

Classical Mechanics Is Lagrangian; It Is Not Hamiltonian[†]

Erik Curiel[‡]

ABSTRACT

One can (for the most part) formulate a model of a classical system in either the Lagrangian or the Hamiltonian framework. Though it is often thought that those two formulations are equivalent in all important ways, this is not true: the underlying geometrical structures one uses to formulate each theory are not isomorphic. This raises the question whether one of the two is a more natural framework for the representation of classical systems. In the event, the answer is yes: I state and sketch proofs of two technical results, inspired by simple physical arguments about the generic properties of classical systems, to the effect that, in a precise sense, classical systems evince exactly the geometric structure Lagrangian mechanics provides for the representation of systems, and none that Hamiltonian mechanics does. The argument not only clarifies the conceptual structure of the two systems of mechanics, their relations to each other, and their respective mechanisms for representing physical systems. It also shows why naively structural approaches to the representational content of physical theories cannot work.

[†]This paper was submitted for review to *British Journal of Philosophy of Science*, July, 2011.

[‡]I thank Robert Geroch for many stimulating conversations in which the seeds of this paper's technical arguments were germinated and, in some cases, fully cultivated. I also thank John Norton for making me explain the technical arguments to him with enough care and thoroughness so as to allow him to offer many useful suggestions about how to give them a simpler, more digestible form, without the need for heavy machinery throughout. I am grateful as well for feedback from receptively critical colloquium audiences at: the mathematics departments at Rutgers and Penn State Universities and at Queen Mary College, London; the Center for Philosophy of Science at the University of Pittsburgh; and the Department of Philosophy, Logic and Scientific Methodology at the London School Of Economics. Email: erik@strangebeautiful.com

Contents

1	Introduction	2
2	Abstract Classical Systems	5
3	The Possible Interactions of a Classical System and the Structure of Its Space of States	8
4	Classical Systems Are Lagrangian	17
5	Classical Systems Are Not Hamiltonian	23
6	How Lagrangian and Hamiltonian Mechanics Respectively Represent Classical Systems	31
7	The Conceptual Structure of Classical Mechanics	39
	References	43

[Lagrange] grasped that he had gained a method of stating dynamical truths in a way which is perfectly indifferent to the particular methods of measurement employed in fixing the positions of the various parts of the system. Accordingly, he went on to deduce equations of motion which are equally applicable whatever quantitative measurements have been made, provided that they are adequate to fix positions. The beauty and almost divine simplicity of these equations is such that these formulae are worthy to rank with those mysterious symbols which in ancient times were held directly to indicate the Supreme Reason at the base of all things.

A. N. Whitehead

Science and the Modern World (p. 63)

1 Introduction

Many physical theories have different mathematical frameworks within which one can construct models of physical systems the theory purports to treat. Quantum mechanics, for example, can be formulated using either Heisenberg's matrix mechanics or Schrödinger's wave mechanics. Classical mechanics is no exception. One has available both the frameworks of Lagrangian mechanics and

Hamiltonian mechanics in which to formulate the equations of motion of classical systems and construct concrete models for them.¹

It is often held that these different frameworks are essentially equivalent. (Even when this is not explicitly stated, the treatment of the two often seems to assume it implicitly.) One will usually be more convenient to work with for certain sorts of investigations, but, the thought goes, there are always methods for passing back and forth between the two, and in this passage nothing of significance is gained or lost. I claim, to the contrary, that not only are they not equivalent in any interesting, physically significant sense, but, more to the point, they are strongly inequivalent (if one may say so) in an interesting, physically significant sense.

Given that is so, the question naturally arises: is there any natural sense in which classical mechanics evinces the character of one rather than the other? In this paper, I argue that there is a natural sense in which this is so, and that, in the event, classical mechanics evinces the character of Lagrangian mechanics: a classical system (roughly speaking, one appropriately modeled by Newton's Second Law) naturally accrues to itself a representation in the framework of Lagrangian mechanics, based on the intrinsic geometry of its family of possible dynamical evolutions. Its representation in Hamiltonian mechanics, by contrast, does not arise naturally based on any structure intrinsic to its representation as a classical system, but must rather be constructed by hand and imposed by fiat, in a sense to be made precise. Most strikingly, my arguments do not depend on vague or obscure concepts such as "simplicity" or "amount of structure", and do not require for their acceptance adherence to any metaphysical system. They are grounded in the geometry of the kinematical and dynamical structures of the theories themselves as determined by empirically founded descriptions of the generic behavior of classical systems

I begin in §§2–3 by sketching a generic representation of classical systems independent of both the Lagrangian and Hamiltonian frameworks but rich enough to allow one to pose questions about the proper empirical descriptions of such systems and to allow one to characterize the intrinsic geometry of important elements of that representation, such as that of the family of a system's allowed dynamical evolutions. Based on the arguments and constructions of those two sections, I show in §4 that the geometric structures characteristic of such a generic representation of classical systems is naturally isomorphic, in a physically significant sense, to the intrinsic geometry of the Euler-Lagrange equation; I capture this idea precisely in a theorem (4.1) about the geometry of tangent bundles. In §5, I show why and how there can be no such isomorphism with the intrinsic geometry of Hamilton's equation. These two results form the heart of the argument that classical mechanics is Lagrangian, not Hamiltonian. I compare the way the two frameworks respectively represent classical systems in §6 to address puzzles raised by the arguments of the previous sections. I conclude in §7 with a summary of results, and a discussion of some of the consequences of the

¹Any decent book on analytical mechanics will cover both. All textbooks on analytical mechanics, however, are not created equal, especially with regard to sensitivity to the sorts of foundational questions I shall address here. For that sort of sensitivity, among the best books I have found are [Whittaker \(1947\)](#), [Lanczos \(1970\)](#), [Pars \(1965\)](#), [Rosenberg \(1977\)](#), and [Arnold \(1978\)](#). I am sorry to say that I find the popular and influential [Goldstein \(1980\)](#) to be dismal in this regard.

paper’s arguments for structural approaches to the understanding of physical theory. For those who do not want to work through all the technical details, I give at the beginning of each of §§3–5 a brief summary of what I intend to accomplish and at the end a summary of what has been achieved along with suggestive comments about its role in the overarching argument.²

I go on at some length in §§2–3 characterizing classical systems and showing the intrinsic empirical structures they generically manifest, with perhaps more caution and care than some readers will like, because those arguments and constructions are among the clearest, most compelling cases I know of the derivation of physically significant, formally precise structure accruing to an entire class of physical systems starting from the simplest of empirical assumptions. It is in fact a sterling example of what structuralist views of physical theory generally live for. It is therefore all the more surprising that the constructions can be used to formulate what I think is a devastating counter-example of naively structural views of physical theories. Indeed, the strength of the putative counter-example to such views is in direct proportion to the strength and naturalness of the constructions that lead one to a case that, at first blush, should do nothing but substantiate it.³

One task of the philosopher of physics is to work out the conceptual structure of a physical theory—or at least an appropriate one, if it is not unique—to isolate and characterize its geometrical and formal structures and show how natural physical interpretation accrues to them from an analysis of their empirical application. Two of the most beautiful examples of this I know are Russell’s *The Analysis of Matter* and Howard Stein’s paper “On the Conceptual Structure of Quantum Mechanics”. I hope the arguments and constructions of this paper offer another example of how such a thing can be done for classical mechanics, starting from a set of strikingly few and weak empirical propositions. Those two works, however, wore their justification on their sleeves—the theories they analyzed and explicated, general relativity and the nascent quantum theory in the former and mature quantum mechanics in the latter, were (and still are, for that matter) recognized as the most fundamental ones we have. Why should a similar project for classical mechanics hold philosophical interest?

Indeed, with few exceptions, philosophers of physics have tended to give short shrift to the philosophical problems that seem peculiar to classical mechanics, and in particular those that arise from the foundations of the classical Lagrangian and Hamiltonian frameworks.⁴ I think the neglect is unjustified, for (at least) three reasons. First, sheer intellectual curiosity and the mettle of philosophy jointly demand that we understand better one of the most magnificent intellectual constructions in human history. That construction is not only magnificent intellectually, however, which brings us to the second reason: classical mechanics still tells us something deep about much of the physical world—it’s what we use for moon-shots, after all. It’s hard for a theory to get more empirically successful than that. It could not do so without reflecting something important about the world, even if not at the deepest level. It is in part the task of philosophy to make clear what that something

²Curiel (2012) gives a rigorous exposition of all the technical matter used and discussed in this paper, along with real proofs of the two primary technical results, theorems 4.1 and 5.1.

³My arguments and conclusions run directly counter to, among others, North (2009) and French (2009), with regard to the bearing of this example on such issues. I discuss North’s arguments at the end of §7.

⁴Butterfield (2003, 2004, 2007) and Belot (2007) are notable exceptions.

important is. Finally, just as classical mechanics is still a rich and deep mine for mathematical investigation in many if not most fields of pure mathematics, much of it inspired by argumentation of a blatantly physical flavor (Thom's work on catastrophe theory, for example, and Arnold's work on partial-differential equations), so it can be in philosophy as well, with regard to traditional questions philosophers tend to examine in the context of quantum mechanics or relativity theory. I do not doubt that investigation of classical mechanics on its own can suggest new approaches to the analysis of deeper theories.⁵

2 Abstract Classical Systems

Our first objective is to construct a framework for the description of classical systems in a way independent of the details of any particular theoretical framework.⁶ The description will include a characterization of a classical system's space of states and its family of dynamical evolutions. I will call a system so represented an *abstract classical system*. This abstract characterization of classical systems provides an appropriate framework for the constructions and arguments we require.

I take as fundamental the idea of a *system*: roughly speaking, something one can look at, interact with. A *quantity* associated with a system or type of system is any property a system may bear amenable to experimental observation; it is a (possibly variable) magnitude that can be thought of as belonging to the system, in so far as it can be measured (at least in principle) by an experimental apparatus designed to interact with that type of system, in a fashion conforming to a particular coupling of the system with its environment, which coupling may be modeled theoretically once a theory is in place.⁷ One assumes that, somehow or other—it does not matter for our purposes how—one has fixed on a set of quantities that play a privileged role in the description and comprehension of the system, those that are physically significant. Linear acceleration and angular momentum are physically significant quantities in the Newtonian mechanics of rigid bodies, for example; the temporal derivative of acceleration (the third temporal derivative of position) and the magnitude computed by adding the numerical values of position and velocity at a point are not. (In this case, the physical significance of the former quantities consists in the fact that the equations of motion of motion, when expressed in terms representing those quantities, take on their simple, canonical form, with no extra spurious terms appearing, *i.e.*, the form of the equations of motion respects the dynamics.)

⁵Earman (1986, 2007) on determinism in classical mechanics as compared to other theories is a beautiful example of this.

⁶I do not claim that one can describe actual classical systems in the context of a physical investigation with the use of no particular, theoretical framework, only that there is a way of abstracting from the details of whatever theoretical apparatus may be involved in any given case, in the way I attempt in this section.

⁷This characterization of quantity involves (at least) one serious over-simplification. Not all quantities' values can be determined by direct preparation or measurement, not even in principle. Some, such as that of entropy, can only be calculated from those of others that are themselves directly preparable or measurable. Other quantities are ambiguous in this regard—does the application of a ruler to a system to measure its length count as a coupling of the system with its environment? These subtleties do not affect the paper's arguments.

A *state* of a system is the aggregation of the values of its physically significant properties at an instant; it is represented by a proposition encapsulating all that can be known of the system physically, at least so far as the pre-theoretical, theoretical and experimental resources one relies on are concerned. If one can distinguish the values of the properties of the system at one time from those at another time by the available resources, then the system is in a state at the first time different from that at the second. A state, therefore, can be thought of as a set of the values of quantities that jointly suffice for the identification of a point of the space of states, which is itself the set containing all states one has identified in practice and all those one extrapolates the system can occupy. Because each state assigns a definite value for each quantity to the system, a quantity is represented by a function on the space of states that assigns to each state a definite value of some mathematical entity, such as a real number or a vector in a vector space, depending on the nature of the quantity.

As a brute fact about the physical world, every abstract classical system we know of has this property: it has associated with it a number, either a single positive, even integer or else infinity, which is the minimum number of independent quantities whose values one must fix in order to individuate and identify a state; this number is the same for all states the system can occupy, no matter the set of quantities whose values one uses to label the states, *viz.*, the system's degrees of freedom.⁸ (Mine is a non-standard usage of 'degrees of freedom', which is often taken to refer to the dimension of configuration space, which I will discuss below, not to that of the total space of states.) These facts allow one to attribute further structure to the space of states, those of a topological and a differential manifold. One derives the topology by requiring all quantities to be continuous (except perhaps at a countable number of points), and one derives the manifold structure by requiring that all quantities be smooth (except perhaps at a countable number of points).⁹ So quantities play a dual role, one local and the other global: they individuate and identify the states, and they determine the topological and differential structures of the space of states.

Every known abstract classical system has the property that at least some of its quantities almost always change in value as time passes, which is to say, the system in general occupies different states at different moments of time. The collection of states it serially occupies during an interval of time, moreover, forms a curve on the space of states parametrized by time, a *kinematically possible evolution* (or just 'possible evolution'), which is in general smooth. From the family of all possible evolutions, one constructs the family of kinematically possible vector fields (or just 'kinematical

⁸As stated, the claim is not correct. I know of exactly one example of a classical system that has an odd number of degrees of freedom: it is a simple device, consisting of two rigid discs joined by a straight, rigid axle connected to each by a universal joint at its center; it rolls without friction or slippage on a curved surface. I claim that system has seven total degrees of freedom. (I'm sure other examples along the same or similar lines can be constructed.) I do not know what to make of such anomalies, so I ignore them for the sake of argument. For what it's worth, I know of no account or discussion of classical mechanics, either in the physical or the philosophical literature, that even remarks on their existence, so I can at least claim that my putting of them aside is no worse practice than that of any other investigator I know of.

⁹To allow countable numbers of discontinuities in the quantities requires the use of straightforward but tedious and unilluminating technical machinery to allow for their use in defining the topology and the differential structure; we can ignore these sorts of problems for our purposes.

vector fields'), those whose integral curves (the curves that "thread the arrows in the vector field") are possible evolutions. Because a vector field on a manifold can be thought of in a natural way as a first-order ordinary differential equation, the kinematical vector fields encode the equations of motion for all possible interactions of the system with its environment. The solutions to the equations of motion are by construction the system's possible evolutions. The family of kinematical vector fields provides a description of the possible histories of an abstract classical system equivalent to that given by the family of possible evolution curves, and one more convenient for our purposes, so I will mostly rely on it in the discussion. Finally, there is a distinguished kinematical vector field, the *free kinematical vector field*, that represents the evolution of the system when it is isolated from all external forces.

I claim that this characterization of abstract classical systems comprises all basic, physically significant structure required to found the investigation of classical systems. It does not provide all the structure that comes into play in all forms of investigations, but it does provide all the tools one requires to define and construct all the other structures. A proper defense of this claim is beyond this paper's scope, though the constructions of §3 will go some way towards providing a sketch of one. Nor do I claim that this characterization is canonical in the sense that no other one comprising other natural structures one takes to be basic can be given. I claim only that it provides one natural, necessary and sufficient toolbox and supply of material for the job. Nonetheless, I do claim that the structure it does provide is essential and minimal: no matter what else is the case, if one is going to represent a classical system, one had better have a space of states and a family of vector fields to represent its possible evolutions, or one had better be able to show that one can reconstruct those two based on what one does have. If one cannot, then one is not talking about a classical system in any sense of the idea I am familiar with.

Before leaving the section, I want to record a virtue of this way of thinking of the framework of classical mechanics: it teases out several of the puzzling features of classical mechanics, ones that are otherwise easy to pass by without remark, so familiar are they to those with even a passing acquaintance with it. Although I do not address them in this paper, I think it is worthwhile to pause for a moment to list them. Why is it that no matter what set of different quantities one uses to characterize the state of a system, one always needs only a certain fixed number of such quantities, the same for all sets of quantities (the degrees of freedom), no matter how different the quantities in each set may be from those in other sets? What ought to count as a physically significant quantity? When ought two *prima facie* distinct quantities be treated as physically different? For a Newtonian particle, the quantity formed by multiplying the numerical value of the position by that of the velocity does not seem to be physically significant; it is often unclear whether 5 times or 5 plus a significant quantity ought to count as a different quantity than the original. It is not the case that two quantities ought to be counted as different only if they can be varied independently of each other, for then momentum and energy for a free particle would not be different quantities. Perhaps the most fundamental question, which many of the others boil down to, is this: why does the space of states of a classical system always have the structure of a connected differentiable manifold? A closely related point is that the following appears to be one of the principles of mechanics (whatever

that may come to) for abstract classical systems: there always exists a smooth tangent vector field whose integral curves are the kinematically possible evolutions of the systems. The meaning of this: “ordinary differential equations are appropriate for the modeling of classical systems”. Thus families of curves on the space of states that do not have associated tangent vector fields simply cannot be possible dynamical evolutions of the system. An example of such a family of curves is that for a particle in a square potential well—presumably here one would say that the square potential is an idealization, and that if one looked closely enough it would really be a very steep but still smooth potential. Another example is a family of curves that intersect each other; in this case, simple determinism would fail. There are no *a priori* reasons why any of these facts should hold. Most strikingly, none of these facts depends on the fixation of a particular theory or framework for their statement or for their substantiation. They seem to reach down to and represent structure at a very deep level of our understanding of classical systems.

3 The Possible Interactions of a Classical System and the Structure of Its Space of States

I have gone into such detail in §2 in the characterization of abstract classical systems because, as I will show in this section, that abstract framework already provides the tools for a construction of startling physical strength and depth: starting from only very weak, almost trivial seeming assumptions, one can recover and describe in the framework of abstract classical systems the family of kinematically possible interactions (or just ‘possible interactions’) any classical system can enter into with any other classical system; even more, one can show that the family of possible interactions has a rich algebraic structure, concomitant with one that will show itself in the family of kinematical vector fields; from those objects, finally, one can show that the abstract classical space of states naturally possesses the structure of a space that plays a foundational role in classical mechanics, *viz.*, the tangent bundle of configuration space, the natural theater in which Lagrangian mechanics plays itself out.¹⁰ The strength of the result derives from the weakness of the system one uses to formulate it in and the weakness of the assumptions one uses to prove it. That result grounds the theorems of §§4–5 whose natural interpretation is that classical systems evince exactly the physically significant structure of Lagrangian mechanics, nothing more and nothing less, and none of the physically significant structure of Hamiltonian mechanics.

In traditional accounts of classical analytical mechanics (that is, Lagrangian and Hamiltonian mechanics), one distinguishes three sorts of quantity, the configuration-like (or ‘configurative’), the velocity-like (or ‘velocital’) and the momentum-like (or ‘momental’). The configurative and the velocital are used in the formulation of Lagrangian mechanics, and the configurative and momental for Hamiltonian mechanics. As their names suggest, the defining properties of the velocital and the momental quantities for a generic abstract classical system are the same as (or, at least, very similar to) those for velocity and momentum, respectively, in Newtonian mechanics. Configurative quantities

¹⁰The core of the construction is due to R. Geroch; I learned of it in conversation with him.

are those having many or all of the same significant properties as position in Newtonian mechanics. Indeed, most expositions of analytical mechanics postulate the differences among these as primitive and foundational. We, however, did not have to distinguish between different types of quantities in the characterization of abstract classical systems in §2, and so nothing like a configuration space (the set of all configurations, which naturally accrues the structure of a differential manifold) showed its face; nor did any but the simplest of algebraic and geometric structures appear in the construction of the theory, nor any set of preferred coordinates (quantities), *etc.* At this stage, therefore, nothing in the description of an abstract classical system seems to militate in favor of a Lagrangian as opposed to a Hamiltonian formulation of it, or vice-versa, if that is the sort of thing one wants. Part of the goal of this section, however, is to distinguish configurative and velocital quantities and explicate their properties. (We will not treat momental ones until §5.)

Now, we want a way using only the concepts sketched in §2 to characterize a property or set of properties (physical, structural, what-have-you) of the quantities a system possesses that will differentiate the configurative from the velocital. With that in hand, we will be able to construct the configuration space, in the traditional sense, of the abstract classical system, which in turn will found the argument for the naturalness of Lagrangian mechanics for classical systems. As we will see, the characterization of the possible interactions a classical system can enter into with other systems provides the the tools to answer the question. To start, let us pose a more concrete question about configuration space. What, for example, should one choose as the configuration coordinates in trying to represent an electromagnetic field as a Lagrangian system, the electric or the magnetic field, and why? These questions would perhaps not seem so pressing to one raised on a diet of traditional text-books on analytical mechanics, in which the author generally starts with configuration space (usually presented in some particular coordinate system, the physical significance of which is itself not discussed) and marches courageously forward.¹¹ But the configuration spaces of physical systems were not handed down by Prometheus with fire, and on its face it is rather a mystery where they come from, what they do and why we need them. Simple measurement of abstract classical systems, it seems, will get one at most the abstract classical space of states and the allowed kinematical vector fields, but it would not seem to get one by itself a preferred way to factor the space of states, so to speak, into configurative and velocital parts.

Consider a simple example, say that of a free particle. Its state can be completely described by giving its position and its velocity, each of which can be thought of as a vector in ordinary, three-dimensional space; its space of states is therefore six-dimensional. Using those two quantities to parametrize the space of states (*i.e.*, to label its points, the states), a representation of a state has the form (\mathbf{x}, \mathbf{v}) , where \mathbf{x} is the particle's position and \mathbf{v} its velocity. We say this representation provides a natural coordinate system for the space, for we know already that these two quantities play a privileged role in our comprehension of Newtonian mechanics: in their terms, the equations

¹¹This happens even in texts written by physicists who normally display a sensitivity to philosophical problems, such as Whittaker (1947, ch. II, §26, pp. 34–39).

of motion representing the particle's evolution take their canonical form

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{v} \\ \dot{\mathbf{v}} &= \mathbf{0}\end{aligned}\tag{3.1}$$

(I use emboldened signs to designate vectors. I abuse the notation in the usual ways, using, *e.g.*, ‘ \mathbf{v} ’ to designate promiscuously either a single vector at a point or a vector field, *etc.* A dot indicates differentiation with respect to time, and ‘ $\mathbf{0}$ ’ designates the zero vector or vector field.) This is just Newton's Second Law, written out in more explicit form than is usual, as a system of coupled first-order ordinary differential equations: the temporal derivative of position is velocity, and that of velocity is acceleration, in this case zero since we have postulated that the system experience no force, and force is directly proportional to acceleration. The components of the kinematical vector field associated with this evolution are, in these coordinates, $(\mathbf{v}, \mathbf{0})$. The first component of the vector field measures the rate of change of the position, and the second that of velocity.

Now, if one turns on an interaction with the environment and pushes the particle around, then during the interaction the equations of motion become

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{v} \\ \dot{\mathbf{v}} &= \mathbf{F}_{\text{push}}\end{aligned}\tag{3.2}$$

where \mathbf{F}_{push} is the force exerted on the particle during the interaction (up to a multiplicative constant, the inverse of the inertial mass, which we ignore), which may be a function of anything one likes; the components of the associated kinematical vector field are $(\mathbf{v}, \mathbf{F}_{\text{push}})$. One is *not* free to postulate *any* new vector field one likes to represent the particle's evolution during the interaction, which is to say one cannot simply dream up just any sort of interaction: one can “directly push around” *only* the velocities. As a brute empirical fact, there is no known interaction for changing the equation of motion of \mathbf{x} directly. (Indeed, that is part of why it is more often thought of as a kinematical constraint than an equation of motion, but nothing in the formal structure of the theory itself divorced from our empirical knowledge allows one to distinguish between the equation for $\dot{\mathbf{x}}$ and that for $\dot{\mathbf{v}}$ in any principled way—we discover the difference by empirical investigation.) In consequence, the velocities need not evolve continuously as one switches interactions on and off, for one can in principle turn the interaction on and off as abruptly as one likes, whereas the position always evolves continuously.¹² This empirical fact does allow us to distinguish between position and velocity as physical quantities, and so between their respective equations of temporal evolution.¹³

¹²This claim may appear to conflict with my earlier stipulation that the quantities specify the manifold structure of the space of states by the requirement that they be smooth. The conflict is illusory, though: there will inevitably be slippage between the rigorous mathematical structure one constructs to represent the system and the results of actual measurements one makes. This slippage and how one deals with it will depend on the particular ends of the project at hand, the approximations and techniques one uses, *etc.* One can ameliorate the slippage by allowing some quantities in the formal representation to be discontinuous at a countable number of points, as already mentioned in footnote 9.

¹³Strictly speaking, the claim that only configurative quantities evolve continuously is not correct if one allows true distributional forces. In that case, for example, the position would evolve discontinuously under the action of a force

Consider now the example of a free electromagnetic field, specified, say, everywhere in space at a given instant of time. The space of states in this case is infinite-dimensional. In a natural coordinate system, a state is of the form $(\nabla \cdot \mathbf{B}, \dot{\mathbf{B}}, \nabla \cdot \mathbf{E}, \dot{\mathbf{E}})$, where \mathbf{E} and \mathbf{B} are respectively the components of the electric and magnetic fields in the fixed coordinate system. (It will be clear in a moment why I use a coordinate system in which the components for the state are $\nabla \cdot \mathbf{B}$ and $\nabla \cdot \mathbf{E}$ rather than simply \mathbf{B} and \mathbf{E} .) The equations of motion are Maxwell's equations,

$$\begin{aligned}\nabla \cdot \mathbf{B} &= 0 \\ \dot{\mathbf{B}} &= -\nabla \times \mathbf{E} \\ \nabla \cdot \mathbf{E} &= 0 \\ \dot{\mathbf{E}} &= \nabla \times \mathbf{B}\end{aligned}$$

The components of the associated kinematical vector field are $(0, -\nabla \times \mathbf{E}, 0, \nabla \times \mathbf{B})$. The only allowed interactions transform the equations thus,

$$\begin{aligned}\nabla \cdot \mathbf{B} &= 0 \\ \dot{\mathbf{B}} &= -\nabla \times \mathbf{E} \\ \nabla \cdot \mathbf{E} &= \rho \\ \dot{\mathbf{E}} &= \nabla \times \mathbf{B} + \mathbf{j}\end{aligned}$$

where ρ is an electrical charge density and \mathbf{j} its associated current. The components of the associated kinematical vector field are $(0, -\nabla \times \mathbf{E}, \rho, \nabla \times \mathbf{B} + \mathbf{j})$. It turns out also to be the case here that the equations of motion of one set of coordinates, \mathbf{B} and $\dot{\mathbf{B}}$, allow their components to evolve only continuously no matter what interaction is turned on or off—the functional form of their components do not change, but remain 0 and $-\nabla \times \mathbf{E}$ respectively throughout the interaction—whereas the components of the other set of coordinates can evolve discontinuously, depending on the form of the interaction—again, we can in principle turn the charge density and the current on and off as quickly as we like. This suggests that, by analogy with the case of the Newtonian particles, we take $(\nabla \cdot \mathbf{B}, \dot{\mathbf{B}})$ to encode the system's configuration and $(\nabla \cdot \mathbf{E}, \dot{\mathbf{E}})$ its phase velocity.

So far as is known, it is a brute fact about the physical world that all systems adequately and appropriately described by classical mechanics have this property: the equations governing the evolution of only some physical quantities associated with the system change in form during its interactions with the environment, whereas the equations of others remain always formally the same no matter the interaction—one can “directly push around” only some of the quantities by coupling the system to another by an interaction. So one generalizes: the configurative quantities are those that one cannot directly push around via any of the allowed interactions of the system with its environment, and so they are the ones that evolve continuously—hit a system with a stick, and the quantities that change continuously across the change in evolution are configurative. It does that had the form of the derivative of a delta-function. Since I do not use this criterion in the formal construction and argument that shows how to rigorously characterize configurative arguments, such technicalities are irrelevant for our purposes. I intend the idea here only to give heuristic motivation for the formal argument.

not seem possible to give a similarly natural characterization of velocal quantities. We cannot say that they are the ones whose governing equations alter in form or the ones whose values can change discontinuously under interactions, for the same holds as well for mixed quantities such as $\mathbf{x} \cdot \mathbf{v}$ for the particle.¹⁴ One is tempted to say that velocal quantities are the ones that “couple directly with external forces”, but I do not see a way of making that idea precise in a way that excludes non-physical, mixed quantities. Perhaps one can say that the physically significant velocal quantities always seem to be characterized by a kinematical constraint expressing them in terms of dynamical derivatives of configurative quantities, but it is not clear to me that even this is true— \mathbf{E} and \mathbf{B} , for instance, are more or less symmetric in this respect. Neither can one say that the velocal quantities are the ones functionally independent of configurative ones, since in fact in many cases they are not, being related to configurative ones by way of a differential relation. We will see below, however, that, though velocal scalar quantities cannot be characterized in a way analogous to that for configurative ones, velocal directions in the space of states, so to speak, can be characterized in a precise sense.¹⁵

This result now justifies thinking of $\dot{\mathbf{x}} = \mathbf{v}$ as a kinematical constraint rather than an equation of motion—the empirical observation gave us enough information to impose a way to differentiate the two formally identical systems of equations (3.1) and (3.2). A kinematical constraint in a theory imposes fixed relations that must hold among the possible values of some set of a system’s physical quantities at all times, no matter the interaction the system actually enters into with its environment, in order for one to be able to apply the theory to appropriately model a physical system. Theories do not predict kinematical constraints; they demand them. I take a prediction to be something that a theory, while meaningfully and appropriately modeling a given system, can still get wrong. Newtonian mechanics, then, does not predict that the kinematical velocity of a Newtonian body equal the temporal rate of change of its position; rather it requires it as a precondition for its own

¹⁴It is my sense that such faux-quantities as $\mathbf{x} \cdot \mathbf{v}$ never play a significant role in an appropriate and adequate model of a system. The present analysis perhaps sheds light on why this should be: the configurative and velocal quantities of an abstract classical system play different roles in the kinematics of the system’s representation, the one independent (“to first-order”) of couplings to the environment and the other entering directly into those couplings and (often) constrained to satisfy a differential equation posed in terms of a configurative quantity, such as $\mathbf{v} = \dot{\mathbf{x}}$. It does not seem possible for a single physical property of a system to play both roles at once. Still, this cannot be entirely right, for the angular momentum of a Newtonian particle with respect to a fixed origin looks just like one of these mixed quantities. So in the end the puzzle seems to remain.

¹⁵In many accounts of Lagrangian mechanics, the configurative quantities are taken as somehow primary and the velocal as derived. My approach has the virtue of making clear that neither configurative nor velocal quantities ought to be treated as primary or prior in any way; they each stand on their own as physically significant and in principle kinematically independent entities. Some philosophers, to the contrary, have taken the standard sort of exposition, in which configurations are primary, to imply metaphysical theses of extraordinary weight. Wallace (2003, p. 164), for example, says, “The only ontologically primary entities in this picture are the configurations and the paths through them: momentum, for instance, is only a derivative property of a path, and (unlike in Hamiltonian mechanics) cannot be regarded as on a par with configuration.” This remark becomes particularly poignant in light of the fact that, as we will see below in §5, one cannot even define the notion of “configuration” in Hamiltonian mechanics in any principled way. These remarks will become clearer after the exposition of Lagrangian mechanics and the arguments connecting it to abstract classical systems, in §4.

applicability. It can't "get it wrong". If the kinematical constraints demanded by a theory do not hold for a family of phenomena, that theory cannot treat it. One easy way to see this: if one uses equations (3.2) to predict the motion of a particle experiencing forces but one's prediction fails for the equation governing $\dot{\mathbf{v}}$, it does not necessarily mean that one is not dealing with a classical particle, for one may simply have not accounted for all the forces the particle experiences; if one's determination fails for $\dot{\mathbf{x}}$, however, one is not dealing with a classical particle, for no error in accounting can explain that failure. Thus, it does seem proper to think of $\dot{\mathbf{x}} = \mathbf{v}$ as a kinematical constraint rather than an equation of motion in our example. This is a serious difference in physical significance, which, among other things, should be reflected in the way an account of the meaning of a theory assigns significance to the theory's structural elements.

We now have a characterization of configurative quantities derived from the intrinsic physics of abstract classical systems. The argument that led to that characterization will now allow us to attribute a rich geometrical structure to the family of kinematical vector fields, and at the same time to introduce an important new family of vector fields. Because the allowed interactions directly affect only the values of the velocital quantities, the difference of two kinematical vector fields will always yield a vector field whose configurative components vanish but whose velocital components do not. It is easy to see this from a quick look at the expressions for the kinematical vector fields of the two examples—the first component is the same, so their difference is zero. For the particle, for example, fix two forces F_1 and F_2 and consider the respective kinematical vector fields for each, $(\mathbf{v}, \mathbf{F}_1)$ and $(\mathbf{v}, \mathbf{F}_2)$. Their difference is

$$(\mathbf{0}, \mathbf{F}_2 - \mathbf{F}_1)$$

Similarly, for a Maxwell field the difference of the kinematical vector fields for two different charge and current distributions (ρ_1, \mathbf{J}_1) and (ρ_2, \mathbf{J}_2) takes the form

$$(0, \mathbf{0}, \rho_2 - \rho_1, \mathbf{J}_2 - \mathbf{J}_1)$$

or, more suggestively,

$$((0, \mathbf{0}), (\rho_2 - \rho_1, \mathbf{J}_2 - \mathbf{J}_1))$$

These difference-vectors point only in velocital directions, as it were. Because the components of vectors on the space of states represent the rates of change of the quantities that form the coordinates, one can also say that these difference vectors encode only non-trivial rates of change for velocital quantities, *viz.*, accelerations.

It is easy to see that all the vectors of that form have the structure of a vector space: if I add two of them or multiply one by a real number, I get another vector of the same form. I shall call such a vector field an 'interaction vector field', because it encodes all and only information about a kinematically possible interaction the system can enter into. The examples, moreover, make it plausible that the addition of any interaction vector field to any kinematical vector field should itself yield a kinematical vector field—this makes physical sense, because I can in principle hit the particle with as big or as small a stick as I want no matter its present dynamical state, and turn

on as large or as small a charged current as I choose to couple with the Maxwell field. Thus the set of kinematical vector fields has the structure of an affine space, modeled on the vector space of interaction vector fields.¹⁶

We are finally in a position to show first that the form of the family of interaction vector fields allows one to construct a system's configuration space in a distinguished geometrical way, and then to show that the abstract classical space of states is naturally isomorphic with a very important space associated with its configuration space, *viz.*, its tangent bundle. That will complete the arguments of this section. (That the space of states turns out to be naturally isomorphic to the configuration space's tangent bundle is important because the tangent bundle is the natural setting for the formulation of Lagrangian mechanics, though this fact is obscured in standard presentations.) Now, because interaction vector fields point, so to speak, only in velocital directions, at any given point they pick out a subspace of the space of states that is half the dimension of the space of states itself.¹⁷ (One can see this in the examples by noting that the representation of a vector always has the same number of velocital as configurative slots.) If I fix a point on the space of states and follow all the interaction vector fields off that point, then, intuitively speaking, I'll end up passing through a subspace of the space of states of half its dimension. This constitutes an equivalence relation, "is connected by an integral curve of one of the interaction vector fields to"; each of those equivalence classes, moreover, does indeed form a subspace of the space of states of half its dimension. Because those subspaces are disjoint, there is a natural representation of the space of states as the collection of all of them suitably "glued together". Now, if I move from point to point in one of those subspaces, the configuration of the system does not change by construction, because, again, all the interaction vector fields defining the subspace point only in directions in which velocital but no configurative quantities change in value. It follows that all the points in a single subspace in some sense represent the system as having the same configuration. We can thus construct the configuration space of the space of states by forming an abstract collection of points in one-to-one correspondence with those equivalence classes—the abstract point associated with a given equivalence class represents the system's configuration that class corresponds to. Configuration space so constructed inherits the structure of a differential manifold from the abstract classical space of states by way of the standard construction on quotient spaces.¹⁸

¹⁶An affine space is roughly speaking a vector space in which one "forgets the zero vector". Only the difference of two elements in an affine space is defined, and it is always defined to be an element of a vector space, the one the affine space is modeled on. The sum of two elements of an affine space is not defined. See, *e.g.*, [Artin \(1957\)](#) for a rigorous exposition.

¹⁷The latter property, obviously, depends on the fact that the space of states is even-dimensional in the finite case; see footnote 8.

¹⁸I ignore technical details about possible obstructions (*e.g.*, topological) to the construction of the quotient space, as we expect spaces representing ordinary classical systems not to manifest such pathologies. This is not to say that such cases present no problems of physical and philosophical interest. In the case of Norton's Dome ([Norton 2008](#)), for example, it follows from the discussion of [Malament \(2008\)](#) that one must exclude from the space of states all points representing the marble as sitting at the dome's apex in order to employ this construction and so realize configuration space as a quotient space. I do not think extreme and pathological examples of that sort bear on this paper's arguments. If one had an axe to grind about whether or not Norton's Dome is "really" a classical system,

It is worth remarking as an aside that, on this conception of classical systems, the kinematical vector fields are more fundamental in a sense than interaction vector fields, as one constructs the latter from the former; one may take this, if one likes, as the geometrical representation of Newton’s methodological imperative, stated in his preface to the *Principia* (Newton 1726, p. 4), to begin with the phenomena, the behavior of bodies, and derive therefrom the forces acting on the bodies: “for all the difficulty of philosophy seems to consist in this—from the phænomena of motions to investigate the forces of nature, and then from these forces to demonstrate the other phænomena.” Likewise, the kinematically possible vector fields are more fundamental than configuration space—and so this analysis yields a surprising conclusion. What counts as a configurative quantity for abstract classical systems cannot be determined by examining a system in isolation; there is rather a deep connection between configurative quantities and how the system can interact with its environment. In other words, configuration space is an implicit description of the allowed interactions of the system with other systems: what counts as a configurative quantity is *not intrinsic to the system*, but is rather a property of the system’s allowed forms of coupling to other systems, the dynamical relations it can stand in to other systems (which, presumably, depends in part on the properties of those other systems).¹⁹ This constitutes much of the physical significance of what it means for a quantity to be “configuration-like”.

Now, to show that the abstract classical space of states is isomorphic to the tangent bundle of configuration space, I first sketch the idea of a tangent bundle. Roughly speaking, a point of a space’s tangent bundle is an ordered pair consisting of a point of the space itself and a vector tangent to the space at that point. Thus, the tangent bundle associates with every point in the original space the vector space of all vectors tangent to the space at that point. To get an intuitive feel for the thing, imagine a globe with a very thin, flat glass plate touching it at exactly one point, such that every ray in the plate originating at the point of contact—the osculating point—makes the same angle with the globe; the plate then is the tangent plane of the point, containing every vector tangent to the globe at that point. We can make this into a vector space by declaring the zero vector to be the osculating point in the plane. Every vector originating at the osculating point and contained in the plane, then, is a vector tangent to the globe at that point. The tangent bundle itself is the space that results from “gluing together” all those tangent planes into a single whole. A point of the globe’s tangent bundle, then, consists of a point of the globe and a vector tangent to the globe at that point, *i.e.*, an element of the point’s tangent plane. The tangent bundle of an arbitrary manifold is the analog of the globe’s tangent bundle: we define a vector tangent to a point of the manifold to be a vector tangent to a curve passing through the point; and a point of the space’s tangent bundle is then an ordered pair consisting of a point of the space and a vector tangent to the space at that point. Thus, one can think of a manifold’s tangent bundle as the collection of all tangent planes over every point of the manifold smoothly glued together into a single space. For a

I reckon one could try to use this paper’s arguments one way or the other. I personally don’t see them militating decisively for either stance.

¹⁹This is one important reason I think points of view such as that of Wallace (2003) (briefly discussed in footnote 15) are not viable.

point p in the manifold, the collection of all points (p, ξ) in the tangent bundle, *i.e.*, all pairs such that ξ is a vector tangent to p , is called the fiber over p .

I construct the isomorphism by showing how to associate a point of the abstract classical space of states with a point of the tangent bundle in such a way that every point of each is associated with exactly one point of the other, all so that the relations of “closeness” among neighboring points in each space is preserved when one moves over to the other space by the mapping. Recall that the family of kinematical vector fields of an abstract classical system has a distinguished member, the free field. Starting from any state, an abstract classical system can freely evolve in any direction, with any velocity; in other words, the free vector field includes all vectors tangent to all configurations, *i.e.*, all the possible rates of change of that configuration starting from that state. Thus the free kinematical field encodes all the system’s possible instantaneous configurations and their dynamical derivatives, and nothing more. It follows that the value of the free vector field at a point of the space of states is naturally associated with the configuration that state attributes to the system and with the dynamical derivative (*i.e.*, the velocity) of the configuration that state attributes to the system. Such an ordered pair, however, is exactly a unique point of the tangent bundle of configuration space, and all points on that tangent bundle are realizable as one of those ordered pairs. It is easy to see that the relations of closeness among points is preserved by construction. It follows that, in so far as the free dynamical vector field is itself a naturally distinguished vector field, one has a naturally distinguished isomorphism from the abstract classical space of states to the tangent bundle of configuration space, completing the construction.

Theorem 3.1 (Geroch, unpublished) *There is a naturally distinguished diffeomorphism from the space of states of an abstract classical system to the tangent bundle of its configuration space.*

(See [Curiel 2012](#) for a precise statement of the theorem and a rigorous proof.)

Before moving on, it is worthwhile to pause to take stock of our progress. From a weak physical assumption inferred directly from observation—that interactions directly couple only with a classical system’s velocital quantities—we have discovered an entirely new structure on the space of states, the vector space of interaction vector fields; we have recovered a rich algebraic structure on the space of kinematical vector fields, that of an affine space; we have discovered a way to characterize configurative quantities in an invariant way that, at the same time, clarifies their meaning (*viz.*, they encode the possible interactions); we have constructed the configuration space of the system; and, by invoking the distinguished kinematical vector field that represents the system’s free evolution, we have shown that the space of states is naturally isomorphic to the tangent bundle over its configuration space. I find these results remarkable for the depth and breadth of physical knowledge they provide of the intrinsic nature of classical systems, especially in light of the weakness of the assumptions we started from.

4 Classical Systems Are Lagrangian

I first describe how the structures of an abstract classical system, when carried over to the tangent bundle of configuration space by the canonical isomorphism constructed in the previous section, allow one to construct a Lagrangian formulation for it. It will follow that abstract classical systems are Lagrangian in a natural, precise sense. I then describe how having in hand a traditional Lagrangian representation of a classical system, in the most minimal sense, allows one to construct its abstract classical representation. In consequence Lagrangian systems are abstract classical systems, in a natural, precise sense.

The Euler-Lagrange equation, the heart of Lagrangian mechanics, takes a scalar field (the Lagrangian) that depends on configurations and velocities and (when well posed) determines as its solution a vector field that gives the evolution of the system.²⁰ That is why Lagrangian mechanics is most naturally formulated on the tangent bundle of configuration space: the function that determines the kinematically possible evolutions has as its domain ordered pairs consisting of a configuration and a velocity at that configuration, which is just a point of the tangent bundle; and the evolution of a body consists of a curve comprising pairs of configurations and velocities at those configurations, which is just a curve on the tangent bundle.

Now, there is a natural way to associate a curve on any manifold, such as configuration space, with a curve on its tangent bundle, a procedure known as lifting the curve. A curve on configuration space by definition has a tangent vector at every point it passes through, the one that represents the rate of change of the curve at that point; a point of configuration space and a vector at that point, however, is just a point in its tangent bundle, so the collection of points forming the curve yields a collection of points on the tangent bundle. It is easy to see that the smooth progression of points along the curve on configuration space ensures that the family of points so lifted to the tangent bundle themselves form a smooth curve. Thus, one can lift a vector field on configuration space to the tangent bundle by lifting its integral curves, yielding the vector field everywhere tangent to the lifted curves. I call such vector fields on the tangent bundle *second-order*, because they represent second-order ordinary differential equations on configuration space, just as a vector field on configuration space itself represents a first-order ordinary differential equation. These vector fields are important because they form the family of possible solutions to the Euler-Lagrange equation: it is a theorem that a vector field represents a possible solution to the Euler-Lagrange equation if and only if it is second-order. This has as an important consequence that not every vector in the tangent plane over a point of the tangent bundle is admissible as initial data for the Euler-Lagrange equation, for not every vector is second-order. Only those that respect the kinematical constraints of Lagrangian mechanics, which as will become clear below is equivalent to being second-order, are admissible.

I claim this makes sense on physical grounds. One can think of a second-order vector as the accel-

²⁰In fact, this is true only of the homogeneous Euler-Lagrange equation; the inhomogeneous Euler-Lagrange equation can include so-called generalized forces, represented by 1-forms on the tangent bundle. Although the difference between the two is important in several of the rigorous, technical arguments, we do not need to worry about it for our purposes.

eration of a curve on configuration space: in so far as a vector on the tangent bundle can be thought of as an infinitesimal change in the configurative directions plus one in the velocital directions, a second-order vector always has its infinitesimal rate of change in configurative directions equal to the kinematical velocity of a body traversing the curve lifted from configuration space, and its infinitesimal rate of change in velocital directions equal to the rate of change of the kinematical velocity along the curve, *i.e.*, the body’s acceleration. In other words, a second-order vector field represents physical evolutions that respect the kinematical constraints connecting configurative quantities to their respective, associated velocital ones. This is why it makes sense on physical grounds for solutions to the Euler-Lagrange equation to be second-order vector fields. Thus, the second-order vector field on the tangent bundle that represents the evolution of a free Newtonian particle, for example, has the same form as the free kinematical vector field on the abstract classical space of states in the example of §3, $(\mathbf{v}, \mathbf{0})$, and likewise for a particle experiencing a force \mathbf{F} , (\mathbf{v}, \mathbf{F}) . This suggests there is an intimate relation between the kinematical vector fields of an abstract classical system and the possible solutions to the Euler-Lagrange equation. That will turn out to be correct.

In order to represent these objects in explicit terms, first note that a coordinate system (q_i) on configuration space \mathcal{C} naturally induces one on its tangent bundle $T\mathcal{C}$, (q_i, v_j) , where the v_i represent vectors tangent to curves on \mathcal{C} when those curves are parametrized in terms of the q_i —in other words, $v_i = \dot{q}_i$. These natural coordinates on the tangent bundle are the generalization of (\mathbf{x}, \mathbf{v}) on the abstract classical space of states of the Newtonian particle. ((q_i, v_i) are natural on the tangent bundle in the same sense as (\mathbf{x}, \mathbf{v}) are on the abstract classical space of states: the Euler-Lagrange equation takes on its canonical form when expressed in their terms—the coordinates respect the dynamics.) We will represent vectors explicitly as sums over the basis of vectors defined by natural coordinate systems. For the standard Cartesian coordinates (x, y) on the plane, for example, the vectors defined by the x coordinate, which we write ‘ $\frac{\partial}{\partial \mathbf{x}}$ ’, are the unit vectors pointing parallel to the x -axis, one at each point of the plane, and the same for the y -coordinates; at every point of the plane, then, any vector \mathbf{k} can be written

$$k_x \frac{\partial}{\partial \mathbf{x}} + k_y \frac{\partial}{\partial \mathbf{y}}$$

where k_x are its x -components, *etc.*

Now, fix a natural coordinate system (q_i, v_j) on $T\mathcal{C}$. Any second-order vector field $\boldsymbol{\xi}$ can be written in the form²¹

$$v_i \frac{\partial}{\partial \mathbf{q}_i} + \xi_j \frac{\partial}{\partial \mathbf{v}_j} \tag{4.1}$$

in any natural coordinate system, where the ξ_i are arbitrary functions of q_i and v_j . The fact that v_i , the kinematical velocity, is the component of the configurative part of the vector encodes the fact that $\boldsymbol{\xi}$ is second-order. There is another class of naturally distinguished vector fields on $T\mathcal{C}$, the vertical ones. A vertical vector field is one whose elements point straight up and down the fibers,

²¹I use the Einstein-summation convention, where one implicitly assumes a sum over the values of all repeated indices. Thus, for example, ‘ $v_i \frac{\partial}{\partial \mathbf{q}_i}$ ’ is short-hand for ‘ $\sum_i v_i \frac{\partial}{\partial \mathbf{q}_i}$ ’.

i.e., that point only in non-configurative directions. Any vertical vector field $\boldsymbol{\eta}$ has the form

$$\eta_i \frac{\partial}{\partial \mathbf{v}_i} \tag{4.2}$$

where the η_i are arbitrary functions of q_i and v_j . The families of vertical and second-order vector fields have natural structures. The vertical vector fields form a vector space; one can see this from the generic expression (4.2). The second-order vector fields form an affine space modeled on the vertical vector fields, as one can see from the generic expression (4.1).

Now, in order to write down the Euler-Lagrange equation, we will need first to characterize some of the intrinsic geometrical structures that live on tangent bundles.²² First, every tangent bundle comes naturally equipped with a distinguished vertical vector field $\boldsymbol{\Lambda}$, the Liouville vector field, which in natural coordinates takes the form

$$\boldsymbol{\Lambda} = v_i \frac{\partial}{\partial \mathbf{v}_i}$$

In effect, this vector tells you “how far a point at which it lives is from the origin” in the bundle’s fibers. Every tangent bundle also comes naturally equipped with a 1-up, 1-down index tensor field \mathbf{J} known as an almost-tangent structure.²³ (As a 1-up, 1-down tensor, one can think of it as a linear map from tangent vectors to tangent vectors, or else as a linear map from differential forms, such as the gradient of a scalar field, to differential forms, or else as a linear map from pairs consisting of a tangent vector and a differential form to a scalar.) In natural coordinates, it has the form

$$\mathbf{J} = \frac{\partial}{\partial \mathbf{v}_i} \otimes \mathbf{d}q_i$$

where $\mathbf{d}q_i$ is the differential of the coordinate function q_i and \otimes is the tensor product. It is easy to see that \mathbf{J} maps every second-order vector field to $\boldsymbol{\Lambda}$. In fact, one can use this property to define the almost-tangent structure if one already knows the family of second-order vector fields, as one can conversely use it to define the family of second-order vector fields if one already knows the almost-tangent structure. In a similar manner, the almost-tangent structure uniquely characterizes the family of natural coordinates on the tangent bundle. A coordinate system (q_i, v_j) is natural if and only if

$$\begin{aligned} \mathbf{J}(\nabla q_i) &= \mathbf{0} \\ \mathbf{J}(\nabla v_i) &= \mathbf{d}q_i \end{aligned} \tag{4.3}$$

²²See, *e.g.*, [Yano and Ishihara \(1973\)](#) and [de León and Rodrigues \(1989\)](#) for a thorough presentation of the geometrical structures I discuss.

²³Technically, an *almost-tangent structure* on a $2n$ -dimensional manifold \mathcal{N} is a smooth tensor field J^a_b satisfying the following conditions:

1. considered as a linear operator on the tangent planes of \mathcal{N} , J^a_b has rank n everywhere
2. $J^a_n J^n_b = 0$

It is not difficult to see that, as a linear operator, the range of J^a_b equals its kernel, an n -dimensional distribution on \mathcal{N} .

Conversely, if one knows the family of natural coordinates on a tangent bundle, one can define the almost-tangent structure. Thus, in a natural sense, the almost-tangent structure completely captures the kinematical relations between the qs and the vs of natural coordinates on the tangent bundle, the analogue of the kinematical constraint $\dot{\mathbf{x}} = \mathbf{v}$ for a Newtonian system.

Now, for a given scalar field on the tangent bundle, the Lagrangian L , the (homogeneous) Euler-Lagrange equation, when written in terms of these natural structures, takes the form

$$\mathcal{L}_{\boldsymbol{\xi}}(\mathbf{J}(\nabla L)) - \nabla L = 0 \quad (4.4)$$

where $\mathcal{L}_{\boldsymbol{\xi}}$ is the Lie derivative with respect to the vector field $\boldsymbol{\xi}$, which, if it exists, is the Lagrangian vector field associated with the Lagrangian L , *i.e.*, the solution to the Euler-Lagrange equation for that Lagrangian. It is straightforward to verify that, when written out in natural coordinates, equation (4.4) has the usual form

$$\frac{d}{dt} \left(\frac{\partial L}{\partial v_i} \right) - \frac{\partial L}{\partial q_i} = 0 \quad (4.5)$$

Not every scalar field yields a well posed Euler-Lagrange equation, however. To state the conditions under which this is true, let us say that a Lagrangian L is *regular* if

$$\det \left| \frac{\partial^2 L}{\partial v^k \partial v^l} \right| \neq 0 \quad (4.6)$$

in any natural coordinate system on $T\mathcal{C}$. Only regular Lagrangians are guaranteed to have unique, second-order solutions. It is not obvious how the condition (4.6) does this. Roughly speaking, the condition guarantees that a certain anti-symmetric matrix the Lagrangian defines be invertible. (In fact, the matrix necessarily has vanishing exterior derivative, and so is a symplectic structure.) Both the invariant form of the Euler-Lagrange equation, (4.4), and the coordinate form, (4.5), mask the presence of this matrix, but it is there under the scenes, and the fact that it is invertible guarantees the existence of a unique solution.²⁴

We are now in a position to address the first problem we set ourselves in this section, the construction of a Lagrangian formulation of an abstract classical system using only the structure the representation of an abstract classical system makes available. We know already from theorem 3.1 that the abstract classical space of states is canonically diffeomorphic to the tangent bundle of configuration space. An argument similar to that used in §3 to prove the theorem shows that under this diffeomorphism the image of the family of interaction vector fields on the space of states is the vector space of vertical vector fields on the tangent bundle, and the image of the family of kinematical vector fields is the affine space of second-order vector fields. Thus, the kinematically possible evolutions of the abstract classical system are exactly the possible solutions to the Euler-Lagrange equation posed on the tangent bundle of the abstract classical system's configuration space. A technical argument given in Curiel (2012) shows that these structures are necessary and sufficient for the construction of the geometrical structures required to formulate the Euler-Lagrange equation.

²⁴One can see it most easily by re-writing the Lie derivative in (4.4) with the use of an arbitrary derivative operator on the tangent bundle; see Curiel (2012).

(It is essentially the same argument that proves theorem 4.1, which I sketch below.) In effect, as I remarked earlier, to know the second-order vector fields is to know the almost-tangent structure, and that is the geometrical structure necessary (and manifestly sufficient) for the formulation of the Euler-Lagrange equation having those vector fields as its family of solutions.²⁵ Equivalently, to know how to formulate the kinematical constraints of abstract classical systems on the tangent bundle, the relations (4.3), is necessary and sufficient for the formulation of the Euler-Lagrange equation. This completes the argument for the first claim of this section: abstract classical systems are Lagrangian.

We now consider the converse problem, as it were, whether in some sense having in hand something like a traditional Lagrangian representation of a classical system allows one to construct its abstract classical representation (in the sense of §§2–3). We will pose the problem in the weakest possible form, to lend correlative strength to the solution: how much, if at all, does the structure of Lagrangian mechanics by itself, that is, the way that the Euler-Lagrange equation associates vector fields with scalar fields, determine the structure of an abstract space of states as a tangent bundle over configuration space? If the Lagrangian structure on its own does allow one to reconstruct an abstract space of states as a tangent bundle over configuration space, then we can avail ourselves of the reverse of the arguments that proved theorem 3.1 to show that the abstract space of states must in fact be diffeomorphic to the abstract classical space of states, that the configuration space of the former is diffeomorphic to that of the latter, that the solutions to the Euler-Lagrange equation on the abstract space of states are the pre-images of the kinematical vector fields, and so on. It would follow that Lagrangian systems are abstract classical ones.

This is our question: if one knew of the tangent bundle of configuration space merely as a differentiable manifold, the abstract space of states (*i.e.*, one did not know that it was a tangent bundle at all, much more the tangent bundle of configuration space in particular), and one also knew the Lagrangian dynamical vector field associated with any given Lagrangian—say one had a black box that spat out the correct Lagrangian vector field three seconds after one fed a Lagrangian into it—would this information alone suffice to reconstruct the abstract space of states as the tangent bundle? If the answer is yes, then we could define a canonical isomorphism from it to the abstract classical space of states by fixing a distinguished vector field on it (presumably, the one representing the free evolution of the system). Then the converse of the argument that proved theorem 3.1 would show that to know how to give a minimal Lagrangian representation of a system on its space of states would *eo ipso* suffice to reconstruct all the structure on the abstract classical space of states accruing to it as an abstract classical system. In the event, the answer to our question is yes.

To state somewhat precisely the theorem that answers the question, let the Euler-Lagrange operator \mathfrak{E} be the (non-linear) functional that takes a scalar field to its associated Lagrangian vector field on a manifold that supports the formulation of the Euler-Lagrange equation in the first place. This means, roughly speaking, that the operator’s space of solutions has the structure of an affine space, which itself can be characterized by the existence of a family of sets of coordinates on the

²⁵North (2009) is thus simply wrong when she claims that one needs a Riemannian metric to formulate Lagrangian mechanics. A Riemannian metric on a tangent bundle by itself is neither necessary nor sufficient for a Lagrangian representation of a system.

manifold, each set satisfying relations that encode the kinematical constraints relating \mathbf{q} and \mathbf{v} considered as natural coordinates on the tangent bundle, *viz.*, (4.3). Thus a manifold has an Euler-Lagrange operator if and only if it has a distinguished family of sets of coordinates that severally satisfy those constraints, which in turn determines a distinguished affine space of vector fields having a few particularly nice properties, the family of solutions of the operator.

Theorem 4.1 *A manifold has an Euler-Lagrange operator if and only if it is a tangent bundle over another space; the operator's action allows one to recover the space over which it is the tangent bundle.*

Roughly speaking, the proof goes like this. If one knows the action of the operator, then one can construct its family of solutions; based on the affine structure of the family of solutions, one can construct the distinguished family of coordinates satisfying the constraints that natural coordinates on the tangent bundle satisfy; those constraints by themselves already determine the manifold to have the structure of a tangent bundle, because they require the existence of an almost-tangent structure with a few particularly nice properties, and it is a theorem that such an almost-tangent structure can exist only on a tangent bundle.²⁶ Thus theorem 4.1 implies that not all manifolds admit the appropriate geometry for the formulation of the Euler-Lagrange equation. One can formulate the Euler-Lagrange equation only on a tangent bundle. It is worth remarking on how surprising this result is: the geometry of the Euler-Lagrange equation is rigid in a way that the geometry of Hamilton's equation is not. Merely being able to *formulate* the equation in the first place, much more solve it, determines the global differential topology of the entire space one would formulate the equation on.

Now all we need to do is to find a distinguished vector field on the tangent bundle that is the analogue to the free kinematical vector field used in the proof of theorem 3.1. The solution to the free Lagrangian is the obvious candidate, the unique solution for which the system experiences no accelerations. It follows by construction that the solutions to the Euler-Lagrange equation map to the kinematical vector fields on the abstract classical space of states and the vertical vector fields map to the interaction vector fields. This completes the argument for the second claim of this section: Lagrangian systems are abstract classical ones.

Before moving on, it is worth remarking once again how strong and deep the results of this section are with regard to our understanding of the physics of classical systems. From the weak premises of §§2–3, themselves founded on entrenched, firm empirical knowledge, we have deduced the fact that classical systems evince the structure intrinsic to Lagrangian mechanics, nothing more and nothing less: the family of possible solutions to the Euler-Lagrange equation is isomorphic to the family of kinematical vector fields of an abstract classical system; the family of possible interaction vector fields of a Lagrangian system is isomorphic to that for an abstract classical ones; and the Lagrangian

²⁶Technically, the conditions are that the almost-tangent structure be integrable in the sense of Fröbenius, and that it be complete in the sense that the flat affine connection it induces on every leaf of the natural foliation of the space it induces be geodesically complete. Both of those conditions are natural ones to impose on a structure representing classical systems. See [Curiel \(2012\)](#) for discussion.

kinematical constraints are the same as for abstract classical systems. It is worth remarking that this deduction would not have been possible had we taken the theory of classical systems to consist only of its family of models, as the semantic view of theories holds, for the constructions relied on structures global in the sense that one can formulate them only as relations among the entire family of models (*e.g.*, the almost-tangent structure), and fundamental in the sense that they are required for a complete account of the meaning of any individual model in the first place, *i.e.*, for the complete semantic interpretation of each of the models (*e.g.*, that the almost-tangent structure encodes the Lagrangian analogue of the kinematical constraints of abstract classical systems).

5 Classical Systems Are Not Hamiltonian

In this section, I briefly review the geometry of Hamiltonian mechanics before discussing the ways it represents classical systems. It will quickly become clear that the structure of an abstract classical system does not give us the resources to construct its Hamiltonian formulation, and so abstract classical systems are not Hamiltonian. Conversely, a minimal Hamiltonian representation of a classical system will not allow one to recover its abstract classical structure, so Hamiltonian systems are not abstract classical ones.

To introduce Hamiltonian mechanics, we first need to review a few definitions. A cotangent vector at a point on a manifold is a linear map from tangent vectors at that point to real numbers. One can think of it as a generalized differential form. (Cotangent vectors are also called ‘1-forms’.) The gradient of a function ∇H is an example of a 1-form—it takes a vector and returns the number that measures the rate of change of the function in the direction the vector determines. The set of all cotangent vectors at a point inherits the structure of a vector space from the set of tangent vectors at the point. A cotangent vector field is a smooth assignment of cotangent vectors to the points of the manifold. The cotangent bundle of a manifold is the same thing as the tangent bundle except that instead of bundling the vector spaces of tangent vectors with each of their respective points on the manifold, one bundles the vector spaces of cotangent vectors with each of their respective points. Thus, a point of the cotangent bundle of configuration space consists of a point of configuration space and a 1-form at the point.

It is well known that one requires only a symplectic structure to formulate Hamilton’s equation. (See, *e.g.*, [Arnold 1978](#).) For our purposes, one can think of a symplectic structure as an anti-symmetric, invertible matrix with vanishing exterior derivative. Thus one can use it to define a linear map from pairs of vectors to scalars, from individual vectors to individual 1-forms, from individual 1-forms to individual vectors, and from pairs of 1-forms to scalars. For a given symplectic structure Ω , Hamilton’s equation is written

$$\xi = \Omega(\nabla H) \tag{5.1}$$

We call ξ the Hamiltonian vector field associated with the Hamiltonian H . (So in this case we treat the symplectic structure as a map from an individual 1-form, ∇H , to an individual tangent vector,

ξ.) It is worth remarking at this point that a symplectic structure is in no way equivalent to an almost-tangent structure, and there is no natural way to construct one from the other.

Now, every cotangent bundle comes equipped with a canonical symplectic structure, Ω . In order to recover the usual formulation in terms of coordinates on the cotangent bundle, fix a natural coordinate system (q_i, p_j) . (The coordinate system is natural in the same sense as those on the tangent bundle: the p_i are the components of the differential forms $\mathbf{d}q_i$ generated by the coordinates (q_i) on configuration space.) In those coordinates, the symplectic structure is

$$\mathbf{d}q_i \wedge \mathbf{d}p_i$$

where \wedge is the exterior (anti-symmetric) product on 1-forms. Hamilton's equation is then written

$$\begin{aligned}\dot{q}_j &= \frac{\partial H}{\partial p_j} \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}\end{aligned}\tag{5.2}$$

It is straightforward to check that a transformation of that coordinate system to any other natural one preserves the form of the symplectic structure and so, *a fortiori*, of equations (5.2) as well. Even though the formulation of Hamiltonian mechanics is often, implicitly, restricted to a phase space (*i.e.*, space of states) isomorphic to a cotangent bundle over configuration space, this need not be the case; any symplectic manifold (a manifold with a symplectic structure) will do, and manifolds not isomorphic to a cotangent bundle can possess symplectic structures. A physically significant example of this occurs in the formulation of the Euler equations of motion for a rigid body as a Hamiltonian system: to construct phase space in this case, one takes the cotangent bundle of the group of spatial rotations and constructs its quotient by the same group of rotations; this space carries a canonical symplectic structure, and Hamilton's equation formulated in its terms is equivalent to Euler's equation written in ordinary configurative and momental coordinates. This space, however, is not diffeomorphic to a cotangent bundle. (See, *e.g.*, Arnold 1978, pp. 318–330, appendix 2.) We therefore generalize natural coordinates on a cotangent bundle to canonical coordinates on a symplectic manifold: a coordinate system on a symplectic manifold is canonical if Hamilton's equations expressed in its terms has the form (5.2).

In a canonical coordinate system, the second half of the coordinates, the p_i , represent momental quantities, the analogue of the velocital quantities in abstract classical systems (and Lagrangian mechanics). That is why (q_i, p_i) for an abstract classical system represents a point of the cotangent bundle: the momentum of a free Newtonian particle is an example of a momental quantity. Such quantities are naturally represented as 1-forms rather than as vectors, as the velocital quantities are. One may wonder why this should be so, especially in light of the fact that the momentum of a Newtonian particle is just $m\mathbf{v}$, the tangent vector representing its velocity multiplied by the scalar representing its mass—surely a scalar multiplying a vector gives another vector and not a 1-form, and so it seems that (q_i, p_i) ought to represent a point of the tangent bundle. To see why momental quantities are properly represented by 1-forms, first note that the momentum of the particle can be

naturally thought of as a linear map from vectors to scalars: it is the map that takes the vector \mathbf{v} to the scalar $\frac{1}{2}mv^2$, the particle's kinetic energy. This interpretation of the momentum may seem odd, too abstract, even unnatural in the context of Newtonian mechanics, but in fact it is exactly what one needs to represent the momental quantities of more complex systems that have more complex relations between the rates of change of its configuration on the one hand and its momenta on the other. Angular momentum, *e.g.*, is not the scalar product of the mass of a rotating body and the rate of change of its configuration. The configuration of a cylinder spinning about its axis does not change at all, but it has non-zero angular momentum. This angular momentum, moreover, does define a linear map from the angular velocity of the cylinder to its rotational kinetic energy. Thus, a 1-form is the proper representation for momental quantities, having the exact form required to capture the relation between generalized momenta and rates of change of generalized configurations. That is the physical meaning of momentum in Hamiltonian mechanics.

Now, the linearity of Hamilton's equation implies that the space of all Hamiltonian vector fields (*i.e.*, those vector fields that are solutions to Hamilton's equation for some scalar field) for all Hamiltonians is a vector space: the sum of two Hamiltonian vector fields associated with two different Hamiltonians is itself the Hamiltonian vector field associated with the sum of the two Hamiltonians, or, more formally, if

$$\begin{aligned}\Omega(\nabla H_1) &= \xi_1 \\ \Omega(\nabla H_2) &= \xi_2\end{aligned}$$

then

$$\Omega(\nabla(H_1 + H_2)) = \xi_1 + \xi_2$$

A straightforward calculation shows, moreover, that the symplectic structure induces the structure of a Lie algebra on that vector space, under the action of the regular Lie bracket of vector fields on a manifold.²⁷ In other words, the vector space of Hamiltonian vector fields is closed under the action of the Lie bracket.

Now, these facts imply that the family of kinematically possible evolutions of an abstract classical system, in so far as they are characterized by interactions with no prior assumption of a geometrical structure as in §3, cannot be naturally represented as Hamiltonian vector fields on phase space, for by definition an affine space is not isomorphic to a Lie algebra over a vector space. It follows that there is no analogous structure in the Hamiltonian representation of a system isomorphic to an abstract classical system's family of interaction vector fields—because the family of Hamiltonian vector fields is not an affine space, one has no way to characterize interactions as independent vector fields defined by the difference between (a Hamiltonian representation of) two kinematical vector fields. One thus loses the capacity to identify configuration space in the Hamiltonian formulation, which had better be the case since phase space may not even be diffeomorphic to a cotangent bundle over configuration space. One way to see this is that one has no way to formulate the kinematical

²⁷A Lie algebra is an anti-symmetric, bilinear product on pairs of vector fields that yields another vector field. The exact definition is not important for our purposes, so if this isn't clear to you, don't worry about it.

constraints appropriate to Lagrangian mechanics, the relations (4.3), for one has no almost-tangent structure, but those kinematical constraints are the natural ones for abstract classical systems. In consequence, not only does the Hamiltonian formulation of a system not allow one to express in a distinguished way the kinematical constraints essential to abstract classical systems, but it does not respect them, for it allows solutions to the equations of motion that violate them. (I discuss this in some detail in §6.)

In sum, then, the family of dynamical evolutions Hamiltonian mechanics admits for systems it represents is not isomorphic to the family of kinematical vector fields of an abstract classical system; the family of interaction vector fields Hamiltonian mechanics admits is not isomorphic to the family of interaction vector fields of an abstract classical system; and Hamiltonian mechanics neither encodes nor even respects the kinematical constraints fundamental to an abstract classical system. Thus Hamiltonian systems are in no way abstract classical ones, and abstract classical ones are not Hamiltonian.

Still, Hamiltonian mechanics does in fact impose its own kinematical constraints among its natural quantities, and it will be illuminating to discuss them. Because quantities are just scalar fields on phase space, we can reframe the idea as the imposition of kinematical constraints on its canonical coordinates. Now, a symplectomorphism is an isometry of a symplectic structure, *i.e.*, a diffeomorphism of phase space that maps the symplectic structure to itself. Let us say that the coordinate vector fields associated with a canonical coordinate system (*e.g.*, $\frac{\partial}{\partial q_i}$ for the q_i) are themselves canonical. Then Hamiltonian mechanics demands of the family of canonical coordinates that all the associated canonical vector fields generate symplectomorphisms (*i.e.*, that the symplectic structure remain constant along the flow-lines of the canonical coordinate axes). Indeed, a stronger statement holds: a coordinate system is canonical if and only if its associated coordinate vector fields are Hamiltonian (*i.e.*, are the solution to Hamilton's equation for some Hamiltonian). That fact gives the precise sense in which the quantities that compose a canonical coordinate system are preferred for the parametrization of the space of states. Not all Hamiltonian vector fields, however, arise in this way (*e.g.*, the zero vector-field). The family of canonical vector fields do, nevertheless, span the vector space of all Hamiltonian vector fields.

It will prove useful to have besides the first two another necessary and sufficient condition for a set of quantities to form a canonical coordinate system. A Poisson bracket, roughly speaking, is an anti-symmetric, bilinear map from pairs of scalar fields to scalar fields that acts in effect like a kind of derivative, measuring the respective rates of change of each function with respect to those of the other. One arises naturally from a symplectic structure Ω : the Poisson bracket for two scalar fields f and g is given by

$$\{f, g\} =_{\text{def}} \Omega(\nabla f, \nabla g)$$

(Here, we treat the symplectic structure as a map from pairs of 1-forms to scalar fields.) Then a

coordinate system (q_i, p_j) is canonical if and only if

$$\begin{aligned}\{q_i, q_j\} &= 0 \\ \{q_i, p_j\} &= \delta_{ij} \\ \{p_i, p_j\} &= 0\end{aligned}\tag{5.3}$$

where δ_{ij} is the Kronecker delta symbol, which equals 1 for $i = j$ and 0 otherwise. It is manifest that families of coordinates satisfying these relations need not satisfy the kinematical constraints of abstract classical (or Lagrangian) systems, and vice-versa.²⁸ Even if phase space is diffeomorphic to a cotangent bundle, for example, one can see by inspection that the constraints allow us to apply a symplectomorphism that does not respect the bundle structure but mixes up the p s and the q s while preserving the form of the symplectic structure. (The simplest one just swaps q_i and p_i —Hamiltonian mechanics doesn't care whether you mind your p s and q s.) In consequence, Hamiltonian mechanics does not respect the kinematical constraints that relate configurative and velocital quantities in abstract classical systems.

One can think of these facts as a way to make precise the idea that configurative quantities and configuration space itself play no distinguished role in Hamiltonian mechanics: one has no tools available to distinguish them in a physically significant way. That makes physical sense, for the symplectic structure, unlike the geometry of the tangent bundle, does *not* allow one to reconstruct the space on which it resides as a cotangent bundle—any even-dimensional, orientable manifold can support a symplectic structure. There is no Hamiltonian form of theorem 4.1.

From a certain perspective, then, Lagrangian mechanics appears to be the more fundamental of the two ways of representing systems in classical analytical mechanics, in the sense that one natural way to describe an abstract classical system is by a manifold and two families of vector fields with appropriate structure, and it so happens that these these are equivalent to Lagrangian mechanics. This is not meant to be a mathematically derived *a priori* proposition, but rather one deduced from a claim about actual classical systems, that their kinematical vector fields always form affine spaces over vector spaces of fields that represent interactions. It is not inconceivable that it could have been the other way, that observation would have shown that the set of kinematical vector fields of classical systems had the structure of a real Lie algebra based on a symplectic structure. In this case, one presumes, an analogous argument would have shown that Hamiltonian mechanics was the more fundamental in this sense. In fact, to try to argue so will allow us to formulate and sketch the proof of the closest analogue in Hamiltonian mechanics to theorem 4.1.

Let us try, then, to construct for Hamiltonian mechanics the analogue of the arguments of §3, to see whether we can recover the Hamiltonian symplectic structure starting with the fundamental

²⁸It is of some interest to note that these constraints still represent the momental quantities as dynamical derivatives of a sort of the configurative quantities, analogous to the relation between velocital and configurative quantities of an abstract classical or Lagrangian system, for the Poisson bracket is, formally, a derivation on the algebra of scalar fields on a symplectic manifold. In the Hamiltonian case, though, the relation is symmetric: one may as well think of the kinematical constraints as saying that the configurative quantities are dynamical derivatives of the momental quantities!

elements analogous to those we used in the construction of configuration space for abstract classical systems. In the event, we can. I give the quick and dirty version.

Let there be given an abstract space of states, a family of kinematical vector fields on it and a set of kinematical constraints. (Behind the curtain, these are really the Hamiltonian structures—the family of kinematical vector fields in this case forms a vector space, for instance—but we do not yet know any of that.) We first determine the class of vector fields that represent possible interactions by requiring, as in the abstract classical case, that the addition of an interaction to a possible evolution yield another possible evolution, as seems plausible on physical grounds. It follows that the interaction vector fields for Hamiltonian mechanics are identical with the Hamiltonian vector fields, because the Hamiltonian vector fields themselves already (behind the scenes) form a vector space. By assumption, therefore, we can discover in the same way as we did for abstract classical systems, by physical probing, observation and generalization, that the family of interaction vector fields on the abstract space of states is identical to the family of kinematical vector fields. Since we demand that the addition of an interaction vector field to a kinematically possible one yield another kinematically possible one, we conclude that the kinematical vector fields form a vector space.

Now, that all canonical vector fields are Hamiltonian implies that every vector at a point of phase space is a member of some Hamiltonian vector field: you give me a point of phase space and a vector at it, and I can produce a Hamiltonian vector field that takes that vector as its value at that point. (In other words, there are no restrictions on what counts as good initial-data for the Hamiltonian initial-value problem.) By assumption, therefore, every vector at every point of the abstract space of states is a member of some kinematical vector field, which we can also discover by physical probing. We say in this case that the vector space of kinematical vector fields spans the tangent planes. (This does not imply that the fixed vector space of kinematical vector fields spans the vector space of all vector fields on the manifold; there are vector fields that cannot be written as a linear sum of fields in the vector space of canonical vector fields, *i.e.*, vector fields that are not the solution to Hamilton's equation for any Hamiltonian.) It follows that starting from any point we can reach any nearby point to first-order along some kinematically possible evolution, *i.e.*, good initial data for the equation of motion consists of a point of the space of states and any tangent vector at that point. (One can also think of this as saying, roughly speaking, that any allowable evolution can be perturbed to any other nearby one by changes of no order higher than the first in natural coordinates.) Thus, the equation of motion itself must be first-order; that it is linear follows from the fact that it must respect the kinematical constraints, which we know by assumption, because the constraints themselves respect the vector-space structure of the kinematical vector fields. We do not yet know, however, what types of entities the equation of motion is formulated in terms of—is it a map from vector fields to vector fields, or from scalar fields or collections of tensor fields to them? And so on. We know only that the equation of motion is encoded in some linear map from some family of entities to the Hamiltonian vector fields.

To address this question, we invoke the given kinematical constraints (which, recall, are really the Hamiltonian ones). By assumption, we know that there is a family of preferred coordinate systems the elements of each of which jointly satisfy the relations (5.3). We use this to define a map Ω from

pairs of canonical vector fields to scalar fields:

$$\begin{aligned}\Omega\left(\frac{\partial}{\partial \mathbf{q}_i}, \frac{\partial}{\partial \mathbf{q}_j}\right) &=_{\text{df}} \{q_i, q_j\} = 0 \\ \Omega\left(\frac{\partial}{\partial \mathbf{q}_i}, \frac{\partial}{\partial \mathbf{p}_j}\right) &=_{\text{df}} \{q_i, p_j\} = \delta_{ij} \\ \Omega\left(\frac{\partial}{\partial \mathbf{p}_i}, \frac{\partial}{\partial \mathbf{p}_j}\right) &=_{\text{df}} \{p_i, p_j\} = 0\end{aligned}\tag{5.4}$$

Restricting consideration to a single point, invoking the linearity and anti-symmetry of the Poisson bracket, and noting that the vectors $\frac{\partial}{\partial \mathbf{q}_i}$ and $\frac{\partial}{\partial \mathbf{p}_j}$ span the tangent-space at that point, we conclude that Ω is a bilinear, non-degenerate, anti-symmetric, linear map from pairs of vectors to scalars. In other words, it is a 2-index, anti-symmetric, invertible covariant tensor, otherwise known as a 2-form, and so we now write it emboldened, ‘ Ω ’, to honor our convention; one can think of it as an invertible, anti-symmetric matrix. Now extend it to a tensor field on a neighborhood of the fixed point by sweeping it along the flow-lines of the canonical vector fields; this guarantees that the canonical vector fields generate isometries of the 2-form. To see that the resulting tensor field is closed, and so a symplectic structure, it suffices to compute its components in the given canonical coordinate system, which turn out to be constant.

Another simple computation shows that the canonical coordinates and vector fields satisfy the equations

$$\Omega(\nabla q_i) = -\frac{\partial}{\partial \mathbf{p}_i}$$

and

$$\Omega(\nabla p_i) = \frac{\partial}{\partial \mathbf{q}_i}$$

One cannot yet think of these as instances of Hamilton’s equation, since the relations are so far confirmed only for canonical coordinates and vector fields. Linearity and the fact that the canonical vector fields span the space of all Hamiltonian vector fields, however, jointly imply that the vector space of all Hamiltonian vector fields is the space of solutions to equations of that form for arbitrary scalar fields. Thus $\Omega(\nabla \cdot)$ is the first-order, linear operator that encodes the equation of motion for all Hamiltonian vector fields, answering our question: the equation of motion takes a scalar field and returns a Hamiltonian vector field. Now it is Hamilton’s equation.

This proves a weak analogue to theorem 4.1.

Theorem 5.1 *Fix an even-dimensional, orientable manifold with a vector space of vector fields on it and a Poisson bracket structure. Then the Poisson bracket arises from a symplectic structure and the vector space includes all and only solutions to Hamilton’s equation formulated with it if and only if the vector space spans the tangent planes, and the manifold has a family of coordinate systems whose coordinate functions satisfy the relations (5.3) and whose associated coordinate vector fields leave the vector space invariant under the action of the Lie bracket.*

The theorem is weaker than 4.1 in so far as it imposes no topological form on phase space, as it should not. It is still strongly analogous to that theorem, however, in the sense that the kinematical

constraints—in this case, the Poisson brackets—are still seen to play a fundamental and constitutive role. They are precisely the relations among a system’s natural quantities that must obtain in order for Hamiltonian mechanics to be an appropriate framework for representing the system.

To sum up, we have seen that the family of Hamiltonian vector fields is not isomorphic to the family of an abstract classical system’s kinematical vector fields, and that Hamiltonian mechanics does not allow one to define an isomorphic analogue to the interaction vector fields of an abstract classical system. Because the abstract classical space of states is diffeomorphic to $T\mathcal{C}$, moreover, and that itself is diffeomorphic to $T^*\mathcal{C}$, though not canonically so, it follows that the abstract classical space of states is also diffeomorphic to $T^*\mathcal{C}$, though again not canonically so. Because one can do Hamiltonian mechanics on any symplectic manifold, however, there are Hamiltonian systems whose phase spaces are not diffeomorphic to the space of states of any abstract classical system, *viz.*, phase spaces that are not cotangent bundles. Finally, and most important, the kinematical constraints of the two frameworks do not encode isomorphic relations: what one needs to assume physically of a system in order to be able to apply the framework to appropriately model the system is different in each case. Thus, abstract classical systems are not Hamiltonian, and Hamiltonian ones are not abstract classical ones.

Before continuing the thread of my argument in the next section, I shall briefly discuss why my conclusions run directly counter to the those of [North \(2009\)](#), who claims that it is most natural to think of classical systems as fundamentally Hamiltonian in character. I have three primary problems with her account, which jointly explain the differences in our conclusions. First, she makes a crucial error in assuming that a Riemannian metric is necessary for the formulation of Lagrangian mechanics. As I explained in footnote 25, a Riemannian metric is neither necessary nor sufficient to formulate a Lagrangian representation of a system. Even when one does use a Riemannian metric in the formulation of a Lagrangian representation of a system, its physical significance is not at all clear, for different Lagrangians that yield the same Lagrangian vector field can induce non-isometric metrics. Second, she relies on a vague, unspecified notion of comparative simplicity of mathematical structure to argue that a symplectic structure, as used in Hamiltonian mechanics, is simpler than a Riemannian metric, as used in Lagrangian mechanics. I have no idea what to make of the idea of comparative simplicity in the vague sense she deploys it in. Depending on how one makes the idea precise—and I can think of a number of more or less natural ways to try to do so—one will get different answers to the question about symplectic structures and Riemannian metrics. For example, say one defines it thus: for two structures definable in the context of a third, that one is simpler than can be defined in a larger class of the third structures. This captures the idea that the structure places fewer constraints on its definitional context, and so is “intrinsically simpler”. By this criterion, Riemannian metrics are far simpler than symplectic structures, for one can define the former on any differential manifold but the latter only on a proper subclass of manifolds, those that are even-dimensional and orientable. Third, her interpretational claim depends on the assumption of a particular metaphysical view, structural realism of a sort. My arguments are based on empirical and mathematical grounds that require no particular metaphysical position either for their formal application in the arguments or for their physical significance in the arguments to be made clear

(though they are compatible with many different metaphysical positions, if that is the sort of thing one likes to take). I think this sort of neutrality is what one should aim for in an analysis of the conceptual structure of a physical theory.

6 How Lagrangian and Hamiltonian Mechanics Respectively Represent Classical Systems

The arguments and conclusions of the previous two sections raise (at least) four deep questions.

1. If Hamiltonian mechanics does not respect the kinematical constraints intrinsic to abstract classical systems, how can it provide adequate representations of classical systems (*e.g.*, the simple harmonic oscillator)?
2. Why does Lagrangian mechanics always respect the kinematical constraints of abstract classical systems?
3. Because we know the Hamiltonian and Lagrangian formulations to be related by the Legendre transform, what happens in the passage from Lagrangian to Hamiltonian mechanics that expunges respect for those constraints?
4. Is any structure in Hamiltonian mechanics isomorphic to any structure in Lagrangian mechanics in a physically significant sense?

I start with the first.

In order to apply Hamiltonian mechanics to model abstract classical systems, we have to impose the kinematical constraints of an abstract classical system more or less by hand. We do this without explicit remark in ordinary practice, by restricting attention to that small class of Hamiltonians that do in fact model abstract classical systems, *viz.*, those that satisfy $p_i = \lambda^j_i \dot{q}_j$, where λ^j_i is a linear, generally invertible transformation—*i.e.*, we demand that the p_i be linear functions of the \dot{q}_j , the kinematical relation between momentum and velocity we expect for classical systems. This condition is weak enough to represent the relation, for instance, between the rate of change of configuration and both linear momentum and angular momentum, and no weaker.

It follows from Hamilton's equation that for this relation to hold H must be a second-degree formula homogeneous in the p_i , *i.e.*, it must satisfy

$$\frac{\partial H}{\partial p_i} = \lambda_i^j p_j \tag{6.1}$$

which in turn implies

$$H = \alpha^{mn} p_m p_n + U(q_j) \tag{6.2}$$

where U and every α^{mn} is each an arbitrary function of the configurative quantities only. Indeed, we must in fact restrict ourselves to this case for the purposes of classical mechanics (as is always done without comment in textbooks), as one can see as follows. In general, that a Hamiltonian has

this restricted form implies that half of the relations in the coordinate formulation of Hamilton's equation, *viz.*,

$$\frac{\partial H}{\partial p^i} = \dot{q}_i \quad (6.3)$$

become physical tautologies, in the sense that they serve only to represent the kinematical constraints of the abstract classical system, $p_i = \lambda^j_i \dot{q}_i$. Thus, that a Hamiltonian has the form (6.2), where all the α^{mn} depend only on the q_i , is the necessary and sufficient condition for it to model an abstract classical system. Nonetheless, it is important to realize that, though the isomorphic relations in the context of abstract classical systems or Lagrangian mechanics are not predictions in the sense I explicated in §3, as they form part of the kinematical infrastructure of the theory and so must hold as necessary preconditions for the applicability of the theory in the first place, here, in Hamiltonian mechanics, they *are* predictions: it is not necessary that they hold in order for the theory to be meaningfully applied to model a system.

This can most easily be seen by the fact that Hamiltonians of any other form do not satisfy this relation. $H = \frac{1}{2}p_i p^i + \sum_j p_j$ is a funny one. It gives $\dot{p}_i = 0$ for the dynamics, like a free system, but $v_i =_{\text{af}} \dot{q}_i = p_i + 1$, which makes no physical sense: the constant number 1 does not have the dimensions of a velocity. (Another way to make the point: for “ $\dot{q}_i = p_i + 1$ ” to make physical sense, there would have to be a canonically distinguished system of units one had to use to express values of position and momentum; otherwise one would get different actual magnitudes for \dot{q}_i depending on whether one used cm/s or km/h.) $H = \frac{1}{2}p_i p^i + \frac{1}{6}(p_i p^i)^{3/2}$ gives even stranger behavior. It yields

$$v_i =_{\text{af}} \dot{q}_i = \frac{\partial H}{\partial p^i} = p_i + \frac{1}{4}p_j p^j$$

Solving this quadratic equation for p_i in terms of v_i gives

$$p_i = \frac{-1 \pm (1 + v_i)^{1/2}}{2}$$

Thus, there are *two* possible solutions for p_i , both of which are complex for values $v_i < -1$. It follows that the Lagrangian one gets from using this H to define the inverse-Legendre transform is also complex for those values of v_i , and so *a fortiori* does not yield a second-order vector field on the tangent bundle. It is impossible to make physical sense of any of this in the realm of the classical world.

Thus Hamiltonian mechanics represents abstract classical systems only in so far as we restrict ourselves to a sub-family of all the formally acceptable Hamiltonians by the *ad hoc* use of conditions foreign to Hamiltonian mechanics itself, for its structures do not provide the appropriate concepts and tools to formulate in their terms the required strictures that treat configurative quantities differently from momental, nor do they provide any natural justification for the restriction to Hamiltonians of the form (6.2). One way to see this is as follows. One might think or even hope that restricting attention to the “physical” Hamiltonians, those that satisfy the relation (6.2), might allow us to recover the structures of an abstract classical system in a natural way. This does not work, however: not only does that family of vector fields not form an affine space, it is not even invariantly defined, precisely because the relation (6.2) is not covariant under all canonical changes of coordinates.

Now, to address the second question listed at the beginning of the section, recall that a curve on the space of states satisfies the Euler-Lagrange equation if and only if it extremizes the variation of the standard action integral. More precisely, the traditional formulation of Lagrangian mechanics poses the following problem:

Given a function $L(q_i, \dot{q}_i)$ of the coordinates and their time-derivatives on configuration space \mathcal{C} , to find a family of curves $\{\gamma_\lambda\}_{\lambda \in \Lambda}$ for some indexing set Λ such that every curve in the family is an extremal, in the sense of the calculus of variations, of the action functional

$$\mathcal{A}[\gamma] = \int_\gamma L(\gamma(t), \dot{\gamma}(t)) dt$$

where $\dot{\gamma}(t)$ is the vector tangent to γ at parameter-value t .

We know already from §4 that any such extremal curve γ must be the canonical lift of a curve from configuration space to its tangent bundle; this is just another restatement of the Lagrangian constraints. What happens if we attempt to drop this restriction? Consider the following completely general variational problem:

Given the scalar field L on a manifold \mathcal{N} , to find a family of curves $\{\gamma_\lambda\}_{\lambda \in \Lambda}$ on \mathcal{N} for some indexing set Λ , such that through each point of \mathcal{N} exactly one curve passes, and each curve γ_λ in the family is an extremal of the action functional

$$\mathcal{S}[\gamma] = \int_\gamma L(\gamma(t)) dt$$

This problem has no non-trivial solution. If the integral must have an extremal value in *every* direction, and not just those directions transverse to those associated with something like lifts from a configuration space, then, roughly speaking, its derivative must be zero in every direction, which implies that the scalar field L must be a constant, and so *every* curve on the space of states is a solution.²⁹ This result is explained by the fact that, locally, all smooth, non-singular vector fields, and so all well behaved families of curves, look exactly alike. If you've seen one, you've seen 'em all.³⁰ This is why variational problems over unrestricted families of curves have no non-trivial solutions. That one is able to find non-trivial solutions to Lagrangian-type problems depends on the fact that one derives the Euler-Lagrange equation not by considering variations over arbitrary curves but only over curves that are canonical lifts on the tangent bundle, *i.e.*, curves along which half the coordinates are the dynamical derivatives of the other half—curves along which the kinematical constraints of an abstract classical system are respected.³¹

²⁹It is not difficult to make this argument precise and rigorous.

³⁰Somewhat more precisely: given two n -dimensional manifolds \mathcal{M} and \mathcal{N} , points $p \in \mathcal{M}$ and $q \in \mathcal{N}$, and smooth vector fields ξ and η that are, respectively, not singular at p and q , then one can always find convex normal neighborhoods O_p and O_q respectively containing p and q and a smooth mapping $\phi : \mathcal{M} \rightarrow \mathcal{N}$ such that $\phi(p) = q$, ϕ is a diffeomorphism of O_p to O_q , and $\phi^*(\xi)|_{\phi(p)} = \eta|_q$.

³¹In the usual derivation of the Euler-Lagrange equations (*e.g.*, [Rosenberg 1977](#), chapter 9), this restriction allows the switching of the order of differentiation and variation of certain terms that in turn allows the crucial integration by parts; the nature and origin of the requirement is masked in that case by the traditional presentation in terms of an arbitrary coordinate system and the lack of recognition that the Lagrangian is a scalar field on $T\mathcal{C}$ not on \mathcal{C} .

Thus, it is built into Lagrangian mechanics from the start, by necessity, that the kinematical constraints of abstract classical systems be respected—one cannot even formulate the theory without it—that is to say, one could not even derive the Euler-Lagrange equation from a variational principle. This is not to say that I think one must conceive of Lagrangian mechanics as based on a variational principle. Indeed, in my preferred presentation of Lagrangian mechanics (adumbrated in §4, presented more fully in [Curriel 2012](#)), one need make no reference, not even implicitly, to a variational principle. One simply imposes the Euler-Lagrange equation as the generic form of the equation of motion and demands that one’s space of states have the geometrical resources for the formulation of the equation. It then turns out that one has those resources on and only on a tangent bundle (the content of theorem 4.1). The point I am making here is only that, in so far as demanding the Euler-Lagrange equation is logically equivalent to demanding that a certain variational problem be well posed, one can get some insight into the need for Lagrangian systems to satisfy the kinematical constraints of an abstract classical system—they are necessary conditions for that variational problem to be well posed.

Consideration of the variational problem also sheds light on why the Lagrangian must be quadratic (at least) in the velocities for there to be a unique solution to the Euler-Lagrange equation, *i.e.*, why it makes physical sense to demand that the Lagrangian be regular as defined by the relation (4.6). Again, the variational problem will not be well posed unless one can take non-trivial derivatives of the Lagrangian up to second-order in the velocity, since the crucial integration by parts that yields the Euler-Lagrange equation demands this, on pain of giving, *e.g.*, the tautology $0 = 0$ (for, say, the Lagrangian $L = k^i v_i$ for a constant vector \mathbf{k}) or the contradiction $1 = 0$ (for, say, the Lagrangian $L = k^i q_i$, for constant \mathbf{k}). This again stands in contradistinction to the case of Hamiltonian mechanics, in which one must impose the quadratic form of the Hamiltonian as an *ad hoc* condition, precisely because *every* smooth scalar field yields a consistent, unique solution to Hamilton’s equation, including $H = k^i q_i$ for constant \mathbf{k} .

We now address the third question, what happens in the passage from a Lagrangian to a Hamiltonian representation of an abstract classical system by way of the Legendre transform: why does the structure of an abstract classical system (in the sense of §§2–3) not get preserved? Fix a natural coordinate system (q_i, v_j) on the tangent bundle, and let (q_i, p_j) be the natural coordinate system on the cotangent bundle based on the same configuration coordinates (q_i) . Then for a given Lagrangian L , the action of the Legendre transform—the natural mapping that takes a Lagrangian model of a system to a Hamiltonian one—is fixed in these coordinates by the condition that $v_i \mapsto \frac{\partial L}{\partial v_i} =_{\text{def}} p_i$, providing a map from $T\mathcal{M}$ to $T^*\mathcal{M}$. (Thus the condition that a Lagrangian be regular, equation (4.6), also guarantees that the p_i form good coordinates on the cotangent bundle.)

In order to see why the affine-space structure of the Lagrangian vector fields does not survive the transform, we need to determine what an arbitrary second-order vector field maps to. What happens, for example, to physically pathological second-order vector fields when they are mapped to the cotangent bundle? Consider the field

$$\xi = v_i \left(\frac{\partial}{\partial q_i} \right) + v_j \left(\frac{\partial}{\partial v_j} \right) \tag{6.4}$$

This represents a system whose acceleration increases in proportion to its velocity, which is to say that its velocity exponentially increases, and so it will be highly unstable and shoot off to infinity in a finite amount of time at the slightest provocation.³² Mapping the vector field (6.4) to $T^*\mathcal{C}$ using the Legendre transform Λ_L defined by $L = \frac{1}{2}v_i v^i$, a manifestly physical Lagrangian,³³ we get

$$\hat{\xi} = \Lambda_L[\xi] = p_i \frac{\partial}{\partial q_i} + p_j \frac{\partial}{\partial p_j}$$

It is easy to see that this cannot be a Hamiltonian vector field on physical grounds: because $\dot{p}_j = p_j$, this system would not conserve energy. (It would again represent a system that goes hurtling off to infinity in a finite time with exponentially increasing velocity.) Thus, the Legendre transform does not map all Lagrangian vector fields to Hamiltonian ones.³⁴

What happens to Hamiltonians that don't respect the Lagrangian kinematical constraints when they get sent back to the Lagrangian formulation via the inverse-Legendre transform? Consider the Hamiltonian $H = p$. This yields a well defined Hamiltonian problem, with the Hamiltonian vector field

$$\begin{aligned}\dot{q} &= 1 \\ \dot{p} &= 0\end{aligned}$$

This seems to represent a free particle moving with velocity 1. (We ignore for the moment that the proposition ‘the velocity equals 1’ makes no physical sense without a fixation of units.) If you try to perform a reverse Legendre transform to put it into Lagrangian form, then you get $\dot{q} = 1$ and $L =_{\text{af}} vp - H = \dot{q} - 1$, which does not yield a well-set Lagrangian problem—more precisely, it yields the identity $0 = 0$ when you plug L into the Euler-Lagrange equation. Thus, not all Hamiltonian vector-fields map to second-order vector fields under the inverse-Legendre transform.

The Legendre transform does not respect the kinematical constraints of Lagrangian mechanics because it, in effect, wipes out any notion of verticality—the idea that the interaction vector fields (the vertical vector fields on the tangent bundle) are different from the kinematical vector fields, and yet define their algebraic structure—in so far as the idea of verticality is defined by the difference of two second-order vector fields on the tangent bundle, in virtue of the affine-space structure of the second-order vector fields based on the vector space of vertical vector fields. In wiping out verticality, it also wipes out the kinematical link between change of position and the momental quantity used to formulate the equation of motion, and the associated link between the kinematical constraints and the form of allowable interactions for abstract classical systems. One can make a stronger statement: any physically significant transformation of a Lagrangian representation of a system into

³²This is an example of a second-order vector field that is not the solution to any Lagrangian for the homogeneous Euler-Lagrange equation; one must use the inhomogeneous form, by including (non-conservative) general forces, to find a Lagrangian that has this as the solution.

³³And, in fact, this L is a Lagrangian that yields the vector field (6.4) as a solution with the appropriate generalized forces, so it is in that sense an appropriate Lagrangian to use to define the Legendre transform.

³⁴One can show this as well using a simple mathematical argument: there is no scalar field $H(q, p)$ such that $-\frac{\partial H}{\partial q} = p$ and $\frac{\partial H}{\partial p} = p$.

a Hamiltonian one *must* wipe out verticality: Hamiltonian mechanics cannot be formulated with it, as Lagrangian mechanics cannot be formulated without it.

This leaves us the fourth and final question I posed on page 31. In fact, there is a promising start to the address of the suggestion: although it is next to impossible to see by looking only at the Euler-Lagrange equation, when posed with a regular Lagrangian it does in fact contain a closed, invertible 2-form—a symplectic structure, in seeming analogy with Hamiltonian mechanics—as part of the equation’s construction, as I discussed immediately after condition (4.6). It is not a fixed symplectic structure as in Hamiltonian mechanics, however, but rather itself depends on the Lagrangian, and so it differs in different instances of the Euler-Lagrange equation. This stands in opposition to the case in Hamiltonian mechanics where the symplectic structure is independent of the Hamiltonian. Thus, although there is a formal isomorphism of the two, they have different physical significances. In Hamiltonian mechanics, the fact that the symplectic structure is fixed once and for all enforces the satisfaction of the kinematical constraints (5.3), and it encodes the existence and uniqueness of solutions to the equations of motion and the conservation of energy, among many other things; all of these features are independent of the individual Hamiltonian used in the equation. In Lagrangian mechanics, by contrast, it is the Lagrangian itself that encodes existence and uniqueness of solutions (which will obtain when and only when the Lagrangian is regular), and there is not even a guarantee that energy is conserved, as one can, *e.g.*, represent dissipative systems in the Lagrangian framework; and the Lagrangian symplectic structure, when it even exists, has no logical relation to the satisfaction of the Lagrangian kinematical constraints, the relations (4.3), which are themselves encoded in a wholly separate geometrical structure (the almost-tangent structure). Indeed, the kinematical constraints are always satisfied even when the Lagrangian is not regular, and so even when there is not even a symplectic structure available, period.

So much for possible isomorphism of kinematical structures. I now briefly address the question whether Lagrangian and Hamiltonian mechanics share physically significant structure in their dynamics. In fact, the question has a simple answer. The dynamical constituents of models of individual systems in Hamiltonian Mechanics, such as the equation of motion for a simple harmonic oscillator, will not be isomorphic to any dynamical constituents of the model for the same system in Lagrangian mechanics, for all the reasons already given. That the global kinematics possess no common structures necessitates that individual models do not either, for the global kinematical structures provides the scaffolding for the construction of the equations of motion. Indeed, that is the physical significance of the difference between theorems 4.1 and 5.1.

One reflection of this, for instance, is the fact that Hamilton’s equation must be first-order and the Euler-Lagrange equation second-order. The former proposition was derived in the proof of theorem 5.1. The latter is suggested by the following argument (which can be made rigorous—see Curiel 2012). For an arbitrary vector field ζ on the tangent bundle, there always exist three second-order vector fields ξ , η and θ such that

$$\zeta = \theta - [\xi, \eta] \tag{6.5}$$

We already expect that a kinematically possible evolution—more precisely, a finite stretch of a curve representing one—connects any two points on the space of states. (If not, then we could divide the space of states into pieces not reachable from each other by any dynamical evolution, and that would thus *prima facie* represent “different systems”—this is the physical significance of the demand that the space of states be pathwise-connected in the sense of topology.) Then equation (6.5) states, roughly speaking, that we can get from any point in the abstract classical space of states to any nearby point “to order no higher than second” along the kinematically possible evolutions. Not every two nearby points in the tangent bundle, *i.e.*, the space of states, are connected by a kinematically possible vector field, because nearby points on the tangent bundle can be separated in directions other than those picked out by the second-order vector fields. This reflects the fact that only evolutions that respect the constraints are kinematically possible, and not all curves on the tangent bundle respect the constraints. On the Hamiltonian space of states, however, every point is connected to every nearby point by a kinematically possible vector field, reflecting the fact that there are no functional constraints among the configurative and momental quantities that reduce the dimensionality of the family of their compossible evolutions. Thus, the dynamical structures of the theories, as encoded in the equations of motion, cannot be isomorphic in any physically significant sense either.

Let’s make this last point concrete and more precise by means of a simple example, the simple harmonic oscillator. For simplicity, I assume the system is one-dimensional (*i.e.*, it can move only in one dimension and so has two degrees of freedom in my sense) and has its mass equal to 1 and its coefficient of elasticity equal to 1/2. Then the Hamiltonian is

$$H = \frac{1}{2}p^2 + \frac{1}{2}q^2$$

and Hamilton’s equations are

$$\begin{aligned}\dot{q} &= \frac{\partial H}{\partial p} = p \\ \dot{p} &= -\frac{\partial H}{\partial q} = -q\end{aligned}\tag{6.6}$$

which has as its phase-portrait (*i.e.*, the integral curves of the associated Hamiltonian vector field) the expected circle on phase space as represented by a Cartesian plane. If we use H to define the inverse-Legendre transform and so pass to the Lagrangian formulation, we get as expected

$$v =_{\text{af}} \frac{\partial H}{\partial p} = p$$

for the velocity, yielding the expected Lagrangian

$$\begin{aligned}L &=_{\text{af}} pv - H \\ &= v^2 - \frac{1}{2}p^2 - \frac{1}{2}q^2 \\ &= \frac{1}{2}v^2 - \frac{1}{2}q^2\end{aligned}$$

This is a regular Lagrangian, having as its solution to the Euler-Lagrange equation the second-order vector field

$$\begin{aligned}\dot{q} &= v \\ \dot{v} &= -q\end{aligned}\tag{6.7}$$

whose integral curves are essentially the same circles as in the Hamiltonian solution, only now on the tangent bundle (also represented as the Cartesian plane); in other words, the integral curves representing the evolution of the system in the two frameworks are isomorphic to each other.

At first sight, the relations (6.6) and (6.7) appear in perfect agreement with each other. Indeed, the perfection of the apparent agreement makes it difficult to see how the general claims of the previous paragraphs could be true, that the dynamics of the Lagrangian and the Hamiltonian frameworks do not have structures that are isomorphic in a physically significant way. Still, the fact that the phase portraits in both frameworks are circles does not falsify the claim. All that fact says is that both frameworks yield brute descriptions of the motion of the simple harmonic oscillator that quantitatively agree. My claim, however, demands more than that. For it to hold, any structure in the model of one framework that is mathematically isomorphic to a structure in the model of the other must yet have a different physical significance. (We have already seen an example of this at the global level: the variable symplectic structure that lies hidden under the Euler-Lagrange equation has a physical significance different than the canonical one that appears in Hamiltonian mechanics, at least in so far as the latter but not the former enforces the kinematical constraints and encodes conservation of energy.) And this is the case at the level of the individual model as well.

To see this, let us look in detail at the formally isomorphic relations, (6.7) and (6.6). As I argued in §3, the first relation in the Lagrangian pair (6.7), $\dot{q} = v$, is not a prediction the theory makes, but rather a stipulation, a proposition necessary for the appropriate application of the theory to model the system in the first place.³⁵ One cannot meaningfully apply the theory while having that proposition fail, for the framework requires the truth of that statement for its own logical consistency. In that sense, it shares much of the flavor, status and behavior of an analytic or *L*-statement in a Carnapian framework, one true, roughly speaking, in virtue solely of the meaning of its terms. It has still this difference from such statements, though: it has non-trivial semantic content. It makes a non-trivial claim about the behavior of physical systems. It is a proposition that can fail to apply to a real physical system, in the sense that the temporal rate of change of a system's configuration can fail to equal the system's kinematical velocity. This happens in a variety of theories, *e.g.*, ordinary quantum mechanics, Bohmian quantum mechanics, some theories of relativistic fluid mechanics,³⁶ *et al.*

³⁵Recall my characterization of a 'prediction' from §3 (p. 12), slightly modified so as to fit the present context:

[S]omething that a theory, while meaningfully and appropriately modeling a given system, can still get wrong. Lagrangian mechanics, then, does not predict that the kinematical velocity of a Newtonian body equal the temporal rate of change of its position; rather it requires it as a precondition for its own applicability. It can't "get it wrong".

³⁶See, *e.g.*, Landau and Lifschitz (1975) and Earman (1978).

The analogous relation in the Hamiltonian framework, the first of the pair (6.6), has none of the same flavor, status or behavior. It is not a proposition that must hold for one to be able meaningfully to apply Hamiltonian mechanics to represent a system. Indeed, it is not even a proposition that always holds in Hamiltonian mechanics, period. It holds only for those systems whose Hamiltonians are of the form (6.2). A system whose Hamiltonian is of any other form will not satisfy it. It therefore is a true prediction of the Hamiltonian framework. (Both of the second relations in (6.6) and (6.7) are also true predictions in my sense, and so have more or less the same physical significance.)

Thus, even when Lagrangian and Hamiltonian mechanics provide individual models of the same physical system, no structure in the individual model of the one is isomorphic to any in the other in a way that has real, non-trivial physical significance (over and above the fact that the two models yield the same brute quantitative description of the behavior of the system). The only way to deny this conclusion is to claim that sameness of solutions to equations of motion by itself—mere sameness in brute description of motion—ensures sameness of physical significance, but that is nothing more than the most naive form of verificationism.

7 The Conceptual Structure of Classical Mechanics

To summarize the primary conclusions of the paper's technical arguments:

1. the global structures of an abstract representation of a classical Newtonian system, as characterized in §§2–3, are necessary and sufficient for the construction of a Lagrangian formulation for the system
2. those structures do not permit one to recover a Hamiltonian formulation for it, nor, conversely, do the structures of Hamiltonian mechanics allow one to recover the structures of a Newtonian system
3. none of the fundamental, global kinematical structures of Lagrangian mechanics is isomorphic to any in Hamiltonian mechanics
4. in consequence, the global dynamical structures of the two are not isomorphic in any way that has physical significance
5. in further consequence, no individual model in one shares physically significant, fundamental, non-trivial structure with any in the other, even when the two are models of exactly the same classical system
6. none of these arguments could have been made by consideration of individual models of physical systems only; they rely crucially on relations among individual models, *i.e.*, on global relations defined on the entire family of a framework's models

In sum, the conceptual structure of classical Newtonian mechanics, both at the level of global structure and at the level of individual models, is that evinced by the Lagrangian framework. This

lesson could not have been learned from looking only to see whether there were theoretical structures isomorphic to empirical structures, *simpliciter*, as is often done in structuralist approaches to the understanding of physical theory. One must also look to see what physical significance accrues to the structure in its theoretical and in its empirical domain—what it requires, what status it has, what roles it plays, what consequences follow from it, both globally and individually—to verify that the formal isomorphism is not empty of true physical content.

To conclude the paper, I shall analyze a little further the light I think this paper’s arguments shed on structuralist approaches to the understanding of physical theory. Roughly speaking, by ‘structuralist approach to the understanding of physical theory’ I mean one that posits as the foundation for the representational content of a theory a relation of structural similarity between the formal apparatus of the theory on the one hand and empirical, phenomenal structure on the other.³⁷ That similarity relation may be one of isomorphism, partial isomorphism, homomorphism, or anything of the sort; for the purposes of my argument, the specific form is irrelevant, so I shall just refer to such a thing as a ‘morphism’. It also does not matter, for the purposes of my argument, whether the empirical structure is thought to consist of a data model or of real stuff in the world or something else of the sort, so I will use the term ‘phenomenal structure’ to refer promiscuously to any such thing used in structuralist accounts. The important fact for our purposes is that, for a structuralist, the morphism by itself characterizes the representational content of theoretical structure. It follows that whatever semantic content accrues to theoretical structure must be grounded on that morphism. Thus, where there is no morphism, there is no representational content. The example of the way that Hamiltonian mechanics represents classical Newtonian systems, however, shows that this position is untenable, for, as the argument of §5 showed, there is no morphism of any kind between the structures one uses to formulate Hamiltonian representations and the phenomenal structures of such systems. But Hamiltonian mechanics does in fact provide good, useful models of such systems. So what is going on?

To try to address that question, it will help to consider a pair of related questions: whether sameness of meaning in two representations entails sameness of structure in them, and whether sameness of structure in two representations entails sameness of meaning. Now, when two frameworks provide good representations of the same system, it is natural to think that, in a strong sense, the two representations have the same empirical meaning. For the structuralist, moreover, it will follow that there must be an appropriate morphism between the structures of the two frameworks, because each is on its own, *ex hypothesi*, appropriately morphic to the same structure, *viz.*, the phenomenal one the system evinces. For the structuralist, then, it is natural to demand that sameness of meaning of representations entails sameness of structure. But this paper’s example shows that cannot be correct. The Lagrangian and the Hamiltonian models of the simple harmonic oscillator in some sense surely “mean the same thing”, but there is no meaningful sameness of structure. The two models agree on all empirical propositions both theories can formulate about it but the formal entities and relations that inform the semantic content of the propositions are heteromorphic. The two models,

³⁷See [Brading and Landry \(2004\)](#) for a particularly clear and thorough exposition of positions of this sort.

for example, agree on the following empirical proposition **P**: “the position and velocity of the simple harmonic oscillator jointly satisfy the kinematical constraints essential to classical Newtonian systems, such as that its velocity is the temporal derivative of its position”. Strictly speaking, however, one cannot even formulate the proposition **P** in the framework of Hamiltonian mechanics, because the framework does not differentiate between configuration-like and velocity-like quantities. The constraints for classical Newtonian systems, however, treat configuration and velocity asymmetrically. Thus, one cannot define a predicate in Hamiltonian terms that represents the constraints. There is still, however, an obvious and important sense in which **P** is not only meaningful in but true of the Hamiltonian model of the simple harmonic oscillator: there is a unique, true proposition **Q** (the relations 6.6) in Hamiltonian mechanics whose translation into Lagrangian terms by the inverse-Legendre transform is (the Lagrangian formulation of) **P** (the relations 6.7). Of course, this translation works, *i.e.*, is physically significant, only when the Hamiltonian is restricted by *ad hoc* measures to take the form (6.2). There is no morphic structure between the two frameworks that underwrites the sameness in empirical meaning: the inverse-Legendre transform does not map relevant structures in the one to morphic structures in the other. Sameness of meaning, then, does not imply sameness of structure.

Neither is the converse slogan true. For the structuralist, if two theoretical representations appropriately model the same system, then it is natural to demand that the structures that are morphic between those representations must mean the same thing, must play the same roles. Otherwise it is difficult to see how morphism of theoretical to phenomenal structure could ground representational content. Again, our example shows this cannot be correct in general. The only structure formally morphic in the Lagrangian and Hamiltonian frameworks—a symplectic structure, for which the Legendre transform and its inverse do in fact act as isomorphisms—has different physical significance, and so different representational content, in each framework. The canonical symplectic structure in Hamiltonian mechanics plays (at least) three representational roles: it encodes all the kinematical constraints; it ensures existence and uniqueness of solutions to the equation of motion; and it ensures that energy (the Hamiltonian) is conserved during the course of possible evolutions. It plays those roles in every model, moreover, independently of the dynamics of any particular model, in the sense that it is independent of the Hamiltonian, and is rather fixed once and for all, the same for all Hamiltonian models. Thus the Hamiltonian symplectic structure encodes all and only the theory’s kinematics for all its models; the Hamiltonian function of a particular model encodes all and only the dynamics of that model and that model alone. In Lagrangian mechanics, to the contrary, only some models have a symplectic structure, those whose associated Lagrangians satisfy a particular condition, relation (4.6), and in those models the symplectic structure’s only function is to ensure existence and uniqueness of solutions to the equation of motion. It encodes none of the kinematical constraints, and it does not ensure conservation of energy. Its role in guaranteeing existence and uniqueness, moreover, depends on the dynamics of the particular model, in so far as the Lagrangian symplectic structure of a particular model is a function of the model’s Lagrangian itself (if the model’s Lagrangian satisfies the relevant condition in the first place). In Lagrangian mechanics, therefore, much of the kinematics is encoded in the algebraic structure of the space of

possible evolutions, but some is encoded in the Lagrangian for each model as well; the dynamics of an individual model is still encoded entirely in the Lagrangian, but now it is not alone. Thus, sameness of structure does not entail sameness of meaning.

Still, there is a strong intuition that, just because the two render the same gross representations of the evolution of a large class of systems, which is to say, Lagrangian and Hamiltonian models yield the same form for the predicted evolution (as for the simple harmonic oscillator), it must follow that the two have some physically significant, morphic structure. Agreement in prediction, however, simply does not imply the existence of morphic, physically significant structure. Even when there are morphic, physically significant structures, moreover, as I emphasized above, they *need not have the same semantic content*, because they may play different roles in the representation of the physical system. The only way to deny this conclusion is to claim that sameness of solutions to equations of motion by itself—mere sameness in brute prediction—ensures sameness of semantical content, but that is nothing more than the most naive form of verificationism. Mere existence of morphisms between structures, then, cannot ground a semantics of physical theory: just because two theories represent the same systems and have morphic structures it does not follow that the empirical meaning of the morphic structures are the same in the two theories—the physical roles the structures play in one theory, and so their semantic content, may be distributed differently among the morphic structures in the other theory, so to speak, or even among entirely different structures in no way morphic to any in the other.

The discussion also shows why Tarskian (or Beth, as the case may be) semantics is not adequate for physical theories. The argument has shown that one can discover the profound differences between the semantics of Lagrangian and Hamiltonian theories only by examination of their global structures, not by the study of any individual model. The semantics of individual models on their own do not suffice for the comprehension of the semantics of their fundamental building blocks such as the preferred quantities by means of which states are to be individuated and identified and which in their internal structure encode the kinematical constraints appropriate for each framework. Relations that hold jointly among all the theory's models must be taken account of—one gets the affine structure of a classical Newtonian system's kinematical vector fields, for instance, *only* by treating the collection of models as a unified space with its own global structure. In particular, one cannot define the space of interaction vector fields when one restricts attention to the collection of models in isolation from any global structure imposed on it in the form of relations among the models, precisely because an interaction vector field is defined as the difference of two kinematical vector fields, *viz.*, the difference of vector fields in two separate models. Each model on its own, in isolation from any relation to the free kinematical vector field, cannot support the idea of an interaction vector field. In this sense, a theory's bare collection of models on its own does not suffice for the recovery of the full semantics of each individual model, much more for the semantics of the theory as a whole. So far as a proper accounting of semantics goes, a physical theory must be more than its collection of models—it must include as well relations among the models. This violates the principles of Tarskian semantics. It is perhaps possible to extend or modify Tarskian semantics so as to take account of relations among models, but that is beside the point. Traditional Tarskian

semantics as in fact deployed by the semantic view of theories does not provide a rich enough set of tools to characterize all the semantic content of a physical theory.

The arguments of the paper notwithstanding, I feel there is something profoundly right about a structuralist point of view—the world as we have known it *does* manifest in its parts clear and beautiful structures that find elegant and verisimilar recapitulation in physical theories. The recovery of Lagrangian mechanics from simple physical assumptions about classical systems provides as good an example of this as one could want. The way that structuralist points of view have generally been expressed up to now, however, with the naively, exuberantly optimistic idea that structure in theories always stands in unambiguous, univocal relation with structure in the world, does not work. In order to represent the structures in the world, the structures in theories must stand in some definite relation to them. That relation, however, cannot in general be mere morphism; it must be something more complex and subtle.

References

- Arnold, V. (1978). *Mathematical Methods of Classical Physics*. Number 60 in Graduate Texts in Mathematics. New York: Springer-Verlag. English translation of the original Russian edition, *Matematicheskie metody klassicheskoi mekhaniki*, Moscow: Nauka, 1974, trans. K. Vogtmann and A. Weinstein.
- Artin, E. (1957). *Geometric Algebra*. Number 3 in Interscience Tracts in Pure and Applied Mathematics. New York: John Wiley & Sons, Inc.
- Belot, G. (2007). The representation of time and change in mechanics. See [Butterfield and Earman \(2007\)](#), pp. 133–228.
- Brading, K. and E. Landry (2004). A minimal construal of scientific structuralism. Delivered as a talk as part of the Symposium “The Semantic View of Theories, Scientific Structuralism and Structural Realism”, at the 20th biennial meeting of the Philosophy of Science Association, 2004, Vancouver, BC. It is available at http://philsci-archive.pitt.edu/archive/00002181/01/Minimal_Structuralism.pdf.
- Butterfield, J. (2003). Some aspects of modality in analytical mechanics. <http://philsci-archive.pitt.edu/archive/00001192/01/ButterfieldForFormal.pdf>.
- Butterfield, J. (2004). Between laws and models: Some philosophical morals of Lagrangian mechanics. <http://philsci-archive.pitt.edu/archive/00001937/01/BetLMLag.pdf>.
- Butterfield, J. (2007). On symplectic reduction in classical mechanics. See [Butterfield and Earman \(2007\)](#), pp. 1–132.
- Butterfield, J. and J. Earman (Eds.) (2007). *Handbook of Philosophy of Physics, Part A*. Dordrecht: North Holland.
- Curiel, E. (2012). The geometry of the Euler-Lagrange equation. Unpublished manuscript.

- de León, M. and P. Rodrigues (1989). *Methods of Differential Geometry in Analytical Mechanics*. Number 158 in Mathematics Studies. Amsterdam: North Holland.
- Earman, J. (1978). Combining statistical-thermodynamics and relativity theory: Methodological and foundations problems. *Philosophy of Science* 2(Proceedings), 157–85.
- Earman, J. (1986). *A Primer on Determinism*. Dordrecht: D. Reidel Publishing Co.
- Earman, J. (2007). Aspects of determinism in modern physics. In J. Butterfield and J. Earman (Eds.), *Handbook of Philosophy of Physics, Part B*, pp. 1369–1434. Dordrecht: North Holland.
- French, S. (2009). Metaphysical underdetermination: Why worry? *Synthese* 171(2), 1–17. doi:10.1007/s11229-009-9598-5.
- Goldstein, H. (1980). *Classical Mechanics* (Second ed.). Reading, MA: Addison-Wesley Publishing Co.
- Lanczos, C. (1970). *The Variational Principles of Mechanics* (Fourth ed.). New York: Dover Publications, Inc. A 1986 republication by Dover of the original 1970 edition by the University of Toronto Press.
- Landau, L. and E. Lifschitz (1975). *Fluid Mechanics* (Second ed.). Oxford: Pergamon Press. An expanded, revised edition of the original 1959 edition. Translated from the Russian by J. Sykes and W. Reid.
- Malament, D. (2008, December). Norton’s slippery slope. *Philosophy of Science* 75(5), 799–816. Proceedings of the 2006 Biennial Meeting of the Philosophy of Science Association Part II: Symposia Papers.
- Newton, I. (1726). *Philosophiæ Naturalis Principia Mathematica* (Third ed.). Amherst, NY: Prometheus Press. The translation by A. Motte of the third edition (1726), originally produced in 1729, and reprinted by Prometheus Press in 1995. The first edition of the *Principia* was published in 1686, the second in 1713.
- North, J. (2009). The “structure” of physics. *The Journal of Philosophy* 106(2), 57–88.
- Norton, J. (2008, December). The dome: An unexpectedly simple failure of determinism. *Philosophy of Science* 75(5), 786–798. Proceedings of the 2006 Biennial Meeting of the Philosophy of Science Association Part II: Symposia Papers.
- Pars, L. (1965). *A Treatise on Analytical Dynamics*. New York: John Wiley & Sons, Inc.
- Rosenberg, R. (1977). *Analytical Dynamics of Discrete Systems*. New York: Plenum Press.
- Russell, B. (1927). *The Analysis of Matter*. New York: Dover Press. A 1954 reprint of the original 1927 edition by Kegan Paul, Trench, Trübner & Co. of London.
- Stein, H. (1972). On the conceptual structure of quantum mechanics. In R. Colodny (Ed.), *Paradigms and Paradoxes: The Philosophical Challenge of the Quantum Domain*, Volume 5 of *University of Pittsburgh Series in the Philosophy of Science*, pp. 367–438. Pittsburgh, PA: University of Pittsburgh Press.

- Wallace, D. (2003). Time-dependent symmetries: The link between gauge symmetries and indeterminism. In K. Brading and E. Castellani (Eds.), *Symmetries in Physics: Philosophical Reflections*, Chapter 9, pp. 163–173. Cambridge: Cambridge University Press.
- Whitehead, A. (1948). *Science and the Modern World*. New York: The New American Library. Originally delivered as the Lowell Lectures, at the Lowell Institute, Boston, 1925.
- Whittaker, E. (1947). *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies* (Fourth ed.). Cambridge: Cambridge University Press. A 1993 reprint of the fourth edition, Cambridge University Press, 1947.
- Yano, K. and S. Ishihara (1973). *Tangent and Cotangent Bundles: Differential Geometry*. Number 16 in Pure and Applied Mathematics. New York: Marcel Dekker.