentally. Formally, a statement s is falsifiable if its negation $\neg s \in \mathcal{S}_v$ is a verifiable statement.

Justification. Note that the informal definition is based on experimentally showing that the statement is false, while the formal definition is based on the falsifiable statement to be the negation of a verifiable one. We have to show they are equivalent.

For a statement $s \in \mathcal{S}$ to be falsifiable, there must be an experimental test $e \in \mathcal{E}$ that

¹³In this case, the old adage “absence of evidence is not evidence of absence” applies.

¹⁴Note that we are on purpose avoiding induction. It does not play any role in our general mathematical theory of experimental science since the decision of when and how to apply induction violates the principle of scientific objectivity.

fails if and only if the statement is false. That is, for all $a \in \mathcal{A}_S$, $\text{result}(e, a) = \text{FAILURE}$ if and only if $a(s) = \text{FALSE}$.

Let $e \in \mathcal{E}$ be an experimental test and consider $e_-(e)$ defined as follows:

1. run test e
2. if e is unsuccessful terminate successfully
3. if e is successful terminate unsuccessfully.

Since e is repeatable and can be executed by anybody, $e_-(e)$ is also repeatable and can be executed by anybody. Therefore we are justified to assume $e_-(e) \in \mathcal{E}$.

Now let s be a verifiable statement, then we can find $e \in \mathcal{E}$ such that $\text{result}(e, a) = \text{SUCCESS}$ if and only if $a(s) = \text{TRUE}$. We also have $\text{result}(e_-(e), a) = \text{FAILURE}$ if and only if $\text{result}(e, a) = \text{SUCCESS}$ and $a(\neg s) = \text{FALSE}$ if and only if $a(s) = \text{TRUE}$. Therefore for all $a \in \mathcal{A}_S$, $\text{result}(e_-(e), a) = \text{FAILURE}$ if and only if $a(\neg s) = \text{FALSE}$. Which means $\neg s$ is falsifiable if and only if s is verifiable. This justifies the definition. \square

Definition 1.29. A *decidable statement* is a statement that can be shown to be either true or false experimentally. Formally, a statement s is decidable if $s \in \mathcal{S}_V$ and $\neg s \in \mathcal{S}_V$. We denote $\mathcal{S}_d \subseteq \mathcal{S}_V$ the set of all decidable statements.

Justification. Note that the informal definition is based on experimentally showing that the statement is either true or false, while the formal definition is based on both the decidable statement and its negation to be verifiable. We have to show they are equivalent.

Let $s \in \mathcal{S}$ be a decidable statement and $e \in \mathcal{E}$ an experimental test that verifies whether the statement is true or false. That is, $\text{result}(e, a) = \text{SUCCESS}$ if and only if $a(s) = \text{TRUE}$ and $\text{result}(e, a) = \text{FAILURE}$ if and only if $a(s) = \text{FALSE}$ for all $a \in \mathcal{A}_S$. Then s is verifiable. We also have $\text{result}(e_-(e), a) = \text{SUCCESS}$ if and only if $a(\neg s) = \text{TRUE}$. Therefore s is also verifiable.

Conversely, let $s \in \mathcal{S}$ be a verifiable statement such that $\neg s$ is also verifiable. Let $e, e_- \in \mathcal{E}$ be their respective experimental tests. We have to be careful as e and e_- may not terminate. Consider the procedure $\hat{e}(e, e_-)$ defined as follows:

1. initialize n to 1
2. run the test e for n seconds
3. if e is successful, terminate successfully
4. run the test e_- for n seconds
5. if e_- is successful, terminate unsuccessfully
6. increment n and go to step 2

The procedure is repeatable and can be executed by anybody therefore $\hat{e}(e, e_-) \in \mathcal{E}$. Both e and e_- are eventually run an arbitrarily long amount of time therefore $\text{result}(\hat{e}(e, e_-), a) \in \{\text{SUCCESS}, \text{FAILURE}\}$, that is the test will always terminate. We have $\text{result}(\hat{e}(e, e_-), a) = \text{SUCCESS}$ if and only if $\text{result}(e, a) = \text{SUCCESS}$ if and only if $a(s) = \text{TRUE}$. We also have $\text{result}(\hat{e}(e, e_-), a) = \text{FAILURE}$ if and only if $\text{result}(e_-, a) = \text{SUCCESS}$ if and only if $a(\neg s) = \text{TRUE}$ if and only if $a(s) = \text{FALSE}$. Therefore s is decidable. This justifies the definition. \square

Corollary 1.30. *Certainties and impossibilities are decidable statements.*

Proof. Let $\top \in \mathcal{S}$ be a certainty and $\perp \in \mathcal{S}$ be an impossibility. By 1.27 they are verifiable statements. We also have $\top \equiv \neg\perp$ therefore they are decidable by definition. \square

We introduce decidable statements here because their definition and related properties clarify what happens during negation, but they do not play a major role in our framework. They represent a special case which will we turn to time and time again over the course of this work.

Combining verifiable statements with conjunction (i.e. the logical AND) is more straightforward. If we are able to verify that “that animal is a swan” and that “that animal is black”, we can verify that “that animal is a black swan” by verifying both. If the tests for both are successful, then the test for the conjunction is successful. That is, if we have two or more verifiable statements, we can always construct an experimental test for the logical AND by running all tests one at a time and check if they are successful. Yet, the number of tests needs to be finite or we would never terminate, so we are limited to the conjunction of a finite number of verifiable statements.

Axiom 1.31 (Axiom of finite conjunction verifiability). *The conjunction of a finite collection of verifiable statements is a verifiable statement. Formally, let $\{s_i\}_{i=1}^n \subseteq \mathcal{S}_v$ be a finite collection of verifiable statements. Then the conjunction $\bigwedge_{i=1}^n s_i \in \mathcal{S}_v$ is a verifiable statement.*

Justification. Let $\{s_i\}_{i=1}^n \subseteq \mathcal{S}_v$ be a finite collection of verifiable statements. Then we can find a corresponding set of experimental tests $\{e_i\}_{i=1}^n \subseteq \mathcal{E}$ such that $\text{result}(e_i, a) = \text{SUCCESS}$ if and only if $a(s_i) = \text{TRUE}$ for all $a \in \mathcal{A}_S$.

Let $\bigwedge_{i=1}^n e_i$ be the experimental procedure defined as follows:

1. for each $i = 1..n$ run the test e_i
2. if all tests terminate successfully then terminate successfully
3. terminate unsuccessfully.

The experimental procedure so defined is repeatable, can be executed by anybody, therefore $\bigwedge_{i=1}^n e_i \in \mathcal{E}$. We have, for every $a \in \mathcal{A}_S$, $\text{result}(\bigwedge_{i=1}^n e_i, a) = \text{SUCCESS}$ if and only if $\text{result}(e_i, a) = \text{SUCCESS}$ for all i if and only if $a(s_i) = \text{TRUE}$ for all i if and only if $a(\bigwedge_{i=1}^n s_i) = \text{TRUE}$.

Therefore $\bigwedge_{i=1}^n s_i$ is a verifiable statement. We are therefore justified to assume that the finite conjunction of verifiable statements is a verifiable statement.

Note that this cannot be generalized to infinite collections as the procedure would not be guaranteed to terminate. In fact, if the only way to test the infinite conjunction is to test each statement individually, then we are guaranteed to take infinite time and the test will never terminate. Therefore we are not justified to assume the infinite conjunction of verifiable statements is verifiable. \square

Combining verifiable statements with disjunction (i.e. the logical OR) is also straightforward. To verify that “the swan is black or white” we can first test that “the swan is black”. If that is verified that’s enough: the swan is black or white. If not, we test that “the swan is white”. That is, if we have two or more verifiable statements we can always construct an experimental test for the logical OR by running all tests and stopping at the first one that is

successful. Because we stop at the first success, the number of tests can be countably infinite. As long as one test succeeds, which will always be the case when the overall test succeeds, it does not matter how many elements we are not going to verify later. But it cannot be more than countably infinite since the only way we have to find if one experimental test in the set is successful is testing them all one by one. Therefore we are limited to the disjunction of a countable number of verifiable statements.

Axiom 1.32 (Axiom of countable disjunction verifiability). *The disjunction of a countable collection of verifiable statements is a verifiable statement. Formally, let $\{s_i\}_{i=1}^{\infty} \subseteq \mathcal{S}_v$ be a countable collection of verifiable statements. Then the disjunction $\bigvee_{i=1}^{\infty} s_i \in \mathcal{S}_v$ is a verifiable statement.*

Justification. Let $\{s_i\}_{i=1}^{\infty} \subseteq \mathcal{S}_v$ be a countable collection of verifiable statements. Then we can find a corresponding set of experimental tests $\{e_i\}_{i=1}^{\infty} \subseteq \mathcal{E}$ such that $\text{result}(e_i, a) = \text{SUCCESS}$ if and only if $a(s_i) = \text{TRUE}$ for all $a \in \mathcal{A}_S$.

In this case, we have to be careful to handle tests that may not terminate. Let $\bigvee_{i=1}^{\infty} e_i$ be the experimental procedure defined as follows:

1. initialize n to 1
2. for each $i = 1..n$
 - a) run the test e_i for n seconds
 - b) if e_i terminates successfully then terminate successfully
3. increment n and go to step 2

The experimental procedure so defined is repeatable, can be executed by anybody, therefore $\bigvee_{i=1}^{\infty} e_i \in \mathcal{E}$. The procedure will eventually run all tests for an arbitrarily long amount of time.

Therefore, for every $a \in \mathcal{A}_S$, $\text{result}(\bigvee_{i=1}^{\infty} e_i, a) = \text{SUCCESS}$ if and only if $\text{result}(e_i, a) = \text{SUCCESS}$ for some i if and only if $a(s_i) = \text{TRUE}$ for some i if and only if $a(\bigvee_{i=1}^{\infty} s_i) = \text{TRUE}$. Therefore

$\bigvee_{i=1}^{\infty} s_i$ is a verifiable statement. We are therefore justified to assume that the countable disjunction of verifiable statements is a verifiable statement.

Note that this cannot be generalized to uncountable infinite collections as the procedure would not be guaranteed to eventually run all tests. In fact, suppose the only way to test the uncountable disjunction is to test each statement individually. We would then need to, at least, run all of them for a finite time, say one minute. Even assuming arbitrary long time, we would only have countably many minutes at our disposal. Since the set of tests is uncountable, we are not going to be able to run each test for one minute. Therefore we are not justified to assume the uncountable disjunction of verifiable statements is verifiable. \square

Proposition 1.33. *The conjunction and disjunction of a finite collection of decidable statements are decidable. Formally, let $\{s_i\}_{i=1}^n \subseteq \mathcal{S}_d$ be a finite collection of decidable statements. Then the conjunction $\bigwedge_{i=1}^n s_i \in \mathcal{S}_d$ and the disjunction $\bigvee_{i=1}^n s_i \in \mathcal{S}_d$ are decidable statements.*

Proof. Let $\{s_i\}_{i=1}^n \subseteq \mathcal{S}_d$ be a finite collection of decidable statements. Then $\{s_i\}_{i=1}^n \subseteq \mathcal{S}_v$ and $\{\neg s_i\}_{i=1}^n \subseteq \mathcal{S}_v$ are verifiable. Consider $\bigwedge_{i=1}^n s_i$: this is the finite conjunction of verifiable statements and is therefore a verifiable statement by 1.31. Consider its negation $\neg \bigwedge_{i=1}^n s_i \equiv \bigvee_{i=1}^n \neg s_i$: this is the finite disjunction of verifiable statements and is therefore a verifiable statement by 1.32. The finite conjunction of decidable statements is decidable by definition.

Similarly, consider $\bigvee_{i=1}^n s_i$: this is the finite disjunction of verifiable statements and is therefore a verifiable statement by 1.32. Consider its negation $\neg \bigvee_{i=1}^n s_i \equiv \bigwedge_{i=1}^n \neg s_i$: this is the finite conjunction of verifiable statements and is therefore a verifiable statement by 1.31. The finite disjunction of decidable statements is decidable by definition.

Note that this cannot be generalized to infinite collection as it would require closure under infinite conjunction. Also note that this result is consistent with the experimental tests given in 1.31 and 1.32. \square

Taken as a whole, finite conjunction and countable disjunction define the **algebra of verifiable statements**. It is limited compared to the algebra of statements and it tells us that, in practice, we are not going to be able in general to construct an experimental test whose success is an arbitrary function of the success of other tests.

Operator	Gate	Statement	Verifiable Statement	Decidable Statement
Negation	NOT	allowed	disallowed	allowed
Conjunction	AND	arbitrary	finite	finite
Disjunction	OR	arbitrary	countable	finite

Table 1.3: Comparing algebras of statements.

Before we continue, it is interesting and useful to stop and understand the interplay between scientific and mathematical constructs.¹⁵ Technically, 1.2, 1.4, 1.9, 1.27, 1.31 and 1.32 are the axioms of our mathematical formalism for statements. Note that the actual content of the statements, the methodology for which an assignment is deemed possible and the procedures for experimental verification are not formally defined: the math simply uses symbols to label and identify them. The only assumptions are that statements exist (organized in logical contexts with possible assignments, one of which is the true one), that some of them are verifiable and that they admit the associated algebra. The mathematical formalism does not know what the symbols actually represent: they may as well be pieces of cardboard painted black or white. The math can only derive consequence given the premise, but it does not know whether the premise makes actual sense. In other words, the way that we are making the framework mathematically precise is not by making everything mathematically precise: it is by omitting the details that are not amenable to a precise specification.

¹⁵It took us many many confusing years to fully understand where the scientific argument ends and the mathematical argument begins, what makes sense to assume physically and what makes sense to prove rigorously. Part of the confusion is that this line is not objective but it is based on what is considered “precise” by a mathematician, which has evolved considerably through the centuries. The rule we follow is: in the mathematical formalism the only objects that can be left unspecified are the elements of a set.

We should stress this for a couple of reasons. First, the part that is not formalized is *the most important part*. Discovering new science is exactly finding new things to study (i.e. new statements), new connections between known objects (i.e. new logical relationships) or devising new measurement techniques (i.e. new experimental tests). The content of the statements, their semantic relationships and the procedure of the experimental tests *is* the actual science. Everything that follows is, in a sense, the trivial bit and that is why it can be done generally. Which leads to the second reason: understanding whether statements and verifiable statements *actually* follow the algebras we defined is crucial. The math just takes it at face value, it does not prove it. The justifications for our axioms, then, are the most critical part of this work and they are not mathematical proofs. If we botch them, we'll have a nice, consistent, rich but meaningless mathematical framework. Lastly, it has to be clear that something gets lost in the formalization. The mathematical framework cannot carry all the physics content: we removed the most important part! Different systems may have the same mathematical description, so the scientific content can never be entirely reconstructed from the math. That is why we always have to carefully bring it along.

Now that we have characterized verifiable statements we want to understand how to characterize groups of them. Consider the verifiable statements

“that animal is a duck”
“that animal is a swan”
“that animal is white”
“that animal is black”
“that animal is a black swan”
“that animal is a white duck”
“that animal is a duck or a swan”

Since some depend on others, we do not need to actually run all the tests. Once we have tested the first four we have gathered enough information for the others. We call this set a basis.¹⁶

Definition 1.34. *Given a set \mathcal{D} of verifiable statements, $\mathcal{B} \subseteq \mathcal{D}$ is a **basis** if the truth values of \mathcal{B} are enough to deduce the truth values of the set. Formally, all elements of \mathcal{D} can be generated from \mathcal{B} using finite conjunction and countable disjunction.*

Note that once we have tested the basis, we have tested any other verifiable statement that can be constructed from it. In the example before, once we tested the first four we have implicitly tested *“that animal is a black duck”*. It is also true that impossibilities and certainties don't really need to be tested. We already know that *“that animal is a duck and a swan”* is false. The idea, then, is to group verifiable statements into experimental domains that can be seen as all the experimental information one can gather for a particular subject. These will include the certainty, the impossibility and any other verifiable statement that can be constructed from a basis. The basis, though, has to be countable so that, by running one test at a time, we can hope to eventually reach any element.

¹⁶The term basis is used in general to define a set of objects from which, through a series of operations, one can construct the full space. It is the same for a vector space: from a basis one can construct any other vector through linear combination. What changes is what objects are combined and what operations are used.

Definition 1.35. An *experimental domain* \mathcal{D} represents a set of verifiable statements that can be tested and possibly verified in an indefinite amount of time. Formally, it is a set of statements, closed under finite conjunction and countable disjunction, that includes precisely the certainty, the impossibility, and a set of verifiable statements that can be generated from a countable basis.

Justification. In principle, indefinite is different from infinite. Having infinite time at our disposal literally means that we can go on forever, which we cannot do. Having indefinite time means that, while at some point we have to stop, we have to be prepared to keep going on because we do not know when we will stop. In practice, in both cases we have to make a plan for an infinite amount of time. In the indefinite case, our plan will be cut short.

As we have already argued, we cannot run uncountably many tests in infinite time. Let $E \subseteq \mathcal{E}$ be an uncountable set of experimental tests. Let t_0 be the amount of time that the shortest test will take to run. Each test must at least run for t_0 time. Given that each test will succeed in finite non-zero time, at best t_0 is a finite number. We can divide time into slots of t_0 time. The slots will be countable. As E is uncountable, we cannot associate each test to a slot. We cannot test uncountably many tests.

We can, however, test countably many tests. Let $\mathcal{B} = \{e_i\}_{i=1}^{\infty} \in \mathcal{E}$ be a countable set of tests. We can proceed as follows:

1. initialize n to 1
2. for each $i = 1..n$
 - a) run the test e_i for n seconds
3. increment n and go to step 2

This will run all tests for an indefinite amount of time.

Now let $\mathcal{D} \subseteq \mathcal{S}$ be the set of statements generated by \mathcal{B} using finite conjunction and countable disjunction. Then $\mathcal{D} \subseteq \mathcal{S}_v$ is a set of verifiable statements. Furthermore, testing \mathcal{B} will eventually verify each true statement in \mathcal{D} .

Therefore we are justified to assume that a set $\mathcal{D} \subseteq \mathcal{S}_v$ of verifiable statements that can be tested in an indefinite amount of time must have a countable basis. We are also justified to assume that it contains the certainty and the impossibility as these are two verifiable statements, and two verifiable statements can always be added to a countable basis and keep it countable. This justifies the definition. \square

We can think of an experimental domain as the enumeration of all possible verifiable answers to a scientific question. For example, the domain related to the question “what is that animal?” would include “it is a mammal”, “it is a dog”, “it is an animal with feathers” and so on. If two statements are possible answers to that question, then their conjunction and disjunction will also be possible answers. For example: “it is a dog or a cat” or “it is a mammal and it lays eggs”.

While each statement only needs finite time to be verified, we allow indefinite time for the domain because we want to capture those questions that can be answered only approximately. The idea is that, given more time, we can always get a better answer so, in principle, we have an infinite sequence of tests to perform and continue indefinitely.

The basis not only serves as a way to constrain the size of the experimental domain, but most of the time it will also serve to define the experimental domain itself. We will typically start by characterizing a set of verifiable statements (e.g. a set of characteristics of animals and how to identify them) and then consider the domain of all the verifiable statements that can be constructed from them (e.g. the set of all animals and groups of animals we can identify).

1.3 Theoretical domains and possibilities

The basis for an experimental domain allows us to create a procedure that will eventually test any verifiable statement. But to fully characterize a domain we want to find those statements, like “*that animal is a cat*” or “*the mass of the photon is exactly 0 eV*”, that, if true, determine the truth value of all verifiable statements in the domain. The main result of this section is that these statements, which we call possibilities for the domain, are not necessarily verifiable themselves. We will therefore need to introduce theoretical domains which consist of those statements that can be associated to a test. These tests are constructed from those associated to verifiable statements from an experimental domain. We will also be able to conclude that the set of possibilities for an arbitrary experimental domain has at most the cardinality of the continuum, thus putting a hard constraint on what type of mathematical objects are useful in science.

Suppose \mathcal{D}_X is the domain of animal species identification. It will contain verifiable statements such as “*that animal has feathers*”, “*that animal has claws*”, “*that animal has paws*”. Some statements are broader and some are narrower. But some statements, like “*that animal is a mute swan (Cygnus olor)*” or “*that animal is a mallard duck (Anas platyrhynchos)*”, are special because if we verify those then we are able to know which other statements are true or false. Once we verify those we are essentially done. These are what we call the possibilities of the domain and enumerating them means characterizing the experimental domain.

Unfortunately, not all possibilities are verifiable statements. Consider the statements $s_1 =$ “*there is extra-terrestrial life*” and $s_2 =$ “*there is no extra-terrestrial life*”. We can create an experimental test to verify the first (i.e. find extra-terrestrial life somewhere) but not the second (it would require us to check every place in the universe which is something we cannot do). So, for this question, the experimental domain $\mathcal{D}_X = \{s_1, \top, \perp\}$ is composed of the first statement, the certainty and the impossibility. But s_2 is conceptually still one of the possibilities: if true we have a complete answer for the domain.

What happens is that while the negation of a verifiable statement is not always a verifiable statement, it is still a statement that can be associated to an experimental test. As such, while we cannot verify the statement, we can still predict the outcome of its test, including non-termination. To be able to find all possibilities, then, we have to create the set of statements that can be associated with experimental tests regardless of termination, which will include negations. We call this set the theoretical domain and theoretical statements its elements.

Definition 1.36. *The **theoretical domain** $\bar{\mathcal{D}}$ of an experimental domain \mathcal{D} is the set of statements constructed from \mathcal{D} to which we can associate a test regardless of termination. We call **theoretical statement** a statement that is part of a theoretical domain. More formally, $\bar{\mathcal{D}}$ is the set of all statements generated from \mathcal{D} using negation, finite conjunction and countable disjunction.*

Justification. The theoretical domain is defined to contain all statements that depend on the verifiable statements for which we can associate an experimental test. A statement in the theoretical domain, therefore, must be associated to a procedure that can be constructed in terms of the tests of the verifiable statements, regardless of whether the procedure is guaranteed to terminate.

First of all, any verifiable statement is a theoretical statement. In fact, let $s \in \mathcal{D}$ be a verifiable statement. Then there is a test $e \in \mathcal{E}$ associated to s . Therefore $s \in \bar{\mathcal{D}}$ is a theoretical statement.

In the previous justifications, we saw how to construct tests for negation, finite conjunction and countable disjunction. We can use these to construct tests of theoretical statements from other theoretical statements. Let $s \in \bar{\mathcal{D}}$ be a theoretical statement. Then there is a test $e \in \mathcal{E}$ associated to s . Consider $\neg s$. This can be associated with the test $e_{\neg}(e) \in \mathcal{E}$. Therefore $\neg s$ is a theoretical statement. Let $\{s_i\}_{i=1}^n \in \bar{\mathcal{D}}$ be a finite collection of theoretical statements. Then there is a finite collection of tests $\{e_i\}_{i=1}^n \subseteq \mathcal{E}$ each associated to the respective statement. Consider $\bigwedge_{i=1}^n s_i$. This can be associated with the test $\bigwedge_{i=1}^n e_i \in \mathcal{E}$.

Therefore $\bigwedge_{i=1}^n s_i$ is a theoretical statement. Let $\{s_i\}_{i=1}^{\infty} \in \bar{\mathcal{D}}$ be a countable collection of theoretical statements. Then there is a countable collection of tests $\{e_i\}_{i=1}^{\infty} \subseteq \mathcal{E}$ each associated to the respective statement. Consider $\bigvee_{i=1}^{\infty} s_i$. This can be associated with the test $\bigvee_{i=1}^{\infty} e_i \in \mathcal{E}$.

Therefore $\bigvee_{i=1}^{\infty} s_i$ is a theoretical statement.

Therefore we are justified to assume the theoretical domain contains the closure of the experimental domain under negation, finite conjunction and countable disjunction. As shown later, this will automatically include countable conjunction as well.

Note that, as before, this cannot be generalized to uncountable operations, as a procedure can be composed only of countably many non-zero-time operations. Therefore we are not justified to close under uncountable operations. \square

Because of its construction, the theoretical domain will also include all the limits of all the sequences of verifiable statements. Consider the experimental domain for the mass of the photon. It will contain verifiable statements such as

$s_1 = \text{“the mass of the photon is smaller than } 10^{-1} \text{ eV”}$

$s_2 = \text{“the mass of the photon is smaller than } 10^{-2} \text{ eV”}$

$s_3 = \text{“the mass of the photon is smaller than } 10^{-3} \text{ eV”}$

...

It will not contain the statement $s = \text{“the mass of the photon is exactly } 0 \text{ eV”}$, though, as we cannot measure a continuous quantity with infinite precision.

Note s can be seen as the limit of the sequence of ever increasing precision, but can also be seen as the conjunction for all those statements $s = \bigwedge_{i=1}^{\infty} s_i$. In fact, the mass of the photon is exactly 0 if and only if all the finite precision measurements will contain 0 in the range. It makes sense, then, that it is not part of the experimental domain because only finite conjunctions of verifiable statements are verifiable. But we expect s to be a possibility for the mass of the photon. Why should it be in the theoretical domain?

Because of the De Morgan properties in 1.19, we can express conjunctions in terms of negation and disjunction. So we have $s = \bigwedge_{i=1}^{\infty} s_i = \neg \bigvee_{i=1}^{\infty} \neg s_i$. Therefore by allowing negation we are also allowing countable conjunction and therefore we are including all the limits of sequences of verifiable statements.

Proposition 1.37. *All theoretical domains are closed under countable conjunction.*

Proof. Any countable conjunction $s = \bigwedge_{i=1}^{\infty} s_i$ is equivalent to the negation of disjunction of the negation: $s = \neg \bigvee_{i=1}^{\infty} \neg s_i$. As the theoretical domain is closed under negation and countable disjunction, so it is closed under countable conjunction. \square

Definition 1.38. *A theoretical statement $s \in \bar{\mathcal{D}}$ is **approximately verifiable** if it is the limit of some sequence of verifiable statements. Formally, $s \in \bar{\mathcal{D}}$ is approximately verifiable if there exists a sequence $\{s_i\}_{i=1}^{\infty} \in \mathcal{D}$ such that $s = \bigwedge_{i=1}^{\infty} s_i$.*

Note that we are closed under countable operations and not arbitrary (e.g. uncountable). Therefore there could be statements that can be constructed from verifiable statements that are not even theoretical statements. One such statement, for example, could be constructed given a set U of possible mass values for a particle that is uncountable, has an uncountable complement, and where the elements are picked arbitrarily and not according to a simple rule.¹⁷ The statement “the mass of the particle expressed in eV is in the set U ” can only be tested by checking each value individually. But since the set is uncountable and a procedure can only be made of countably many steps, it will be impossible to construct a test for such a statement.

A theoretical statement, then, is one for which we can at least conceive an experimental test. This may not always terminate if the statement is true or it may not always terminate if the statement is false, but at least we have one. The statements that depend on the experimental domain but are not part of the theoretical domain do not even hypothetically allow for a procedure, regardless of the fact that it can terminate, and therefore we do not consider them part of our scientific discourse, even theoretically.

In general, given a theoretical statement \bar{s} , we would like to characterize what experimental test can be associated to it. Ideally, we want the experimental test for that statement that terminates, successfully or unsuccessfully, in the most cases. Consider all the verifiable statements that are narrower than \bar{s} . If we take their disjunction we get the broadest verifiable statement that is still narrower than \bar{s} . We call this the verifiable part of \bar{s} , noted $\text{ver}(\bar{s})$. Testing $\text{ver}(\bar{s})$ means running the test that is guaranteed to terminate successfully in the broadest situations in which \bar{s} is true. In fact, if \bar{s} is itself verifiable then $\text{ver}(\bar{s})$ will be exactly \bar{s} .

Conversely, consider all the verifiable statements that are incompatible with \bar{s} . If we take their disjunction we get the broadest verifiable statement that is still incompatible with \bar{s} . We call this the falsifiable part of \bar{s} , noted $\text{fal}(\bar{s})$. Testing $\text{fal}(\bar{s})$ means running the test that is guaranteed to terminate successfully in the broadest situations in which \bar{s} is false. In fact, if \bar{s} is itself falsifiable then $\text{fal}(\bar{s})$ will be exactly $\neg \bar{s}$.

¹⁷Mathematically, we are looking for a set of real numbers that is not a Borel set.

To each theoretical statement, then, we associate the experimental test constructed by returning successfully if the test for the verifiable part succeeds and returning unsuccessfully if the test for the falsifiable part succeeds. We will not be able to terminate if either of those doesn't terminate, which will correspond to the statement $\neg \text{ver}(\bar{s}) \wedge \neg \text{fal}(\bar{s})$ being true. We call this the undecidable part of \bar{s} , noted $\text{und}(\bar{s})$.

In light of this, consider $s =$ "the mass of the photon is rational as expressed in eV". It is the disjunction of all possibilities with rational numbers, which is countable, and therefore is a theoretical statement. Since we can only experimentally verify finite precision intervals, each verifiable statement will include infinitely many rational (and irrational) numbers. Therefore no verifiable statement is narrower than s and therefore $\text{ver}(s) \equiv \perp$. But for the same reason no verifiable statement is incompatible with s and therefore $\text{fal}(s) \equiv \perp$. Which means $\text{und}(\bar{s}) \equiv \top$. This means that the experimental test for s will never terminate either successfully or unsuccessfully. We call this type of statements undecidable as we will never be able to experimentally test anything about them.

Definition 1.39. Let $\bar{s} \in \bar{\mathcal{D}}$ be a theoretical statement. We call the **verifiable part** $\text{ver}(\bar{s}) = \bigvee_{s \in \mathcal{D} | s \leq \bar{s}} s$ the broadest verifiable statement that is narrower than \bar{s} . We call the **falsifiable part** $\text{fal}(\bar{s}) = \bigvee_{s \in \mathcal{D} | s \not\leq \bar{s}} s$ the broadest verifiable statement that is incompatible with \bar{s} . We call the **undecidable part** $\text{und}(\bar{s}) = \neg \text{ver}(\bar{s}) \wedge \neg \text{fal}(\bar{s})$ the broadest statement incompatible with both the verifiable and the falsifiable part.

Justification. Let $e_{\bar{s}}$ be the optimal test for $\bar{s} \in \bar{\mathcal{D}}$, that is the experimental test constructible from those associated to \mathcal{D} that terminates under most condition. Let $\text{ver}(\bar{s})$ be the verifiable statement associated with that test. We must have $\text{ver}(\bar{s}) \in \mathcal{D}$ since it is constructible from elements of the domain. Consider $s_v = \bigvee_{s \in \mathcal{D} | s \leq \bar{s}} s$. We have $s_v \leq \bar{s}$ and therefore the test e associated with s_v must terminate successfully whenever \bar{s} is true. Since $e_{\bar{s}}$ terminates under most conditions, it must terminate successfully whenever e terminates successfully. If it didn't, we could construct the test $e_{\bar{s}} \vee e$ which would terminate in more cases, and therefore $e_{\bar{s}}$ would not be optimal. Therefore we must have $\text{ver}(\bar{s}) \geq s_v$. However, we cannot have $\text{ver}(\bar{s}) > s_v$. Since $\text{ver}(\bar{s}) \in \mathcal{D}$, this would imply $\text{ver}(\bar{s}) > \text{ver}(\bar{s})$. Therefore we must have $\text{ver}(\bar{s}) = \bigvee_{s \in \mathcal{D} | s \leq \bar{s}} s$.

Along the same lines, let $\text{fal}(s)$ be the verifiable statement associated with $e_{\neg}(\bar{s})$. Noting that $s \leq \bar{s}$ if and only if $s \not\leq \bar{s}$, we find $\text{fal}(\bar{s}) = \bigvee_{s \in \mathcal{D} | s \not\leq \bar{s}} s$. This justifies the definitions. \square

Corollary 1.40. The verifiable, falsifiable and undecidable part partition the certainty. That is, for every $\bar{s} \in \bar{\mathcal{D}}$ we have $\top \equiv \text{ver}(\bar{s}) \vee \text{und}(\bar{s}) \vee \text{fal}(\bar{s})$ while $\text{ver}(\bar{s}) \not\leq \text{und}(\bar{s})$, $\text{ver}(\bar{s}) \not\leq \text{fal}(\bar{s})$ and $\text{und}(\bar{s}) \not\leq \text{fal}(\bar{s})$.

Proof. We have $\text{und}(\bar{s}) = \neg \text{ver}(\bar{s}) \wedge \neg \text{fal}(\bar{s}) \equiv \neg(\text{ver}(\bar{s}) \vee \text{fal}(\bar{s}))$. Therefore $\top \equiv \text{und}(\bar{s}) \vee \neg \text{und}(\bar{s}) \equiv \text{und}(\bar{s}) \vee \text{ver}(\bar{s}) \vee \text{fal}(\bar{s})$.

Since $\text{ver}(\bar{s}) \leq \bar{s}$ and $\text{fal}(\bar{s}) \not\leq \bar{s}$ we have $\text{ver}(\bar{s}) \not\leq \text{fal}(\bar{s})$. Since $\text{und}(\bar{s}) \equiv \neg \text{ver}(\bar{s}) \wedge \neg \text{fal}(\bar{s}) \leq \neg \text{ver}(\bar{s})$ and therefore $\text{und}(\bar{s}) \not\leq \text{ver}(\bar{s})$. Similarly, we have $\text{und}(\bar{s}) \leq \neg \text{fal}(\bar{s})$ and therefore $\text{und}(\bar{s}) \not\leq \text{fal}(\bar{s})$. \square

Corollary 1.41. A theoretical statement $\bar{s} \in \bar{\mathcal{D}}$ is verifiable if and only if $\bar{s} \equiv \text{ver}(\bar{s})$. It is falsifiable if and only if $\neg \bar{s} \equiv \text{fal}(\bar{s})$. It is decidable if and only if $\text{und}(\bar{s}) \equiv \perp$.

Proof. Let \bar{s} be a verifiable statement. Then \bar{s} itself is the broadest verifiable statement compatible with itself and therefore $\bar{s} \equiv \text{ver}(\bar{s})$. Conversely, let $\bar{s} \equiv \text{ver}(\bar{s})$, then \bar{s} is equivalent to a verifiable statement and is therefore verifiable.

Let \bar{s} be a falsifiable statement. Then $\neg\bar{s}$ itself is the broadest verifiable statement incompatible with \bar{s} and therefore $\neg\bar{s} \equiv \text{fal}(\bar{s})$. Conversely, let $\neg\bar{s} \equiv \text{fal}(\bar{s})$, then $\neg\bar{s}$ is equivalent to a verifiable statement and \bar{s} is therefore falsifiable.

Let \bar{s} be a decidable statement. Then $\text{und}(\bar{s}) = \neg\text{ver}(\bar{s}) \wedge \neg\text{fal}(\bar{s}) \equiv \neg\bar{s} \wedge \neg\neg\bar{s} \equiv \neg\bar{s} \wedge \bar{s} \equiv \perp$. Conversely, let $\text{und}(\bar{s}) \equiv \perp$. Then $\text{ver}(\bar{s})$ and $\text{fal}(\bar{s})$ are two incompatible statements whose disjunction is a certainty. This means $\text{ver}(\bar{s}) \equiv \neg\text{fal}(\bar{s})$. Since $\neg\text{fal}(\bar{s}) \equiv \text{ver}(\bar{s}) \leq \bar{s} \leq \neg\text{fal}(\bar{s}) \equiv \text{ver}(\bar{s})$ then $\bar{s} \equiv \text{ver}(\bar{s})$ and $\neg\bar{s} \equiv \text{fal}(\bar{s})$. Therefore \bar{s} is decidable. \square

Corollary 1.42. *For every theoretical statement \bar{s} , the undecidable part $\text{und}(\bar{s})$ is a falsifiable statement.*

Proof. Since $\text{und}(\bar{s}) \equiv \neg(\text{ver}(\bar{s}) \vee \text{fal}(\bar{s}))$ then it is the negation of a verifiable statement and is therefore falsifiable. \square

Definition 1.43. *A theoretical statement $\bar{s} \in \bar{\mathcal{D}}$ is **undecidable** if $\text{und}(\bar{s}) \equiv \top$.*

The theoretical domain, then, contains all sort of edge cases. Similar to undecidable statements, we could have statements that can just never be verified experimentally (i.e. $\text{ver}(\bar{s}) \equiv \perp$) or just never falsified (i.e. $\text{fal}(\bar{s}) \equiv \perp$). The possible presence of these types of statements means we have to be cautious in giving physical significance to all theoretical statements, as the only thing we may be able to say about some is that nothing can be said about them.¹⁸

In particular, we want to be able to understand when two statements represent two distinct situations that we can tell apart experimentally. If the two statements are verifiable (e.g. “the mass is between 1 and 2 Kg” and “the mass is between 2 and 3 Kg”), incompatibility will be enough to tell us that the two cases are distinct (only one can happen at a time) since verifying one will be enough to exclude the other. If the statements are not both verifiable (e.g. “the mass is between 1 and 2 Kg” and “the mass is exactly 2 Kg”) being incompatible is not enough: one may lie in the undecidable part of the other in which case we cannot verify one and exclude the other at the same time. In that case, we would not be able to distinguish the two cases experimentally.

For two statements \bar{s}_1 and \bar{s}_2 to be experimentally distinguishable, then, we need to have an experimental test that always terminates successfully in one case and always terminates unsuccessfully in the other. That is, we need to have a third statement \bar{s} whose verifiable part is broader than one (e.g. $\text{ver}(\bar{s}) \geq \bar{s}_1$) and whose falsifiable part is broader than the other (e.g. $\text{fal}(\bar{s}) \geq \bar{s}_2$). This way if \bar{s} is found to be true experimentally, we know that \bar{s}_1 may be true and \bar{s}_2 must be false; if \bar{s} is found to be false experimentally, we know that \bar{s}_1 must be false and \bar{s}_2 may be true; if the test for \bar{s} does not terminate, then both statements must be false so we are in neither of those cases.

¹⁸It will be a recurrent theme of this work to make a precise distinction between mathematical objects that represent physical entities (e.g. verifiable statements), those that represent idealizations of physical entities (e.g. theoretical statements) and those that do not have scientific standing (e.g. statements that are neither verifiable nor theoretical).

Definition 1.44. Let $\bar{s}_1, \bar{s}_2 \in \bar{\mathcal{D}}$ be two theoretical statements. We say they are **experimentally distinguishable** if there is an experimental test that can tell them apart. Formally, we can find a theoretical statement $\bar{s} \in \bar{\mathcal{D}}$ such that $\bar{s}_1 \preceq \text{ver}(\bar{s})$ and $\bar{s}_2 \preceq \text{fal}(\bar{s})$.

Justification. Suppose we have an experimental test e constructible from those associated to \mathcal{D} that can tell \bar{s}_1 and \bar{s}_2 apart. Then we must have that, without loss of generality, $\text{result}(e, a) = \text{SUCCESS}$ for all $a \in \mathcal{A}_{\mathcal{S}}$ for which $a(\bar{s}_1) = \text{TRUE}$ and $\text{result}(e, a) = \text{FAILURE}$ for all $a \in \mathcal{A}_{\mathcal{S}}$ for which $a(\bar{s}_2) = \text{TRUE}$. Without loss of generality, we can take e to be an optimal test, since we could make it optimal by combining it with other tests. Let \bar{s} be a statement associated with that optimal test. Then we have $\bar{s}_1 \preceq \text{ver}(\bar{s})$ and $\bar{s}_2 \preceq \text{fal}(\bar{s})$. This justifies the definition. \square

Proposition 1.45. Two theoretical statements $\bar{s}_1, \bar{s}_2 \in \bar{\mathcal{D}}$ are experimentally distinguishable if and only if we can find two verifiable statements $s_1, s_2 \in \bar{\mathcal{D}}$ such that $s_1 \not\preceq s_2$, $\bar{s}_1 \preceq s_1$ and $\bar{s}_2 \preceq s_2$.

Proof. Let $\bar{s}_1, \bar{s}_2 \in \bar{\mathcal{D}}$ be experimentally distinguishable. Then we can find \bar{s} such that $\bar{s}_1 \preceq \text{ver}(\bar{s})$ and $\bar{s}_2 \preceq \text{fal}(\bar{s})$. Moreover, $\text{ver}(\bar{s}) \not\preceq \text{fal}(\bar{s})$. Therefore $\text{ver}(\bar{s})$ and $\text{fal}(\bar{s})$ are verifiable statements that satisfy the condition in the proposition.

Conversely, let $s_1, s_2 \in \bar{\mathcal{D}}$ such that $s_1 \not\preceq s_2$, $\bar{s}_1 \preceq s_1$ and $\bar{s}_2 \preceq s_2$. We have $\bar{s}_1 \preceq s_1 \equiv \text{ver}(s_1)$. We also have $\bar{s}_2 \preceq s_2 \preceq \text{fal}(s_1)$ because $\text{fal}(s_1)$ by definition is broader than all verifiable statements incompatible with s_1 . Therefore $\bar{s}_1, \bar{s}_2 \in \bar{\mathcal{D}}$ are experimentally distinguishable. \square

To sum up, verifiable statements are the only ones that, perhaps under some simplifying assumption, we can think of as tangible scientific objects (e.g. the idea that we can measure mass with finite precision). From these we construct theoretical statements that represent our idealizations (e.g. the infinitely precise value for the mass of a particle or the idea that said value is a rational number expressed in a particular unit). And then there are statements that are not even physically meaningful as they have no well defined experimental consequences.

We have seen that the theoretical domain may contain more statements, but we want to make it clear that it never contains more information. That is, if we knew which verifiable statements were true and which weren't, we would automatically know which theoretical statements would be true or not. So it is not adding extra cases. It is essentially completing the list of answers by adding a “no” if only “yes” can be experimentally verified and vice-versa. To see that this is the case, we can show that a basis for an experimental domain \mathcal{D} is also a basis for its theoretical domain $\bar{\mathcal{D}}$. That is, all verifiable and theoretical statements can be expressed as functions of the same basis. The difference is that a theoretical statement can be a function of the negation of the element of a basis.

Proposition 1.46. The truth values of the statements of a basis \mathcal{B} for an experimental domain \mathcal{D} are enough to determine the truth values for all statements in the associated theoretical domain $\bar{\mathcal{D}}$. More formally, all statements in the theoretical domain $\bar{\mathcal{D}}$ can be generated by negation, countable conjunction and countable disjunction from a basis \mathcal{B} of \mathcal{D} .

Proof. By definition of basis, any verifiable statement within the experimental domain

\mathcal{D} can be generated from \mathcal{B} using only finite conjunction and countable disjunction. The certainty may be generated through negation and disjunction from any verifiable statement therefore \mathcal{B} generates \mathcal{D} which in turn generates $\bar{\mathcal{D}}$ by definition. This means that \mathcal{B} , through negation, countable conjunction and countable disjunction, generates all of $\bar{\mathcal{D}}$. \square

Having defined what a theoretical domain is, we can finally define what the possibilities of a domain are: those statements that if known to be true determine the truth value of all other statements.

Definition 1.47. A *possibility* for an experimental domain \mathcal{D} is a statement $x \in \bar{\mathcal{D}}$ that, when true, determines the truth value for all statements in the theoretical domain. Formally, $x \neq \perp$ and for each $s \in \bar{\mathcal{D}}$, either $x \leq s$ or $x \not\leq s$. The **full possibilities**, or simply the **possibilities**, X for \mathcal{D} are the collection of all possibilities.

A possibility represents a complete answer for a scientific question. Only one of them can be true and one of them must be true since the theoretical domain contains all negations. But how can we construct them? Suppose \mathcal{D} is the experimental domain for animal species identification. There will be a set of possible truth assignments for it. Suppose $\mathcal{B} \subset \mathcal{D}$ is a basis of statements, like “that animal has feathers”, “that animal has claws”, “that animal has paws” and so on, that allows us to fully identify the animal species. Then each possible assignment for the basis will correspond to one and only one assignment for the domain. Consider a minterm, a conjunction where each statement appears once either negated or not. For example, “that animal has feathers” \wedge “that animal has claws” $\wedge \neg$ “that animal has paws” $\wedge \dots$. If that statement is true, it will determine the truth value of all the basis, it will select one possible truth assignment for the whole domain. Then a possibility for the domain is simply a minterm of a basis that corresponds to a possible assignment for the domain.

Proposition 1.48. Let \mathcal{D} be an experimental domain. A possibility for \mathcal{D} is any minterm of a basis that is not impossible.

Proof. Let $\mathcal{B} \subseteq \mathcal{D}$ be a basis for \mathcal{D} . Let x be a minterm of \mathcal{B} . Any theoretical statement $s \in \bar{\mathcal{D}}$ can be expressed as the disjunction of minterms of \mathcal{B} by 1.26 and by 1.46. x is either within the minterms needed to express s , or not. If it is, $x \wedge s \equiv x$ and therefore $x \leq s$. If it’s not, $x \wedge \neg s \equiv x$ and therefore $x \not\leq s$. Therefore a minterm is either narrower or incompatible with all theoretical statements and, if it is not impossible, it is a possibility by definition.

Conversely, suppose $x \in \bar{\mathcal{D}}$ is a possibility. As it is a theoretical statement, it can be expressed as a disjunction of minterms of a basis \mathcal{B} . Suppose it is the disjunction of more than one minterm. Then each minterm would be narrower than x , which cannot be. x must be expressed by a single minterm. Therefore any possibility is a minterm. \square

Proposition 1.49 (No other possibilities). All statements that determine and only determine the truth value of all statements in a theoretical domain $\bar{\mathcal{D}}$ are possibilities of \mathcal{D} . Formally, there is no statement $s \in \mathcal{S}$ that has all the properties of a possibility except that $s \notin \bar{\mathcal{D}}$.

Proof. Let $x \in \mathcal{S}$ be a statement that determines and only determines all truth values of all statements in a theoretical domain $\bar{\mathcal{D}}$. This is equivalent to determining the truth

values and only the truth values of all elements of a basis $\mathcal{B} \subseteq \mathcal{D}$. As we can find a countable basis, the statement x is equivalent to the countable conjunction of statements of \mathcal{B} or their negation. Therefore $x \in \bar{\mathcal{D}}$ as it is generated by the statements of the basis by negation and countable conjunction. But a statement in $\bar{\mathcal{D}}$ that determines all truth values of the statements in $\bar{\mathcal{D}}$ is a possibility by definition. Therefore x is a possibility. \square

There is one possibility that is often forgotten and sometimes needs special handling. Suppose one is trying to identify an illness by going through a series of known markers. It may happen that no match for the disease is found because we are dealing with a new kind of illness. In the same way, we may fail to measure the value of a quantity because it lies outside the sensitive range of our equipment. In other words, it may be possible that none of our tests succeed and none of the verifiable statements is verified. We call this possibility the residual because it's what remains after we went through all the cases we already know.

Note, though, that the residual possibility does not exist for all domains. Suppose we have a basket of fruit and we want to count how many whole apples there are. There can only be a finite number of them, and we can successfully identify all finite numbers: there is no "something else" to be found in this case.

Definition 1.50. The *residual possibility* \hat{x} for an experimental domain \mathcal{D} is, if it exists, the possibility that predicts that no test will be successful. Formally, let $\mathcal{B} \subseteq \mathcal{D}$ be a basis then $\hat{x} = \bigwedge_{e \in \mathcal{B}} \neg e$ if it is not impossible. An experimental domain is **complete** if it doesn't admit a residual possibility.

Definition 1.51. The *established possibilities* $\hat{X} = X \setminus \{\hat{x}\}$ for an experimental domain is the set of all possibilities excluding the residual possibility.

As we refine our understanding and techniques for a domain of knowledge, we may find that the residual possibility actually corresponds to multiple cases that weren't previously cataloged. So, intuitively, it is better thought of as a bucket that contains all that is yet to be experimentally discovered within the particular domain of knowledge. Because of its special nature, the residual possibility sometimes behaves differently than the other possibilities. Therefore we will find that some theorems are more elegantly expressed in terms of the established possibilities and some in terms of the full possibilities.

To sum up the mathematical structure of the domain, consider the following table:

$\overbrace{\hspace{14em}}^{\bar{\mathcal{D}}}$													
$\overbrace{\hspace{9em}}^{\mathcal{D}}$									$\overbrace{\hspace{5em}}^X$				
$\overbrace{\hspace{3em}}^{\mathcal{B}}$			$\overbrace{\hspace{3em}}$			$\overbrace{\hspace{3em}}$			$\overbrace{\hspace{4em}}$				
b_1	b_2	...	s_1	s_2	...	\bar{s}_1	\bar{s}_2	...	x_1	x_2	x_3	x_4	...
T	F	...	T	F	...	F	F	...	T	F	F	F	...
T	T	...	T	F	...	F	T	...	F	T	F	F	...
F	F	...	T	T	...	F	F	...	F	F	T	F	...
T	F	...	F	F	...	T	T	...	F	F	F	T	...

On the left we find the basis elements $b \in \mathcal{B}$. From these, through finite conjunction and

countable disjunction we construct all verifiable statements $s \in \mathcal{D}$. From these, through negation, countable conjunction and countable disjunction, we construct all theoretical statements $\bar{s} \in \bar{\mathcal{D}}$. Within these we find the possibilities $x \in X$. Each row corresponds to a possible truth assignment for the domain. Each possibility is true in one and only one possible assignment. The mathematical structure, then, is simply there to keep track of the logical relationships between all the statements, which assignments are possible and which statements are verifiable.

With our definitions in mind, we can answer the following fundamental question: what is the maximum number of possibilities that an experimental domain can have? In other words: what is the maximum number of cases among which we can distinguish experimentally? As we saw before, a possibility is a statement that defines the truth values of a basis. Since a basis is countable, we can uniquely identify a possibility by a countable sequence of TRUE or FALSE. Note that a real number expressed in a binary basis (e.g. 0.10110001...) will also be uniquely identified by such a sequence: the cardinality of the possibilities is at most that of the continuum.

Theorem 1.52. *The possibilities X for an experimental domain \mathcal{D} have at most the cardinality of the continuum.*

Proof. Let $\mathcal{B} = \{e_i\}_{i=1}^{\infty} \subseteq \mathcal{D}$ be a countable basis. Let $2^{\mathbb{N}}$ denote the set of infinite binary sequences. We define the function $F : X \rightarrow \mathbb{B}^{\mathbb{N}}$ such that $F(x) = \{F(x)_i\}_{i=1}^{\infty}$ is given by:

$$F(x)_i = \begin{cases} \text{TRUE} & x \approx e_i \\ \text{FALSE} & x \not\approx e_i \end{cases}$$

For each $x \in X$ we have $x = \bigwedge_{i=1}^{\infty} \neg^{F(x)_i} e_i$. Suppose $x_1 \neq x_2$, then $F(x_1)_i \neq F(x_2)_i$ for some i , therefore F is injective. We then have $|X| \leq |2^{\mathbb{N}}| = |\mathbb{R}|$. X has at most the cardinality of the continuum. \square

This means we have an upper bound on how many cases can be distinguished experimentally: only up to the continuum. We are not going to be able to tell apart experimentally more possibilities than those. This result gives us a basic requirement for any mathematical object we want to use in a scientific theory: if the cardinality is greater than the continuum, it cannot have a well defined experimental meaning. For example, while the set of all continuous functions between real numbers has the cardinality of the continuum, the set of all functions (including discontinuous ones) has greater cardinality. We can already conclude that the first set may be useful to represent physical objects and the second may not.

As we are looking into the cardinality of the possibilities of a domain, it should not be surprising that the only way we can have infinitely many possibilities is if we are given infinitely many verifiable statements. As each verifiable statement can distinguish between two cases (i.e. true or false), we need infinitely many of these distinctions to reach infinitely many cases.

Proposition 1.53. *Let \mathcal{D}_X be an experimental domain. The following are equivalent:*

1. *the set of possibilities is finite*

2. *the experimental domain contains finitely many verifiable statements*
3. *there exists a finite basis*

Proof. To prove 2 from 1, let \mathcal{D}_X be an experimental domain and X its possibilities. Suppose X is finite. Since any statement in \mathcal{D}_X can be generated from X using finite conjunction, \mathcal{D}_X must be finite as well.

To prove 3 from 2, let \mathcal{D}_X contain finitely many verifiable statements. Then a basis will consist of finitely many statements as it is a subset of the domain.

To prove 1 from 3, suppose \mathcal{B} is a finite basis for the domain. Since the possibilities are minterms of the basis, we have $|X| \leq |2^{\mathcal{B}}|$. Since the basis is finite, X is finite as well. \square

We have now presented the fundamental objects of our general mathematical theory of experimental science. In our framework, a “scientific theory” or a “scientific model” is an experimental domain: a set of statements, what they mean and how to verify them experimentally.

The starting point is often a basis: a set of verifiable statements which defines all the domain knowledge that we can gather experimentally. The content of the statements determines what combinations can be true at the same time, which defines the possibilities for our domain. Everything in the domain is grounded within the verifiable statements: there is nothing else in it. This also maps well to the practice of most scientific fields, where one defines states and other physical objects based on what can be measured.

As we’ll see later, some verifiable statements may be idealizations. For example, we may assume that a quantity can be measured with arbitrary level of precision, which we know not to be strictly true. We may assume a volume of gas to have a well defined temperature, which we know not to be true if it is not at equilibrium. This type of simplification makes a domain applicable only within the realm of validity of those idealizations, but does not change the formal structure we have identified here. In fact, the focus of much of this work will be deriving the details of different experimental domains under different physical assumptions.

The main point of our framework is that this conceptual structure is inescapable once we set the principle of scientific objectivity. We will always need a set of statements and a way to test them and, if we are given those, we have all we need. These elements are necessary and sufficient to be able to do science. And by being clear about which statements can be considered physically meaningful, and which are an idealization, we can then be more precise on the physical status of each component of a particular scientific theory.

1.4 Topological spaces

Now that we have defined what experimental domains are, we want to explore the link between them and some fundamental mathematical structures. The main result of this section is that an experimental domain provides a natural topology for its possibilities. Each verifiable statement can be seen as the disjunction of a set of possibilities. Performing finite conjunction and countable disjunction of verifiable statements means performing finite union and countable intersection on those sets of possibilities.

Topological spaces were developed in the first half of the 1900s as a generalization of metric spaces. The idea is to define a notion of closeness without having to define an actual

distance.¹⁹ Other branches of math (e.g. metric spaces, differential geometry, Lie algebras) now see their foundation on topological spaces, which therefore play a very important role in mathematics as a whole. In our case, this notion of closeness will map to how hard it is to tell possibilities apart. That is, possibilities that are topologically closer are more difficult to distinguish experimentally.

Let's first review what a topology is. The general idea is that we have a set X of elements, which we call points, and a collection of subsets of X such that it is closed under finite intersection and arbitrary union, contains the empty set and contains the whole set X . For example, suppose $X = \{1, 2, 3\}$ then $\{\{\}, \{1\}, \{2\}, \{1, 2, 3\}\}$ is not a topology while $\{\{\}, \{1\}, \{2\}, \{1, 2\}, \{1, 2, 3\}\}$ is. The first one is missing the union of $\{1\}$ and $\{2\}$.

Definition 1.54. *Let X be a set. A **topology** on X is a collection \mathbb{T}_X of subsets of X closed under finite intersection and arbitrary union such that it contains X and \emptyset . A **topological space** is a tuple (X, \mathbb{T}_X) of a set and a topology defined on it.*

Mathematicians designed this abstract mathematical structure because it is a useful and general tool to study the notion of continuity. It also happens that all the mathematical structures used in science are topological spaces. Why is that? What is it that topological spaces capture?

Let's go back to our verifiable statements and possibilities. For example, consider $s_1 =$ "the mass of the photon is less than 10^{-10} eV". This can be expressed as $s_1 = \bigvee_{0 \leq x < 10^{-10}}$ "the mass of the photon is precisely x eV": the precise value must be in the given range of possibilities. Consider $s_2 =$ "the mass of the photon is greater than 10^{-20} eV" = $\bigvee_{x > 10^{-20}}$ "the mass of the photon is precisely x eV". The conjunction is the intersection of the possible values: $s_1 \wedge s_2 =$ "the mass of the photon is between 10^{-20} and 10^{-10} eV" = $\bigvee_{10^{-20} < x < 10^{-10}}$ "the mass of the photon is precisely x eV". The disjunction is the union of the possible values: $s_1 \vee s_2 =$ "the mass of the photon can be anything" = $\bigvee_{x \geq 0}$ "the mass of the photon is precisely x eV".

This is something that works in general. In Proposition 1.26 we saw that, if a statement is a function of other statements, it can be expressed as the disjunction of minterms of the arguments. A verifiable statement is a function of basis, so it can be expressed as the disjunction of minterms of a basis. But we have also seen that the minterms of a basis are the possibilities, so each verifiable statement can be expressed as the disjunction of possibilities.

Therefore each statement in the experimental domain defines a set of possibilities, which we call a verifiable set. Since certainty and impossibility are in the domain, the empty set and the full set of possibilities are verifiable sets. Since we can take finite conjunction and countable disjunction of verifiable statements, we can take finite intersection and countable union of verifiable sets. The collection of all verifiable sets forms a topology on the set of possibilities.

¹⁹In science and engineering, one talks about topology also when discussing the structure of a molecule, an electronic circuit or a computer network. These types of structures (i.e. nodes connected by vertices) are studied by graph theory and should not be confused with point-set topology.

Definition 1.55. Let \mathcal{D} be an experimental domain and X its possibilities. We define the map $U : \mathcal{D} \rightarrow 2^X$ that for each statement $s \in \mathcal{D}$ returns the set of possibilities compatible with it. That is: $U(s) \equiv \{x \in X \mid x \approx s\}$. We call $U(s)$ the **verifiable set** of possibilities associated with s .

Proposition 1.56. A statement $s \in \mathcal{D}$ is equivalent to the disjunction of the possibilities in its verifiable set $U(s)$. That is, $s \equiv \bigvee_{x \in U(s)} x$.

Proof. First we show each statement is the disjunction of some set of possibilities. Let \mathcal{D} be an experimental domain, $s \in \mathcal{D}$ a verifiable statement and $\mathcal{B} \subseteq \mathcal{D}$ a basis. Since s is a function of the basis, it can be expressed as a disjunction of minterms of \mathcal{B} . The minterms of \mathcal{B} that are impossible can be ignored since $s \vee \perp \equiv s$. But the minterms of \mathcal{B} that are not impossible are possibilities of \mathcal{D} so $s = \bigvee_{x \in U} x$ for some $U \subseteq X$.

Now we show it is the disjunction of its verifiable set. Let $x \in X$ be a possibility and consider $x \wedge s$. If $x \in U$ then $x \wedge s = x \wedge \bigvee_{\hat{x} \in U} \hat{x} = \bigvee_{\hat{x} \in U} (x \wedge \hat{x}) \equiv x \neq \perp$. Therefore $x \approx s$. If $x \notin U$ then $x \wedge s \equiv \perp$. Therefore $x \not\approx s$. This means $U = U(s)$ as it contains and only contains all the possibilities compatible with s . \square

Proposition 1.57. Let X be the set of possibilities for an experimental domain \mathcal{D} . X has a natural topology given by the collection of all verifiable sets $\mathcal{T}_X = U(\mathcal{D})$.

Proof. The verifiable sets for the certainty and the impossibility correspond to the full set and empty set respectively. Formally, $U(\top) = \{x \in X \mid x \approx \top\} = X$ while $U(\perp) = \{x \in X \mid x \approx \perp\} = \emptyset$. Therefore $X, \emptyset \in U(\mathcal{D})$ since $\top, \perp \in \mathcal{D}$.

The finite intersection of verifiable sets corresponds to the verifiable set of the finite conjunction and therefore it is a verifiable set. Formally, $U(s_1 \wedge s_2) = \{x \in X \mid x \approx (s_1 \wedge s_2)\} = \{x \in X \mid x \approx s_1 \text{ and } x \approx s_2\} = \{x \in X \mid x \approx s_1\} \cap \{x \in X \mid x \approx s_2\} = U(s_1) \cap U(s_2)$.

The countable union of verifiable sets corresponds to the verifiable set of the countable disjunction and therefore it is a verifiable set. Formally, $U(s_1 \vee s_2) = \{x \in X \mid x \approx (s_1 \vee s_2)\} = \{x \in X \mid x \approx s_1 \text{ or } x \approx s_2\} = \{x \in X \mid x \approx s_1\} \cup \{x \in X \mid x \approx s_2\} = U(s_1) \cup U(s_2)$. This generalizes to countable disjunctions. Arbitrary disjunctions can be re-expressed as countable disjunctions, since any verifiable statement can always be expressed in terms of a countable basis.

The collection $\mathcal{T}_X = U(\mathcal{D})$ is therefore a topology by definition since it satisfies all its properties. \square

Mainly for historical reasons, the sets in a topology are called open sets. The complements of open sets are called closed sets. In metric spaces, such as the Euclidean space with the standard topology, these will map to the standard notion of open and closed intervals. But, in general, they do not and this may lead to confusion. For example, if we take the integers with their standard topology, any subset is both open and closed.

Given that we are only interested in the natural topologies of possibilities, we are going to refer to the sets in our topology as verifiable sets and we will occasionally call falsifiable sets their complements. For example, when counting apples a subset of the integers is both verifiable and falsifiable: we can test whether the apple count is within or outside that set of

possible numbers. While this terminology does not follow math convention, we find it more intuitive and meaningful in the context of this work.

We can also re-express the semantic relationships between statements in terms of set operations on the verifiable sets. For example, “*that animal is a cat*” is narrower than “*that animal is a mammal*” because the set of possibilities for which the first is true is a subset of the possibilities for which the second is true. Conversely, “*that animal is a cat*” and “*that animal is a dog*” are incompatible because the set of possibilities in which both are true is empty.

In the following table we summarize how statement operations and relationships are expressed in terms of operations and relationships between sets of possibilities.

	Statement relationship		Set relationship
$s_1 \wedge s_2$	(Conjunction)	$U(s_1) \cap U(s_2)$	(Intersection)
$s_1 \vee s_2$	(Disjunction)	$U(s_1) \cup U(s_2)$	(Union)
$\neg s$	(Negation)	$U(s)^C$	(Complement)
$\text{ver}(s)$	(Verifiable part)	$\text{int}(U(s))$	(Interior)
$\text{fal}(s)$	(Falsifiable part)	$\text{ext}(U(s))$	(Exterior)
$\text{und}(s)$	(Undecidable part)	$\partial U(s)$	(Boundary)
$s_1 \equiv s_2$	(Equivalence)	$U(s_1) = U(s_2)$	(Equality)
$s_1 \preceq s_2$	(Narrower than)	$U(s_1) \subseteq U(s_2)$	(Subset)
$s_1 \succeq s_2$	(Broader than)	$U(s_1) \supseteq U(s_2)$	(Superset)
$s_1 \approx s_2$	(Compatibility)	$U(s_1) \cap U(s_2) \neq \emptyset$	(Intersection not empty)

Table 1.4: Correspondence between statement operators and set operators.

Let’s review the definition of basis and sub-basis for a topology: a collection of sets from which we can generate the whole topology through finite intersection and countable union (for a sub-basis) or just through countable union (for a basis). Bases are important since they are often used in proofs and calculations. Moreover, many properties of topologies can be shown to be equivalent to properties of one of their bases. Countability, and in particular second-countability, is one such property which characterizes the number of verifiable sets in the topology.

Definition 1.58. A collection $\mathcal{B} \subseteq \mathbb{T}_X$ of verifiable sets of X is a **sub-basis** if every verifiable set in X is the union of finite intersections of elements of \mathcal{B} . It is a **basis** if every verifiable set in X is the union of elements of \mathcal{B} .

Definition 1.59. A topology for X is **second-countable** if it admits a countable basis.

There is a link between the basis of an experimental domain and a sub-basis of a topology. If every statement in the experimental domain can be constructed from a basis through finite conjunction and countable disjunction, then each corresponding verifiable set can be generated through intersection and union of the verifiable set corresponding to the basis. Therefore the verifiable set corresponding to the basis of the experimental domain forms a sub-basis in the topology. Since experimental domains must have a countable basis to make sure we can test any verifiable statement given enough time, the topologies we’ll be interested in must be second-countable.

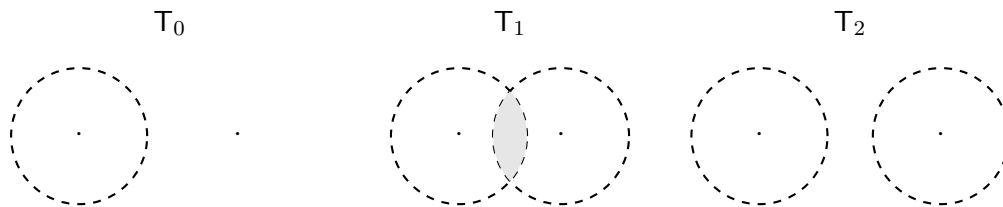
Proposition 1.60. *Let X be the set of possibilities of an experimental domain \mathcal{D} . Let $\mathcal{B} \subseteq \mathcal{D}$ be a basis for the domain, then the collection of verifiable sets $U(\mathcal{B}) \cup \{X\}$ forms a sub-basis for the natural topology of X .*

Proof. Since every verifiable statement of a domain can be generated by finite conjunction and countable disjunction from a basis $\mathcal{B} \subseteq \mathcal{D}$, its corresponding verifiable set can be generated by finite intersection and countable union from the verifiable sets $U(\mathcal{B})$ corresponding to the basis. Note that the certainty, though, is not necessarily the union of $U(\mathcal{B})$: if the domain is not complete, the residual possibility is not contained in any verifiable sets. Therefore $U(\mathcal{B}) \cup \{X\}$ can generate all verifiable sets, including the one for the certainty, and is a sub-basis. \square

Proposition 1.61. *The natural topology for the possibilities of an experimental domain is second-countable.*

Proof. Since each experimental domain admits a countable basis, its verifiable sets form a countable sub-basis for the natural topology. We can close the sub-basis over finite intersection, forming a countable basis for the topology. The natural topology is therefore second-countable as it admits a countable basis. \square

Another important property to classify topological spaces is the degree of separation of their elements: how well one can use verifiable sets to tell points and sets apart. A Kolmogorov, or T_0 , space is one in which for every pair of points there is always a verifiable set that contains one but not the other. This property is significant because it allows all points to be distinguished through verifiable sets. A Fréchet, or T_1 , space is one in which for every pair of points we can find two verifiable sets each containing only one. In a Hausdorff, or T_2 , space the two verifiable sets containing the points are disjoint. This implies the uniqueness of limits of sequences of points.



Definition 1.62. *A topology for X is **Kolmogorov** (or T_0) if for every two elements $x_1, x_2 \in X$ there exists a verifiable set $U \in \mathcal{T}_X$ containing one element but not the other. That is: either $x_1 \in U$ while $x_2 \notin U$ or $x_1 \notin U$ while $x_2 \in U$.*

Definition 1.63. *A topology for X is **Fréchet** (or T_1) if for every two elements $x_1, x_2 \in X$ there exist two, not necessarily disjoint, verifiable sets $U_1, U_2 \in \mathcal{T}_X$ each containing only one element. That is: $x_1 \in U_1$ and $x_2 \in U_2$.*

Definition 1.64. *A topology for X is **Hausdorff** (or T_2) if for every two elements $x_1, x_2 \in X$ there exist two disjoint verifiable sets $U_1, U_2 \in \mathcal{T}_X$ each containing one element. That is:*

$U_1 \cap U_2 = \emptyset$, $x_1 \in U_1$ and $x_2 \in U_2$.

Remark. Note that T_2 implies T_1 which in turn implies T_0 .

How do these properties relate to experimental domains? Consider two possibilities for a domain, for example “that is a cat” and “that is a swan”. We can always find a verifiable statement, such as “that animal has feathers”, that we can use to distinguish one possibility from the other. This means that, given two different possibilities, we can always find a verifiable set that contains one and not the other: the natural topology for any set of possibilities is always T_0 .

Now suppose two possibilities are approximately verifiable as we defined in Definition 1.38. For example, “the mass of the photon is exactly 0 eV” or “the mass of the photon is exactly 10^{-20} eV”. We can find two verifiable statements “the mass of the photon is less than 10^{-25} eV” and “the mass of the photon is more than 10^{-25} eV” each compatible with only one possibility. This means that, given two approximately verifiable possibilities, we can find two verifiable sets each containing only one possibility: if all possibilities are approximately verifiable then the natural topology is T_1 .

Now suppose two possibilities are experimentally distinguishable as we defined in Definition 1.44. Then, by 1.45, we can find two disjoint approximations. In the example before, the two verifiable statements were in fact incompatible. This means that, given two approximately verifiable possibilities, we can find two disjoint verifiable sets each containing only one possibility: if all possibilities experimentally distinguishable then the natural topology is T_2 .

Proposition 1.65. *The natural topology of a set of possibilities is Kolmogorov (or T_0).*

Proof. Let X be the set of possibilities for an experimental domain \mathcal{D} . Let $x_1, x_2 \in X$ be two distinct possibilities. Each of them can be expressed as a minterm of a basis $\mathcal{B} \subseteq \mathcal{D}$. Since the two possibilities are distinct, there must exist a verifiable statement $e \in \mathcal{B}$ that appears negated in one conjunction but not the other. That is, e is compatible with only one possibility. Since the verifiable set associated with a verifiable statement contains only the possibilities compatible with said statement, the verifiable set of e either contains x_1 or x_2 but not both. The topology is therefore Kolmogorov (or T_0). \square

Proposition 1.66. *The natural topology of a set of possibilities is Fréchet (or T_1) if and only if all possibilities are approximately verifiable.*

Proof. Suppose all possibilities in X for an experimental domain \mathcal{D}_X are approximately verifiable. Let $x_1, x_2 \in X$ be two possibilities, then we can find two sequences of verifiable statements $\{\mathfrak{s}_i^1\}_{i=1}^\infty, \{\mathfrak{s}_j^2\}_{j=1}^\infty \in \mathcal{D}_X$ such that $x_1 = \bigwedge_{i=1}^\infty \mathfrak{s}_i^1$ and $x_2 = \bigwedge_{j=1}^\infty \mathfrak{s}_j^2$. We can assume the sequences are monotone with respect to narrowness, that is $\mathfrak{s}_{i+1}^1 \preccurlyeq \mathfrak{s}_i^1$, as we can always create a monotone sequence from one that is not by taking the sequence of finite conjunction, that is $\hat{\mathfrak{s}}_k^1 = \bigwedge_{i=1}^k \mathfrak{s}_i^1$. If $x_1 \neq x_2$, then $x_1 \wedge x_2 \equiv \perp$ since different possibilities are incompatible. Therefore we must have $x_1 \wedge \mathfrak{s}_j^2 \equiv \perp$ and $\mathfrak{s}_i^1 \wedge x_2 \equiv \perp$ from some $i, j \geq 1$ or the limits would not be incompatible. In terms of verifiable sets we have $x_2 \notin U(\mathfrak{s}_i^1)$ and $x_1 \notin U(\mathfrak{s}_j^2)$. For any

two distinct possibilities we can find two verifiable sets each containing one: the natural topology is T_1 .

Conversely, suppose the natural topology T_X for the possibilities X for an experimental domain \mathcal{D}_X is T_1 . Let $x \in X$ be a possibility. Consider the collection, not necessarily countable, of all verifiable sets $\{U_i\}_{i \in I} \subset \mathsf{T}_X$ such that they contain x . Consider their intersection $U_x = \bigcap_{i \in I} U_i$. It will contain x since all U_i contain x . It will not contain anything else: since the topology is T_1 , for every other possibility \hat{x} there is always an open set U_i that does not contain it. Therefore $U_x = \{x\}$. Because the natural topology is second countable, we can find a countable basis \mathcal{B} and rewrite the arbitrary intersection into $\{x\} = \bigcap_{i=1}^{\infty} V_i$ a countable intersection of elements $V_i \in \mathcal{B}$ of the basis. Let $\{\mathbf{s}_i\}_{i=1}^{\infty}$ be the sequence of verifiable statements such that $U(\mathbf{s}_i) = V_i$ for every i . Then $x = \bigwedge_{i=1}^{\infty} \mathbf{s}_i$ which means x is approximately verifiable. □

Proposition 1.67. *The natural topology of a set of possibilities is Hausdorff (or T_2) if and only if all possibilities are pairwise experimentally distinguishable.*

Proof. Suppose all possibilities in X for an experimental domain \mathcal{D}_X are pairwise experimentally distinguishable. Then, by 1.45, given two possibilities $x_1, x_2 \in X$ we can find two verifiable statements $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D}_X$ such that $\mathbf{s}_1 \not\# \mathbf{s}_2$, $x_1 \leq \mathbf{s}_1$ and $x_2 \leq \mathbf{s}_2$. In terms of verifiable sets we have $U(\mathbf{s}_1) \cap U(\mathbf{s}_2) = \emptyset$, $x_1 \in U(\mathbf{s}_1)$ and $x_2 \in U(\mathbf{s}_2)$. The topology is T_2 .

Conversely, suppose the natural topology T_X for the possibilities X for an experimental domain \mathcal{D}_X is T_2 . Given two possibilities $x_1, x_2 \in X$ we can find two verifiable sets $U_1, U_2 \in \mathsf{T}_X$ such that $U_1 \cap U_2 = \emptyset$, $x_1 \in U_1$ and $x_2 \in U_2$. Since $U_1, U_2 \in \mathsf{T}_X$, we can find two corresponding verifiable statements $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D}_X$ such that $U(\mathbf{s}_1) = U_1$ and $U(\mathbf{s}_2) = U_2$. We have $\mathbf{s}_1 \not\# \mathbf{s}_2$, $x_1 \leq \mathbf{s}_1$ and $x_2 \leq \mathbf{s}_2$ and by 1.45 the possibilities are pairwise experimentally distinguishable. □

1.5 Sigma-algebras

In the same way that experimental domains find a natural mathematical representation as topological spaces, theoretical domains find a natural mathematical representation in σ -algebras. The main result of this section is that a theoretical domain provides a natural σ -algebra on its possibilities.

Like topologies, σ -algebras are fundamental in mathematics as they allow us to construct measures (i.e. assigning sizes to sets), limits for sequences of sets and probability spaces. It is again fitting that theoretical domains are associated to such a fundamental mathematical structure.

Let's first review what a σ -algebra is. The general idea is that we have a set X of elements which we call points, and we have a collection of subsets of X such that it is closed under complement and countable union, contains the empty set and contains the whole set X . For example, suppose $X = \{1, 2, 3\}$ then $\{\{\}, \{1\}, \{1, 2, 3\}\}$ is not a σ -algebra while $\{\{\}, \{1\}, \{2, 3\}, \{1, 2, 3\}\}$ is. The first one is missing the complement of $\{1\}$.

Definition 1.68. Let X be a set. A σ -algebra on X is a collection Σ_X of subsets of X closed under complement and countable union such that it contains X .

Note that σ -algebras are also closed under countable intersections, since these can be expressed in terms of complements and countable unions.

In the previous section we saw how each verifiable statement can be expressed as the conjunction of a set of possibilities, how the operations on statements can be expressed as operations on the verifiable sets and how all the verifiable sets form a topology. The same is true for theoretical statements, with the only difference being that we will end up with a collection of sets that is closed under complement and countable union since the theoretical domain is closed under negation and countable disjunction.

Definition 1.69. Let $\bar{\mathcal{D}}$ be a theoretical domain and X its possibilities. We define the map $A : \bar{\mathcal{D}} \rightarrow 2^X$ that for each theoretical statement $s \in \bar{\mathcal{D}}$ returns the set of possibilities compatible with it. That is, $A(s) \equiv \{x \in X \mid x \approx s\}$. We call $A(s)$ the **theoretical set** of possibilities associated with s

Proposition 1.70. Let X be the set of possibilities for a theoretical domain $\bar{\mathcal{D}}$. X has a natural σ -algebra given by the collection of all theoretical sets $\Sigma_X = A(\bar{\mathcal{D}})$.

Proof. The theoretical sets for the certainty and the impossibility correspond to the full set and empty set respectively. Formally, $A(\top) = \{x \in X \mid x \approx \top\} = X$ while $A(\perp) = \{x \in X \mid x \approx \perp\} = \emptyset$. Therefore $X, \emptyset \in A(\bar{\mathcal{D}})$ since $\top, \perp \in \bar{\mathcal{D}}$.

The complement of a theoretical set corresponds to the theoretical set of the negation and therefore it is a theoretical set. Formally, $A(s)^C = \{x \in X \mid x \not\approx s\} = \{x \in X \mid x \approx \neg s\} = A(\neg s)$.

The countable union of verifiable sets corresponds to the verifiable set of the countable disjunction and therefore it is a theoretical set. Formally, $A(s_1 \vee s_2) = \{x \in X \mid x \approx s_1 \vee s_2\} = \{x \in X \mid x \approx s_1 \text{ or } x \approx s_2\} = \{x \in X \mid x \approx s_1\} \cup \{x \in X \mid x \approx s_2\} = A(s_1) \cup A(s_2)$. This generalizes to countable disjunctions.

The collection $\Sigma_X = A(\bar{\mathcal{D}})$ is therefore a σ -algebra by definition since it satisfies all its properties. \square

There is also a special link between topologies and σ -algebras. As one may want to construct measures and probability spaces on topological spaces, there is a standard way to construct a σ -algebra from a topology. This object, called Borel algebra, is the smallest σ -algebra that contains all verifiable sets defined by the topology. The σ -algebra defined by a theoretical domain is none other than the Borel algebra of the topology defined by the corresponding experimental domain.

Definition 1.71. Let (X, τ) be a topological space. Its **Borel algebra** is the collection Σ_X of subsets of X generated by countable union, countable intersection and complement from the verifiable sets.

Proposition 1.72. The natural σ -algebra for a set of possibilities is the Borel algebra of

its natural topology.

Proof. Since the theoretical domain can be generated by a basis of the experimental domain, the natural σ -algebra can be generated by a sub-basis of the natural topology. This means that it is also generated by countable union, countable intersection and negation from the verifiable sets of the natural topology. \square

This fundamental link between experimental domains and topology on one side and theoretical domains and σ -algebra on the other is important for multiple reasons.

From a practical standpoint, it guarantees that these mathematical tools can always be used in science. Since experimental and theoretical domains are general constructs, any branch of scientific investigation can use techniques and results from topology and σ -algebras for calculations or for characterizing the domain at hand.

From a conceptual standpoint it provides a Rosetta stone, i.e a way to translate, between the mathematical concepts and the scientific ones. It gives a precise scientific meaning to the mathematical tools and everything built on top of them. Every single step in a calculation, every single argument in a proof can be given a clear, and possibly insightful, physical meaning. It grounds the abstract mathematical language in more concrete scientific objects. This in turn helps clarify the science described by common mathematical tools, unearthing possible hidden assumptions or simplifications about the physical systems being studied.

This connection explains why these mathematical tools have found such successful application in the physical sciences.

1.6 Decidable domains

We conclude this chapter by analyzing decidable domains, those for which we can experimentally test both the truth and the falsehood of all statements. For example, the domain for animal identification and the domain for the amount of money in my pocket can be considered decidable as we can typically tell experimentally whether “*this animal has whiskers*” or not, or whether “*there is more than one dollar and fifty cents in my pocket*” or not.

Decidable domains have special characteristics and are easier to study. Since we can verify the negation, any theoretical statement is also verifiable. And since all theoretical statements are verifiable, so are the possibilities. That is, we can verify that “*this animal is a cat*” and that “*there are two dollars and thirty cents in my pocket*”. As all statements can be expressed as a disjunction of possibilities, the possibilities themselves form a countable basis. For example, “*there is more than one dollars and fifty cents in my pocket*” can be expressed as the disjunction of the appropriate statements of the form “*there are x dollars and y cents in my pocket*”.

Definition 1.73. *An experimental domain \mathcal{D}_X is **decidable** if all statements in the domain are decidable. Formally, for every $s \in \mathcal{D}_X$ we have $\neg s \in \mathcal{D}_X$.*

Proposition 1.74. *Let \mathcal{D}_X be an experimental domain. The following are equivalent:*

1. *the experimental domain is decidable*
2. *the experimental domain and its theoretical domain coincide*

3. all possibilities are verifiable
4. the possibilities form a countable basis.

Proof. To prove 2 from 1, suppose \mathcal{D}_X is a decidable experimental domain. As \mathcal{D}_X is decidable, it is already closed under negation and therefore all statements in its theoretical domain $\bar{\mathcal{D}}_X$ are already in \mathcal{D}_X .

To prove 3 from 2, suppose \mathcal{D}_X coincides with its theoretical domain $\bar{\mathcal{D}}_X$. As each possibility is a theoretical statement, it is also a verifiable statement.

To prove 4 from 3, suppose the possibilities are verifiable. Note that the possibilities can generate all other statements through disjunction. To show X is countable, consider a countable basis $\mathcal{B} \subseteq \mathcal{D}_X$. Because the possibilities are verifiable statements, they can be generated from \mathcal{B} by finite conjunction and countable disjunction. Moreover, since the possibilities are the narrowest statements that are not impossible, they can be generated from \mathcal{B} using finite conjunction only. Since \mathcal{B} is countable and X is generated by \mathcal{B} through finite conjunction, X can be at most countable. Therefore X is a countable basis.

To prove 1 from 4, suppose the possibilities X form a countable basis. Then each possibility is verifiable and so is their countable union. The negation of a verifiable statement can be expressed as the countable union of possibilities, and is therefore verifiable. All statements in the experimental domain are decidable and therefore the domain is decidable. \square

As the possibilities for a decidable domain must form a countable basis, their cardinality can't be greater than countable. That is: only domains that are non-decidable can have possibilities with cardinality of the continuum. In this sense they are more constrained and simpler to study.

Mathematically, the natural topology corresponds to the discrete topology: the one for which any subsets of the possibilities is a verifiable set. That is, the topology is simply the set of all possible sets of possibilities. The cardinality of the possibility is therefore enough to determine the topology of the space, which means that one number is enough to characterize the space.

Definition 1.75. A topology \top_X on a set X is called **discrete** if it contains every subset of X .

Theorem 1.76 (Decidability is discreteness). *The natural topology of the possibilities X for a domain \mathcal{D}_X is discrete if and only if the domain is decidable.*

Proof. Suppose \mathcal{D}_X is decidable. Let $U \subseteq X$ be a subset of possibilities. The statement $s = \bigvee_{x \in U} x$ is generated from X through countable disjunction. Since \mathcal{D}_X is decidable, X is a countable basis and s is verifiable. Therefore U is a verifiable set and it is contained in the natural topology. The natural topology of X is discrete by definition.

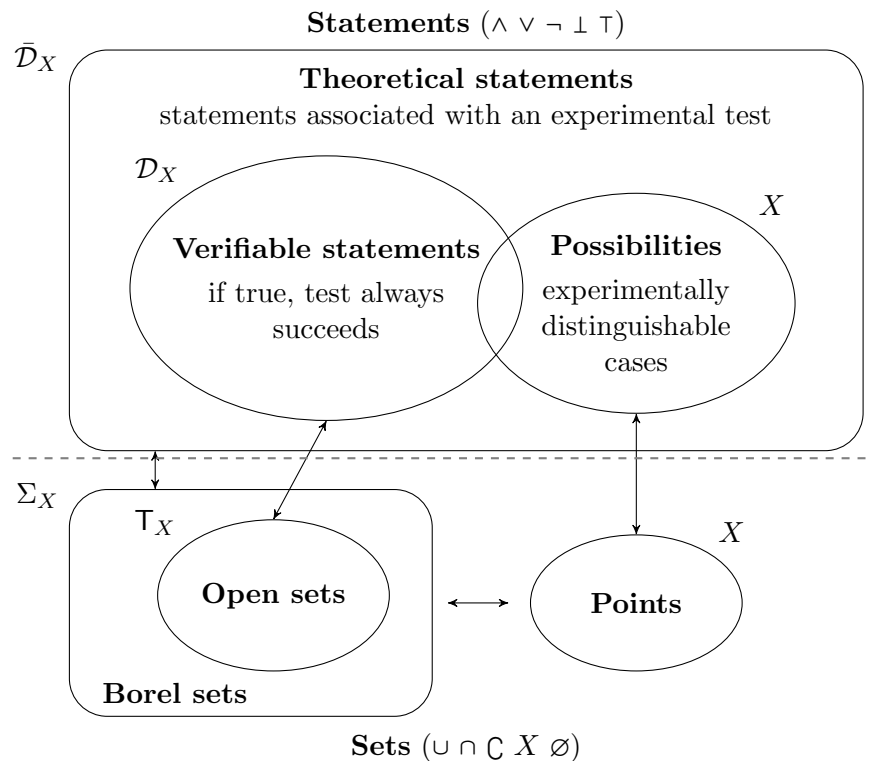
Now suppose \mathcal{D}_X is such that the natural topology for its possibilities X is discrete. Let $s = \bigvee_{x \in U} x$ be a statement. Since the topology is discrete, U is part of the topology and s is verifiable. Consider its negation $\neg s = \bigvee_{x \in U^C} x$. Since the topology is discrete, U^C is also

part of the topology and $\neg s$ is verifiable. This means s is decidable. Since every statement in \mathcal{D}_X is decidable, the domain is decidable. \square

Note, though, that discrete does not imply finite or vice-versa. The domain for extra-terrestrial life is finite but is not decidable as we cannot verify that “*there is no extra-terrestrial life*”. The domain for the amount of money in my pocket, instead, is decidable but not necessarily finite as I could potentially have an arbitrarily large amount.

1.7 Summary

In this first chapter we have laid down the foundations for our general mathematical theory of experimental science. We have seen how it is grounded in the logic of verifiable statements, which is more limited than the logic of pure statements as it has to deal with the practical constraints introduced by the termination of the tests.



We saw that we can group verifiable statements into experimental domains which must have a countable basis to allow us to test any statement within an indefinite amount of time. We saw how to construct theoretical domains to find all the theoretical statements that can be associated to an experimental test. And we saw how the possibilities are those statements that, if true, give a complete prediction for all statements in the domain.

We have seen that, because of the disjunctive normal form, each verifiable and theoretical statement is equivalent to a set of possibilities and how logic operations and relationships become set operations and relationships. As such, the experimental and theoretical domains respectively provide a natural topology and σ -algebra for the possibilities.

What we have ended up with is a conceptual framework that captures the necessary elements of scientific practice and codifies them into a symbolic representation with a well defined meaning. There is no guesswork as to what the points of our spaces are: they are the possibilities, statements that provide a complete description for the domain. We do not have to provide an “interpretation” as to what the sets of a topology represent: they correspond to verifiable statements. All the objects have a clear definition and meaning from the start, we know which ones are necessary and to what extent they are physical or idealized. This will provide a much more solid foundation to the rest of the work, which will ultimately allow us to understand much better the fundamental physical theories and the connections between them and to other areas of scientific thought.

1.8 Reference sheet

	Name	Meaning
\mathbb{B}	Boolean domain	the set of possible truth values i.e. $\mathbb{B} = \{\text{TRUE}, \text{FALSE}\}$
\mathcal{S}	logical context	a set of statements with well defined logical relationships
$s \in \mathcal{S}$	statement	an assertion with a well defined truth value
$\text{truth} : \mathcal{S} \rightarrow \mathbb{B}$	the truth function	returns whether a statement is true or not
$\mathcal{A}_{\mathcal{S}} \subseteq \mathbb{B}^{\mathcal{S}}$	possible assignments	the logically consistent truth value combinations that can be assigned to the statements
\top	certainty	a statement that can only be true (i.e. it is true in all possible assignments)
\perp	impossibility	a statement that can never be true (i.e. it is false in all possible assignments)
	contingent statement	a statement that can be either true or false depending on the possible assignment
$\neg s$	negation (logical NOT)	the statement whose truth value is always opposite
$s_1 \wedge s_2$	conjunction (logical AND)	the statement that is true only if all statements are true
$s_1 \vee s_2$	disjunction (logical OR)	the statement that is true if any of the statements is true
$s_1 \equiv s_2$	equivalence	whether each statement is a logical consequence of the other (i.e. they must have the same value in every possible assignment)
$s_1 \preceq s_2$	narrower than	whether the first statement is more specific than the second (i.e. in every possible assignment, if the first is true than the second must be also true)
$s_1 \succeq s_2$	broader than	whether the second statement is narrower than the first
$s_1 \hat{\wedge} s_2$	compatibility	whether both statement can be true at the same time (i.e. there is a possible assignment in which they are both true)
$s_1 \perp\!\!\!\perp s_2$	independence	whether both statement can be true at the same time (i.e. there is a possible assignment for each combination of their possible truths)
	minterm	a conjunction where each statement appears once, either negated or not
$s \in \mathcal{S}_v$	verifiable statement	a statement that can be validated experimentally
\mathcal{D}	experimental domain	a set of verifiable statement that can be tested in an indefinite amount of time (i.e. a set of statements closed under finite conjunction and countable disjunction, that precisely contains the certainty, the impossibility and a set of verifiable statements generated by a countable basis)
$\mathcal{B} \in \mathcal{D}$	basis	a set of verifiable statements from which all others can be constructed

	Name	Meaning
\mathcal{D}	theoretical domain	the set of all statements constructed from an experimental domain that can be associated with an experimental test
	approximately verifiable	when a statement is not verifiable but is the limit of a sequence of statements that are
X	possibilities of a domain	those statements that, if true, determine the value of all verifiable statements of a domain
\dot{x}	established possibility	a possibility for which at least a verifiable statements is true (i.e. it can be established experimentally)
\dot{x}	residual possibility	if it exists, the possibility for which all verifiable statements are false (i.e. the remaining case that cannot be established experimentally)

Chapter 2

Domain combination and relationships

We continue our investigation of the fundamental mathematical structures for experimental science by studying what happens when we have more than one experimental domain. We will define **experimental relationships** between experimental domains, which capture either causal or inference relationships between them. We will see that these correspond to continuous functions in the natural topology.

We will take two or more domains and merge all the experimental information that can be gathered through them into a **combined domain**. We will study how the set of possibilities of the combined domain depends not only on the original domains, but also on the relationships between them. These will also determine the natural topology that can vary from the product topology all the way to the disjoint union topology.

We will also show that experimental relationships, under suitable conditions, can themselves be verified experimentally by constructing the **relationship domain** for which its possibilities correspond to the possible relationships.

2.1 Dependence and equivalence between domains

The first thing we want to be able to characterize, when dealing with more than one domain, is when there exists a relationship between them. For example, consider the domains for the temperature and height of a mercury column or the domains for the temperature and density of water. How do we express, in this framework, the fact that these domains are connected?

We have two ways to define these relationships between domains. The first is in terms of inference: any measurement on the height of a mercury column is an indirect measurement on its temperature; any experimental test on the density of water is an indirect experimental test on its temperature. The second is in terms of causes: the height of the mercury column depends on its temperature; the density of water is a function of its temperature. The main result of this section is to show that these definitions are equivalent and that the dependent domain can be seen as a sub-domain of the other.

Suppose \mathcal{D}_X represents the domain for the temperature of a mercury column while \mathcal{D}_Y represents the domain for its height. Since we know that an increase in temperature makes the metal expand, we can infer the temperature of the mercury column by looking at its height. For example, if we verify that *“the height of the mercury column is between 24 and*

25 millimeters” we will be able to infer that “the temperature is between 24 and 25 Celsius”. That is, given a verifiable statement s_Y we have another verifiable statement s_X that is going to be true if and only if the first one is, that is $s_Y \equiv s_X$.

Note that the inference is between verifiable statements and not intervals. For example, the verifiable statement “the water density is between 999.8 and 999.9 kg/m³” will correspond to “the water temperature is between 0 and 0.52 Celsius” \vee “the water temperature is between 7.6 and 9.12 Celsius” as water is most dense at 4 Celsius. The disjunction of verifiable statements is still a verifiable statement so we are still inferring one verifiable statement from the other. For each verifiable statement in \mathcal{D}_Y we can find a verifiable statement in \mathcal{D}_X that is verified if and only if the first is. That is: an inference relationship is a map from \mathcal{D}_Y to \mathcal{D}_X that preserves equivalence.

Definition 2.1. An ***inference relationship*** between two experimental domains establishes that testing a verifiable statement in one means testing a verifiable statement in the other. Formally, an inference relationship between two experimental domains \mathcal{D}_X and \mathcal{D}_Y is a map $r : \mathcal{D}_Y \rightarrow \mathcal{D}_X$ such that $r(s_Y) \equiv s_Y$. In other words: it is an equivalence-preserving map between experimental domains.

An inference relationship is essentially an injection that preserves equivalence instead of identity. In terms of equality, the two statements “the height of the mercury column is between 24 and 25 millimeters” and “the temperature is between 24 and 25 Celsius” are different, but they are the same in terms of equivalence. In this sense, the dependent domain is already contained within the other domain. This means we can define domain inclusion and equivalence based on inference relationships.

Definition 2.2. An experimental domain \mathcal{D}_Y is ***dependent*** on another experimental domain \mathcal{D}_X , noted $\mathcal{D}_Y \subseteq \mathcal{D}_X$, if there exists an inference relationship $r : \mathcal{D}_Y \rightarrow \mathcal{D}_X$.

Corollary 2.3. Let \mathcal{D}_X be an experimental domain. Let \mathcal{D}_Y be a subset of statements of \mathcal{D}_X that form an experimental domain (i.e. contains impossibility, certainty and is closed under finite conjunction and countable disjunction). Then $\mathcal{D}_Y \subseteq \mathcal{D}_X$.

Proof. Let $\iota : \mathcal{D}_Y \rightarrow \mathcal{D}_X$ be the inclusion map. This is an inference relationship since $\iota(s_Y) = s_Y \equiv s_Y$ therefore \mathcal{D}_Y depends on \mathcal{D}_X . \square

Definition 2.4. Two experimental domains \mathcal{D}_X and \mathcal{D}_Y are ***equivalent*** $\mathcal{D}_X \equiv \mathcal{D}_Y$ if \mathcal{D}_X depends on \mathcal{D}_Y and vice-versa.

Corollary 2.5. Domain equivalence satisfies the following properties:

- reflexivity: $\mathcal{D} \equiv \mathcal{D}$
- symmetry: if $\mathcal{D}_X \equiv \mathcal{D}_Y$ then $\mathcal{D}_Y \equiv \mathcal{D}_X$
- transitivity: if $\mathcal{D}_X \equiv \mathcal{D}_Y$ and $\mathcal{D}_Y \equiv \mathcal{D}_Z$ then $\mathcal{D}_X \equiv \mathcal{D}_Z$

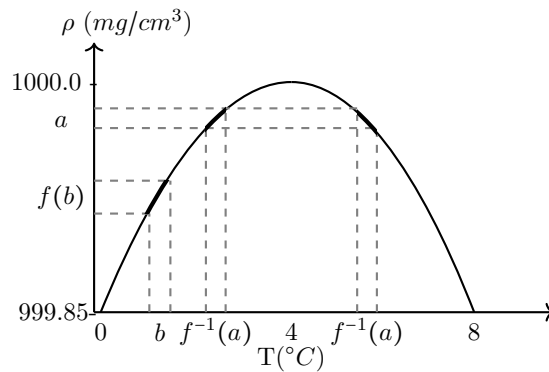
and is therefore an ***equivalence relationship***.

Proof. For reflexivity, \mathcal{D} is a subset of \mathcal{D} that is an experimental domain, therefore $\mathcal{D} \subseteq \mathcal{D}$ by 2.3. Equivalence follows by symmetry.

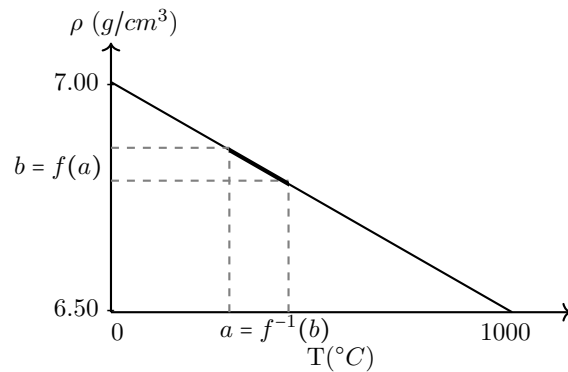
For symmetry, suppose $\mathcal{D}_X \equiv \mathcal{D}_Y$, then $\mathcal{D}_Y \subseteq \mathcal{D}_X$ and $\mathcal{D}_X \subseteq \mathcal{D}_Y$ and therefore $\mathcal{D}_Y \equiv \mathcal{D}_X$.

For transitivity, suppose $\mathcal{D}_X \equiv \mathcal{D}_Y$ and $\mathcal{D}_Y \equiv \mathcal{D}_Z$. Then we have the following inference relationships: $r_{XY} : \mathcal{D}_X \rightarrow \mathcal{D}_Y$, $r_{YX} : \mathcal{D}_Y \rightarrow \mathcal{D}_X$, $r_{YZ} : \mathcal{D}_Y \rightarrow \mathcal{D}_Z$, $r_{ZY} : \mathcal{D}_Z \rightarrow \mathcal{D}_Y$. We can define the function compositions $r_{XZ} = r_{YZ} \circ r_{XY}$ and $r_{ZX} = r_{YX} \circ r_{ZY}$. These are inference relationships since $s_X \equiv r_{XY}(s_X) \equiv r_{YZ}(r_{XY}(s_X))$ and $s_Z \equiv r_{ZY}(s_Z) \equiv r_{YX}(r_{ZY}(s_Z))$. Therefore $\mathcal{D}_X \equiv \mathcal{D}_Z$. \square

Water Temperature vs. Density



Steel Temperature vs. Density



It should be evident that we cannot impose inference relationships between any two domains: it's something that the domains allow or not. The domains for the temperature of two different mercury columns are in general not related: testing the value of one does not tell us anything about the other. The topologies of the two domains, however, are going to be the same because we'll have the same possible values and the same way to experimentally test them. Equivalence between experimental domains is a much stronger relationship than equivalence of the natural topology. It carries enough of the semantic to be able to tell what spaces are truly scientifically equivalent.

Let's continue with our example. We can re-express a relationship between domains in terms of causal relationship between the two domains. If x is the value of temperature of the mercury column (i.e. a possibility for \mathcal{D}_X) and y is the height of the mercury column (i.e. a possibility for \mathcal{D}_Y), then we can write $y = f(x)$ since the height is determined by the temperature.

Note that the direction of the causal relationship is the opposite of the inference. X causes Y and from \mathcal{D}_Y we can infer \mathcal{D}_X . Chains of events are in terms of possibilities and start with the cause and end with the effect. Chains of inferences are in terms of verifiable statements and start with the result and end with the origin.

The other directions do not work in general. Even if we know the final possibility, we may not be able to reconstruct the initial possibility: if the water density is exactly 999.9 kg/m^3 , the temperature could be either 0.52 or 7.6 Celsius because density peaks at 4 Celsius. For the same reason, a measurement of the cause is not always equivalent to a measurement of the effect: verifying that "the water temperature is between 0 and 0.52 Celsius" will mean that "the water density is between 999.8 and 999.9 kg/m^3 " but not the other way around.

Because of the peak in density, the statement about the temperature tells us more (i.e. it is narrower) than the statement about the density and therefore they are not equivalent. That is: we can learn more about the temperature by measuring it directly than indirectly through the density.

Another important consideration is that, in order to be consistent, the function $y = f(x)$ has to be continuous. The general idea is the following: if we say that we can only measure both the temperature and height of a mercury column with finite precision, we have to make sure that we can use the causal relationship for inference. Therefore a finite precision measurement of height will correspond to a finite precision measurement of temperature. This means a small change in height has to correspond to a small change in temperature: the function is continuous.

More precisely, consider the verifiable statement $s_Y = \text{“the height of the mercury column is between 24 and 25 millimeters”}$. The height of the mercury column y is within the verifiable set $U(s_Y) = (24, 25)$ millimeters. We can then infer that the temperature must be in the reverse image of the possible heights $f^{-1}(U_Y(s_Y)) = (24, 25)$ Celsius. But this means that, indirectly, we have experimentally verified that x is in $f^{-1}(U_Y(s_Y))$. And if \mathcal{D}_X is really the domain of the verifiable statements for the temperature, then it must contain one that matches *“the temperature of the mercury column is between 24 and 25 Celsius”*. In other words, $f^{-1}(U_Y(s_Y))$ must be a set in the topology of X and the function is continuous.¹

Definition 2.6. Let (X, \mathcal{T}_X) and (Y, \mathcal{T}_Y) be two topological spaces. A **continuous function** is a map $f : X \rightarrow Y$ such that given any verifiable set $U_Y \in \mathcal{T}_Y$ its reverse image $f^{-1}(U_Y) \in \mathcal{T}_X$ is a verifiable set. A **homeomorphism** is a continuous bijective map such that its inverse is also continuous.

Definition 2.7. A **causal relationship** between two experimental domains establishes that determining which possibility is true in the first domain also determines which possibility is true in the second. Formally, a causal relationship between two experimental domains \mathcal{D}_X and \mathcal{D}_Y is a function $f : X \rightarrow Y$ between the possibilities of the respective domains such that $x \leq f(x)$ for all $x \in X$.

Corollary 2.8. All causal relationships are continuous functions over the respective natural topologies.

Proof. Let $f : X \rightarrow Y$ be a causal relationship between \mathcal{D}_X and \mathcal{D}_Y . Let $y \in Y$ be a possibility for \mathcal{D}_Y . Consider $f^{-1}(y)$: this is the set of all possibilities in X that are compatible with y which is, by definition, the theoretical set of y . Therefore we have $y \equiv \bigvee_{x \in f^{-1}(y)} x$. Now consider a verifiable statement $s_Y \in \mathcal{D}_Y$ and the associated verifiable set $U_Y(s_Y) \subseteq Y$. We have $s_Y \equiv \bigvee_{y \in U_Y(s_Y)} y \equiv \bigvee_{y \in U_Y(s_Y)} \bigvee_{x \in f^{-1}(\{y\})} x \equiv \bigvee_{x \in f^{-1}(U_Y(s_Y))} x$. Given that s_Y is verifiable, $\bigvee_{x \in f^{-1}(U_Y(s_Y))} x$ is also verifiable because it is equivalent to a verifiable statement. Since $f^{-1}(U_Y(s_Y))$ is the set of possibilities compatible with that statement,

¹In topology, continuity is defined in terms of the sets in the topology and not in terms of small changes as in analysis. When using the standard topology on real numbers, the two coincide but not in general.

$f^{-1}(U_Y(\mathfrak{s}_Y))$ must be a verifiable set and therefore $f^{-1}(U_Y(\mathfrak{s}_Y)) \in \mathbb{T}_X$ must be in the natural topology of \mathcal{D}_X . Therefore f is a continuous function over the respective natural topologies. \square

Corollary 2.9. *A causal relationship between two domains is unique if it exists.*

Proof. Suppose $f_1 : X \rightarrow Y$ and $f_2 : X \rightarrow Y$ are two causal relationships. Let $x \in X$. We have $x \leq f_1(x)$ and $x \leq f_2(x)$. This means $f_1(x) \approx f_2(x)$. But these are two possibilities of the same domain, so they are either incompatible or are the same possibilities. Therefore $f_1(x) \equiv f_2(x)$ for all $x \in X$. The causal relationships are the same. \square

Theorem 2.10 (Experimental Relationship Theorem). *Inference and causal relationships are equivalent. More formally, let \mathcal{D}_X and \mathcal{D}_Y be two experimental domains. An inference relationship $r : \mathcal{D}_Y \rightarrow \mathcal{D}_X$ exists between them if and only if a causal relationship $f : X \rightarrow Y$ also exists.*

Proof. First we show that a causal relationship exists between the independent and the dependent domain. Suppose \mathcal{D}_Y depends on \mathcal{D}_X . Given that for each statement in \mathcal{D}_Y there exists an equivalent statement in \mathcal{D}_X , \mathcal{D}_Y is effectively a subset of \mathcal{D}_X . Moreover, since the theoretical domains for both experimental domains are generated by completing under negation, the theoretical domain $\bar{\mathcal{D}}_Y$ will effectively be a subset of $\bar{\mathcal{D}}_X$. This means that a possibility $x \in \bar{\mathcal{D}}_X$, if true, will determine all the truth values of all statements in $\bar{\mathcal{D}}_Y$, including its possibilities. Because one possibility of Y must be true and because all possibilities are incompatible with each other, there must be one and only one possibility $y \in \bar{\mathcal{D}}_Y$ compatible with x . Therefore we can define $f : X \rightarrow Y$ the function that given a possibility $x \in X$ returns the only possibility $y = f(x) \geq x$ that is compatible with it.

We still need to show that f is continuous. Consider a verifiable statement $\mathfrak{s}_Y \in \mathcal{D}_Y$. Let $U_Y(\mathfrak{s}_Y) \in \mathbb{T}_Y$ be its verifiable set. Since \mathcal{D}_Y depends on \mathcal{D}_X , we can find $\mathfrak{s}_X \in \mathcal{D}_X$ such that $\mathfrak{s}_X \equiv \mathfrak{s}_Y$. Let $U_X(\mathfrak{s}_X) \in \mathbb{T}_X$ be its verifiable set. This is also the set of all possibilities in X that are compatible with \mathfrak{s}_Y , which means $U_X(\mathfrak{s}_X)$ contains all the possibilities that are compatible with a possibility in $U_Y(\mathfrak{s}_Y)$. Since f returns the only possibility in Y compatible with a possibility in X , $f^{-1}(U_Y(\mathfrak{s}_Y))$ will return all the possibilities in X that are compatible with a possibility in $U_Y(\mathfrak{s}_Y)$. That means $f^{-1}(U_Y(\mathfrak{s}_Y)) = U_X(\mathfrak{s}_X)$ and that f^{-1} maps verifiable sets to verifiable sets. Therefore f is continuous.

Now we show that a causal relationship implies dependence between domains. Suppose we have a causal relationship $f : X \rightarrow Y$ between \mathcal{D}_X and \mathcal{D}_Y . Let $y \in Y$ be a possibility for \mathcal{D}_Y . Consider $f^{-1}(\{y\})$: this is the set of all possibilities in X that are compatible with y which is, by definition, the theoretical set of y . Therefore we have $y \equiv \bigvee_{x \in f^{-1}(\{y\})} x$. Now

consider a verifiable statement $\mathfrak{s}_Y \in \mathcal{D}_Y$. We have $\mathfrak{s}_Y \equiv \bigvee_{y \in U_Y(\mathfrak{s}_Y)} y \equiv \bigvee_{y \in U_Y(\mathfrak{s}_Y)} \bigvee_{x \in f^{-1}(\{y\})} x \equiv \bigvee_{x \in f^{-1}(U_Y(\mathfrak{s}_Y))} x$. Because f is continuous, the reverse image of a verifiable set is a verifiable

set. Therefore there is an $\mathfrak{s}_X \in \mathcal{D}_X$ such that $U_X(\mathfrak{s}_X) = f^{-1}(U_Y(\mathfrak{s}_Y))$. The two verifiable statements $\mathfrak{s}_X \equiv \bigvee_{x \in U_X(\mathfrak{s}_X)} x \equiv \mathfrak{s}_Y$ are equivalent. For each $\mathfrak{s}_Y \in \mathcal{D}_Y$ we can find an equivalent

$\mathfrak{s}_X \in \mathcal{D}_X$ so \mathcal{D}_Y depends on \mathcal{D}_X . \square

Corollary 2.11. *Two experimental domains \mathcal{D}_X and \mathcal{D}_Y are equivalent if and only if there exists a homeomorphism $f : X \rightarrow Y$ between the possibilities such that $x \equiv f(x)$.*

Proof. Let \mathcal{D}_X and \mathcal{D}_Y be two equivalent experimental domains. Then we can find the causal relationship $f : X \rightarrow Y$ and $g : Y \rightarrow X$. We have $x \preceq f(x) \preceq g(f(x))$. Since $g(f(x)) \in X$ is a possibility, and since x is the only possibility compatible with itself, we must have $x \equiv g(f(x))$. Therefore g is the inverse of f and it is continuous. Therefore f is a homeomorphism. We also have $x \preceq f(x) \preceq x$, therefore $f(x) \equiv x$.

Now let $f : X \rightarrow Y$ be a homeomorphism between the possibilities of \mathcal{D}_X and \mathcal{D}_Y such that $x \equiv f(x)$. Then f is a causal relationship and \mathcal{D}_Y depends on \mathcal{D}_X . Let g be the inverse of f , which is continuous since f is a homeomorphism. We have $y \equiv f(g(y)) \equiv g(y)$. Then g is also a causal relationship and \mathcal{D}_X depends on \mathcal{D}_Y . Therefore $\mathcal{D}_X \equiv \mathcal{D}_Y$. \square

Since for each inference relationship we have a causal relationship and vice-versa, we will simply use the term **experimental relationship** to describe the link between the two domains.

We should stress that causal relationships and inference relationships are defined on spaces that have, in a sense, a different status in a physical theory. Inference relationships are defined on verifiable statements, on finite precision measurements, which are the objects that are directly defined experimentally. In this sense, inferences have a higher status as they are more directly related to experimental verification. However, the map $r : \mathcal{D}_Y \rightarrow \mathcal{D}_X$ is over-complicated and redundant precisely because it maps all possible finite precision measurements from one domain to the other. Conversely, the causal relationship $f : X \rightarrow Y$ is only defined on the possibilities, on the points, therefore there is no redundancy. However, the possibilities are not verifiable statements in general and are often the product of idealizations. In this sense, causal relationships have a lower status as they are only indirectly defined by experimental verification.

The perfect correspondence between causal and inference relationships is what rescues and justifies the predominant focus on causal relationships to describe experimental relationships: since studying one is mathematically the same as studying the other, why should we use the more complicated object? Therefore, while the inference relationship is more directly physically meaningful, the causal relationship is a much more convenient object to study and characterize. That is why, in the end, all relationships will be predominantly defined by a function over the possibilities.

We now turn our attention back to domain equivalence. We have seen that two experimental domains \mathcal{D}_X and \mathcal{D}_Y are equivalent if they consist of equivalent statements: if they allow a one to one correspondence that preserves the equivalence of their statements. This implies, for example, that the possibilities are also equivalent but there is more to it.

Suppose we define some type of operation on one domain. For example, on the experimental domain for the temperature of a mercury column we define an increase by one Celsius; or on the experimental domain for the amount of gasoline in a tank we define the sum of two possible amounts. These will correspond either to operations on the domain (e.g. $f : \mathcal{D}_X \rightarrow \mathcal{D}_X$) or on its possibilities (e.g. $+$: $X \times X \rightarrow X$). But by doing so we are also defining them on all equivalent domains as well: in the end, they are made of equivalent statements. Therefore we are also defining an increase of the height of the mercury column and the sum of the monetary value of the gasoline.

This means that if we capture some physical feature using some mathematical structure on one domain, then all equivalent domains will inherit the same structure. Moreover, the causal relationship is a function that preserves that structure. If the possibilities of one domain form a vector space, then the possibilities of an equivalent domain form a vector space and the causal relationship is an invertible linear transformation. If the possibilities of one domain form a group, then the possibilities of an equivalent domain form a group and the causal relationship is an isomorphism.² As we'll see much later, this is fundamental since deterministic and reversible evolution means equivalence of the domains describing the past, present and future states. Therefore deterministic and reversible motion is not “just” a one to one map.

Theorem 2.12 (Domain Equivalence is Isomorphism). *Let $\mathcal{D}_Y \equiv \mathcal{D}_X$ be two equivalent experimental domains. Suppose \mathcal{D}_X is endowed with some mathematical structure. Then \mathcal{D}_Y is also endowed with an equivalent structure and the experimental relationship preserves said structure.*

Proof. Since an experimental domain is really defined not on the statements themselves but on their equivalence classes, the mathematical structure will also be defined on the equivalence classes. But this means that a structure defined on \mathcal{D}_X is also defined on \mathcal{D}_Y since they contain the same equivalence classes. Therefore \mathcal{D}_Y is also endowed with an equivalent structure.

The experimental relationship can either expressed as a map between possibilities $f : X \rightarrow Y$ or as a map between verifiable statements $r : \mathcal{D}_X \rightarrow \mathcal{D}_Y$. This means that the mathematical structure defined on \mathcal{D}_Y can be transported to \mathcal{D}_X using the experimental relationship. But since the mathematical structure defined on \mathcal{D}_X already contains the mathematical structure defined on \mathcal{D}_Y , the transported mathematical structure has to be the same. That is, the experimental relationship must preserve the mathematical structure defined on \mathcal{D}_X . \square

Note that the converse is not true: two domains that are endowed with the same mathematical structure are not necessarily equivalent. Consider two similarly constructed thermometers: their respective experimental domains are not equivalent since knowing something about one tells us nothing about the other. Yet, their natural topologies are equivalent because the way we can measure temperature for both is the same. One way to look at it is that the mathematical structures “forget” the full equivalence between statements and only look at a particular aspect. Topological spaces capture how the possibilities are distinguished in terms of verifiable statements. Therefore, while the domains for temperature of two different thermometers are not equivalent, their natural topologies are equivalent because the way we characterize all possible measurements is the same (i.e. the value is within a finite precision interval). Similarly, the σ -algebra only cares about what statements can be associated to experimental tests. We'll see that, in some cases, Abelian groups will capture how distributions can be composed into other distributions, that non-Abelian groups will capture how transformations can be composed into other transformations, and so on.

²Domain equivalence is an isomorphism in whatever category (e.g. topological space, group, vector space, ...) used to model the experimental domain.

2.2 Combining domains

In this section we want to understand what happens when we combine statements from different domains. For example, suppose we have the experimental domain for the pressure of an ideal gas and the experimental domain for its temperature. We can mix and match verifiable statements with conjunction and disjunction as in “the pressure is between 1 and 1.1 KPa” \wedge “the temperature is between 20 and 21 C” creating a new domain. How can we characterize this combined experimental domain?

The main result of this section is that the possibilities of the combined domain depend on how compatible the verifiable statements of the domains are. In particular, if the verifiable statements are independent across domains (e.g. the horizontal and vertical position of an object), then the possibilities of the combined domain will be the Cartesian product of those for the individual domains. On the other hand, if the verifiable statements are incompatible across domains (e.g. plant identification and animal identification), then the possibilities for the combined domain will be the disjoint union of the possibilities of the individual domains.

Suppose \mathcal{D}_X is the experimental domain generated by the two verifiable statements “the patient is dead” and “the patient is alive” and a second one \mathcal{D}_Y generated by the two verifiable statements “the patient is not in a coma” and “the patient is in a coma”. The given verifiable statements also correspond to the possibilities for the respective domains.

We can construct the combined domain $\mathcal{D}_X \times \mathcal{D}_Y$ by taking all possible disjunctions and conjunctions. What are the possibilities for the new domain? Since by 1.48 the possibilities are minterms, we have the following cases to consider:

- “the patient is alive” \wedge “the patient is in a coma”
- “the patient is alive” \wedge “the patient is not in a coma”
- “the patient is dead” \wedge “the patient is in a coma”
- “the patient is dead” \wedge “the patient is not in a coma”

The third one is impossible: the patient cannot be dead and in a coma. Therefore the combined domain has only three possibilities. The possibilities of the combined domain are, in general, the subset of all possible combinations (i.e. the Cartesian product) of the possibilities of the domains we are combining, those that are not impossible.

Definition 2.13. Let $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ be a countable set of experimental domains. The **combined experimental domain** $\mathcal{D}_X = \times_{i=1}^{\infty} \mathcal{D}_{X_i}$ is the experimental domain generated from all statements in $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ by finite conjunction and countable disjunction.

Proof. We need to show that the combined experimental domain is indeed an experimental domain. It will contain the certainty and the impossibility since any of the original experimental domains contains them. It is closed under finite conjunction and countable disjunction by construction. To show that it has a countable basis, for each $i = 1.. \infty$ let $\mathcal{B}_i \in \mathcal{D}_{X_i}$ be a countable basis for the respective domain. Consider $\mathcal{B} = \bigcup_{i=1}^{\infty} \mathcal{B}_i$. From this set we can generate any \mathcal{D}_{X_i} and therefore we can also generate all of \mathcal{D}_X . \mathcal{B} is a basis and it is countable since it is the union of a countable set of countable elements. Note that it is precisely because the basis needs to remain countable that we cannot extend the operation to an uncountable set of domains. \square

Proposition 2.14. *The possibilities for a combined domain are a subset of the Cartesian product of the possibilities for the individual domains. Formally, let $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ be a countable set of experimental domains and $\{X_i\}_{i=1}^{\infty}$ their respective possibilities. Let X be the set of possibilities for the combined domain $\mathcal{D}_X = \prod_{i=1}^{\infty} \mathcal{D}_{X_i}$. Then $X = \{x = \bigwedge_{i=1}^{\infty} x_i \mid \{x_i\}_{i=1}^{\infty} \in \prod_{i=1}^{\infty} X_i, x \neq \perp\}$.*

Proof. A possibility x of the combined domain is a minterm of a basis $\mathcal{B} \subseteq \mathcal{D}_X$. Since we can choose $\mathcal{B} = \bigcup_{i=1}^{\infty} \mathcal{B}_i$ where $\mathcal{B}_i \subseteq \mathcal{D}_{X_i}$ is a countable basis for each domain, x is the conjunction $x \equiv \bigwedge_{i=1}^{\infty} x_i$ of minterms x_i of \mathcal{B}_i . Since x is a possibility, it is not impossible and therefore none of the x_i can be impossible. Since each x_i is a minterm of the respective basis \mathcal{B}_i that is not impossible, it is a possibility by 1.48. Therefore a possibility x of the combined domain is the conjunction of the possibilities x_i of the original domains that is not impossible. \square

Proposition 2.15. *Let $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ be a countable set of incomplete experimental domains and \mathcal{D}_X their combined experimental domain. Let $\{\hat{x}_i\}_{i=1}^{\infty}$ be the residual possibility of each respective domain. Let $\hat{x} = \bigwedge_{i=1}^{\infty} \hat{x}_i$. The combined domain \mathcal{D}_X is incomplete if and only if $\hat{x} \neq \perp$ in which case \hat{x} is the residual possibility.*

Proof. For each domain, the residual possibility is the conjunction of the negation of its basis. Therefore we have $\hat{x} \equiv \bigwedge_{e \in \mathcal{B}} \neg e \equiv \bigwedge_{i=1}^{\infty} \bigwedge_{e \in \mathcal{B}_i} \neg e \equiv \bigwedge_{i=1}^{\infty} \hat{x}_i$. If $\hat{x} \neq \perp$ then it is the residual possibility. \square

Corollary 2.16. *The residual possibility of the combined domain is narrower than the ones of the original domains. If one of the domains is complete then the combined domain is also complete.*

Proof. Let \hat{x} be the residual possibility for the combined domain and \hat{x}_j the one for one of the original domains. We have $\hat{x}_j \wedge \hat{x} \equiv \hat{x}_j \wedge \bigwedge_{i=1}^{\infty} \hat{x}_i \equiv \bigwedge_{i=1}^{\infty} \hat{x}_i \equiv \hat{x}$. Therefore, by 1.22 $\hat{x} \leq \hat{x}_j$.

Suppose that one of the original domains \mathcal{D}_{X_j} is complete. Then the conjunction of the negation of its basis \hat{x}_j is an impossibility. This means the conjunction of the negation of the basis of the combined domain \hat{x} is also an impossibility and the combined domain is complete. \square

Independent domains

A special case is when combining two independent domains. For example, the domain for the pressure and the domain for the volume of an ideal gas are independent because a measurement on one tells us nothing about the other. Similarly, the domain for the shape and the domain for the color of an object are independent. In these cases, we can have any combination of possibilities: any pressure with any volume or any color with any shape.

In terms of topology, the possibilities of the combined domain are the Cartesian product of

the possibilities of the original domains and their natural topology is the product topology.³

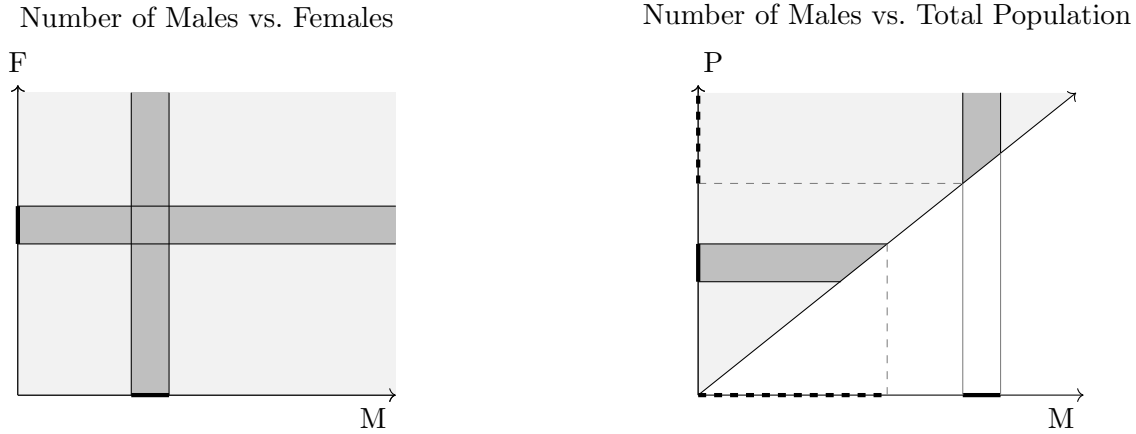


Figure 2.1: Domain independence and projections. Given a population, the number of male and female members form independent domains: knowing something about one value tells us nothing about the other. Pictorially, a constraint on one side gives us a vertical or horizontal band, which projected on the other axis, gives us the full axis (i.e. the certainty). On the other hand, the total population and the number of males are not independent domains: given the total population, the number of males cannot exceed that number; given the number of males, the total population cannot be lower than that number. Pictorially, we see that the combined domain does not span the whole plane. A constraint on one domain gives us, when projected, a constraint on the other domain.

Definition 2.17. *The experimental domains of a countable set $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ are **independent** if taking one verifiable statement $s_i \in \mathcal{D}_{X_i}$ from each domain always gives an independent set of statements.*

Proposition 2.18. *Let $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ be a countable set of independent experimental domains and X_i their respective possibilities. The set of possibilities X of the combined experimental domain $\prod_{i=1}^{\infty} \mathcal{D}_{X_i}$ consists of all the possible conjunctions of the possibilities of each domain.*

That is: $X = \left\{ \bigwedge_{i=1}^{\infty} x_i \mid x_i \in X_i \right\}$. Notationally, we write $\mathcal{D}_X = \mathcal{D}_{\prod_{i=1}^{\infty} X_i}$.

Proof. A possibility $x \equiv \bigwedge_{i=1}^{\infty} x_i$ for the combined domain is the conjunction of possibilities of each individual domain by 2.14. Since the domains are independent and since possibilities are neither certainties nor impossibilities, by 1.21 there exists an assignment $a \in \mathcal{A}_S$ such

³Note that the topology is quite naturally the product topology and not the box topology. The box topology would require countable conjunction and is therefore discarded. The fact that the correct topology is the one most natural to define confirms again the appropriateness of our framework.

that $a(x_i) = \text{TRUE}$ for all $i = 1..∞$ and therefore $a(x) = \text{TRUE}$. That is, each conjunction $x = \bigwedge_{i=1}^{\infty} x_i$ is not impossible and therefore is a possibility. \square

Definition 2.19. Let $\{(X_i, \mathbb{T}_i)\}_{i=1}^{\infty}$ be a countable set of topological spaces. Let $X = \prod_{i=1}^{\infty} X_i$ be the Cartesian product of the points. Let \mathcal{B} be the collection of sets of the form $\prod_{i=1}^{\infty} U_i$, with $U_i \in \mathbb{T}_i$ and $U_i \neq X_i$ only finitely many times. The topology generated by \mathcal{B} is called the **product topology**.

Proposition 2.20. Let $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ be a countable set of independent experimental domains. The natural topology for the possibilities of the combined experimental domain $\mathcal{D}_{\prod_{i=1}^{\infty} X_i}$ is the product topology of the natural topology for the possibilities of each domain.

Proof. Let $U_i : \mathcal{D}_{X_i} \rightarrow \mathbb{T}_{X_i}$ be the map from a verifiable statement of a domain to its verifiable set in the respective topology. Let $U : \mathcal{D}_X \rightarrow \mathbb{T}_X$ be the same map for the combined domain. Let $s_i \in \mathcal{D}_{X_i}$ be a verifiable statement from a particular domain and $U_i(s_i)$ its verifiable set in that domain. Since we also have $s_i \in \mathcal{D}_X$, the statement is also associated with the verifiable set $U(s_i)$ in the combined domain. Because the domains are independent, every possibility in $U_i(s_i)$ is compatible with any possibility $x_j \in X_j$ for all $j \neq i$. This means that $U(s_i) = X_1 \times \dots \times X_{i-1} \times U_i(s_i) \times X_{i+1} \times \dots$. Given that a verifiable statement in the combined domain can be generated using finite conjunction and countable disjunction from the verifiable statements of the independent domains, the topology of the combined space can be generated by all sets of the form $\prod_{i=1}^{\infty} U_i$, with $U_i \in \mathbb{T}_i$ and $U_i \neq X_i$ only once. Using finite conjunction, this includes those sets where $U_i \neq X_i$ finitely many times. The natural topology of the combined domain is the product topology by definition. \square

Dependent domains

Another special case is combining a domain \mathcal{D}_X with another $\mathcal{D}_Y \subseteq \mathcal{D}_X$ that is dependent on it. For example, combining the domain for the temperature of a mercury column with the domain for its height. Since the height can be determined by the temperature, no new possibilities are added. The combined domain is equivalent to the independent domain \mathcal{D}_X since all the verifiable statements in \mathcal{D}_Y have equivalents in it.

Proposition 2.21. Let \mathcal{D}_X and \mathcal{D}_Y be two experimental domains such that $\mathcal{D}_Y \subseteq \mathcal{D}_X$ depends on the first. Then $\mathcal{D}_X \times \mathcal{D}_Y \equiv \mathcal{D}_X$.

Proof. Since \mathcal{D}_Y is dependent on \mathcal{D}_X , any statement in \mathcal{D}_Y is equivalent to one in \mathcal{D}_X . Therefore no statement can be generated from them that is not equivalent to one already contained in \mathcal{D}_X . Therefore $\mathcal{D}_X \times \mathcal{D}_Y$ is equivalent to \mathcal{D}_X . \square

Corollary 2.22. Let \mathcal{D}_X and \mathcal{D}_Y be two experimental domains such that $\mathcal{D}_Y \subseteq \mathcal{D}_X$ depends on the first. The possibilities of $\mathcal{D}_X \times \mathcal{D}_Y$ are the possibilities of \mathcal{D}_X .

Proof. The possibilities of $\mathcal{D}_X \times \mathcal{D}_Y$ are the possibilities of \mathcal{D}_X since they are equivalent domains. To be consistent with 2.14, we additionally show that they are also the subset of the Cartesian product. Let $f : X \rightarrow Y$ be the causal relationship between the domains, let x and y be two possibilities of the respective domains. We have $x \wedge y \neq \perp$ if and only if $y = f(x)$, so the possibilities of $\mathcal{D}_X \times \mathcal{D}_Y$ are $x \wedge f(x)$ for all $x \in X$. We also have $x \wedge f(x) \equiv x$ since $x \preceq f(x)$. \square

Incompatible domains

The last special case we consider is when the domains are incompatible, that is all verifiable statements of one are incompatible with the verifiable statements of the others. This is one case where the residual possibility behaves differently from all the others.⁴ Suppose \mathcal{D}_X is the domain to classify a particular specimen as an animal and \mathcal{D}_Y is the domain to classify it as a plant. If we take a verifiable statement from the first, such as “that specimen has fur”, then it will be incompatible with a verifiable statement from the other, such as “that specimen has lobed leaves”. The only way we can combine the possibilities is to take an established possibility of one (e.g. “this specimen is a cat”) and combine it with the residual possibility of the other (e.g. “this specimen is not a plant”). In other words, the combined possibilities are the union of the possibilities of the two domains (e.g. all possible plants and all possible animals).

In terms of the topology, the established possibilities of the combined domain are the disjoint union of the established possibilities of the original domains and their natural topology is the disjoint union topology (or co-product topology).

Definition 2.23. *Two experimental domains \mathcal{D}_X and \mathcal{D}_Y are **incompatible** if all verifiable statements in one are incompatible with all verifiable statements of the other. Formally, $s_X \not\# s_Y$ for each pair of verifiable statements $s_X \in \mathcal{D}_X$ and $s_Y \in \mathcal{D}_Y$.*

Corollary 2.24. *Let \mathcal{D}_X and \mathcal{D}_Y be two incompatible experimental domains. Then they must be incomplete and admit a residual possibility.*

Proof. Let \mathcal{B}_X and \mathcal{B}_Y be a countable basis for the respective domain. Since we have $e_X \not\# e_Y$ for all choices of $e_X \in \mathcal{B}_X$ and $e_Y \in \mathcal{B}_Y$, we also must have $\bigvee_{e_X \in \mathcal{B}_X} e_X \not\# \bigvee_{e_Y \in \mathcal{B}_Y} e_Y$. Therefore $\bigvee_{e_X \in \mathcal{B}_X} e_X \neq \top$. Which means the residual possibility $\dot{x} = \bigwedge_{e_X \in \mathcal{B}_X} \neg e_X = \neg \bigvee_{e_X \in \mathcal{B}_X} e_X \neq \perp$. Therefore \mathcal{D}_X is not complete and, by symmetry, neither is \mathcal{D}_Y . \square

Proposition 2.25. *Let $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ be a countable set of experimental domains pair-wise incompatible and \dot{X}_i their respective established possibilities. The set of established possibilities \dot{X} of the combined experimental domain $\bigotimes_{i=1}^{\infty} \mathcal{D}_{X_i}$ consists of the disjoint union of the possibilities of each domain. That is: $\dot{X} = \prod_{i=1}^{\infty} \dot{X}_i = \bigcup_{i=1}^{\infty} \dot{X}_i$ as $\dot{X}_i \cap \dot{X}_j = \emptyset$ for all $i, j \geq 1$ and $i \neq j$. Notationally, we write $\mathcal{D}_X = \mathcal{D}_{\prod_{i=1}^{\infty} X_i}$.*

⁴In fact, this is what prompted us to introduce the residual possibility.

Proof. Consider two incompatible domains \mathcal{D}_X and \mathcal{D}_Y . Let $\hat{x} \in X$ and $\hat{y} \in Y$ be two established possibilities. Then they both correspond to a minterm of the respective basis where at least one element is taken without negation. This also means that their conjunction will include the conjunction of one element of each of the basis. Since the elements of one basis are incompatible with the elements of the other, we have $\hat{x} \not\# \hat{y}$.

Now let $\hat{x} \in X$ be an established possibility, $\hat{y} \in Y$ the residual and $\dot{Y} = Y \setminus \{\hat{y}\}$ the established possibilities. We have $\hat{x} \equiv \hat{x} \wedge \top \equiv \hat{x} \wedge \bigvee_{y \in Y} y \equiv \hat{x} \wedge \left(\bigvee_{\hat{y} \in \dot{Y}} \hat{y} \vee \hat{y} \right) \equiv \bigvee_{\hat{y} \in \dot{Y}} (\hat{x} \wedge \hat{y}) \vee (\hat{x} \wedge \hat{y}) \equiv \perp \vee (\hat{x} \wedge \hat{y}) \equiv \hat{x} \wedge \hat{y}$. By symmetry, $\hat{y} \equiv \hat{x} \wedge \hat{y}$.

To conclude, let $\hat{x} \in X$ and $\hat{y} \in Y$ be the two residual possibilities. The conjunction $\hat{x} \wedge \hat{y}$ corresponds to a minterm where all the elements of the basis are negated. Therefore $\hat{x} \wedge \hat{y}$ is the residual possibility of the combined domain, if it is not impossible.

Generalizing to a countable set of incompatible domains, the conjunction of all the residual possibilities $\hat{x} = \bigwedge_{i=1}^{\infty} \hat{x}_i$ is the residual possibility of the combined domain, if it is not impossible. The only other conjunctions of the form $x = \bigwedge_{i=1}^{\infty} x_i$ with $x_i \in X_i$ for all $i \geq 1$ that are not impossible are those where only one element is not a residual possibility. Those correspond to the established possibilities. But each of those conjunctions will be equivalent to the only element that is not a residual possibility. Therefore each established possibility of the combined domain is equivalent to an established possibility of one of the original domains: $\dot{X} = \bigcup_{i=1}^{\infty} \dot{X}_i$. Given that the established possibilities of two incompatible domains are incompatible and therefore different, we have $\dot{X}_i \cap \dot{X}_j = \emptyset$ for all $i, j \geq 1$ and $i \neq j$. The established possibilities of the combined domains are the disjoint union of the established possibilities of the individual domains. \square

Definition 2.26. Let $\{(X_i, \mathbb{T}_i)\}_{i=1}^{\infty}$ be a countable set of topological spaces. Let $X = \bigsqcup_{i=1}^{\infty} X_i$ be the disjoint union of the points. The **disjoint union topology** \mathbb{T} is the topology for which $U \in \mathbb{T}$ if and only if $U \cap X_i \in \mathbb{T}_i$ for all $i \geq 1$.

Proposition 2.27. The disjoint union topology is generated by closing the topologies of the initial spaces under disjoint union.

Proof. First we show that all disjoint unions of verifiable sets are part of the disjoint union topology. Let $\{(X_i, \mathbb{T}_i)\}_{i=1}^{\infty}$ be a countable set of topological spaces and (X, \mathbb{T}) their disjoint union with the disjoint union topology. Any disjoint union of verifiable sets can be put in the form $U = \bigsqcup_{i=1}^{\infty} U_i$ with $U_i \in \mathbb{T}_i$ for all $i \geq 1$. We have $U \cap X_i = \bigsqcup_{j=1}^{\infty} U_j \cap X_i = U_i$ for all $i \geq 1$. Therefore $U \in \mathbb{T}$ by definition.

Now we show that any set in the disjoint union topology is the disjoint union of verifiable sets of the individual topologies. Let $U \in \mathbb{T}$. Since $U \subseteq X$ and X is the disjoint union for all X_i , we can write $U = \bigsqcup_{i=1}^{\infty} U_i$. As before $U \cap X_i = U_i$ for all $i \geq 1$ and, since $U \cap X_i \in \mathbb{T}_i$ for all $i \geq 1$ by definition of the disjoint union topology, $U_i \in \mathbb{T}_i$ for all $i \geq 1$. Therefore any $U \in \mathbb{T}$ is the disjoint union of verifiable sets.

This also shows that the disjoint union topology is a topology since it is closed under arbitrary union by definition, it is closed under finite intersection within each topology and

it is closed under finite intersection across topologies since their intersection is always the empty set. \square

Proposition 2.28. *Let $\{\mathcal{D}_{X_i}\}_{i=1}^{\infty}$ be a countable set of pair-wise incompatible domains. The natural topology for the established possibilities of the combined experimental domain $\mathcal{D}_{\coprod_{i=1}^{\infty} X_i}$ is the disjoint union topology of the natural topology for the established possibilities of each domain.*

Proof. Given that the domains are incompatible, they are also incomplete. The only statement in each domain \mathcal{D}_{X_i} compatible with the respective residual possibility \hat{x}_i will be the certainty \top . This means the natural topology restricted to the established possibilities contains all the verifiable sets associated to the impossibilities and to all verifiable statements. This will also be true for the combined domain $\mathcal{D}_{\coprod_{i=1}^{\infty} X_i}$, if it is incomplete.

If it is complete, then the established possibilities are the full possibilities and their natural topology coincides. Note that all verifiable statements in the combined domain can be generated from the verifiable statements of the individual domains. Also note that the conjunction between the different domains is an impossibility and therefore does not yield a new statement. Therefore all the statements whose verifiable sets form the topology on the established possibilities on the combined domain can be generated by the disjunction of all the statements whose verifiable sets form the topology on the established possibilities on the individual domains. Therefore the topology of the combined space is simply closing the individual topologies under the disjoint union. The topology of the combined space is therefore the disjoint union topology by 2.27 \square

For topologies, as well as for other mathematical structures, we can choose between two types of products. If we have two one dimensional euclidean spaces we can decide whether to take their product (i.e. the plane) or their co-product (i.e. the disjoint union of the two lines). For experimental domains we do not choose: it is what it is. There is no way to combine the two experimental domains for temperature and height of the same mercury column and get the Cartesian product of their possibilities. Though combining two independent domains mimics the categorical product and combining two incompatible domains mimics the categorical co-product, it is the semantic (and ultimately physical) relationship between them that decides which product we have. This is another case where the mathematical structures “forget” the full equivalence. The topology only captures how the temperature and height can be measured, and not whether they are independent, dependent or incompatible. That information lies outside of the topology and therefore needs to be added by choosing the correct combination.

2.3 Experimental domain for experimental relationships

Now that we have seen how to describe relationships between domains, we should ask: are experimental relationships themselves something we can experimentally verify? We may know that there is a relationship between the temperature of a mercury column and its height, but how can we confirm experimentally which one it is?

The main result of this section is to show that, given two related experimental domains, we can always mathematically construct from them another experimental domain for which the possibilities are continuous functions between the possibilities of the original domains. This

means that, since we can recursively create relationship domains about relationship domains, the universe of discourse of our mathematical framework is closed. Yet, the availability of the experimental tests (i.e. whether the statements we construct are actually verifiable) is not guaranteed.

First of all we have to clarify within our framework what it means to experimentally verify a relationship. Suppose \mathcal{D}_X is the domain for the state of a light switch, up or down being the two possibilities, and \mathcal{D}_Y is the domain for the state of the light, which can be on or off. Suppose we could have three cases: the one where the switch is wired correctly, up corresponding to on, the one where the switch is wired incorrectly, up corresponding to off, and the one where the switch is not actually wired to the light, the two domains are independent. Each of these cases is represented by a different set of logical relationships, a different set of possible assignments.

X		Y	
up	down	on	off
T	F	T	F
F	T	F	T

Table 2.1: Case 1: switch wired correctly to the light

X		Y	
up	down	on	off
T	F	F	T
F	T	T	F

Table 2.2: Case 2: switch wired incorrectly to the light

X		Y	
up	down	on	off
T	F	T	F
T	F	F	T
F	T	T	F
F	T	F	T

Table 2.3: Case 3: switch not wired to the light

Each case, then, is a different model: a different logical context. To experimentally verify which model is the right one means verifying statements of the type “*it is possible for the switch to be up while the light is on*”. This does not correspond to “*the switch is up*” \wedge “*the light is on*” but to “*the switch is up*” \approx “*the light is on*”. These are statements about the table and they cannot be columns of the table itself. What we need to create is a metacontext, in which each line corresponds to one choice of model, to one context.

Each line will tell us what statements are compatible and given the truth values of the line we can reconstruct the model. The verifiable statements that distinguish between different models, therefore, live in this metacontext and so will the experimental domain for the relationships.

	up $\hat{=}$ on	down $\hat{=}$ on	up $\hat{=}$ off	down $\hat{=}$ off
Case 1	T	F	F	T
Case 2	F	T	T	F
Case 3	T	T	T	T

Table 2.4: Possible assignments for the metacontext

Definition 2.29. Let \mathcal{S}_1 and \mathcal{S}_2 be two logical contexts. Let $S_1 \subseteq \mathcal{S}_1$ and $S_2 \subseteq \mathcal{S}_2$ be sets of statements. We say $f : S_1 \rightarrow S_2$ is a **logic homomorphism** if it preserves the logical structure. That is:

- possible assignments on S_2 correspond to possible assignments on S_1 : if $a_2 \in \mathcal{A}_{S_2}$ then there exists $a_1 \in \mathcal{A}_{S_1}$ such that $a_1(\mathfrak{s}) = a_2(f(\mathfrak{s}))$ for all $\mathfrak{s} \in S_1$
- a verifiable statement of S_1 is mapped to a verifiable statement in S_2

Additionally, if there exists a logic homomorphism $g : S_2 \rightarrow S_1$ such that $f \circ g = \text{Id}_{S_1}$ and $g \circ f = \text{Id}_{S_2}$ then f is a **logic isomorphism** and g is the **inverse** of f . Two logic isomorphisms $f : S_1 \rightarrow S_2$ and $g : S_1 \rightarrow S_2$ are equivalent if $f(\mathfrak{s}) \equiv g(\mathfrak{s})$ for all $\mathfrak{s} \in S_1$.

Corollary 2.30. Let $h : S_1 \rightarrow S_2$ be a logic homomorphism, then for every $\hat{\mathfrak{s}} \in S_1$, $\hat{S} \subseteq S_1$ and $f_{\mathbb{B}} : \mathbb{B}^{\hat{S}} \rightarrow \mathbb{B}$ such that $a_1(\hat{\mathfrak{s}}) = f_{\mathbb{B}}(\{a_1(\mathfrak{s})\}_{\mathfrak{s} \in \hat{S}})$ for all $a_1 \in \mathcal{A}_{S_1}$, then $a_2(h(\hat{\mathfrak{s}})) = f_{\mathbb{B}}(\{a_2(h(\mathfrak{s}))\}_{\mathfrak{s} \in \hat{S}})$ for all $a_2 \in \mathcal{A}_{S_2}$. In particular:

- $h(\neg \mathfrak{s}) \equiv \neg h(\mathfrak{s})$ for all $\mathfrak{s} \in S_1$
- $h(\bigwedge_{\mathfrak{s} \in S} \mathfrak{s}) \equiv \bigwedge_{\mathfrak{s} \in S} h(\mathfrak{s})$ for all $S \subseteq S_1$
- $h(\bigvee_{\mathfrak{s} \in S} \mathfrak{s}) \equiv \bigvee_{\mathfrak{s} \in S} h(\mathfrak{s})$ for all $S \subseteq S_1$

Proof. If $a_1(\hat{\mathfrak{s}}) = f_{\mathbb{B}}(\{a_1(\mathfrak{s})\}_{\mathfrak{s} \in \hat{S}})$ for all $a_1 \in \mathcal{A}_{S_1}$ then, by the first property of logic homomorphisms, there is no $a_2 \in \mathcal{A}_{S_2}$ such that $a_2(h(\hat{\mathfrak{s}})) \neq f_{\mathbb{B}}(\{a_2(h(\mathfrak{s}))\}_{\mathfrak{s} \in \hat{S}})$. Therefore $a_2(h(\hat{\mathfrak{s}})) = f_{\mathbb{B}}(\{a_2(h(\mathfrak{s}))\}_{\mathfrak{s} \in \hat{S}})$ for all $a_2 \in \mathcal{A}_{S_2}$.

In particular, since the relationship holds for any arbitrary $f_{\mathbb{B}}$, it will hold for negation, arbitrary conjunction and arbitrary disjunction. \square

Proposition 2.31. Let \mathcal{S}_1 and \mathcal{S}_2 be two logical contexts. Let $S_1 \subseteq \mathcal{S}_1$ and $S_2 \subseteq \mathcal{S}_2$ be sets of statements and $f : S_1 \rightarrow S_2$ a logic homomorphism. Let $\bar{S}_1 \subseteq \mathcal{S}_1$ and $\bar{S}_2 \subseteq \mathcal{S}_2$ be the closure of S_1 and S_2 respectively under arbitrary conjunction, arbitrary disjunction and negation. Then if there exists a logic homomorphism $\bar{f} : \bar{S}_1 \rightarrow \bar{S}_2$ such that $\bar{f}(\mathfrak{s}_1) \equiv f(\mathfrak{s}_1)$ for all $\mathfrak{s}_1 \in S_1$ it is unique.

Proof. Let $\bar{\mathfrak{s}}_1 \in \bar{S}_1$ but $\bar{\mathfrak{s}}_1 \notin S_1$. Then $\bar{\mathfrak{s}}_1$ depends on S_1 through some $g_{\mathbb{B}} : \mathbb{B}^{S_1} \rightarrow \mathbb{B}$. Let $\bar{f}_1 : \bar{S}_1 \rightarrow \bar{S}_2$ and $\bar{f}_2 : \bar{S}_1 \rightarrow \bar{S}_2$ be two logic homomorphisms such that $\bar{f}_1(\mathfrak{s}_1) \equiv \bar{f}_2(\mathfrak{s}_1) \equiv f(\mathfrak{s}_1)$ for all $\mathfrak{s}_1 \in S_1$. Then by 2.30 both $\bar{f}_1(\bar{\mathfrak{s}}_1)$ and $\bar{f}_2(\bar{\mathfrak{s}}_1)$ depend on S_2 through $g_{\mathbb{B}}$. This means that in every possible assignment $a \in \mathcal{A}_{\bar{S}_2}$ we have $a(\bar{f}_1(\bar{\mathfrak{s}}_1)) = a(\bar{f}_2(\bar{\mathfrak{s}}_1))$ and therefore $\bar{f}_1(\bar{\mathfrak{s}}_1) \equiv \bar{f}_2(\bar{\mathfrak{s}}_1)$. Then \bar{f}_1 and \bar{f}_2 are equivalent. \square

Definition 2.32. Let $\mathcal{D}_X, \mathcal{D}_Y \subseteq \mathcal{S}$ be two experimental domains. The **relationship meta-context** $\mathcal{S}_{C(X,Y)}$ between \mathcal{D}_X and \mathcal{D}_Y over $\{\mathcal{S}_i\}_{i \in I}$ is defined by following construction. Let I be an index set. Let $\{\mathcal{S}_i\}_{i \in I}$ be an indexed set of logical contexts, let $\{\mathcal{D}_{X_i}\}_{i \in I}$ and $\{\mathcal{D}_{Y_i}\}_{i \in I}$ be two indexed sets of experimental domains such that $\mathcal{D}_{X_i}, \mathcal{D}_{Y_i} \subseteq \mathcal{S}_i$ for all $i \in I$. Let $\{f_i\}_{i \in I}$ be a set of logic isomorphisms such that $f_i : \mathcal{D}_X \rightarrow \mathcal{D}_{X_i}$. Let $\{g_i\}_{i \in I}$ be a set of logic isomorphisms such that $g_i : \mathcal{D}_Y \rightarrow \mathcal{D}_{Y_i}$.

Let S be the set of ordered pairs of \mathcal{D}_X and \mathcal{D}_Y and “ $\mathbf{s}_x \approx \mathbf{s}_y$ ” be the notation of an element of S for some $\mathbf{s}_x \in \mathcal{D}_X$ and $\mathbf{s}_y \in \mathcal{D}_Y$. For each $i \in I$, let $a_i \in \mathbb{B}^S$ be such that $a_i(\text{“}\mathbf{s}_x \approx \mathbf{s}_y\text{”}) = \text{TRUE}$ if and only if $f_i(\mathbf{s}_x) \approx g_i(\mathbf{s}_y)$. Define $\mathcal{A}_S = \{a_i \mid \forall i \in I\}$. Assume $a_i \neq a_j$ if $i \neq j$; if not, given a subset of I that leads to the same possible assignment, remove all elements but one. Pick $t \in I$ and define truth = a_t .

Given an arbitrary function $f_{\mathbb{B}} : \mathbb{B}^S \rightarrow \mathbb{B}$, we can extend the set S with a new element for which the possible assignments are calculated according to the function. The relationship context $\mathcal{S}_{C(X,Y)}$ is the closure of S under all possible functions $f_{\mathbb{B}} \in \mathbb{B}^S$.

Let $S_v \subseteq S$ be set of statements “ $\mathbf{s}_x \leq \mathbf{s}_y$ ” = “ $\mathbf{s}_x \not\approx \neg \mathbf{s}_y$ ” = “ $\neg(\mathbf{s}_x \approx \neg \mathbf{s}_y)$ ” such that $\mathbf{s}_x \in \mathcal{D}_X$ and $\mathbf{s}_y \in \mathcal{D}_Y$ are verifiable. The set of verifiable statements $S_v \subseteq \mathcal{S}$ for the relationship context is the set of statements generated by S_v using finite conjunction and countable disjunctions.

Justification. By constructions, the relationship context satisfies the axioms for a logical context. We have a set of elements and a set of possible assignments of which one is the truth, this satisfies axioms 1.2 and 1.4. Some of the statements are verifiable and satisfy 1.27. Closure for axioms 1.9, 1.31 and 1.32 are satisfied by constructions. We need to clarify why and what cases this construct makes physical sense.

The set $\{\mathcal{S}_i\}_{i \in I}$ consists of different contexts, each containing a copy of the domains \mathcal{D}_X and \mathcal{D}_Y with a different relationship between them. For example, if \mathcal{D}_X is the domain for a light switch, up or down being the possibilities, and \mathcal{D}_Y is the domain for the light, being on or off, in one context the two domain may have a causal relationship (i.e. the switch is wired to the light), in another they may have a different causal relationship (i.e. the switch is wired in the wrong way), and in yet another they may be independent (i.e. switch is not wired with the light). The logic isomorphisms $\{f_i\}_{i \in I}$ and $\{g_i\}_{i \in I}$ allow us relate the statements that represent the same assertion in the different contexts.

Note that while mathematically we can always construct the relationship metacontext, it may not exist in reality. Specifically, there is nothing that guarantees that statements of the form “ $\mathbf{s}_x \leq \mathbf{s}_y$ ” are actually verifiable. \square

Now that we have the right context, we have to understand what is the right basis to generate the experimental domain. Suppose we have a way to verify experimentally statements of the type “the temperature of the mercury column is between 24 and 25 Celsius” \leq “the height of the mercury column is between 24 and 25 millimeters”. That is, we can verify that whenever \mathbf{s}_X = “the temperature of the mercury column is between 24 and 25 Celsius” is verified, then \mathbf{s}_Y = “the height of the mercury column is between 24 and 25 millimeters” is verified. In that case, we can explore the connection between the two domains within different ranges at ever increasing precision. This means we can narrow the range of possible functions in the same way that we can narrow the range of possible values for a quantity. These types of statements, then, form an experimental domain where each possibility corresponds to a possible continuous function between the two initial experimental domains.

Definition 2.33. Let $\mathcal{S}_{C(X,Y)}$ be the relationship metacontext between \mathcal{D}_X and \mathcal{D}_Y over $\{\mathcal{S}_i\}_{i \in I}$. Let $\mathcal{D}_{Y_i} \subseteq \mathcal{D}_{X_i}$ for all $i \in I$. Let $\mathcal{B}_X \subseteq \mathcal{D}_X$, $\mathcal{B}_Y \subseteq \mathcal{D}_Y$ be the two countable basis of the respective domains and let $\mathcal{B} = \{“e_x \leq e_y” \mid e_x \in \mathcal{B}_X, e_y \in \mathcal{B}_Y\}$. Then the **relationship domain** $\mathcal{D}_{C(X,Y)}$ is the experimental domain generated by \mathcal{B} .

Proof. The only thing that needs to be checked is that the basis that generates $\mathcal{D}_{C(X,Y)}$ is countable. Since \mathcal{B}_X and \mathcal{B}_Y are countable, the set of statements of the form $\leq(e_x, e_y)$ with $e_x \in \mathcal{B}_X$ and $e_y \in \mathcal{B}_Y$ is also countable. \square

Proposition 2.34. The possibilities for a relationship domain $\mathcal{D}_{C(X,Y)}$ coincide with the possible experimental relationships between X and Y . That is, for each possibility z of $\mathcal{D}_{C(X,Y)}$ there exists a continuous function $f_z : X \rightarrow Y$ such that $z \equiv \bigwedge_{x \in X} “x \leq f_z(x)”$ and an $i \in I$ such that $\bar{g}_i(f_z(\bar{f}_i^{-1}(x_i))) \equiv f_c(x_i)$ for all $x_i \in X_i$ where f_c is the causal relationship between \mathcal{D}_{X_i} and \mathcal{D}_{Y_i} .

Proof. First we show that for each possibility z of the relationship domain there exists a set $P_z \subseteq X \times Y$ such that $z \equiv \bigwedge_{(x,y) \in P_z} “x \approx y” \wedge \bigwedge_{(x,y) \notin P_z} “x \not\approx y”$. Let z be a possibility of a relationship domain. This will be a minterm of the basis therefore $z \equiv \text{minterm}(\mathcal{B})$. Let $x \in X$ and $y \in Y$ be two possibilities and consider the statement “ $x \approx y$ ”. If z is true in a possible assignment, then the truth value for all statements of the form “ $e_x \leq e_y$ ” where $(e_x, e_y) \in \mathcal{B}_X \times \mathcal{B}_Y$ will be set. But $x \not\approx y$ if and only if there exists at least one pair $(e_x, e_y) \in \mathcal{B}_X \times \mathcal{B}_Y$ such that $x \leq e_x \leq e_y$ and $y \not\leq e_y$. Therefore if z is true in an assignment it will tell us whether $x \not\approx y$ is true or not, so we either have $z \leq “x \approx y”$ or $z \leq \neg “x \approx y”$. Let $P_z = \{(x, y) \in X \times Y \mid z \leq “x \approx y”\}$ and $\hat{z} = \bigwedge_{(x,y) \in P_z} “x \approx y” \wedge \bigwedge_{(x,y) \notin P_z} \neg “x \approx y”$. Then $z \leq \hat{z}$.

Conversely, if we suppose \hat{z} to be true in an assignment, then we will know the truth assignment for all statements of the form “ $x \approx y$ ” where $(x, y) \in X \times Y$. Let $(e_x, e_y) \in \mathcal{B}_X \times \mathcal{B}_Y$ and consider the statement “ $e_x \leq e_y$ ”. We have $e_x \leq e_y$ if and only if there is no pair (x, y) such that $x \leq e_x$, $y \not\leq e_y$ and $x \approx y$. Therefore if \hat{z} is true in an assignment it will tell us whether $e_x \leq e_y$ is true or not, so we either have $\hat{z} \leq “e_x \leq e_y”$ or $\hat{z} \leq \neg “e_x \leq e_y”$. As z is a minterm of statements of the form “ $e_x \leq e_y$ ”, we’ll either have $\hat{z} \leq z$ or $\hat{z} \not\leq z$. Since $z \leq \hat{z}$, $\hat{z} \approx z$ and therefore $\hat{z} \leq z$. Since we also have $z \leq \hat{z}$, then we have $z \equiv \hat{z}$.

Now we want to show that for each possibility z there is a function $f_z : X \rightarrow Y$ such that $z \equiv \bigwedge_{x \in X} “x \leq f_z(x)”$. We note that statements of the form “ $s_x \approx s_y$ ” are precisely the statements used to generate the relationship metacontext. Statements of the form “ $x \approx y$ ”, where $x \in X$ and $y \in Y$, are enough to determine the truth value of statements of the form “ $s_x \approx s_y$ ”. Therefore for each possibility z there is one and only one possible truth assignment in which z is true. That is, each x correspond to particular logical context \mathcal{S}_i . Since in that context $\mathcal{D}_{Y_i} \subseteq \mathcal{D}_{X_i}$ there will be a causal relationship $f_i : X_i \rightarrow Y_i$ such that $x_i \leq f_i(x_i)$ which means $x_i \approx y_i$ if and only if $y_i = f_i(x_i)$. Since causal relationships are unique and the possible assignments are distinct, to each z will correspond a unique $f_z : X \rightarrow Y$ such that $\bar{g}_i \circ f_z \circ \bar{f}_i^{-1} = f_i$ and $z \equiv \bigwedge_{x \in X} “x \leq f_z(x)”$. Moreover f_z is continuous since \bar{f}_i and \bar{g}_i are logical isomorphisms. \square

Note that the definition for the metacontext simply declares some statements to be veri-

fiable: it does not guarantee that this is actually possible. So when can we do it? Or better, what do we need in practice to be able to build the necessary confidence? Let's think how we would test the relationship between temperature and height of a mercury column. We would prepare different mercury samples with different values of temperatures in many different conditions, measure height and temperature with different devices, repeat many times, ask someone else to do it independently, compare results, and so on. At some point we will have explored enough of the possible cases, checked and tried anything that could invalidate the result and we will have the confidence to say that *"if the temperature of the mercury column is between 24 and 25 Celsius then its height is between 24 and 25 millimeters"*. We should stress that the procedure is not at all to observe a few values and then generalize. It is not mere induction. It is the ability to prepare and control the system in many different conditions and our inability to violate the relationship that gives us the confidence needed to reach the conclusion.

Suppose, in fact, that we want to experimentally verify the link between inflation and money supply. As long as we cannot create new countries in different economic conditions, the only thing we can do is gather data for as many nations as we can throughout history.⁵ Since we cannot purposely explore different conditions and we can't even replicate older ones, the best we can do is show that there was a correlation between those specific values. It is the inability to freely and fully explore the problem space that may not enable us to experimentally verify the causal relationship.⁶

The point is that we cannot give a general purpose algorithm for how to construct experimental tests for the relationships starting from the experimental tests of the individual domains. We cannot formalize in general what it means that we have explored the space "enough" to consider the relationship verified, no more than we can formalize when the data collected is "enough" to consider a statement verified or when a statement is specific "enough" to consider the semantics well defined. This is where the practice of experimental science comes in.

But, while we may not know in general whether we can experimentally verify a relationship between two specific domains, we do know that the relationship domain can always be constructed in principle and therefore our mathematical framework is complete. That is: we can take two experimental domains (e.g. \mathcal{D}_t for time and \mathcal{D}_x for position), construct a relationship domain between them (e.g. $\mathcal{D}_{x(t)}$ for trajectories in space), then take another experimental domain (e.g. $\mathcal{D}_{(q,p)}$ for the states) and construct another relationship domain between the two (e.g. $\mathcal{D}_{(q,p) \rightarrow x(t)}$ for the relationship between states and trajectories). Our general mathematical theory of experimental science is therefore closed, since we can recursively create relationship domains about relationship domains indefinitely while remaining within its bounds. In other words, experimentally distinguishable objects and their relationships will never lie outside of the theory.

⁵Computer simulations can sometimes alleviate this problem, though they are only as good as the model one uses.

⁶In this sense, experimental sciences allow for more rigorous results than observational sciences precisely because it is more feasible to experimentally test relationships between domains.

2.4 Summary

In this chapter we have seen the first important set of consequences of our general mathematical theory of experimental science. We have seen that experimental relationships between domains can be defined in terms of inference between verifiable statements or equivalently in terms of causal relationship between possibilities. The causal relationship always corresponds to a continuous function in the natural topology. Moreover, it will need to preserve any additional mathematical structure that formalizes a physical characteristic of the domain.

We have also seen how to combine a set of countably many experimental domains and how the combined possibilities depend on the logical and semantic relationships that exist across domains. Depending on the case, we can have the Cartesian product of the possibilities (with the product topology) all the way to the disjoint union of the established possibilities (with the disjoint union topology).

We have shown that it is possible to construct experimental domains for the relationships themselves. This means that within our theory we can describe relationships of arbitrarily higher order (i.e. relationships about relationships) and therefore we never go outside of our framework.

As we have only explored simple constructions based on the original concepts, these conclusions are of a general nature and therefore must apply to any area of science.

Chapter 3

Properties and quantities

In this chapter we will introduce the idea of **properties**, which we use to label the possibilities of an experimental domain. For example, we may use names to distinguish among people, fur color to distinguish among cats, pressure to distinguish the state of a gas. In particular, we'll define a **quantity** as a property that has a magnitude: we can compare two different values and determine which one is greater or smaller.

Instead of simply assuming we have quantities, we will need to construct them from the verifiable statements of an experimental domain. To do that, we will prove three theorems that will give necessary and sufficient conditions for that construction, clarifying how quantities are defined experimentally. These theorems will allow us to understand how the ordering of the values for a quantity are linked to the logical relationships between the verifiable statements, and in particular how the ordering of values is linked to the ordering defined by narrowness.

After characterizing quantities in general, we will see the cases of **discrete** and **continuous** quantities. We will see how discrete quantities are linked to decidability, the ability to verify both a statement and its negation, while continuous quantities are linked to a lack thereof. We will then discuss how the requirements for continuous quantities cannot be truly realized in practice and what happens when that idealization fails.

3.1 Properties

Whether the result of direct observation or not, a physical object is most often identified by its properties and their values. A person may be identified by name and date of birth; a bird may be characterized by the color of its beak; the state of a particle may be qualified by its position and momentum.

In this section we define how in general we can use properties to label the possibilities of an experimental domain. Each property will specify a continuous function from the possibilities to the topological space of the possible property values. Some properties may fully identify each possibility, as they assign distinct values, and some may also fully characterize which statements are verifiable.

Suppose \mathcal{D}_X is a domain for animal identification and X , its possibilities, are all animal species. Providing good names and definitions for species is a whole scientific subject by itself (i.e. taxonomy) with its own rules (i.e. the International Code of Zoological Nomenclature). The ICZN assigns each species a name composed of two Latin words. For example, "*Passer domesticus*" is the official name for the house sparrow while "*Passer italiae*" is the one for

the Italian sparrow. This means that identifying an animal species is equivalent to identifying its name. Therefore the domain to identify the species name $\mathcal{D}_Q \equiv \mathcal{D}_X$ is equivalent to the one for identifying the actual species and we have an experimental relationship $q : X \rightarrow \mathcal{Q}$ between them. As each species is given a unique name this is the most discriminating type of property: one that covers the whole range of possibilities and fully identifies each of them. But this is a special case.

The ICZN also assigns a genus (pl. genera) to each species, which corresponds to the first part of the species name. For example, both aforementioned sparrows are of the genus “*Passer*” while all swans, black or white, are of the genus “*Cygnus*”. Now suppose that \mathcal{Q} is the set of all names for all genera and $q : X \rightarrow \mathcal{Q}$ is the function that gives us the genus name for each species. We can still use this as a label for our possibilities, but it does not fully identify them.

On a different note, we could decide to distinguish species by their morphological attributes. For example, if \mathcal{Q} is the set of all colors, we can imagine a function that gives us the beak color for each species. The function is partial, not defined on all elements, as not all species have a beak or its color may not be unique. We have $q : U \rightarrow \mathcal{Q}$ where $U \subseteq X$ is the set where the property is defined. To be consistent, though, we must at least be able to tell in what cases the property is actually defined. For example, we must be able to tell whether an animal has a beak. Therefore there must be a verifiable statement that is true if and only if the property is well defined. The set U , then, must correspond to the verifiable set of said statement.

Other examples of properties include: postal addresses for buildings, tax ID numbers for people, generation for fundamental particles, position of the center of mass of a body. In all these cases, the general pattern is that we assign a label to a possibility from an established set.

Definition 3.1. A *property* for an experimental domain \mathcal{D}_X is any attribute we can use to distinguish between its possibilities. Formally, it is a tuple (\mathcal{Q}, q) where \mathcal{Q} is a topological space and $q : U \rightarrow \mathcal{Q}$ is a continuous function where $U \in \mathcal{T}_X$ is a verifiable set of possibilities.

Justification. Suppose q is a physical property that can be assigned to distinguish the set of possibilities $U \subseteq X$. We must be able to tell experimentally whether the property is defined for each possibility. Therefore we are justified to assume that U is a verifiable set. Let \mathcal{Q} be the set of possible values for the property. Each element U will be matched with a possible value. Therefore we are justified to assume there is a function $q : X \rightarrow \mathcal{Q}$ that specifies that mapping.

As we must be able to experimentally distinguish between the values in \mathcal{Q} , there must be a collection \mathcal{T}_Q of verifiable sets that corresponds to the possible verifiable statements associated to the quantity. Therefore we are justified to assume \mathcal{Q} is a topological space. Moreover, since the property value is determined by each possibility, the function q establishes a causal relationship between the two and therefore we are justified to assume q is a continuous function. \square

Note that we have not defined a property as an experimental domain. The possibility of the domain for a property would be a statement like “*the species name for this animal is *Passer domesticus**” while the value is just the object “*Passer domesticus*”. The idea is that the property and its values are defined independently of the particular object. For example,

distance in meters is defined independently of whether we later want to measure the earth diameter or its distance to the moon.¹

The fact that the property and its values are defined more generally leads to some subtle issues: even if each possibility of the domain corresponds to one property value and each verifiable statement of the property is equivalent to a verifiable statement of the domain, the reverse may not be true. Consider the experimental domain for identifying negatively charged fundamental particles. In this case, each particle will correspond to a unique value of mass, and by measuring the mass we can infer the particle. But for each possible value of mass we will not have a possible fundamental particle. Moreover, the verifiable statement “*this particle is an electron*” would correspond to an infinitely precise statement of the form “*the mass of the particle is exactly 510998.9461... eV*” which is not verifiable. The reason is that identifying a negatively charged particle is not the same as measuring its mass with arbitrary precision: in the first case we can stop when there is only one possible particle within the finite range of our measurement. Now consider the experimental domain for the position of the center of mass of a particle along a particular direction. Not only is the distance from a fixed point enough to identify the position, but for each value of the distance we also have a possible position of the particle.

We say that a property fully characterizes an experimental domain if each possible value of a property corresponds to a possibility of the domain, like in the case of the distance from a reference and position of a particle. The idea is that it tells us what cases are possible and how they are experimentally verified. Instead, we say that a property fully identifies an experimental domain if each possibility corresponds to a unique value, like in the case of the mass of a negatively charged particle. The idea is that the value of the property allows us to uniquely identify the possibility but nothing more.

Definition 3.2. *The possibilities of an experimental domain are **fully identified** by a property if its value is enough to uniquely determine a possibility. Formally, let (\mathcal{Q}, q) be a property for an experimental domain \mathcal{D}_X . Then the possibilities X are fully identified if $q : X \rightarrow \mathcal{Q}$ is injective.*

Definition 3.3. *An experimental domain is **fully characterized** by a property if all verifiable statements in the domain correspond to verifiable statements on the property and vice-versa. Formally, let (\mathcal{Q}, q) be a property for an experimental domain \mathcal{D}_X . Then it is fully characterized if $q : X \rightarrow \mathcal{Q}$ is a homeomorphism.*

3.2 Quantities and ordering

We now focus our attention on those properties that can be quantified. The number of people in a group can be quantified by an integer, the distance between two objects can be quantified in meters, the force acting on a body can be quantified by the magnitude of a vector expressed

¹While in principle we could envision a formal construction of properties and their values purely built on top of the notion of statement (e.g. define the value *red* to be the set of all statements that claim that something is of that color) we are not going to do so. While appealing in some respects, it is not clear whether such a construction would actually help to clarify underlying assumptions while it is clear it would distract from other more critical issues.

in Newtons. The defining characteristic for a quantity is that a value can be greater, smaller or equal to another.

In this section we will define a quantity as a property with a linear order which we assume given a priori, and study its relationship with the experimental domain. We will define the order topology, which assumes one can experimentally verify whether a value is before or after another one. We will see that a domain is fully characterized by a quantity only if the possibilities 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 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 vector will not strictly be a quantity but a collection of quantities.²

We also have to define how this 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 511 ± 0.5 keV*” which we take to be equivalent to “*the mass of the electron is more than 510.5 keV but less than 511.5 keV*”.³ For integers, this also allows us to verify particular numbers as “*the earth has one natural satellite*” is equivalent to the “*the earth has more than zero natural satellites and fewer than two*”. Therefore we will define the order topology as the one generated by sets 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 \mathcal{Q} is a relationship $\leq: \mathcal{Q} \times \mathcal{Q} \rightarrow \mathbb{B}$ 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 (\mathcal{Q}, \leq) be a linearly ordered set. The **order topology** is the topology generated by the collections of sets of the form:

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

²In other languages, there are two words to differentiate quantity as in “physical quantity” (e.g. grandezza, Grösse, grandeur) and as in “amount” (e.g. quantità, Menge, quantité). It is the second meaning of quantity that is captured here.

³The sentence “*the mass of the electron is 511 ± 0.5 keV*” could instead be referring to statistical uncertainty instead of an accuracy bound and would constitute a different statement when the different meaning is attached. We will be treating these types of statistical statements later in the book, but suffice it to say that they cannot be defined before statements that identify bounds.

Definition 3.6. A *quantity* for an experimental domain \mathcal{D}_X is a linearly ordered property. Formally, it is a tuple (\mathcal{Q}, \leq, q) where (\mathcal{Q}, q) is a property, $\leq: \mathcal{Q} \times \mathcal{Q} \rightarrow \mathbb{B}$ is a linear order and \mathcal{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. The set \mathcal{Q} may correspond to the integers, real numbers or a set of words ordered alphabetically.⁴ The units are not captured by the numbers themselves: they are captured by the function q 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 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 1 Kg” precedes “the mass of the system is 2 Kg”. This ordering of the possibilities will be linked to the natural topology as “the mass of the system is less than 2 Kg”, the disjunction 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 assign a quantity to an experimental domain if it already has a natural ordering of the same type.

Definition 3.7. Let \mathcal{D}_X be an experimental domain and X its possibilities. We say $\leq: X \times X \rightarrow \mathbb{B}$ is a **natural order** for the possibilities of the domain if the order topology constructed from that ordering coincides with the natural topology. We say the domain and the possibilities are **naturally ordered** if they admit a natural order.

Definition 3.8. Let (\mathcal{Q}, \leq) and (\mathcal{P}, \leq) be two ordered sets. A function $f: \mathcal{Q} \rightarrow \mathcal{P}$ is **increasing** if for every $q_1, q_2 \in \mathcal{Q}$, $q_1 \leq q_2$ implies $f(q_1) \leq f(q_2)$. It is an **order isomorphism** if it is bijective and, for every $q_1, q_2 \in \mathcal{Q}$, $q_1 \leq q_2$ if and only if $f(q_1) \leq f(q_2)$.

Theorem 3.9 (Property ordering theorem). An experimental domain \mathcal{D}_X is fully characterized by a quantity (\mathcal{Q}, \leq, q) if and only if it is naturally ordered and the possibilities are order isomorphic to the quantity.

Proof. Suppose \mathcal{D}_X is fully characterized by a quantity (\mathcal{Q}, \leq_q, q) . Since q is a bijection, we can define an ordering $\leq_x: X \times X \rightarrow \mathbb{B}$ such that $x_1 \leq_x x_2$ if and only if $q(x_1) \leq_q q(x_2)$. As q is now an order isomorphism, sets of the type (q_1, ∞) will be mapped to $(q^{-1}(q_1), \infty)$, and sets of the type $(-\infty, q_2)$ will be mapped to $(-\infty, q^{-1}(q_2))$ for all $q_1, q_2 \in \mathcal{Q}$. As q is also a homeomorphism, it will map a basis of one space to and only to a basis of the other. This means the collection of sets of the type (x_1, ∞) and $(-\infty, x_2)$ form a basis, and therefore the ordering \leq_x is a natural order.

⁴When consulting the dictionary, we use the fact that we can experimentally tell whether the word we are looking for is before or after the one we randomly selected.

We can run the argument in reverse, by assuming we have a naturally ordered experimental domain, finding a set \mathcal{Q} that is order isomorphic to the possibility and showing that the topology induced by the isomorphism is the order topology. \square

Now that we have established that the linear ordering is something already present in the domain, we can show it is actually based on the ordering of verifiable statements in terms of narrowness. That is, “*the mass of the system is 1 Kg*” is before “*the mass of the system is 2 Kg*” precisely because “*the mass of the system is less than 1 Kg*” is narrower than “*the mass of the system is less than 2 Kg*”. The set of statements \mathcal{B}_b of the form “*the mass of the system is less than q_1 Kg*” are linearly ordered by narrowness, and their ordering is the same as the ordering of the possibilities/values. Similarly, the set of statements \mathcal{B}_a of the form “*the mass of the system is more than q_1 Kg*” 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 1 Kg*” with $s_2 =$ “*the mass of the system is less than 1 Kg*”. We have $s_2 \leq s_1$. In fact, if we replace the value in s_2 with anything less than 1 Kg we’ll still have $s_2 \leq s_1$. Instead if we use a value greater than 1 Kg we’d have $s_1 \leq s_2$. In other words, if we call B the set that includes both the less-than-or-equal and less-than statements this is also linearly ordered by narrowness. But “*the mass of the system is less than or equal to 1 Kg*” is equivalent to \neg “*the mass of the system is greater than 1 Kg*”. In other words, $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$ contains all the statements like “*the mass of the system is less than q_1 Kg*” and \neg “*the mass of the system is more than q_1 Kg*” and these are all linearly ordered by narrowness.

The ordering of B can be further characterized. Note that $s_1 =$ “*the mass of the system 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 is broader than s_2 but narrower than s_1 since they differ for 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 copy is immediately followed by an element of the other copy. Moreover, if a statement in B has an immediate successor, there must be only one case that separates the two. If we call q_1 the value of that case, then the statement must be of the form “*the mass 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 associated with q_1 . Therefore statements in B that have an immediate successor must be in \mathcal{B}_b 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 two parts \mathcal{B}_b and \mathcal{B}_a such that $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$ is linearly ordered by narrowness with those characteristics, then the experimental domain is naturally ordered. The possibilities can be written as $x \equiv \neg s_b \wedge \neg s_a$ where $s_b \in \mathcal{B}_b$ and $\neg s_a$ is the immediate successor. In fact, “*the mass of the system is 1 Kg*” is equivalent to “*the mass of the system is not less than 1 Kg and not more than 1 Kg*”. Note that, because of negation, the possibilities may not in general be verifiable.

Definition 3.10. Let \mathcal{D}_X be a naturally ordered experimental domain and X its possi-

bilities. We define the following notation: “ $x < x_1$ ” = $\bigvee_{\{x \in X \mid x < x_1\}} x$, “ $x > x_1$ ” = $\bigvee_{\{x \in X \mid x > x_1\}} x$, “ $x \leq x_1$ ” = $\bigvee_{\{x \in X \mid x \leq x_1\}} x$ and “ $x \geq x_1$ ” = $\bigvee_{\{x \in X \mid x \geq x_1\}} x$. That is, they represent statements like “the value of the quantity x is less than x_1 ”.

Corollary 3.11. *Based on the previous notation, “ $x \leq x_1$ ” \equiv \neg “ $x > x_1$ ” and “ $x \geq x_1$ ” \equiv \neg “ $x < x_1$ ”.*

Proof. Note that “ $x \leq x_1$ ” \vee “ $x > x_1$ ” $\equiv \bigvee_{x \in X} x \equiv \top$ while “ $x \leq x_1$ ” \wedge “ $x > x_1$ ” $\equiv \bigvee_{x \in \emptyset} x \equiv \perp$. Therefore one is the negation of the other. Similarly “ $x < x_1$ ” \vee “ $x \geq x_1$ ” $\equiv \bigvee_{x \in X} x \equiv \top$ while “ $x < x_1$ ” \wedge “ $x \geq x_1$ ” $\equiv \bigvee_{x \in \emptyset} x \equiv \perp$. \square

Definition 3.12. *Let \mathcal{D}_X be a naturally ordered experimental domain and X its possibilities. Define $\mathcal{B}_b = \{“x < x_1” \mid x_1 \in X\}$, $\mathcal{B}_a = \{“x > x_1” \mid x_1 \in X\}$ and $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$.*

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

Proposition 3.14. *Let \mathcal{D}_X be a naturally ordered experimental domain. Then (\mathcal{B}_b, \leq) , (\mathcal{B}_a, \geq) and (B, \leq) are linearly ordered sets. Moreover (\mathcal{B}_b, \leq) , (\mathcal{B}_a, \geq) are order isomorphic to (X, \leq) .*

Proof. Let $f : X \rightarrow \mathcal{B}_b$ be defined such that $f(x_1) = “x < x_1”$. As there is one and only one statement “ $x < x_1$ ” for each $x_1 \in X$, f is a bijection. Suppose $x_1 \leq x_2$, we have $f(x_2) \equiv \bigvee_{\{x \in X \mid x < x_2\}} x \equiv \left(\bigvee_{\{x \in X \mid x < x_1\}} x \right) \vee \left(\bigvee_{\{x \in X \mid x < x_2\}} x \right) \equiv f(x_1) \vee f(x_2)$ and therefore $f(x_1) \leq f(x_2)$. On the other hand if $f(x_1) \leq f(x_2)$ then as sets $(-\infty, x_1) \subseteq (-\infty, x_2)$ which means $x_1 \leq x_2$. This means that f is an order isomorphism between (\mathcal{B}_b, \leq) and (X, \leq) .

Similarly, let $g : X \rightarrow \mathcal{B}_a$ be defined such that $g(x_1) = “x > x_1”$. As there is one and only one statement “ $x > x_1$ ” for each $x_1 \in X$, g is a bijection. Suppose $x_1 \leq x_2$, we have $g(x_1) \equiv \bigvee_{\{x \in X \mid x > x_1\}} x \equiv \left(\bigvee_{\{x \in X \mid x > x_1\}} x \right) \vee \left(\bigvee_{\{x \in X \mid x > x_2\}} x \right) \equiv g(x_1) \vee g(x_2)$ and therefore $g(x_1) \geq g(x_2)$. On the other hand if $g(x_1) \geq g(x_2)$ then as sets $(x_1, \infty) \supseteq (x_2, \infty)$ which means $x_1 \leq x_2$. This means that g is an order isomorphism between (\mathcal{B}_a, \geq) and (X, \leq) .

To show that B is linearly ordered, let $x_1, x_2 \in X$. If they both come from either \mathcal{B}_b or $\neg(\mathcal{B}_a)$ then they are already ordered by narrowness. If not, consider the two statements “ $x < x_1$ ” and “ $x \leq x_2$ ” \equiv \neg “ $x > x_2$ ”. As X is linearly ordered, either $\{x \in X \mid x < x_1\} \subseteq \{x \in X \mid x \leq x_2\}$ or $\{x \in X \mid x \leq x_2\} \subseteq \{x \in X \mid x < x_1\}$. Therefore either “ $x < x_1$ ” \leq “ $x \leq x_2$ ” or “ $x \leq x_2$ ” \leq “ $x < x_1$ ”. Which means $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$ is linearly ordered by \leq . \square

Proposition 3.15. *Let \mathcal{B}_b and \mathcal{B}_a be two sets of verifiable statements such that $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$ is linearly ordered by narrowness. Let \mathcal{D}_b and \mathcal{D}_a be the experimental domains they respectively generate and $D = \mathcal{D}_b \cup \neg(\mathcal{D}_a)$. Then (\mathcal{D}_b, \leq) , (\mathcal{D}_a, \geq) and (D, \leq) are linearly ordered.*

Proof. First we show that (\mathcal{D}_b, \leq) is linearly ordered. We have that \mathcal{B}_b is linearly ordered by narrowness because it is a subset of B which is linearly ordered by narrowness. Note that the conjunction of a finite set of statements linearly ordered by narrowness will return the narrowest element and the disjunction of a finite set of statements linearly ordered by narrowness will return the broadest element. The countable disjunction, instead, can introduce new elements. But using those elements again will not introduce new ones: the disjunction of countable disjunctions will still be a countable disjunction; the finite conjunction of countable disjunctions is the countable disjunction of finite conjunctions, which returns elements of the original set and therefore reduces to countable conjunctions. Therefore, when forming \mathcal{D}_b the only new elements will be the countable disjunctions.

Consider two countable sets $B_1, B_2 \subseteq \mathcal{B}_b$. Their disjunctions $\mathbf{b}_1 = \bigvee_{\mathbf{b} \in B_1} \mathbf{b}$ and $\mathbf{b}_2 = \bigvee_{\mathbf{b} \in B_2} \mathbf{b}$ represent the narrowest statement that is broader than all elements of the respective set. Suppose that for each element of B_1 we can find a broader element in B_2 . Then \mathbf{b}_2 , being broader than all elements of B_2 , will be broader than all elements of B_1 . But since \mathbf{b}_1 is the narrowest element that is broader than all elements in B_1 , we have $\mathbf{b}_2 \geq \mathbf{b}_1$. Conversely, suppose there is some element in B_1 for which there is no broader element in B_2 . Since the initial set is fully ordered, it means that that element of B_1 is broader than all the elements in B_2 . This means that element is broader than \mathbf{b}_2 and since \mathbf{b}_1 is broader than all elements in B_1 we have $\mathbf{b}_1 \geq \mathbf{b}_2$. Therefore the domain \mathcal{D}_b generated by \mathcal{B}_b is linearly ordered by narrowness.

Now we show that (\mathcal{D}_a, \geq) is linearly ordered. The basis \mathcal{B}_a is linearly ordered by broadness because the negation of its elements are part of B and are ordered by narrowness. Note that broadness is the opposite order of narrowness and therefore a set linearly ordered by one is linearly ordered by the other. Therefore \mathcal{B}_a is also linearly ordered by narrowness and so is \mathcal{D}_a by the previous argument. Therefore \mathcal{D}_a is ordered by broadness.

To show that $D = \mathcal{D}_b \cup \neg(\mathcal{D}_a)$ is linearly ordered by narrowness, we only need to show that the countable disjunctions of elements of \mathcal{B}_b are either narrower or broader than the countable conjunctions of the negations of elements of \mathcal{B}_a . Let $B_1 \subseteq \mathcal{B}_b$ and $A_2 \subseteq \mathcal{B}_a$. The disjunction $\mathbf{b}_1 = \bigvee_{\mathbf{b} \in B_1} \mathbf{b}$ represents the narrowest statement that is broader than all elements of B_1 while the conjunction $\neg \mathbf{a}_2 = \neg \bigvee_{\mathbf{a} \in A_2} \mathbf{a} = \bigwedge_{\mathbf{a} \in A_2} \neg \mathbf{a}$ represents the broadest statement that is narrower than all elements of $\neg(A_2)$. Suppose that for one element of $\neg(A_2)$ we can find a broader statement in B_1 . Then \mathbf{b}_1 , being broader than all elements in B_1 , will be broader than that one element in $\neg(A_2)$. But since $\neg \mathbf{a}_2$ is narrower than all elements in $\neg(A_2)$, we have $\neg \mathbf{a}_2 \leq \mathbf{b}_1$. Conversely, suppose that for no element of $\neg(A_2)$ we can find a broader statement in B_1 . As B is linearly ordered, it means that all elements in $\neg(A_2)$ are broader than all elements in B_1 . This means that all elements in $\neg(A_2)$ are broader than \mathbf{b}_1 and therefore $\mathbf{b}_1 \leq \neg \mathbf{a}_2$. Therefore D is linearly ordered by narrowness. \square

Theorem 3.16 (Domain ordering theorem). *An experimental domain \mathcal{D}_X is naturally ordered if and only if it is the combination of two experimental domains $\mathcal{D}_X = \mathcal{D}_a \times \mathcal{D}_b$ such that:*

- (i) $D = \mathcal{D}_b \cup \neg(\mathcal{D}_a)$ is linearly ordered by narrowness

- (ii) all elements of D are part of a pair $(s_b, \neg s_a)$ such that $s_b \in \mathcal{D}_b$, $s_a \in \mathcal{D}_a$ and $\neg s_a$ is either the immediate successor of s_b in D or $s_b \equiv \neg s_a$
- (iii) if $s \in D$ has an immediate successor, then $s \in \mathcal{D}_b$

Proof. Let \mathcal{D}_X be a naturally ordered experimental domain. Let \mathcal{B}_b and \mathcal{B}_a be defined as in 3.12 which means $\mathcal{B} = \mathcal{B}_b \cup \mathcal{B}_a$ is the basis that generates the order topology. Let \mathcal{D}_b be the domain generated by \mathcal{B}_b and \mathcal{D}_a be the domain generated by \mathcal{B}_a . Then \mathcal{D}_X is generated from \mathcal{D}_b and \mathcal{D}_a by finite conjunction and countable disjunction and therefore $\mathcal{D}_X = \mathcal{D}_b \times \mathcal{D}_a$.

To prove (i), we have that \mathcal{B}_b and \mathcal{B}_a are linearly ordered by 3.14. We need to show that the linear ordering holds across the sets. Let $x_1, x_2 \in X$ and consider the two statements “ $x < x_1$ ” and “ $x \leq x_2$ ” \equiv “ $x > x_2$ ”. As X is linearly ordered, either $\{x \in X \mid x < x_1\} \subseteq \{x \in X \mid x \leq x_2\}$ or $\{x \in X \mid x \leq x_2\} \subseteq \{x \in X \mid x < x_1\}$. Therefore either “ $x < x_1$ ” \preceq “ $x \leq x_2$ ” or “ $x \leq x_2$ ” \preceq “ $x < x_1$ ”. Which means $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$ is linearly ordered by \preceq . By 3.15 the set $D = \mathcal{D}_b \cup \neg(\mathcal{D}_a)$ is also linearly ordered.

To prove (ii), let $s_b \in \mathcal{D}_b$. Take $s_a \in \mathcal{D}_a$ such that $\neg s_a$ is the narrowest statement in $\neg(\mathcal{D}_a)$ that is broader than s_b . This exists because \mathcal{D}_a is closed by infinite disjunction. As $\neg s_a \succeq s_b$, let X_1 be the set of possibilities compatible with $\neg s_a$ but not compatible with s_b . The set cannot have more than one element, or we could find an element $x_1 \in X_1$ such that $s_b \preceq “x \leq x_1” < \neg s_a$. If X_1 contains one possibility, then $\neg s_a$ is the immediate successor. If X_1 is empty then $s_b \equiv \neg s_a$. Similarly, we can start with $s_a \in \mathcal{D}_a$ and find $s_b \in \mathcal{D}_b$ such that s_b is the broadest statement in \mathcal{D}_b that is narrower than $\neg s_a$. Let X_1 be the set of possibilities compatible with $\neg s_a$ but not compatible with s_b . If X_1 contains one possibility, then $\neg s_a$ is the immediate successor and if X_1 is empty then $s_b \equiv \neg s_a$.

To prove (iii), let $s_1, s_2 \in D$ such that s_2 is the immediate successor of s_1 . This means we can write $s_2 \equiv s_1 \vee x_1$ for some $x_1 \in X$. This means $s_1 \equiv “x < x_1”$ while $s_2 \equiv “x \leq x_1”$ and therefore $s_1 \in \mathcal{B}_b$.

Now we need to prove the opposite direction. Now let $\mathcal{D}_X = \mathcal{D}_b \times \mathcal{D}_a$ be an experimental domain as described in the theorem. Let X be the set of all statements of the form $x = \neg s_b \wedge \neg s_a$ for which $s_b \in \mathcal{D}_b$, $s_a \in \mathcal{D}_a$ and $\neg s_a$ is the immediate successor of s_b in D . All statements in X are possibilities. In fact, take $x = \neg s_b \wedge \neg s_a \in X$ and $s \in D$. It is not impossible because $s_b < \neg s_a$. Since $\neg s_a$ is the immediate successor of s_b , we either have $s \preceq s_b \preceq s_b \vee s_a \equiv \neg x$ or $s \succeq \neg s_a \succeq x$. That is, $x \not\preceq s$ or $x \preceq s$. And since the theoretical domain of \mathcal{D}_X can be generated from D , x is either narrower than or incompatible with any other statement in the theoretical domain. Therefore x is a possibility.

Now we show that X , as defined, covers all possibilities. Let x be a possibility for \mathcal{D}_x . Let $F_x = \{s \in D \mid x \not\preceq s\}$ and $T_x = \{s \in D \mid x \preceq s\}$. Since x is a possibility $F_x \cup T_x = D$ and since D is linearly ordered, $s_1 < s_2$ for all $s_1 \in F_x$ and $s_2 \in T_x$. Let $f_x = \bigvee_{s \in F_x} s$ and $t_x = \bigwedge_{s \in T_x} s$. To see that $f_x \in D$, let $f'_x = \bigvee_{s \in F_x \cap \mathcal{D}_b} s$. This will be in F_x as it is in \mathcal{D}_b and will be false if x is true. Because of (ii), there can be only one statement in $F_x \cap \neg(\mathcal{D}_a)$ that is broader than f'_x but narrower than the other elements of \mathcal{D}_b that are not in F_x , therefore f_x is in D as it has been reduced to a finite disjunction. Similarly, $t_x \in D$. Let $t'_x = \bigwedge_{s \in T_x \cap \neg(\mathcal{D}_a)} s$. This will be in T_x as it is in $\neg(\mathcal{D}_a)$ and will be true if x is true. Because of (ii), there can be only

one statement in $T_x \cap \mathcal{D}_b$ that is narrower than \mathbf{t}'_x but broader than the other elements of \mathcal{D}_a that are not in T_x , therefore \mathbf{t}_x is in D as it has been reduced to a finite conjunction.

Consider $\neg \mathbf{f}_x \wedge \mathbf{t}_x$: if true then \mathbf{f}_x will be false, and so will all statements in F_x since they are narrower; also \mathbf{t}_x will be true, and so will all statements in T_x since they are broader. We have $x \equiv \neg \mathbf{f}_x \wedge \mathbf{t}_x$. Since x is not impossible, $\mathbf{f}_x \neq \mathbf{t}_x$. Since all statements in D are either in F_x or T_x , \mathbf{t}_x is the immediate successor of \mathbf{f}_x . Therefore by (iii) $\mathbf{f}_x \in \mathcal{D}_b$, $\neg \mathbf{t}_x \in \mathcal{D}_a$ and $x \in X$.

Now we show that X can be given a natural ordering. Let $\mathcal{B}_b \subseteq \mathcal{D}_b$ be the set of statements that have an immediate successor in D and $\mathcal{B}_a \subseteq \mathcal{D}_a$ be the set of the negation of the immediate successors. Let $(\cdot)^{++} : \mathcal{B}_b \rightarrow \mathcal{B}_a$ be the function such that $\neg(\mathbf{b}^{++}) = \neg \mathbf{b}^{++}$ is the immediate successor of \mathbf{b} . Let $\mathbf{b} : X \rightarrow \mathcal{B}_b$ be the function such that $x \equiv \neg \mathbf{b}(x) \wedge \neg \mathbf{b}(x)^{++}$. On X define the ordering \leq such that $x_1 \leq x_2$ if and only if $\mathbf{b}(x_1) \leq \mathbf{b}(x_2)$. Since (\mathcal{B}_b, \leq) is linearly ordered so is (X, \leq) . To show that the ordering is natural, suppose $x_1 < x_2$ then $\mathbf{b}(x_1) < \neg \mathbf{b}(x_1)^{++} \leq \mathbf{b}(x_2)$ and therefore $x_1 \leq \mathbf{b}(x_2)$. We also have $\neg \mathbf{b}(x_1)^{++} \leq \mathbf{b}(x_2) < \neg \mathbf{b}(x_2)^{++}$ and therefore $x_2 \leq \mathbf{b}(x_1)^{++}$. This means that given a possibility $x_1 \in X$, all and only the possibilities lower than x_1 are compatible with $\mathbf{b}(x_1)$ and therefore $\mathbf{b}(x_1) \equiv "x < x_1"$, while all and only the possibilities greater than x_1 are compatible with $\mathbf{b}(x_1)^{++}$ and therefore $\mathbf{b}(x_1)^{++} \equiv "x > x_1"$. The topology is the order topology and the domain has a natural ordering. \square

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 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 cases. 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 target position (e.g. between 1.2 and 1.3 cm). To measure weight, we can use a balance and a set of equally 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 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 tell whether what we are measuring is before (e.g. shorter, lighter, sooner) or after (e.g. longer, heavier, later) the reference.

Note that determining whether the quantity is exactly equal to the reference is not as easy: the mark on the ruler has a width, the balance has friction, the tick of our clock will last a finite amount of time. That is, the reference itself can only be compared up to a finite level of precision. This may be a problem when constructing the references themselves: how do we know that the marks on our ruler are equally spaced, or that the weights are equally prepared, 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 still say whether the value is well before or 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 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.

In our general mathematical theory of experimental science, we can capture the above discussion with the following definitions. A reference is represented by a set of three statements: they tell us 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 an impossibility. 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 reference has an extent and so does the object being measured.

We can compare the extent of two references and say that one is finer 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 ruler 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 cases are distinct and can't be true at the same time.

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

1. *we can verify whether the object is before or after the reference: \mathbf{b} and \mathbf{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.

Corollary 3.18. Let (b, o, a) be a reference. Then $b \vee o \vee a \equiv \top$.

Proof. By definition, we have $\neg b \wedge \neg a \leq o$ and by 1.22 $\neg(\neg b \wedge \neg a) \vee o \equiv \top \equiv b \vee a \vee o$. \square

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 $b_1 \geq b_2$, $o_1 \leq o_2$ and $a_1 \geq a_2$.

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

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

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

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

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

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

Mathematically, this type of ordering is strict in the sense that it defines what is strictly before and strictly after. It does not define what happens in the overlap, in between cases. One may be tempted to define the ordering based on how the references overlap, but that requires refining the references and, in the end, it means we are defining an ordering on those refined references, not the original ones.

Definition 3.22. A reference is before another if whenever an object is found before or on the first it cannot be on or after the second. Formally, $r_1 < r_2$ if and only if $b_1 \vee o_1 \not\leq o_2 \vee a_2$.

Proposition 3.23. Reference ordering satisfies the following properties:

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

and is therefore a **strict partial order**.

Proof. For irreflexivity, since the on statement can't be impossible, we have $o \approx o$ and therefore $b \vee o \approx o \vee a$. Therefore a reference is not before itself and the relationship is irreflexive.

For transitivity, if $r_1 < r_2$, we have $b_1 \vee o_1 \not\approx o_2 \vee a_2$ and therefore $\neg(b_1 \vee o_1) \geq o_2 \vee a_2$ by 1.22. Since $b_1 \vee o_1 \vee a_1 \equiv \top$, we have $a_1 \geq \neg(b_1 \vee o_1)$. Similarly if $r_2 < r_3$ we'll have $a_2 \geq \neg(b_2 \vee o_2) \geq o_3 \vee a_3$. Putting it all together $\neg(b_1 \vee o_1) \geq o_2 \vee a_2 \geq a_2 \geq \neg(b_2 \vee o_2) \geq o_3 \vee a_3$, which means $b_1 \vee o_1 \not\approx o_3 \vee a_3$. \square

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. But references can overlap in different ways.

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

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

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

Definition 3.25. *Two references $r_1 = (b_1, o_1, a_1)$ and $r_2 = (b_2, o_2, a_2)$ are **aligned** if for any $s_1 \in \{b_1, \neg a_1\}$ and $s_2 \in \{b_2, \neg a_2\}$ we have $s_1 \leq s_2$ or $s_2 \leq s_1$. A set R of references is aligned if any pair of references is aligned.*

Proposition 3.26. *Let $R = \{(b_i, o_i, a_i)\}_{i \in I}$ be a set of aligned strict references. Let $\mathcal{B}_b = \{b_i\}_{i \in I}$ and $\mathcal{B}_a = \{a_i\}_{i \in I}$ be respectively the sets of before and after statements. Then the set $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$ is linearly ordered by narrowness.*

Proof. Let R be a set of aligned strict references. Let $s_1, s_2 \in B$. Suppose they are taken from the same reference. If they are both before statements or both after statements, we

have $s_1 \equiv s_2$ and therefore $s_1 \leq s_2$. If one is the before statement and the other is the not-after statement, since the reference is strict, we have $b \not\leq a$ and $b \leq \neg a$ by 1.22 and therefore $s_1 \leq s_2$ or $s_2 \leq s_1$. Now suppose they are taken from different references. Since they are aligned we have $s_1 \leq s_2$ or $s_2 \leq s_1$ by definition. \square

Proposition 3.27. *Let $r_1 = (b_1, o_1, a_1)$ and $r_2 = (b_2, o_2, a_2)$ be two references. If $r_1 < r_2$ then $b_1 \leq b_2$, $a_2 \leq a_1$, $b_1 \leq \neg a_2$, $\neg a_1 \leq b_2$ and therefore the two references are aligned.*

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

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

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

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

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. \square

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

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

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

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

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

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

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

If we have a set of references, we can generate an experimental domain by using the before and after statements as the basis. More specifically we can take all the before statements and generate the before domain \mathcal{D}_b and all the after statements and generate the after domain \mathcal{D}_a . If the references are all strict and aligned, the set $D = \mathcal{D}_b \cup \neg(\mathcal{D}_a)$ of all the before and not-after statements will be linearly ordered by narrowness. We recognize this as the first requirement of the domain ordering theorem 3.16.

Definition 3.32. Let $R = \{(b_i, o_i, a_i)\}_{i \in I}$ be a set of references. Let $\mathcal{B}_b = \{b_i\}_{i \in I}$ be the set of all before statements and $\mathcal{B}_a = \{a_i\}_{i \in I}$ be the set of all after statements. The experimental domain \mathcal{D} generated by R is the one generated by all before and after statements $\mathcal{B}_b \cup \mathcal{B}_a$. The before domain \mathcal{D}_b is the domain generated only by the before statements \mathcal{B}_b and the after domain \mathcal{D}_a is the domain generated only by the after statements \mathcal{B}_a .

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

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

Proof. By 3.26 we have that $B = \mathcal{B}_b \cup \neg(\mathcal{B}_a)$ is aligned by narrowness. By 3.15 the ordering extends to D . \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 overlap, we can break them apart into finer ones that do not overlap and one is after the other.

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

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

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

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

1. If $b_1, b_2 \in \mathcal{D}_b$ such that $b_1 < b_2$, then there exists $a \in \mathcal{D}_a$ such that $b_1 < \neg a \preceq b_2$.
2. If $a_1, a_2 \in \mathcal{D}_a$ such that $\neg a_1 < \neg a_2$, then there exists $b \in \mathcal{D}_b$ such that $\neg a_1 \preceq b < \neg a_2$.
3. If $a_1 \in \mathcal{D}_a$ and $b_2 \in \mathcal{D}_b$ such that $\neg a_1 < b_2$, then there exists $b \in \mathcal{D}_b$ and $a \in \mathcal{D}_a$ such that $\neg a_1 \preceq b < \neg a \preceq b_2$.

Proof. For the first, suppose $b_1, b_2 \in \mathcal{D}_b$ such that $b_1 < b_2$. Then $r_1 = (b_1, \neg b_1, \perp)$ and $r_2 = (b_2, \neg b_2, \perp)$ are strict references aligned with the domain such that $\neg b_2 < \neg b_1$. This means we can find $r_3 = (b_1, \neg b_1 \wedge \neg a, a)$ for some $a \in \mathcal{D}_a$ such that $r_3 < r_2$ and therefore $b_1 < \neg a \leq b_2$.

For the second, suppose $a_1, a_2 \in \mathcal{D}_a$ such that $\neg a_1 < \neg a_2$. Then $r_1 = (\perp, \neg a_2, a_2)$ and $r_2 = (\perp, \neg a_1, a_1)$ are strict references aligned with the domain such that $\neg a_1 < \neg a_2$. This means we can find $r_3 = (b, \neg b \wedge \neg a_2, a_2)$ for some $b \in \mathcal{D}_b$ such that $r_3 < r_2$ and therefore $\neg a_1 \leq b < \neg a_2$.

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

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

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

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

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

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

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

Now suppose \mathcal{D}_X is naturally ordered. Define the set $\mathcal{B}_b, \mathcal{B}_a$ and D as in 3.12. Let $R = \{(b, \neg b \wedge \neg a, a) \mid b \in \mathcal{B}_b, a \in \mathcal{B}_a, b < \neg a\}$ be the set of all references constructed from the basis. First let us verify they are references. The before and after statements are verifiable

since they are part of the basis. The on statement $\neg\mathbf{b} \wedge \neg\mathbf{a}$ is not impossible since $\mathbf{b} < \neg\mathbf{a}$ means $\mathbf{b} \neq \mathbf{a}$ and $\mathbf{b} \neq \neg\mathbf{a}$. The on statement is broader than $\neg\mathbf{b} \wedge \neg\mathbf{a}$ as they are equivalent and it is broader than $\mathbf{b} \wedge \mathbf{a}$ as that is impossible since $\mathbf{b} < \neg\mathbf{a}$. Therefore R is a set of references. Since the before and after statements of R coincide with the basis of the domain, \mathcal{D}_X is generated by R .

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

To show R is refinable, note that each reference can be expressed as (“ $x < x_1$ ”, “ $x_1 \leq x \leq x_2$ ”, “ $x > x_2$ ”) where $x_1, x_2 \in X$ and “ $x_1 \leq x \leq x_2$ ” \equiv “ $x \geq x_1$ ” \wedge “ $x \leq x_2$ ”. That is, every 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 “ $x \leq x_2$ ” $<$ “ $x < x_3$ ”. That is, we can find $x_5 \in X$ such that $x_2 < x_5 < x_3$ which means “ $x \leq x_2$ ” \leq “ $x < x_5$ ” and “ $x \leq x_5$ ” \leq “ $x < x_3$ ”. Therefore the reference $\mathbf{r}_3 \in R$ identified by (x_5, x_5) is between the two references. On the other hand, assume the second reference is finer than the first. Then $x_1 \leq x_3$ and $x_4 \leq x_2$ with either $x_1 \neq x_3$ or $x_4 \neq x_2$. 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 $\mathbf{r}_2 < \mathbf{r}_4$. Also note that the before statements of \mathbf{r}_1 and \mathbf{r}_3 are the same and the after statements of \mathbf{r}_1 and \mathbf{r}_4 are the same. Therefore we satisfy all the requirements and the set R is refinable 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, and therefore a quantity, only when the set of references meets special conditions. The references must be strict, meaning that before, on and after are mutually exclusive. They must be aligned, meaning that the before and not-after statement must be ordered by narrowness. They must be refinable, meaning when they overlap we can always find finer references with well defined before/after relationships. If all these conditions apply, we have a linear order. If any of these conditions fail, a linear order cannot be defined.

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

3.4 Discrete quantities

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

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

The first thing we want to do is characterize the ordering of the integers. That is, we want

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

In fact, if an ordered set is sparse we can always go from any element to another in finitely many steps. Therefore we can pick one element, call it zero, go forward one element at a time and assign a positive integer to all the following elements or go backward one element at a time and assign a negative integer to all the preceding elements.

Definition 3.39. A *chain* is a linearly ordered subset of an ordered set. A chain between two elements is a chain where the two elements are the greatest and smallest elements.

Definition 3.40. An ordered set is *sparse* if every chain between any two elements is finite.

Corollary 3.41. Every element in a sparse linearly ordered set that has a predecessor (or successor) has an immediate predecessor (or successor).

Proof. Let \mathcal{Q} be a sparse ordered set. Suppose $q_1 \in \mathcal{Q}$ has a predecessor (or successor) q_0 . Then we have a finite chain between the two, and the immediate predecessor (or successor) of q_1 is the second greatest (or smallest) element. \square

Remark. The converse of the corollary, that a linearly ordered set in which every element that has a predecessor (or successor) has an immediate predecessor (or successor) is sparse, is not true. Consider the integers with the following ordering $\{0, 1, 2, 3, \dots, -3, -2, -1\}$. All elements have an immediate predecessor/successor or no predecessor/successor, yet 3 and -3 have infinitely many elements in between.

Proposition 3.42. A linear order is sparse if and only if it is isomorphic to a contiguous subset of the integers.

Proof. Let \mathcal{Q} be a sparse linearly ordered set. Pick an element $q_0 \in \mathcal{Q}$. Let $q : \mathcal{Q} \rightarrow \mathbb{Z}$ such that it returns 0 for q_0 , 1 for its immediate successor (if it exists), -1 for its immediate predecessor (if it exists), 2 for the immediate successor of the immediate successor (if it exists) and so on. Since the order is sparse, all elements will eventually be reached through a chain of immediate successors/predecessors and will be assigned a value. The function is injective and order preserving, so it is an isomorphism over $q(\mathcal{Q})$, which, by construction, is a contiguous subset of the integers

Conversely, let \mathcal{Q} be an ordered set isomorphic to a contiguous subset of the integers and let $q : \mathcal{Q} \rightarrow \mathbb{Z}$ be the isomorphism. The number of elements between two elements of the set will be equal to the number of elements between the two corresponding integers, which is always finite. \square

We can now define a discrete quantity as one for which the ordering is sparse. While this may typically correspond to a set of contiguous integers, it is not necessary. For example, a set of names ordered alphabetically, the set of orbitals for a hydrogen atom, the possible energies for a quantum harmonic oscillator, these are all discrete quantities even if the label we use is not an integer.

The natural question now is: under what conditions are the possibilities of a domain ordered like the integers? The answer is straightforward: when the references within the domain have a sparse order. That is, between two references we can only put finitely many ordered references.

Definition 3.43. A *discrete quantity* for an experimental domain \mathcal{D}_X is a quantity (\mathcal{Q}, \leq, q) for which the ordering is sparse.

Theorem 3.44 (Discrete ordering theorem). *Let \mathcal{D}_X be an experimental domain. Then the following are equivalent:*

1. *the domain has a natural sparse order*
2. *the domain is fully characterized by a discrete quantity*
3. *the domain is generated by a set of refinable aligned strict references with a sparse order*

Proof. For (1) to (2), let \mathcal{D}_X be an experimental domain with a natural sparse order and let X be its possibilities. Pick an ordered set (\mathcal{Q}, \leq) that is order isomorphic to the possibilities and let $q : X \rightarrow \mathcal{Q}$ be an order isomorphism. By 3.9 \mathcal{D}_X is fully characterized by (\mathcal{Q}, \leq, q) . Since the order on X is sparse, X will be order isomorphic to a contiguous set of integers and so will \mathcal{Q} . Therefore \mathcal{Q} has a sparse order as well and is therefore a discrete quantity.

For (2) to (3), let \mathcal{D}_X be an experimental domain fully characterized by \mathcal{Q} . Then by 3.9 and by 3.38 it is generated by a set of refinable aligned strict references. Let \mathbf{r}_1 and \mathbf{r}_2 be two references aligned with the domain such that $\mathbf{r}_1 < \mathbf{r}_2$. Then the after statement of \mathbf{r}_1 will be of the form “ $x > q^{-1}(q_1)$ ”, the before statement of \mathbf{r}_2 will be of the form “ $x < q^{-1}(q_2)$ ” for some $q_1, q_2 \in \mathcal{Q}$ such that $q_1 < q_2$. Let $\mathbf{r} : \mathcal{Q} \rightarrow \mathcal{D}_X \times \bar{\mathcal{D}}_X \times \mathcal{D}_X$ be the function such that $\mathbf{r}(q_i) = (“x < q^{-1}(q_i)”, “x \geq q^{-1}(q_i)” \wedge “x \leq q^{-1}(q_i)”, “x > q^{-1}(q_i)”)$. Let $C = \{\mathbf{r}_1\} \cup \{\mathbf{r}(q_i) \mid q_i \in \mathcal{Q}, q_1 < q_i < q_2\} \cup \{\mathbf{r}_2\}$. This is the longest chain between \mathbf{r}_1 and \mathbf{r}_2 and it is finite because \mathcal{Q} has a sparse ordering. The set of references that generate the domain, then, must have a sparse ordering.

For (3) to (1), let \mathcal{D}_X be an experimental domain generated by a set of refinable aligned strict references with a sparse order. By 3.38 \mathcal{D}_X has a natural order. Let $\mathbf{r} : X \rightarrow \mathcal{D}_X \times \bar{\mathcal{D}}_X \times \mathcal{D}_X$ be the function such that $\mathbf{r}(x_i) = (“x < x_i”, “x \geq x_i” \wedge “x \leq x_i”, “x > x_i”)$. Let $R = \{\mathbf{r}(x_i) \mid x_i \in X\}$. Then R is order isomorphic to X . As the order on R is sparse then the order on X is sparse as well. \square

Now, consider the examples above of discrete quantities: in each case we can experimentally test whether we have a particular value or not. For example, we are always able to tell whether there are exactly three apples on the table or not.⁵ This is not a coincidence: there is a direct link between the ability to have consecutive references and decidability.

As we saw before in 3.31, two consecutive references are such that the before statement of one is equal to the negation of the after statement of the other. But since before and after statements are both verifiable, it means their negation is also verifiable: they are decidable.

⁵Recall that this is not the case with continuous quantities. Because of finite precision, we are able to exclude that a given particle has exactly zero mass but it is not possible to conclusively show that it has zero mass.

And since before and after statements generate the domain, all statements in the domain are decidable. It turns out that this will work in reverse as well: whenever we have a domain consisting of only decidable statements, we can always create a discrete quantity that fully characterizes the experimental domain.

Proposition 3.45. *The order topology for the integers is discrete.*

Proof. Each singleton $\{z\} \subseteq \mathbb{Z}$ is in the order topology since $\{z\} = (z-1, \infty) \cap (-\infty, z+1)$. Each arbitrary set of integers is the union of singletons and is therefore in the order topology as well. The order topology on the integers is discrete. \square

Proposition 3.46. *An experimental domain is decidable if and only if it is fully characterized by a discrete quantity.*

Proof. Let \mathcal{D}_X be a decidable domain. Then by 1.74 the set of possibilities X is countable and by 1.76 the natural topology is discrete. Since X is countable, there exists a bijective map $q : X \rightarrow \mathbb{Z}$. The map is a homeomorphism since the topology on both X and \mathbb{Z} is discrete. The domain \mathcal{D}_X is fully identified by (\mathbb{Z}, \leq, q) .

Let \mathcal{D}_X be fully characterized by (\mathbb{Z}, \leq, q) . This means that $q : X \rightarrow \mathbb{Z}$ is a homeomorphism. The natural topology for X is therefore discrete and the domain is decidable by 1.76. \square

Note that, since any discrete order leads to a decidable domain and a discrete topology, any reordering of the possibilities will give the same exact domain. This means that, while we can always assign a natural order, the order itself may not necessarily be meaningful. For example, we can always take a finite group of objects and arbitrarily assign each a unique integer to identify it. In the discrete case the domain itself is not enough to pick a unique order, though the set of aligned references that are used to generate the domain is.

Also note that for the link between decidability and discrete quantities to apply, it is crucial the quantity is measurable: that we can actually experimentally ascertain its values. Consider the domain with the possibilities “*there is no extra-terrestrial life*” and “*there is extra-terrestrial life*”. We can arbitrarily label 0 the former and 1 the latter. But since we cannot verify the first statement, we cannot really “measure” 0. In that case, the domain is fully identified by the discrete quantity, but not fully characterized.

3.5 Arbitrary precision and continuous quantities

The second type of quantities we want to consider are continuous ones: the average wingspan for a species, the population density of a country or the mass of a proton are all continuous quantities. These are quantities that are fully characterized by real numbers.

We will see that also continuous quantities have a simple characterization: between two references there can always be an infinite number of other references.

Similar to what we did for the integers, we want to characterize the ordering of the real numbers. This will be a little bit more involved as we will need a few more requirements. First of all we note that between two real numbers there are always infinitely many real numbers. Let’s call dense an ordered set that has that property. This is not enough to identify the real numbers, though; the rational numbers are also dense.

One reason real numbers are used over the rationals is that they contain all the limits. This property can be restated in terms of ordering in the following way. Suppose (\mathcal{Q}, \leq) is a linearly ordered set. Take a set $A \subset \mathcal{Q}$ that is bounded. That is, the set $B = \{q \in \mathcal{Q} \mid a \leq q \forall a \in A\}$ of elements that are greater than all the elements of A is not empty. Then we say that (\mathcal{Q}, \leq) is complete if B has a smallest element. One can show that the rationals are not complete. Say A contains all rationals less than π , then there is no smallest rational value that is greater than all elements of A .

Dense and complete linear orders exclude both the integers and the rationals, but they don't pick out only the reals. To do that we take advantage of two results of order theory. The first is that all dense countable ordered sets are isomorphic to the rational numbers. The second is that from any linear order one can construct its completion (i.e. you add all the missing limits), which is unique up to an order isomorphism. Suppose that (\mathcal{Q}, \leq) has a countable subset $Q \subset \mathcal{Q}$ that is dense in \mathcal{Q} . That is, for every two distinct elements $q_1, q_2 \in Q$ where $q_1 < q_2$ we can find an element $q \in Q$ such that $q_1 < q < q_2$. If \mathcal{Q} is dense, Q will also be and therefore will be order isomorphic to the rationals, since it is countable. If \mathcal{Q} is complete, then it will be the completion of its dense set Q , and therefore it will be order isomorphic to the reals.

Definition 3.47. An ordered set is said to be **dense** if between any two elements there exists an infinite chain.

Corollary 3.48. An ordered set is dense if and only if between two elements we can always find another one.

Proof. We give a definition of dense that is different from the typical definition because we want it to be formally similar to our definition of sparse. Here we show that our definition of dense is equivalent to the standard one.

Let (\mathcal{Q}, \leq) be a dense ordered set. Let $q_1, q_2 \in \mathcal{Q}$ then we can find an infinite chain between them. Take an element q within that chain that is not an endpoint. We have $q_1 < q < q_2$.

Now let (\mathcal{Q}, \leq) be a linearly ordered set such that between two elements we can always find another one. Let $q_a, q_b \in \mathcal{Q}$ and let $C \subset \mathcal{Q}$ contain q_a, q_b , an element q_1 such that $q_a < q_1 < q_b$, an element q_2 such that $q_1 < q_2 < q_b$, an element q_3 such that $q_2 < q_3 < q_b$ and so on. Then C is a chain between q_a and q_b that contains infinitely many elements. \square

Definition 3.49. A subset $Q \subset \mathcal{Q}$ of a linearly ordered set is **dense** in \mathcal{Q} if given $q_1, q_2 \in \mathcal{Q}$ such that $q_1 < q_2$ we can find $q \in Q$ such that $q_1 \leq q \leq q_2$.

Definition 3.50. A linearly ordered set \mathcal{Q} is **complete** if every non-empty bounded subset of \mathcal{Q} has a supremum. This is, given $A \subset \mathcal{Q}$ such that $B = \{q \in \mathcal{Q} \mid a \leq q \forall a \in A\}$ is not empty, then B has a smallest element.

Theorem 3.51. A linearly ordered set is dense, complete and has a countable dense subset if and only if it is order isomorphic to a contiguous subset of real numbers.

Remark. Proving this theorem would go beyond the scope of this book so we will take it as a given. The general idea is as follows. Show that the completion of an ordered set is

unique up to an isomorphism. Show that the given set is the completion of the countable dense subset. Show that the countable dense subset is order isomorphic to a subset of the rational numbers. Show that the real numbers are the completion of the rational numbers. Then the set is isomorphic to a subset of the real numbers.

We can now define a continuous quantity as one for which the values are a contiguous subset of the real numbers. While in principle we could define it on a generic set that is order isomorphic to the real numbers, we do not have any other example of such a set.

The natural question now is: under what conditions are the possibilities of a domain ordered like the real numbers? The answer is straightforward: when the references within the domain have a dense order. That is, between two references we can put infinitely many ordered references. Intuitively, between two marks of a ruler we can keep putting finer and finer marks.

Note that the dense order on the references is enough to get a dense complete order on the possibilities that has a countable dense subset. The completion comes without extra conditions because experimental domains are closed under countable disjunctions. New references, in fact, can be constructed as limits of others and, since they will be just other references they will have the same properties. The countable dense subset simply corresponds to the countable basis of the domain.

Definition 3.52. A *continuous quantity* for an experimental domain \mathcal{D}_X is a quantity (U, \leq, q) where $U \subseteq \mathbb{R}$ is a contiguous subset of the real numbers.

Theorem 3.53 (Continuous ordering theorem). *Let \mathcal{D}_X be an experimental domain. Then the following are equivalent:*

1. *the domain has a natural dense complete order that has a countable dense subset*
2. *the domain is fully characterized by a continuous quantity*
3. *the domain is generated by a set of refinable aligned strict references with a dense order*

Proof. For (1) to (2). Let \mathcal{D}_X be a domain with a natural dense complete order with a countable dense subset. Then by 3.51 it is order isomorphic to a contiguous subset of the real numbers. Then by 3.9 the domain is fully characterized by a continuous quantity (U, \leq, q) where $U \subseteq \mathbb{R}$.

For (2) to (3). Let \mathcal{D}_X be a domain fully characterized by a continuous quantity. Then by 3.9 and by 3.38 it is generated by a set of refinable aligned strict references. Let \mathbf{r}_1 and \mathbf{r}_2 be two references aligned with the domain such that $\mathbf{r}_1 < \mathbf{r}_2$. Then the after statement of \mathbf{r}_1 will be of the form “ $x > q^{-1}(q_1)$ ”, the before statement of \mathbf{r}_2 will be of the form “ $x < q^{-1}(q_2)$ ” for some $q_1, q_2 \in \mathcal{Q}$ such that $q_1 < q_2$. Let $\mathbf{r} : \mathcal{Q} \rightarrow \mathcal{D}_X \times \bar{\mathcal{D}}_X \times \mathcal{D}_X$ be the function such that $\mathbf{r}(q_i) = (“x < q^{-1}(q_i)”, “x \geq q^{-1}(q_i)” \wedge “x \leq q^{-1}(q_i)”, “x > q^{-1}(q_i)”)$. Let $C = \{\mathbf{r}_1\} \cup \{\mathbf{r}(q_i) \mid q_i \in \mathcal{Q}, q_1 < q_i < q_2\} \cup \{\mathbf{r}_2\}$. As \mathcal{Q} is dense, this chain will be infinite. Therefore the domain can be generated by a set of refinable aligned strict references with a dense order.

For (3) to (1). Let \mathcal{D}_X be an experimental domain generated by a set of refinable aligned strict references with a dense order. By 3.38 \mathcal{D}_X has a natural order. Let $\mathbf{r} : X \rightarrow$

$\mathcal{D}_X \times \bar{\mathcal{D}}_X \times \mathcal{D}_X$ be the function such that $\mathbf{r}(x_i) = (“x < x_i”, “x \geq x_i” \wedge “x \leq x_i”, “x > x_i”)$. Let $R = \{\mathbf{r}(x_i) \mid x_i \in X\}$. Then R is order isomorphic to X . As the order on R is dense then the order on X is dense as well. Let \mathcal{D}_b be the before domain. By 3.14 it is order isomorphic to X . Let $\mathcal{B}_b \subseteq \mathcal{D}_b$ be a countable basis for the domain. As it is a subset of a linearly ordered set, it will also be a linearly ordered set. As the basis is ordered by narrowness, finite conjunctions and disjunctions of basis elements will return a basis element. Therefore every element in \mathcal{D}_b is equivalent to the disjunction of a countable set of elements of \mathcal{B}_b . Let $\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{D}_b$ be two statements such that $\mathbf{b}_1 < \mathbf{b}_2$. Let $B_1, B_2 \subseteq \mathcal{B}_b$ be the set of basis elements such that $\mathbf{b}_1 = \bigvee_{s \in B_1} s$ and $\mathbf{b}_2 = \bigvee_{s \in B_2} s$. Since $\mathbf{b}_1 \neq \mathbf{b}_2$, there must be a $\mathbf{b} \in B_2$ such that $\mathbf{b} \notin B_1$. Because of the ordering of the basis, we have $\mathbf{b}_1 < \mathbf{b} \leq \mathbf{b}_2$. The basis \mathcal{B}_b is dense in \mathcal{D}_b . Moreover, \mathcal{D}_b is complete. Let $B \subseteq \mathcal{D}_b$ then $\bigvee_{s \in B} s$ is in \mathcal{D}_b and, since it is the narrowest statement that is broader than any statement in B , it is the supremum of B . As \mathcal{D}_b is complete and has a countable dense subset so does X and therefore it has a continuous order. \square

Another way to think about the verifiable statements of a domain characterized by a continuous quantity is in terms of finite but arbitrarily small precision. That is, when we measure a continuous quantity we can verify statements of the form “the value is 1 ± 0.5 ”. The verifiable statement is in terms of a range and the extremes are rational numbers. It is instructive to know, then, that the topology for the real numbers can be generated by those types of statements. In terms of references this means that all our verifiable statements could be expressed as verifying that the value is between two references from a countable set of possible references. This, again, maps well to what one can and does do in scientific practice.

Mathematically, we call the standard topology on the real numbers the one generated by the open intervals between rational numbers, and we can show that this is exactly the order topology. Moreover, every verifiable set is the disjoint union of open intervals.

Definition 3.54. We call **standard topology** on the real numbers \mathbb{R} the one generated by the collections of sets $\mathcal{B} = \{(a, b) \in \mathbb{R} \mid a, b \in \mathbb{Q}\}$ of all open intervals between rational numbers \mathbb{Q} .

Proposition 3.55. The order topology on the real numbers is the standard topology.

Proof. To show that they are equivalent, we show that the basis of one generates the basis of the other. Let $a, b \in \mathbb{Q}$ be two rationals. The sets $(-\infty, b)$ and (a, ∞) are in the basis of the order topology. Their intersection is the set (a, b) of the standard topology. The basis of the standard topology can be generated by the order topology.

Conversely, let $a \in \mathbb{R}$ be a real number. Let $\{U_i\}_{i \in I}$ be the collection of all sets $U_i = (a_i, b_i)$ such that $a_i, b_i \in \mathbb{Q}$ and $a < a_i$. These sets are in the basis of the standard topology. We have $\bigcup_{i \in I} U_i = (a, \infty)$. In the same way, let $b \in \mathbb{R}$ be a real number. Let $\{V_j\}_{j \in J}$ be the collection of all sets $V_j = (a_j, b_j)$ such that $a_j, b_j \in \mathbb{Q}$ and $b_j < b$. These sets are in the basis of the standard topology. We have $\bigcup_{j \in J} V_j = (-\infty, b)$. The basis of the order topology can be generated by the standard topology. \square

Proposition 3.56. *Let $U \in \mathbb{T}_{\mathbb{R}}$ be a set in the standard topology on the reals. Then $U = \bigcup_{i=1}^{\infty} V_i$ is the countable disjunction of open intervals $V_i = (a_i, b_i)$ where $a_i, b_i \in \mathbb{R} \cup \{-\infty, \infty\}$.*

Proof. The set U can be expressed as the union of some collection $B \subseteq \mathcal{B}$ of open rational intervals. From B construct $B_1 \subseteq B$ by picking an element of B and keep adding any element that is not-disjoint from an element in B . If there are elements left over, construct $B_2 \subseteq B$ with the same procedure and continue until there are no elements left.

Take $V_1 = \bigcup_{V \in B_1} V$ the union of all elements of B_1 . Since we are taking the countable union of open rational intervals, and because of the overlap of these intervals, the result will be an open interval over the real numbers. Repeating for all B_i will give us a collection of disjoint open intervals $\{V_i\}_{i=1}^{\infty}$ such that $U = \bigcup_{i=1}^{\infty} V_i$. \square

In the previous section we saw how the ability to have consecutive references is linked to decidability. In the case of continuous quantities, the inability to have consecutive references means we cannot have decidable before or after statements, otherwise we could use them to create consecutive references. So, while for integer quantities references have immediate successors and predecessors and all statements are decidable, for a continuous quantity references do not have immediate successors and predecessors and the statements are only verifiable.

We stress here that it is the immediate successors of the references that matter, not the immediate successors of possibilities/values. Take the rationals, for example. As values, they do not have immediate successors or predecessors: their order is dense. But we can construct a reference with “the rational quantity is more than π ” as an after statement and a reference that has “the rational quantity is less than π ” as a before statement. Since π is not a rational number, there are no values in between: the two references are consecutive. This can never happen with the reals, because all limit values are possible values. So, while the rationals do not admit consecutive values, it is only the reals that never admit consecutive references.⁶

Proposition 3.57. *A naturally ordered experimental domain \mathcal{D}_X is characterized by a continuous quantity if and only if none of the verifiable statements are decidable except for the certainty and the impossibility.*

Proof. Let \mathcal{D}_X be a naturally ordered domain for which none of the verifiable statements are decidable except for the certainty and the impossibility. Then it is generated by a refinable set of aligned strict references R . Let $\mathbf{r}_1 = (\mathbf{b}_1, \neg\mathbf{b}_1 \wedge \neg\mathbf{a}_1, \mathbf{a}_1)$ be a reference aligned with the domain that admits a successor. Then \mathbf{a}_1 is a verifiable statement that is not impossible and therefore it is not decidable given the properties of \mathcal{D}_X . This means there can't be an immediate successor for \mathbf{r}_1 : by 3.31 the before statement of the immediate successor would have to be equivalent to $\neg\mathbf{a}_1$ which is not verifiable. Therefore if \mathbf{r}_3 is a reference such that $\mathbf{r}_1 < \mathbf{r}_3$, we can find yet another reference \mathbf{r}_2 aligned with the domain such that $\mathbf{r}_1 < \mathbf{r}_2 < \mathbf{r}_3$. The order on R is dense and by 3.53 the domain is fully characterized by a continuous quantity.

⁶Mathematically, this construction corresponds to the Cauchy limits or the Dedekind cuts that one uses to construct the reals. The idea is that experimental domains already have them built into their structure in a way that corresponds better to physical concepts.

Conversely, let \mathcal{D}_X be a domain characterized by a continuous quantity (U, \leq, q) . Take a contingent verifiable statement $s \in \mathcal{D}_X$ and a possibility $x_0 \in X$ such that $x_0 \approx s$. Consider the value $q_0 \in \mathbb{R}$ such that $q(x_0) = q_0$. Because s is verifiable, it will correspond to a set V of the topology of the quantity, which by 3.56 is the union of open intervals. Since $x_0 \approx s$ then $q_0 \in V$. Find the interval (q_1, q_2) that contains q_0 . We can find $q_4, q_5 \in (q_1, q_2)$ such that $q_4 < q_0 < q_5$. We'll also have $\hat{q} \in V$ for all $q_4 < \hat{q} < q_5$. That is, for any value that is compatible with s we can find a contiguous finite interval surrounding the value with all elements compatible with s .

Consider now the statement s . Because it is contingent, it will not be compatible with at least one possibility and therefore one value in U . The set V will contain some interval with at least one finite endpoint q_0 . Since it is the endpoint of an open interval, we cannot find $q_1 < q_0 < q_2$ such that $\hat{q} \notin V$ for all $q_1 < \hat{q} < q_2$. That is, there is at least one value incompatible with s that does not admit a finite interval surrounding it with all elements incompatible with s .

Now consider $\neg s$. There will be at least one value compatible with $\neg s$ that does not admit a finite interval surrounding it that is all compatible with $\neg s$. But since this cannot happen for a verifiable statement, then $\neg s$ is not a verifiable statement and s is not decidable. \square

The integers and the reals, then, are the only two possible orders defined experimentally that are, in a sense, regular. That is, the order relationships look the same no matter where you are in the order. You have an immediate successor (or not) regardless of what reference you have. You can only put finitely many references (or not) between any two references. For any other ordering, instead, some references will have an immediate successor and some won't. In this sense integers and reals are very special and that is why they are fundamental in physics.

3.6 When ordering breaks down

As we have identified the necessary and sufficient conditions to define an order from experimental verification, we can reflect on whether these are achievable in practice or they are idealizations. We need to be able to create references that are strict (before/on/after mutually exclusive), aligned (references can be fully before or after each other) and refinable (overlapping references can be divided into finer sequential ones). Are these always reasonable expectations?

In the case of discrete quantities, they indubitably are. The domain is decidable so the possibilities themselves are verifiable statements. We can actually confirm that *"there are 3 ducks on the table"*. Therefore it is clear that this can be and is achieved.

In the case of continuous quantities, instead, things are a lot more problematic. In some cases, you start with what you think is a continuous quantity but then you realize that it was a discrete one. For example, an amount of water seems continuous but if you keep refining your references you see that it consists of discrete molecules. That is, you can't really go on refining references indefinitely. But that's not the only way the continuous order can break down.

Suppose you want to refine the marks on a ruler over and over, getting more and more precise position measurements. You start with 1 mm thick lines, you reduce them to 0.1 mm and make more of them and so on. At some point you'll reach single molecules or single atoms,

but those have spatial extension as well. We can reach fundamental particles, but those too have an extent (i.e. their wavefunctions can overlap). We can imagine using more wavelike features, but the spatial resolution will be linked to higher and higher energies through higher wave-numbers. So it is not clear we continue having ever finer references.

On a similar note, if a reference is, in the end, realized through a set of particles, it stands to reason that if we reduce the number of particles the reference will become finer. If single particles are the finest reference, we don't have that many choices. But how are we going to be able to tell our references apart, since, for example, all electrons are indistinguishable from each other? How can we place them ever increasingly close to each other so that they don't scatter and switch place?

Moreover, it is not clear how our references can be strict. That is, that the object we measure is always either before, on or after the reference. If the object we are measuring is of smaller extent than our references then we can reasonably pretend they are strict. But if both the references and what we are measuring are single particles, it would seem we have a problem.

It would appear, then, that the conditions for a continuous quantity can never really be ultimately met. And if those conditions can't be met, it's not that we don't have the real numbers: we don't have ordering at all. Strictness, alignment and refinability are requirements for order in general. And if space and time can't be truly given an ordering, other derived quantities, like velocity, acceleration, mass, energy and so on would inherit the problems.

In both cases the real numbers are just an approximation we can make by pretending we can get finer and finer strict and aligned references. This may be contrary to the way many people see the relationship between mathematical and physical objects. Some may feel that the geometric description, with its infinite precision, is the perfect one while the physical one, with the inherent measurement bounds, is the less precise one. Actually, it is quite the opposite: the bounds of a measurement better qualify our description and knowledge while the geometrical description provides a simplified, idealized and therefore less precise account. In other words, 3.14 ± 0.005 is an exact physical description while π is the approximation.

We again stress the fact that the approximation can break down in different ways. It may break down because we reach a finest element or it may break down because we do not have finer, strict and aligned references anymore, which is what we expect to happen for space and time. We'd have a structure where coarse references, at some point, are to a good approximation refinable/strict/aligned and therefore approximately ordered while the really fine ones are not. Note that a lot of physical ideas and mathematical tools rest on the idea that there is a well defined ordering. Causality and deterministic motion require that time is linearly ordered. Differentiation and integration also require the reals to be linearly ordered. All these tools, then, need to be fundamentally reshaped. The general theory, then, is telling us that there is a lot of work that needs to be redone if we are to construct a physical theory that works in those regimes.

3.7 Summary

In this chapter we have seen how our general mathematical theory of experimental science handles properties and quantities. These are what we typically use in practice to distinguish between the possible cases and are what we measure experimentally.

Mathematically, each possibility is mapped to an element of a topological space whose verifiable sets correspond to verifiable statements. This construction maps to what happens in manifolds, where the points of the space are in one-to-one correspondence with a suitable Euclidean space. Our construction is more general and works for non-numeric properties as well.

We have seen that quantities are particular types of properties characterized by a linear order. In this case the topology is the order topology given by the linear order, which represents the ability to experimentally compare two different values and tell which one is greater. The property ordering theorem and the domain ordering theorem give us the necessary and sufficient conditions under which an experimental domain is fully characterized by a quantity.

To construct a system of measurement for a quantity, we saw that all we need is to define a set of references: objects that partition the cases into a before and after. In general, though, references can overlap as they will have some physical extent and may not be aligned. The reference ordering theorem tells us that an ordering on the possibilities emerges only if the references are refinable (we can always break apart overlapping references), aligned (the before and not-after statements are ordered by narrowness) and strict (the value is always either before, on or after the reference).

We defined discrete quantities as the ones that can be associated with integers and continuous quantities as the ones that can be associated with real numbers. Physically, the defining characteristic of the first is that between two references we can only put finitely many references while the defining characteristic of the second is that between two references we can always put infinitely many. Mathematically, the ordering in the second case is automatically complete because experimental domains will already contain all the limits in the form of countable disjunctions.

It is important to note that the requirements for continuous quantities cannot really be physically realized. Continuous quantities, then, should really be thought of as an idealization: the limit of an infinite process of subdivision. To go past the idealization, either we lose the idea of having infinitely many references between two, in which case we revert to discrete quantities, or we lose the ordering altogether.

Part III

Blueprints for the work ahead

This part is dedicated to preliminary ideas and work in progress, meaning that it includes conjectures, current thinking and approaches that may or may not turn out to be good ideas. As the work matures, sections and chapters from this part will be merged into the main part.

The first chapter will give an overview of rough general ideas within reverse physics. The second chapter will do the same for physical mathematics. The subsequent chapters capture an in-depth stab at a particular problem that, if solved, it would likely turn out to be a standalone chapter in either reverse physics or physical mathematics.

Chapter 1

Reverse Physics

1.1 Classical mechanics

The work on classical mechanics is considered mostly concluded, in the sense that suitable initial assumptions have been identified. There are still a few open issues, such as the case of variable mass, the generalization of the directional DOF to the relativistic case, or clarifying the nature or the generalization to infinite DOFs (i.e. field theories).

Curvature for particle dynamics

The assumption of kinematic equivalence already gives us relativistic Hamiltonians. Does it also give us a relationship between the curvature of the metric tensor and the forces acting on the particles?

The setup is the following. Suppose we have two vectors in the extended phase space $d\xi^a = \{dq^\alpha, 0\}$ and $d\nu^a = \{0, dp_\alpha\}$. Using the symplectic form we have the invariant $d\xi^a \omega_{ab} d\nu^b = dq^\alpha dp_\alpha$. Under the kinematic assumption we have $dq^\alpha = dx^\alpha$ and $dp_\alpha = mg_{\alpha\beta} du^\beta + \mathbf{q}A_\alpha$. We have $d\xi^a \omega_{ab} d\nu^b = dx^\alpha mg_{\alpha\beta} du^\beta + dx^\alpha \mathbf{q}A_\alpha$.

Since the two terms have to match at each point and the symplectic form has the same components at each point, can we constrain the change of the components of $g_{\alpha\beta}$? The general idea would be that components of $g_{\alpha\beta}$ may have to change in space/time coordinately with A_α as to make $d\xi^a \omega_{ab} d\nu^b = dq^\alpha dp_\alpha$ remain the same. Note that derivatives in q^α are taken at constant p_α while derivatives in x^α are taken at constant u^α .

1.2 Thermodynamics

Process entropy

The key to recover thermodynamics is finding a definition of entropy that applies in very general cases and recovers the usual definition. Instead of using the logarithm of the count of states, we use the logarithm of the count of possible evolutions. That is, the ways a system can evolve under a specific process. The entropy of the system is automatically relativistic (i.e. we are essentially counting “worldlines” of the overall system in its state space) and is process dependent (i.e. contextual).

In the case of deterministic and reversible evolution, the count of states is equivalent to the count of evolutions, and therefore the usual definition is recovered. In the case of stochastic steady state over continuous time, that is when the probability distribution stabilizes, the

states will traverse infinitely many states within a small time difference dt . The count of evolutions, then, can be shown to reduce to the permutations of infinite sequences which recovers the Gibbs/Shannon entropy.

As for the behavior of entropy, the idea is that for a specific process, the state at a particular time identifies a set of possible evolutions. This would be the entropy of that state. Over the continuum, where states are points, the entropy would become a density of the count of evolutions. In essence, the entropy of a system at a particular time tells us how much or how little the evolution is constrained. In other words, it tells us how much the system is expected to fluctuate. As time evolves, the state changes, and the count of evolutions changes as well. If the evolution is deterministic, the evolutions can never split, in the sense that all the evolutions that end up in a particular state must all go to another state. This means that for a deterministic process the count can never decrease. If the evolution is deterministic and reversible, then the count must stay the same. This recovers the feature of entropy to be a non-decreasing quantity, which is conserved during reversible processes.

If the evolution allows equilibria, the evolutions will concentrate around states of equilibria. Given that states cannot go out of equilibrium once it has been reached, the count of evolutions is maximized at equilibrium. This recovers another feature of entropy.

Lastly, if two systems are independent, the way one evolves does not constrain the other. The total count of evolutions, then, is simply the product of the count of evolutions of the two systems. Since the entropy is the logarithm of the count of evolutions, it sums over independent systems. This recovers the last property of entropy.

Equation of state

If we study the space of equilibria, each state will have a well defined entropy. Therefore we have an equation of state $S(\xi^a)$ where ξ^a form a set of variables that fully identify the state. Moreover, as noted before, entropy is additive under system composition of independent systems.

In a process with equilibria different evolutions must converge to the same final state, which means the process entropy increases and is maximized at equilibria. This gives the general idea that entropy increases during an irreversible process. These results are therefore valid in general, no matter what type of system is being described.

To find thermodynamics specifically, we need an additional set of assumptions. First, all states are equilibria. Second, all state variables ξ^a are additive under system composition. Third, one of the them, which we call internal energy U , is conserved under any evolution, including irreversible evolution. We can then write the equation of state as $S(U, x^i)$ and define the following quantities:

$$\begin{aligned}\frac{\partial S}{\partial U} &= \beta = \frac{1}{k_B T} \\ \frac{\partial S}{\partial x^i} &= -\beta X_i\end{aligned}\tag{1.1}$$

We can then express the differentials as:

$$\begin{aligned}dS &= \frac{\partial S}{\partial U} dU + \frac{\partial S}{\partial x^i} dx^i = \beta dU - \beta X_i dx^i \\ dU &= T k_B dS + X_i dx^i\end{aligned}\tag{1.2}$$

This is essentially Gibbs' approach to thermodynamics.

Thermodynamic laws

To recover the laws, we need a few more definitions. We define a reservoir R as a system for which the internal energy U_R is the only state variable and the state entropy S_R is a linear function of U_R . That is, $\frac{\partial S_R}{\partial U_R} = \beta_R = \frac{1}{k_B T_R}$ is a constant. We call heat $Q = -\Delta U_R$ the energy lost by the reservoir during a transition.

We define a purely mechanical system M as a system for which the state entropy is zero for each state. That is, $S_M(U_M, x_M^i) = 0$. We call work $W = \Delta U_M$ the energy acquired by a purely mechanical system during a transition.

Now, consider a composite system made of a generic system A , a reservoir R and a purely mechanical system M . Consider a transition where we go to a new equilibrium. Since energy is additive under system composition, let us call U the total energy. Since energy is conserved we have:

$$\begin{aligned}\Delta U = 0 &= \Delta U_A + \Delta U_R + \Delta U_M = \Delta U_A - Q + W \\ \Delta U_A &= Q - W\end{aligned}\tag{1.3}$$

Since entropy is extensive, let us call S the total entropy. Since the process is going to an equilibrium, the entropy can only increase. We have:

$$\begin{aligned}0 \leq \Delta S &= \Delta S_A + \Delta S_R + \Delta S_M = \Delta S_A + \beta_R \Delta U_R + 0 = \Delta S_A + \frac{-Q}{k_B T_R} \\ k_B \Delta S_A &\geq \frac{Q}{T_R}\end{aligned}\tag{1.4}$$

1.3 Quantum mechanics and irreducibility

Quantum mechanics can be recovered by swapping reducibility with irreducibility as shown in diagram 1.1, which can be used as a guide throughout this section.

The assumptions lie on the left column. Each assumption leads to one or two key insights that progressively lead to the physical concepts in the middle column. Each of these is then mapped to its corresponding formal framework on the right. Note that “quasi-static process” and “conserved density” both independently lead to the same result of “unitary evolution”.

Irreducibility

The state space of quantum mechanics can be recovered under the:

Assumption V (Irreducibility). *The state of the system is irreducible. That is, giving the state of the whole system says nothing about the state of its parts.*

Under this assumption the state of the system is automatically an ensemble over the state of the parts as preparation of the whole leaves the parts unspecified. For the same reason, the entropy of these ensembles must be the same, or some ensembles would provide more or less information about the parts. The whole task, then, is to characterize these ensembles without making specific assumptions on the parts.

Let \mathcal{C} be the state space of the irreducible system. Let us call fragment a part of the irreducible system. The state of a fragment will be associated with a random variable uniformly distributed over the possible fragment states. As discussed in the context of classical

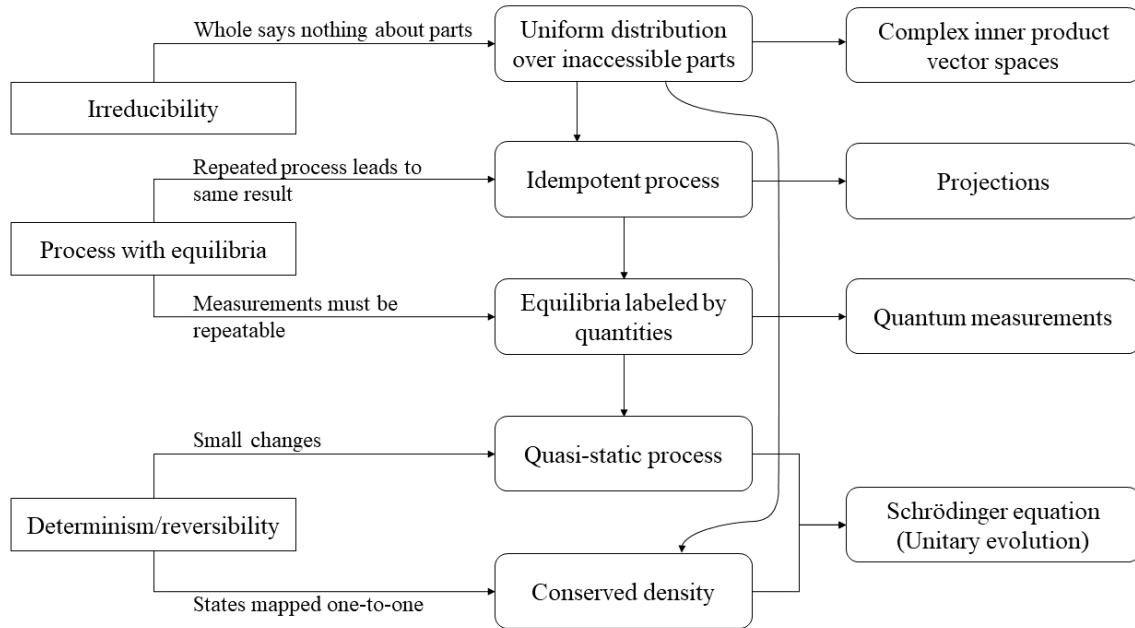


Figure 1.1: Assumptions for quantum mechanics

mechanics, distributions over states must be invariant and symplectic manifolds are the only manifolds over which invariant distributions can be defined. As we cannot say anything about the state of the fragments, the dimensionality of this manifold must be irrelevant as long as it is even dimensional. For simplicity, we can choose a two-dimensional one. Therefore we are interested in the space of bi-dimensional uniform distributions formed by a pair of two random variables A and B .

The values of the variables themselves are not relevant, as they are not physically accessible by assumption. However, the size of the system $\mu = \int \rho dA \wedge dB$ is relevant. Without loss of generality, we can rescale A and B such that the density ρ is not only uniform but unitary: $\rho = 1$. This way the size of the system is directly proportional to the area covered by the random variables. In other words: the more fragments there are, the more each fragment can swap its state with another without changing the whole, the more uncertainty there is on the state of the fragment, the higher the variance of the random variables.

Since only linear transformations will preserve the uniform distribution, we look to those. These are translations, stretches and rotations. Translations do not lead to other physically distinguishable states since the exact values of A and B are not physically accessible. Stretching of the distribution will correspond to an increase of the size of the system, which is physically accessible. However, only the stretching of the area is of interest. So, without loss of generality, we can set $\sigma_A = \sigma_B = \sigma$ and we have $\mu \propto \sigma^2$. Rotations just change the correlations which, by themselves, are not physically accessible. However, under addition the correlations still result in differences in variance and, indirectly, the size of the system, and therefore are physically interesting. The space of transformations is therefore given by two

parameters a and b such that:

$$\begin{aligned} C &= aA + bB \\ D &= -bA + aB \end{aligned} \tag{1.5}$$

Equivalently, we can use the complex number $c = a + ib$ to characterize the transformation, which we can note as $\tau(c)$. The increase/decrease in size is given by $a^2 + b^2 = (a - ib)(a + ib) = c^*c$ and the change in correlation is given by the Pearson correlation coefficient $\rho_{A, \tau(c)A} = \cos \arg c$.

Putting it all together, we can characterize the state space \mathcal{C} with a complex vector space. The linear combination represents the mixing of the different stochastic descriptions. Two vectors that only differ by a total phase are physically equivalent since a global change of correlation does not change the distribution.

We can define a scalar product $\langle \cdot | \cdot \rangle$ where the square norm induced corresponds to the size of the system (or equivalently to the strength of the random variable) and the phase difference corresponds to the correlations (the Pearson correlation coefficient). To see this, note the formal equivalence between the variance and norm rules under linear composition:

$$\begin{aligned} \sigma_{X+Y}^2 &= \sigma_X^2 + \sigma_Y^2 + 2\sigma_X\sigma_Y\rho_{X,Y} \\ |\psi + \phi|^2 &= |\psi|^2 + |\phi|^2 + 2|\psi||\phi|\cos(\Delta\theta) \end{aligned} \tag{1.6}$$

The quadratic form, again, reflects the fact that the size of the system is proportional to the variance of a random variable. Since the size of the system is fixed, we use unitary vectors to represent actual states. The state of the system, then, is represented by a ray in a complex inner product space.

Lastly, we need to define an expectation operator that returns the average value for each physical quantity. This operator will have to be linear under linear combination of quantities:

$$E[aX + bY|\psi] = aE[X|\psi] + bE[Y|\psi]. \tag{1.7}$$

It will not be linear under linear combination of states:

$$E[X|\psi + \phi] \neq E[X|\psi] + E[X|\phi]. \tag{1.8}$$

Yet, it will have to be proportional to the increase in size and invariant under a total change in correlation: $E[X|\tau(c)\psi] = c^*cE[X|\psi]$. This leads us to associate to each physical quantity a linear Hermitian operator X where $E[X|\psi] = \langle \psi | X | \psi \rangle$. An eigenstate ψ_0 of X corresponds to a state where all the elements of the ensemble have exactly the same value. That is, $E[(X - \bar{x})^2|\psi_0] = 0$.

Note that an inner product space can always be completed into a Hilbert space. This may, however, bring in objects that may not correspond to physical objects (i.e. infinite expectation for some quantities). In general, we believe it is better to regard the (possibly incomplete) inner product space as the physical state space and regard the completion as a mathematical device for calculation. For example, the Schwartz space seems more physically meaningful than the standard L^2 space as it gives finite expectation of all polynomials of position and momentum and, moreover, it is closed under Fourier transform.

Process with equilibria

The first type of process we consider is one with equilibria. The measurement process is recovered as a special case.

Assumption VI (Process with equilibria). *Given an initial ensemble (i.e. mixed state), the final ensemble is uniquely determined and remains the same if the process is applied again.*

Under this assumption, the process can be characterized by a projection operator. Let ρ_1 be the density matrix that characterizes a mixed state. Since the final mixed state must be uniquely determined by ρ_1 , it will be $\mathcal{P}(\rho_1)$ for some operator \mathcal{P} . Similarly, if ρ_2 is another initial mixed state, its final operator will be $\mathcal{P}(\rho_2)$. Note that, given any observable X the expectation $E[X|\rho_1] = \text{tr}(X\rho_1)$ is the trace of $X\rho_1$. Similarly $E[X|\mathcal{P}(\rho_1)] = \text{tr}(X\mathcal{P}(\rho_1))$.

We can always create statistical mixtures of the ensembles and we must have $E[X|a\rho_1 + b\rho_2] = aE[X|\rho_1] + bE[X|\rho_2]$ since these are classical mixtures. But since these are classical mixtures, the final state will also need to obey $E[X|a\mathcal{P}(\rho_1) + b\mathcal{P}(\rho_2)] = aE[X|\mathcal{P}(\rho_1)] + bE[X|\mathcal{P}(\rho_2)]$ for all possible X . Which means $\mathcal{P}(a\rho_1 + b\rho_2) = a\mathcal{P}(\rho_1) + b\mathcal{P}(\rho_2)$ Therefore the operator \mathcal{P} is a linear operator. Moreover, the process applied twice must lead to the same result, which means $\mathcal{P}(\mathcal{P}(\rho)) = \mathcal{P}(\rho)$ for any ρ . That is, $\mathcal{P}^2 = \mathcal{P}$. Therefore \mathcal{P} is a projection.

Suppose, now, that we want to measure a quantity X . We want the final outcome, the final ensemble, to be determined by the initial state, the initial ensemble. We also want the measurement to be consistent in the sense that, if it is repeated immediately after, it should yield the same result. Therefore the process will be a projection. We will also want that the process does not distort the quantity. That is, $E[X|\rho] = E[X|\mathcal{P}(\rho)]$. This means that the eigenstates of X will correspond to equilibria of the process. Moreover, subsequent measurements must give the same value, not just the same mixture. That is, if X_1 is the random variable after the first instance of the process and X_2 is the random variable after the second instance, $P(X_2 = x|X_1 = x) = 1$. This means that $E[(X - \bar{x})^2|\mathcal{P}(\rho)] = 0$ which means the eigenstates of X are the only equilibria.

The measurement process is therefore simply a special case of a process with equilibria.

Deterministic and reversible evolution

The second type of process we consider is one that is deterministic and reversible, which is the same as assumption DR.

Under this assumption, the process can be characterized by unitary evolution (i.e. the Schrodinger equation). There are multiple different ways to see this. The first relates to the more general idea that all deterministic and reversible processes must be isomorphisms in the category of states. Since the state space is an inner product space, the isomorphism is unitary evolution.

The second, is that if there is a set of quantities X_0 at time t_0 that fully identify the state (i.e. the state is the only eigenstate of those quantities), then there must be a corresponding set of quantities X_1 that fully identify the state at time t_1 . This means that the evolution maps basis to basis. Moreover, given the linearity of statistical mixtures, this will also mean that a statistical distribution over X_0 will have to map to the same distribution over X_1 . Therefore the evolution must map linear combinations of that basis to the same linear combination. The evolution is a linear operator. Since the total size of the irreducible system cannot change, the operator must be unitary.

The third, is by constructing a quasi-static process from processes with equilibria, much like one does in thermodynamics. The idea is that we have an infinitesimal time step, an initial state ψ_t and a final state ψ_{t+dt} . We want $P(\psi_{t+dt}|\psi_t) = 1$. This means that $|\langle\psi_{t+dt}|\psi_t\rangle|^2 = 1$. This can happen only if the difference between initial and final states is infinitesimal. That is, $\langle\psi_{t+dt}|\psi_t\rangle = 1 + \iota\epsilon dt$ where ϵ is a real number. Therefore, by convention, we can write $|\psi_{t+dt}\rangle = I + \frac{Hdt}{i\hbar}|\psi_t\rangle$ where H is a Hermitian operator.

Putting these perspectives together, time evolution is a unitary operator which can be written as $U = e^{\frac{H\Delta t}{i\hbar}}$. If we start in an eigenstate of X , that is $X|\psi_t\rangle = x_0|\psi_t\rangle$ we will end in an eigenstate $\hat{X}|\psi_{t+\Delta t}\rangle = x_0|\psi_{t+\Delta t}\rangle$ of another operator $\hat{X} = e^{\frac{H\Delta t}{i\hbar}}Xe^{-\frac{H\Delta t}{i\hbar}}$.

In fact:

$$\begin{aligned}
e^{\frac{H\Delta t}{i\hbar}}Xe^{-\frac{H\Delta t}{i\hbar}}|\psi_{t+\Delta t}\rangle &= e^{\frac{H\Delta t}{i\hbar}}Xe^{-\frac{H\Delta t}{i\hbar}}U|\psi_t\rangle \\
&= e^{\frac{H\Delta t}{i\hbar}}Xe^{-\frac{H\Delta t}{i\hbar}}e^{\frac{H\Delta t}{i\hbar}}|\psi_t\rangle \\
&= e^{\frac{H\Delta t}{i\hbar}}X|\psi_t\rangle \\
&= e^{\frac{H\Delta t}{i\hbar}}x_0|\psi_t\rangle \\
&= x_0U|\psi_t\rangle \\
&= x_0|\psi_{t+\Delta t}\rangle
\end{aligned} \tag{1.9}$$

This is consistent with assuming there is a quasi-static process that, at every t , has equilibria identified by $e^{\frac{H(t-t_0)}{i\hbar}}Xe^{-\frac{H(t-t_0)}{i\hbar}}$. Note that, unlike thermodynamics, the equilibria during the evolution are not set by external constraints but by the system itself. That is, X depends on the initial state of the system.

In this light, the measurement processes and the unitary processes can be seen as particular cases of the same type of processes, those with equilibria, which are defined as a black-box from initial to final state. This is consistent with the irreducibility assumption as the inability to describe the dynamics of the parts implicitly assumes that the dynamics of the parts is at equilibrium and sets a time-scale under which the further description of the system (i.e. non-equilibrium dynamics) would require describing the internal dynamics.

Chapter 2

Physical mathematics

This chapter presents the areas that still need to be covered to conclude the general mathematical theory of experimental science and a summary of the preliminary work done on them.

2.1 Experimental verifiability

This first part is already well developed and has been presented in chapters one to three. Possible improvements are discussed in section 2.4.

2.2 Informational granularity

The general goal of this part is to recover elements of measure theory, differential geometry, probability theory and information theory. The central theme is the ability to compare and then quantify the granularity of the description provided by different statements. The idea is to have a single unified structure which can be, in some cases, reduced to the more familiar mathematical structures.

Statement fineness

Conceptually, we want to be able to compare two statements to see which one provides a more refined description, which one provides more information. For this, we need to establish a new axiom.

Note that a theoretical domain \bar{D} comes with a partial order \leq that indicates whether one statement gives a **narrower**, more specific, description than the other. For example:

- “The position of the object is between 0 and 1 meters” \leq “The position of the object is between 0 and 1 kilometers”
- “The fair die landed on 1” \leq “The fair die landed on 1 or 2”
- “The first bit is 0 and the second bit is 1” \leq “The first bit is 0”

In these cases, the first statements are “contained” in the second ones, which are more general.

We need to define an additional preorder $\leq: \bar{D} \times \bar{D} \rightarrow \mathbb{B}$ that compares two statements and tells us if the first provides a description with finer granularity than the second. Saying $s_1 \leq s_2$ means that the description provided by s_1 is **finer**, gives more information, is more precise, than the description provided by s_2 . For example:

- “The position of the object is between 0 and 1 meters” \leq “The position of the object is between 2 and 3 kilometers”
- “The fair die landed on 1” \leq “The fair die landed on 3 or 4”
- “The first bit is 0 and the second bit is 1” \leq “The third bit is 0”

In these cases, the first statement may not be contained or overlap with the second. The existence of this operator and its property would be an additional axiom. Fineness is a preorder, rather than an order, because it does not satisfy antisymmetry: if $s_1 \leq s_2$ and $s_2 \leq s_1$ then it is not necessarily true that $s_1 \equiv s_2$. In that case, we will say that the two statements are **equigranular**, noted $s_1 \doteq s_2$.

Note how statements about geometry, probability and information all satisfy the same concept. In fact, each of these structures will generate a preorder on the statements. The general question is what are the necessary and sufficient conditions on the preorder to be able to recover those structures.

Measure theory

Conceptually, a measure allows one to assign a size to a set. For us, a theoretical set is really a statement, so we want to assign sizes to statements that represent the coarseness of the description they provide.

The construction should, roughly, proceed as follows. Let $\bar{\mathcal{D}}_X$ be a theoretical domain. We select a unit statement $u \in \bar{\mathcal{D}}_X$. We define, in some way, the set $\bar{\mathcal{D}}_u \subseteq \bar{\mathcal{D}}_X$ which contains all statements that are comparable to u . We then try and construct a measure $\mu_u : \bar{\mathcal{D}}_u \rightarrow \mathbb{R}$ such that $\mu_u(u) = 1$. By a measure, we mean that μ_u is additive over incompatible statements (i.e. disjoint sets of possibilities). That is, if $s_1 \not\# s_2$, we have $\mu_u(s_1 \vee s_2) = \mu_u(s_1) + \mu_u(s_2)$. We want the measure to respect the fineness preorder, to be monotonic. That is, if $s_1 \leq s_2$ then $\mu_u(s_1) \leq \mu_u(s_2)$.

Originally, we thought that these measures would have to be always additive and therefore we starting adding suitable axioms on fineness. However, we realized that, in the context of quantum mechanics, the measure cannot be additive if it has to agree with the von Neumann entropy. Worse, it is not even monotonic (i.e. a broader statement is not necessarily coarser). More conceptual work needs to be done to understand the issue.

Note that we have essentially one measure for each equivalence class defined by fineness. This is intended. One reason a single measure is not sufficient for our work is because we need to compare statements of “different infinities”. If we have a single measure, we can only compare objects with a finite measure. All objects with zero measure (or infinite measure) are indistinguishable. For example, we want to say:

- $s_1 =$ “The horizontal position of the object is exactly 0 meters”
- $s_2 =$ “The horizontal position of the object is exactly 1 or 2 meters”
- $s_3 =$ “The horizontal position of the object is between 0.5 and 1.5 meters”
- $s_4 =$ “The horizontal position of the object is between 1.5 and 3.5 meters”
- $s_1 \leq s_2 \leq s_3 \leq s_4$
- $s_1 \not\# s_2 \not\# s_3 \not\# s_4$

Fineness may also capture the concept of physical dimension. In fact, two descriptions in the same units are “finitely comparable” in the sense that one gives a finer description than the other by a finite factor. Descriptions of different units are either “infinitely comparable”

(e.g. areas are always bigger than lengths) or not comparable (e.g. position and momentum). Consider a two dimensional phase space of a classical system. Points should be comparable and in fact should be equigranular \doteq so that we can compare sets of finitely many points. Areas are also comparable to each other, and are comparable to points (i.e. they are infinitely bigger). However, vertical lines (i.e. ranges in momentum alone) are not comparable to horizontal lines (i.e. ranges in position alone). Symplectic geometry, in fact, gives a size to areas and not to lines. Mathematically, this should be clarified when one is trying to define the domain of the measure \bar{D}_u .

Probability

Conceptually, probability is recovered as a measure restricted to a particular subset. The idea is that you take two statements, such as “*the die landed on 2*” given that “*the die has 6 sides and it is fair*”, and you ask what fraction of the possibilities compatible with the second is also compatible with the first. This defines the conditional probability.

Let $s_1, s_2 \in \bar{D}$ be two theoretical statements. Then the probability of s_2 given s_1 is

$$P(s_2|s_1) = \mu_{s_1}(s_1 \wedge s_2) = \frac{\mu_u(s_1 \wedge s_2)}{\mu_u(s_1)} \quad (2.1)$$

which quantifies the fraction of possibilities compatible with s_1 that are also compatible with s_2 .

If we take the certainty \top as a unit, we have a probability measure for the whole space. However, since we can take different statements as a unit, we will be able to distinguish between the following cases:

- $P(\text{“}n \text{ is odd”} \mid \text{“}n \text{ is picked fairly from all integers”}) = 1/2$
- $P(\text{“}n \text{ is between } 0 \text{ and } 9\text{”} \mid \text{“}n \text{ is picked fairly from all integers”}) = 0$
- $P(\text{“}n \text{ is } 3\text{”} \mid \text{“}n \text{ is picked fairly from all integers”}) = 0$
- $P(\text{“}n \text{ is } 3\text{”} \mid \text{“}n \text{ is between } 0 \text{ and } 9\text{”} \wedge \text{“}n \text{ is picked fairly from all integers”}) = 1/10$

Differentiability

We want to construct a notion of differentials and differentiability that is the same for all spaces, even infinite dimensional ones. When introducing derivatives, this is typically done by taking limits of differences, and therefore differentiability is the existence of those limits. In differential topology, this notion is used to define differentiability of manifolds in terms of differentiability of coordinates, and then differentials are defined as linear functions of vectors. That is, the differentials defined on the coordinates of a particular chart are technically not the same objects as the differentials defined on the space.

The idea is to define differentiability on the vector space structure alone. That is, given two vector spaces V and W , a map $f : V \rightarrow W$ is differentiable if it becomes linear in the neighborhood. We would first define a differential as a sequence of vectors $\{v_i\}_{i=1}^{\infty} \in V$ such that there exists a vector $t \in V$ and a sequence of non-zero elements $\{a_i\}_{i=1}^{\infty} \in \mathbb{R}$ that converges to 0 for which

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

We call t the **tangent vector** of the differential and $\{a_i\}_{i=1}^{\infty}$ its **convergence envelope**. Note that, given a sequence v_i , these are not unique. We note $dv[a_i t]$ the differential with its

tangent vector and convergence envelope. One can show that every differential can be written as $v_i = a_i t_i$ where t_i converges to t .

We can now study how a map $f : V \rightarrow W$ maps differentials. Given a sequence $\{v_i\}_{i=1}^{\infty} \in V$, we can define $w_i = f(v_i)$. If, additionally, we have a differential $dv[a_i t]$, we can define the sequence $\{w_i\}_{i=1}^{\infty} = \{f(v_i + a_i t_i) - f(v_i)\}_{i=1}^{\infty}$. Now, the observation here is that if the map is linear, the sequence $\{w_i\}_{i=1}^{\infty}$ will be a differential with tangent vector $f(t)$ and convergence envelope a_i . But any map that is locally linear will have the same property, given that differentials are local objects. Therefore we say f is differentiable at $v \in V$ if there exists a map $d_v f|_{v_0} : V \rightarrow W$ such that $\{w_i\}_{i=1}^{\infty} = d_v f|_{v_0}(t)$.

From a preliminary study, this would work on any vector space, regardless of dimension or field (i.e. real, complex, rational, ...).

Differential geometry/geometric measure theory

In the reverse physics chapter about classical mechanics we have seen that forms can be understood as modeling additive functionals of subregions. We need to connect those ideas to the rest of the formal framework.

Conceptually, we want to assign quantities to regions instead of points. If we assume these quantities are additive, the idea is that we can decompose them into the sum of infinitesimal contributions at each point. Therefore the differential objects exist as the limit of infinitesimal decomposition. This, again, reflects the overall spirit of the project that compels us to start from physically well defined entities (in this case the quantities associated with finite regions) and derive the theoretical ones (in this case the infinitesimal contributions that are integrated).

Let \bar{D}_X be a theoretical domain and $U \in \Sigma_X$ a theoretical set. This represents the region associated to our measurement. Let \bar{D}_Y be a theoretical domain and $R \in \Sigma_Y$ a theoretical set. This represents the possible values found. Our starting point consists of statements like:

- “the amount of mass inside volume U is within range R ”
- “the force applied to surface U is within range R ”
- “the energy used to move the object along the line U is within range R ”

These are finite precision statements of a quantity associated to a region of finite size.

The first step is to group statements within the same region U into subdomains $\bar{D}_{U \rightarrow Y}$. We can then show how the possibilities for each $\bar{D}_{U \rightarrow Y}$ reduce to statements like:

- “the amount of mass inside volume U is precisely y ”
- “the force applied to surface U is precisely y ”
- “the energy used to move the object along the line U is precisely y ”

These are infinite precision statements of a quantity associated to a region of finite size. We define $S \subseteq \Sigma_X$ as the type of region (i.e. volumes vs surfaces vs lines) upon which the functional is defined and therefore we have a functional $f : S \rightarrow Y$ which tells us the exact value of the quantity in each region.

Then we study the case where f is a real linear k -functional, meaning:

- the possibilities X are identified by a set of real values; that is, X with the natural topology is a manifold
- the domain is all k -dimensional surfaces S^k ; that is, the submanifolds of dimension k

- the co-domain is the reals; so we have $f : S^k \rightarrow \mathbb{R}$
- the functional is additive over disjoint sets; that is, $F(U_1 \cup U_2) = F(U_1) + F(U_2)$ if $U_1 \cap U_2 = \emptyset$
- the functional commutes with the limit; that is, $\lim_{i \rightarrow \infty} F(U_i) = F(\lim_{i \rightarrow \infty} U_i)$

Under these conditions (and possibly others) one can express the functional as a sum of infinitesimal contributions. That is, $f(U) = \int_U \omega(dU)$, where ω represents a suitable k -form.

Note that there is not a unique way to perform this decomposition. For example, if $f(U)$ is the total mass in the volume, $\omega(dU)$ is the density in the infinitesimal volume. If we change the density at a single point, the integral does not change and only the integral is physical. These are the types of issues that still need to be solved.

Stokes' theorem and exterior derivatives. One interesting application of this viewpoint is that we can understand things like Stokes' theorem, exterior derivative and the difference between closed and exact forms directly on the finite functionals.

Let $\partial : S^k \rightarrow S^{k-1}$ be the boundary operator that, given a surface σ^k , returns the boundary $\partial\sigma^k$ which is of dimension $k - 1$. We have $\partial\partial\sigma = \emptyset$ for any surface of any dimensionality.

Let F_k be the space of linear k -functionals. We can define the boundary functional operator $\partial : F_k \rightarrow F_{k+1}$ such that $\partial f(\sigma) = f(\partial\sigma)$. That is, given a functional that acts on k -surfaces we can always construct one that acts on $k+1$ -surfaces by taking the boundary of the $k+1$ -surface and giving it to the first functional. Note that $\partial\partial f(\sigma) = \partial f(\partial\sigma) = f(\partial\partial\sigma) = f(\emptyset) = 0$, so the boundary functional of the boundary functional is the null functional, the one that returns zero for every k -surface. What we should be able to prove is that if ω is the k -form associated with f , $d\omega$ is the $k + 1$ -form associated with ∂f . In other words, Stokes' theorem essentially becomes a definition of the boundary functional and the calculation of the expression for $d\omega$.

We say a surface is contractible if it can be reduced to a point with a continuous transformation. A functional is closed if it is zero for all closed contractible surfaces. It is exact if it is zero for all closed surfaces. All boundary functionals are exact since $\partial f(\partial\sigma) = f(\partial\partial\sigma) = 0$. The form associated to a closed functional will be closed while the form associated to an exact functional will be exact.

2.3 States and processes

The general goal of this part is to give general definitions of states and processes that are always valid and are captured by a fundamental mathematical framework. Different theories would then specialize these basic definitions for different circumstances.

Processes

A process is an experimental domain \mathcal{P} that contains all the possible statements of the systems under study for all possible times. We call evolutions the possibilities E of the domain, as they represent the complete description of all systems at all times.

We define a time parameter $t \in T \subseteq \mathbb{R}$. We group all statements relative to a system of interest at a particular time into a time domain \mathcal{D}_t . We call snapshots the possibilities X_t of each time domain. A possible trajectory is a sequence $\{x_t\}_{t \in T}$ such that $x_t \in X_t$ for all $t \in T$ and $e \leq \bigwedge_{t \in T} x_t$ for some $e \in E$. That is, there is an evolution for which the system will be described by that sequence of snapshots.

A process is deterministic if for all possible trajectories $x_{t_0} \preceq x_{t_1}$ for all $t_0 \leq t_1$. A process is reversible if for all possible trajectories $x_{t_1} \preceq x_{t_0}$ for all $t_0 \leq t_1$. Recall that narrowness between the possibilities of two domains means there is an experimental relationship. Therefore, if the process is deterministic, we can write a causal relationship $f : X_{t_0} \rightarrow X_{t_1}$ such that $x_0 \preceq f(x_0)$.

Once we derive a measure $\mu_u : \bar{\mathcal{P}}_u \rightarrow \mathbb{R}$, we can define the evolution entropy as $\log \mu_u$. As the measure is multiplicative for independent systems, the evolution entropy will be additive making it an extensive property. The evolution entropy of a system at a time is defined to be the evolution entropy $\log \mu_u(x_t)$ of the snapshot at that time. Under a deterministic process, the evolution entropy can never decrease: $\log \mu_u(x_{t_0}) \leq \log \mu_u(x_{t_1})$ since $x_{t_0} \preceq x_{t_1}$ for all $t_0 \leq t_1$ and therefore $\mu_u(x_{t_0}) \leq \mu_u(x_{t_1})$. If the process is also reversible, then $\log \mu_u(x_{t_0}) = \log \mu_u(x_{t_1})$.

These definitions give a very general setting to describe a process and already find a quantity that cannot decrease during deterministic evolution.

States

Conceptually, states represent description of the system, and only of the system, regardless of time. Therefore the state space is not a set of statements, but a “template” for a set of statements that can be “instantiated” at different times.

The idea is that a state space \mathcal{S} comes equipped with a function $\iota : \mathcal{S} \times T \rightarrow \bar{\mathcal{P}}$ such that $\iota(\mathcal{S}, t) = \bar{\mathcal{D}}_t$. That is, it maps the state space and its statements to the particular time domain that represents the system at that particular time. Specifically, states of the system will be mapped to snapshots of the system.

The structure of the state space will not be, in general, isomorphic to each particular time domain. In a particular process at a particular time some states may not be accessible, so some states will be mapped to an impossibility. Or there may be correlations with other system, so the snapshot will provide more information (will be narrower) than the states themselves.

The relationships defined on the state space will be equivalent to the ones in the time domain if and only if the time domain of the system is independent from the time domain of the other systems. In other words: the state space represents the system and its properties when the system is independent. This also means that, to be able to define a system, we need to have a process that renders it independent from other systems.

When the system is independent from all others, the description is coarser than in the case of when there are correlations. Note that to a coarser description is associated a higher process entropy. Processes that render the system independent are exactly the ones that maximize the process entropy. We can associate a state entropy to each state, which is the process entropy associated to that description when the system is independent.

While it is still not clear what can be derived and what must be imposed, the overall goal is to understand what assumptions are needed to construct state spaces. One result should be that processes that isolate the system are implicitly needed, which forms the basis of requiring entropy maximization. All states are therefore equilibria of those processes (i.e. symmetries of the group of processes). Conceptually, this maps well with all branches of physics as all state spaces come equipped with some structure which, in the end, is connected to entropy/probability/measure.

2.4 Open questions and possible extensions

Here we note some thoughts and ideas about open problems and possible extensions to the general theory.

Homogeneity of an experimental domain

It may be interesting to characterize some notion of homogeneity that makes all possibilities in a domain “equally verifiable”, that no possibility is “special” compared to the others in terms of experimental verifiability. For example:

- the “extra-terrestrial life” domain is not homogeneous because one possibility can be verified while the other cannot
- the integers and reals are the only linearly ordered quantities where all contingent statements are the same experimentally: all decidable and none decidable
- phase transitions are special, as knowing whether a system is in a mixed state is decidable, so a domain with phase transitions is not homogeneous

It is not clear how this notion should be implemented and how exactly it would be useful. It may give a reason to expect a complete domain (the residual possibility is the only one that is not compatible with any contingent verifiable statement, so the domain would not be homogeneous) and also that all possibilities are approximately verifiable (if one is able to prove that, in any domain, at least one possibility is approximately verifiable).

Predictive relationships

Another way to characterize relationships between domains could be in terms of predictions, what statements of one domain can tell about the other. That is, we give a theoretical statement on one and look for the best prediction (i.e. narrowest theoretical statement broader than the original) for the other.

For example, if a domain is independent from another, any theoretical statement should predict the certainty on the other. If a domain is dependent on another, any theoretical statement should predict an equivalent statement.

A possible approach. Let \mathcal{D}_X and \mathcal{D}_Y be two experimental domains and $\bar{\mathcal{D}}_X$ and $\bar{\mathcal{D}}_Y$ their respective theoretical domains. Now we construct the function $\pi : \bar{\mathcal{D}}_X \rightarrow \bar{\mathcal{D}}_Y$ such that given $s_X \in \bar{\mathcal{D}}_X$ and $s_Y \in \bar{\mathcal{D}}_Y$ such that $s_X \preceq s_Y$ we always have $s_X \preceq \pi(s_X) \preceq s_Y$. In other words, it should map to the narrowest broader statement in $\bar{\mathcal{D}}_Y$.

In principle, we can even extend $\pi : \mathcal{S} \rightarrow \bar{\mathcal{D}}_Y$ to be defined on the whole context. In that case, π can be proven to be a projection. This map should be able to characterize the relationship between domains. For example, if $\pi(\bar{\mathcal{D}}_X) = \{\top, \perp\}$ then the domains should be independent. If $\pi(\bar{\mathcal{D}}_X) = \bar{\mathcal{D}}_Y$ the domains should be dependent.

Defining structures on experimental domains

Some mathematical structures are defined on points (i.e. vector spaces, ordering) and others on their σ -algebras. In our context, verifiable statements are the only elements that are actually physical, therefore it would be nice to always define the structures on the experimental domain (i.e. the topology) and show that it induces a unique structure on the theoretical statements and possibilities (and vice-versa).

We have already implemented this approach in a couple of areas. Theoretical domains are constructed from experimental domains (see 1.36) and so are the possibilities (see 1.47). Theorem 2.10 shows that causal relationship on the possibilities is equivalent to an inference relationship on the verifiable domain. Theorem 3.16 shows that ordering of the possibilities is equivalent to the ordering of the basis according to narrowness.

We need to understand how this can be achieved for other structures, such as measures, metrics, groups, vector spaces, inner products, ...

Space of possible combined domains

It should be possible to better characterize the space of all possible combined domains. As we show in 2.34 that the space of the possible experimental relationships is the space of topologically continuous functions, there should be an analogue for the space of all possible combined domains. For example, one should be able to show that the combined domain is an immersion within the product topology. Is that the only constraint? How can that be characterized? Can we create an experimental domain to distinguish them?

Limited precision

One area we could explore for new physical ideas is what happens if we assume that the precision cannot be arbitrarily decreased. How is it different from the continuous case? Here are some preliminary ideas.

The limited precision case cannot simply lead to a discrete topology. The standard topology of the reals is not the limit of the integer topology since it is not discrete. Most likely, the limited precision case will need to have uncountable possibilities so that the limit to arbitrary precision can work well.

The main cause of confusion is that, in the continuous case, whether the precision of two statements overlap determines whether the statements are compatible. For example, “the position is between 0 and 1 meters” and “the position is between 2 and 3 meters” are both incompatible and not overlapping. This cannot be the case for limited precision. The possibilities themselves must be incompatible with each other but some of them must overlap, or we would simply have a discrete topology. That is, suppose that 1 unit is the precision limit, the statements “the position is between 0 and 1” and “the position is between 0.5 and 1.5” are incompatible because if we verify one we cannot verify the other. If we could verify them both, we would measure at a smaller precision. So, overlapping cannot be defined in terms of incompatibility.

Whether two statements overlap cannot be determined through incompatibility but must be recovered from the precision of the disjunction. Suppose we have the following arbitrary precision statements.

$$\begin{aligned} s_1 &= \text{“the position is between 0 and 1 meter”} \\ s_2 &= \text{“the position is between 0.5 and 1.5 meter”} \\ s_3 &= \text{“the position is between 2 and 3 meter”} \end{aligned}$$

The precision associated to all statements will be one meter. The precision for $s_1 \vee s_2$ will be one meter and a half while the precision for $s_1 \vee s_3$ will be two meters. That is: the precision for non overlapping statements sums. It may even be the case that if the precision sums, statements must be incompatible but the converse is what fails.

This means that the measure we put on the possibilities cannot represent the precision anymore. That is, $d\mu \neq dx$. We can imagine a relationship like $dx^2 = d\mu^2 + 1$. This would both make the precision go to 1 when the measure goes to 0 and $dx \simeq d\mu$ for large μ .

Chapter 3

Ensemble spaces

Here we summarize the work in progress for a general theory of states and processes that is applicable to any physical system. The tentative core concept is the ensemble, therefore we want to find definitional requirements for ensembles and then find suitable assumptions to recover the different theories (e.g. classical, quantum and thermodynamics).

Conceptually, we want to argue that physical theories are, and can only be, about ensembles. At a practical level, most of the time we can only prepare and measure statistical properties as we do not have perfect control over any system. The cases where properties can be prepared with one hundred percent reliability can still be understood as ensembles of identical preparations, though one can argue for an ensemble of external conditions or internal dynamics. At a conceptual level, the goal of physics is to write laws that apply all the time: every time that one prepares a system according to a particular procedure, lets it evolve in particular conditions, he will obtain a particular result. That is, the idea of repeatability of experimental results implicitly assumes that the objects of scientific inquiry are not single instances, but the infinite collections of all reproducible instances.

3.1 Ensembles

TODO: Preamble that argues the justification with physics examples.

Axiom 3.1 (Axiom of ensembles). *The state of a system is represented by an **ensemble**, which represents all possible preparations of equivalent systems prepared according to the same procedure. The set of all possible ensembles for a particular system is an **ensemble space**. Formally, an ensemble space is a T_0 second countable topological space where each element is called an ensemble.*

Justification. In physics, ensembles are usually defined as probability distributions over states. This is conceptually problematic. Any actual physical preparation is always associated with some uncertainty. That is, the system will never be replicated exactly. This is true even in quantum mechanics: while we can always prepare a hydrogen atom in the ground state, we can never prepare an atom in the same exact position with the same exact velocity. The ability to prepare ensembles in which the position and the excitation of the atom can be assumed to be uncorrelated is what allows us to concentrate on one of the properties. Therefore, ensembles are the only thing that can actually be prepared in prac-

tice. Pure states, then, should not be taken as primitive notions but should be understood as an idealized ensemble.

From a more conceptual level, reproducibility is baked into the requirement of experimental verifiability. A physical law, then, must be understood as describing a relationship that always exist whenever the same set of circumstances is replicated. Given that we need to always be able to replicate those circumstances “one more time”, the relationship is about countably infinite similar preparations: an ensemble. Therefore, to the extent that physics is about reproducible experimental results, the basic description of a system is in terms of ensembles. This justifies the use of ensembles as the fundamental object to describe the state of a system.^a

Ensembles are experimentally defined objects, and therefore they are possibilities of an experimental domain. Therefore an ensemble space is a T_0 second countable topological space where each element is an ensemble and the topology is induced by the verifiable statements. \square

^aNote that reproducibility also already implies that all properties that characterize an ensemble must be relative to the procedure. If the properties depended, for example, on absolute space or absolute time, then different practitioners would not be able to prepare the same ensemble.

A general theory of ensembles will need to be able to recover classical and quantum spaces under appropriate physically meaningful assumptions. Here we review these cases and give the appropriate mathematical definitions, highlighting potential issues, as these serve as models and examples for a general theory.

There are two aspects that are common to all cases and will be the target of generalization: statistical mixing (captured by a convex structure) and entropy (captured by a strictly concave function). Note: all logs are assumed in base 2.

Discrete classical theory

Definition 3.2. *The ensemble space of a **discrete classical theory** is a simplex with the standard convex structure and the Shannon entropy. The simplex can have finitely or countably many vertices.*

That is, a discrete classical ensemble space is a set \mathcal{E} containing n points $\{s_i\}_{i=1}^n$ and for which every element $\mathbf{e} \in \mathcal{E}$ is characterized by a unique convex combination $\mathbf{e} = \sum_i p_i s_i$ such that $\sum_i p_i = 1$. The space is closed under convex combinations of finitely many elements. The entropy is given by $S(\mathbf{e}) = -\sum_i p_i \log p_i$.

Remark. The correct closure under convex combinations of countably many elements is not clear.

Finite case

The space of classical distributions over a discrete space corresponds to a simplex (<https://en.wikipedia.org/wiki/Simplex>). In the finite case, the pure states $\mathcal{S} = \{s_i\}_{i=1}^n \subset \mathcal{E}$ are finitely many and each ensemble $\mathbf{e} = \sum p_i s_i$ is uniquely identified by a decomposition of pure states. Effectively, each ensemble is a probability distribution over the pure states. Mathematically, each point of the space is a convex combination of the vertices. The simplex has a center point, which corresponds to the maximally mixed state, a uniform distribution over all pure states.

The entropy is given by the Shannon entropy $-\sum_i p_i \log p_i$. This means that the entropy of each pure state is zero and the entropy of the maximally mixed state is $\log n$ where n is the number of pure states. The entropy increases as we go from pure states to the maximally mixed state. The level sets (i.e. the fibers) of the entropy form a series of concentric “shells”.

The fact that all pure states have the same entropy is an additional hypothesis that cannot work in general. To see this, consider the case where the state is defined by the number of molecules for two substances. This space is the product of two independent variables n_a and n_b . If we have a uniform distribution over N_a cases of n_a and N_b cases of n_b , the total number of cases is $N_a N_b$. Therefore the entropy of the joint state is the sum of the entropy of the marginals. However, if we pair n_a with the total number of molecules $n_{(a+b)}$ we have a problem. The issue is that variable $n_{(a+b)}$ corresponds to a variable number of joint cases. Therefore the case where the ensemble space is a simplex but the entropy is not the Shannon entropy (e.g. is the Shannon entropy plus the contributions of entropy from each vertex) is a physically meaningful case that should be possible in the general theory.

Countable case

The countable case is not well-defined mathematically.

The obvious extension is to include all series $\{p_i\} \in [0, 1]$ that converge to one. That is, the space of all probability measures over a countable discrete space. Since we cannot create a uniform distribution over infinitely many cases, there is no center point. Effectively, there is a “hole” in the middle. More precisely, the space is not topologically closed in the sense it does not contain all the limit points.

[TODO] It may be useful to characterize this “hole” and the limit points. There should be at least one limit point for each series $\{p_i\}$ that converges to a finite $p < 1$. Intuitively, we can keep that part of the distribution constant while we spread the rest uniformly to all cases. Each should reach a different limit point.

However, the space of all probability measures is too large. Note that the entropy is not finite for all convergent p_i (i.e. infinite convex combinations). For details, see <https://arxiv.org/pdf/1212.5630.pdf>. Given that we want the entropy to exist and be finite for all ensembles, this generalization (like in <https://ncatlab.org/nlab/show/superconvex+space>) does not seem physically warranted.

Also note that expectation values are not guaranteed to be finite either, and requiring a particular observable to be finite further restricts the space. This restriction may be desirable for another reason: a discrete ensemble space has no notion of the ordering of the pure states. Physically, this would mean that the states with 1, 100, or 1 trillion particles are “equally distant” (i.e. all infinite permutations are allowed). Requiring the expectation of the number of particles to be finite (e.g. $\sum_i N(i)p_i < \infty$) should effectively encode the infinite ordering in the rate of convergence of the probability distributions (i.e. not all infinite permutations would be allowed).

No uncountable case

The uncountably infinite case is not physically relevant (i.e. no second countable discrete topology, cases are not experimentally decidable). Also note that any set of real numbers whose sum is finite can have only countably many non-zero elements. To understand why, note that there can only be finitely many terms above any particular positive value if their sum is to remain finite. Effectively, the uncountable case would be stitching together infinitely

many countable cases.

Continuous classical theory

Definition 3.3. *The ensemble space of a **continuous classical theory** is the convex space given by the probability distributions (i.e. probability measures) over phase space (i.e. a symplectic manifold) that allow a continuous probability density (i.e. the Radon-Nikodym derivative exists - the probability measures are absolutely continuous with respect to the Liouville measure - and is continuous). Formally, let (X, ω) be a symplectic manifold. The ensemble space is given by $\mathcal{E} = \{\rho \in C(X) \mid \int_X \rho d\mu = 1\}$. The entropy is given by the Shannon/Gibbs entropy calculated using the probability density (i.e. the Radon-Nikodym derivative between the probability measure and the Liouville measure). That is, $S(\rho) = - \int_X \rho \log \rho d\mu$.*

In the continuous case, the space of ensembles is the space of continuous integrable functions over a symplectic manifold (e.g. over phase space) that integrate to one. That is, if X is a symplectic manifold, then $\mathcal{E} = \{\rho \in C(X) \mid \int_X \rho(x) d\mu = 1\}$ where $\mu = \int \omega^n$ is the Liouville measure. The space does not include the pure states, as they correspond to delta distributions $\mathcal{S} = \{\delta_x\}_{x \in X}$: they are limit points. In some sense, a distribution can be understood as a convex integral of pure states: $\rho(x) = \int_X \rho(y) \delta_x(y) d\mu$. Even in this case, each ensemble can be understood as a unique probability distribution (density in this case) over all pure states. The symplectic nature of the manifold is required to assign a frame invariant density to states and a frame invariant notion of independence between DOFs, as we saw in the classical mechanics section of reverse physics.

The fact that pure states are not part of the space, but limits, means we need a more sophisticated definition of pure states than simply the extreme points of the convex space of ensembles.

The entropy is given by $-\int \rho \log \rho d\mu$ where μ is the Liouville measure and ρ is the probability density over canonical coordinates. If a different measure is used, or if the coordinates are not canonical, the formula gives the wrong result.

Similarly to the infinite discrete case, the entropy can be infinite and expectation values can be infinite. The added complication is the frame invariance: it would not make sense to have the expectation for position in one frame to be finite while infinite in another. Requiring all functions of position/momentum to have finite expectation restricts the distributions to those with finite support. Requiring all polynomial functions of position/momentum to have finite expectation restricts the distributions to those that decay faster than any polynomial.

Unlike the discrete classical case, subspaces and dimensionality of subspaces cannot be defined without the entropy. The issue is that we need a measure on the set of pure states, and the convex structure cannot provide it. The entropy, however, does as the supremum of the entropy for all distributions with support U is $\log \mu(U)$. As we will see, the entropy can be used to both identify subspaces and recover the Liouville measure.

Note that we exclude discontinuous distributions as starting from the space of discontinuous distributions does not allow one to recover the topology of the base space. Suppose, in fact, to have the space of all distributions over the real line that integrate to one. Now imagine switching the interval $[0, 1)$ with the interval $[1, 2)$. This is an isomorphism on the space of discontinuous distributions. Therefore given the space of distributions, we cannot recover the topology of the real line. Now take the space of continuous functions that integrate to one. Take those with support $[0, 1) \cup [2, 3)$. This is the set of convex combinations of the distri-

butions with support $[0, 1)$ and $[2, 3)$ precisely because the sets are not contiguous and the distributions will have to go to zero at the boundary. However, the distributions with support $[0, 2)$ are not just the convex combinations of those with support $[0, 1)$ and $[1, 2)$ because this will include the continuous functions that have a non-zero value at 1.

[TODO] It may be interesting to study the shell of zero entropy states. For example, it should not be path connected: all uniform distributions with support of the same finite size (in terms of the Liouville measure) will have the same entropy. The region, however, need not be contiguous. Since we cannot continuously transform a single region into two disjoint regions, there will be different distributions at zero entropy that cannot be transformed continuously.

Quantum theory

Definition 3.4. *The ensemble space of a **quantum theory** is modeled by the convex space given by the density operators (i.e. positive semi-definite self-adjoint operators with trace one) of a Hilbert space equipped with the von Neumann entropy.*

That is, given a Hilbert space \mathcal{H} , the ensemble space is the space of positive semi-definite self-adjoint operators with trace one $M(\mathcal{H})$. The space of pure states is given by the projective space $P(\mathcal{H})$. The entropy of an ensemble $\rho \in M(\mathcal{H})$ is given by the von Neumann entropy $S(\rho) = -\text{tr}(\rho \log \rho)$.

Finite dimensional case

The simplest non-trivial case is the qubit, for which the Bloch ball is the space of ensembles $M(\mathcal{H})$. The interior of the Bloch ball corresponds to mixtures while the surface corresponds to the pure states $\mathcal{S} = P(\mathcal{H}) = \{|\psi\rangle\langle\psi|\}_{\psi \in \mathcal{H}}$. In quantum ensemble spaces there is no unique decomposition. Note that the space is exactly characterized by knowing which different mixtures provide the same ensemble.

The multiple decompositions make the ensemble space behave in a way that is a hybrid between the classical discrete and continuous. Pure states are properly a part of the ensemble space, as in the discrete case, and we can describe each mixture in terms of finitely many pure states. However, the pure states are a continuum, therefore we can also define probability densities over the space, convex integrals. For example, for a single qubit, the maximally mixed state (the center of the ball) can be equally described as the equal mixture of two opposite states (e.g. spin up and spin down, or spin left and spin right). However, it can also be described as the equal mixture of the whole sphere.

Note that complex projective spaces are symplectic, which is what allows one to define frame invariant densities. The goal is to have one argument applied to the generic definition as to why the space of pure states must be symplectic. Also note that the two dimensional sphere is the only symplectic sphere. By homogeneity, we should be able to argue that the space is symmetric around the maximally mixed state, and is therefore a sphere. The symplectic requirement would select dimension two. Note that real and quaternionic spaces would be excluded by this argument.

The von Neumann entropy for the maximally mixed state is $\log n$ where n is the dimensionality of the Hilbert space. Again we see that the maximum entropy gives us a measure of the size of the space. Note that, to calculate the von Neumann entropy, we are diagonalizing the density matrix ρ . This means finding a set of orthogonal pure states s_i such that $\rho = \sum p_i s_i$ is a convex combination. Note that the convex hull of a set of n orthogonal pure states is an n -dimensional simplex whose center is the maximally mixed state. Therefore, we are looking

for a simplex that contains ρ and the maximally mixed state. In the two dimensional case, ρ is an interior point of the Bloch ball. Take the line that connects ρ to the center. The two points of the sphere are the extreme points for the decomposition. The distance from the points will be proportional to the probability. Because of this property, the von Neumann entropy is the smallest Shannon entropy between all possible decompositions.

Countably infinite dimensional case

The countable infinite dimensional case presents similar problems as the classical case, and adds others. As in the classical infinite cases, the maximally mixed state (i.e. uniform distribution) is not in the convex space and the entropy is not finite for all infinite convex combinations. As in the classical continuous case, there is the issue of finite expectation of position/momentum in all frames. The problem is compounded by the fact that one cannot require finite expectation for all functions of position and momentum: finite support in position automatically implies infinite support on momentum, since the distribution in momentum is the Fourier transform of that in position.

The Hilbert space for a discrete variable with infinite range (e.g. number of particles) and a continuous variable (e.g. position/momentum) is the same. The first is defined as the space of square convergent complex sequences l^2 while the second is the space of square integrable complex functions L^2 . Given that L^2 allows a countable basis, the two are isomorphic. This also means that all spaces with finitely many degrees of freedom are also isomorphic. This makes the problem of infinite expectations even more problematic.

Note that Schwartz spaces have finite expectation for all polynomial functions of position and momentum. Given that infinite permutations can change the rate of convergence, the Schwartz space has an idea of what is further away from the origin, unlike Hilbert spaces. [TODO] It is not yet clear to us whether the Schwartz space for one DOF is the same as the one for two DOFs.

3.2 Mixtures and convex spaces

Definition 3.5. Given a real number $p \in [0, 1]$, its complement is defined as $\bar{p} = 1 - p$.

Axiom 3.6 (Axiom of mixture). 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 $pe_1 + \bar{p}e_2$, with the following properties:

- **Continuity:** the map $(p, e_1, e_2) \rightarrow pe_1 + \bar{p}e_2$ is continuous (the products $[0, 1] \times \mathcal{E} \times \mathcal{E} \rightarrow \mathcal{E}$ have the product topology)
- **Identity:** $1e_1 + 0e_2 = e_1$
- **Idempotence:** $pe_1 + \bar{p}e_1 = e_1$ for all $p \in [0, 1]$
- **Commutativity:** $pe_1 + \bar{p}e_2 = \bar{p}e_2 + pe_1$ for all $p \in [0, 1]$
- **Associativity:** $p_1e_1 + \bar{p}_1 \left(\left(\frac{p_3}{p_1} \right) e_2 + \frac{p_3}{p_1} e_3 \right) = \bar{p}_3 \left(\frac{p_1}{p_3} e_1 + \left(\frac{p_1}{p_3} \right) e_2 \right) + p_3 e_3$ for all $p_1, p_3 \in [0, 1]$

Justification. This axiom captures the ability to create a mixture merely by selecting between the output of different processes.

Given that mixing represents an experimental relationship, and all experimental relationships must be continuous in the natural topology, mixing must be a continuous function. Note that p is a continuous quantity, and therefore the natural topology is the one of the reals.

Let \mathbf{e}_1 and \mathbf{e}_2 be the ensembles that represent the output of two different processes P_1 and P_2 . Let a selector S_p be a process that outputs two symbols, the first with probability p and the second with probability \bar{p} . Then we can create another process P that, depending on the selector, outputs either the output of P_1 or P_2 . Therefore we are justified in assuming the mixing operation among ensembles.

If $p = 1$, the output of P will always be the output of P_1 . This justifies the identity. If P_1 and P_2 are the same process, then, again, the output of P will always be the output of P_1 . This justifies the idempotence. As long as the same probability is matched to the same output, the process P is identical. This justifies commutativity. If we are mixing three processes P_1 , P_2 and P_3 , as long as the final probabilities are the same, it does not matter if we mix P_1 and P_2 first or P_2 and P_3 . This justifies associativity. \square

Remark. Given symmetry and associativity, we can write $p_1\mathbf{e}_1 + p_2\mathbf{e}_2 + p_3\mathbf{e}_3$ where $p_2 = \bar{p}_1\left(\frac{p_3}{p_1}\right) = \bar{p}_3\left(\frac{p_1}{p_3}\right) = 1 - p_1 - p_3 = \overline{(p_1 + p_3)}$. We can also extend to mixtures of finitely many elements $\sum p_i\mathbf{e}_i$ where $\sum_i p_i = 1$.

Corollary 3.7. *An ensemble space is a convex space.*

Remark. Basic definitions for convex spaces are taken from <https://ncatlab.org/nlab/show/convex+space> and <https://arxiv.org/abs/0903.5522>. The notation and terminology will be slightly different to better map to physics ideas.

Definition 3.8. *Let \mathcal{E} be an ensemble space. Let $\rho = \sum_i p_i\mathbf{e}_i$ where $\rho, \{\mathbf{e}_i\} \in \mathcal{E}$ and $p_i \in (0, 1]$ such that $\sum p_i = 1$. We say that ρ is a **mixture** of $\{\mathbf{e}_i\}$ and each \mathbf{e}_i is a **component** of ρ .*

Remark. In terms of the convex space, all the mixtures between two elements corresponds to the segment between them; all the mixtures between three elements correspond to the triangle formed by the three elements and so on. An element \mathbf{e}_1 is a component of a different element \mathbf{e}_2 if the segment can be continued on the side of \mathbf{e}_2 . Note that two elements can be a component of each other. If two elements are not a component of each other, then they are the extreme points of the line that connect the two.

Definition 3.9. *Let \mathcal{E} be an ensemble space and $\mathbf{e}_1, \mathbf{e}_2 \in \mathcal{E}$. We say that they **have a common component** if we can find $\mathbf{e}_3 \in \mathcal{E}$ such that $\mathbf{e}_1 = p_1\mathbf{e}_3 + \bar{p}_1\mathbf{e}_4$ and $\mathbf{e}_2 = p_2\mathbf{e}_3 + \bar{p}_2\mathbf{e}_5$ for some $\mathbf{e}_4, \mathbf{e}_5 \in \mathcal{E}$ and $p_1, p_2 \in (0, 1)$. They are **distinct**, noted $\mathbf{e}_1 \perp \mathbf{e}_2$, otherwise.*

Remark. In terms of convex spaces, two distinct ensembles are always the extreme points of the line that connects them. Additionally, for any other ensemble \mathbf{e} , one of the two ensembles must be the extreme point of the line that connects it to \mathbf{e} .

Corollary 3.10. *The previous definitions obey the following:*

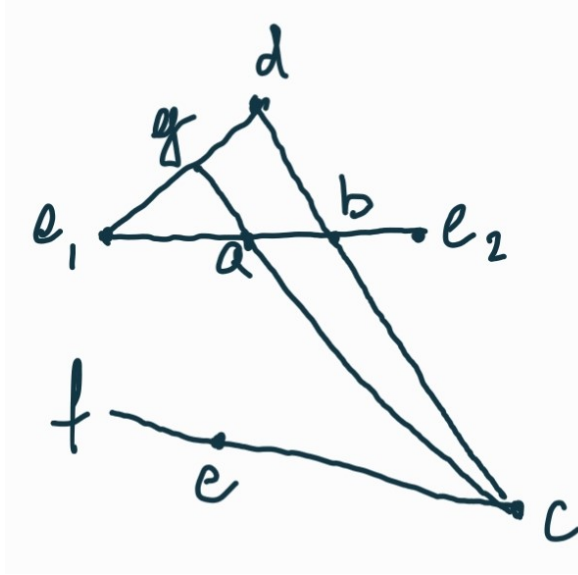
1. every ensemble has a common component with itself, therefore every ensemble is not distinct from itself
2. distinctness is a irreflexive symmetric relation
3. if \mathbf{e}_1 is a component of \mathbf{e}_2 , then \mathbf{e}_1 and \mathbf{e}_2 have a common component

Proof. 1. Since by idempotence $\mathbf{e} = p\mathbf{e} + \bar{p}\mathbf{e}$ for any p , we can satisfy the definition by setting $\mathbf{e}_1 = \mathbf{e}_2 = \mathbf{e}_3 = \mathbf{e}_4 = \mathbf{e}_5 = \mathbf{e}$. Therefore every ensemble has a common component with itself and every ensemble is not distinct from itself.

2. The previous property shows that distinctness is irreflexive. The definition of common component is symmetric and therefore so is distinctness.

3. By idempotence, we can write $e_1 = p_1 e_1 + \bar{p}_1 e_1$ for some $p_1 \in (0, 1)$. Since e_1 is a component of e_2 , we can write $e_2 = p_2 e_1 + \bar{p}_2 e_3$ for some $p_2 \in (0, 1)$ and $e_3 \in \mathcal{E}$. Therefore e_1 and e_2 have a common component. \square

Proposition 3.11. *Let $e, e_1, e_2 \in \mathcal{E}$. If e is distinct from a mixture of e_1 and e_2 then it is distinct from all mixtures of e_1 and e_2 and from both e_1 and e_2 . That is, if $e \perp p e_1 + \bar{p} e_2$ for some $p \in (0, 1)$ then $e \perp p e_1 + \bar{p} e_2$ for all $p \in [0, 1]$. However, if e is distinct from e_1 and e_2 , it is not necessarily true that e is distinct from a mixture of e_1 and e_2 .*



Proof. Let $e \perp a = p e_1 + \bar{p} e_2$ for some $p \in (0, 1)$. Let $b = \alpha e_1 + \bar{\alpha} e_2$ with $0 \leq \alpha < p$. We can do so without loss of generality since switching e_1 with e_2 would lead to the case where $1 \geq \alpha > p$. Suppose b is not distinct from e . Then we can find $c \in \mathcal{E}$ such that $b = \beta c + \bar{\beta} d$ and $e = \gamma c + \bar{\gamma} f$ for some $d, f \in \mathcal{E}$ and $\beta, \gamma \in (0, 1)$.

Setting $\epsilon = \frac{p-\alpha}{\bar{\alpha}}$ and $\lambda = \bar{\epsilon}\beta$ we have:

$$\begin{aligned} a &= p e_1 + \bar{p} e_2 = \left(p - \frac{\bar{p}}{\bar{\alpha}} \alpha\right) e_1 + \frac{\bar{p}}{\bar{\alpha}} \alpha e_1 + \frac{\bar{p}}{\bar{\alpha}} \bar{\alpha} e_2 \\ &= \left(\frac{p\bar{\alpha} - \bar{p}\alpha}{\bar{\alpha}}\right) e_1 + \frac{\bar{p}}{\bar{\alpha}} (\alpha e_1 + \bar{\alpha} e_2) = \left(\frac{p - p\alpha - \alpha + p\alpha}{\bar{\alpha}}\right) e_1 + \frac{1 - p + \alpha - \alpha}{\bar{\alpha}} (\alpha e_1 + \bar{\alpha} e_2) \\ &= \frac{p - \alpha}{\bar{\alpha}} e_1 + \left(1 - \frac{p - \alpha}{\bar{\alpha}}\right) (\alpha e_1 + \bar{\alpha} e_2) = \epsilon e_1 + \bar{\epsilon} (\alpha e_1 + \bar{\alpha} e_2) = \epsilon e_1 + \bar{\epsilon} b \\ &= \epsilon e_1 + \bar{\epsilon} (\beta c + \bar{\beta} d) = \bar{\epsilon} \beta c + \epsilon e_1 + \bar{\epsilon} \bar{\beta} d = \lambda c + \bar{\lambda} g \end{aligned}$$

where $g = \frac{1}{\bar{\lambda}} (\epsilon e_1 + \bar{\epsilon} \bar{\beta} d)$. This means a and e have a common component, which is a contradiction. Therefore $b \perp e$ and $e \perp p e_1 + \bar{p} e_2$ for all $p \in [0, 1]$

As a counter-example for the converse, consider a circle and three points on the circumference. These are all distinct from each other. Now take two points. A mixture will be on

the chord that connects the two. Connecting any mixture with the third point will give a line that will intersect the circumference at a fourth point. This means that the given mixture is has a common component with the third point, even though the third point is distinct from the first two. \square

Remark. Note that, in classical mechanics, $\mathbf{e} \perp p\mathbf{e}_1 + \bar{p}\mathbf{e}_2$ for some p does imply that $\mathbf{e} \perp \mathbf{e}_1$ and $\mathbf{e} \perp \mathbf{e}_2$. However, this is not the case in quantum mechanics. For example, consider the Bloch sphere and the states x^+ , x^- , z^+ , and z^- . These are all distinct ensembles. We have $\frac{1}{2}x^+ + \frac{1}{2}x^- = \frac{1}{2}z^+ + \frac{1}{2}z^-$. Therefore a mixture of x^+ and x^- has a common component with z^+ .

Vector space embedding

Definition 3.12. Let $\rho = p\mathbf{e}_1 + \bar{p}\mathbf{e}_2$ with $p \in (0, 1)$. Then we say that \mathbf{e}_2 is a p -**complement** of \mathbf{e}_1 towards ρ . An ensemble space is **complemented** if all p -complements are unique for all $p \in (0, 1)$.

Proposition 3.13. An ensemble space is complemented if and only if it is a convex subset of a real vector space for which mixtures are linear combination.

Proof. Theorem 4 in <https://arxiv.org/abs/1105.1270> states that a convex space embeds into a real vector space with $c_\lambda(x, y) = \lambda x + \bar{\lambda}y$ if and only if

$$c_\lambda(x, y) = c_\lambda(x, z) \quad \forall \lambda \in (0, 1) \implies y = z.$$

This property, called cancellative property, is exactly the condition that the λ -complement is unique. \square

Proposition 3.14. Discrete classical ensemble spaces, continuous classical ensemble spaces and quantum ensemble spaces satisfy the axiom of mixture and are complemented ensemble spaces.

Proof. The space of probability measures, discrete or continuous, \mathcal{E} is a convex subset of the vector space of finite signed measures. This means that it is closed under convex combinations: $\sum p_i \mathbf{e}_i$ for all $\{\mathbf{e}_i\} \subseteq \mathcal{E}$ and $\sum p_i = 1$. The properties of mixing are inherited from the properties of linear combinations. Therefore the discrete and continuous classical ensemble spaces satisfy the axiom of mixture. Moreover, \mathcal{E} is a convex subset of a vector space, it is complemented by proposition 3.13.

Similarly, the space of positive semi-definite self-adjoint operators with trace one is a convex subset of the vector space of self-adjoint operators. Therefore it is closed under convex combinations and it will satisfy the axiom of mixture. Moreover, it is complemented convex by proposition 3.13. \square

Remark. Since all current physical ensemble spaces are complemented, it is not yet clear whether an ensemble space must be complemented or not. It is difficult to justify as it is a claim on existence of more refined ensembles. It may very well be that this is just an assumption on the decomposition, which may fail in new physical conditions (e.g. at Planck scale). It may also be something that the strict concavity of the entropy requires.

Remark. Assuming that the mixing operation is continuous does not guarantee that we can vector space that embeds the convex space is a topological vector space. See <https://math.stackexchange.com/questions/4921905/cancellative-convex-spaces-and-topological-vector-spaces>. The issue is we have no guarantee of continuity on the inverse (i.e. subtraction, multiplication for a negative number).

3.3 Quantities

In the same way that we define ensembles first, we define quantities through their expectation over ensembles. The possible value for pure states will need to be recovered with constructions that are connected to spectral theory.

Definition 3.15. A *statistical variable*, or simply *variable*, is a numerical quantity that is well-defined over all ensembles. Formally, it is a continuous linear real valued operator on \mathcal{E} . That is, $F : \mathcal{E} \rightarrow \mathbb{R}$ such that $F(p\mathbf{e}_1 + \bar{p}\mathbf{e}_2) = pF(\mathbf{e}_1) + \bar{p}F(\mathbf{e}_2)$.

Justification. Continuity is required since the value of a variables is experimentally verifiable. The linearity of the operator associated to the quantity stems for the expectation. When mixing two different processes, the measured value will only depend on the selection at that particular time, and therefore each process will contribute to the final value based on its fraction.

Note that the value cannot be infinite. The only way to measure infinity is if the measured average keeps increasing in time. This would be a contradiction on the assumption that the ensemble is at equilibrium. \square

Remark. Note that variables will have a contiguous range on the ensembles because, given two ensembles with different values, we can mix them to obtain any intermediate value. This does not mean that the variable can take all possible values on pure states. For example, the number of particles in a pure state will necessarily be a non-negative integer, but we can mix those to create ensembles that have non-integer average number of particles.

Proposition 3.16. Let \mathcal{E} be a complemented ensemble space. Then a statistical variable F induces a semi-norm on the vector space that embeds \mathcal{E} .

Proof. Let F be a variable on \mathcal{E} and let V be the vector space that embeds \mathcal{E} . Extend F on V such that $F(a\mathbf{e}) = |a|F(\mathbf{e})$. Then, according to the definition given after 1.1 in <https://personal.math.ubc.ca/~cass/research/pdf/TVS.pdf> F is a semi-norm. \square

Definition 3.17. A *quantifiable ensemble space* is an ensemble space where each ensemble can be identified by a set of statistical variables. That is, there is family of statistical variables $F_i : \mathcal{E} \rightarrow \mathbb{R}$ such that, given $\mathbf{e}_1, \mathbf{e}_2 \in \mathcal{E}$, $F_i(\mathbf{e}_1) = F_i(\mathbf{e}_2)$ for all i if and only if $\mathbf{e}_1 = \mathbf{e}_2$. Moreover, the topology is generated by those variables.

Proposition 3.18. A quantifiable complemented ensemble space is embedded into a Hausdorff locally compact topological vector space.

Proof. The ensemble is complemented, therefore it is embedded into a vector space. The topology of the ensemble space is generated by the statistical variables, which induce semi-norms on the vector space. This means that the topology of the vector space is generated by a countable set of semi-norms, and is therefore a locally convex topological vector space. See Prop 2.2 in <https://personal.math.ubc.ca/~cass/research/pdf/TVS.pdf>.

The statistical variables fully identify each ensemble, which means that, if two elements are different, one semi-norm will give a non-zero value for the distance function associated with it. This means that the topology is Hausdorff. See Prop 2.6 in <https://personal.math.ubc.ca/~cass/research/pdf/TVS.pdf>. \square

Corollary 3.19. *A quantifiable ensemble spaces is fully determined by countably many statistical variables.*

Proof. The topology of an ensemble space is second countable. A Hausdorff second countable locally compact topological vector space can always be generated by countably many semi-norms. \square

Remark. Note that we are missing completeness in terms of the semi-norms to obtain a Fréchet space. It is not clear whether this is required since, for example, $L^1(\mathbb{R}^{2n}) \cap C(\mathbb{R}^{2n})$ is not Fréchet.

Conjecture 3.20. *Discrete/continuous classical ensemble spaces and quantum ensemble spaces are quantifiable.*

Proof. For discrete classical spaces, the expectation of the indicator of each extreme point defines a countable set of quantities that fully identifies the distribution.

For continuous classical spaces, $L^1(\mathbb{R}^{2n}) \cap C(\mathbb{R}^{2n})$ is a Hausdorff locally convex topological spaces.

TODO Quantum (note that we are looking at the space of density operators with finite expectation value for position and momentum). \square

Conjecture 3.21. *Quantifiable spaces are complemented.*

Remark. Not clear whether this is true. The idea is that one can create open balls with quantities, which are, in a sense, symmetric neighborhoods. These may define unique inverses.

Conjecture 3.22. *If all ensembles can be connected by linear transformations parameterized by real quantities, then the ensemble is quantifiable.*

Remark. Note clear whether this is true. The idea is that the time interval can be used to define variables. A more sophisticated conjecture may look at the space of generators of Lie groups. This would link the topology of time to the topology of the ensemble space. If the topology of time is not that of the real numbers, then, the ensemble space must change significantly.

Proof. Check that the all maximal sequences give the same value. \square

3.4 Biggest component and probability

Convex hull

Definition 3.23. *Let $U \subseteq \mathcal{E}$ be the subset of an ensemble space. The **convex hull** of U , noted $\text{hull}(U)$ is the set of all possible mixtures that can be constructed with elements contained in U .*

Corollary 3.24. *The convex hull has the following properties*

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

and is therefore a closure operation

Proof. 1. Every element of U is trivially a mixture of elements of U . Therefore $U \subseteq \text{hull}(U)$.

2. Let $e \in \text{hull}(U)$. Then it is a mixture of some elements of U . Since $U \subseteq V$, then e is also the mixture of some elements of V and therefore $e \in \text{hull}(V)$.

3. Since mixing is associative and commutative, a mixture of mixtures of U can be reexpressed as a mixture of U . Therefore for all $e \in \text{hull}(\text{hull}(U))$ we have $e \in \text{hull}(U)$. \square

Maximal component sequence

Definition 3.25. Let $e \in \mathcal{E}$ be an ensemble. A sequence of ensembles $\{a_i\} \subseteq \mathcal{E}$ is an **increasing component sequence of e** if there exists a sequence of ensembles $\{b_i\} \subseteq \mathcal{E}$ and an increasing sequence of $\{p_i\} \subseteq (0, 1]$ such that $e = p_i a_i + \bar{p}_i b_i$. The limit $p = \lim_{i \rightarrow \infty} p_i$ is called the **size of the component sequence**.

Remark. Since the sequence of p_i is increasing and is bounded, it must converge.

Definition 3.26. Let $e \in \mathcal{E}$ be an ensemble and $A \subseteq \mathcal{E}$ a set of ensembles. Then the **A -components of e** are the components of e that are mixtures of A . That is, $A_e = \{e_1 \in \text{hull}(U) \mid \exists p \in (0, 1], e_2 \in \mathcal{E} \text{ s.t. } e = p e_1 + \bar{p} e_2\}$. An **increasing A -component sequence of e** is an increasing component sequence of e such that $\{a_i\} \subseteq \text{hull}(A)$. The sequence $\{a_i\} \subseteq \text{hull}(A)$ is **maximal** if the limit $p_i \rightarrow p$ is the supremum of all such sequences.

Proposition 3.27. Let $e \in \mathcal{E}$ be an ensemble and $A \subseteq \mathcal{E}$ a set of ensembles, then there exists a maximal A -component sequence of e .

Proof. Consider the set of all sizes of all increasing A -component sequence of e . This is a subset of $(0, 1]$ that contains all the limits of all increasing sequences. Therefore, it must have an upper bound. \square

Mixing probability

Definition 3.28. Let $e \in \mathcal{E}$. The **mixing probability measure** for e is defined as $p_e(A) = \sup(\{p \in [0, 1] \mid \exists e_1 \in \text{hull}(A), e_2 \in \mathcal{E} \text{ s.t. } e = p e_1 + \bar{p} e_2\})$ if $A \neq \emptyset$ and $p_e(A) = 0$ otherwise.

Corollary 3.29. The mixing probability is the size of a maximal A -component sequence of e . That is, let $\{a_i\} \subseteq \text{hull}(A)$ be a maximal sequence and $\{p_i\} \subseteq (0, 1]$ its associated coefficients. Then $p_i \rightarrow p_e(A)$.

Proof. Suppose that $p_i \not\rightarrow p_e(A)$. Then there exists an $a \in \text{hull}(A)$ that is a bigger component than all the elements of the sequence. This means that the sequence is not maximal. Therefore $p_i \rightarrow p_e(A)$. \square

Corollary 3.30. The mixing probability measure uniquely identify an ensemble. That is, let $e_1, e_2 \in \mathcal{E}$ such that $e_1 \neq e_2$. Then $p_{e_1} \neq p_{e_2}$.

Proof. Note that $e = 1e_1 + 0e_2$ if and only if $e = e_1$. Therefore, let $e_1, e_2 \in \mathcal{E}$ such that $e_1 \neq e_2$. We have $p_{e_1}(\{e_1\}) = 1$ and $p_{e_2}(\{e_1\}) \neq 1$. Which means $p_{e_1} \neq p_{e_2}$. \square

Proposition 3.31. The mixing probability measure for an ensemble is

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

3. *subadditive*: $p_e(U \cup V) \leq p_e(U) + p_e(V)$

Proof. 1. The probability measure takes a subset of \mathcal{E} and returns a real value and is therefore a set function. The coefficients used in mixtures are real values between zero and one. The supremum of a set of numbers between zero and one is between zero and one, and therefore the probability measure for an ensembles is non negative and unit bounded.

2. Let $U, V \subseteq \mathcal{E}$ such that $U \subseteq V$. If $U = \emptyset$, we have $p_e(U) = 0$. Since $p_e(V)$ is non-negative, $p_e(U) \leq p_e(V)$. If $U \neq \emptyset$, $\text{hull}(U) \subseteq \text{hull}(V)$ and therefore $\{p \in [0, 1] \mid \exists \mathbf{e}_1 \in \text{hull}(U), \mathbf{e}_2 \in \mathcal{E} \text{ s.t. } \mathbf{e} = p\mathbf{e}_1 + \bar{p}\mathbf{e}_2\} \subseteq \{p \in [0, 1] \mid \exists \mathbf{e}_1 \in \text{hull}(V), \mathbf{e}_2 \in \mathcal{E} \text{ s.t. } \mathbf{e} = p\mathbf{e}_1 + \bar{p}\mathbf{e}_2\}$. This means that the supremum of the first set cannot be greater than the supremum of the second set, and therefore $p_e(U) \leq p_e(V)$. The probability measure is a monotone set function.

3. Let $U, V \subseteq \mathcal{E}$ and let $p \in [0, 1]$ such that $\mathbf{e} = p\mathbf{e}_1 + \bar{p}\mathbf{e}_2$ for some $\mathbf{e}_1 \in \text{hull}(U \cup V)$ and $\mathbf{e}_2 \in \mathcal{E}$. Since $\mathbf{e}_1 \in \text{hull}(U \cup V)$, we can write $\mathbf{e}_1 = \lambda\mathbf{u} + \bar{\lambda}\mathbf{v}$ for some $\lambda \in [0, 1]$, $\mathbf{u} \in U$ and $\mathbf{v} \in V$. Therefore we have $\mathbf{e} = p\lambda\mathbf{u} + p\bar{\lambda}\mathbf{v} + \bar{p}\mathbf{e}_2$. By the definition of probability measure, we must have $p\lambda \leq p_e(U)$ and $p\bar{\lambda} \leq p_e(V)$, therefore $p = p\lambda + p\bar{\lambda} \leq p_e(U) + p_e(V)$. Since $p_e(U \cup V)$ is the supremum for a set of p s for which the expression always holds, we have $p_e(U \cup V) \leq p_e(U) + p_e(V)$. The probability measure is subadditive. \square

Conjecture 3.32. *Let $\{U^i\}_{i=1}^n \subseteq \mathcal{E}$ be a family of sets of ensembles such that $\mathbf{e} \in \text{hull}(\cup U^i)$. Then $p_e(\cup U^i) = \sum p_e(U^i)$ if and only if we can find a maximal component sequence within each set such that its elements are distinct from those of any other sequence.*

Proof. TODO \square

3.5 Entropy

Definition 3.33. *Given the coefficients $\{p_i\} \in [0, 1]$ such that $\sum p_i = 1$, the **entropy of the coefficients** (also known as Shannon entropy) is defined as $I(\{p_i\}) = -\sum p_i \log p_i$.*

Axiom 3.34 (Axiom of entropy). *An ensemble space \mathcal{E} is equipped a function $S : \mathcal{E} \rightarrow \mathbb{R}$ called **entropy** with the following properties*

- **Continuity**
- **Strict concavity**: $S(p_1\mathbf{e}_1 + p_2\mathbf{e}_2) \geq p_1S(\mathbf{e}_1) + p_2S(\mathbf{e}_2)$ with the equality holding if and only if $\mathbf{e}_1 = \mathbf{e}_2$
- **Upper variability bound**: $S(p_1\mathbf{e}_1 + p_2\mathbf{e}_2) \leq I(p_1, p_2) + p_1S(\mathbf{e}_1) + p_2S(\mathbf{e}_2)$; if the equality hold then, \mathbf{e}_1 and \mathbf{e}_2 are **disjunct**, noted $\mathbf{e}_1 \perp \mathbf{e}_2$

Justification. The entropy quantifies the variability of the elements within the ensemble. This justifies the existence of the entropy.

Small changes in the ensemble should produce small changes in the variability, which justifies continuity. Mixing different ensembles will increase the variability, which justifies concavity. Mixing the same ensemble, however, should not increase the variability, which justifies strict concavity.

The variability increase is maximal when the two ensembles are disjunct, meaning they have no elements in common. In this case, the increase in variability is given only by the coefficients and must follow the Shannon entropy as it is the only continuous indicator of variability that is linear in probability. This justifies the upper bound. \square

Remark. In <https://arxiv.org/abs/1912.02012> we showed how the Shannon entropy quantifies the variability of elements within a distribution. Is it possible to rederive I from requirements on convex spaces instead of distributions in particular? This would make the axiom less ad-hoc.

The difference is that in that proof we are automatically assuming that things can be decomposed into disjunct elements, and that mixture of disjunct elements remain disjunct (i.e. the axiom of entropic disjunctness). This should imply that mixture of many disjunct element saturate the upper bound. Once that is assumed, then one can find the entropy difference for a generic uniform distribution, then for a rational distribution and then, by continuity, for all reals.

Conjecture 3.35. *An ensemble space embeds in a topological vector space if and only if it is complemented (i.e. cancellative).*

Proof. The idea is to use the entropy to show that open balls are part of the topology, and then use this to show that multiplication for a scalar is continuous. The continuity of addition should come from the continuity of the mixing operation. \square

Axiom 3.36 (Axiom of entropic disjunctness). *The entropy function of an ensemble space obeys the following properties*

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

Justification. If two ensembles are disjunct, they have no elements in common. Therefore there will be no common component with the two. We are justified to assume that disjunctness implies distinctness.

Suppose an ensemble e_1 is disjunct from the mixture of two ensembles e_2 and e_3 . That it has no elements in common with the mixture, which means it has no elements in common with either of them. Therefore e_1 is disjunct from both e_2 and e_3 . Now suppose e_1 is disjunct from both e_2 and e_3 . That it means it has no elements in common with either of them, and therefore it does not have any elements in common with their mixture. Therefore e_1 is disjunct from the mixture of two ensembles e_2 and e_3 . We are justified to assume that mixtures preserve disjunctness. \square

Remark. It is unclear whether or not these properties can be proven by the axiom of mixture. It is possible to prove that two disjunct ensembles are not one the component of the other. The entropy of the coefficients (i.e. the Shannon entropy) has a vertical asymptote at each end. If two elements are disjunct, the entropy along the segment that connects the two follows that function plus a linear function, which does not change the vertical asymptotes. Therefore the segment cannot be extended while preserving the convexity of the entropy, and therefore the two distinct elements must be the extremes of the line that connects the two. This is necessary for distinctness, but not sufficient. Maybe, extending the argument, one can at least prove that disjunctness implies distinctness.

Proposition 3.37. *The discrete classical ensemble spaces, continuous classical ensemble spaces and quantum ensemble spaces satisfies the axiom of mixture and the axiom of entropic disjunctness.*

Proof. In both classical cases, the entropy is a function of only the probabilities of the decomposition in terms of pure states. The function is continuous in terms of those variables as $x \log x$ is a continuous function. [TODO find short proof of convexity and maximum bound]

In the quantum case, the entropy is a continuous function of the density operator as the trace, multiplication and logarithm of an operator are all continuous functions. [TODO find short proof of convexity and maximum bound] \square

Remark. There should be an interplay between convexity and entropy which limits either the space or the possible entropy. For example, consider a line. It is a convex space with no extreme points. We can parameterize the space with a variable x such that the mixture of two ensembles is identified by the average x . Entropy can be written as $S(x)$. Since $S(x)$ is strictly concave over an infinite range, it will tend to minus infinity either when x tends to infinity or minus infinity. Now take a line that intersects the curve. This will intersect it at only two points. By strict convexity, there is only going to be a single point with maximum vertical distance to the line, which has to be greater than zero. Now imagine moving the line down by one vertical unit. The vertical distance now is greater than one. This means that the average entropy can increase more than one during mixture, which violates the upper variability bound. Therefore a line cannot be an ensemble space.

Note that an open segment is absolutely fine. We can, in fact, imagine a two state discrete classical ensemble space with the endpoint removed. This suggests that the convex space knows something about the geometry (probably parallelism and ratio of length - i.e. the affine structure).

Conjecture 3.38. *Let $e_1, e_2 \in \mathcal{E}$ be such that $pe + \bar{p}e_1 = pe + \bar{p}e_2$ for some $p \in (0, 1)$ and $e \in \mathcal{E}$. Then e_1 and e_2 are not disjunct.*

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

Definition 3.39. *Two sets of ensembles $U, V \subseteq \mathcal{E}$ are disjunct if $e_1 \perp e_2$ for all $e_1 \in U$ and $e_2 \in V$.*

Reducible (classical) ensembles

Definition 3.40. *An ensemble space \mathcal{E} is **reducible** (or **classical**) if distinct ensembles are also disjunct. That is, $e_1 \perp e_2$ implies $e_1 \perp\!\!\!\perp e_2$.*

Justification. Given that $e_1 \perp\!\!\!\perp e_2$ implies $e_1 \perp e_2$, classical spaces add the opposite implication. Therefore they exclude the case where two ensembles are disjunct but not distinct. This case describes two ensembles that have elements in common, but there is no ensemble corresponding to those common elements. That is, we cannot refine the ensembles into three distinct ones: one with only elements of the first, one with only elements of the second and one with elements of both. In other words, we cannot reduce the coarser description of the system into finer distinct descriptions. In the excluded case, then, the coarser description is irreducible into finer ensembles. This justifies the definition.

As we prove below, that case does not exist in classical mechanics, but exists in quantum mechanics. This justifies calling the property classical. \square

Remark. The above property should ultimately be responsible for the single decomposition in terms of pure states.

Proposition 3.41. *Continuous and discrete classical ensemble spaces are reducible.*

Proof. In both cases, ensembles are probability measures: the first over the Borel algebra of a symplectic manifold; the second over the power set of countably many elements. In both cases, the upper bound of the entropy is maximized if and only if the probability measures being combined have disjoint support. This means that two ensembles are disjunct if and only if the respective probability measures have disjoint support. Intuitively, two probability distributions can have a common sub-distribution if and only if they overlap. Therefore two ensembles representing probability measures with disjoint support are exactly ensembles that are distinct. Therefore, in both cases, the ensemble space is reducible. \square

3.6 Effective count

Proposition 3.42 (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 disjunct and $p = \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}$.*

Proof. If p is fixed, the upper variability bound of entropy is saturated only if \mathbf{e}_1 and \mathbf{e}_2 are disjunct by definition. The entropy maximum for the mixed ensemble can only be achieved when the elements are disjunct, for some value of p .

Now fix S_1 and S_2 . The entropy of the mixture depends only on p , so we need to find the p that maximizes the expression. Since \mathbf{e}_1 and \mathbf{e}_2 are disjunct, $S(p\mathbf{e}_1 + \bar{p}\mathbf{e}_2) = -p \log p - \bar{p} \log \bar{p} + pS_1 + \bar{p}S_2$ which is a smooth function of p .

$$\begin{aligned}
 0 &= \frac{dS}{dp} = \frac{d}{dp} S(\mathbf{e}) = \frac{d}{dp} (-p \log p - \bar{p} \log \bar{p} + pS_1 + \bar{p}S_2) \\
 &= -\log p - 1 + \log \bar{p} + 1 + S_1 - S_2 \\
 \log \frac{p}{\bar{p}} &= \log 2^{S_1} - \log 2^{S_2} \\
 \log \frac{p}{1-p} &= \log \frac{2^{S_1}}{2^{S_2}} \\
 p2^{S_2} &= (1-p)2^{S_1} \\
 p(2^{S_1} + 2^{S_2}) &= 2^{S_1} \\
 p &= \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}
 \end{aligned} \tag{3.43}$$

Having found the value of p that maximizes the entropy, we can calculate the maximum

entropy.

$$\begin{aligned}
\bar{p} &= 1 - \frac{2^{S_1}}{2^{S_1} + 2^{S_2}} = \frac{2^{S_2}}{2^{S_1} + 2^{S_2}} \\
S &= S(\mathbf{e}) = -p \log p - \bar{p} \log \bar{p} + pS_1 + \bar{p}S_2 \\
&= -\frac{2^{S_1}}{2^{S_1} + 2^{S_2}} \log \frac{2^{S_1}}{2^{S_1} + 2^{S_2}} - \frac{2^{S_2}}{2^{S_1} + 2^{S_2}} \log \frac{2^{S_2}}{2^{S_1} + 2^{S_2}} \\
&\quad + \frac{2^{S_1}}{2^{S_1} + 2^{S_2}} \log 2^{S_1} + \frac{2^{S_2}}{2^{S_1} + 2^{S_2}} \log 2^{S_2} \\
&= \frac{2^{S_1}}{2^{S_1} + 2^{S_2}} \log (2^{S_1} + 2^{S_2}) + \frac{2^{S_2}}{2^{S_1} + 2^{S_2}} \log (2^{S_1} + 2^{S_2}) \\
&= \frac{2^{S_1} + 2^{S_2}}{2^{S_1} + 2^{S_2}} \log (2^{S_1} + 2^{S_2}) \\
\log 2^S &= \log (2^{S_1} + 2^{S_2}) \\
2^S &= 2^{S_1} + 2^{S_2}
\end{aligned} \tag{3.44}$$

Therefore the maximum entropy obtainable through a mixture is $S = \log(2^{S_1} + 2^{S_2})$ which is obtained when \mathbf{e}_1 and \mathbf{e}_2 are disjunct and $p = \frac{2^{S_1}}{2^{S_1} + 2^{S_2}}$. \square

Definition 3.45. Let $U \subseteq \mathcal{E}$ be the subset of an ensemble space. The **effective count** of U is defined as $\text{efc}(U) = \sup(2^{S(\text{hull}(U))})$ if $U \neq \emptyset$ and $\text{efc}(U) = 0$ otherwise.

Proposition 3.46. The effective count is a set function that is

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

Proof. 1. The effective count takes a subset of \mathcal{E} and returns a real value and is therefore a set function. The exponential can only return non-negative values, therefore the effective count of a set is non-negative.

2. Let $U, V \subseteq \mathcal{E}$ such that $U \subseteq V$. If $U = \emptyset$, we have $\text{efc}(U) = 0$. Since $\text{efc}(V)$ is non-negative, $\text{efc}(U) \leq \text{efc}(V)$. If $U \neq \emptyset$, $\text{hull}(U) \subseteq \text{hull}(V)$ and therefore $2^{S(\text{hull}(U))} \subseteq 2^{S(\text{hull}(V))}$. This means that the supremum of the first set cannot be greater than the supremum of the second set, and therefore $\text{efc}(U) \leq \text{efc}(V)$. The effective count is a monotone set function.

3. Let $U, V \subseteq \mathcal{E}$ and let $\mathbf{e} \in \text{hull}(U \cup V)$. Then we can write $\mathbf{e} = p\mathbf{u} + \bar{p}\mathbf{v}$ for some $p \in [0, 1]$, $\mathbf{u} \in U$ and $\mathbf{v} \in V$. By 3.42 and the definition of effective count, $2^{S(\mathbf{e})} = 2^{S(\mathbf{u})} + 2^{S(\mathbf{v})} \leq \text{efc}(U) + \text{efc}(V)$. Since this is true for any element of $\text{hull}(U \cup V)$, the supremum of the exponential entropy cannot exceed the sum of the effective counts. Therefore $\text{efc}(U \cup V) \leq \text{efc}(U) + \text{efc}(V)$, the effective count is subadditive.

4. Let $U, V \subseteq \mathcal{E}$ be two disjunct subsets. Let $\{\mathbf{u}_i\} \subset U$ be a sequence of ensembles such that $2^{S(\mathbf{u}_i)} \rightarrow \text{efc}(U)$ and let $\{\mathbf{v}_i\} \subset V$ be a sequence of ensembles such that $2^{S(\mathbf{v}_i)} \rightarrow \text{efc}(V)$. Consider $\mathbf{e}_i = p_i\mathbf{u}_i + \bar{p}_i\mathbf{v}_i$ where $p_i = \frac{2^{S(\mathbf{u}_i)}}{2^{S(\mathbf{u}_i)} + 2^{S(\mathbf{v}_i)}}$. Then, by 3.42, $2^{S(\mathbf{e}_i)} = 2^{S(\mathbf{u}_i)} + 2^{S(\mathbf{v}_i)}$. This means that $2^{S(\mathbf{e}_i)} \rightarrow \text{efc}(U) + \text{efc}(V)$. Therefore $\text{efc}(U \cup V) \geq \text{efc}(U) + \text{efc}(V)$. Combining with the previous result, $\text{efc}(U \cup V) = \text{efc}(U) + \text{efc}(V)$. Therefore effective count, as set function, is additive over disjunct sets of ensembles. \square

3.7 Entropic geometry

We now show that entropy imposes a geometric structure on the ensemble space.

Definition 3.47. *Given two ensembles $\mathbf{e}_1, \mathbf{e}_2 \in \mathcal{E}$, the **mixing entropy**, also called *Jensen-Shannon divergence*, is the increase in entropy associated to their mixture. That is:*

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

Corollary 3.48. *The mixing entropy obeys the following bounds*

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

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

Proof. The bounds descend directly from the bounds on entropy. By strict concavity, $S(p_1\mathbf{e}_1 + p_2\mathbf{e}_2) \geq p_1S(\mathbf{e}_1) + p_2S(\mathbf{e}_2)$, which means $S(p_1\mathbf{e}_1 + p_2\mathbf{e}_2) - p_1S(\mathbf{e}_1) - p_2S(\mathbf{e}_2) \geq 0$, and in particular $S(\frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_2) - \frac{1}{2}S(\mathbf{e}_1) - \frac{1}{2}S(\mathbf{e}_2) = MS(\mathbf{e}_1, \mathbf{e}_2) \geq 0$. By strict concavity, the equality holds if and only if $\mathbf{e}_1 = \mathbf{e}_2$

By the upper variability bound, $S(p_1\mathbf{e}_1 + p_2\mathbf{e}_2) \leq I(p_1, p_2) + p_1S(\mathbf{e}_1) + p_2S(\mathbf{e}_2)$, which means $S(p_1\mathbf{e}_1 + p_2\mathbf{e}_2) - p_1S(\mathbf{e}_1) - p_2S(\mathbf{e}_2) \leq I(p_1, p_2)$, and in particular $S(\frac{1}{2}\mathbf{e}_1 + \frac{1}{2}\mathbf{e}_2) - \frac{1}{2}S(\mathbf{e}_1) - \frac{1}{2}S(\mathbf{e}_2) = MS(\mathbf{e}_1, \mathbf{e}_2) \leq I(\frac{1}{2}, \frac{1}{2}) = 1$. Given that \mathbf{e}_1 and \mathbf{e}_2 are disjunct by definition if they saturate the upper variability bound, the equality holds if and only if \mathbf{e}_1 and \mathbf{e}_2 are disjunct. \square

Proposition 3.49. *In discrete and continuous classical cases, the mixing entropy coincides with the Jensen-Shannon divergence. In quantum spaces it coincides with the quantum Jensen-Shannon divergence.*

Proof. Looking at the definitions [Jensen-Shannon divergence](#), one can see that

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

. The same is true for the quantum case:

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

.

\square

Definition 3.50. *An ensemble space is **geometric** if it is complemented and has a twice differentiable entropy with respect to the mixing coefficients.*

Conjecture 3.51. *A geometric ensemble space is a convex subset of a real vector space. Moreover, every finite dimensional subspace is a smooth manifold.*

Proof. Since a geometric ensemble space is complemented, it embeds into a real vector space. If conjecture [3.35](#) holds, then it is a topological vector space. Every finite dimensional subspace V , then, is locally R^n with a topology. This would be enough to claim V is a manifold. The differentiable structure is imposed by requiring coordinates that coordinates are twice differentiable with respect to the mixing coefficients. \square

Remark. The above restriction is useful to recover standard elements of differential geometry. It is an open question to understand whether other conditions already imply these requirements, or whether a more general setting can be developed. For example, in a classical space every ensemble can be understood as the mixture of orthogonal elements. Therefore the entropy can be proven the Shannon/Gibbs entropy, and therefore it is infinitely smooth with respect to the coefficients. A classical space, then, would be automatically geometric. Similar arguments may exist for a quantum space. Additionally, it may be that the convex structure may be enough to define a generalized version of tangent space.

Definition 3.52. Let $V \subseteq \mathcal{E}$ be a differentiable manifold embedded in the ensemble space. Let $\mathbf{e} \in \mathcal{E}$ be an ensemble and $T_{\mathbf{e}}$ the tangent space at that point. The **norm** of $\delta\mathbf{e} \in T_{\mathbf{e}}$ is given by

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

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

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

Theorem 3.53. Let \mathcal{E} be a geometric ensemble space and $V \subseteq \mathcal{E}$ a differentiable manifold embedded in the ensemble space. Then

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

and

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

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

Proof. To recover the first two expressions, we simply have to calculate the leading term. Since the entropy is twice differentiable, we can expand it as

$$S(\mathbf{e} + \delta\mathbf{e}) = S(\mathbf{e}) + \frac{\partial S}{\partial \mathbf{e}}\delta\mathbf{e} + \frac{1}{2}\frac{\partial^2 S}{\partial \mathbf{e}^2}\delta\mathbf{e}\delta\mathbf{e} + O(\delta\mathbf{e}^3). \quad (3.54)$$

Expanding the definition of MS , we have

$$\begin{aligned} MS(\mathbf{e}, \mathbf{e} + \delta\mathbf{e}) &= S\left(\frac{1}{2}\mathbf{e} + \frac{1}{2}(\mathbf{e} + \delta\mathbf{e})\right) - \frac{1}{2}S(\mathbf{e}) - \frac{1}{2}S(\mathbf{e} + \delta\mathbf{e}) \\ &= S\left(\mathbf{e} + \frac{1}{2}\delta\mathbf{e}\right) - \frac{1}{2}S(\mathbf{e}) - \frac{1}{2}S(\mathbf{e} + \delta\mathbf{e}) \\ &= S(\mathbf{e}) + \frac{\partial S}{\partial \mathbf{e}}\frac{1}{2}\delta\mathbf{e} + \frac{1}{2}\frac{\partial^2 S}{\partial \mathbf{e}^2}\frac{1}{2}\delta\mathbf{e}\frac{1}{2}\delta\mathbf{e} + O(\delta\mathbf{e}^3) \\ &\quad - \frac{1}{2}S(\mathbf{e}) - \frac{1}{2}\left(S(\mathbf{e}) + \frac{\partial S}{\partial \mathbf{e}}\delta\mathbf{e} + \frac{1}{2}\frac{\partial^2 S}{\partial \mathbf{e}^2}\delta\mathbf{e}\delta\mathbf{e} + O(\delta\mathbf{e}^3)\right) \\ &= S(\mathbf{e}) + \frac{1}{2}\frac{\partial S}{\partial \mathbf{e}}\delta\mathbf{e} + \frac{1}{8}\frac{\partial^2 S}{\partial \mathbf{e}^2}\delta\mathbf{e}\delta\mathbf{e} \\ &\quad - S(\mathbf{e}) - \frac{1}{2}\frac{\partial S}{\partial \mathbf{e}}\delta\mathbf{e} - \frac{1}{4}\frac{\partial^2 S}{\partial \mathbf{e}^2}\delta\mathbf{e}\delta\mathbf{e} + O(\delta\mathbf{e}^3) \\ &= -\frac{1}{8}\frac{\partial^2 S}{\partial \mathbf{e}^2}\delta\mathbf{e}\delta\mathbf{e} + O(\delta\mathbf{e}^3). \end{aligned} \quad (3.55)$$

Therefore

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

We can now substitute the norm to the definition of metric tensor. We have

$$\begin{aligned} g_{\mathbf{e}}(\delta \mathbf{e}_1, \delta \mathbf{e}_2) &= \frac{1}{2} (\|\delta \mathbf{e}_1 + \delta \mathbf{e}_2\|^2 - \|\delta \mathbf{e}_1\|^2 - \|\delta \mathbf{e}_2\|^2) \\ &= \frac{1}{2} \left(-\frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_1 + \delta \mathbf{e}_2, \delta \mathbf{e}_1 + \delta \mathbf{e}_2) + \frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_1, \delta \mathbf{e}_1) + \frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_2, \delta \mathbf{e}_2) \right) \\ &= -\frac{1}{2} \left(\frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_1, \delta \mathbf{e}_1) + \frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_1, \delta \mathbf{e}_2) + \frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_2, \delta \mathbf{e}_1) + \frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_2, \delta \mathbf{e}_2) \right) \\ &\quad - \frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_1, \delta \mathbf{e}_1) - \frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_2, \delta \mathbf{e}_2) \\ &= -\frac{\partial^2 S}{\partial \mathbf{e}^2}(\delta \mathbf{e}_1, \delta \mathbf{e}_2) \end{aligned} \quad (3.56)$$

Let us now take an embedded manifold $V \subseteq \mathcal{E}$. It will have a smooth structure inherited from the ensemble space.

Let us now show that g is a metric tensor. Restricted to V , $g|_V$ is the Hessian of the entropy S . Therefore it is a symmetric rank two tensor which means a symmetric bilinear function of the arguments. Since the entropy is strictly concave, the Hessian is negative definite and therefore g is positive definite. Therefore $(V, g|_V)$ is a Riemannian manifold. \square

Proposition 3.57. *For a classical ensemble space, the metric corresponds to the Fisher-Rao metric.*

Proof. Recall that, given a manifold of probability distributions over X parametrized by θ^i , the [Fisher-Rao metric](#) is defined as:

$$g_{ij} = - \int_X \frac{\partial^2 \log \rho}{\partial \theta^i \partial \theta^j} \rho dx. \quad (3.58)$$

For a classical ensemble, the entropy is given by

$$S(\rho) = - \int_X \rho \log \rho dx. \quad (3.59)$$

Let us now calculate the first two terms in the Taylor expansion around ρ with a variation $\delta \rho$. Recall that:

$$\begin{aligned} \log(x + dx) &= \log(x) + d_x \log x dx + \frac{1}{2} d_x d_x \log x dx^2 + O(dx^3) \\ &= \log(x) + \frac{1}{x} dx - \frac{1}{2} \frac{1}{x^2} dx^2 + O(dx^3) \end{aligned} \quad (3.60)$$

We have:

$$\begin{aligned}
S(\rho + \delta\rho) &= - \int_X (\rho + \delta\rho) \log(\rho + \delta\rho) dx \\
&= - \int_X (\rho + \delta\rho) \left[\log \rho + \frac{1}{\rho} \delta\rho - \frac{1}{2\rho^2} \delta\rho^2 + O(\delta\rho^3) \right] dx \\
&= - \int_X \rho \log \rho dx - \int_x [\log \rho + 1] dx \delta\rho - \int_X \left[\frac{1}{\rho} - \frac{\rho}{2\rho^2} \right] dx \delta\rho^2 + \int_X dx O(\delta\rho^3) \\
&= - \int_X \rho \log \rho dx - \int_x [\log \rho + 1] dx \delta\rho - \frac{1}{2} \int_X \frac{1}{\rho} dx \delta\rho^2 + \int_X dx O(\delta\rho^3)
\end{aligned} \tag{3.61}$$

$$\begin{aligned}
\frac{\partial^2 S}{\partial \rho^2}(\delta\rho, \delta\rho) &= - \int_X \frac{1}{\rho} \delta\rho^2 dx = - \int_X \frac{1}{\rho} \delta\rho^2 dx + 0 = - \int_X \frac{1}{\rho} \delta\rho^2 dx + \delta^2(1) \\
&= - \int_X \frac{1}{\rho} \delta\rho^2 dx + \delta^2 \int_X \rho dx = - \int_X \frac{1}{\rho} \delta\rho^2 dx + \int_X \delta^2 \rho dx \\
&= \int_X \rho dx \left[-\frac{1}{\rho^2} \delta\rho^2 + \frac{1}{\rho} \delta^2 \rho \right] = \int_X \rho dx \delta \left[\frac{1}{\rho} \delta\rho \right] \\
&= \int_X \rho dx \delta^2 \log \rho
\end{aligned} \tag{3.62}$$

Let us consider a family of ensembles charted by a set of parameters θ^i . We have:

$$g_e(d\theta^1, d\theta^2) = - \frac{\partial^2 S}{\partial \rho^2} \left(\frac{\partial \rho}{\partial \theta^1} d\theta^1, \frac{\partial \rho}{\partial \theta^2} d\theta^2 \right) = - \int_X \rho dx \frac{\partial^2 \log \rho}{\partial \theta^1 \partial \theta^2} d\theta^1 d\theta^2 \tag{3.63}$$

which recovers the Fisher-Rao metric. □

Remark. For quantum mechanics, the situation is more complicated as there are different definitions of Fisher metrics (see [RLD Fisher Information Bound for Multiparameter Estimation of Quantum Channels](#) or [Quantum Estimation Theory](#)). We will therefore just show a general connection, without worrying about the mathematical details.

Proposition 3.64. *For a quantum ensemble space, the metric corresponds to the Bures metric and the Quantum Fisher information metric.*

Proof. We can repeat the calculations from the classical case and obtain

$$\frac{\partial^2 S}{\partial \rho^2}(\delta\rho, \delta\rho) = - \text{tr}(\rho^{-1} \delta\rho \delta\rho) \tag{3.65}$$

Let us consider a family of ensembles charted by a set of parameters θ^i . We have:

$$g_e(d\theta^1, d\theta^2) = - \frac{\partial^2 S}{\partial \rho^2} \left(\frac{\partial \rho}{\partial \theta^1} d\theta^1, \frac{\partial \rho}{\partial \theta^2} d\theta^2 \right) = \text{tr} \left(\rho^{-1} \frac{\partial \rho}{\partial \theta^1} \frac{\partial \rho}{\partial \theta^2} \right) d\theta^1 d\theta^2 \tag{3.66}$$

which recovers one version of the Fisher-Rao metric.

Recalling that the right logarithmic derivative L^R satisfies $\frac{\partial \rho}{\partial \theta} = \rho L^R$, we can write

$$g_e(d\theta^1, d\theta^2) = \text{tr} \left(\rho^{-1} \frac{\partial \rho}{\partial \theta^1} \frac{\partial \rho}{\partial \theta^2} \right) d\theta^1 d\theta^2 = \text{tr}(\rho^{-1} \rho L_1^R \rho L_2^R) d\theta^1 d\theta^2 = \text{tr}(\rho L_1^R L_2^R) d\theta^1 d\theta^2 \tag{3.67}$$

which recovers the RLD Fisher information. □

Conjecture 3.68. *The mixing entropy is the square of a distance function.*

Proof. This is something that is true for both classical and quantum version of the JSD. It is not clear whether this generalizes. The arc length provided by the metric is indeed a distance function, which has a square root over the metric applied to the differentials. The JSD will correspond to the energy (or action) of the curve. This does not seem to be, in general, the square of a distance function. \square

Remark. How exactly to make these statements precise in the general case depends on finding a suitable generalized notion of differentiability. It is unclear whether this notion has already been mathematically formulated.

3.8 Ensemble subspaces

We now want to define a notion of subspace in a way that recovers the notions of subspaces we have in both classical and quantum mechanics. Given that orthogonality in all three cases corresponds to ensembles being disjoint, disjoint ensembles will need to belong to different subspaces. The idea is to reconstruct subspaces only from the disjoint relationship.

Irreflexive symmetric relations and topped \cap -structures

We will first prove a more general set of results, on a generic set (not necessarily an ensemble space) with an irreflexive symmetric relation (not necessarily disjointness of ensembles). This will give us a more general tool to investigate possible different notion of subspaces.

Definition 3.69. *Let X be a set and $R \subseteq X \times X$ a symmetric relation. Given a subset $U \subseteq X$, we define the R -complement to be*

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

Proposition 3.70. *Let X be a set and $R \subseteq X \times X$ a symmetric relation. Then:*

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

Proof. 1. Suppose $a \in V^{RC}$. Then, by definition, $\forall b \in V, aRb$. Since $U \subset V$, it is also true that $\forall b \in U, aRb$. Therefore $a \in U^{RC}$ by definition. Since a was arbitrary, $V^{RC} \subseteq U^{RC}$.

2. By expanding the definition of complement, we have $(U^{RC})^{RC} = \{a \in X \mid \forall b \in U^{RC}, aRb\} = \{a \in X \mid \forall b \in \{c \in X \mid \forall d \in U, cRd\}, aRb\} = \{a \in X \mid \forall b \in X \text{ s.t. } (\forall d \in U, bRd), aRb\}$.

Let $a \in U$ and let $b \in X$ such that $\forall d \in U, bRd$. Since $a \in U$ and bRd for all $b \in U$, we have bRa in particular. Since R is symmetric, aRb . Given that b was arbitrary, we conclude that $\forall b \in X \text{ s.t. } (\forall d \in U, bRd), aRb$. Therefore $a \in (U^{RC})^{RC}$ by definition of complement. Given that a was arbitrary, $U \subseteq (U^{RC})^{RC}$.

3. We again expand the definition and have $((U^{RC})^{RC})^{RC} = \{a \in X \mid \forall b \in (U^{RC})^{RC}, aRb\}$.

Let $x \in ((U^{RC})^{RC})^{RC}$. Then $\forall b \in (U^{RC})^{RC}, xRb$ by definition of the complement. Since by 1. $U \subset (U^{RC})^{RC}$, we can restrict the previous expression to only the elements of U , and

therefore $\forall b \in U, xRb$. But this means that $x \in U^{RC}$ by definition of the complement. Since x was arbitrary, $((U^{RC})^{RC})^{RC} \subseteq U^{RC}$. But by 1., we also have $U^{RC} \subseteq ((U^{RC})^{RC})^{RC}$ since U^{RC} is just a set onto which we can apply the complement twice. By two-way containment, we have $U^{RC} = ((U^{RC})^{RC})^{RC}$.

4. Let $U, V \subseteq X$ such that $U^{RC} = (V^{RC})^{RC}$. Applying the complement on each side, $(U^{RC})^{RC} = ((V^{RC})^{RC})^{RC}$. By the previous property $((V^{RC})^{RC})^{RC} = V^{RC}$ and therefore $(U^{RC})^{RC} = V^{RC}$. Switching U and V proves the other direction.

5. We have:

$$\begin{aligned} \left(\bigcup_{i \in I} U_i\right)^{RC} &= \{a \in X \mid \forall b \in \bigcup_{i \in I} U_i, aRb\} \\ &= \{a \in X \mid \forall U_i, \forall b \in U_i, aRb\} \\ &= \{a \in X \mid \forall U_i, a \in \{c \in X \mid \forall b \in U_i, cRb\}\} \\ &= \bigcap_{i \in I} \{c \in X \mid \forall b \in U_i, cRb\} \\ &= \bigcap_{i \in I} (U_i)^{RC} \end{aligned}$$

Note that this property does not rely on the symmetry of R .

6. Let $a \in X$. There is no $b \in \emptyset$ such that aRb . Therefore $\forall b \in \emptyset, aRb$. This means that $a \in \emptyset^{RC}$. Since a was arbitrary, $X = \emptyset^{RC}$. \square

Proposition 3.71. *Let X be a set and $R \subseteq X \times X$ a symmetric and irreflexive relation. Then*

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

Proof. 1. Let $a \in U$. Since R is irreflexive, aRa is false. Therefore it is not true that, for all $b \in U, aRb$. This means that $a \notin U^{RC}$. Since that a was arbitrary, $U \cap U^{RC} = \emptyset$.

2. Suppose $a \in X^{RC}$. Then for all $b \in X, aRb$. In particular, we would have aRa , which can't be true since R is irreflexive. Therefore $a \notin X^{RC}$ and, since a is arbitrary, $X^{RC} = \emptyset$. \square

Definition 3.72. *Let X be a set and $R \subseteq X \times X$ a symmetric relation. Let $U \subseteq X$. The **R -subspace generated by U** is $\langle U \rangle_R = (U^{RC})^{RC}$. An **R -subspace** of X is a set $U \subseteq X$ such that $U = \langle U \rangle_R$. The **lattice of R -subspaces** is the set $\mathfrak{L} = \{U \subseteq X \mid U = \langle U \rangle_R\}$ ordered by inclusion.*

Corollary 3.73. *The lattice of R -subspaces \mathfrak{L} is a topped \cap -structure on X and therefore is also a complete lattice.*

Proof. The set \mathfrak{L} is a collection of subsets of X . Let $\{U_i\}_{i \in I} \subseteq \mathfrak{L}$ be a non-empty family. Then, using the definition of subspace, and the third and fifth properties of 3.70, we have

$$\begin{aligned} \bigcap_{i \in I} U_i &= \bigcap_{i \in I} (U_i^{RC})^{RC} = \left(\bigcup_{i \in I} U_i^{RC}\right)^{RC} \\ &= \left(\left(\left(\bigcup_{i \in I} (U_i^{RC})^{RC}\right)^{RC}\right)^{RC}\right)^{RC} = \left(\left(\bigcap_{i \in I} (U_i^{RC})^{RC}\right)^{RC}\right)^{RC} \\ &= \left(\left(\bigcap_{i \in I} U_i\right)^{RC}\right)^{RC}. \end{aligned}$$

Therefore $\bigcap_{i \in I} U_i \in \mathfrak{L}$. This means \mathfrak{L} is an \bigcap -structure. Using the second property of 3.71, $X = \emptyset^{RC} = (X^{RC})^{RC}$. Therefore \mathfrak{L} is a topped \bigcap -structure. This also means that it is a complete lattice. \square

Corollary 3.74. *The R -subspace operator satisfies the following properties*

1. $U \subseteq \langle U \rangle_R$
2. $U \subseteq V \implies \langle U \rangle_R \subseteq \langle V \rangle_R$
3. $\langle \langle U \rangle_R \rangle_R = \langle U \rangle_R$

and is therefore a closure operation.

Proof. 1. The first property is true by the second property of 3.71.

2. Using the first property of 3.71 we have $U \subseteq V$ implies $V^{RC} \subseteq U^{RC}$ which in turns implies $(U^{RC})^{RC} \subseteq (V^{RC})^{RC}$. Therefore $\langle U \rangle_R \subseteq \langle V \rangle_R$.

3. Using the fifth property of 3.71 we have $\langle \langle U \rangle_R \rangle_R = (((U^{RC})^{RC})^{RC})^{RC} = (U^{RC})^{RC} = \langle U \rangle_R$. \square

Proposition 3.75. *Let X be a set and $R \subseteq X \times X$ a symmetric and irreflexive relation. Then*

1. $\langle U \rangle_R$ is the smallest R -subspace containing U
2. if $U, V \in \mathfrak{L}$ then $U = V^{RC} \iff U^{RC} = V$

Proof. 1. Let $U \subseteq X$ and $V \in \mathfrak{L}$ such that $U \subseteq V$ and $V \subseteq \langle U \rangle_R$. Since $U \subseteq V$, using the second property of 3.74, $\langle U \rangle_R \subseteq \langle V \rangle_R = V$. Since $V \subseteq \langle U \rangle_R$ and $\langle U \rangle_R \subseteq V$, $\langle U \rangle_R = V$. This means that no R -subspace that contains U is smaller than $\langle U \rangle_R$.

2. Since $U, V \in \mathfrak{L}$, $U = (U^{RC})^{RC} = V^{RC}$ which, by the fourth property of 3.70, implies $U^{RC} = (V^{RC})^{RC} = V$. Switching U and V proves the other direction. \square

Proposition 3.76. *Let X be an inner product space and $R = \{(a, b) \mid \langle a, b \rangle = 0\}$. Then:*

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

Proof. 1. Given that the inner product is symmetric, so will be R . Given that no vector is orthogonal to itself, R is irreflexive.

2. The orthogonal complex is defined as $U^\perp = \{a \in X \mid \forall b \in U, \langle a, b \rangle = 0\}$. Since $aRb \iff \langle a, b \rangle = 0$, $U^{RC} = U^\perp$.

3. We have $\langle U \rangle_R = (U^\perp)^\perp$ which returns the smallest closed subspace that contains U . \square

This shows that full structure of subspaces of an inner product space can be fully recovered only from pairwise orthogonality.

Subspaces of ensemble spaces

We will now use the previous results where X is an ensemble space and R is the disjunct relation between two ensembles.

Proposition 3.77. *Let \mathcal{E} be an ensemble space, disjunctness \perp is an irreflexive symmetric relation.*

Proof. Since any ensemble is not distinct from itself, it is neither disjunct from itself. Therefore disjunctness is irreflexive. Two elements are disjunct if they saturate the upper entropy bound. This does not depend on their order. Therefore disjunctness is symmetric. \square

Definition 3.78. Let \mathcal{E} be an ensemble space and $X \subseteq \mathcal{E}$ be a subset. The **disjunct complement** $X^\perp \subseteq \mathcal{E}$ is the set of all ensembles that are disjunct from all elements of X . An **ensemble subspace** is a subset $X \subseteq \mathcal{E}$ such that $X = (X^\perp)^\perp$.

Remark. Since subspaces are subsets of the ensembles space, the dimension is automatically defined on subspaces.

Proposition 3.79. Let $X \subseteq \mathcal{E}$ be an ensemble subspace. The $\text{hull}(U) \subseteq X$ for all $U \subseteq X$.

Proof. Let $U \subseteq X$. Then $U \perp X^\perp$. But since mixtures preserve disjunctness, we also have $\text{hull}(U) \perp X^\perp$. Therefore $\text{hull}(U) \subseteq X$. \square

Proposition 3.80. Let \mathcal{E} be a discrete classical theory. Then

1. the lattice of subspaces \mathfrak{L} is the set of all possible simplexes generated by all possible combinations of extreme points
2. the effective count of each subspace is the number of extreme points

Proof. Let \mathcal{E} be a discrete classical ensemble space. Then each ensemble is a sequence p_i such that $\sum_i p_i = 1$. The dimensionality of the space fixes the range of i . Note that, given that the space is discrete, the collection p_i can be understood as a continuous function of the pure elements.

Note that two probability distributions will be disjunct if and only if their supports are disjoint. Therefore disjunctness for a discrete classical ensemble space is the same as having disjoint support. This means that a subspace is given by all possible probability distributions over a subset of all pure states.

If U is the set of all probability distributions over n cases, the distribution with highest entropy is achieved by a continuous distribution and will be $S(p_i) = \log n$. Therefore the dimension of U will be n . \square

Remark. By support, we will mean the set-theoretic support, that is the set of points for which the function is non-zero.

Conjecture 3.81. Let \mathcal{E} be a continuous classical theory. Then

1. the lattice of subspaces \mathfrak{L} is the set of all subspaces of continuous probability densities with support within an open set that is the interior of its own closure
2. the effective count of each subspace is the Liouville measure

Proof. Let \mathcal{E} be a continuous classical ensemble space. Then each ensemble is a continuous function over a symplectic manifold X that integrates to one. Since the function is continuous, the support is an open set.

Consider two continuous functions that integrate to one. They will be disjunct if and only if their supports are disjoint. Therefore disjunctness for a continuous classical ensemble space is the same as having disjoint support.

Take a set of ensembles $U \subseteq \mathcal{E}$. An ensemble $e \in \mathcal{E}$ will be disjunct from all ensembles in U if and only if its support is disjoint from the union of all the supports of all the elements of U .

Therefore U^\perp is the set of all ensembles whose support is a subset of the exterior of the union of all supports of elements of U . With a similar logic, $(U^\perp)^\perp$ is the set of all ensembles whose support is a subset of the exterior of the support of elements of U^\perp . This corresponds to the interior of the closure of the union of all supports of elements of U . Therefore, any subspace is a set of all continuous probability densities with support within an open set $V \subseteq X$ that is the interior of its own closure.

Now take an open set $V \subseteq X$ that is the interior of its own closure and consider the space U of all continuous probability densities that have support within V . Then U^\perp is the subspace of all continuous probability densities that have support within $\text{ext}(V)$ and $(U^\perp)^\perp$ is the subspace of all continuous probability densities that have support within $\text{ext}(\text{ext}(V)) = \text{int}(\overline{V}) = V$. Therefore U is a subspace.

Let $U \in \mathcal{L}$ be a ensemble subspace. If $\mathbf{e} \in U$, then $S(\mathbf{e}) \leq \text{efc}(U)$. Moreover, there must be a sequence \mathbf{e}_i such that $S(\mathbf{e}_i) \rightarrow \text{efc}(U)$. Let $V \subseteq X$ be the open set corresponding to the largest support of all elements in U . Then $S(\mathbf{e}) \leq \log \mu(V)$ as the highest entropy is given by a uniform distribution over V . The uniform distribution is discontinuous, and therefore is not a member of \mathcal{E} . However, we can construct a sequence of continuous functions whose limit is the uniform distribution. This would correspond to a sequence of ensembles for which $S(\mathbf{e}_i) \rightarrow \log \mu(V)$. Therefore $\log \mu(V)$ is the supremum of the entropy for the subspace, and therefore the effective count is $\mu(V)$, the Liouville measure. \square

Conjecture 3.82. *Let \mathcal{E} be a quantum ensemble space. Then*

1. *the lattice of subspaces \mathcal{L} is the set of all possible density operators for each subspace of the Hilbert space*
2. *the effective count of each subspace is the dimension of the corresponding Hilbert space*

Proof. Let \mathcal{E} be a quantum ensemble space. Then each ensemble is a density operator defined over a Hilbert space \mathcal{H} .

Consider two density operators. They will be disjunct if and only if they are defined on orthogonal subspaces. That is, $\mathbf{e}_1|\psi\rangle \neq 0$ only if $\mathbf{e}_2|\psi\rangle = 0$ and vice-versa. This means that a subspace of the ensemble space corresponds to a subspace of the vector space \mathcal{H} .

Let $U \in \mathcal{L}$ be a ensemble subspace. If it is finite dimensional, the ensemble with maximal entropy will correspond to the maximally mixed state I/n where n is the dimension of the subspace. The entropy will be given by $\log n$, which means $\text{efc}(U) = n$. If U is infinite dimensional, then there is no upper bound to the entropy and $\text{efc}(U) = +\infty$. \square

3.9 Contexts and probability measure

Definition 3.83. *Let \mathcal{L} be the lattice of subspaces and $\wedge, \vee, (\cdot)^\perp$ be respectively the join, meet and disjunct complement. A **context** is a lattice of subspaces $C \subseteq \mathcal{L}$ such that*

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

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

Proposition 3.84. *Let $C \subseteq \mathfrak{L}$ be a context. Then efc is an additive measure over the context. That is, $\text{efc}(X \vee Y) = \text{efc}(X) + \text{efc}(Y)$ for all $X, Y, X \vee Y \in C$ such that $X \cap Y = \emptyset$.*

Proof. Since $X \vee Y$ contains both X and Y , we have $X \cup Y \subseteq (X \vee Y)$. By 3.79, $\text{hull}(X \cup Y) \subseteq (X \vee Y)$. By \square

Definition 3.85. *An ensemble $\mathbf{e} \in \mathcal{E}$ is **compatible** with a context $C \subseteq \mathfrak{L}$ if for any $X \in C$ we can write $\mathbf{e} = p_i \mathbf{x}_i + \lambda_i \mathbf{y} + \epsilon_i \mathbf{e}_i$ where $\mathbf{x}_i \in X$ and $\mathbf{y}_i \in X^\perp$ are maximal component sequences.*

Corollary 3.86. *Let $\mathbf{e} \in \mathcal{E}$ an ensembles compatible with a context $C \subseteq \mathfrak{L}$. Then $p_{\mathbf{e}}(X^\perp) = 1 - p_{\mathbf{e}}(X)$.*

Proof. Since \mathbf{e} is compatible with C , we can write $\mathbf{e} = p_i \mathbf{x}_i + \lambda_i \mathbf{y} + \epsilon_i \mathbf{e}_i$ where $\mathbf{x}_i \in X$ and $\mathbf{y}_i \in X^\perp$ are maximal component sequences. Note that $p_{\mathbf{e}}(X^\perp) = 1 - p_{\mathbf{e}}(X)$ if and only if $\epsilon_i \rightarrow 0$. Suppose it didn't. Then every \mathbf{e}_i would have a component that is neither in X or X^\perp . That would mean it is disjunct from all the element of X and of X^\perp . But X contains all the elements disjunct from X^\perp and vice-versa. Therefore such component cannot exist, $\epsilon_i \rightarrow 0$ and $p_{\mathbf{e}}(X^\perp) = 1 - p_{\mathbf{e}}(X)$. \square

Conjecture 3.87. *Let $C \subseteq \mathfrak{L}$ be a context and $\mathbf{e} \in \mathcal{E}$ an ensemble compatible with the context. Then $p_{\mathbf{e}}$ is an additive measure over the context. That is, $p_{\mathbf{e}}(X \vee Y) = p_{\mathbf{e}}(X) + p_{\mathbf{e}}(Y)$ for all $X, Y, X \vee Y \in C$ and $X \cap Y = \emptyset$.*

Remark. The additivity of probability is more of a special condition than one may expect. Naively, one would expect that if U and V are distinct, then $p_{\mathbf{e}}(U \cup V) = p_{\mathbf{e}}(U) + p_{\mathbf{e}}(V)$. That is, the component of \mathbf{e} in $U \cup V$ is simply the sum of the components within U and V , which are going to be distinct (i.e. will not overlap). However this is not true in quantum mechanics.

Take a single qubit and let \mathbf{e} be the maximally mixed state. Then $p_{\mathbf{e}}(\{x\}) = \frac{1}{2}$ for any pure state x . However, if we take two pure states x and y that are not orthogonal, the maximally mixed state is not a mixture of x and y , which means $p_{\mathbf{e}}(\{x, y\}) < 1 = p_{\mathbf{e}}(\{x\}) + p_{\mathbf{e}}(\{y\})$.

Conjecture 3.88. *If \mathfrak{L} is a context, then all probability measure are additive and the ensemble space is classical probability space (i.e. a space of additive measures).*

Conjecture 3.89. *Let \mathcal{E} be a discrete finite classical ensemble space. Then*

1. *the lattice of subspaces \mathfrak{L} is the set of all possible simplexes generated by all possible combinations of extreme points*
2. *the dimension of each subspace is the number of extreme points*
3. *the pure states are the extreme points.*

Conjecture 3.90. *Let \mathcal{E} be a continuous classical ensemble space. Then*

1. *the lattice of subspaces \mathfrak{L} is the set of all subspaces of probability measures with support over an open set U*
2. *the dimension of each subspace is the Liouville measure*
3. *the pure states are the points of the symplectic manifold.*

Conjecture 3.91. *Let \mathcal{E} be a quantum ensemble space. Then*

1. the lattice of subspaces \mathfrak{L} is the set of all possible density operators for each subspace of the Hilbert space
2. the dimension of each subspace is the dimension of the corresponding Hilbert space
3. the pure states are the subspaces of dimension one (i.e. the rays of the Hilbert space).

Conjecture 3.92. Let \mathcal{E} be a reducible ensemble space and $X, Y \subseteq \mathcal{E}$ two subspaces. Then $\dim(X \vee Y) = \dim(X) + \dim(Y)$.

Definition 3.93. An ensemble is **decomposable** if it is the mixture of two or more components. An ensemble space is **infinitesimally decomposable** if all ensembles are decomposable. An ensemble space is **finitely decomposable** if any ensemble can be decomposed into elements that are not decomposable.

Proposition 3.94. The discrete classical ensemble space and the quantum ensemble space are finitely decomposable. The continuous classical space is infinitesimally decomposable.

Proof. Pure states in discrete classical and quantum ensembles can no longer be decomposed, and every other ensemble is the mixture of pure states. Therefore these spaces are finitely decomposable.

The ensembles of a continuous classical ensemble spaces are probability distributions that are continuous and therefore must have finite support. However, given any finite support, we can always find distribution with smaller support and therefore decompose any distribution in the sum of two distribution. Therefore the space is infinitesimally decomposable. \square

Conjecture 3.95. An ensemble space \mathcal{E} is a finite discrete classical ensemble space if and only if the following properties are all satisfied

1. \mathcal{E} is reducible (i.e. classical)
2. every state with zero entropy is non-decomposable
3. the entropy is bounded

Proof. Let us first show that all finite discrete classical \square

Remark. The non-decomposable ensembles are exactly the extreme points of the convex space.

Pure states and state space

Definition 3.96. A **pure state** is a limit of an ensemble with the smallest entropy. Formally, a pure state $s \subset \mathfrak{L}$ is a non-empty collection of subspaces such that

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

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

Remark. The above should correspond to an [ultrafilter](#).

Definition 3.97. The set of states of a subspace $X \in \mathfrak{L}$ is the set $\text{supp}(X) = \{s \in S(\mathcal{E}) \mid X \in s\}$. The topology of the pure states is the topology over the set of pure states generated by the set of states of all subspaces.

Definition 3.98. Given an ensemble space \mathcal{E} , its **state space** is the set of all pure states $S(\mathcal{E})$ equipped with

1. the topology generated by the set of states of all subspaces
2. the measure given by the dimension on the Borel algebra

Conjecture 3.99. The state space of a discrete classical theory is the set of extreme points with a discrete topology and the counting measure.

Conjecture 3.100. The state space of a continuous classical theory is the phase space manifold with the standard topology and the Liouville measure.

Conjecture 3.101. The state space of a quantum theory is the projective Hilbert space with its topology. The measure applied to a subspace returns its dimensionality.

3.10 Lessons learned

Subspaces from convex structure

Insight 3.102. Convex structure, by itself, cannot define subspaces.

Initially, we tried recovering the notion of subspaces purely from the convex structure. We did make some progress in the finite dimensional case, but ultimately this does not work. The ultimate problem is that it cannot be generalized to the classical continuous case. The problem of rate of convergence is described below in the example. Even if that problem is fixed, there would be nothing to determine the dimensionality of U , and we could create convex maps that “stretch” the space.

The attempt was the following:

Definition 3.103. Let \mathcal{E} be a convex space and $X \subseteq \mathcal{E}$ be a subset. We say that X is a **subspace** of \mathcal{E} if it contains all the convex combinations and all the components of its elements. That is, for every $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \in \mathcal{E}$ and $\lambda \in (0, 1)$ such that $p\mathbf{e}_1 + \bar{p}\mathbf{e}_2 = \mathbf{e}_3$ we have:

- $\mathbf{e}_1, \mathbf{e}_2 \in X$ implies $\mathbf{e}_3 \in X$
- $\mathbf{e}_3 \in X$ implies $\mathbf{e}_1, \mathbf{e}_2 \in X$

Definition 3.104. Let \mathcal{E} be a convex space and $X \subseteq \mathcal{E}$. The **convex span** of X , noted $\text{cospan}(X)$, is the smallest subspace containing X .

Remark. As defined, the convex span of two elements will include all their possible mixtures (i.e. the segment that connects them), all possible decompositions (i.e. all lines that pass through them) plus, recursively, all other mixtures and decompositions that can be reached from those. Physically, the idea is that not all ensembles, pure states in particular, cannot be physically realized. Therefore, the inclusion of pure states in our convex space stems from a theoretical idealization useful to decompose and study the problem. It makes sense, then, that a subspace comes with all its idealizations.

Example 3.105. Classical discrete subspaces. Let S be a set of n possible discrete states and let \mathcal{E} the space of probability distributions over the set S (i.e. \mathcal{E} is an n -simplex and S are its extreme points). A subspace X of \mathcal{E} is a convex hull of a subset U of S . That is, a subspace

of \mathcal{E} is the space of probability distributions over a subset of the cases. Geometrically, it is one of the sides (possibly recursively) of the simplex.

To see this, first note that the convex hull X of any subset U of extreme points S is a subspace. In fact, it will contain all convex combinations of U , and any element can only be decomposed in convex combinations of U . Second, note that only convex hulls of a subset of extreme points can be a subspace. In fact, any element of \mathcal{E} can be expressed as a non-trivial convex combination of a set of extreme points U . Therefore, if an element is present in a subspace X , then $U \subset X$, which means all elements of the convex hull of U are in X .

Remark. The definition does not entirely work in the continuous case. Let S be a symplectic manifold and let \mathcal{E} be the space of probability distributions over S (i.e. the space of continuous functions over S that are integrable and integrate to one). We would like to have a definition for which a subspace of \mathcal{E} is a set of functions whose support is a subset of an open region $U \subseteq S$. That is, a subspace contains all probability distributions defined over a subset of the cases. One problem is that continuity forces the function to go to zero on the boundary of U , and the speed of the convergence cannot be change with a finite convex combination. That is, the convex combination of functions that go down like an exponential will also go down like an exponential. The issue seems to be related to finding the correct closure for infinite convex combinations in the definition of convex space.

Example 3.106. Quantum subspaces. Let \mathcal{H} be an n -dimensional Hilbert space and let \mathcal{E} be the space of density matrices (i.e. positive semi-definite self-adjoint operators with trace one). A subspace X of \mathcal{E} is the space of density matrices of a subspace U of \mathcal{H} . That is, a subspace of \mathcal{E} is the space of mixed states over a subspace of pure states.

To see this, first note that the space of density matrices X of a subspace U of \mathcal{H} is a subspace of \mathcal{E} . In fact, X it will contain all convex combinations of its elements. Moreover, any element $x \in X$ can only be decomposed in a convex combination of pure states of U . Therefore any convex decomposition of x has all its elements in X . Second, note that only the space of density matrices X of a subspace U of \mathcal{H} is a subspace of \mathcal{E} . In fact, any element x of \mathcal{E} can be expressed as a non-trivial convex combination of orthogonal pure states, its eigenstates. These elements will span a subspace U of \mathcal{H} . From those elements, we can construct an equal mixture which represents the maximally mixed states and, mathematically, is the identity operator I/m divided by the number of elements $m \leq n$ of U . The equal mixture of any orthogonal basis of U will also give the maximally mixed state. Therefore, given an element x , any subspace that contain x will also contain a basis of U , the maximally mixed state I/n , all possible basis of U , which means all the pure states, and finally all convex combinations of the pure states, which means all possible density matrices, all possible mixed states.

Definition 3.107. A convex space \mathcal{E} is **closed** if it contains all its extreme points.

Definition 3.108. Let \mathcal{E} be a convex space and $X \subset \mathcal{E}$ be a subspace of \mathcal{E} . The **dimension** of X , noted $\dim(X)$, is the minimum number of extreme points whose convex span is X .

Conjecture 3.109. Let \mathcal{E} be closed convex space of finite dimensions. Then:

- \mathcal{E} is a simplex if and only if every two dimensional subspace is a line segment
- the set of extreme points of \mathcal{E} is a complex projective space if and only if every two dimensional subspace is a two dimensional sphere (\mathcal{E} is the space of density matrices).

Remark. The only if is easy to show based on the discussion above. The if part does not actually work. Take 6 points in a 4 dimensional euclidean space such that they form two triangles that intersect at a single point. All two dimensional convex spaces will be segments. However, the whole convex hull is not a simplex.

Part IV
Appendix

Appendix A

Reference sheets for math and physics

A.1 Set theory

	Name	Meaning
$A = \{1, 2, 3\}$	set	a collection of elements
$\mathbb{N} = \{0, 1, 2, \dots\}$	natural numbers	the set of numbers one uses to count
$\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$	integers	the set of all whole numbers
\mathbb{Q}	rationals	the set of all fractions
\mathbb{R}	reals	the set of numbers with infinite precision
\mathbb{C}	complex	the set of numbers that represent a two dimensional vector or rotation
$a \in A$	in	whether the element a is contained in A
$A \subseteq B$	subset	a set that only contains elements of the other set
$A \subset B$	proper subset	a set that only contains elements of the other set but not all of them; it is a subset but is not the same set
$A \supseteq B$	superset	a set that contains all elements of the other set
$A \supset B$	proper superset	a set that contains all elements of the other set but not just them; it is a superset but is not the same set
$A \cup B$	union	the set of all elements contained in either sets
$A \cap B$	intersection	the set of all elements contained in both sets
$A \setminus B$	subtraction	the set of elements in A that are not in B
A^C	complement	the set of all elements that are not in A it is equal to $A \setminus U$ where U is the set of all elements, which depends on context
$A \times B$	Cartesian product	the set of all ordered pairs (a, b) with $a \in A$ and $b \in B$
2^A	power set	the set of all possible subsets of A

	Name	Meaning
$f : A \rightarrow B$	function	a map that for every element A returns an element of B
	injective function	a function that every distinct element of A map that for every element A returns an element of B
B^A		the set of all possible functions $f : A \rightarrow B$
$C(A, B)$		the set of all continuous functions $f : A \rightarrow B$

Credits

Created by: Gabriele Carcassi
Written by: Gabriele Carcassi and Christine A. Aidala

Subject-matter advisors (Math): Mark Greenfield (Ch. II.1,II.2,II.3)
review prompted significant technical changes
Additional subject-matter advisors (Phil): Josh Hunt (Ch. II.1)
review prompted significant technical improvements

Diagrams and figures: Matteo Carcassi (Ch. I.1), Saja Gherri (Ch. II.1,II.2)
contributed one or more

Test readers: Chami Amarasinghe, Andre Antoine, Saja Gherri, Uriah Israel, Micah Johnson, Sean Kelly, Dan McCusker, Pietro Monticone, Everardo Olide
reviewed a full chapter or more

Additional test readers: Josce Kooistra, Armin Nikkhah Shirazi, Ayla Rodriguez, Alex Takla, Tobias Thrien, Allan Vanzandt
review prompted corrections and clarifications