

Overview and Status of Assumptions of Physics*

(Extended version)

Gabriele Carcassi, Christine A. Aidala

November 2019

Abstract

This article gives an overview of the project “Assumptions of Physics” which aims to rederive the known laws from a few physically meaningful starting points. It presents the motivations behind the project, a summary of the main findings and current status of the research.

1 Introduction

The aim of this article is to give a broad overview of the project “Assumptions of Physics”: its motivations, its status and a list of the main findings. It is intended to be a good start for anyone interested in simply learning more about the project or finding an area for collaboration.

As we plan to update this document, please make sure you have a recent version. Each subsection will be fairly self-contained to allow the reader to skim through the document and focus on the parts of interest. It will contain a summary section at the beginning and a titled paragraph for each of the main idea or finding.

2 Goals and method

A better understanding of physics. The overall goal of the project is to better understand physics: to understand why the laws of quantum mechanics, or classical mechanics, are what they are; to understand

the limit of the validity of the different theories; to understand what the Poisson brackets, or any other mathematical operation, represent physically; to understand which mathematical objects correspond to actual physical entities and which are just artifacts; to understand what new theories are possible.

Ideally, we want to achieve a single unified framework for all science that is physically meaningful (i.e. it is clear what each objects represents experimentally and why must it be represented that way), mathematically precise (i.e. the physical assumptions are captured through axioms and definitions and then the arguments are carried out with the rigor of modern mathematics) and philosophically consistent (i.e. the concepts and viewpoints taken fit with the mathematical results and the practice of science).

2.1 Objectives

The overall goal of a better understanding of physics can be broken down into the following five objectives.

Clarify the assumptions. Physics is currently a patchwork of different theories (e.g. classical mechanics, quantum mechanics, thermodynamics, ...) that are used for different systems or for the same system in different contexts. When should we use one instead of another? What are the limits of validity of each theory? We want to understand what are the assumptions one has to make on a system such that it can be studied with a particular theory. The only way to make sure our assumptions are necessary and sufficient is to show the theory can be rederived from them. This, in turn, gives us a better idea of what

*<http://assumptionsofphysics.org>

each theory describes, how it fails, and what new directions one could take.

Put physics back at the center of the discussion. Physics used to try and identify basic “laws” or “principles” for itself. Current physical theories, instead, just postulate a mathematical structure relegating physics to a mere after-the-fact interpretation. We want to go back to the old approach: we want to start with the physics and derive the math suitable to capture those physical concepts. Ideally, no mathematical construct should be introduced if it is not physically motivated. Why is phase-space formed with conjugate pairs? What does the commutator represent? If these are really representing physical objects, we should have a better understanding than a vague “you can think of it as...” or a cryptic mathematical “it is the left action of a fiber bundle on a...”. Most of our understanding should come from physical intuition and not from the mathematics it is used to represent it.

Give science sturdier mathematical grounds. There is a general sense within physics that, at some point, mathematical details are physically not interesting and should be handled “appropriately” by mathematicians. Therefore even mathematical structures that originally came from physics are now formalized by mathematicians to solve their own needs, and not the physicist’s. This, in turn, reinforces the idea that those details are not interesting to the physicists. The reality is that a well-posed physical problem must also be a well-posed mathematical problem, but not the other way around, so leaving the details to the mathematicians means having fundamental structures that are less physical.¹ Mathematicians will come up with definitions so that theorems are easier to prove or calculations are easier to perform. They will not come up with definitions based on whether they are physically meaningful. If we give a theory, as we said before, by simply stating a mathematical structure, then there is no guarantee that what we give is a fully physical theory. Unphysical elements simply means the physical problem was not specified correctly. If we truly are able to give a

¹For example, quantum states are modeled with Hilbert spaces even though we know they contain mathematical objects that are unphysical.

precise physical meaning to every mathematical object, then our formal structures will map one-to-one to our physical understanding. This means no unphysical mathematical artifacts and a more precise mathematical treatment of physics.

Foster connections between different fields of knowledge. Knowledge is increasingly specialized and fewer and fewer people are well versed in more than a couple of subjects. Yet, connections between different fields of science, mathematics and engineering are routinely found to be useful for one simple reason: nature is one and does not care about such divisions. A more holistic vision of scientific knowledge, then, is a natural byproduct of our effort. The notions of state, environment, process and equilibrium, for example, are intertwined and are fundamental to most fields of science and engineering. Proper characterization of such basic concepts will give a common language and mathematical tools that span across different disciplines.

Provide a solid basis for new theories. Most of the attention in fundamental physics is focused on the development of new theory or the search for new effects. A better understanding of the current theories, recast in a single broader and more precise framework, will probably facilitate that search. It makes explicit what ingredients went into the theories, so we know what assumption can fail and what principles cannot be changed.² If all physical theories can be seen as instances of a more general structure, they prove that the general structure is sound and provide templates for new theories.

While the objectives cover many areas and are very broad, our experience tells us that they are ultimately connected and cannot be pursued to their fullest independently. Therefore we are always looking for people with diverse background and interest to help us cover the different areas.

²For example, we may abandon the assumption that space is measurable at an indefinite precision, leaving the realm of manifold and real numbers; yet we cannot abandon the notion that space is experimentally accessible, staying in the realm of second countable T_0 topological spaces.

2.2 Methodology

As each topic is investigated by the project, it generally goes through three phases. Reverse engineering, where we try to find what assumptions are beneath a theory. Forward engineering, where we try to check whether the theory can be recovered from the assumptions we found. Formalization, where we try to codify in mathematical language the assumptions and the derivation.

Reverse engineering. The goal of this phase is to deconstruct an already existing theory or mathematical structure to understand what each piece is supposed to represent physically. One way to conduct such investigation is to ask what would happen if a particular characteristic of the framework would fail.³ This usually leads to a set of necessary assumptions about the physical system one is describing, the ones one must take for the particular description to apply.

Forward engineering. The goal of this phase is to construct an argument from the necessary assumptions outlined in the previous phase to rederive the framework at hand. This checks whether the necessary assumptions are also sufficient⁴. It ends when a set of necessary and sufficient conditions to rederive the theory are found, meaning that the full theory has been physically characterized.

Formalization. The goal of this phase is to take the previous set of arguments and formalize them in pure mathematical language. This forces us to clarify all the starting points and what parts of the arguments are truly formal and what are physical. Unfortunately, this can ultimately be done only by starting from the very foundations⁵ though we are at a point where an overall tentative structure is emerging and tentative formalization can be attempted on

³For example, a line of reasoning could be as follows: Hamilton's equations are differential; if they weren't the mapping of densities and areas would not be well defined; maybe Hamiltonian mechanics is not about mapping points but about mapping densities or areas.

⁴For example, if we started simply by assuming we have density distributions mapped in time, this would give us all possible differentiable evolutions, not necessarily Hamiltonian ones.

⁵For example, before formalizing Hamiltonian mechanics we have to formalize differentiable manifolds, real numbers and topological spaces.

higher structures.

2.3 Organization

The overall project is organized into the following macro-categories.

The general theory. This formalizes the basic mathematical framework that is the basis for all science. As every algebraic structure in mathematics is a specialization of the generic structures axiomatized by set theory, all scientific theory are ultimately specialization of the generic structure provided by the general theory. This defines the basic requirements imposed on a theory by logical consistency and experimental verifiability. It defines properties and quantities, accuracy and other concepts of general applicability.

Physics core. This defines the notions that are fundamental to physics specifically, such as states, processes, environment, equilibrium, determinism and so on. The idea is to define these concepts in the most general way possible such that common requirements can be identified. Every physical theory is a further refinement of this structure in that it will choose to study a particular set of states over a particular process.

Physical theories. This studies the different assumptions that are needed to recover the known physical theories. The assumptions typically describe what level of description is available for the system and its parts, how does it change in time and whether trajectories are enough to recover it.

3 The general theory

This part of the project focuses on the development of a rigorous mathematical framework that can provide the building blocks that are common to all scientific theories. As logic and set theory provides the foundation for all other mathematical structures, which are essentially sets with associated operations, the general theory will define common tools (e.g. logic of experimental verification, causal relationships, measurable quantities) that other theories will specialize in different ways.

3.1 The Principle of Scientific Objectivity

Our guiding principle is that “Science is universal, non-contradictory and evidence based.” Any scientific theory must be logically consistent, its content be equally true or false for everybody and it must deal with what can be established experimentally. Different branches of science will specialize on different systems and topics, but all the theories and models they develop must satisfy those basic constraints. The general theory deals with the fundamental mathematical structure needed to realize that principle.

Not every subject can be studied by science. The principle indirectly states that anything that is subjective, contradictory or not evidence based cannot be the subject of scientific investigation. The properties of prime numbers, moral and existential questions or the aesthetics of music fail in at least one of those respects and are excluded. Only the parts of nature that are accessible through consistent experimental verification can be studied scientifically. New experimental techniques have been extending that reach over time.

3.2 The Logic of Verifiable Statements

The basic element in our framework is the idea of a verifiable statement: an assertion that is either true or false for everybody and for which we have an experimental test that will terminate successfully in finite time if and only if the statement is true. The most basic structures therefore need to provide a logic framework to keep track of what statements are verifiable and their relationships.

3.2.1 Logical contexts

A logical context consists of a set of statements with well defined logical relationships. It is the most fundamental structure and it is the only one that is axiomatically introduced. A scientific theory will consist of a set of statements taken within a logical con-

text with particular logical relationships.⁶ The context defines basic logical relationships (e.g. equivalence, narrowness, compatibility and independence) which form the basis for higher level constructs (e.g. ordering, linear and statistical independence). Every other structure imposed on the statements, to be logically consistent, will need to “play” nice with these fundamental relationships.

Note on terminology. The terminology may depart slightly from what logicians may be accustomed to, simply because our aim is different. Namely, our goal is not to study the rule of inference or study what can be formally proven, but simply to keep track of the logical relationships between statements that are assumed to be already given.

Statements as primary object. Every formal theory needs prime objects, elements that will not further be described by the theory. We will take statements, assertion that are either true or false for everybody, as these objects. The formal system will specifically:

not be propositional; the language of the statements, its syntax, its grammar, will be left unspecified; statements will represent the content regardless of how it is expressed;

be algebraic; operations on statements will not form new statements, but describe relationships within a given set of statements; note that most mathematical structures used in science are algebraic (computer science is a notable exception)

purely formal; the semantic of the statements is not captured, though the semantic will impose logical relationships that will be formally captured

Statements cannot exist in isolation. Statements should, in general, not be regarded as independent of the context they are part of. Take the statement “The mass of the electron is 510 ± 0.5 KeV”. If we are measuring the mass of the electron, that is a statement that may or may not be true depending on experiment. If we are performing particle identification

⁶For example, Newtonian mechanics will use statements like “the mass of the object is 1 Kg” and “the acceleration of the object is 1 m/s²”. If both of these are true, then the statement “the force on the object is 1 N” is also true.

in a detector, that statement is assumed to be true. The truth values allowed by the same statement depends on context.

Logical context as the fundamental structure. A logical context is a set of statements with well defined logical relationships. It defines which truth assignments are possible (i.e. are consistent with the semantic of the statements). It identifies one of the possible assignments as the true one. It allows to find statements whose truth value depends on others through a particular Boolean function. The existence of this structure is introduced by axioms as they are crossing the lines between physical objects and mathematical ones. We currently do not assume there is a single universal context.

Tautologies, contradictions and contingent statements. We define tautologies as those statements that are true in every assignment, contradictions as those statements that are false in every assignment and contingent all the remaining.

Statement equivalence and Boolean algebra. Within our formal system, we can distinguish between:

statement equality: two statements that re-express the same fact in different languages or in different words

statement equivalence: two statements have the same truth value in all possible assignments

statement material equivalence: two statements have the same truth value

A logical context is a complete Boolean algebra under statement equivalence.

Narrowness and ordering. A statement is narrower than another if whenever it is true then the other is also true. Narrowness can often be thought as implication, except in a few corner case (e.g. a statement that is always false is narrower than any other statement). Narrowness imposes a partial order to the structure. This is the same order associated to the fact that every Boolean algebra is a complemented distributed lattice.

Compatibility and independence. Two statements are said to be compatible if they can both be true in

the same assignment. Two statements are independent if fixing the truth value of one does not restrict the truth value of the other. Independence, like linear independence or statistical independence of which it is the foundation, is a property of a set of statements and is not transitive.

3.2.2 Verifiable statements

Logical contexts need to keep track which statements are experimentally verifiable. A statement is experimentally verifiable if we are provided with a test that, if the statement is true, will always terminate successfully in a finite time. The fact that the test has to terminate in a finite time in the positive case and that it may not terminate in the negative case has profound implications. The negation of a verifiable statement is not in, general, a verifiable statement. The finite conjunction (i.e. logical AND) of verifiable statements is a verifiable statement, but not an infinite one. The countable disjunction (i.e. logical OR) of verifiable statements is a verifiable statement, but not a more than countable infinite one. Therefore verifiable statements are not closed under the standard logical connectors

Different algebra. Verifiable statements do not follow the standard Boolean algebra. The negation of a verifiable statement is not necessarily verifiable since the test is not guaranteed to terminate. The conjunction of a finite set of verifiable statements is necessarily verifiable since we can run the tests one by one and, if all are successful, the conjunction is also verified. This cannot be extended to an infinite set since the test would not terminate. The disjunction of a countable set of verifiable statements is necessarily verifiable since, once one test terminates, the disjunction is verified and we can stop. This cannot be extended to an infinite set that is more than countable since we would not be able to find the test that terminates in a finite time.

Decidable statements. These are statements that can be verified and falsified. That is, their test always terminates. These can be defined from verifiable statements and they follow a finite Boolean algebra: the negation, finite conjunction and finite disjunction of decidable statements is also decidable.

3.2.3 Experimental and theoretical domains

We define an experimental domain as a set of verifiable statements that can be expressed as the combination of a countable subset (i.e. a countable base). Because of finite time verifiability, this is the biggest space we can experimentally probe given an indefinite amount of time. An experimental domain, then, represents all the information that can be gathered experimentally about a particular subject.

From each experimental domain we construct a theoretical domain by allowing negation as well. This will include all statements for which an experimental test is in principle possible, though there is no guarantee of termination. A theoretical domain, then, represents all statements to which we can attach a prediction.

Within a theoretical domain we define the set of possibilities as those statements that, once known to be true, will set the truth value of all other statements. Each possibility represents a possible case that is distinguishable experimentally and corresponds to a unique possible assignment for the experimental and theoretical domain.⁷ Because of the countable base, the set of possibilities can never be greater than the continuum.

Maximum cardinality for the possibilities. Because the domain is generated by a countable set of verifiable statements, the possibilities must be distinguished with a countable sequence of boolean values. This means that the cardinality of the possibilities cannot exceed that of the continuum. This means that higher order infinities and all associated problems do not play a role in science.

3.2.4 Topologies and σ -algebras

An experimental domain provide a natural topology over the possibilities. Each verifiable statements can

⁷For example, an experimental domain consists of a set of statements that can be tested experimentally (e.g. “the animal has whiskers”, “the mass of the photon is less than $10^{-13}eV$ ”). The theoretical domains extends to statements that may not be tested experimentally (e.g. “the mass of the photon is exactly $0eV$ ”). The possibilities consists of all the possible cases (e.g. “the animal is a cat”, “the mass of the photon is exactly $0.10210^{-34}eV$ ”).

be identified with the set of possibilities compatible with it.⁸ Since verifiable statements are closed under finite conjunction and countable disjunction, the sets corresponding to verifiable statements form a T_0 second countable topology.

On the other hand, theoretical domains provide a natural σ -algebra over the possibilities. Each theoretical statement can also be identified with the set of possibilities compatible with it. Since theoretical statements are closed under negation and countable disjunction, the sets corresponding to theoretical statements form a σ -algebra, which is the Borel algebra of the natural topology.

Topologies and σ -algebras provide the foundations for differential geometry, Lie algebras, measure theory, probability theory and many other mathematical tools used in physics and the sciences. As we now have a precise understanding of what they represent, all concepts and proofs in those subject can be understood in terms of experimental verifiability.

Interior, exterior and boundary. Given a test associated to a theoretical statement, it will either terminate successfully, terminate unsuccessfully or never terminate. Each Borel set is associated with a theoretical statement, which means it is associated with a test. The interior of the Borel set represents the cases in which the test will terminate successfully; the exterior represents the cases in which the tests will terminate unsuccessfully; the boundary represents the cases in which the test will not terminate.

Decidable and undecidable statements. A theoretical statement is decidable if the associated test always terminates, which means the boundary is empty. This corresponds to a clopen set. It is undecidable if the associated test never terminates, which means the boundary is the full set.⁹

3.2.5 Status and open issues

This part of the work is very well developed. More work could be done in finding meaning for all math-

⁸For example, the statement “the mass of the photon is less than $10^{-13}eV$ ” is equal to the disjunction of all statement of the form “the mass of the photon is exactly $x eV$ ” with $x < 10^{-13}$

⁹For example, “there are three apples in the bowl” is a decidable statement, while “the mass of the electron is a rational number when expressed in eV ” is undecidable.

ematical concepts (e.g. compact sets, all separability axioms, ...). It would be interesting to reach out to individuals from related fields (i.e. logic, foundations of mathematics, philosophy of science, ...) to see whether any aspect of this section would be novel and interesting to their respective communities.

3.3 Domain relationships

The study of experimental domains gives us basic constructions that investigate the way different domains can be related or combined.

3.3.1 Relationships and equivalence between domains

We define two relationships between domains: inference relationships and causal relationships. An inference relationship establishes that testing a verifiable statement in one domain is the same as testing a verifiable statement in the other. Mathematically it is a map that takes a verifiable statement from one and returns a verifiable statement from the other that is equivalent to the first. A causal relationship establishes that determining which possibility is true in one domain also determines which possibility is true in the second. Mathematically it is a map that given a possibility of the first returns a possibility of the second that is broader than (i.e. less specific, true in more cases) the first.

One result is that causal relationships must be continuous function in terms of the natural topologies, which justifies why functions in science are always assumed to be “well-behaved” (i.e. analytically continuous with at most countable discontinuities). Another result is that there exists an inference implication between two domains if and only if there exists a causal relationship between them. The direction of the inference is the opposite of the causal direction. Therefore we say that a domain depends on another if there exists an inference relationship between the first and the second or if there exists a causal relationship between the second and the first. If the relationships are invertible (i.e. each domain depends on the other) then the domains are equivalent.

Continuity and inference. An inference relationship establishes that testing a verifiable statement in one domain tells whether a statement in the other is true or false. As these statements are equivalent, the experimental test on one constitutes an experimental test on the other, which means the second statement is verifiable. One cannot impose an inference relationship between a verifiable and non-verifiable statement. In terms of the natural topology, these correspond to open sets, therefore, by definition, the map is continuous as it maps open sets to open sets. The typical requirement that functions are “well behaved”, then, is actually physically significant.

3.3.2 Composite domains

Given two experimental domains, we can create the composite by including all the verifiable statements that can be constructed from them. A possibility of the composite domain will determine the truth assignment for all statements of both original domains. Depending on the logical relationships between the two domain, we have different cases. If the two domains are independent (e.g. position along two different directions), then the possibilities are the scalar product of the possibilities. If one of the domains depends on the other (e.g. the temperature of a mercury column and its height), then the possibilities are the ones of the independent domain. If the domains are incompatible (e.g. plant species and animal species) then the possibilities are the disjoint union.

3.3.3 Relationship domains

Given two experimental domains where one depends on the other, we would like to characterize how the relationship is identified experimentally.¹⁰ The main difficulty is that, within a single context, there can only be one such relationship or we would introduce logical inconsistencies. Choosing an inference relationship, in fact, means fixing the logical relationship between statements¹¹ which ultimately means fixing

¹⁰For example, knowing that there is a causal relationship between the temperature and the height of a mercury column, we would like to measure how the two are related.

¹¹For example, “the temperature is 12 C” is incompatible with “the height is 23 mm”

a logical context. Identifying the relationship, then, means identifying the right context.

Given a set of contexts, each describing a possible relationship, we can construct an experimental domain for the relationship in which each possibility corresponds to a possible relationship and its context. Whether or not the experimental test will exist in practice, the mathematical construction can always be performed resulting in just another domain. This means that we can construct relationship domains over relationship domains for any arbitrary higher order relationship, meaning that the mathematical framework is closed.

Tests for the relationship domain. The experimental verifiability of relationship domain statements cannot be proven from the verifiability of each domain. This formalizes the problem that being able to measure a single point (i.e. the temperature and height in a particular circumstance) is not enough to verify the relationship (i.e. the temperature and height are always linked in a particular way). The ability to experimentally explore the space is what builds confidence, which may not be always possible. What constitutes enough evidence is left, as before, outside of the formalism.

3.3.4 Status and open issues

The overall mechanics of composite domains and relationships domains are fairly well developed and understood. There are details, though, that can be understood better. The connection to category theory can be also better developed. Other common concepts and constructions, like similarities or sequences of domains, could be formulated.

3.4 Properties and Quantities

Another basic tool is the ability identify the possibilities of a domain through the values of its properties, since these are the ones that, in practice, we measure. Part of the general theory is the formalization of these concepts.

3.4.1 Properties and values

A property is defined as a map from the possibilities to the set of values that the property can have. The map has to be topologically continuous as a measurement on the property should correspond to a measurement over the domain. We distinguish between various cases depending on whether the property is defined over all the possibilities and how much is able to distinguish between all the elements

3.4.2 Quantities and ordering

Quantities are defined to be ordered properties, values have magnitudes that can be compared. The objective is to understand how this ordering is fully defined by the logical relationships between verifiable statement. The key insight is that $3 \leq 4$ precisely because “the quantity is less than 3” is narrower (i.e. more specific) than “the quantity is less than 4”.

Ordered domains are constructed through references (e.g. ticks of a clock, marks of a ruler or levels of a graded recipient) that allow us to tell whether the quantity is before or after itself. To have an ordered quantity, the set of references have to satisfy a set of necessary and sufficient conditions that are laid out by three ordering theorems.

Once those conditions are satisfied, the real and integer numbers (i.e. continuous and discrete ordering) respectively emerge by requiring either that between two references there is always at least one more or that there are only finitely many.

Of note is that the requirements for ordering itself are quite demanding and are ultimately unlikely to be satisfied for quantities like space and time. Which would mean that not only time (or space) would not be chartable by the real line, but that ordering would not be well defined at the finest scale.

3.4.3 Status and open issues

Properties and quantities are fairly well developed. The extension to manifolds should be straight forward (require independent quantities to identify each possibility within a “small” open set).

3.5 Further work

The general theory at this point covers only the topological aspects and needs to be extended to the geometrical/measure theoretical ones. Our current thinking is that diverse aspects, such as geometry, measure theory, probability theory, information theory, can be recovered by characterizing the granularity (i.e. accuracy) of the statements. Group theoretical aspects may be recovered

Granularity as foundational concept. As each statement corresponds to a Borel set of possibilities, assigning a size to these sets effectively means giving a size to the statements themselves. For geometry, two sets have the same volume because they contain the same number of possible spatial points. For probability, borrowing idea from Cox probable inference, we would assign probability only on conditional statements, therefore the $P(s_1|s_2)$ would measure what fraction of the possibilities compatible with s_2 are also compatible with s_1 . For information theory, given two statements the logarithm of the ratio of the size of the possibilities corresponds to the number of yes/no questions that are needed to go from one statement to the other, the relative information.

Recover differential geometry or geometry measure theory. The idea is to start with finite functions of finite k-dimensional surfaces that have the following properties: they are linear if the surface is broken into disjoint parts (the value for the whole is the sum of the values for the parts) and they respect the limit (the value for the limit surface is the limit of the sequence of values for the surfaces). Under these, and possibly others, conditions we expect these functions to be expressed through an integral of local functions (i.e. k-forms) over infinitesimal surfaces (i.e. k-vectors).

4 Physics core

This part has not yet been formally developed, though some common ideas are starting to emerge. The main idea comes from the realization that, when defining a system, the notions of states, processes, time, evolution laws, interaction with the environ-

ment and equilibria are all interrelated and can't be given independently. Roughly speaking, to define a system we have to choose a boundary that separates it from the environment. Only the quantities that are unaffected by the interaction with the environment can, in those circumstances, be thought as properties of the system. This in turns is linked to what we can define as state for the system, which is linked to the laws we can write. The processes and the laws are naturally linked to what is assumed about time.

The idea, then, is that giving a set of states is actually giving a much richer structure. It means defining a set of interaction at the boundary, it means defining equilibria with such interactions, it means defining time-scales below which and above which our description fails or is not interesting, it means defining how probing the system may affect those assumptions.¹² The plan is, at some point, to formalize these objects and relationships to provide a basic mathematical structure that all physical theories must adhere to.

All states are equilibrium states. In thermodynamics equilibria are simply assumed to exist, without specifying exactly how they are achieved. However, in all physical theories, states are assumed to exist, without specifying what are the conditions upon which the system can be characterized by such states. The idea that, at the very least, all states must be equilibrium states of processes that happen at a faster scale than the processes than we are considering.¹³ This sets the scale of infinitesimal time intervals.

States are objective but not absolute. The same system can be found and studied in different conditions. Depending on the circumstances we may be interested in studying different properties of the system.¹⁴ Therefore the state will change depending on

¹²If we are studying a molecule, we may impose boundaries and processes such that that molecule and its constituent remain stable, which means, for example, interaction energies lower than the chemical bonds.

¹³For example, the position and velocity of a book makes sense on the surface of the earth, at a temperature of 293 Kelvin. On the surface of the sun, we would not be able to talk about the position and velocity of a book.

¹⁴Consider a ball. If we are studying its motion, the state can be typically be captured by position and momentum of the center of mass. If the ball is rotating at high speeds in air,

the situation and our focus, and in this sense it is not absolute. However, once the situation and focus are well defined, the state is objective in the sense that everybody will come to the same characterization.

The tension between isolation and interaction. Properties of a system can be fully characterized if they are not perturbed by the interaction with the environment, if they are in stable equilibrium.¹⁵ To fully characterize a system, then, we would want it to be in perfect isolation.¹⁶ However, to be able to measure those quantities we would like to interact with the system. To fully access the system, then, we would want it to be in constant interaction with the environment. Conceptually, there is a trade-off. Specifying the state space means characterizing this trade-off for the specific system in the specific circumstances.

Determinism and reversibility is isomorphism.
 TODO

[To be completed]

5 Physical theories

5.1 Classical mechanics

Classical mechanics can be derived from three main assumptions: infinitesimal reducibility, determinism and reversible evolution and kinematic equivalence. From the first assumption we can show that the state space of the infinitesimal parts of a system (i.e. particles) has the structure of phase space. From the second we can show that the time evolution follows

the rotation speed and axis will also need to be part of the state because of the Magnus effect. If we are studying the gas within the ball, the state may be given by volume, pressure and temperature. If the ball is so tiny and light that random impact of the fluids on the surface will affect the motion, the state will be a statistical ensemble. And so on.

¹⁵For example, the position of the center of mass of a cannonball is not affected by the air and can therefore be characterized by a set of numbers; however, the position of the center of mass of a speck of dust is affected by the air and must be characterized by a probability distribution over the possible values.

¹⁶Note that full isolation is not possible as this would mean no interaction, not even gravitation, between the system and the environment, to the point that the system has essentially vanished.

Hamiltonian mechanics. From the third we can show that the system is Lagrangian and is restricted to the case of massive particles under potential forces.

5.1.1 Classical state spaces

Infinitesimal reducibility. A system is said reducible if giving the state of the whole system is equivalent to giving the state of its parts and vice-versa. For example, given a ball, we can throw it and study the motion of the ball. Alternatively, one can take a red marker, make a red dot on the ball and study the motion of the dot. The system is reducible if studying the system is equivalent to studying the motion of all possible red dots. The system is infinitesimally reducible if we can keep reducing the system to smaller and smaller parts. We call particle an infinitesimal part of the system, the limit of recursive subdivision. The system is then described by a distribution of the state of its parts over phase space, which is the only structure that allows coordinate invariant distributions over continuous variables.

Composite states are distributions over particle states. Under infinitesimal reducibility, if \mathcal{C} is the state space of the whole system and \mathcal{S} is the state space of the particles, then each state $c \in \mathcal{C}$ for the full system is identified by a distribution $\rho : \mathcal{S} \rightarrow \mathbb{R}$ over the states of infinitesimal parts. The function is real to signify that we can associate arbitrarily small amounts to each particle states. The distribution tells us how many particles can be found in each particle state.

Continuous time implies particle state space is a manifold. If time is assumed to be continuous, then the state space \mathcal{S} is a manifold. That is, each particle state $s \in U \subseteq \mathcal{S}$ can be identified by a set of continuous quantities $\xi^a : U \rightarrow \mathbb{R}$ which we call state variables. It is a consequence of the general theory that, if we wanted to write a trajectory $\xi^a(t)$, then this must be a topologically continuous function. Therefore, conceptually, once we assume time to be a real number, the state space of the particles must be charted by a set of real quantities: the topology that is given to time reasserts itself on all the quantities that will depend on time.

Densities imply differentiability. As we can write

the state $s(\xi^a)$ as a function of the state variables ξ^a , we will also want to write the density $\rho(s(\xi^a)) = \rho(\xi^a)$ as a function of the state variables. This requires the state variables to be differentiable with respect to the distribution and with each other. Recall that if the Jacobian is not well defined during a change of variable $\rho(\hat{\xi}^b) = \left| \frac{\partial \hat{\xi}^b}{\partial \xi^a} \right| \rho(\xi^a)$ then the density wouldn't be expressible in the new coordinates. Mathematically, this means that the manifold is equipped with a differentiable structure. Note that any time evolution that would map a density to another density will not only have to be continuous, but also differentiable. In short: assuming density distributions as our primary objects, instead of points, justifies why differentiable structures and differentiable time evolution is present in physics.

Coordinate invariant densities imply phase-space. As particles states must be defined independently of coordinate systems, they must be invariant under coordinate transformations. That is, $s(\xi^a) = s(\hat{\xi}^b(\xi^a))$. As the density distribution depends only on the state, it should be invariant as well, that is $\rho(s(\xi^a)) = \rho(s(\hat{\xi}^b(\xi^a)))$. We have two seemingly contradicting requirement: ρ should both change like a density and be invariant. The only way to satisfy both requirements is the following: particle states are fully identified by a set of coordinates q^i and by conjugate set of variables k_i whose units are the inverse of the ones defined by q^i . This way the product $dq^i dk_i$ is a pure number that is invariant under coordinate transformations. A coordinate transformation $\hat{q}^j = \hat{q}^j(q^i)$ induces a transformation $\hat{k}_j = \frac{\partial q^i}{\partial \hat{q}^j} k_i$ such that the area element $d\hat{q}^j d\hat{k}_j = dq^i \frac{\partial \hat{q}^j}{\partial q^i} \frac{\partial q^i}{\partial \hat{q}^j} dk_i = dq^i dk_i$ is invariant. An independent degree of freedom, then, is a pair (q, k) for which the units can be changed independently from all other state variables. Mathematically, symplectic manifolds are the only spaces that allow us to define coordinate invariant densities. We can introduce a constant \hbar such that $p_i = \hbar k_i$ which at this point is simply an arbitrary choice of unit.

Precise physics-math correspondence. These results gives us a precise understanding of what each layer of the mathematical structure corresponds to

physically:

Set of points: possible states for the particles.

Topological space: ability to experimentally distinguish the states.

Manifold: states can be distinguished with a set of real valued quantities or state variables.

Differentiable manifold: we can define a distribution over the states and the state variables.

Symplectic manifold: we can define a distribution over the states that are coordinate independent.

5.1.2 Hamiltonian mechanics

Determinism and reversibility. A system undergoes deterministic and reversible evolution if given the initial state one can predict the final state and given the final state one can reconstruct the initial state. In other words, the system is predictable and retrodictable. Densities at each past state must be equal to the density of the corresponding future state, which leads to Hamiltonian dynamics.

Reversibility as reconstructing the past. This definition of reversibility should not be confused with time symmetry or the ability to undo the change. The first one means that if one substitute $t \rightarrow -t$ the dynamics is invariant (i.e. the same) or form invariant (i.e. different but described by the same equations) depending on the author. The second one means that we can find another process that brings the final state back to the original state. These concepts do not seem to work well in the general case. For example, when discussing reversibility in the sense of undoing the changes, one typically discusses changing the direction of the velocities to bring about the initial state with inverted velocities. In the presence of forces or an arbitrary dynamics, this operation because more complicated or ill-defined. The core problem with this definition is that it is not a property of the process itself, but on the existence of other processes (i.e. the process that undoes the changes or the process that flips the velocities, therefore changing the state). The requirement of time-invariance is too restrictive as it excludes all time dependent processes. Form invariance on an arbitrary set of equations, instead, can be broken by a suitably complicated change of

variables, so it is only well-defined for a limited number of systems (e.g. particles under potential forces). The definition of reversibility we use is preferable because it works more generally and, when appropriate conditions apply, reduces to the other two which are better regarded as special cases. It is mathematically more elegant as it is the dual of determinism. It connects directly with information theory, as it states that the present has enough information to reconstruct the past, which gives a logical bridge to connect reversibility with information entropy and thermodynamics/statistical mechanics entropy.

Determinism and reversibility imply Hamiltonian mechanics. In the context of infinitesimal reducibility, determinism and reversibility means that all the particles and only the particle within a given initial state $\rho(s(t)) = \rho(s(t + \Delta t))$ will be mapped to a particular final state. That is, the densities will be preserved over time. Time evolution, then, is a canonical transformation in phase space. As time is continuous, we can find a function $H : \mathcal{S} \rightarrow \mathbb{R}$ which acts as a time generator, which we recognize as being the Hamiltonian. This is essentially Louisville's theorem applied in reverse.

Invariant densities under time transformations imply extended phase-space. So far, time has been treated as a parameter. If we introduce coordinate transformation that mix time, however, this is no longer possible. If $d\hat{q} = dq + vdt$ then the densities need to be defined on dt as well. To do that, we have to introduce a conjugate variable ω of unit inverse time so that $dt d\omega$ is dimensionless and invariant. The invariant area element becomes $\sum_i dq^i dk_i - dt d\omega$ or $\sum_i dq^i dp_i - dt dE$ in units of \hbar . For surfaces at equal time $dt = 0$ it reduces to the previous case. For surfaces that are not at equal time, since the temporal degree of freedom does not introduce new states, we need to correct the area element to account for the larger areas measured by the state variables that does not correspond to new states. If we set $q^\alpha = [t, q^i]$ and $p_\alpha = [-E, p_i]$ the area element becomes $dq^\alpha dp_\alpha$. These are elements of special relativity that come just by requiring invariance under time transformation. They do not come from invariance of the speed of light, they do not require a metric tensor. In other words, there is no way to have time transformations

in Hamiltonian mechanics without requiring a four-momentum co-vector. Special relativity is not an independent choice: it is already partially baked into the two assumptions we took.

Extended phase-space imply classical anti-particles. The trajectories in the extended phase space are in terms of an affine parameter s and not time: $q^\alpha(s)$. As the motion is still deterministic and reversible, though, time $t(s)$ must be monotonic with respect to the affine parameter. For each state, the direction of time can be either aligned or anti-aligned with the affine parameter. Because the relationship is monotonic, aligned states can only connect to other aligned states and anti-aligned states can only connect to other anti-aligned states. We call aligned states the particle states and the anti-aligned states the anti-particle states.

Time evolution is better expressed by the Hamiltonian constraint. The temporal degree of freedom introduces two variables but, since one is energy, there is an additional constraint since the value of the energy is linked to the other variables through the Hamiltonian. This link is better expressed by the Hamiltonian constraint $\mathcal{H} : \mathcal{S} \rightarrow \mathbb{R}$ which is a function over the extended phase space, just like the Hamiltonian was a function over the standard phase space. The Hamiltonian constraint has two roles: it gives us the equation of motions in the same way of the standard Hamiltonian, and it constraints the motion over the surface $\mathcal{H} = 0$. To give context, the Hamiltonian constraint for a free particle is $\mathcal{H} = \frac{|p^\alpha|^2}{2m} - \frac{1}{2}mc^2$, which constraints the norm of the four-momentum to mc . The Hamiltonian constraint, unlike the standard Hamiltonian, is invariant under time transformation. It is the generator of the evolution in terms of an affine parameter.

Hamiltonian constraint as operator on distributions. If ρ is the distribution for the whole system, then $\mathcal{H}\rho = 0$. In fact, the distribution can only be non-zero on the constrain, for which the Hamiltonian constraint is zero. So, for each point in the extended phase space, either the distribution is zero or the Hamiltonian constrain is zero. Note that this is the same form one has in quantum mechanics for the Klein-Gordon and Dirac equations $\mathcal{H}|\psi\rangle = 0$, which

in fact can be understood as Hamiltonian constraints that fix the norm of the momentum four-vector and four-velocity.

Determinism and reversibility is equivalent to conservation of information entropy. Given a distribution, its information entropy is given by $I = -\int \rho \log \rho d\xi^n$. In general, information entropy is not invariant under change of variable: $-\int \rho(\xi^b) \log \rho(\xi^b) d\xi^n = -\int \rho(\xi^a) \log \rho(\xi^a) d\xi^n - \int \rho(\xi^i) \log \left| \frac{\partial \xi^b}{\partial \xi^a} \right| d\xi^n$. Under a canonical transformation over phase space, though, the Jacobian determinant is unitary everyone therefore information entropy is invariant. That is, invariance of the density and invariance of its information entropy are one and the same. Requesting that densities are conserved in time is therefore equivalent to requesting that information entropy is conserved in time. This should be intuitive: the amount of information needed to identify an element within the initial or the final distribution is the same if the evolution is deterministic and reversible. Deterministic and reversible evolution, Hamiltonian evolution and conservation of information entropy, then, are one and the same.

Conservation of information entropy leads to classical uncertainty principle. As information entropy is conserved, a Hamiltonian process can only map a distribution to one that has the same information entropy. If we consider a single degree of freedom (q, p) and fix the information entropy I_0 , then the distribution that minimizes the product of the standard deviation is the Gaussian. The information entropy for the Gaussian is given by $I_G = \ln(2\pi e \sigma_q \sigma_p)$. During the evolution, we have:

$$\sigma_q \sigma_p \geq \frac{e^{I_0}}{2\pi e}$$

5.1.3 Lagrangian mechanics and massive particles

Kinematic equivalence. A system is said to satisfy kinematic equivalence if giving the state of the system is equivalent to giving the trajectory in space and vice-versa. That is, for each possible state there is one and only one trajectory associated with it.

Note that this is not true in general: for example, given the trajectory of a photon we cannot reconstruct its momentum since all photons travel at the same speed. Under this assumption the relationship between position/momentum and position/velocity is invertible and the system admits a Lagrangian. Moreover, given the transformation rules of momentum and velocity, we find that there must be a linear relationship between the two, which constrains the dynamics to the one of massive particle under potential forces.

Note on notation. We use (q^α, p_α) for position and momentum and (x^α, u^α) for position and velocity even though $q^\alpha = x^\alpha$. We do this because $\frac{\partial}{\partial q^\alpha} \neq \frac{\partial}{\partial x^\alpha}$ since one is taken at constant p_α and the other at constant u^α . This is a tremendous source of confusion which is avoided by the use of the two symbols.

Kinematic equivalence means having an invertible relationship between velocity and momentum. If the state of the system (q^α, p_α) is enough to determine the trajectory $x^\alpha(s)$, then it will be enough to determine the position and velocity (x^α, u^α) . These will be enough to reconstruct the state as they span a space of equal dimension. As $q^\alpha = x^\alpha$, then $u^\alpha = u^\alpha(q^\beta, p_\gamma)$ is invertible.

Kinematic equivalence leads to Lagrangian mechanics. Since $u^\alpha(q^\beta, p_\gamma)$, we can write $\mathcal{L} = \dot{q}^\alpha p_\alpha - \mathcal{H}$ as a function of (x^α, u^α) . That is the Lagrangian of the system. Note that this cannot be done if we do not have kinematic equivalence. For example, for a photon treated as a particle, we have $H = c|p|$. Using Hamilton's equations, we have $u^i = \frac{dx^i}{dt} = c \frac{p^i}{|p|}$ which is not invertible. While we can write the Lagrangian, this cannot be expressed in terms of position and velocity.

Kinematic equivalence leads to linear relationship between velocity and momentum. Being able to go from (q^α, p_α) to (x^α, u^α) and vice-versa is not enough. We need to express the density ρ in terms of position and velocity as well. This means expressing area elements $dq^\alpha dp_\alpha$ in terms of dx^α and du^α . Since $dx^\alpha = dq^\alpha$, we must have $dp_\alpha = m g_{\alpha\beta} du^\beta$ where $g_{\alpha\beta}$ is an arbitrary linear function and m is an arbitrary constant of proportionality.

Linear relationship between velocity and momen-

tion leads to massive particles under potential forces. We can integrate the relationship $dp_\alpha = mg_{\alpha\beta}du^\beta$ and find $p_\alpha = mg_{\alpha\beta}u^\beta + A_\alpha(x^\gamma)$ where A_α are arbitrary functions. Since $u^\alpha = \frac{dq^\alpha}{ds} = \frac{\partial \mathcal{H}}{\partial p_\alpha}$, we can take the previous relationship, integrate it again and find $\mathcal{H} = \frac{1}{2m}(p_\alpha - A_\alpha)g^{\alpha\beta}(p_\beta - A_\beta) + V$. We recognize this to be the Hamiltonian for a massive particle under potential forces.

Inertial mass measures the number of states per unit of velocity. Inertial mass is usually introduced using the Newtonian relationship $F = ma$: the lower the mass, the easier is accelerate the body. If the mass were zero, then, we would expect the body to be infinitely easy to accelerate. If the mass is zero, however, the velocity is fixed and the body cannot be accelerated. The derivation offers a different insight: inertial mass is the constant of proportionality between ranges of velocity and ranges of states, as measured by momentum. The higher the mass, the more possible states are there per unit of velocity. If there are more states per unit of velocity, then the body is more difficult to accelerate as we have to go through more states, which recover the Newtonian intuition. However, if the mass is zero, set of states defined in a unit of velocity has measure zero: we cannot change state, we cannot accelerate.

The speed of light is a ration between the possible states in a unit of space over the possible states in a unit of time. Regarding the constant c as a speed is problematic. It is unclear why a speed, which is a property of motion, would be part of the metric of the space-time, which is defined regardless of motion and of what is there in the space itself. The derivation offers a different insight. As we move a state in space, it will go through different possible states: this can be quantified in phase space. As we move a state in time, it will also go through different possible states: this can be quantified on the extended phase space. The ratio between the two is the constant c . If we imagine a particle moving in both space and time, the position may remain the same but the time will always need to change. As the motion is deterministic and reversible, at most we can have one position in space for every position in time. Therefore, over a finite time, we must have $\frac{\Delta x}{\Delta t} \leq c$. Thinking about

the constant c as relating the count of states in space over time is more fundamental.

The Planck constant, the speed of light and the inertial mass serve to transfer the measure of states from one quantity to another. As we saw initially, (q, k) are the quantities on which we can measure the number of states. The constant \hbar serves to transfer the count from dk to dp . The inertial mass m serves to transfer the count from du to dp . The speed of light served to transfer the count from dt to dx . They are book-keeping tools we use as we express states over different quantities which allow us to express in different ways the fundamental relationship coming from determinism and reversibility: the count of states must remain the same.

5.1.4 Status and open issues

We consider classical mechanics to be essentially forward engineered. There are some details, though, that could be better understood. The formalization will need a lot more time as it has many mathematical prerequisites.

Physical meaning of the Poisson brackets. We still miss an understanding of the Poisson brackets as a generalization of Jacobian determinants. In principle, they should tell us how densities over a degree of freedom formed by two variables transform. We haven't found the right approach.

Densities as measure theoretic derivatives. As we want finite objects to be our starting points, densities should be the ration of amount of "stuff" divided by the size of the region. This, in measure theory, is known as the Radon-Nikodym derivative. We want this idea to be combined with statistics and differential geometry, so that we can talk about marginal distributions. It seems that geometric measure theory at least partially addresses that, which would be the first step for a full formalization of this section of the project.

Relationship between metric tensor and symplectic form. The way the derivation works, the metric tensor is a secondary object that should be understood as being applied to a differential of position and a differential of velocity (e.g. $dx^\alpha g_{\alpha\beta} du^\beta$). This gives us the number of states on an area defined by the

differentials. It is unclear whether the curvature of $g_{\alpha\beta}$ is linked to the geometry set by the symplectic form.

5.2 Quantum mechanics

Quantum mechanics follows the same strategy as classical mechanics with the difference that it assumes irreducibility instead of infinitesimal reducibility. From this, one can show that the state of the overall system is a complex inner product vector space. The deterministic and reversible evolution corresponds to unitary evolution. From kinematic equivalence one recovers Lagrangian and potential forces as in the classical case

[To be completed]

5.2.1 Quantum state space

Irreducibility. A system is said irreducible if giving the state of the whole system tells us nothing about the state of its parts. For example, given an electron, we can scatter a photon off of it to learn its state. Yet, we cannot scatter with only a part of the electron, we only scatter with the whole electron. The system is irreducible because we cannot study, and assign a state, to its parts.

Divisibility vs reducibility vs decomposability. There are three concepts that are often confused to one another. Divisibility is the possibility to divide a system into two independent ones. That is, we have a time evolution such that we start from a state of a system and we end up with two states of independent systems. Reducibility is the ability to describe a system as the composition of more systems. That is, giving the state of the system at one time is the same as giving the state of the parts at the same time. These are independent properties.¹⁷ Decomposability is the ability to treat one object as the composition of others. We take two objects or descriptions and combine them formally to give a new one. This

¹⁷For example, the Planarian worm is divisible into two worms but is not reducible to two worms. A muon is divisible (i.e. it decays) into an electron and two neutrinos but is not reducible to them. A magnetic is reducible to a north and a south pole but is not divisible into them. A proton is reducible to quarks and gluons but is not divisible into them.

is again independent from the rest and often is just mathematical convenience.¹⁸

[To be completed]

5.2.2 Status and open issues

While we consider quantum mechanics to be essentially forward engineered, that is we know the assumptions are necessary and sufficient, there are some conceptual aspects and details that, if clarified, would paint a more satisfactory picture.

Link between quantum hydrodynamics, stochastic mechanics, Bohmian mechanics. We should be able to reconcile the fact that quantum mechanics can be reframed in these different ways. The fact that hydrodynamics-like equations emerge should be linked to the idea that a quantum is a finite size object that can be thought as made of infinitesimal parts, as a fluid. Each part, though, is subjected by the Brownian motion associated with the unknowable internal dynamics. Stochastic mechanics, then, characterizes those fluctuations while Bohmian mechanics characterizes the average paths.

Deriving quantum from classical. If the above characterizations are understood, we should be able, in line of principle, to take a classical system, couple it with a suitable purely stochastic dynamic and recover quantum mechanics.

5.3 Thermodynamics and statistical mechanics

Reverse engineering work for thermodynamics and statistical mechanics is ongoing.

5.4 Further work

The work still needs to be extended to field theories, classical (electromagnetism and general relativity) and quantum (quantum electrodynamics, quantum chromodynamics, ...). Some ideas have been explored but only at a preliminary stage.

¹⁸For example, velocity in any direction can be decomposed into the components along the axis, but is not divisible or reducible to those components. A pure state is decomposable into a linear combination of a basis but is not reducible or divisible into the basis.

We have not yet attempted to work with field theories, either classical or quantum. On the quantum side, there are two obstacles we would need to solve. The first is how to characterize relativistic spin and the second is the right mathematical tools to use (e.g. algebraic geometry?). On the classical side, the main conceptual issue is that we start with distributions over q and p so it is not clear what these would represent as distributions over field values.

General relativity. Classical mechanics as derived in this project is already relativistic, in the sense that we derive a metric tensor that is locally Minkowskian. However, it is not clear whether that metric has a link to the energy/momentum tensor or it is independent. One idea to explore is that the presence of a field changes how states are counted over kinetic momentum (i.e. $\{mu^\alpha, mu^\beta\} = F^{\alpha\beta}$). It remains to be seen whether the curvature of the metric can be understood in terms of this insight.