The Geometry of Dynamics

A. IBORT

Departamento de Física Teórica, Universidad Complutense de Madrid, 28040 Madrid, Spain

AMS Subject Class. (1991): 58F05, 49N45, 70A05

Contents

1. Introduction
2. Hamilton’s equations of motion and Poisson brackets.
4. The action functional and Hamilton’s principle.
5. Invariance in dynamics. Constants of the motion and reduction.

1. INTRODUCTION

One of the main purposes of these notes is to offer a basic dictionary of the arcana used by practitioners of geometrical mechanics to researchers on other areas of the study of dynamical systems. To achieve this in a reasonable way, we will try to expose the rationale behind some of them. We believe that it could be a good starting point to summarize some of the assumptions and/or creeds that most Geometrical Mechanics People (GMP for short) have:

1. Geometrical modeling of physical systems is useful.
2. The Geometry involved in such process is natural.

These principles are inspired by Physics and Mathematics (it is interesting to point it out that GMP’s are mainly recruited from the ranks of theoretical physicists and differential geometers)\(^1\).

\(^1\)We should point out here that the comments, remarks and ideas contained in this introduction reflect only a partial description of the different viewpoints held by the people that have contributed to the foundations of Mechanics, continuum mechanics and classical Field Theories. In this sense we probably should add a third postulate to (1) and (2):

3) Each GMP has a optimal set of reasons to support his/her viewpoints on the foundations of the subject.
The first assumption comes naturally from the pursuit for fundamental laws of nature. This quest has led to the introduction of a variety of geometrical ideas in Physics. As a consequence of this it has become evident the power of geometry and topology to unify and clarify concepts and results, and this makes the second assumption unescapable. This is, the geometrical ideas that have been arising along this process are the natural geometrical notions introduced and studied by differential geometers (in some occasions the same ideas have been discovered almost simultaneously and unconnectedly in Physics and Mathematics).

A rough simplification of how this mechanism operates is offered for instance by Relativity Principles [Be76]. Relativity principles in Physics state that physical phenomena are independent on the particular frame of reference we use to describe them within a previously defined class of them. The right tool to express mathematically this idea is to use intrinsic geometry with respect to some a priori chosen group and this is another way of stating Klein's “erlangen program” for the geometrization of mathematics. Eventually the general covariance principle forces us to use intrinsic geometry on manifolds to formulate physical laws.

From the discussion above we are tempted to extend the formulations of ordinary relativity principles and state a metarelativity principle saying that the description of physical systems must be done on the realm of manifolds (or differential geometry). Then, from now on we will assume that we are dealing with systems that can be modeled using smooth manifolds (either finite or infinite dimensional).

We must remark here that there are obvious situations were the previous discussion does not apply and the principles (1) and (2) have to be reformulated. For instance it is not obvious what is the right foundational setting to describe some quantum (finite and infinite dimensional) mechanical systems [Ja68], [Ba91].

Assuming in what follows that differential geometry could be a good tool to model dynamical systems, we can ask ourselves, what do we actually want to model?

If we are describing evolution systems defined using some space of parameters $M$, usually another smooth manifold, the simplest choice for modeling is the space of all trajectories or histories of the system [So70]. Usually such trajectories $\gamma$ are solutions of an ordinary or partial differential equation. Such space of all trajectories of a dynamical system (physical or not) will have a local description using initial data or Cauchy data, and we will obtain in this
way a manifold structure, possibly with singularities, for it. An alternative modeling can also be considered. Instead of thinking in time evolution and initial value problems, we can consider atemporal observers, i.e., observers “seeing” the histories of a given physical system as a whole. To describe such global behaviour, variational principles are the right tool. The calculus of variations provides the bridge between the two approaches sketched here (see Fig. 1.).

We must point it out that not all systems are parametrized by smooth manifolds. Many interesting situations arise where this is not possible but we will not discuss the meaning of geometrically modeling this kind of systems in these notes.

It can seems odd the effort to model the set of solutions of the system that we want precisely to solve. This is not so because a substantial amount of information about the structure of such solutions could come from such modelization. Nevertheless, it is much more frequent to make a strong simplification to the modelization problem by considering as fundamental entities the initial data of the systems we observe. This is very much attached to our psychological perception of real, or fundamental, as those things that we can

\footnote{Notice that this is not equivalent to postulate Lorentz covariance.}
actually touch, control, measure, etc. Then, one of the main task for GMP’s is to understand the space of trajectories of a (family of) dynamical system(s) using the geometry and/or the topology of the space of initial data\(^3\).

The previous declaration makes also possible to select a priori some particular ODE’s and PDE’s as more interesting than others. We will discuss this point in the coming sections.

2. Hamilton’s Equations of Motion and Poisson Brackets

Inspired from Physics and in particular by Classical Mechanics [Ar76], [Ab78], [Ma85], [Lb87]\(^4\), it is a good idea to consider systems whose initial data are parametrized by \(q^i, p_i\), where \(q^i\) are local coordinates describing the points \(q\) of a smooth manifold \(Q\) and \(p_i\), the canonical conjugate momenta, are local coordinates describing covectors on such manifold, i.e., points \(p\) in the contangent bundle \(T^*Q\) of \(Q\). An arbitrary dynamical system in such space will have the general form

\[
\dot{q}^i = f^i(q,p); \quad \dot{p}_i = g_i(q,p). \tag{2.1}
\]

We can write the previous system of ODE’s in an operational way using the first order differential operator or vector field,

\[
X = f^i \frac{\partial}{\partial q^i} + g_i \frac{\partial}{\partial p_i}. \tag{2.2}
\]

The vector field \(X\) is a linear operator on the space of smooth functions on \(T^*Q\) that encodes the infinitesimal evolution of any quantity defined on it, this is, if \(F \in C^\infty(T^*Q)\), then the infinitesimal change \(\dot{F}\) of \(F\) along the solutions of the systems of eqs. (2.1) is given by:

\[
\dot{F} = f^i \frac{\partial F}{\partial q^i} + g_i \frac{\partial F}{\partial p_i} = X(F). \tag{2.3}
\]

In this sense the previous equation, eq. (2.3), can be considered as an intrinsic description of the system (2.1).

\(^3\)It is also a common attitude to postulate the existence of a space–time arena where the systems under study evolve and to derive their geometrical properties from this basic structure [Tr70], [La69], [Ar76].

\(^4\)We will not refer here to the extraordinary good classical literature on Mechanics because we are concentrating ourselves in the geometrical modeling problem. Thus the texts we are quoting are those making an special emphasis on the geometrical aspects of Mechanics.
A fundamental axiom in the description of physical systems is what we could call the Energy paradigm, that can be stated as follows:

“For every physical system there is a function defined on its space of states, called the energy or Hamiltonian of the system, containing all its dynamical information”.

Just forcing slightly the previous statement we can equivalently consider that, in the particular case that $T^*Q$ models the state space of a family of dynamical systems, there is an assignment to any function (the possible energies) $H$ on $T^*Q$ of a vector field $X_H$ describing a dynamical system. For such assignment to be meaningful it would have to satisfy some fundamental physical requirements. For instance it should be:

1. Linear. This is, if we add energies the corresponding dynamical system will be the sum of the corresponding factors:

$$X_{H_1+H_2} = X_{H_1} + X_{H_2}; \ X_{\lambda H} = \lambda X_H \ \forall H, H_1, H_2 \in C^\infty(T^*Q), \ \forall \lambda \in \mathbb{R}. \tag{2.4}$$

2. Energy conservation. The dynamical system defined by a given energy will preserve its own energy$^5$:

$$\dot{H} = X_H(H) = 0. \tag{2.5}$$

3. Compatibility of dynamical evolution. If we have two energy functions, $H_1$ and $H_2$, then the infinitesimal change on an arbitrary function when we act first with $X_{H_2}$ and then with $X_{H_1}$, is the same as the change suffered by $F$ when we act in the reverse order plus the change induced by the action of $X_{H_1}$ on $H_2$:

$$X_{H_1}(X_{H_2}(F)) = X_{H_2}(X_{H_1}(F)) + X_{X_{H_1}(H_2)}(F). \tag{2.6}$$

A more transparent way of expressing this property is to think on the difference between the infinitesimal variations of a quantity along two different dynamical systems $X_{H_1}$ and $X_{H_2}$. It is natural to think that such difference is caused by the variation of energy of the second along the action of the first, i.e., the difference term will be caused by the modified hamiltonian $X_{H_1}(H_2)$, and this is what eq. (2.6) states. It is clear that the order of the dynamical systems can be reversed and the same argument will lead us to conclude

$^5$We will not consider time-dependent dynamics here, even though there are nice geometrical frameworks for them.
that the difference will have to be generated by the Hamiltonian $X_{H_2}(H_1)$, and obviously this has to be $-X_{H_1}(H_2)$. We will see immediately that this is precisely what happens.

2.1. Poisson Brackets  From these few assumptions on the nature of the assignment $H \mapsto X_H$, we discover the following important consequences. Let us introduce first some traditional notation and denote by $\{H,\_\}$ the vector field $X_H$. Then, eq. (2.4) implies that the map $\{\_,\_\} : C^\infty(T^*Q) \times C^\infty(T^*Q) \rightarrow C^\infty(T^*Q)$ defined by $(F, G) \mapsto \{F, G\} = X_F(G)$ is bilinear. From energy conservation, eq. (2.5), we obtain that $\{\_,\_\}$ is skew-symmetric. In fact, consider $F, G$ two arbitrary functions. Then

$$0 = \{F + G, F + G\} = \{F, G\} + \{G, F\}. \quad (2.7)$$

Because $X_F$ is a vector field (hence a derivation) we get Leibnitz’s rule

$$\{F, GH\} = G\{F, H\} + H\{F, G\}, \quad (2.8)$$

and finally, from the third requirement, eq. (2.6), we get Jacobi’s identity

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0. \quad (2.9)$$

Notice that $X_{H_1}(H_2) = \{H_1, H_2\} = -\{H_2, H_1\} = -X_{H_2}(H_1)$ as it has to be. We have been led to the fundamental notion of Poisson brackets. Let $P$ be now an arbitrary smooth manifold. A bilinear map on $C^\infty(P)$ satisfying eqs. (2.7), (2.8) and (2.9) will be called a Poisson bracket on $P$ and $(P, \{\_,\_\})$ a Poisson manifold. Thus, we have shown that the energy paradigm forces us to choose a Poisson bracket on $T^*Q$. See [Jo64], [Li77], [We83], [Lb87], [Va93] and references therein for a more detailed description of Poisson manifolds.

Several Poisson brackets have been discovered in the last two hundred years. The most well-known is

$$\{F, G\}_0 = \sum_{i=1}^n \left( \frac{\partial F}{\partial q^i} \frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i} \frac{\partial G}{\partial q^i} \right), \quad (2.10)$$

which is the one mostly used in Hamiltonian mechanics\(^6\).

\(^6\)It is interesting to point out here that it was J.L. Lagrange the first who used a Poisson bracket in discussing the structure of the equations describing the evolution of the elements of a planetary orbit.
Another Poisson bracket of great interest is the linear Poisson brackets discovered by S. Lie and rediscovered several times in the last 30 years. It is defined by the fundamental commutation relations,

\[ \{x_i, x_j\} = C_{ij}^k x_k, \quad (2.11) \]

where \( C_{ij}^k \) are the structure constants of a Lie algebra (see [Ca94] and references therein).

It is important to point it out that in principle, unless some further physical requirements are added to the list above (see §3 for more comments on this), there are no privileged Poisson brackets and any other Poisson bracket on \( T^*Q \) could be used to describe Hamiltonian mechanics.

It is not known yet how to classify the Poisson brackets existing on a given manifold, not even in the nondegenerate case that we will discuss in the following section [Gr93].

### 2.2. Symplectic versus Poisson Geometry

A Poisson bracket \{., .\} defines a contravariant 2-tensor by means of

\[ \{z^a, z^b\} = \Lambda^{ab}(z), \quad (2.12) \]

where \( z^a \) denotes a local system of coordinates on our space of initial data conditions (that we can consider not to be necessarily \( T^*Q \)). Under changes of coordinates it is easy to check using the linearity of \{., .\} and eq. (2.8) that \( \Lambda^{ab} \) transforms as the components of a contravariant skew-symmetric 2-tensor called the Poisson tensor and denoted by \( \Lambda \). The Jacobi identity is equivalent to an integrability property of this tensor, to be precise \( \Lambda \) satisfies that \([\Lambda, \Lambda] = 0\), where \([., .\) denotes the Schouten-Nijenhuis bracket of multivectors [Li77], [Va93].

We will say that the Poisson bracket \{., .\} is nondegenerate if \( \det \Lambda^{ab} \neq 0 \). In such case, we can define the inverse tensor \( \omega_{ab} = (\Lambda^{-1})_{ab} \) where \( \Lambda^{ab}(\Lambda^{-1})_{bc} = \delta^a_c \). The 2-form

\[ \omega = \omega_{ab} dz^a \wedge dz^b, \quad (2.13) \]

is nondegenerate and because of the Jacobi identity, eq. (2.9), it satisfies,

\[ \frac{\partial \omega_{ab}}{\partial z^c} + \text{cyclic} = 0, \]

\( i.e., \omega \) is closed

\[ d\omega = 0. \]
A nondegenerate closed 2-form $\omega$ is called a symplectic form and the manifold where it is defined is called a symplectic manifold. The well-known Darboux theorem [Go69], [Ab78] states that locally any symplectic structure is like the symplectic structure obtained from $\{.,.\}_0$, i.e.,

$$\omega_0 = \sum_{i=1}^n dq^i \wedge dp_i.$$ (2.14)

Thus the local classification problem of Poisson structures is completely solved in the nondegenerate case. In the degenerate case there is a local splitting theorem by Weinstein [We83]. It is also true that any Poisson manifold is foliated by symplectic manifolds. These symplectic manifolds are integral submanifolds of the (singular) integrable distribution defined by the image of the Poisson tensor $\Lambda$. The leaves of maximal dimension correspond to the level sets of the Casimirs of the Poisson brackets, i.e., those functions that commute with all the others.

Some global theorems are known [Gr85], [Gr86] and very recently there has been new results on 4-manifolds but the global classification problem is widely open.

Thus we can conclude these remarks noticing that if we are describing a physical dynamical system $\Gamma$ on $T^*Q$ the energy paradigm imply that using the Poisson bracket $\{.,.\}_0$, it will have the form

$$\dot{q}^i = \frac{\partial H}{\partial p_i}; \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}. \quad (2.15)$$

Intrinsically, eq. (2.15), can be written using the symplectic form (2.14), as

$$i_\Gamma \omega_0 = dH. \quad (2.16)$$

Even if the parameter space is not $T^*Q$ but the Poisson structure that we use is nondegenerate, the dynamical system $\Gamma$ will be written again in appropriate coordinates as in eq. (2.15). Eqs. (2.15) will be called in what follows Hamilton's equations of motion and the intrinsic expression provided by eq. (2.16) will be refered to as the dynamical equation.

It is clear from what we said that the same ideas apply both in finite and infinite dimensional systems. Many applications of these principles and ideas can be found for instance in [Ab78], [Gu77], [Gu84], [Mr74].
3. EULER–LAGRANGE EQUATIONS AND THE GEOMETRY OF SECOND ORDER DIFFERENTIAL EQUATIONS

In spite of the attractiveness of the previous discussion, it must be noticed that experimentally (and historically) things happen in a different way. In fact it is easy to convince oneself from direct everyday observations that accelerations are proportional to the forces acting on the systems\(^7\), i.e., the equations describing the dynamics of physical systems are second order differential equations that will be written as,

\[
\ddot{q}^i = f^i(q, \dot{q}, t),
\]

and they will be called Newton’s equations of motion. Here, \(q^i\) denote as before any set of coordinates parametrizing the configurations of our system. The initial data for Newton’s equations are given by \(q^i, \dot{q}^i\) and geometrically they define the so called tangent bundle of the configuration space \(Q\). It will be denoted in what follows by \(TQ\) which is also commonly called the velocity phase space (its points corresponds to dynamical states or trajectories of the system). Of course, it is possible to transform the previous system, Eqs. (3.17), into a first order system introducing a new set of variables \(v^i\). Then,

\[
\dot{q}^i = v^i; \quad \dot{v}^i = f^i(q, v, t).
\]

This makes transparent the independence of the initial data \(q^i, v^i\), but adds nothing to the problem of solving Newton’s equations (3.17).

Some help will be obtained if we would be able to use the geometry of Hamilton’s equations to study them. But we do not have a Poisson bracket available in \(TQ\) (as the bracket \(\{.,.\}_0\) is in \(T^*Q\)) to set them in Hamiltonian form or, in other words, we do not have a realization of the energy paradigm on \(TQ\). Looking for a Poisson bracket \(\{.,.\}\) and a Hamiltonian function \(H\) on \(TQ\) such that Eqs. (3.17) would be written as \(\dot{F} = \{F, H\}\), will be called the weak inverse problem of the calculus of variations. If the Poisson bracket \(\{.,.\}\) is required to be nondegenerate\(^8\), there will be a symplectic form \(\omega\) on \(TQ\) such that the second order vector field

\[
\Gamma = v^i \frac{\partial}{\partial q^i} + f^i \frac{\partial}{\partial v^i},
\]

\(^7\)This property plays the role of a constitutive equation for a dynamical system, in fact for mechanical systems this is equivalent to \(p_i = mq_i^i\).

\(^8\)This requirement is equivalent to the nonexistence of global Casimirs, i.e., there are no superselection rules for the dynamical system.
will be hamiltonian with hamiltonian $H$, i.e., it will satisfy the dynamical equation (2.16) $i_{\gamma} \omega = dH$. This problem, slightly stronger than the one before, could be called the symplectic inverse problem of the calculus of variations.

These problems stated in such a generality have been only partially solved (see for instance [Du93], [Ca95] and references therein). On the other hand the Poisson brackets that usually arise in physical systems of mechanical type have the property that

$$\{q^{i}, q^{j}\} = 0, \tag{3.18}$$

for any pair of configuration space coordinates $q^{i}, q^{j}$. This property reflects the localizability of mechanical systems. Upon quantization, following Dirac's prescription [Di64], Poisson brackets become commutators of the corresponding quantum observables, then, the previous equation tells that the positions in configuration space of the system can be simultaneously measured, i.e., the system is localizable in configuration space. For this reason Eq. (3.18) will be called the localizability condition. In consequence, if Newton's equations are describing a \textit{bona fide} physical system, we can ask if there exists a nondegenerate localizable Poisson bracket $\{., .\}$ such that with respect to it the vector field $\Gamma$ is Hamiltonian. This problem is called, for reasons that will become evident in what follows, the inverse problem of the calculus of variations. Under different perspectives this problem has been the subject of intensive research. The first solution to it can be traced back to Helmholtz [He87] and its geometrical presentation first appeared in [Ba82]. Apart from these, other main contributions to it are due to Douglas, Sarlet, Anderson, Crampin, Henneaux, etc. (see for instance [Cr81],[He82], [An92] and [Mo90] and references therein).

The main result related to this problem can be stated as follows:

\textbf{Theorem 1.} (The inverse problem of the calculus of variations [Ca95]) A second order differential equation $\Gamma$ is (locally) hamiltonian with respect to a localizable symplectic form iff there exists a (local)$^9$ Lagrangian $L$ for it.

What is the meaning of the existence of the Lagrangian $L$? To answer this question we will discuss succinctly the proof of this statement. We will work locally (the theorem is a local statement) and we will assume that there is a closed 2-form $\omega$ of the form

$$\omega = a_{ij} dq^{i} \wedge dq^{j} + b_{ij} dq^{i} \wedge dv^{j} + c_{ij} dv^{i} \wedge dv^{j}. \tag{9}$$

$^9$See [Ba83] for the discussion of many examples of interest possessing locally defined Lagrangians and [IB90] for the description of the obstructions to find global Lagrangians.
The localizability of \( \omega \) implies that the coefficients \( c_{ij} \) must vanish and the nondegeneracy of \( \omega \) implies that \( \det b_{ij} \neq 0 \). Because \( \omega \) is closed by Poincaré's lemma there must exist (locally) a 1-form \( \Theta = A_i dq^i + B_i dv^i \) such that \( -d\Theta = \omega \). In consequence we get the following equations for the coefficients \( A_i, B_i \) and \( a_{ij}, b_{ij} \):

\[
a_{ij} = \frac{\partial A_i}{\partial q^j} - \frac{\partial A_j}{\partial q^i}, \tag{3.19}
\]

\[
b_{ij} = \frac{\partial A_i}{\partial v^j} - \frac{\partial B_j}{\partial q^i}, \tag{3.20}
\]

\[
0 = \frac{\partial B_j}{\partial v^k} - \frac{\partial B_k}{\partial v^j}. \tag{3.21}
\]

From the last of the previous equations, eq. (3.21), we get that it must exist a function \( \phi(q,v) \) such that

\[
B_j = \frac{\partial \phi}{\partial v^j}.
\]

Then, the 1-form \( \Theta' = \Theta - d\phi \) satisfies that \( d\Theta' = \omega \) and it has the form

\[
\Theta' = \left(A_i - \frac{\partial \phi}{\partial q^i}\right) dq^i.
\]

Writting again \( \Theta' \) as \( \Theta \), what we have shown is that there exists a 1-form \( \Theta = A_i dq^i \) such that \( \omega = -d\Theta \), and

\[
b_{ij} = \frac{\partial A_i}{\partial v^j}. \tag{3.22}
\]

Because the vector field \( \Gamma \) is locally hamiltonian with respect to \( \omega \), then \( d(i_\Gamma \omega) = 0 \) and if \( \Gamma \) is written locally as

\[
\Gamma = v^i \frac{\partial}{\partial q^i} + f^i \frac{\partial}{\partial v^i},
\]

the previous condition is equivalent to the existence of a local function \( H \) such that

\[
\frac{\partial H}{\partial q^i} = v^ia_{ij} - f^ib_{ij}; \tag{3.23}
\]

\[
\frac{\partial H}{\partial v^i} = v^ib_{ij}. \tag{3.24}
\]
Then, from eqs. (3.22), (3.24) we get

\[ \frac{\partial H}{\partial v^j} = v^i \frac{\partial A_i}{\partial v^j}, \]

hence the function

\[ L = v^i A_i - H + V, \]

whith \( V = V(q) \), is such that

\[ \frac{\partial L}{\partial v^i} = A_i, \]

and in consequence

\[ \Theta = \frac{\partial L}{\partial v^i} dq^i. \]

Finally, eq. (3.23), fixes \( V \) up to a constant, and then we conclude that \( \omega \) has the form

\[ \omega = \frac{\partial^2 L}{\partial v^i \partial v^j} dq^i \wedge dv^j + \frac{\partial^2 L}{\partial q^i \partial q^j} dq^i \wedge dq^j, \quad (3.25) \]

that will be denoted by \( \omega_L \). The function

\[ E_L = H = v^i \frac{\partial L}{\partial v^i} - L. \quad (3.26) \]

is the hamiltonian of the second order vector field \( \Gamma \), and the Newton’s equations defined by it take the Euler–Lagrange form

\[ \frac{d}{dt} \frac{\partial L}{\partial v^i} = \frac{\partial L}{\partial q^i}. \quad (3.27) \]

We also learn in this way that Euler–Lagrange equations (3.27) are intrinsic and coordinate independent.

Summarizing, we have found that looking for a Hamiltonian formulation for Newton’s equations is equivalent to the existence of a function \( L \) called the Lagrangian that as we will see in the next section is deeply related to the calculus of variations (and it will give sense to the name of the theorem 1). We can push forward the previous discussion and try to find out if the previous conditions have implications on the possible forms of the Lagrangians arising in the descriptions of Newton’s equations. This amounts to determine a priori which kind of interactions can occur among the particles constituting a given system. This question was raised and solved by Feynman (see [Wi50], [Dy90], [Hu92] and [Ca95] for a recent review on the subject).
3.1. The Geometry of Euler–Lagrange Equations In the search for ideas helping us to solve Newton’s equations we have arrived to a class of Hamilton’s equations defined by means of a particular kind of symplectic forms, the Cartan 2–forms derived from Lagrangian functions (3.25). These forms possess the property of being exact (as the canonical symplectic structure on $T^*Q$). Then if $\omega_L$ is a Cartan 2–form, there is a canonical symplectic potential, called the Poincaré–Cartan 1-form, which is given by

$$\Theta_L = \frac{\partial L}{\partial v^i} dq^i.$$  

(3.28)

It is obvious that $\omega_L = -d\Theta_L$. We can think of $\Theta_L$ as the result of the action of the operator $dq^i \partial/\partial v^i$ on functions $L$. This operator is similar to the action of the exterior differential $d$ on functions on a manifold. In fact, the exterior differential has associated a $(1,1)$-tensor $I = dx^i \otimes \partial/\partial x^i$ on the manifold $M$ and in the same sense the operator before defines a $(1,1)$–tensor on the manifold $TQ$,

$$S = \frac{\partial}{\partial v^i} \otimes dq^i,$$  

(3.29)

which is a “twisted exterior differential” on $TQ$. This tensor is natural and has been called the vertical endomorphism of $TQ$. The tensor field $S$ is integrable and $\ker S = \text{Im}S$. Its crucial role in the analysis of second order differential equations has been pointed out by many people, J. Klein, M. Crampin, G. Marmo, M. de León, etc. ([Cr83], [Kl62], [Mo90]). In particular it has been proved that the tensor field $S$ together with a vector field $\Delta$ called the Liouville vector field characterizes a tangent bundle (modulo some topological conditions). The Liouville vector field is the generator of dilations along the fibres of the tangent bundle $TQ$ and is locally written as $\Delta = v^i \partial/\partial v^i$.

Using the tensor $S$ the geometry of Euler–Lagrange equations (3.27) becomes apparent. Multiplying them by $dq^i$ we get,

$$dq^i \left( \frac{d}{dt} \frac{\partial L}{\partial v^i} - \frac{\partial L}{\partial q^i} \right) = 0,$$

and using that $d/dt(dq^i) = dv^i$ we obtain,

$$\frac{d}{dt} \left( dq^i \frac{\partial L}{\partial v^i} \right) - dL = 0.$$

Then, because

$$\Theta_L = dL \circ S,$$  

(3.30)
THE GEOMETRY OF DYNAMICS

we get as in the proof of Thm. 1,

$$\mathcal{L}_\Gamma \Theta_L = dL,$$

(3.31)

which is the intrinsic form of Euler–Lagrange equations. If $\Gamma$ is a SODE, then $S(\Gamma) = \Delta$ and developing the Lie derivative in (3.31) we get

$$i_\Gamma d\Theta_L = d(L - i_\Gamma \Theta_L),$$

and because $i_\Gamma \Theta_L = \Delta(L)$, we get the Lagrangian version of the dynamical equation (2.16),

$$i_\Gamma \omega_L = dE_L.$$

It is obvious from what it has been said that if $\omega_L$ is nondegenerate, $\Gamma$ is a second order differential equation whose solutions satisfy Euler–Lagrange equations of motion. Most of the previous construction can be generalized to higher order Lagrangian dynamical systems as in [Le85].

We must remark here that $\omega_L$ can be degenerate. Lagrangians such that $\omega_L$ is not symplectic are called singular\(^{10}\). They are commonplace in physical theories [Ha73], [Di64] and their analysis introduces a whole new field of research in the modeling of physical systems, its main characteristic is the nonexistence a priori of a dynamical characterization of the evolution of the system (i.e., generically there is no initial value problem because they lead to implicit differential equations) [Ca90], [Ca86]. The analysis of such systems involve the introduction of different algorithms to construct a well-posed evolution system. Such algorithms usually consist in the recursive construction of a family of submanifolds on which the system has a better behaviour and eventually they define a submanifold where the system has a genuine evolution description. These submanifolds are defined by constraint functions\(^{11}\) that have to be carefully analyzed [Go80], [Ma83], [He92], [Tu95].

Further research in the geometrical description of Euler–Lagrange equations has led to the important observation that there exists a unique nonlinear connection $\nabla$ associated to any SODE $\Gamma$ [Cr83]. We shall recall that if $L_0$ denotes the kinetic energy Lagrangian function defined by a Riemannian metric $g$ on $Q$,

$$L_0 = \frac{1}{2} g_{ij} v^i v^j,$$

\(^{10}\)The inverse problem for some systems of equations arising from singular Lagrangians has been discussed in [Ib91].

\(^{11}\)These constraints arise from compatibility conditions on the dynamical content of the theory and they have not to be mistaken with the study of mechanical systems with holonomic or nonholonomic constraints imposed externally.
then, the corresponding Euler-Lagrange’s equations define the geodesic spray of $g$ which is the SODE,

$$
\Gamma_0 = v^i \frac{\partial}{\partial q^i} - \Gamma^k_{ij} v^i v^j \frac{\partial}{\partial v^k},
$$

(3.32)

where $\Gamma^k_{ij}$ are the Christoffel symbols of the Levi–Civita connection defined by $g$. In this particular case the connection associated to $\Gamma_0$ is nothing else but the linear Levi-Civita connection of the metric $g$. It was Crampin’s remark that not only the SODE’s of the form (3.32) define connections but any SODE defines a (nonlinear) connection [Cr95]. In the same way as the Levi–Civita connection has been used to investigate Riemannian geometry Crampin’s nonlinear connection is being used to investigate the properties of SODE’s [CMI].

4. The Action Functional and Hamilton’s Principle

We turn now our attention to the alternative viewpoint sketched in §1 regarding the geometrization programme for dynamical systems. Instead of modeling the initial conditions of the system we will try to model the space of trajectories. To show how we can do that, we will start with the simplest possible situation. We will consider the first order differential equation defined on the parameter space $M$ given by,

$$
\dot{x}^i = f^i(x, t).
$$

(4.33)

We want to characterize its integral curves with some prescribed properties, for instance, we will impose that they will start at the point $x_0$ at time $t_0$ and end at $x_1$ at time $t_1$ ($> t_0$). We will define the path space $\Omega(x_0, x_1; M)$ as the space of $C^1$ curves\footnote{For analytical reasons it is convenient sometimes to use piecewise $C^1$–differentiable curves or even better curves of Sobolev class 1.} $\gamma: [t_0, t_1] \to M$ such that $\gamma(t_0) = x_0$ and $\gamma(t_1) = x_1$. The curves $\gamma$ such that for each $t$ they satisfy eq. (4.33) form a subspace of $\Omega(x_0, x_1; M)$. In fact, defining $\Phi^i_t$ as the family of functions on $\Omega(x_0, x_1; M)$, $i = 1, \ldots, n$, $t \in [t_0, t_1]$ by

$$
\Phi^i_t(\gamma) = \dot{\gamma}^i(t) - f^i(\gamma(t), t),
$$

then, the curves we are looking for are precisely the set $\cap_{t \in [t_0, t_1]} \Phi^{-1}(0)$. We can describe this set using a variational principle, i.e., as the critical set of a function defined in $\Omega(x_0, x_1; M)$. To achieve this in the simplest possible
way we introduce Lagrange multipliers \( \xi_i(t) \) and define the functional \( S \) on the space of paths \( \Omega(x_0, x_1; M) \) as,

\[
S_\xi(\gamma) = \int_{t_0}^{t_1} \xi_i(t) \left( \dot{\gamma}^i(t) - f^i(\gamma(t), t) \right) dt. \tag{4.34}
\]

It is obvious that the Lagrange multipliers \( \xi_i \) must transform as covectors in \( M \) in order to define a scalar inside the integral in (4.34). Moreover they can vary in an arbitrary way as functions of \( t \), then the previous functional is actually defined in the space of curves on \( T^*M \) with endpoints lying over \( x_0 \) and \( x_1 \) respectively, i.e.,

\[
\Omega(x_0, x_1; T^*M) = \{ \sigma: [t_0, t_1] \to T^*M \mid \sigma(t) = (x(t), \xi(t)), \ x(t_0) = x_0, x(t_1) = x_1 \},
\]

and

\[
S(\sigma) = \int_{t_0}^{t_1} \xi_i(t) \left( \dot{x}^i(t) - f^i(x(t), t) \right) dt.
\]

In fact, the critical points of \( S \) are not exactly the solutions of eq. (4.33) as expected but they contain more (geometrical) information. Computing \( \delta S(\sigma) \) we obtain:

\[
\delta S(\sigma) = \int_{t_0}^{t_1} \left[ \xi_i(t) \delta(\dot{x}^i(t) - f^i(x(t), t)) + (\dot{x}^i(t) - f^i(x(t), t)) \delta \xi_i(t) \right] dt
\]

\[
= \int_{t_0}^{t_1} \left[ \xi_i(t) \frac{d}{dt} \delta x^i(t) - \xi_i(t) \frac{\partial f^i}{\partial x^j}(x(t), t) \delta x^j(t) + (\dot{x}^i(t) - f^i(x(t), t)) \delta \xi_i(t) \right] dt
\]

\[
= \int_{t_0}^{t_1} \left[ \left( -\dot{\xi}_i(t) - \xi_j(t) \frac{\partial f^j}{\partial x^i}(x(t), t) \right) \delta x^i(t) + (\dot{x}^i(t) - f^i(x(t), t)) \delta \xi_i(t) \right] dt,
\]

and then, the critical points of \( S \) are given by the solutions of the system of ODE's

\[
\dot{x}^i = f^i(x(t), t); \quad \dot{\xi}_i = -\frac{\partial f^j}{\partial x^i} \xi_j, \tag{4.35}
\]

with boundary conditions \( x(t_0) = x_0, x(t_1) = x_1 \) and \( \xi(t_0), \xi(t_1) \) free.

We must point out that the set of solutions of eq. (4.35) is mapped onto the set of solutions of eq. (4.33). This map is defined simply by taking the \( x \)-component of the solutions \( (x(t), \xi(t)) \) of (4.35). It is in this sense that we have reformulated the problem (4.33) with boundary conditions as a variational problem.
An important observation is that the set of equations (4.35) are actually Hamilton’s equations on $T^*M$ for a (time dependent) Hamiltonian which is linear on the momenta $\xi$’s,

$$H_f(x, \xi) = \xi_i f^i(x, t).$$

It is also remarkable that the functional $S$ is built using the symplectic geometry of $T^*M$. In fact, we can rewrite $S$ as

$$S = \int_{t_0}^{t_1} (\xi_i \dot{x}^i - \xi_i f^i) dt = \int_{t_0}^{t_1} \Theta_0 - H_f dt,$$

where $\Theta_0 = \xi_i dx^i$ is the canonical Liouville 1–form on $T^*M$. The vector field on $T^*M$ defined by eqs. (4.35) is the cotangent lifting $X^*$ of the vector field $X = f^i \partial / \partial x^i$ on $M$. Obviously we can replace $H_f$ by an arbitrary Hamiltonian function on $T^*M$ and the functional $S$ before by the action functional

$$S(\sigma) = \int_x \Theta_0 - \int_{t_0}^{t_1} H(\sigma(t), t) dt. \quad (4.36)$$

Then we can state:

**Theorem 2.** (Hamilton’s principle) The integral curves of Hamilton’s equations (2.15) with endpoints $x_0$ and $x_1$ are the critical points of the action functional (4.36) on the space of paths $\Omega(x_0, x_1; T^*M)$. Moreover, the periodic solutions of Hamilton’s equations are the critical points of the action functional on the loop space of $T^*M$.

This main result constitute Hamilton’s principle and provides the direct link between the initial value geometry and global modeling of dynamical systems. All classical textbooks on Mechanics include a detailed description of variational principles [Gl50], [Su74] (see [Yo68] for a historical account of Hamilton’s principle).

4.1. Legendre Transformation As we discussed in §3 there is a particularly meaningful set of Hamilton’s equations, the so called SODEs obtained from a regular Lagrangian function. A similar analysis as the one carried out in the previous subsection leading to Hamilton’s principle, Theor. 2, can be repeated to show that the solutions to the SODE’s equations

$$\dot{x}^i = v^i; \quad \dot{v}^i = f^i(x, v), \quad (4.37)$$
with Lagrangian $L$ are the critical points of the action functional

$$S_L(\sigma) = \int_{t_0}^{t_1} \Theta_L - \int_{t_0}^{t_1} E_L \, dt,$$

(4.38)
on the space of paths $\Omega(x_0, x_1; TM)$. An straightforward analysis shows that the critical points of the functional $S_L$ in (4.38) are given by solutions of the equations

$$\frac{\partial^2 L}{\partial \dot{x}^i \partial \dot{x}^j} (v^j - \dot{x}^j) = 0; \quad \frac{\partial^2 L}{\partial v^i v^j} \dot{v}^i + \frac{\partial^2 L}{\partial v^i x^j} \dot{x}^j = \frac{\partial L}{\partial x^i}.$$

Thus, if $L$ is regular, we get $v^i = \dot{x}^i$ and the Euler–Lagrange equations follows. Then we can restrict our attention to curves satisfying the second order condition $v^i = \dot{x}^i$, i.e., to curves on $M$ lifted naturally to $TM$. Then, the action functional $S_L$ becomes

$$S_L(\gamma) = \int_{t_0}^{t_1} L(\gamma(t), \dot{\gamma}(t), t) \, dt,$$

which is the well-known action functional of Lagrangian mechanics [La60], [Ge63], [Ar76]. It is in this form that the action functional is generalized to field theories even though its geometrical content is partially lost.

It is noticeable that if we fix the energy of the system $E_L = c$, which is a constant of the motion for time-independent systems, we can restrict the space of paths to those contained in the level set of energy $c$. Then, the variational principle will become Maupertuis principle

$$S_L(\sigma) = \int_{t_0}^{t_1} \Theta_L.$$

The previous ideas can be thought as the translation to the velocity phase space $TM$ of the results in the phase space $T^*M$ using the identification $p_i = \partial L / \partial v^i$. This map is called the Legendre transformation and will be denoted by $F_L$. If the Legendre transformation is a diffeomorphisms it allows to use both descriptions, Lagrangian and Hamiltonian, for a given Lagrangian dynamical system. Lagrangian functions such that this happens are called hyperregular. If the Lagrangian $L$ is regular it is easy to see that $F_L$ is a local diffeomorphism. In general it is not hard to see that $F_L^* \omega_0 = \omega_L$ and $F_L^* \Theta_0 = \Theta_L$, and (locally) $F_L \Gamma = X^H$ with $E_L = F_L^* H$.

\footnote{The Legendre transformation can also be defined starting from the Hamiltonian formalism as $v^i = \partial H / \partial p_i$ and gives rise to the same kind of considerations as the Legendre transformation $F_L$.}
4.2. DIRECT METHODS  Once we have obtained a variational characterization of solutions of Hamilton’s equations one can get the wrong impression that it will be straightforward to compute them because they are simply the critical points of a smooth function. In fact this idea has been pursued many times in the last two hundred years and even if the way is paved with extreme difficulties, there has been big successes. Nowadays this is an intensive field of research that keeps providing wonderful surprises.

We will just mention some of the main ideas involved in the direct analysis of the critical points of the action functional.

- Morse theory. It was envisioned by M. Morse as a way of proving the existence of periodic geodesics in compact Riemannian manifolds. The main idea of the theory consists in linking the critical points to the topology of the underlying manifold. This idea turns to be extremely deep and succesful and it leads naturally to a remarkable set of inequalities called Morse inequalities [Mo65], [Mi69].

- Lusternik-Schnirelmann theory. Similar to Morse theory, it uses a topological invariant of a manifold, called the category of the manifold, to obtain information on the existence of critical points of functions on it [Li66].

- Minimax techniques. A very deep idea introduced by Birkhoff to characterize the eigenvalues of a quadratic form. It provides extremely powerful tools to prove the existence of critical values of a function (see the recent book [Wi89] and references therein).

- Floer’s theory. A modern refinement of Morse theory used by A. Floer to prove the existence of periodic orbits of Hamilton’s equations on compact symplectic manifolds. It was inspired by ideas coming from S. Smale, J. Milnor, and E. Witten [Fi89] and was used to partially prove a conjecture by V. Arnold on the minimum number of periodic orbits of Hamiltonian systems on compact symplectic manifolds.

5. INVARIANCE IN DYNAMICS. CONSTANTS OF THE MOTION AND REDUCTION

We will finish this overview of some of the main ideas entering the geometrical modeling of dynamics by sketching the use of symmetries to investigate properties and solutions of differential equations.

It is intuitive clear that the existence of symmetries of dynamical systems must be helpful to understand its solutions, this idea comes at least from S. Lie [Li88]. What it is not clear at all is why symmetries gives rise to numerical
invariants. We will see immediately that the energy paradigm provides the link between numerical invariants (or constants of the motion) and symmetries of dynamical systems.

Noether's theorem

\[ \text{SYMMETRIES} \iff \text{CONSTANTS OF THE MOTION} \]

Energy paradigm

Symmetries versus constants of the motion

Let \( X \) be a vector field on a manifold \( P \) modeling a dynamical system. A symmetry of the system is a transformation \( \phi: P \to P \) leaving invariant \( X \), i.e., \( \phi_* X = X \circ \phi \), or in other words, if \( \psi_t \) denotes the flow of \( X \) then \( \psi_t \circ \phi = \phi \circ \psi_t \), and \( \phi \) sends solutions of \( X \) into solutions of \( X \). Infinitesimally, a symmetry of \( X \) is a vector field \( Y \) such that \( [X, Y] = 0 \). The flow \( \phi_s \) of \( Y \) is made of symmetries of \( X \), \( \phi_s \circ \psi_t = \psi_t \circ \phi_s \).

On the other hand a numerical invariant of the dynamical system \( X \) is a function \( F \) defined on \( P \) such that is constant along the evolution of the system, i.e., \( F \circ \psi_t = F \). This is equivalent to \( X(F) = 0 \).

If we restrict our attention to dynamical systems for which the energy paradigm applies, we can assume that our manifold \( P \) has a Poisson bracket and \( X \) is a Hamiltonian vector field with Hamiltonian \( H \). If we have a transformation \( \phi \) of the theory and we ask it to respect the energy paradigm, this implies that it must send any Hamiltonian vector field into another Hamiltonian vector field. Transformations satisfying this property have not been characterized for general Poisson manifolds. For nondegenerate Poisson structures, it is the content of Lee-Hwa-Chung theorem that such transformations must be conformal symplectic [Le47], [Go84], [Li88].

We will assume in what follows that the symmetry transformations we are considering are Poisson, and that their infinitesimal generators will be represented by Hamiltonian vector fields. Even if this is not the most general situation is relevant enough in the sense that most of the known applications fall within this hypothesis. Then, we can summarize the previous discussion stating that in what follows by an infinitesimal symmetry of the dynamical system \( X \) we will mean a Hamiltonian vector field \( Y = X_F \) such that \( X(F) = 0 \).

**Theorem 3.** (Noether's theorem [No71]) \(^{14}\) There is a one-to-one corre-

\(^{14}\)Usually for historical reasons, Noether's theorem is established in the Lagrangian setting. It is an straightforward exercise to translate this theorem to the geometrical setting described in §3.
spondence among infinitesimal symmetries and constants of the motion of the Hamiltonian vector field $X_H$.

The proof of Noether’s theorem is a straightforward consequence of the following identities and simply reflects the skewsymmetry of $\{\cdot, \cdot\}$,

$$Y(H) = X_F(H) = \{F, H\} = -\{H, F\} = -X_H(F),$$

thus the l.h.s. of the previous equation is zero iff the r.h.s. vanishes, but this is equivalent to $F$ being a constant of the motion and the former condition is equivalent to $Y$ being an infinitesimal symmetry of the system.

We can go further and ask about the set of all symmetries of $X_H$. It is easy to check that they form a Lie algebra. Computing $[[Y_1, Y_2], X_H]$ and using Jacobi’s identity,

$$[[Y_1, Y_2], X_H] = -[[Y_2, X_H], Y_1] + [[Y_1, X_H], Y_2] = 0.$$}

5.1. THE MOMENTUM MAP This Lie algebra of infinitesimal symmetries can be finite or infinite dimensional. Suppose either that it is finite dimensional or that we are just interested in a finite dimensional subalgebra of it. Let us denote it by $\mathfrak{g}$ and let $Y_i$, $i = 1, \ldots, r$ be a set of generators of $\mathfrak{g}$. As before we will assume that the symmetry generators $Y_i$ are Hamiltonian vector fields $X_{J_i}$ for some constants of the motion $J_i$. Combining together this constants of the motion we construct a map $J = (J_1, \ldots, J_r)$ that can be described intrinsically in a better way as a map $J : P \to \mathfrak{g}^*$,

$$\langle J(m), Y_i \rangle = J_i(m); \ \forall m \in P, \ \forall i = 1, \ldots, r.$$

The map $J$ is called the momentum map of the symmetry algebra $\mathfrak{g}$ acting on the Poisson manifold $P$, and it indeed generalizes the notion of angular momentum and linear momentum in mechanical systems [So70], [Ar76], [Ab78], [Ma85], [Mr94].

As it was pointed out in the beginning of this section that the existence of symmetries can help us in solving the problem under consideration. The momentum map $J$ indicates how to do that. Note that the momentum map is a constant of the motion, thus if we take a trajectory of our dynamical system with initial data $m$, it always remains in the level set of the value $\mu = J(m)$. On the other hand on the level set $J^{-1}(\mu)$, there is still a residual symmetry of $\mathfrak{g}$. In fact, notice that the commutation relations $\text{Ad}_{Y_i}(Y_j) = [Y_i, Y_j] = c_{ij}^k Y_k$ induce an action of $\mathfrak{g}$ on $\mathfrak{g}^*$ called the coadjoint action and defined by

$$\text{Ad}^*_{Y_i} \theta^j = -c_{ij}^k \theta^k,$$
where \( \theta^i \) denotes the dual basis of \( Y_i \) on \( \mathfrak{g}^* \), i.e., \( \langle \theta^i, Y_i \rangle = \delta_i^j \). Then there is a subalgebra \( \mathfrak{g}_\mu \) leaving invariant \( \mu \) that consists of the elements \( Y \in \mathfrak{g} \) such that \( \text{Ad}_{Y^*} \mu = 0 \). This algebra is called the isotropy algebra of \( \mu \) and under rather general conditions it leaves invariant the level set \( \mu \) of the momentum map.

To be more specific, we can rephrase the previous discussion as follows. There is a unique connected and simply connected Lie group \( G \) possessing \( \mathfrak{g} \) as a Lie algebra. The group \( G \) will act in a natural way on \( P \) and it also has a natural action on \( \mathfrak{g}^* \) by the coadjoint action. If the second cohomology group of \( G \) is trivial then the momentum map \( J \) is equivariant with respect to both actions, i.e.,

\[
J(g \cdot m) = \text{Ad}_g^* J(m); \quad \forall g \in G, m \in P.
\]

The same is true if \( G \) is compact (a general discussion about this can be found in [Mr94] and references therein). The other crucial property of \( J \) is that under the conditions above, \( J \) is a Poisson map with respect to the canonical linear Poisson structure existing on \( \mathfrak{g}^* \) which is precisely the linear Poisson structure defined by eq. (2.11).

Denoting by \( G_\mu \) the subgroup of \( G \) whose Lie algebra is \( \mathfrak{g}_\mu \) we have that
$X_H$, which is contained in $J^{-1}(\mu)$, is invariant under $G_\mu$ and thus it projects to the quotient space $P_\mu = J^{-1}(\mu)/G_\mu$ called the Marsden–Weinstein reduced phase space which will be assumed to be a smooth manifold and the canonical projection $J^{-1}(\mu) \to P_\mu$ a submersion (see figure 2.).

**Theorem 4.** (Symplectic reduction theorem [Mr74]) Under the conditions above\(^\text{15}\), the reduced phase space is a Poisson manifold and the Hamiltonian vector $X_H$ induces a vector field on the reduced phase space which is also Hamiltonian.

For symplectic manifolds, it is not hard to prove that the restriction of $\omega$ to the level set $J^{-1}(\mu)$ is degenerate and its kernel is spanned at each point by the tangent vectors along the orbits of $G_\mu$. Thus, this form projects to $P_\mu$ and the projection is nondegenerate and closed, hence symplectic. On the other hand the projected vector field induced by $X_H$ is obviously hamiltonian with respect to the projected symplectic form. □

Examples of this are abundant and a vast literature has been devoted to explore and discuss its implications\(^\text{16}\) and applications like relative equilibrium and stability, plasma and fluid dynamics, etc. [Mr94].

**Acknowledgements**

The author AI wishes to acknowledge the partial financial support provided by CICYT under the programme PB92-0197 as well as the NATO collaborative research grant 940195.

**References**


\(^\text{15}\)These conditions can be somewhat relaxed even to include singular values of the momentum map.

\(^\text{16}\)An important implication of the previous theorem has been obtained by M. Gotay and G. Tuynman showing that essentially all symplectic manifolds can be obtained from $\mathbb{R}^{2n}$, for some $n$, by a (generalized) symplectic reduction procedure [Go89].


