Classical Mechanics Is Lagrangian; It Is Not Hamiltonian; The Semantics of Physical Theory Is Not Semantical

Erik Curiel†

November 01, 2008

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 prove 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 provides a decisive counter-example to the semantical view of physical theories, and one, moreover, that shows its crucial deficiency: a theory must be, or at least be founded on, more than its collection of models (in the sense of Tarski), for a complete semantics requires that one take account of global structures defined by relations among the individual models. The example also shows why naively structural accounts of theory cannot work: simple isomorphism of theoretical and empirical structures is not rich enough a relation to ground a semantics.

†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 to Chris Pincock for stimulating and edifying conversations about the semantical view of physical theories, in which he set straight a few misunderstandings I had. Email: erik@strangebeautiful.com

Contents

1 Introduction 2

2 Classical Systems 3

3 The Possible Interactions of a Classical System and the Structure of Its Space of States 6

4 Classical Systems Are Lagrangian 13

5 Classical Systems Are Not Hamiltonian 18

6 How Lagrangian and Hamiltonian Mechanics Respectively Represent Classical Systems 24

7 A Counter-Example to the Semantical View of Physical Theory 32

1 Introduction

The semantical view of physical theories relies on models to provide the bridge between a theory's formal structures and the structure of phenomena in the world. In weak versions of the view, a theory is characterized by its models; in stronger views, the theory simply is the collection of its models.[1] In either case, however, it is a tenet of the view that a theory's model adequately represents a physical system only when the model and the system possess relevant, isomorphic structures. It follows that, at a minimum, when two theories provide sound models for the same physical systems, then the models of those two theories should share the same structure in some important sense, including at a minimum that the shared structure have essentially the same physically significant semantic content in the respective representations of the system. The case of Lagrangian and Hamiltonian mechanics shows that the tenet cannot be correct. Both yield sound models of the same systems, but neither possesses any structure isomorphic in a physically significant way to any in the other.

I go on at some length in §§2–3 chracterizing 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, empirically grounded, formally precise structure accruing to an entire class of physical systems starting from the simplest of assumptions. It is a sterling example of what the semantical view of physical theories lives for. Thus, the strength of the putative counter-example to that view 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.

[1] See, *e.g.*, Brading and Landry (2004) for a concise and elegant summary of the semantical view and a comparison with the syntactical view.

Because the primary philosophical claims of the paper are surprising, moreover, and will I am sure be controversial, I give in §§4–6 the exposition of Lagrangian and Hamiltonian mechanics and of their relations in some detail, to substantiate the claims. I do this as well because, with few exceptions, philosophers of physics have tended to give short shrift to the philosophical problems that lie at the foundation of classical mechanics.[2] I think the neglect is unjustified. Just as the mathematical theories of classical mechanics still today provide frameworks and fodder for deep investigations in several fields of pure mathematics, so they can in philosophy, both with regard to traditional questions philosophers tend to examine in the context of quantum mechanics or relativity theory, and with regard to questions peculiar to classical mechanics itself. I hope that this paper's attempt to clarify the conceptual structure of Lagrangian and Hamiltonian mechanics makes that case by example.

For those who do not want to work through all the technical details, I give at the beginning of each of §§4–6 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 about the semantics of physical theories. §7 contains the gathering together and summation of that argument in non-technical terms.

Finally, I must remark that some of what I put down when I use or discuss technical machinery will be, strictly speaking, incorrect or meaningless, but I think it still paints an accurate picture without the need for deep mathematics, which justifies the abuses I heap on it. Curiel (2009) gives a rigorous exposition of all the technical matter used and discussed in this paper, along with proofs of the two primary technical results, theorems 4.1 and 5.1.

2 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.[3] 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 a *dynamical 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

[2]Butterfield (2003, 2004) is a notable exception.

[3]I 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.

theory is in place.[4] 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.

A *state* of a system is the aggregation 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 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 dynamical 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.

As a brute fact about the physical world, every dynamical 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.[5] (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 finite number of points), and one derives the manifold structure by requiring that all quantities be smooth (except perhaps at a finite number of points).[6] So quantities play a dual role, one local

[4]This 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.

[5]As stated, the claim is not correct. I know of exactly one example of a dynamical 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.

[6]To allow finite 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

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 dynamical system has the property that at least a subset 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, form a curve on the dynamical 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 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 a dynamical 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 dynamical 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 the beginnings of 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 quite natural, necessary and sufficient toolbox and supply of material for the job.

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 distinct 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 distinct or significant quantities? 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

can ignore these sorts of problems for our purposes.

of each other, for then momentum and energy for a free particle would not be different quantities. Perhaps the most fundamental question is this: why does the space of states of a classical system always have the structure of a 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 dynamical 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 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 dynamical 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 dynamical 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 dynamical 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.[7] 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

[7] The construction is due to R. Geroch; I learned of it in conversation with him.

for Hamiltonian mechanics. As their names suggest, the defining properties of the velocital and the momental quantities for a generic dynamical system are the same as (or, at least, very similar to) those for velocity and momentum, respectively, in Newtonian mechanics. Configurative quantities 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 dynamical systems in §2, and so nothing like a configuration space (the set of all configurations, which naturally accrues the structure of a 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, *etc.* At this stage, therefore, nothing in the description of a dynamical 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.) As we will see, the characterization of the possible interactions a classical system can enter into with other systems provides the key.

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 dynamical system, which in turn will found the argument for the naturalness of Lagrangian mechanics for classical systems. As it turns out, the investigation of a system's possible interactions provides 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.[8] 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 dynamical systems, it seems, will get one at most the dynamical 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

[8]This 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).

play a privileged role in our comprehension of Newtonian mechanics. The equations of motion representing the particle's evolution take the 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 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: 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. The kinematical vector field associated with this evolution is, 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

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

where $\mathbf{F}_{\mathbf{push}}$ is the force exerted on the particle during the interaction, which may be a function of anything one likes; the associated kinematical vector field is $(\mathbf{v}, \mathbf{F}_{\mathbf{push}})$. 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.) 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.[9] This empirical fact does allow us to distinguish between position and velocity as physical quantities, and so between their respective equations of temporal evolution. The families of solutions the respective equations have differ in the form of the functions they contain: velocital quantities include discontinuous functions, whereas configurative ones do not.

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

[9]This 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 finite number of point, as already mentioned in footnote 6.

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 $\nabla\cdot\mathbf{B}$ and $\nabla\cdot\mathbf{B}$ for the components of the state 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 associated kinematical vector field is $(0,\ -\nabla\times\mathbf{E},\ 0,\ \nabla\times\mathbf{B})$. The only allowed interactions take the form

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

where ρ is an electrical charge density and $\jmath$ is its associated current. The associated kinematical vector field is $(0,\ -\nabla\times\mathbf{E},\ \rho,\ \nabla\times\mathbf{B}+\jmath)$. 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: only some physical quantities associated with the system can evolve discontinuously during its interactions with the environment, whereas others always evolve continuously. 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 not seem possible to give a similarly natural characterization of velocital quantities. We cannot say *simpliciter* that they are the ones that can change discontinuously under interactions, for mixed quantities such as $\mathbf{x}\cdot\mathbf{v}$ for the particle can as well.[10] One is tempted to say that velocital quantities

[10] It 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 sheds light on why this should be: the configurative and velocital quantities of a dynamical 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

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 velocital 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 velocital 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 velocital scalar quantities cannot be characterized in a way analogous to that for configurative ones, velocital directions in the space of states, so to speak, can be characterized in a precise sense.[11]

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 equations that formed the sets (3.1) and (3.2). A kinematical constraint in a theory imposes relations that must hold among the possible values of some set of a system's physical quantities 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. Newtonian mechanics 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. If the kinematical constraints demanded by a theory do not hold for a family of phenomena, that theory cannot treat it. 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.

We now have a characterization of configurative quantities derived from the intrinsic physics of dynamical 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

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. I shall give a more precise statement of this in §4.

[11] In many accounts of Lagrangian mechanics, the configurative quantities are taken as somehow primary and the velocital as derived. My approach has the virtue of making clear that neither configurative nor velocital 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 dynamical systems, in §4.

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 $(\rho_1,\, \boldsymbol{\jmath}_1)$ and $(\rho_2,\, \boldsymbol{\jmath}_2)$ takes the form

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

or, more suggestively,

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

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 interact 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.[12]

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 dynamical 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, they are in one-to-one correspondence, roughly speaking, with half the dimensions of the space of states at any given point.[13] (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

[12] An affine space is essentially 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.

[13] The latter property, obviously, depends on the fact that the space of states is even-dimensional in the finite case; see footnote 5.

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 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, therefore, 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 do not point in configurative directions. 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 the 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 dynamical space of states.

It is worth remarking before moving on that, on the assumption that the kinematical vector fields are fundamental in some sense, as indeed they seem to be—in this picture, one gets the dynamical space of states and the kinematical vector fields before one gets the interactions and configuration space—this analysis yields a surprising conclusion. What counts as a configurative quantity for dynamical 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 its environment: what counts as a configurative quantity is *not intrinsic to the system,* but is rather a property of the system's allowed interactions with the environment.[14]

Now, to show that the dynamical 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 line 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. A point of the globe's tangent bundle 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

[14]This is one important reason I think points of view such as that of Wallace (2003) (briefly discussed in footnote 11) are not viable.

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 point p in the manifold, the collection of all points $(p, \boldsymbol{\xi})$ in the tangent bundle, *i.e.*, all pairs such that $\boldsymbol{\xi}$ 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 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. Recall that the family of kinematical vector fields of a dynamical system has a distinguished member, the free field. Starting from any state, a dynamical 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) that state attributes to the system. Such an ordered pair, however, is exactly a point of the tangent bundle of configuration space. It follows that, in so far as the free dynamical vector field is itself a distinguished vector field, one has a distinguished isomorphism from the dynamical space of states to the tangent bundle of configuration space, completing the construction.

Theorem 3.1 (R. Geroch) *There is a canonical diffeomorphism from the space of states of a dynamical system to the tangent bundle of its configuration space.*

Before moving on, it is worthwhile pausing to take stock of our progress. From a weak physical assumption inferred directly from observation—that interactions can directly affect only 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; 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 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. It is on its face a peculiarly strong and clear example of the sort that gives the structural-semantical view of physical theories the appeal it has.

4 Classical Systems Are Lagrangian

I first describe how the structures of a dynamical system, when carried over to the tangent bundle of configuration space by the canonical isomorphism, allow one to construct a Lagrangian formulation for it. It will follow that dynamical 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 dynamical representation. In consequence, Lagrangian systems are dynamical 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 determines as its solution a vector field that gives the kinematical evolution of the system.[15] 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 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, the Lagrangian vector fields: a vector field represents a possible solution to the Euler-Lagrange equation if and only if it is second-order.

I claim this makes sense on physical grounds. One can think of a second-order vector as the acceleration 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 that represents the evolution of a free Newtonian particle, for example, has

[15] In fact, this is true only of the homogeneous Euler-Lagrange equation; the inhomogeneous Euler-Lagrange equation can include so-called generalized forces. 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.

the same form as the free kinematical vector field in the example of §3, $(\mathbf{v},\, \mathbf{0})$, as does a particle experiencing a force $\mathbf{F}$, $(\mathbf{v},\, \mathbf{F})$. This suggests there is an intimate relation between the kinematical vector fields of a dynamical 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 are the generalization of $(\mathbf{x},\, \mathbf{v})$ as used on the dynamical space of states of the Newtonian particle. 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 '$\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{x}}$', 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

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

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 form[16]

$$v_i \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{q_i}} + \xi_j \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{v}_j} \tag{4.1}$$

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

$$\eta_i \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{v_i}} \tag{4.2}$$

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).

For a given scalar field, the Lagrangian L, the (homogeneous) Euler-Lagrange equation in natural coordinates has the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial v_i}\right) - \frac{\partial L}{\partial q_i} = 0 \tag{4.3}$$

The solution to this equation (if there is one), the Lagrangian vector field associated with the Lagrangian L, is a vector field $\boldsymbol{\xi}$ on the tangent bundle. Not every scalar field yields a well posed

[16] I use the Einstein-summation convention, where one implicitly assumes a sum over the values of all repeated indices. Thus, for example, '$v_i \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{q_i}}$' is short-hand for '$\sum_i v_i \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{q_i}}$'.

Euler-Lagrange equation, however. To state the conditions under which this is true, let us say that a Lagrangian L is regular if

$$\det \left| \frac{\partial^2 L}{\partial v^k \, \partial v^l} \right| \neq 0 \tag{4.4}$$

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 (4.4) does this. Roughly speaking, the condition guarantees that a certain anti-symmetric matrix the Lagrangian defines be invertible. The coordinate form of the Euler-Lagrange equation, (4.3), masks the presence of this matrix,[17] but it is there under the scenes, and the fact that it is invertible guarantees the existence of a unique solution.

We are now in a position to address the first problem we set ourselves in this section, the construction of a Lagrangian formulation of a dynamical system using only the structure the representation of a dynamical system makes available. We know already from theorem 3.1 that the dynamical 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 dynamical system are exactly the possible solutions to the Euler-Lagrange equation posed on the tangent bundle of the dynamical system's configuration space. This completes the argument for the first claim of this section: dynamical 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 dynamical representation (in the sense of §§2–3). In fact 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 dynamical 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.

This is our problem: 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 dynamical

[17] One can see it most easily in the geometrical formulation of Lagrangian mechanics; see Curiel (2009).

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 the first assertion of this section would show that to know how to give a completely abstract Lagrangian formulation of a system on its abstract space of states would *eo ipso* suffice to reconstruct all the structure on the space of states accruing to it as a dynamical system. In the event, the answer 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.

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.*

The theorem states, in other words, that not all manifolds admit the appropriate geometry for the formulation of the Euler-Lagrange equation; see Curiel (2009) for a precise statement and proof of the theorem.

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 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 dynamical.

We can now give an alternate characterization of a dynamical system's kinematical constraints that will be useful in §5; it will at the same time show why quantities such as the one I mentioned in footnote 10 can play no physically significant role in the representation of dynamical systems. Recall that physical quantities, as scalar fields on the space of states, can be used to parametrize the space of states in the sense of assigning values to every point so as to individuate them. The kinematical constraints can be expressed as relations among quantities that any set of them must satisfy in order that the quantities yield a parametrization of the space of states that satisfies the kinematical constraints as originally formulated. Call such a set of quantities 'appropriate'. Then the kinematical constraints can be expressed by demanding that the only sets of physical quantities appropriate for a parametrization are those such that half are configurative and the other half are the velocital ones formed from the respective dynamical derivatives of those configurative 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, understanding at both the semantic and the cognitive levels. 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. 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 semantical 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, 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.

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 become clear almost immediately that Lagrangian models and Hamiltonian models of the exact same systems in the exact same empirical regimes must have different semantics, whether or not there is any physically significant structure the two models isomorphically share. This already contravenes the semantical view of physical theories, because the two models are of the same state of the same system, evolving in the same environment, coupled with other systems in the same ways, *etc.*: they are not models in competing, mutually exclusive theories, nor is one the approximation, in a preceding, superannuated theory, of another successor theory, nor anything else of the like. Because the two models are not isomorphic in a strong sense, however, the semantical view cannot attribute the same semantical content to them.

Now, 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 $\boldsymbol{\nabla} H$ 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. For our purposes, one can think of a symplectic structure as an anti-symmetric, invertible matrix. Thus one can use it to define a map from pairs of vectors to scalars, from individual vectors to individual 1-forms, and from individual 1-forms to individual vectors. Every cotangent bundle comes equipped with a canonical symplectic structure, $\boldsymbol{\Omega}$. Hamilton's equation is then

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

We call $\boldsymbol{\xi}$ 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, $\boldsymbol{\nabla} H$, to an individual tangent vector, $\boldsymbol{\xi}$.) In order to recover the usual formulation in terms of coordinates, fix a natural coordinate system (q_i, p_j) on the cotangent bundle. (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 $\mathbf{dq_i}$ generated by the coordinates (q_i) on configuration space.) In those coordinates, the symplectic structure is

$$\mathbf{dq_i} \wedge \mathbf{dp_i}$$

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. We therefore generalize natural coordinates on a cotangent bundle to canonical coordinates on a symplectic manifold: a coordinate system is canonical if Hamilton's equations expressed in their 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 dynamical systems (and Lagrangian mechanics). That is why (q_i, p_i) 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 $m\mathbf{v}$ 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 rate of change of its velocities 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 generalized velocities. That is the physical meaning, the semantics, of momentum in Hamiltonian mechanics. One can see this only by taking account of the global structure of the entire theory, not by the examination of any single model.

Now, the linearity of Hamilton's equation implies that the space of all Hamiltonian vector fields 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

$$\mathbf{\Omega}(\mathbf{\nabla} H_1) = \boldsymbol{\xi_1}$$
$$\mathbf{\Omega}(\mathbf{\nabla} H_2) = \boldsymbol{\xi_2}$$

then

$$\mathbf{\Omega}(\mathbf{\nabla}(H_1 + H_2)) = \boldsymbol{\xi_1} + \boldsymbol{\xi_2}$$

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.[18] 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 a dynamical 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 a dynamical 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, which had better be the case since phase space may not even be diffeomorphic to a cotangent bundle over configuration space. In consequence, not only does the Hamiltonian formulation of a system not allow one to express the kinematical constraints essential to dynamical systems, but it does not respect them, for it allows solutions to the equations of motion that violate them in the sense that they cannot be the images of second-order vector fields on the tangent bundle under a Legendre transform. (I discuss the Legendre transform in §6.)

Still, Hamiltonian mechanics does in fact impose its own kinematical constraints among its analogues to configurative and momental quantities. 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.*, $\frac{\boldsymbol{\partial}}{\boldsymbol{\partial} \mathbf{q_i}}$ 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

[18] A Lie algebra is an anti-symmetric, bilinear product of two vector fields that yields a third vector field; it also satisfies the Jacobi identity. The exact definition is not important for our purposes, so if this isn't clear to you, don't worry about it.

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. This is the analogue in Hamiltonian mechanics of the expression in Lagrangian mechanics of its kinematical constraints by the proposition that the unit vector fields generated by the natural coordinate systems are Lagrangian. One difference is that not all Hamiltonian vector fields arise in this way (the zero vector field does not, *e.g.*), whereas all Lagrangian vector fields do arise in that way. The family of canonical vector fields do, however, 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 the symplectic structure $\mathbf{\Omega}$: the bracket for two scalar fields f and g is given by

$$\{f,\, g\} \equiv \mathbf{\Omega}(\mathbf{\nabla} f,\, \mathbf{\nabla} g)$$

(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. This is a far weaker condition than demanding that the momental quantities be dynamical derivatives of the configurative ones, the analogue of the kinematical constraint demanded by Lagrangian mechanics. 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 ps and the qs 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 ps and qs.) In consequence, Hamiltonian mechanics does not respect the kinematical constraints that relate configurative and momental quantities in dynamical 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 one. 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 a dynamical 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 dynamical 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 do so will allow us to formulate and sketch the proof of the closest analogue in Hamiltonian mechanics to theorem 4.1.

It is instructive to think about this idea in more detail by trying to recapitulate for Hamiltonian mechanics the analogue of the arguments of §3, to see whether we can recover the Hamiltonian symplectic structure starting with the fundamental elements anologous to those we used in the construction of configuration space for dynamical 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 dynamical 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 dynamical 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 discover by physical probing. We say in this case that the vector space of 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.) It follows that starting from any point we can reach any nearby point to first-order in 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 by changes of no order higher than the first.) 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 the elements of any preferred coordinate system satisfy $\{q_i, p_i\} = 1$. We use this to define a map Ω from pairs of canonical vector fields to scalar fields:

$$\begin{aligned} \Omega(\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{q_i}}, \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{q_j}}) &\equiv \{q_i, q_j\} = 0 \\ \Omega(\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{q_i}}, \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{p_j}}) &\equiv \{q_i, p_j\} = \delta_{ij} \\ \Omega(\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{p_i}}, \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{p_j}}) &\equiv \{p_i, p_j\} = 0 \end{aligned} \tag{5.4}$$

Restricting consideration to a single point, invoking the linearity and anti-symmetry of the Poisson bracket, and noting that the vectors $\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{q_i}}$ and $\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{p_j}}$ 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, '$\boldsymbol{\Omega}$', 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

$$\boldsymbol{\Omega}(\boldsymbol{\nabla} q_i) = -\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{p_i}}$$

and

$$\boldsymbol{\Omega}(\boldsymbol{\nabla} p_i) = \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\mathbf{q_i}}$$

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 $\boldsymbol{\Omega}(\boldsymbol{\nabla}\cdot\,)$ 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 group 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.

To sum up, we have discovered that the semantics of the appropriate quantities is different in Hamiltonian mechanics from those in Lagrangian mechanics. We have seen that the family of Hamiltonian vector fields is not isomorphic to the family of a dynamical system's kinematical vector fields, and that Hamiltonian mechanics does not allow one to define an isomorphic analogue to the interaction vector fields of a dynamical system. Because the dynamical 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 dynamical 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 dynamical system, *viz.*, phase spaces that are not cotangent bundles. An 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. (See, *e.g.*, Arnold 1978, pp. 318–330, appendix 2.) Finally, and most important, the kinematical constraints of the two frameworks do not encode isomorphic relations. Thus, dynamical systems, in so far as they are Lagrangian, are in no way Hamiltonian, and Hamiltonian ones are not dynamical.

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 dynamical systems, how can it provide adequate representations of classical systems (*e.g.*, the simple harmonic oscillator)?

2. Why does Lagrangian mechanics always respect the constraints of dynamical 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?

I start with the first.

In order to apply Hamiltonian mechanics to model dynamical systems, we have to impose the kinematical constraints of a dynamical 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 dynamical systems, *viz.*, those that satisfy $p_i \triangleright \dot{q}_j$—that the p_i are linear functions of the 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 to both linear momentum and to angular momentum, and no weaker.

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

$$\frac{\partial H}{\partial p_i} \triangleright p_j \tag{6.1}$$

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, in the examples we are familiar with from text-books, all the α^{mn} are constants (not necessarily all the same), and we must in fact restrict ourselves to this case. In general, that a Hamiltonian has this restricted form implies that half of the relations in the coordinate formulation of Hamilton's equation, *viz.*,

$$\frac{\partial H}{\partial p^i} = \dot{q}_i \tag{6.3}$$

become physical tautologies, in the sense that they serve only to represent the kinematical constraints of the dynamical system, $p_i \triangleright \dot{q}_i$. Thus, that a Hamiltonian has the form (6.2), where all the α^{mn} are constants, is the necessary and sufficient condition for it to model a dynamical system.

Hamiltonians of any other form do not do this. $H = \frac{1}{2} p_i p^i + \sum_j p_j$ is a funny one. It gives $\dot{p}_i = 0$ for the dynamics, like a free system, but $\dot{q}_i = p_i + 1$, 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

$$v_i \equiv \frac{\partial H}{\partial p^i} = p_i + \frac{1}{4} p_i p^i$$

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* is not 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 dynamical 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).

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 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 lifts from configuration space, so to speak, 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.[19] This result is explained by the fact that, locally, all smooth vector fields, and so all well behaved families of curves, look exactly alike. If you've seen one, you've seen 'em all. 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 has to do with 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 a dynamical system are respected.[20]

Thus, it is built into Lagrangian mechanics from the start, by necessity, that the kinematical constraints of dynamical 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. Not only this, though—it also builds in from the start that the Lagrangian must be quadratic (at least) in the velocities for there to be a unique solution, *i.e.*, that the Lagrangian be regular as defined by the relation (4.4). One can see this from the fact that 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 k^i). 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.

We now address the third question, what happens in the passage from a Lagrangian to a Hamiltonian representation of a dynamical system by way of the Legendre transform: why does the structure of a dynamical 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 $v^i \mapsto \dfrac{\partial L}{\partial v^i} \equiv p_i$, providing a map from $T\mathcal{M}$ to $T^*\mathcal{M}$. (Thus the condition that a Lagrangian be regular, equation (4.4), 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 non-physical second-order vector fields when they are mapped to the cotangent bundle? Consider the field

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

[19]It is not difficult to make this argument precise and rigorous.

[20]In 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 $\mathcal{C}$.

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.[21] 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,[22] we get

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

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 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.

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

$$\dot{q} = 1$$
$$\dot{p} = 0$$

This seems to represent a free particle moving with velocity 1. If you try to perform a reverse Legendre transform to put it into Lagrangian form, then you get $\dot{q} = 1$ and $L = vp - H = \dot{q} - 1$, 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.

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 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. 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 dynamical systems. One can make a stronger statement: any physically significant transformation of a Lagrangian representation of a system into a Hamiltonian one *must* wipe out verticality: Hamiltonian mechanics cannot be formulated with it, as Lagrangian mechanics cannot be formulated without it.

Still, even in the face of all the contravening evidence I have marshaled, 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 a dynamical system in a natural way. This suggestion

[21] This 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.

[22] And, in fact, this L is a Lagrangian that yields the vector field (6.4) as a solution with the appropriate generalized forces.

leads us to address the fourth and final question we posed on page 24. 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 coordinate-form of the Euler-Lagrange equation, *viz.*, equation (4.3), the Euler-Lagrange equation 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 we intimated immediately after the condition (4.4). It is not a fixed symplectic structure as in Hamiltonian mechanics, however, but rather it 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. I will discuss the significance of this difference in §7, where I argue that this seeming sameness of structure does not support the semantical view of physical theories, because the formally isomorphic structures play different semantical roles in each framework.

There is one more possibility for defending the semantical view against these arguments. Most of the examples and arguments so far have treated only the kinematical structures of theories. 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 models of individual systems such as the equation of motion for a simple harmonic oscillator in Hamiltonian Mechanics will not be isomorphic to the model for the same system in Lagrangian mechanics, for all the reasons given here. 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 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 2009). For an arbitrary vector field $\boldsymbol{\zeta}$ on the tangent bundle, there always exist three second-order vector fields $\boldsymbol{\xi}$, $\boldsymbol{\eta}$ and $\boldsymbol{\theta}$ such that

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

We already expect that a kinematically possible evolution 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 thus that would *prima facie* represent "different systems".) Then equation (6.5) states, roughly speaking, that we can get from any point in the dynamical 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 either.

It will be useful for the arguments of §7 to make this last point concrete 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 to 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

$$v \equiv \frac{\partial H}{\partial p} = p$$

for the velocity, yielding the expected Lagrangian

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

This is a regular Lagrangian, having as its solution 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 (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 physically significant isomorphic structures. I put off until §7 the argument that the sets of relations (6.6) and (6.7), though apparently the same in form, do not in fact have the same semantic content. For now, let us restrict attention to the way the solutions for the simple harmonic oscillator change under small perturbations. This will illustrate the physical significance of the difference in the respective orders of the equations of motion in the two frameworks (first-order for Hamiltonian, second-order for Lagrangian).

We expect on physical grounds that a slight perturbation to the forces acting on the system should yield an evolution quite close to the original. Consider, then, the Hamiltonian

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

for $1 \gg \epsilon > 0$. Hamilton's equations give

$$\begin{aligned} \dot{q} &= p - \epsilon \\ \dot{p} &= -q \end{aligned} \tag{6.8}$$

For ϵ small enough, this will be as close to a circle as one likes, for any reasonable sense of 'close'. Now, however, we are faced with a dilemma. If we use the inverse-Legendre transform defined by the perturbed Hamiltonian, then the derived Lagrangian system is unphysical in the sense that it does not respect the kinematical constraint $v \equiv \dot{q} = p$; rather,

$$v \equiv \dot{q}\frac{\partial H'}{\partial p} = p - \epsilon$$

If we enforce the kinematical constraint $v = p$ (say, by using the Hamiltonian of the ordinary simple harmonic oscillator to define the Legendre transform), then the derived Lagrangian in this case has the form

$$L' = \frac{1}{2}v^2 - \frac{1}{2}q^2 + \epsilon v + \epsilon^2$$

for which, it is easily checked, the Lagrangian vector field is exactly equivalent to that for an ordinary simple harmonic oscillator, which is not isomorphic to the non-circular, perturbed Hamiltonian solution. In other words, the Lagrangian vector field in this case is not isomorphic to the Hamiltonian vector field. (We ignore, for the sake of argument, the fact that it is not clear from a physical point of view what it means to use the inverse-Legendre transform defined by another Hamiltonian to map a solution of a given Hamiltonian to a Lagrangian formulation.)

The divergence of the solutions in this case can be explained by the difference in the order of the fundamental equations and the difference in the kinematical constraints in the respective frameworks. That Hamilton's equation is first-order guarantees that one can reach from any point in phase space to any nearby point directly along a Hamiltonian vector field simply by perturbing any given Hamiltonian; because the point is reached along a Hamiltonian vector field, moreover, it follows *a fortiori* that the Hamiltonian kinematical constraints are not violated. Because the Euler-Lagrange equation is second-order, one cannot do this while preserving the Lagrangian kinematical constraints.

[*** Geroch (lecture notes on geometrical formulation of quantum mechanics, "math-physics/other-minds/geroch/qm/geroch-geom-qm.pdf", §18, pp. 47) shows that one can derive the quantum Hamiltonian from the classical one unambiguously, according to the "standard" method, when and only when the classical Hamiltonian is regular, *i.e.*, quadratic in momentum; see his discussion there ***]

7 A Counter-Example to the Semantical View of Physical Theory

In order to draw out the full philosophical relevance of the technical arguments of the previous sections, I first sum them up.

1. the global structures of a theory of classical systems, as characterized in §§2–3, are necessary and sufficient to give the theory a Lagrangian formulation
2. those structures do not permit one to recover a Hamiltonian formulation
3. none of the fundamental, global kinematical structures of Lagrangian mechanics is isomorphic to any in Hamiltonian mechanics
4. in consequence, the dynamical structures of the two are not isomorphic in any way that preserves semantical content
5. in futher consequence, no individual model in one shares physically significant, fundamental structure with any in the other, even when the two are models of exactly the same classical system

These facts, I shall argue, imply that the semantical view of physical theories, in the forms it has been propounded, cannot be correct.

van Fraassen (1980, p. 64) concisely sums up what I take to be the three fundamental tenets of the semantical view:[23]

> To present a theory is to specify a family of structures, its *models*; and secondly, to specify certain parts of those models (the *empirical substructures*) as candidates for the direct representation of observable phenomena. The structures which can be described in experimental and measurement reports we can call *appearances*: the theory is empirically adequate if it has some model such that all appearances are isomorphic to empirical substructures of that model.

I think this constitutes a minimal construal of the semantical view. This is shown by the fact that philosophers as dramatically different in their goals and metaphysical and epistemological predilections as Suppes, Suppe, and Worrall, *inter multa alia*, in essence endorse the tenets.

These tenets have a simple implication that will ground my argument: if two theories appropriately and adequately represent the same physical systems under the same conditions in the same regimes, then the relevant structures in the models of each should be isomorphic to each other, since they are both, *ex hypothesi*, isomorphic to the same empirical structure. Any particular account of the semantical view can have as sophisticated, nuanced and clever a view of the relations between theoretical and empirical structures as it wants, but in so far as at bottom it endorses those tenets,

[23] Italics are van Fraassen's.

it cannot be correct. The case of Lagrangian and Hamiltonian mechanics provides, I shall argue, a decisive counter-example to the third fundamental tenet, that the semantics of a model (and so, by the lights of the semantical view, of a theory) is founded on isomorphism of theoretical and empirical structures. The explanation of the details of the way the case works as a counter-example, moreover, will at the same time show why the first tenet cannot be correct: a theory neither is nor can be fully characterized by its family of models. (If the first and the third tenet are not correct, then the second is not so much incorrect as irrelevant.)

Now, when two theories provide good models of the same system, the semantical view demands that sameness of meaning for the models implies sameness of structure between them. But that does not always happen. The Lagrangian and the Hamiltonian models of the simple harmonic oscillator 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, 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 dynamical systems, *viz.*, that its velocity is the temporal derivative of its position". Both the structures one uses to formulate **P** and the grounds for its truth, however, could not be more different between the two frameworks. In Lagrangian mechanics, the proposition is a necessary truth, in the sense that every Lagrangian system satisfies it by construction; it is not a prediction in the framework about the possible behavior of a system, but rather a pre-condition that systems must satisfy in order for Lagrangian mechanics to model them appropriately. In Hamiltonian mechanics, the proposition is not even true in most models, much more necessarily true for all. Hamiltonian mechanics predicts that a simple harmonic oscillator satisfy the kinematical constraints of a dynamical system; it does not require it as a pre-condition for its propriety as a representation of the thing. In Lagrangian mechanics, the relation that makes the proposition true is encoded in a global, fundamental, kinematical structure that defines the kinematics of every model, *viz.*, the algebra encoded in the family of models representing all possible evolutions. In Hamiltonian mechanics, it is encoded in the idiosyncratic form of a particular Hamiltonian, the local dynamics of that particular model. This explains why the example of the simple harmonic oscillator worked out at the end of §6 does not provide grist for the semantical theorist's mill. Even though the formal representation of the motion of the simple harmonic oscillator, as integral curves of the kinematical vector field on the respective space of states, is the same in both frameworks, the semantical content of the details of the representation are different. The models in the two frameworks make the same set of empirical propositions about the simple harmonic oscillator true; in that sense, each represents the system as having the same evolution as does the other; but the two do *not* make the same predictions about the system, precisely because $\dot{q} = v$ is analytic in one and not a prediction, but is a prediction in the other. *A fortiori*, they cannot make the same predictions.

A finer analysis puts the point more trenchantly. Strictly speaking, 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 dynamical 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).

Sameness of meaning, then, does not imply sameness of structure. Neither is the converse slogan true. To ground a semantics, the things that are isomorphic must mean the same things in their respective isomorphic systems, must play the same roles. The only structure formally isomorphic in the two frameworks has different physical significance, and so different semantic content, in each one. The canonical symplectic structure in Hamiltonian mechanics plays three semantical 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 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 (4.4), 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.

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, isomorphic structure. Agreement in prediction, however, simply does not imply the existence of isomorphic, physically significant structure. Even when there are isomorphic, 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 verificationist empiricism. Mere existence of isomorphisms between mathematical structures, then, cannot ground a semantics of physical theory: just because two theories model the same systems and have isomorphic structures it does not follow that the empirical meaning of the isomorphic structures are the same in the two theories—the physical roles the structures play in one theory may be distributed differently among the isomorphic structures in the other theory, so to speak, or even among entirely different structures in no way isomorphic to any in the other.

It is now clear why the first tenet of the semantical view cannot be correct as well: 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 the two theories only by examination of the global structures of the theory, 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 dynamical 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.

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 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, cannot stand. 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 isomorphism; 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.

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. Available online at the philsci archive: http://philsci-archive.pitt.edu/archive/00001192/01/ButterfieldForFormal.pdf.

Butterfield, J. (2004). Between laws and models: Some philosophical morals of lagrangian mechanics. Available online at the philsci archive: http://philsci-archive.pitt.edu/archive/00001937/01/BetLMLag.pdf.

Curiel, E. (2009). The geometry of the Euler-Lagrange equation and the Legendre transform. Unpublished manuscript in preparation for submission.

Rosenberg, R. (1977). *Analytical Dynamics of Discrete Systems*. New York: Plenum Press.

van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University 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.

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.