Anales Acad. Nac. de Cs. Ex., Fís. y Nat., tomo 64 (2012): 117-156.

THE DIRAC THEORY OF CONSTRAINTS, THE GOTAY-NESTER THEORY AND POISSON GEOMETRY

Hernán Cendra [a], María Etchechoury [b] and Sebastián J. Ferraro [c]

[a] Departamento de Matemática
Universidad Nacional del Sur, Av. Alem 1253
8000 Bahía Blanca, and CONICET, Argentina.
uscendra@criba.edu.ar
[b] Laboratorio de Electrónica Industrial, Control e Instrumentación,
Facultad de Ingeniería, Universidad Nacional de La Plata and
Departamento de Matemática, Facultad de Ciencias Exactas, Universidad Nacional de La Plata.
CC 172, 1900 La Plata, Argentina.
marila@mate.unlp.edu.ar
[c] Departamento de Matemática and Instituto de Matemática Bahía Blanca
Universidad Nacional del Sur, Av. Alem 1253
8000 Bahía Blanca, and CONICET, Argentina
sferraro@uns.edu.ar

Abstrac

The Dirac theory of constraints has been widely studied and applied very successfully by physicists since the original works by Dirac and by Bergmann. From a mathematical standpoint, several aspects of the theory have been exposed rigorously afterwards by many authors. However, many questions related to, for instance, singular or infinite dimensional cases remain open. The work of Gotay and Nester presents a mathematical generalization in terms of presymplectic geometry, which introduces a dual point of view. We present a study of the Dirac theory of constraints emphasizing the duality between the Poisson-algebraic and the geometric points of view, related respectively to the work of Dirac and of Gotay and Nester, under strong regularity conditions. We deal with some questions insufficiently treated in the literature: a study of uniqueness of solution; avoiding almost completely the use of coordinates; the role of the Pontryagin bundle. We also show how one can globalize some results usually treated locally in the literature. For instance, we introduce the *global* notion of *second class submanifold* as being tangent to a *second class subbundle*. A general study of global results for Dirac and Gotay-Nester theories remains an open question in this theory.

Keywords: Dirac's theory of constraints, presymplectic manifolds, Poisson geometry.

Resumen

La Teoría de ligaduras de Dirac, la teoría de Gotay-Nester y geometría de Poissin. La teoría de Dirac ha sido ampliamente estudiada y aplicada muy exitosamente por los físicos desde los trabajos originales de Dirac y de Bergmann. Desde un punto de vista matemático, varios aspectos de la teoría han sido expuestos rigurosamente por varios autores. Sin embargo, aún quedan abiertas varias preguntas relacionadas, por ejemplo, con casos singulares o infinito-dimensionales. El trabajo de Gotay y Nester presenta una generalización matemática en términos de la geometría presimpléctica, lo cual introduce un punto de vista dual. Presentamos un estudio de la teoría de ligaduras de Dirac enfatizando la dualidad entre los puntos de vista de las álgebras de Poisson y de la geometría presimpléctica, relacionados respectivamente con los trabajos de Dirac y de Gotay-Nester, bajo condiciones de regularidad fuertes. Abordamos algunas cuestiones insuficientemente tratadas en la literatura: un estudio de la unicidad de solución; evitar casi completamente el uso de coordenadas; el rol del fibrado de Pontryagin. También mostramos cómo se pueden globalizar algunos resultados usualmente tratados localmente en la literatura. Por ejemplo, introducimos la noción *global* de *subvariedad de segunda clase* como variedad tangente a un *subfibrado de segunda clase*. Un estudio general de resultados globales para las teorías de Dirac y de Gotay-Nester sigue siendo una pregunta abierta en esta teoría.

Palabras clave: teoría de ligaduras de Dirac, variedades presimplécticas, geometría de Poisson.

* Académico Correspondiente de la Academia Nacional de Ciencias Exactas, Físicas y Naturales

1. Introduction

References, preliminaries and a description of the main works related to the present paper will be given in section 2. Here we will give a brief overview. The Dirac theory of constrained Hamiltonian systems was written by Dirac in terms of the canonical Poisson brackets in the space of classical observables (functions on the phase space) avoiding the notion of *constraint submanifolds*, which, on the other hand, is naturally present in the theory. For instance, instead of introducing the *final constraint submanifold*, as we do, the notion of *weak equality* of functions is preferred. This indicates a deliberate decision of Dirac to study the Poisson-algebraic aspect of constrained Hamiltonian systems, a point of view that is appropriate for quantization, which gave a very successful theory.

However, from the point of view of classical mechanics the states are points in phase space and the constraint submanifolds also play an important role in understanding the geometry of the equations of motion and solution curves. This point of view also suggests, for instance, that one can view a constrained Hamiltonian system as an IDE (Implicit Differential Equation). Then the Dirac algorithm, as well as several questions about dynamics like the existence of solutions for the initial condition problem, has a meaning also in the context of IDEs, which is an active field of study.

In many papers after Dirac's work, for instance the work of Gotay and Nester, cited in the next section, the geometric side of his theory has been developed and proved very useful. The geometric side is, in a sense, dual to the algebraic side and this duality is apparent in the commutative diagrams of subsection 3.2 where geometric diagrams have a Poisson-algebraic counterpart. This algebro-geometric possibility of approaching questions is present throughout the paper. Even though there are many beautiful works emphasizing the Poisson-algebraic or the geometric aspects our main references will be the works of Dirac and those of Gotay and Nester, respectively.

The notion of Dirac structure (Courant and Weinstein [10], Courant [11], Bursztyn and Crainic [1]) is originated in part in Dirac's work and gives a new possibility to understand and extend the theory. The present paper should be followed soon by a generalization in the realm of Dirac geometry.

In section 2 we review the Gotay-Nester and the Dirac algorithms and define the notion of secondary constraint submanifolds, in particular, the final constraint submanifold.

Both the primary and the final constraint submanifolds are important for writing equations of motion.

In subsection 3.2 we perform a careful study of the primary and final constraint submanifolds and various quotient manifolds and commutative diagrams, which helps to understand important aspects of the dynamics, for instance, the question of uniqueness of solutions. The latter is related to the notion of *physical variables*. We also study the dual point of view using the Poisson algebra of first class constraints and various quotients and commutative diagrams, which shows the duality between the geometric and the Poisson-algebraic points of view. In subsection 3.3 we show that the second class constraint submanifolds are submanifolds of the phase space tangent to a second class vector subbundle along the final constraint submanifold, which may lead to a classification, at least in some examples, of second class constraint submanifolds modulo tangency. The second class vector subbundle carries enough information to write the Dirac bracket at points of the final constraint submanifold. We write equations of motion in terms of the Dirac bracket in subsection 3.4.

Another feature of our work is that notions such as second class constraint submanifolds and Dirac brackets, or at least their restriction to them, are defined globally (in the sense explained in lemma 3.19 and theorem 3.20), and notions such as the standard Dirac bracket appear as coordinate expressions of a global object.

Along this paper we assume strong regularity conditions that lead, for instance, to the fact that the Dirac bracket is locally constant (in coordinates) which can be established using the Weinstein splitting theorem.

2. Constraint Algorithms

Implicit Differential Equations. We now briefly review some basic results concerning general IDEs and constraint algorithms. We do this just because we find useful to realize that some aspects and concepts of Dirac's theory are of a more general nature, not necessarily related to mechanics or Poisson geometry. Let M be a given differentiable manifold. An IDE on M, written as

$$\varphi(x,\dot{x}) = 0, \qquad (1)$$

of which ODEs $\dot{x} - f(x) = 0$ or algebraic equations $\phi(x) = 0$ are considered trivial particular cases, appear naturally in science and technology. A *solution of (1) at a point* x is a vector $(x,\dot{x}) \in T_x M$ satisfying (1). A *solution curve*, say $x(t)$, $t \in (a,b)$, must satisfy, by definition, that $(x(t),\dot{x}(t))$ is a solution at $x(t)$ for all $t \in (a,b)$. In the local case, M is an open set of R^n and the IDE can be written equivalently in the form

$$\begin{aligned} \dot{x} &= u \\ 0 &= \varphi(x,u), \end{aligned}$$

that is,

$$A(y) \cdot \dot{y} = f(y),$$

where $y=(x,u)$ and $f(y)=(u,\varphi(x,u))$. This is by definition a *quasilinear IDE*.

Basic questions such as existence, uniqueness or extension of solutions are not completely answered yet, although many partial results in this direction have been established. One of the common features of those results is that they show, at least under certain regularity conditions, how to transform, using a certain *constraint algorithm*, a given IDE into an equivalent parameter dependent ODE on a certain *final constraint manifold*.

Very briefly, the idea consists in finding a decreasing sequence of constraint submanifolds $M \supseteq M_1 \supseteq \cdots \supseteq M_c$, which appears naturally by imposing the condition of existence of a solution $(x,\dot{x}) \in TM_k$ to the given IDE at a each point $x \in M_{k+1}$. Under certain additional conditions, like locally constant rank conditions, the original IDE is reduced to an equivalent ODE depending on parameters on the *final constraint submanifold* M_c, which, by construction, has the fundamental property that it must contain all solutions curves of the given IDE.

In spite of the simplicity of the general algorithm, there are fundamental examples where extra meaningful structures are used to build the submanifolds M_k and to write the equations of motion. Moreover, in the Dirac approach the sequence of submanifolds M_k is not emphasized and the methods of Poisson geometry are used, for good reasons.

We can compare Dirac [15], Gotay et al. [21], Pritchard [34], Rabier and Rheinboldt [35], Cendra and Etchechoury [7], to see how the idea of the algorithm works in different contexts. In Cendra and Etchechoury [7], one works in the realm of subanalytic sets; in Gotay et al. [21] one works with presymplectic manifolds; in Pritchard [34] one works with complex algebraic manifolds; Dirac [15] uses Poisson brackets; in Rabier and Rheinboldt [35] some degree of differentiability of the basic data is assumed, and, besides, some constant rank hypothesis is added, essentially to ensure applicability of some constant rank theorem. Some relevant references for general IDEs connected to physics or control theory, which show a diversity of geometric or analytic methods or a combination of both are Cariñena and Rañada [5], de León and Martín de Diego [12], Delgado-Téllez and Ibort [14], Gràcia and Pons [22, 23], Ibort et al. [25], Marmo et al. [30], Mendella et al. [31].

In the present paper we will concentrate on the Dirac and the Gotay-Nester points of view (which represent the algebraic and the geometric side), see Dirac [15, 16, 17] and Gotay et al. [21]. One may say that some aspects of Dirac's idea have been nicely formalized and generalized in Gotay et al. [21] in the context of presymplectic geometry on reflexive Banach manifolds. Both the algebraic and the geometric aspects of Dirac's theory have been treated by many people, cited below, with different ideas. The Dirac algorithm is not the same as the Gotay-Nester algorithm although the two methods are essentially equivalent in fundamental examples, like degenerate Lagrangian systems, as shown in Gotay et al. [21]. The Dirac algorithm provides *explicit* equations of motion written in terms of the canonical bracket of the ambient symplectic manifold and a *total Hamiltonian* depending on parameters. Besides, the Dirac approach yields the *Dirac bracket* defined in a neighborhood of the final constraint submanifold. Equations of motion written in terms of the Dirac bracket are specially simple and elegant, as we will see. On the other hand, the IDE obtained on the final constraint submanifold by the Gotay-Nester algorithm does not depend on any parameters or an embedding in a symplectic manifold.

General assumptions. All manifolds involved will be finite-dimensional smooth manifolds and all maps will be smooth, unless otherwise specified. Several arguments in this paper are of a local character, but they can be regarded as coordinate versions of global results. For instance, this is the case for the notion of second class constraints, which represent a second class constraint submanifold.

The Gotay-Nester and the Dirac constraint algorithms studied in this paper can be considered as particular ways of writing the general constraint algorithm for (quasilinear) IDEs mentioned above, using the special structure available in each case (presymplectic structures, symplectic and Poisson structures, respectively). Therefore, the sequence of secondary constraints M_k, $k=1,\ldots$ is the same for all these algorithms. It is important to have a criterion to ensure that this sequence stops. We will assume that each M_{k+1} is a closed submanifold of M_k defined by equations. Also, we assume that for each k and each $x \in M_{k+1}$, if $\dim_x M_k = \dim_x M_{k+1}$, then $x \in M_{k+p}$ and $\dim_x M_k = \dim_x M_{k+p}$ for all $p \in \mathrm{N}$. This implies that the sequence stabilizes, that is there is some c such that $M_c = M_{c+p}$, for all $p \in \mathrm{N}$.
For each M_k one has the corresponding k-th IDE, which can be written as

$$A_k(x) \cdot \dot{x} = f_k(x),$$

where $A_k(x) = A(x) \mid T_x M_k$ and $f_k = f \mid M_k$. A point x will be in M_{k+1}, by definition, iff this equation has a solution $(x, \dot{x})$. We will assume throughout the paper that the rank of $A_k(x)$ is locally constant on M_{k+1}. This implies that the rank of $A_c(x)$ is locally constant on M_c, and on each point of M_c there is at least one solution that is tangent to M_c. The main property of M_c from the dynamical point of view is that every solution curve to the original system must lie on M_c. Since we are assuming the locally constant rank condition, the final system on M_c can be converted, at least locally, into a parameter-dependent family of ODEs.

2.1 A brief review of Dirac's theory

Dirac's theory of constraints has been extensively studied from many different points of view and extended in many directions. An important part of those developments is contained in Cantrijn et al. [2], Cariñena et al. [3, 6], Cariñena [4], Cariñena and Rañada [5], de León et al. [13], Gotay and Nester [18, 19], Henneaux and Teitelboim [24], Ibort et al. [26], Krupková [27, 28], Marmo et al. [30], Mukunda [32, 33], Skinner [36], Skinner and Rusk [37, 38], Sudarshan and Mukunda [40], van der Schaft [41].

There is a certain duality between the Dirac approach, in which the role of constraints as being functions on the phase space and the canonical bracket is essential, and the approach of many other authors, starting with Gotay and Nester, where, in addition, the geometry behind the canonical Poisson algebra on phase space is emphasized. This duality is present along this paper, and in this sense, our main references will be Dirac [17] and Gotay et al. [21].

We will recall some essential aspects of the Dirac theory of constraints, following Dirac [17], but using a more modern language, adapted to our purposes, and assuming explicitly certain regularity conditions.

Dirac's theory starts with a given singular Lagrangian system $L:TQ\to \mathrm{R}$, since in the case of a regular Lagrangian the theory becomes trivial. Then, in order to construct a Hamiltonian theory one must consider from the beginning the image of the Legendre transformation, which may be a very complicated subset $M_0\subseteq T^*Q$. By definition, solution curves to the Hamiltonian system must be exactly the image of solutions to Euler-Lagrange equations under the Legendre transformation. Natural questions such as existence of solution curves to the Euler-Lagrange equations for a given initial condition are not completely solved in general, to the best of our knowledge. In order to obtain this kind of results one would need to choose mathematically precise hypotheses, a topic not considered in Dirac's work.

In this paper we will assume a general hypothesis about ***regularity***, under which this kind of problem is easier. Regularity means, among other hypotheses to be established along the paper as they are needed, that certain sets $M_0\supseteq M_1\cdots$ are *submanifolds* of T^*Q defined regularly by equations $\phi_i^{(k)}=0$, $i=1,\ldots,a_k$, $k=0,1,\ldots$. The $\phi_i^{(k)}$ are functions defined on T^*Q constructed by the ***Dirac constraint algorithm*** and called ***constraints***. The submanifolds M_k are called the ***constraint submanifolds***.

As usual, one assumes that the algorithm stops for $k=c$. One also assumes that the ranks of the matrices $(\{\phi_i^{(c)},\phi_j^{(c)}\}(x))$ and $(\{\phi_i^{(c)},\phi_j^{(0)}\}(x))$ are locally constant on the final constraint submanifold M_c.

Each $\phi_i^{(0)}$ is called a ***primary constraint*** and each $\phi_i^{(c)}$ is called a ***final constraint***. Generically, $\phi_i^{(k)}$, $k=1,\ldots,c$, are called ***secondary constraints***. The main property of the final constraint submanifold M_c is that any motion of the classical particle, that is, any solution $(q(t),p(t))$, must remain in M_c, and Dirac shows how to write Poisson equations of motion in terms of position and momentum using the canonical Poisson bracket on T^*Q and the ***total Hamiltonian*** H_T. For given initial conditions belonging to M_c, solutions are not necessarily unique and Dirac interprets this fact as being due to the ***nonphysical character*** of some of the variables. Of fundamental importance for Dirac's theory, especially for quantization, are the classification of constraints into ***first class*** and ***second class*** in terms of certain commutation relations, and the construction of a Poisson bracket $\{,\}^*$ called the ***Dirac bracket***. An important result is that with respect to the Dirac bracket all final constraints $\phi_i^{(c)}$ appear to be first class constraints, in other words, $\{\phi_i^{(c)},\phi_j^{(c)}\}^*(x)=0$, for all $x\in M_c$. Dirac's procedure also shows how to deal with the nonphysical variables and find the correct notion of state of the system. One shows that there are ***physically meaningful variables*** in terms of which the evolution for a given initial state is determined. This is important from the classical and also from the quantum mechanics point of view.

Now, we will be more precise. The image of the Legendre transformation $FL:TQ\to T^*Q$, that is, $M_0=FL(TQ)$, contained in the canonical symplectic manifold T^*Q, is *assumed* to be defined by equations $\phi_i^{(0)}=0$, $i=1,\ldots,a_0$, where each $\phi_i^{(0)}:T^*Q\to \mathrm{R}$ is a *primary constraint*, by definition.

In the case in which the Legendre transformation is degenerate the Hamiltonian $H: T^*Q \to \mathrm{R}$ is not uniquely defined from the formulas $pv - L(q,v) = H(q,p)$, $p = \partial L(q,v)/\partial v$, but in Dirac's theory one assumes that such a function can be conveniently defined on M_0 (which can be done in examples using ideas akin to the Pontryagin maximum principle, like the fact that for each $(q,p) \in M_0$ the derivative of $pv - L(q,v)$ with respect to v is 0) and then extended, more or less arbitrarily, to T^*Q. Then one defines, following Dirac, the *total Hamiltonian* $H_T = H + \lambda^i_{(0)}\phi^{(0)}_i$, with arbitrary parameters $\lambda^i_{(0)}$ to be determined. The *Dirac constraint algorithm* goes as follows. The preservation of the primary constraints is written $\{\phi^{(0)}_i, H_T\}(x) = 0$, $i = 1,\ldots,a_0$, $x \in M_0$, or

$$\{\phi^{(0)}_i, H\}(x) + \lambda^j_{(0)}\{\phi^{(0)}_i, \phi^{(0)}_j\}(x) = 0,\ i,j = 1,\ldots,a_0,\ x \in M_0.$$

Then M_1 is defined by the condition that $x \in M_1$ if and only if there exist $\lambda_{(0)} = (\lambda^1_{(0)},\ldots,\lambda^{a_0}_{(0)})$ such that the system of equations $\phi^{(0)}_i(x) = 0$, $\{\phi^{(0)}_i, H_T\}(x) = 0$, $i = 1,\ldots,a_0$, is satisfied. Clearly, $M_0 \supseteq M_1$, and one *assumes* that M_1 is a submanifold regularly defined by equations, say, $\phi^{(1)}_i = 0$, $i = 1,\ldots,a_1$, where each $\phi^{(1)}_i$ is a *secondary constraint*, by definition. By proceeding iteratively one obtains a sequence $M_0 \supseteq M_1 \supseteq \cdots$, and we will *assume* that this sequence stops. Then there are *final constraints*, say $\phi^{(c)}_i$, $i = 1,\ldots,a_c$, defining regularly a (nonempty by assumption) submanifold M_c by equations $\phi^{(c)}_i = 0$, $i = 1,\ldots,a_c$, called the *final constraint submanifold*, and the following condition is satisfied. For each $x \in M_c$ there exists $(\lambda^1_{(0)},\ldots,\lambda^{a_0}_{(0)})$ such that

$$\{\phi^{(c)}_i, H\}(x) + \lambda^j_{(0)}\{\phi^{(c)}_i, \phi^{(0)}_j\}(x) = 0,\ i = 1,\ldots,a_c,\ j = 1,\ldots,a_0. \tag{2}$$

For each $x \in M_c$ the space of solutions of the linear system of equations (2) in the unknowns $\lambda^j_{(0)}$ is an affine subspace of R^{a_0}, called $S^{(c)}_x$ whose dimension is a locally constant function $d^{(c)}(x) = a_0 - \mathrm{rank}(\{\phi^{(c)}_i, \phi^{(0)}_j\}(x))$. One can locally choose $d^{(c)}(x)$ unknowns as being free parameters and the rest will depend affinely on them. Then the solutions of (2) form an affine bundle $S^{(c)}$ over M_c. After replacing $\lambda_{(0)} \in S^{(c)}$ in the expression of the total Hamiltonian, the corresponding Hamiltonian vector field,

$$X_{H_T}(x) = X_H(x) + \lambda^j_{(0)} X_{\phi^{(0)}_j}(x),$$

$x \in M_c$, which will depend on the free unknowns, will be tangent to M_c. Its integral curves, for an arbitrary choice of a time dependence of the free unknowns, will be solutions of the equations of motion, which is the main property of the final constraint submanifold M_c from the point of view of classical mechanics. The lack of uniqueness of solution for a given initial condition in M_c, given by the presence of free parameters, indicates, according to Dirac, the nonphysical character of some of the variables. In our context the physical variables can be defined on a quotient manifold.

Remark. Dirac introduces the notion of weak equality for functions on T^*Q. Two such functions are ***weakly equal***, denoted $f \approx g$, if $f \mid M_c = g \mid M_c$. Then, for instance $\phi_j^{(k)} \approx 0$. If $f \approx 0$ then $f = v^i \phi_i^{(c)}$, for some functions v^i on T^*Q and conversely. Since we have introduced the notion of a constraint submanifold, in particular the final constraint submanifold, we prefer not to use the notation $\approx$.

Now let us make some comments on the notions of first class and second class constraints. The rank of the matrix $(\{\phi_i^{(c)}, \phi_j^{(c)}\}(x)), i, j = 1, \ldots, a_c$, is necessarily even, say, $2s$, and it is assumed to be constant. Then, using elementary properties of determinants (like adding to a row or column a linear combination of the other rows or columns) one can find, at least locally in a neighborhood of each point $x \in M_c$, functions ψ_i, $i = 1, \ldots, a_c - 2s$, and χ_j, $j = 1, \ldots, 2s$, such that the equations $\psi_i = 0$, $\chi_j = 0$, define M_c regularly and, besides, $\{\psi_i, \psi_{i'}\}(x) = 0$, $\{\psi_i, \chi_j\}(x) = 0$, $\det(\{\chi_j, \chi_{j'}\}(x)) \neq 0$, for $i, i' = 1, \ldots,$ $a_c - 2s$, $j, j' = 1, \ldots, 2s$ and $x \in M_c$. In fact, we will assume that this is can be done globally, for simplicity. The $\phi_j^{(c)}$ are linear combinations with smooth coefficients of the χ_j and ψ_i, and conversely. The functions χ_j, $j = 1, \ldots, 2s$, are called ***second class constraints*** and the functions ψ_i, $i = 1, \ldots, a_c - 2s$, are called ***first class constraints***.

More generally, any function ρ on T^*Q satisfying $\rho \mid M_c = 0$, $\{\rho, \psi_i\} \mid M_c = 0$, $\{\rho, \chi_j\} \mid M_c = 0$, is a first class constraint with respect to the submanifold M_c, by definition. Any function g on T^*Q satisfying $\{g, \psi_i\} \mid M_c = 0$, $\{g, \chi_j\} \mid M_c = 0$, is a ***first class function***, by definition. For instance, the total Hamiltonian H_T is a first class function.
Now define the Hamiltonian h_c in terms of ψ_i, χ_j, $i = 1, \ldots, a_c - 2s$, $j = 1, \ldots, 2s$, as

$$h_c = H + \lambda^i \psi_i + \mu^j \chi_j.$$

The preservation of the constraints for the evolution generated by h_c can be rewritten as $\{\psi_i, h_c\}(x) = 0$, which is equivalent to $\{\psi_i, H\}(x) = 0$ for all $x \in M_c$, and $\{\chi_j, h_c\}(x) = 0$, for all $x \in M_c$. The latter is equivalent to

$$\{\chi_i, H\}(x) + \mu^j \{\chi_i, \chi_j\}(x) = 0, \ i, j = 1, \ldots, 2s,$$

for all $x \in M_c$, which determines the μ^j as well-defined functions on M_c. Then the solutions $(\mu(x), \lambda)$ form an affine bundle with base M_c and whose fiber, parametrized by the free parameters λ, has dimension $a_c - 2s$.

Any section $(\mu(x), \lambda(x))$ of this bundle determines h_c as a first class function. This means that $X_{h_c}(x) \in T_x M_c$, for each $x \in M_c$, and therefore a solution curve of X_{h_c} is contained in M_c provided that the initial condition belongs to M_c. We will show in this section that there is a symplectic manifold $\bar{M}_c$ such that one can pass to the quotient $M_c \to \bar{M}_c$ and also h_c passes to the quotient $\bar{h}_c$ in such a way that solution curves of h_c become solutions curves of the Hamiltonian $\bar{h}_c$. Moreover, we will show that there is a manifold $\tilde{M}_c$ and natural maps $M_c \to \tilde{M}_c \to \bar{M}_c$ such that the Hamiltonian h_c passes to the quotient to a function $\tilde{h}_c$ on $\tilde{M}_c$. The *extended Hamiltonian* defined by Dirac is related to $\tilde{h}_c$. One can show that $\tilde{h}_c$ passes to the quotient via $\tilde{M}_c \to \bar{M}_c$ to the function $\bar{h}_c$ defined above.

Dirac defines an interesting bracket, now called the ***Dirac bracket***,

$$\{F,G\}^* = \{F,G\} - \{F,\chi_i\}c^{ij}\{\chi_j,G\},$$

which is defined on an open set in T^*Q containing M_c, where c^{ij}, which by definition is the inverse matrix of $\{\chi_i,\chi_j\}$, is defined. The Dirac bracket is a Poisson bracket and has the important property that, for *any* function F on T^*Q, the condition $\{F,\chi_j\}^* = 0$, $j = 1,\ldots,2s$, is satisfied on a neighborhood of M_c, which implies that $\dot{F} = \{F,h_c\} = \{F,h_c\}^*$, for any function F. Besides, $\{\psi_j,\psi_i\}^* = 0$, $i,j = 1,\ldots,a_c - 2s$, on M_c. Because of this, one may say that, with respect to the Dirac bracket, all the constraints χ_j, $j = 1,\ldots,2s$ and ψ_i, $i = 1,\ldots,a_c - 2s$, are first class with respect to M_c. This is important for purposes of quantization.

2.2 A brief Review of the Gotay-Nester Theory

In this section we recall some aspects of the Gotay-Nester theory which we need. This theory was developed in Gotay et al. [21] to deal geometrically with the Dirac-Bergmann theory of constraints. The main equation studied is an IDE of the type

$$\mathrm{i}_{\dot{x}}\omega(x) = \alpha(x), \qquad (3)$$

where ω is a closed 2-form on a manifold M and $\alpha \in \Omega^1(M)$ is a closed 1-form on M. As we have indicated before this kind of equation appears naturally in classical Lagrangian mechanics, in fact, we will show later that the Euler-Lagrange equations can be rewritten equivalently in the form

$$\mathrm{i}_{\dot{x}}\omega(x) = d\mathrm{E}(x), \qquad (4)$$

which is clearly of the type (3).

Description of the Gotay-Nester Algorithm. As we mentioned in the Introduction, in order to deal with IDEs one can apply a basic idea which consists in building a sequence of constraint submanifolds.

Let us first describe that basic approach for a system like (3) without using explicitly the presymplectic form, and later on we will briefly explain how the presymplectic form can be used to write equations for the constraint submanifolds explicitly. The latter is an important contribution of the Gotay-Nester algorithm.

We want to find solution curves to (3). Let $x(t)$ be such a solution curve; then for each t the *linear algebraic system*

$$\mathrm{i}_{v(t)}\omega(x(t)) = \alpha(x(t)),$$

has at least one solution, namely, $v(t) = \dot{x}(t)$. This implies that, for each t, $x(t)$ must belong to the subset

$$M_1 = \{x \in M \mid \mathrm{i}_v\omega(x) = \alpha(x) \text{ has at least one solution } v \in T_xM\}.$$

Assume, as in Gotay et al. [21], that M_1 is a submanifold of M. Since $x(t) \in M_1$ for all t we must have that $\dot{x}(t) \in T_{x(t)}M_1$ for all t. This implies that, for each t, $x(t)$ must belong to the subset

$$M_2 = \{x \in M_1 \mid \mathrm{i}_v\omega(x) = \alpha(x) \text{ has at least one solution } v \in T_xM_1\}.$$

We can continue in a similar way and define M_{k+1} recursively as

$$M_{k+1} = \{x \in M_k \mid \mathrm{i}_v\omega(x) = \alpha(x) \text{ has at least one solution } v \in T_xM_k\}.$$

This sequence stabilizes, under the *General Assumptions* described at the beginning of this section. Under the assumption that the map

$$\omega^{\flat} : T_xM_c \to T^*M \mid M_c$$

has *locally constant rank* on the *final constraint manifold* M_c, existence of local solution curves to (3) for each initial condition in M_c is guaranteed. For given local coordinates $(x_1,\ldots,x_m)$ on M_c and for a given initial condition $x_0 \in M_c$ one can fix some appropriate coordinates as functions of t, say $x_1(t),\ldots,x_r(t)$, where $r = \dim\ker\omega$, in a neighborhood of x_0 and then solve (3) uniquely for $x_{r+1}(t),\ldots,x_m(t)$. More precisely, in local coordinates our equation becomes

$$\omega_{ij}(x)\cdot\dot{x}^j = \alpha_i(x), \qquad (5)$$

$i = 1,\ldots,\dim M$, $j = 1,\ldots,m$. Since $r = \dim\ker\omega$, in a neighborhood of a given point x_0 one can solve, after relabeling the coordinates if necessary, $x_{r+k} = x_{r+k}(x_1,\ldots,x_r)$, $k = 1,\ldots,m-r$. After choosing arbitrarily the curves $x_1(t),\ldots,x_r(t)$ and replacing in (5) one obtains a time-dependent

ODE in the remaining variables $x_{r+1},\ldots,x_m$. We can also interpret the previous arguments by saying that the implicit differential equation

$$\omega(x)(\dot{x},) = \alpha(x), x \in M_c$$

is an ODE on M_c, depending on r parameters.

The following geometric description will be useful later on. Since by assumption r does not depend on $x \in M_c$, at least locally, then the equation on M_c

$$\omega(x)(X,) = \alpha(x), \qquad (6)$$

where $X \in TM_c$, defines an affine distribution on M_c of locally constant rank. More precisely, one has an affine bundle $S^{(c)}$ with base M_c whose fiber $S_x^{(c)}$ at a given point $x \in M_c$ is, by definition,

$$S_x^{(c)} = \{X \in T_x M_c \mid (6) \text{ is satisfied}\}. \qquad (7)$$

Remark. (a) If rank $\omega(x)$ is not locally constant we still have a distribution $S^{(c)}$ on M_c, but it may be singular. The analysis of existence of solution curves in this case may be difficult, see Cendra and Etchechoury [7], Pritchard [34] and references therein. The algorithm developed in Cendra and Etchechoury [7] for a general system of the type $a(x)\cdot\dot{x} = f(x)$, with analytic data, represents an improvement of the previous basic ideas also in the sense that the final system obtained after applying the algorithm *always* has locally constant rank, and that singular cases are also studied using desingularization methods.

(b) In Gotay et al. [21] it is explained how solutions can be expressed using brackets, as in Dirac's work.

Example. Let $L : TQ \to \mathrm{R}$ be a Lagrangian, degenerate or not. Since the problem is of a local nature we can use local coordinates. Let $\mathrm{E}(q,v,p) = pv - L(q,v)$ and let $\omega \in \Omega^2(TQ \oplus T^*Q)$ be the presymplectic form $\omega = dq^i \wedge dp_i$ on the Pontryagin bundle $M = TQ \oplus T^*Q$. Then Euler-Lagrange equations are written equivalently in the form of equation (4) with $x = (q,v,p)$. In fact, we have

$$\mathrm{i}_{(\dot{q},\dot{v},\dot{p})} dq^i \wedge dp_i = \dot{q}^i dp_i - \dot{p}_i dq^i \qquad (8)$$

$$d\mathrm{E} = \frac{\partial \mathrm{E}}{\partial q^i} dq^i + \frac{\partial \mathrm{E}}{\partial p_i} dp_i + \frac{\partial \mathrm{E}}{\partial v^i} dv^i \qquad (9)$$

$$= -\frac{\partial L}{\partial q^i}(q,v) dq^i + v^i dp_i + \left(p_i - \frac{\partial L}{\partial v^i} \right) dv^i \qquad (10)$$

Using equations (8)–(10) we can easily see that (4) is equivalent to

$$\begin{aligned}
\dot{q}^i &= v^i \\
\dot{p}_i &= \frac{\partial L}{\partial q^i}(q,v) \\
0 &= p_i - \frac{\partial L}{\partial v^i}(q,v),
\end{aligned}$$

which is clearly equivalent to the Euler-Lagrange equations. The idea of using the Pontryagin bundle to write important equations of physics like Euler-Lagrange or Hamilton's equations appears in Cendra et al. [9], Livens [29], Skinner [36], Skinner and Rusk [37, 38], Yoshimura and Marsden [43, 44].

Describing the Secondary Constraints Using ω. The constraint manifolds M_k defined by the algorithm can be described by *equations written in terms of the presymplectic form* ω, which is a simple but important idea. Depending on the nature of ω one may obtain analytic, smooth, linear, etc., equations, which may simplify matters in given examples. This idea is also important in the context of reflexive Banach manifolds, as remarked in Gotay et al. [21]. Besides, those equations will obviously be invariant under changes of coordinates preserving ω.

The condition defining the subsets M_{k+1}, $k = 0,1,\ldots$ (calling $M_0 = M$ to uniformize the notation) namely,

$$\mathrm{i}_v\omega(x) = \alpha(x) \text{ has at least one solution } v \in T_xM_k,$$

is equivalent to $\alpha(x) \in \left(T_xM_k\right)^{\text{ç}}$. Since $\left(T_xM_k\right)^{\text{ç}} = \left(\left(T_xM_k\right)^{\omega}\right)^{\circ}$, we have

$$M_{k+1} = \{x \in M_k \mid \langle \alpha(x), \left(T_xM_k\right)^{\omega}\rangle = \{0\}\}.$$

3 Main Results

3.1 Preliminaries

We will need the following results about linear symplectic geometry which are an essential part of many of the arguments in our treatment of Dirac and Gotay-Nester theories. This is because under our strong regularity assumptions those theories are, to a certain extent, linear.

Lemma 3.1. *Let* (E,Ω) *be a symplectic vector space of dimension* $2n$, $V \subseteq E$ *a given subspace. For a given basis* α_i, $i = 1,\ldots,r$ *of* V°, *let* $X_i = \alpha_i^{\#}, i = 1,\ldots r$. *Then the rank of the matrix* $[\alpha_i(X_j)]$ *is even, say* $2s$, *and* $X_i, i = 1,\ldots,r$ *form a basis of* V^{Ω}. *Moreover, the basis* $\alpha_i, i = 1,\ldots,r$ *can be chosen such that for all* $j = 1,\ldots,r$

$$\begin{aligned}
&\alpha_i(X_j) = \delta_{i,j-s},\ 1 \le i \le s \\
&\alpha_i(X_j) = -\delta_{i-s,j},\ s+1 \le i \le 2s \\
&\alpha_i(X_j) = 0,\ 2s < i \le r.
\end{aligned}$$

Proof. Consider the subspace $V^{\Omega} = (V^{\circ})^{\#}$. By a well-known result there is a basis $X_i, i = 1,\ldots,r$ of V^{Ω} such that for all $j = 1,\ldots,r$

$$\Omega(X_i, X_j) = \delta_{i,j-s},\ 1 \le i \le s$$
$$\Omega(X_i, X_j) = -\delta_{i-s,j},\ s+1 \le i \le 2s$$
$$\Omega(X_i, X_j) = 0,\ 2s < i \le r$$

then take $\alpha_i = X_i^{\flat}$. The first part of the lemma is easy to prove using this.

Lemma 3.2. *Let* α_i, $i = 1,\ldots,r$ *be a basis of* V° *having the properties stated in Lemma 3.1. Then* X_i, $i = 2s+1,\ldots,r$ *form a basis of* $V \cap V^{\Omega}$.

Proof. Let $X = \lambda^i X_i$ be an arbitrary vector in V^{Ω}. Now $\lambda^i X_i \in V \cap V^{\Omega}$ iff $\alpha_j(X) = \lambda^i(\alpha_j(X_i)) = 0, j = 1,\ldots,r$. Since the first $2s$ columns of the matrix $[\alpha_i(X_j)]$ are linearly independent and the rest are zero, we must have $\lambda^i = 0$, for $1 \le i \le 2s$, and λ^i, $i = 2s+1,\ldots,r$ are arbitrary. This means that $V \cap V^{\Omega}$ is generated by X_i, $i = 2s+1,\ldots,r$.

Corollary 3.3. $\dim V \cap V^{\Omega} = r - 2s$.

Proof. Immediate from lemma 3.2.

Let ω be the pullback of Ω to V via the inclusion. Then (V,ω) is a presymplectic space. In what follows, the $^{\flat}$ and $^{\#}$ operators are taken with respect to Ω unless specified otherwise.

Lemma 3.4. $V^{\omega} = V \cap V^{\Omega}$.

Proof. $X \in V^{\omega}$ iff $\omega(X,Y) = 0, \forall Y \in V$ iff $\Omega(X,Y) = 0, \forall Y \in V$. This is equivalent to $X \in V \cap V^{\Omega}$.

Lemma 3.5. *Let* $\gamma_i, i = 1,\ldots r$ *be a given basis of* V° *and let* $Y_i = \gamma_i^{\#}, i = 1,\ldots r$. *Let* $\beta \in E^*$ *be given. Then the following conditions are equivalent.*

(i) $\beta(V^{\omega}) = 0$.

(ii) *The linear system*

$$\beta(Y_i) + \lambda^j \gamma_j(Y_i) = 0 \qquad (11)$$

has solution $\lambda = (\lambda^1,\ldots,\lambda^r)$.

Proof. Let us show that (11) has solution $(\lambda^1,\ldots,\lambda^r)$ iff the system

$$\beta(X_k) + \mu^l \alpha_l(X_k) = 0 \qquad (12)$$

has solution $(\mu^1,\ldots,\mu^r)$, where $k,l=1,\ldots r$ and α_l is a basis satisfying the conditions of lemma 3.1. Since $Y_i, i=1,\ldots r$ and $X_k, k=1,\ldots r$ are both bases of V^Ω there is an invertible matrix $[A_k^i]$ such that $X_k = A_k^i Y_i$. Let $[B_i^l]$ be the inverse of $[A_k^i]$, so $Y_i = B_i^l X_l$. Assume that (11) has solution $\lambda^j, j=1,\ldots r$. We can write (11) as

$$\beta(Y_i)+\lambda^j\Omega(Y_j,Y_i)=0,\ i=1,\ldots,r.$$

Using this we have that for $k=1,\ldots,r$

$$\begin{aligned}0 &= \beta(A_k^iY_i)+\lambda^j\Omega(Y_j,A_k^iY_i)=\beta(X_k)+\lambda^j\Omega(Y_j,X_k)\\ &= \beta(X_k)+\lambda^j\Omega(B_j^lX_l,X_k)=\beta(X_k)+\mu^l\Omega(X_l,X_k)\end{aligned}$$

where $\mu^l=\lambda^jB_j^l$. This means that the system (12) has solution. The converse is analogous. Using this, lemmas 3.2 and 3.4, and the form of the coefficient matrix $[\alpha_l(X_k)]$ in lemma 3.1, the proof that (12) has solution $\mu=(\mu^1,\ldots,\mu^r)$ iff $\beta(V^\omega)=0$ is easy and is left to the reader. □

Lemma 3.6. *Consider the hypotheses in lemma 3.5. Then the solutions to*

$$i_X\omega=\beta\,|\,V \qquad (13)$$

(if any) are precisely $X=\beta^\#+\lambda^jY_j$*, where* $(\lambda^1,\ldots,\lambda^r)$ *is a solution to (11). A solution to (13) exists if and only if* $\beta(V^\omega)=0$*. If* ω *is symplectic then (11) and (13) have a unique solution and if, in addition,* $\beta^\#\in V$*, then* $\lambda^1=0$*, ...,* $\lambda^r=0$ *and* $\beta^\#$ *coincides with* $X=(\beta\,|\,V)^{\#_\omega}$ *defined by (13).*

Proof. Since Y_j, $j=1,\ldots,r$ form a basis of V^Ω we have that $(\lambda^1,\ldots,\lambda^r)$ is a solution to (11) iff $(\beta+\lambda^j\gamma_j)(V^\Omega)=0$ iff $\beta+\lambda^j\gamma_j\in V^\circ$ iff $\beta^\#+\lambda^jY_j\in V$. Now, let $X=\beta^\#+\lambda^jY_j$, where $(\lambda^1,\ldots,\lambda^r)$ satisfies (11). Then we have $X\in V$ as we have just seen and we also have

$$i_X\omega=(i_X\Omega)\,|\,V=X^\flat\,|\,V=(\beta+\lambda^j\gamma_j)\,|\,V=\beta\,|\,V,$$

since $\gamma_j, j=1,\ldots,r$ generate V°. We have proven that X is a solution to (13). To prove that every solution X to (13) can be written as before, we can reverse the previous argument. Using this, it is clear that if ω is symplectic then (11) has unique solution, in particular, we have that $\det(\gamma_j(Y_i))\neq 0$. If, in addition, $\beta^\#\in V$ then $\lambda^jY_j=X-\beta^\#\in V$. Since Y_j, $j=1,\ldots,r$ is a basis of V^Ω, using lemma 3.4 and the fact that $V^\omega=\{0\}$ we get that $\lambda^j=0$ for $j=1,\ldots,r$.

Corollary 3.7. *Let* $\Lambda=\{\lambda \mid \lambda \text{ satisfies } (11)\}$. *Then* $\dim\Lambda = r-2s = \dim\ker\omega$.

Proof. $\ker\omega = V^{\omega}$, which has dimension $r-2s$ from corollary 3.3 and lemma 3.4. On the other hand the dimension of the subspace of λ satisfying (11) is clearly also $r-2s$, since the coefficient matrix has rank $2s$.

3.2 A Poisson-Algebraic and Geometric Study of the Primary and Final Constraint Submanifolds

The Dirac algorithm, briefly explained in the previous subsection, can be applied to any given ***constrained Hamiltonian system*** (P,Ω,H,M) where (P,Ω) is a symplectic manifold, the primary constraint submanifold M is a given submanifold of P defined regularly by an equation $\phi=0$ and H is a Hamiltonian defined on P. This is because the particular cotangent bundle structure of the symplectic manifold T^*Q is not essentially used in the Dirac algorithm.

For instance, an interesting variant of the Dirac algorithm for a degenerate Lagrangian system is the following. Consider the canonical symplectic manifold $N=T^*TQ$ with the canonical symplectic form Ω, and let the primary constraint be $M=TQ\oplus T^*Q$, canonically embedded in N via the map given in local coordinates (q,v,p,ν) of N by $\varphi(q,v,p)=(q,v,p,0)$. In particular, M is defined regularly by the equation $\nu=0$. If ω is the presymplectic form on M obtained by pulling back the canonical symplectic form of T^*Q, then $\varphi^*\Omega=\omega$. This embedding is globally defined (see Appendix for details).

Remark. For a given presymplectic manifold (M,ω) one can always find an embedding φ into a symplectic manifold (P,Ω) such that $\varphi^*\Omega=\omega$. Moreover, this embedding can also be chosen such that it is coisotropic, meaning that M is a coisotropic submanifold of P (see Gotay and Sniatycki [20]). However, we should mention that the embedding given above is not coisotropic.

The number pv is a well-defined function on M and it can be naturally extended to a function on a chart with coordinates (q,v,p,ν), but this does not define a function on N consistently. In any case, it can be extended to a smooth function on N and any such extension will give the same equations of motion. The Dirac theory of constraints is essentially a local theory. However, we will see a global version of the notion of Dirac bracket, in a sense, as well as its local descriptions.

Consider the function $\mathrm{E}: M\to\mathrm{R}$ given by $\mathrm{E}=pv-L(q,v)$. Using the fact that E can be extended naturally on a chart with coordinates (q,v,p,ν) and taking an appropriate partition of unity we can choose once and for all an extension to a smooth function E on N called the ***Energy***. Then we can apply the Dirac algorithm to the constrained Hamiltonian system (N,Ω,E,M).

There are some interesting features in this approach, as compared with the one described in subsection 2.1, where the symplectic manifold is T^*Q and the primary constraint is the image of the Legendre transformation. For instance, the primary constraint $M = TQ \oplus T^*Q$ is well defined in a natural way as a closed submanifold of the symplectic manifold N. Besides, the comparison with the Gotay-Nester approach becomes clear from the beginning and the Euler-Lagrange equations are derived quickly as a differential-algebraic equation (DAE). On the other hand, this approach may have the disadvantage of introducing the extra variable v, which may lead to longer calculations in some examples.

We shall start with the constrained Hamiltonian system $(N, \Omega, \mathrm{E}, M)$, and we will work locally, for simplicity. We will call $\phi^i_{(0)}$, $i = 1, \ldots, r_0$, the primary constraints v^i defining $M_0 \equiv M$ regularly by equations.

We will emphasize the Gotay-Nester point of view and we will see how it combines with the Dirac procedure.

Accordingly, we shall study the Dirac dynamical system on the manifold $M = TQ \oplus T^*Q$, already considered in the example at the end of subsection 2.2 where the Dirac structure is associated to a presymplectic form ω which is the pullback of the symplectic form Ω on $N = T^*TQ$ via the inclusion $M \subseteq N$. Then M is the *primary constraint*. Then the equation to be solved, according to the Gotay-Nester algorithm, is the equation

$$\omega(x)(X,) = d\mathrm{E}(x) \,|\, T_x M, \qquad (14)$$

where $X \in T_x M_c$ and $x \in M_c$, M_c being the *final constraint*. Let ω_c be the pullback of Ω via the inclusion of M_c in N. Since ω_c is presymplectic, $\ker \omega_c$ is an involutive distribution.
From now on we will assume the following.

Assumption K_1. The distribution $\ker \omega_c$ has constant rank and defines a regular foliation K_c, that is, the natural map $p_{K_c} : M_c \to \bar{M}_c$, where $\bar{M}_c = M_c / K_c$ is a submersion.

Lemma 3.8. *The following assertions hold:*

(a) *There is a uniquely defined symplectic form* $\bar{\omega}_c$ *on* $\bar{M}_c$ *such that* $p^*_{K_c} \bar{\omega}_c = \omega_c$.

(b) *Let* $\bar{X}$ *be a given vector field on* $\bar{M}_c$. *Then there is a vector field* X *on* M_c *that is* p_{K_c}*-related to* $\bar{X}$.

(c) *Let* $\bar{f} \in \mathcal{F}(\bar{M}_c)$. *Then there exists a vector field* X *on* M_c *such that* X *is* p_{K_c}*-related to* $X_{\bar{f}}$, *and for any such vector field* X *the equality* $\omega_c(x)(X,) = d(p^*_{K_c} \bar{f})(x)$ *holds for all* $x \in M_c$.

(d) *Let* $X_{x_0} \in T_{x_0} M_c$. *Then one can choose the function* $\bar{f} \in \mathrm{F}(\bar{M}_c)$ *and the vector field* X *in* ***(c)*** *in such a way that* $X(x_0) = X_{x_0}$.

Proof. **(a)** By definition, the leaves of the foliation K_c are connected submanifolds of M_c, that is, each $p_{K_c}^{-1}(z)$, $z \in \bar{M}_c$, is connected. For $z \in \bar{M}_c$, let $x \in M_c$ such that $p_{K_c}(x) = z$. For $\bar{A}, \bar{B} \in T_z \bar{M}_c$, as p_{K_c} is a submersion, there are $A, B \in M_c$ such that $T_x p_{K_c} A = \bar{A}, T_x p_{K_c} B = \bar{B}$. We define $\bar{\omega}_c(z)(\bar{A}, \bar{B}) = \omega_c(x)(A, B)$. To prove that this is a good definition observe first that it is a consistent definition for fixed x, which is easy to prove, using the fact that $\ker \omega_c(x) = \ker T_x p_{K_c}$. Now choose a Darboux chart centered at x, say $U \times V$, such that, in this chart, $p_{K_c} : U \times V \to U$ and $\omega_c(x^1, x^2) = \bar{\omega}_c(x^1)$, where $\omega_c(x^1, x^2)$ and $\bar{\omega}_c(x^1)$ are independent of (x_1, x_2). This shows that $\bar{\omega}_c$ is well defined on the chart. Using this and the fact that one can cover the connected submanifold $p_{K_c}^{-1}(z)$ with charts as explained above, one can deduce by a simple argument that $\bar{\omega}_c(z)$ is well defined.

(b) Let g be a Riemannian metric on M_c. Then for each $x \in M_c$ there is a uniquely determined $X(x) \in T_x M_c$ such that $X(x)$ is orthogonal to $\ker T_x p_{K_c}$ and $T_x p_{K_c} X(x) = \bar{X}(x)$, for all $x \in M_c$. This defines a vector field X on M_c which is p_{K_c}-related to $\bar{X}$.

(c) Given $\bar{f}$ and using the result of **(b)** we see that there is a vector field X on M_c that is p_{K_c}-related to $X_{\bar{f}}$. Then, for every $x \in M_c$ and every $Y_x \in T_x M_c$,

$$\begin{aligned} \omega_c(x)(X(x), Y_x) &= \bar{\omega}_c(p_{K_c}(x))\left(X_{\bar{f}}(p_{K_c}(x)), T_x p_{K_c} Y_x\right) = d\bar{f}(p_{K_c}(x))(T_x p_{K_c} Y_x) \\ &= d(p_{K_c}^* \bar{f})(x)(Y_x). \end{aligned}$$

(d) One can proceed as in **(b)** and **(c)** and choosing $\bar{f}$ such that $\left(d\bar{f}\left(p_{K_c}(x_0)\right)\right)^{\#} = T_{x_0} p_{K_c} X_{x_0}$ and, besides, the metric g such that X_{x_0} is perpendicular to $\ker T_{x_0} p_{K_c}$. □

Definition 3.9. (a) For any subspace $A \subseteq \mathrm{F}(N)$ define the distribution $\Delta_A \subseteq TN$ by $\Delta_A(x) = \{X_f(x) \mid f \in A\}$.

(b) The space of ***first class functions*** is defined as

$$R^{(c)} = \{f \in \mathrm{F}(N) \mid X_f(x) \in T_x M_c, \text{ for all } x \in M_c\}.$$

In other words, $R^{(c)}$ is the largest subset of $\mathrm{F}(N)$ satisfying

$$\Delta_{R^{(c)}}(x) \subseteq T_x M_c,$$

$x \in M_c$.

Remark. **(a)** From the point of view of classical mechanics, the *constraint submanifolds* M and M_c seem to be at least as important as the functions $\phi_i^{(0)}$ and $\phi_i^{(c)}$ (called *constraints* by Dirac) defining them by equations $\phi_i^{(0)} = 0$ and $\phi_i^{(c)} = 0$, respectively. Dirac was interested in classical mechanics, where states are points in phase space, as well as in quantum mechanics where functions are observables and states are not points in phase space. In the present paper we focus mainly in classical mechanics, and therefore we need to concentrate on the constraint submanifolds. In particular, M and M_c are the only ones that play an important role. The other secondary constraints submanifolds seem to be less important.

(b) The total Hamiltonian H_T is a first class function, by construction.

Lemma 3.10. ***(a)*** *$R^{(c)}$ is a Poisson subalgebra of $(\mathcal{F}(N), \{,\})$.*

(b) *M_c is an integral submanifold of $\Delta_{R^{(c)}}$. Moreover, for any vector field X on M_c that is p_{K_c}-related to a vector field $X_{\bar{f}}$ on $\bar{M}_c$ there exists a function $f \in R^{(c)}$ such that $f \mid M_c = p_{K_c}^* \bar{f}$ and $X = X_f \mid M_c$. In particular, any vector field X on M_c satisfying $X(x) \in \ker \omega_c(x)$ for all $x \in M_c$ is p_{K_c}-related to the vector field 0 on the symplectic manifold $\bar{M}_c$, which is associated to the function $\bar{f} = 0$, therefore there exists a function $f \in R^{(c)}$, which satisfies $f \mid M_c = 0$, such that $X(x) = X_f(x)$, $x \in M_c$.*

Proof. **(a)** Let $f, g \in R^{(c)}$. Then $X_f(x)$ and $X_g(x)$ are both tangent to M_c at points x of M_c which implies that $-X_{\{f,g\}}(x) = [X_f, X_g](x)$ is also tangent to M_c at points of x of M_c. This shows that $\{f, g\} \in R^{(c)}$. It is easy to see that any linear combination of f, g and also fg belong to $R^{(c)}$.

(b) By definition $\Delta_{R^{(c)}} \subseteq TM_c$. We need to show the converse inclusion. Let $X_{x_0} \in T_{x_0} M_c$, we need to find $f \in R^{(c)}$ such that $X_f(x_0) = X_{x_0}$. Choose the function $\bar{f}$ and the vector field X on M_c as in lemma 3.8, **(d)**. Choose any extension of $p_{K_c}^* \bar{f}$ to a function g on N. For each $x \in M_c$, we can apply lemmas 3.5 and 3.6 with $E := T_x N$, $V := T_x M_c$, $\beta := dg(x)$, $\gamma_i(x) := d\phi_i^{(c)}(x)$, $i = 1, \ldots, r_c$. We obtain that in a neighborhood $U \subseteq N$ of each point x_0 of M_c we can choose C^∞ functions $\lambda_{(c)}^i(x)$, $i = 1, \ldots, r_c$, such that

$$X(x) = \left(dg(x) + \lambda_{(c)}^i(x) d\phi_i^{(c)}(x) \right)^{\#},$$

for all $x \in M_c \cap U$. Let $f_U(x) = g(x) + \lambda^i_{(c)}(x)\phi^{(c)}_i(x)$, for all $x \in U$. Then we have that $X_{f_U}(x) = X(x)$, for all $x \in M_c \cap U$.

Now, consider a partition of unity ρ_i, $i \in I$, on N, where each ρ_i is defined on an open set U_i, $i \in I$. Let $J \subseteq I$ be defined by the condition $i \in J$ if and only if $U_i \cap M_c \neq \varnothing$. Using standard techniques of partitions of unity and the above result one can assume without loss of generality that for each $i \in J$ there is a function f_{U_i} defined on U_i such that $X_{f_{U_i}}(x) = X(x)$, for all $x \in M_c \cap U_i$. Let $f = \sum_{i \in J} \rho_i f_{U_i}$, which can be naturally extended by 0 on N. Then it is easy to see, using the fact that $f_{U_i}(x) = p^*_{K_c}\bar{f}(x) = g(x)$, for each $x \in M_c$, that $X(x) = X_f(x)$, for each $x \in M_c$, and in particular $X_0 = X(x_0) = X_f(x_0)$. □

Lemma 3.11. ***(a)*** *Each function* $f \in R^{(c)}$ *is locally constant on the leaves of* K_c *therefore, since they are connected, for each* $f \in R^{(c)}$ *there is a uniquely determined* $\bar{f} \in \mathcal{F}(\bar{M}_c)$, *called* $(p_{K_c})_* f$, *such that* $f \mid M_c = p^*_{K_c}\bar{f}$. *Moreover, the vector fields* $X_f(x)$, $x \in M_c$, *on* M_c *and* $X_{\bar{f}}$ *on* $\bar{M}_c$ *are* p_{K_c}*-related.*

(b) *For each* $\bar{f} \in \mathcal{F}(\bar{M}_c)$ *there exists* $f \in R^{(c)}$ *such that* $f \mid M_c = p^*_{K_c}\bar{f}$ *and the vector fields* $X_f(x)$, $x \in M_c$, *on* M_c *and* $X_{\bar{f}}$ *on* $\bar{M}_c$ *are* p_{K_c}*-related.*

Proof. **(a)** Let $f \in R^{(c)}$, we only need to show that f is constant on the leaves of K_c, which is equivalent to showing that $df(x) \mid \ker \omega_c(x) = 0$, for all $x \in M_c$. For a given $x \in M_c$, let $v_x \in \ker \omega_c(x)$; then using lemma 3.10 **(b)** one sees that there is a function $g \in R^{(c)}$ such that $v_x = X_g(x)$. Then we have

$$0 = \omega_c(x)\left(X_f(x), X_g(x)\right) = \Omega(x)\left(X_f(x), X_g(x)\right) = df(x)X_g(x).$$

Now we shall prove that X_f and $X_{\bar{f}}$ are p_{K_c}-related. For each $x \in M_c$ and each $Y_x \in T_x M_c$, we have

$$\omega_c(X_f(x), Y_x) = \Omega(X_f(x), Y_x) = df(x)(Y_x) = d(p^*_{K_c}\bar{f})(x)(Y_x).$$

Using this we obtain

$$\omega_c(x)(X_f(x), Y_x) = \bar{\omega}_c(x)(p_{K_c}(x))\left(T_x p_{K_c} X_f(x), T_x p_{K_c} Y_x\right) = d\bar{f}(p_{K_c}(x))(T_x p_{K_c} Y_x),$$

which shows that $X_{\bar{f}}(p_{K_c}(x)) = T_x p_{K_c} X_f(x)$, because $\bar{\omega}_c$ is symplectic and $T_x p_{K_c} Y_x$ represents an arbitrary element of $T_{p_{K_c}(x)}\bar{M}$.

(b) To find f we choose a vector field X that is p_{K_c}-related to $X_{\bar{f}}$ according tolemma 3.8 and then use 3.10, **(b)**. □

Definition 3.12.

$$I^{(c)} = \{f \in R^{(c)} \mid f \mid M_c = 0\},$$
$$Z_{I^{(c)}} R^{(c)} = \{f \in R^{(c)} \mid \{f,h\} \in I^{(c)}, \text{for all } h \in R^{(c)}\}.$$

Elements of $I^{(c)}$ are called **first class constraints.**

Lemma 3.13. ***(a)*** *$I^{(c)}$ is a Poisson ideal of $R^{(c)}$, that is, it is an ideal of the ring $R^{(c)}$ such that if $f \in I^{(c)}$, then $\{f,h\} \in I^{(c)}$, for all $h \in R^{(c)}$.*

(b) *$Z_{I^{(c)}} R^{(c)}$ is a Poisson subalgebra of $R^{(c)}$.*

Proof. **(a)** Let $f,g \in I^{(c)}$ and $h \in R^{(c)}$. Then it is immediate that $f+g$ and hg belong to $I^{(c)}$. For any $h \in R^{(c)}$, we have $\{f,h\} \mid M_c = X_h(f) \mid M_c = 0$.

(b) Follows from **(a)**, using basic Poisson algebra arguments. □

Lemma 3.14. *The following conditions are equivalent for a function $f \in R^{(c)}$.*

(i) $f \in Z_{I^{(c)}} R^{(c)}$.
(ii) $f \mid M_c$ *is locally constant.*
(iii) $X_f(x) \in \ker \omega_c(x)$ *for* $x \in M_c$.

Proof. Assume **(i)**. Then $\{f,h\} \mid M_c = 0$ for all $h \in R^{(c)}$, that is, $df(x) X_h(x) \mid M_c = 0$. By lemma 3.10, **(b)**, we know that $X_h(x)$ represents any vector in $T_x M_c$. We can conclude that $f \mid M_c$ is locally constant, so **(ii)** holds. Now we will prove that **(ii)** implies **(iii)**. Let $f \mid M_c$ be locally constant. Then for all $g \in R^{(c)}$ and all $x \in M_c$,

$$0 = X_g(f)(x) = \Omega(x)(X_f, X_g)(x) = \omega_c(x)(X_f(x), X_g(x)).$$

Since, again by lemma 3.10, $X_g(x)$ represents any element of $T_x M_c$, we can conclude that $X_f(x) \in \ker \omega_c(x)$, so **(iii)** holds true. Now we will prove that **(iii)** implies **(i)**. Assume that $X_f(x) \in \ker \omega_c(x)$, $x \in M_c$. Then for all $g \in R^{(c)}$ and all $x \in M_c$,

$$\{g,f\}(x) = \Omega(x)(X_g, X_f)(x) = \omega_c(x)(X_g, X_f)(x) = 0,$$

that is, $\{g,f\} \in I^{(c)}$. Using this and the definitions, we see that $f \in Z_{I^{(c)}} R^{(c)}$.

Lemma 3.15. *The map* $(p_{K_c})_* : R^{(c)} \to \mathcal{F}(\bar{M}_c)$ *defined in lemma 3.11 is a surjective Poisson map and its kernel is* $I^{(c)}$*, therefore there is a natural isomorphism of Poisson algebras* $(p_{K_c})_{*I^{(c)}} : R^{(c)} / I^{(c)} \to \mathcal{F}(\bar{M}_c)$.

Proof. Surjectivity of $(p_{K_c})_*$ and the fact that its kernel is $I^{(c)}$ follows immediately from lemma 3.11 and the definitions. This implies that $(p_{K_c})_{*I^{(c)}}$ is an algebra isomorphism. Also, using the definitions, for $f, g \in R^{(c)}$ and any $x \in M_c$ we can prove easily that

$$\begin{aligned}
\{f,g\}(x) &= \Omega(x)(X_f(x), X_g(x)) \\
&= \omega_c(x)(X_f(x), X_g(x)) \\
&= \bar{\omega}_c\left(p_{K_c}(x)\right)\left(X_{\bar{f}}\left(p_{K_c}(x)\right), X_{\bar{g}}\left(p_{K_c}(x)\right)\right) \\
&= \{\bar{f}, \bar{g}\}(p_{K_c}(x)),
\end{aligned}$$

where $\bar{f} = (p_{K_c})_* f$, $\bar{g} = (p_{K_c})_* g$. Denote by $\pi_{I^{(c)}} : R^{(c)} \to R^{(c)} / I^{(c)}$ the natural homomorphism of Poisson algebras. Then from the previous equalities we obtain $(p_{K_c})_{*I^{(c)}} \left\{\pi_{I^{(c)}}(f), \pi_{I^{(c)}}(g)\right\} = \{\bar{f}, \bar{g}\}$, which shows that $(p_{K_c})_{*I^{(c)}}$ is a Poisson isomorphism. In other words, we have the commutative diagram

$$\begin{tikzcd}
R^{(c)} \arrow[r, "\pi_{I^{(c)}}"] \arrow[dr, "(p_{K_c})_*"'] & R^{(c)}/I^{(c)} \arrow[d, "(p_{K_c})_{*I^{(c)}}"] \\
& \mathcal{F}(\bar{M}_c)
\end{tikzcd}$$

All the arrows are defined in a natural way and they are surjective Poisson algebra homomorphisms.

Equations of Motion and Physical Variables. It is immediate to see from the definitions that

$$\ker \omega(x) \cap T_x M_c \subseteq \ker \omega_c(x), \qquad (15)$$

for all $x \in M_c$.

From now on we will assume the following.

Assumption K_2**. (a)** $\ker \omega(x)$ is a regular distribution, that is, it determines a regular foliation K and the natural projection $p_K : M \to \bar{M}$, where $\bar{M} = M / K$, is a submersion.

(b) The distribution $\ker \omega(x) \cap T_x M_c$ is a distribution of constant rank.

Theorem 3.16. *The distribution* $\ker \omega(x) \cap T_x M_c$ *is regular and has rank* $d^{(c)}(x)$. *Its integral manifolds are* $S \cap M_c$, *where* S *is an integral manifold of* $\ker \omega$. *Moreover, these integral manifolds give a foliation* $\tilde{K}_c$ *of* M_c *which is regular, that is, the natural map* $p_{\tilde{K}_c} : M_c \to \tilde{M}_c$, *where* $\tilde{M}_c = M_c / \tilde{K}_c$ *is a submersion. Besides, each leaf of the foliation* K_c *is foliated by leaves of* $\tilde{K}_c$, *which gives a naturally defined submersion* $p_{K_c \tilde{K}_c} : \tilde{M}_c \to \bar{M}_c$. *In other words, we obtain the commutative diagram*

$$\begin{tikzcd}
M_c \arrow[r, "p_{\tilde{K}_c}"] \arrow[dr, "p_{K_c}"'] & \tilde{M}_c \arrow[d, "p_{K_c\tilde{K}_c}"] \\
 & \bar{M}_c
\end{tikzcd}$$

where each arrow is a naturally defined submersion.

Proof. The first assertion, about the rank of the distribution $\ker \omega(x) \cap T_x M_c$, is easy to prove. Let $x_0 \in M_c$. Then there exists a uniquely determined integral manifold S of the distribution $\ker \omega$ such that $x_0 \in S$. Using that, by assumption, $\ker \omega(x) \cap T_x M_c$ is a distribution of constant dimension and that $\dim\left(\ker \omega(x) \cap T_x M_c\right) = \dim\left(T_x S \cap T_x M_c\right)$ we can conclude that the intersection $S \cap M_c$ coincides with the integral leaf of the integrable distribution of $\ker \omega \cap TM_c$ containing x_0. So we obtain the foliation $\tilde{K}_c$ of M_c. Using (15) we can deduce that each leaf of the foliation K_c is foliated by leaves of $\tilde{K}_c$. The rest of the proof follows by standard arguments.

Lemma 3.17. ***(a)*** *The following diagram is commutative*

$$\begin{tikzcd}
M \arrow[r, "p_K"] & \bar{M} & T\bar{M} \arrow[l, "\tau_{\bar{M}}"'] & \\
M_c \arrow[u, "f_c"] \arrow[r, "p_{\tilde{K}_c}"] \arrow[dr, "p_{K_c}"'] & \tilde{M}_c \arrow[u, "\tilde{f}_c"'] \arrow[d, "p_{K_c\tilde{K}_c}"] & T\bar{M}|\tilde{M}_c \arrow[l, "\tau_{\bar{M}}"] \arrow[u, "\tilde{F}_c \circ \epsilon_c"] \arrow[r, "\epsilon_c"] & \tilde{f}_c^* T\bar{M} \arrow[ul, "\tilde{F}_c"'] \\
 & \bar{M}_c & &
\end{tikzcd}$$

where the arrows are defined as follows. The maps p_K, p_{K_c}, $p_{\tilde{K}_c}$ *and* $p_{K_c \tilde{K}_c}$ *are defined in Assumption* K_1, *Assumption* K_2 *and theorem 3.16. By definition, the map* f_c *is the inclusion. The map* $\tilde{f}_c$ *is an embedding defined by* $\tilde{f}_c(S \cap M_c) = S$, *where* S *is a leaf of the foliation* K. *We will think of* $\tilde{f}_c$ *as being an inclusion. The vector bundle* $T\bar{M} \mid \tilde{M}_c$ *is the tangent bundle* $T\bar{M}$

restricted to $\tilde{M}_c$. *In other words, since* $\tilde{f}_c$ *is an inclusion,* $T\bar{M} \,|\, \tilde{M}_c$ *is identified via some isomorphism, called* ε_c, *with the pullback of* $T\bar{M}$ *by* $\tilde{f}_c$. *We call* $\tilde{F}_c$ *the natural map associated to the pullback.*

(b) *The presymplectic form* ω *on* M *passes to the quotient via* p_K *giving a uniquely defined symplectic form* $\bar{\omega}$ *on* $\bar{M}$, *satisfying* $p_K^* \bar{\omega} = \omega$. *The presymplectic form* ω_c, *which, by definition is* $f_c^* \omega$, *defines uniquely a presymplectic form* $\tilde{\omega}_c$ *on* $\tilde{M}_c$ *via* $p_{\tilde{K}_c}$ *satisfying* $p_{\tilde{K}_c}^* \tilde{\omega}_c = \omega_c$, $\tilde{\omega}_c = \tilde{f}_c^* \bar{\omega}$. *The energy function* E *on* M *satisfies* $d\mathrm{E}(x) \,|\, \ker \omega(x) = 0$, *for all* $x \in M_c$, *therefore it defines uniquely a 1-form on* $T\bar{M} \,|\, \tilde{M}_c$, *called* $(\tilde{F}_c \circ \varepsilon)^* d\mathrm{E} \in \Gamma\left((T\bar{M} \,|\, \tilde{M}_c)^*\right)$. *Since* E *is constant on each leaf of* $\tilde{K}_c$, *it also defines a function* $\tilde{\mathrm{E}}_c$ *on* $\tilde{M}_c$. *Since* $T\tilde{M}_c \subseteq T\bar{M}_c$ *via the inclusion* $T\tilde{f}_c$ *we have* $(\tilde{F}_c \circ \varepsilon)^* d\mathrm{E} \,|\, T_{\tilde{x}} \tilde{M}_c = d\tilde{\mathrm{E}}_c(\tilde{x})$, *for all* $\tilde{x} \in \tilde{M}_c$.

(c) *Equation of motion (14) on* M_c *passes to the quotient* $\tilde{M}_c$ *as*

$$\bar{\omega}(\tilde{x})(\tilde{X}(\tilde{x}),) = (\tilde{F}_c \circ \varepsilon)^* d\mathrm{E}(\tilde{x}), \qquad (16)$$

where $\tilde{X}(\tilde{x}) \in T_{\tilde{x}} \tilde{M}_c$. *This means that if* $X(x) \in T_x M_c$ *is a solution of (14) then* $\tilde{X}(\tilde{x}) := T_x p_{\tilde{K}_c} X(x)$, *where* $\tilde{x} = p_{\tilde{K}_c}(x)$, *is a solution of (16). Therefore, a solution curve* $x(t)$ *of (14) projects to a solution curve* $\tilde{x}(t) = p_{\tilde{K}_c}\left(x(t)\right)$ *of (16) on* $\tilde{M}_c$. *Equation (16) has unique solution* $\tilde{X}(\tilde{x})$ *for each* $\tilde{x} \in \tilde{M}_c$. *This solution also satisfies the equation*

$$\tilde{\omega}_c(\tilde{x})(\tilde{X}(\tilde{x}),) = d\tilde{\mathrm{E}}(\tilde{x}). \qquad (17)$$

However solutions to equation (17) are not necessarily unique, since $\ker \tilde{\omega}_c(\tilde{x})$ *is not necessarily* 0.

(d) *The restriction of the energy function* $\mathrm{E} \,|\, M_c$ *satisfies*

$$d(\mathrm{E} \,|\, M_c)(x) \,|\, \ker \omega_c(x) = 0,$$

for all $x \in M_c$, *therefore there is a uniquely defined function* $\bar{\mathrm{E}}_c$ *on* $\bar{M}_c$ *such that* $p_{K_c}^* \bar{\mathrm{E}}_c = \mathrm{E} \,|\, M_c$. *The equation*

$$\bar{\omega}_c(\bar{x})(\bar{X}(\bar{x}),) = d\bar{\mathrm{E}}_c(\bar{x}) \qquad (18)$$

has unique solution $\bar{X}(\bar{x})$ *for* $\bar{x} \in \bar{M}_c$. *If* $\tilde{X}(\tilde{x})$ *is a solution of (16) then* $\bar{X}(\bar{x}) = T_{\tilde{x}} p_{K_c \tilde{K}_c} \tilde{X}(\tilde{x})$ *is a solution of (18). Therefore, a solution curve* $\tilde{x}(t)$ *of (16) projects to a solution curve* $\bar{x}(t) = p_{K_c \tilde{K}_c}(\tilde{x}(t))$ *of (18) on* $\bar{M}_c$.

Proof. **(a)** The equality $p_{K_c} = p_{K_c \tilde{K}_c} \circ p_{\tilde{K}_c}$ was proven in theorem 3.16. The equality $\tau_{\bar{M}} \circ \tilde{F}_c \circ \varepsilon_c = \tilde{f}_c \circ \tau_{\bar{M}}$ results immediately from the definitions. The equality $p_K \circ f_c = \tilde{f}_c \circ p_{\tilde{K}_c}$ results by applying the definitions and showing that, for given $x \in M_c$, $p_K \circ f_c(x) = S_x = \tilde{f}_c \circ p_{\tilde{K}_c}(x)$, where S_x is the only leaf of K containing x.

(b) Existence and uniqueness of $\bar{\omega}$ and $\tilde{\omega}$ is a direct consequence of the definitions and standard arguments on passing to quotients. For any $x \in M_c$ we know that there exists a solution X of equation (14), from which it follows immediately that $d\mathrm{E}(x) \,|\, \ker \omega(x) = 0$. The rest of the proof of this item consists of standard arguments on passing to quotients.

(c) We shall omit the proof of this item which is a direct consequence of the definitions and standard arguments on passing to quotients.

(d) If $X(x)$ is a solution of $\omega(x)(X(x),) = d\mathrm{E}(x)$ then it is clear that it also satisfies $\omega_c(x)(X(x),) = d(\mathrm{E} \,|\, M_c)(x)$. It follows that

$$d(\mathrm{E} \,|\, M_c)(x) \,|\, \ker \omega_c(x) = 0,$$

for all $x \in M_c$. The rest of the proof is a consequence of standard arguments on passing to quotients.

Remark. Recall that the locally constant function $d^{(c)}(x)$ on M_c is the dimension of the distribution $\ker \omega(x) \cap T_x M_c$ on M_c and also the dimension of the fiber of the bundle $S^{(c)}(x)$. If $d^{(c)}(x)$ is nonzero then there is no uniqueness of solution to equation of motion (14), since solution curves to that equation satisfy, by definition,

$$\omega(x)(\dot{x},) = d\mathrm{E}(x) \,|\, T_x M,$$

where $(x, \dot{x}) \in T_x M_c$. Passing to the quotient manifold $\tilde{M}_c$ eliminates this indeterminacy and uniqueness of solution is recovered. This is related to the notion of *physical variables* mentioned by Dirac.

3.3 First Class and Second Class Constraints and Constraint Submanifolds. The Tangent Bundle V^{Ω} to a Second Class Submanifold.

As we have said before, an important topic in the theory of constraints as developed by Dirac is the distinction between *first class* and *second class constraints*. His treatment is intended to solve systems with constraints coming from degeneracies in the Lagrangian from both the classical and the quantum mechanics point of view. The Poisson algebra structure of functions on a symplectic manifold is the context in which this theory is developed and it is not very geometric and almost no attention is paid to the *constraint submanifolds* defined by the several equations involved. Among several interesting references we cite Sniatycki [39] which has several points of contact with our work.

In this paragraph we shall give a geometric framework and describe its close relationship to the Poisson algebra point of view to deal with the notions of first class and second class constraints and functions and also first class and second class submanifolds. These notions only depend on the final constraint submanifold M_c and the ambient symplectic manifold N and do not depend on the primary constraint $M_0 = M$ or the Hamiltonian $H : N \to \mathbb{R}$. Accordingly, in this paragraph we will adopt an abstract setting, where only an ambient symplectic manifold and a submanifold are given. This kind of abstract setting was studied in Sniatycki [39], in particular the notion of second class constraint submanifold and its connection with the Dirac bracket.

Then we will go back to equations of motion in the next subsection, where the role of both the final and the primary constraint is essential.
The definitions given at the beginning of this section inspire the following one.

Definition 3.18. Let (P,Ω) be a symplectic manifold and $S \subseteq P$ a given submanifold. Then, by definition,

$$R^{(S,P)} := \{f \in \mathcal{F}(P) \mid X_f(x) \in T_xS, \text{for all } x \in S\}$$
$$I^{(S,P)} := \{f \in R^{(S,P)} \mid f \mid S = 0\}$$

Elements of $R^{(S,P)}$ are called ***first class functions.*** Elements of $I^{(S,P)}$ are called ***first class constraints.*** The submanifold S is called a ***first class constraint submanifold*** if for all $f \in \mathcal{F}(P)$ the condition $f \mid S = 0$ implies $f \in I^{(S,P)}$, that is, $I_{(S,P)} \subseteq I^{(S,P)}$, where $I_{(S,P)}$ is the ideal of the ring $\mathcal{F}(P)$ of all functions vanishing on S.

Obviously, using the notation introduced before in the present section, $R^{(M_c,N)} = R^{(c)}$ and $I^{(M_c,N)} = I^{(c)}$. All the properties proven for $R^{(c)}$ and $I^{(c)}$ hold in general for $R^{(S,P)}$ and $I^{(S,P)}$. For instance, T_xS is the set of all $X_f(x), f \in R^{(S,P)}$. Every function $f \in R^{(S,P)}$ satisfies $df(x)(X_x) = 0$, for all $X_x \in \ker(\Omega(x) \mid T_xS)$ and $\ker(\Omega(x) \mid T_xS)$ is the set of all $X_f(x),\ f \in I^{(S,P)}$.

The following lemma 3.19 is one of our main results. It studies the vector subbundles $V \subseteq TP \mid S$ which classify all second class submanifolds S^V containing S at a linear level, that is, V^Ω is tangent to the second class submanifold. For such a second class submanifold, which is a symplectic submanifold, the Dirac bracket of two functions F and G at points $x \in S$ can be calculated, by definition, as the canonical bracket of the restrictions of F and G. This has a global character. A careful study of the *global* existence of a bracket defined on sufficiently small open sets $U \subseteq P$ containing S which coincides with the previous one on the second class submanifold will not be considered in this paper. However, to write global equations of motion on the final constraint submanifold one only needs to know the vector bundle V^{Ω}, which carries a natural fiberwise symplectic form. We will also describe the Dirac brackets locally, on an open neighborhood of any point of S, essentially as Dirac does, but in a more geometric way.

All these are fundamental properties of second class constraints and constraint submanifolds, and theorem 3.20 collects some essential parts of them; we suggest to take a look at it before reading lemma 3.19.

Lemma 3.19. *Let* (P,Ω) *be a symplectic manifold of dimension* $2n$ *and* $S \subseteq P$ *a given submanifold of codimension* r. *Let* ω *be the pullback of* Ω *to* S *and assume that* $\ker \omega(x)$ *has constant dimension. Assume that* S *is defined regularly by equations* $\phi_1 = 0, \ldots, \phi_r = 0$ *on a neighborhood* $U \supseteq S$ *and assume that we can choose a subset* $\{\chi_1, \ldots, \chi_{2s}\} \subseteq \{\phi_1, \ldots, \phi_r\}$ *such that* $\det(\{\chi_i, \chi_j\}(x)) \neq 0$ *for all* $x \in S$, *where we assume that* $2s = \operatorname{rank}(\{\phi_i, \phi_j\}(x))$, *for all* $x \in S$. *We shall often denote* $c_{ij}^{\chi}(x) = \{\chi_i, \chi_j\}(x)$ *and* $c_{\chi}^{ij}(x)$ *the inverse of* $c_{ij}^{\chi}(x)$. *Moreover, we will assume that the following stronger condition holds, for simplicity. Equations* $\phi_1 = B_1$, ..., $\phi_r = B_r$ *and* $\chi_1 = C_1$, ..., $\chi_{2s} = C_{2s}$ *define submanifolds of* U *regularly, for small enough* $B_1, \ldots, B_r$ *and* $C_1, \ldots, C_{2s}$.
Then

(a) $2s = r - \dim \ker \omega = 2n - \dim S - \dim \ker \omega$. *There are* $\psi_k \in I^{(S,P)}$, $k = 1, \ldots, r - 2s$, *which in particular implies* $\{\psi_k, \psi_l\}(x) = 0$, $\{\psi_k, \chi_i\}(x) = 0$, *for* $k, l = 1, \ldots, r - 2s$, $i = 1, \ldots, 2s$, *such that* $d\psi_1(x), \ldots, d\psi_{r-2s}(x), d\chi_1(x), \ldots, d\chi_{2s}(x)$, *are linearly independent for all* $x \in S$. *Moreover,* $X_{\psi_1}(x), \ldots, X_{\psi_{r-2s}}(x)$ *form a basis of* $\ker \omega(x)$, *for all* $x \in S$ *and* $d\psi_1(x), \ldots, d\psi_{r-2s}(x), d\chi_1(x), \ldots, d\chi_{2s}(x)$ *form a basis of* $(T_x S)^{\circ}$.

(b) *The vector subbundle* $V^{\chi} \subseteq TP \mid S$ *with base* S *and fiber*

$$V_x^{\chi} = \operatorname{span}\left(X_{\chi_1}(x), \ldots, X_{\chi_{2s}}(x)\right) \subseteq T_x P,$$

satisfies

$$V_x^{\chi} \cap T_x S = \{0\} \qquad (19)$$

$$V_x^{\chi} \oplus \ker \omega(x) = (T_x S)^{\Omega} \qquad (20)$$

$$(V_x^{\chi})^{\Omega} \cap (\ker \omega(x))^{\Omega} = T_x S, \qquad (21)$$

$x \in S$.

(c) *There is a neighborhood* U *of* S *such that the equations* $\chi_1 = 0, \ldots, \chi_{2s} = 0$ *on* U *define a symplectic submanifold* S^{χ} *such that* $S \subseteq S^{\chi}$ *and*

$$T_x S^\chi = \left(V_x^\chi\right)^\Omega$$
$$T_x S^\chi \oplus V_x^\chi = T_x P,$$

for $x \in S^\chi$*, where we have extended the definition of* V_x^χ *for* $x \in S^\chi$ *using the expression*

$$V_x^\chi = \operatorname{span}\left(X_{\chi_1}(x),\ldots,X_{\chi_{2s}}(x)\right) \subseteq T_x P,$$

for $x \in S^\chi$*. The submanifold* S^χ *has the property* $I_{(S,S^\chi)} \subseteq I^{(S,S^\chi)}$*, that is,* S *is a first class constraint submanifold of* S^χ*, defined regularly by* $\psi_i \,|\, S^\chi = 0$*,* $i = 1,\ldots,r-2s$*, and* $\psi_i \,|\, S^\chi \in I^{(S,S^\chi)}$*,* $i = 1,\ldots,r-2s$*. Moreover, it has the only possible dimension, which is* $\dim S^\chi = \dim S + \dim\ker\omega = 2n-2s$*, for symplectic submanifolds having that property. It is also a minimal object in the set of all symplectic submanifolds* $P_1 \subseteq P$*, ordered by inclusion, satisfying* $S \subseteq P_1$.

(d) *Let* V *be any vector subbundle of* $TP\,|\,S$ *such that*

$$V \oplus \ker\omega = (TS)^\Omega, \qquad (22)$$

or equivalently,

$$V^\circ \oplus (\ker\omega)^\circ = (TS)^\circ. \qquad (23)$$

Then $\dim V_x = 2s$*, for* $x \in S$*. Let* S^V *be a submanifold of* P *such that* $T_x S^V = V_x^\Omega$*, for each* $x \in S$*. Then* S *is a submanifold of* S^V*. Such a submanifold* S^V *always exists. Moreover, for such a submanifold there is an open set* $U \subseteq P$ *containing* S *such that* $S^V \cap U$ *is a symplectic submanifold of* P.

Let $\bar{x} \in S$ *and let* $\chi_1' = 0,\ldots,\chi_{2s}' = 0$ *be equations defining* $S^V \cap U'$ *for some open neighborhood* $U' \subseteq P$ *and satisfying that* $d\chi_1'(x),\ldots,d\chi_{2s}'(x)$ *are linearly independent for* $x \in S^V \cap U'$*. Then,* $d\chi_1'(x),\ldots,d\chi_{2s}'(x), d\psi_1(x),\ldots,d\psi_{r-2s}(x)$ *are linearly independent and* $\det(\{\chi_i',\chi_j'\}(x)) \neq 0$*, for* $x \in S^V \cap U'$*. All the properties established in* ***(a)****,* ***(b)****,* ***(c)*** *for* $\chi_1,\ldots,\chi_{2s}$ *on* S *hold in an entirely similar way for* $\chi_1',\ldots,\chi_{2s}'$*, on* $S \cap U'$*. In particular,* $S^V \cap U' = S^{\chi'}$.

(e) *Let* ω^χ *be the pullback of* Ω *to* S^χ *and* $\{,\}_\chi$ *the corresponding bracket. For given* $F, G \in \mathcal{F}(P)$ *define* $F_\chi := F - \chi_i c_\chi^{ij}\{\chi_j, F\}$ *and also*

$$\{F,G\}_{(\chi)} := \{F,G\} - \{F,\chi_i\} c_\chi^{ij} \{\chi_j, G\},$$

which is the famous bracket introduced by Dirac, called ***Dirac bracket****, and it is defined for* x *in the neighborhood* U *where* $c_{ij}^\chi(x)$ *has an inverse* $c_\chi^{ij}(x)$*. Then, for any* $x \in S^\chi$*,*

$$\{F_\chi, \chi_i\}(x) = 0$$

for $i = 1, \ldots, 2s$ *, and also*

$$\{F_\chi, G_\chi\}(x) = \{F, G\}_{(\chi)}(x) = \omega^\chi(x)\left(X_{F|S^\chi}(x), X_{G|S^\chi}(x)\right) = \{F \mid S^\chi, G \mid S^\chi\}_\chi(x).$$

If we denote $X_{(\chi),F}$ *the Hamiltonian vector field associated to the function* $F \in \mathcal{F}(P)$*, with respect to the Dirac bracket* $\{,\}_{(\chi)}$ *then the previous equalities are equivalent to*

$$X_{(\chi),F}(x) = X_{F_\chi}(x) = X_{F|S^\chi}(x).$$

The Jacobi identity is satisfied for the Dirac bracket $\{F, G\}_{(\chi)}$ *on* S^χ*, that is,*

$$\{\{F, G\}_{(\chi)}, H\}_{(\chi)}(x) + \{\{H, F\}_{(\chi)}, G\}_{(\chi)}(x) + \{\{G, H\}_{(\chi)}, F\}_{(\chi)}(x) = 0,$$

for $x \in S^\chi$.

(f) *Let* U *be an open neighborhood of* S *such that* $c_{ij}^\chi(x)$ *is invertible for* $x \in U$*. For each* $C = (C_1, \ldots, C_{2s}) \in \mathbb{R}^{2s}$ *let* $\chi_i^C = \chi_i - C_i$ *and define* S^{χ^C} *by the equations* $\chi_i^C(x) = 0$*,* $i = 1, \ldots, 2s$*,* $x \in U$*. For any* C *in a sufficiently small neighborhood of* 0*,* S^{χ^C} *is a nonempty symplectic submanifold of* P*. Define the matrix* $c_{ij}^{\chi^C}(x) = \{\chi_i^C, \chi_j^C\}(x)$*, and also* $c_{\chi^C}^{ij}(x)$ *as being its inverse,* $x \in U$*. Then, the equalities*

$$c_{ij}^\chi(x) = \{\chi_i, \chi_j\}(x) = \{\chi_i^C, \chi_j^C\}(x) = c_{ij}^{\chi^C}(x), \qquad (24)$$

and also,

$$\{F, G\}_{(\chi)}(x) = \{F, G\}_{(\chi^C)}(x) \qquad (25)$$

are satisfied for all $x \in U$*. All the definitions and properties proved in **(e)** for the case* $C = 0$ *hold in general for any* C *in a neighborhood of* 0 *small enough to ensure that* S^{χ^C} *is nonempty. In particular, the equalities*

$$\begin{aligned}\{F_{\chi^C}, G_{\chi^C}\}(x) &= \{F, G\}_{(\chi^C)}(x) = \omega^{\chi^C}(x)\left(X_{F|S^{\chi^C}}(x), X_{G|S^{\chi^C}}(x)\right) \\ &= \{F \mid S^{\chi^C}, G \mid S^{\chi^C}\}_{\chi^C}(x)\end{aligned} \qquad (26)$$

and

$$X_{(\chi^C),F}(x) = X_{F_{\chi^C}}(x) = X_{F|S^{\chi^C}}(x) \qquad (27)$$

hold for $x \in S^{\chi^C}$*, and any* C *in such a neighborhood. The Dirac bracket* $\{F, G\}_{(\chi)}$ *satisfies the Jacobi identity for* $F, G \in \mathcal{F}(U)$ *and the symplectic submanifolds* S^{χ^C} *are the symplectic leaves of the Poisson manifold* $(\mathcal{F}(U), \{,\}_{(\chi)})$*. By shrinking, if necessary, the open set* U *and for* C *in a* *sufficiently small neighborhood of* $0 \in \mathrm{R}^{2s}$*, the equations* $\psi_k \,|\, S^{\chi^C} = 0$, $k = 1, \ldots, r-2s$*, define regularly a first class constraint submanifold* $S^C \subseteq S^{\chi^C} \subseteq U$*, and the functions* $\psi_k \,|\, S^{\chi^C} \in R^{(S^C, S^{\chi^C})} \subseteq \mathrm{F}\,(S^{\chi^C})$ *are first class constraints, that is,* $\psi_k \,|\, S^{\chi^C} \in I^{(S^C, S^{\chi^C})}$ $k = 1, \ldots, r-2s$*. We have that* $\dim S^{\chi^C} = \dim S^C + \dim \ker \omega^C$*, where* ω^C *is the pullback of* Ω *to* S^C*. One has* $\dim S^C = \dim S$ *and* $\dim S^{\chi^C} = \dim S^{\chi}$*, therefore* $\dim \ker \omega = \dim \ker \omega^C$*.*

Proof. **(a)** Let $x \in S$. We are going to use lemmas and corollaries 3.1–3.7 with $E := T_x P;\ V := T_x S;\ \gamma_i := d\phi_i(x)$, $i = 1, \ldots, r;\ \Omega := \Omega(x);\ \omega := \omega(x);\ \beta = 0$.

Elements $\psi = \lambda^i \phi_i$, $\lambda^i \in \mathrm{F}\,(P)$ such that $\psi \in I^{(S,P)}$, which implies $X_\psi(x) \in T_x S$ for $x \in S$, must satisfy $\{\psi, \phi_j\}(x) = 0$, or, equivalently, $\lambda^i(x) d\phi_i(x)\left(X_{\phi_j}(x)\right) = 0$, for $j = 1, \ldots, r$, $x \in S$. Using lemma 3.6 we see that $X_\psi(x) \in \ker \omega(x)$. Since one can choose $r-2s$ linearly independent solutions, say $\lambda_i = (\lambda_i^1, \ldots, \lambda_i^r)$, $i = 1, \ldots, r-2s$, we obtain elements $\psi_i \in I^{(S,P)}$, namely, $\psi_i = \lambda_i^j \phi_j$, such that $(d\psi_1(x), \ldots, d\psi_{r-2s}(x))$ are linearly independent, or, equivalently, taking into account lemma 3.7, that $(X_{\psi_1}(x), \ldots, X_{\psi_{r-2s}}(x))$ is a basis of $\ker \omega(x)$ for $x \in S$. If $d\psi_1(x)$, ..., $d\psi_{r-2s}(x)$, $d\chi_1(x)$, ..., $d\chi_{2s}(x)$, were not linearly independent, then there would be a linear combination, say $\bar{\chi} = a^i \chi_i$, with at least one nonzero coefficient, and some $x \in S$, such that $d\bar{\chi}(x) = \mu^k d\psi_k(x)$ for some μ^k, $k = 1, \ldots, r-2s$. But then, for any $j = 1, \ldots, 2s$, $\{\bar{\chi}, \chi_j\}(x) = d\bar{\chi}(x) X_{\chi_j}(x) = \mu^k d\psi_k(x) X_{\chi_j}(x) = -\mu^k d\chi_j(x) X_{\psi_k}(x) = 0$, which contradicts the fact that $\det\left(\{\chi_i, \chi_j\}(x)\right) \neq 0$. Using this and the fact that $\psi_i = 0$, $\chi_j = 0$, $i = r-2s$, $j = 1, \ldots, 2s$ define S regularly, we can conclude that $d\psi_1(x)$, ..., $d\psi_{r-2s}(x)$, $d\chi_1(x)$, ..., $d\chi_{2s}(x)$ is a basis of $(T_x S)^\circ$.

(b) If $\lambda^i X_{\chi_i}(x) \in T_x S$ then $d\chi_j(x) \lambda^i X_{\chi_i}(x) = 0$, $j = 1, \ldots, 2s$, which implies $\lambda^i \{\chi_j, \chi_i\} = 0$, $j = 1, \ldots, 2s$, then $\lambda^i = 0$, $i = 1, \ldots, 2s$, which proves (19). To prove (20) we apply the operator $^{\circ}$ to both sides and obtain the equivalent equality $\mathrm{span}(d\chi_1(x), \ldots, d\chi_{2s}(x)) \oplus \mathrm{span}(d\psi_1(x), \ldots, d\psi_{r-2s}(x)) = (T_x S)^\circ$, which we know is true, as proven in **(a)**. To prove (21) we apply the orthogonal operator $^{\Omega}$ to both sides of (20).

(c) Since $d\chi_1(x), \ldots, d\chi_{2s}(x)$ are linearly independent for $x \in S$ they are also linearly independent for x in a certain neighborhood U of S. Then $\chi_1(x) = 0, \ldots, \chi_{2s}(x) = 0$ define a submanifold S^χ of U containing S. To see that it is a symplectic submanifold choose $x \in S^\chi$ and apply corollary 3.7 with $E := T_x P;\ V := T_x S^\chi;\ \beta := 0;\ \gamma_i := d\chi_i$, $i = 1, \ldots, 2s$; $\omega := \Omega(x) \,|\, T_x S^\chi$. We can conclude that $\dim(\ker \Omega(x) \,|\, T_x S^\chi) = 0$. Now let us prove that

$T_x S^\chi = \left(V_x^\chi\right)^\Omega$, namely, $T_x S^\chi = \mathrm{span}\left(d\chi_1(x),\ldots,d\chi_{2s}(x)\right)^\circ = \left(\left(V_x^\chi\right)^\circ\right)^\circ = \left(V_x^\chi\right)^\Omega$. From this, using that S^χ is symplectic one obtains $T_x S^\chi \oplus V_x^\chi = T_x P$. To prove that $S \subseteq S^\chi$ is a first class constraint submanifold defined by first class constraints $\psi_i \mid S^\chi$, $i=1,\ldots,r-2s$, on S^χ, we observe first that it is immediate that $\psi_i \mid S^\chi = 0$, $i=1,\ldots,r-2s$, define S regularly. It remains to show that $X_{\psi_i|S^\chi}(x) \in T_x S$, $i=1,\ldots,r-2s$, $x \in S$, where $X_{\psi_i|S^\chi}$ is the Hamiltonian vector field associated to the function $\psi_i \mid S^\chi$ with respect to the symplectic form ω^χ. This is equivalent to showing that

$$X_{\psi_i|S^\chi}(x)\left(\psi_j \mid S^\chi\right) = 0,$$

for $x \in S$ or, equivalently,

$$\omega^\chi(x)\left(X_{\psi_i|S^\chi}(x), X_{\psi_j|S^\chi}(x)\right) = 0,$$

for $x \in S$. We know that ω^χ is the pullback of Ω to S^χ and $\psi_i \mid S^\chi$ is the pullback of ψ_i, via the inclusion $S^\chi \subseteq U$, then we have

$$\omega^\chi(x)\left(X_{\psi_i|S^\chi}(x), X_{\psi_j|S^\chi}(x)\right) = \Omega(x)\left(X_{\psi_i}(x), X_{\psi_j}(x)\right) = 0,$$

for $x \in S$, since ψ_i are first class constraints, $i=1,\ldots,r-2s$. Finally, using the definitions we can easily see that $\dim S^\chi = 2n-2s$ and that $\dim S = 2n-r$ and from **(a)** we know that $\dim\ker\omega = r-2s$. We can conclude that $2n-2s = \dim S + \dim\ker\omega$.

(d) We know that, for $x \in S$, $\dim T_x S = 2n-r$, and $\dim\ker\omega(x) = r-2s$; then using (22) we obtain $\dim V_x = 2s$. Also from (22) we immediately deduce applying $^\Omega$ to both sides,

$$V^\Omega \cap (\ker\omega)^\Omega = TS,$$

in particular $TS \subseteq V^\Omega$. Let g be a given Riemannian metric on P and let W_x be the g-orthogonal complement of $T_x S$ in V_x^Ω, in particular, $W_x \oplus T_x S = V_x^\Omega$, for each $x \in S$. Define

$$S^V = \{\exp(tw_x) \mid w_x \in W_x, g(x)(w_x,w_x) = 1, |t| < \tau(x), x \in S\}.$$

By choosing $\tau(x)$ appropriately one can ensure that S^V is a submanifold and, moreover, it is easy to see from the definition of S^V that $T_x S^V = W_x \oplus T_x S = V_x^\Omega$, for each $x \in S$. We leave for later the proof that $S^V \cap U$ is a symplectic submanifold of P, for an appropriate choice of the open set U, which amounts to choosing $\tau(x)$ appropriately.

Assume that $d\chi_1'(x),\ldots,d\chi_{2s}'(x)$ are linearly independent for $x \in S^V \cap U'$. Since $\langle d\chi_i'(x), V_x^\Omega\rangle = \langle d\chi_i'(x), T_x S^V\rangle = 0$ for $x \in S$ and $i = 1,\ldots,2s$, we can deduce that $d\chi_i'(x) \in \left(V_x^\Omega\right)^\circ$, that is, $d\chi_i'(x) \in V_x^\varsigma$. Then using (23), we see that $d\chi_1'(x),\ldots$, $d\chi_{2s}'(x)$, $d\psi_1(x),\ldots,d\psi_{r-2s}(x)$ are linearly independent and span $V^\varsigma \oplus (\ker\omega)^\varsigma = (TS)^\circ$. If $\det(\{\chi_i', \chi_j'\}(x)) = 0$ for some $x \in S$ then $\lambda^i\{\chi_i', \chi_j'\}(x) = 0$, where at least some $\lambda^i \neq 0$, $i = 1,\ldots,2s$. Let $\lambda^i\chi_i' = \varphi$, then $\{\varphi, \chi_j'\}(x) = 0$, $j = 1,\ldots,2s$. On the other hand, since $\varphi \mid S = 0$, then $\{\varphi, \psi_j\}(x) = 0$, $j = 1,\ldots,r-2s$. We can conclude that $\varphi \in I^{(S,U')}$ and then $X_\varphi(x) \in \ker\omega(x)$, in particular, $X_\varphi(x) = \mu^j X_{\psi_j}(x)$, which implies $\lambda^i d\chi_i'(x) = d\varphi(x) = \mu^j d\psi_j(x)$, contradicting the linear independence of $d\chi_1'(x),\ldots,d\chi_{2s}'(x), d\psi_1(x),\ldots,d\psi_{r-2s}(x)$.

It follows from which precedes that by replacing χ_i by χ_i', $i = 1,\ldots,2s$ and S by $S \cap U'$ all the properties stated in **(a)**, **(b)** and **(c)** are satisfied. In particular, $S \cap U = S^{\chi'}$ and $S \cap U'$ is symplectic. It is now clear that by covering S with open subsets like the U' we can define U as being the union of all such open subsets and one obtains that $S \cap U$ is a symplectic submanifold.

(e) Let $x \in S^\chi$. Then, since $F_\chi = F - \chi_i c_\chi^{ij}\{\chi_j, F\}$, we obtain

$$\{F_\chi, \chi_k\}(x) = \{F, \chi_k\}(x) - \{\chi_i, \chi_k\}(x) c_\chi^{ij}(x)\{\chi_j, F\}(x) = \{F, \chi_k\}(x) + \{\chi_k, F\}(x) = 0.$$

Using this we obtain

$$\begin{aligned}\{F_\chi, G_\chi\}(x) &= \{F_\chi, G - \chi_k c_\chi^{kl}\{\chi_l, G\}\}(x) = \{F_\chi, G\}(x)\\ &= \{F, G\}(x) - \{\chi_i, G\} c_\chi^{ij}\{\chi_j, F\} = \{F, G\}_{(\chi)}(x).\end{aligned}$$

For any $F \in \mathcal{F}(P)$, $x \in S^\chi$ and $k = 1,\ldots,2s$, we have $X_{F_\chi}(x)\chi_k = \{\chi_k, F_\chi\}(x) = 0$, so $X_{F_\chi}(x) \in T_x S^\chi$. Therefore, for any $Y_x \in T_x S^\chi$,

$$\begin{aligned}\omega^\chi(x)\left(X_{F_\chi}(x), Y_x\right) &= \Omega(x)\left(X_{F_\chi}(x), Y_x\right) = dF_\chi(x) Y_x = d\left(F_\chi \mid S^\chi\right)(x) Y_x\\ &= d\left(F \mid S^\chi\right)(x) Y_x = \omega^\chi(x)\left(X_{F|S^\chi}(x), Y_x\right),\end{aligned}$$

which shows that $X_{F_\chi}(x) = X_{F|S^\chi}(x)$, where both Hamiltonian vector fields are calculated with the symplectic form ω^χ. Using this, for any $G \in \mathcal{F}(P)$ and any $x \in S^\chi$, one obtains

$$\{G_\chi, F_\chi\}(x) = X_{F_\chi}(x) G_\chi = X_{F_\chi}(x) G = X_{F|S^\chi}(x) G \mid S^\chi = \{G \mid S^\chi, F \mid S^\chi\}_\chi(x).$$

The equality $X_{(\chi),F}(x) = X_{F_\chi}(x) = X_{F|S^\chi}(x)$ is an immediate consequence of the previous ones. The Jacobi identity for the bracket $\{,\}_{(\chi)}$ follows using the previous formulas, namely, for $x \in S^\chi$, one obtains

$$\{\{F,G\}_{(\chi)},H\}_{(\chi)}(x)=\{\{F,G\}_{(\chi)}\,|\,S^{\chi},H\,|\,S^{\chi}\}_{\chi}(x)=\{\{F\,|\,S^{\chi},G\,|\,S^{\chi}\}_{\chi},H\,|\,S^{\chi}\}_{\chi}(x),$$

where the bracket in the last term is the canonical bracket on the symplectic manifold S^{χ}, for which the Jacobi identity is well known to be satisfied.

(f) The equalities (24) and (25) are proven in a straightforward way. The equations (26) and (27) follow easily using a technique similar to the one used in **(e)**. Using all this, the proof of the Jacobi identity for the bracket $\{,\}_{(\chi)}$ on U goes as follows. Let $x\in U$ and let C be such that $x\in S^{\chi^C}$. For $F,G,H\in \mathcal{F}(U)$ using **(e)** we know that the Jacobi identity holds for $\{,\}_{(\chi^C)}$ on S^{χ^C}. But then, according to (25) it also holds for $\{,\}_{(\chi)}$ for all $x\in S^{\chi^C}$. Now we will prove that S^{χ^C} are the symplectic leaves. Since they are defined by equations $\chi_i^C=0$, $i=1,\ldots,2s$ on U we need to prove that $\{F,\chi_i^C\}_{\chi}(x)=0$, $x\in S^C$, for all $F\in \mathcal{F}(U)$, $i=1,\ldots,2s$. Using (25) and (26) we see that $\{F,\chi_i^C\}_{(\chi)}(x)=\{F,\chi_i^C\}_{(\chi^C)}(x)=\{F\,|\,S^{\chi^C},\chi_i^C\,|\,S^{\chi^C}\}_{\chi^C}(x)=0$. To finish the proof, observe first that, since $\chi_i=0$, $\psi_i=0$, $i=1,\ldots,2s, j=1,\ldots,r-2s$ define regularly the submanifold $S\subseteq U$, by shrinking U if necessary and for all C sufficiently small, we have that $\chi_i^C=0$, $\psi_i=0$, $i=1,\ldots,2s, j=1,\ldots,r-2s$ define regularly a submanifold $S^C\subseteq U$ and therefore $\psi_j\,|\,S^{\chi^C}=0$, $j=1,\ldots,r-2s$, define regularly S^C as a submanifold of S^{χ^C}. To prove that it is a first class constraint submanifold and that $\psi_j\,|\,S^{\chi^C}$, $j=1,\ldots,r-2s$, are first class constraints, that is $\psi_j\,|\,S^{\chi^C}\in I^{(S^C,S^{\chi^C})}$, $j=1,\ldots,r-2s$, we proceed in a similar fashion as we did in **(c)**, replacing χ by χ^C. The fact that $\psi_j\,|\,S^{\chi^C}$, $j=1,\ldots,r-2s$, are first class constraints defining S^C implies that $\dim S^{\chi^C}=\dim S^C+\dim\ker\omega^C$. From the definitions one can deduce that $\dim S^C=\dim S$ and $\dim S^{\chi^C}=\dim S^{\chi}$, therefore $\dim\ker\omega=\dim\ker\omega^C$.

The following theorem summarizes some essential parts of the previous lemma.

Theorem 3.20. *Let (P,Ω) be a symplectic manifold, $S\subseteq P$ and let ω be the pullback of Ω to S. Assume that $\ker\omega$ has constant rank. Let V be a vector subbundle of $TP\,|\,S$ such that $V\oplus\ker\omega=(TS)^{\Omega}$. Then there is a symplectic submanifold S^V containing S of dimension $\dim S+\dim\ker\omega$ such that the condition $T_xS^V=V_x^{\Omega}$, for all $x\in S$ holds. The vector bundle V^{Ω} is called the* ***second class subbundle*** *tangent to the second class submanifold S^V. For given functions F, G on P, one defines the Poisson bracket $\{F,G\}^V(x):=\{F\,|\,S^V,G\,|\,S^V\}(x)$, $x\in S$. We call this the V* ***-Dirac bracket on*** *S.*

Remark. **(a)** Since we are interested mainly in describing equations of motion we will not consider the definition of a global Poisson bracket on a neighborhood of S such that one of its symplectic leaves coincides with S^V. By choosing a local regular description $\chi_k=0$ of

S^V one obtains the usual expression for the Dirac bracket, as it was shown in lemma 3.19. Under our strong regularity conditions the symplectic leaves of the Dirac bracket give a (local) regular foliation of a neighborhood of the final constraint submanifold S. This implies by the Weinstein splitting theorem (Weinstein [42]) that there are local charts where the Dirac bracket is constant.

(b) The tangent second class subbundle V^Ω in a sense (modulo tangency) classifies all the possible second class constraint submanifolds containing a given submanifold $S \subseteq P$. It carries enough information to write the Dirac brackets along the final constraint submanifold S and therefore also equations of motion, as we show in subsection 3.4.

3.4 Equations of motion

We are going to describe equations of motion in the abstract setting of subsection 3.3, that is, a symplectic manifold (P,Ω) and a submanifold $S \subseteq P$, defined regularly by equations $\phi_i = 0$, $i = 1,\ldots,a$. We are going to assume all the results, notation and regularity conditions of that subsection. We need to introduce in this abstract setting, by definition, the notions of *primary constraints*, *primary constraint submanifold* and the *Hamiltonian*.

The ***primary constraint submanifold*** is a given submanifold $S' \subseteq P$ containing S, and in this context, S will be called the final constraint. We will assume without loss of generality that S' is defined regularly by the equations $\phi_i = 0$, $i = 1,\ldots,a'$, with $a' \le a$, where each ϕ_i, $i = 1,\ldots,a'$ will be called a ***primary constraint*** while each ϕ_i, $i = a'+1,\ldots,a$ will be called a ***secondary constraint***, for obvious reasons. In this abstract setting the ***Hamiltonian*** is by definition a given function $H \in \mathcal{F}(P)$.

The equations of motion can be written in the Gotay-Nester form,

$$\Omega(x)(\dot{x},\delta x) = dH(x)(\delta x),$$

where $(x,\dot{x}) \in T_x S$, for all $\delta x \in T_x S'$.

Now we will transform this equation into an equivalent Poisson equation using the Dirac bracket.

The condition $\{H,\psi\}(x) = 0$, for all $x \in S$ and all first class constraints ψ will appear later as a necessary condition for existence of solutions for any given initial condition in S, so we will assume it from now on.

The ***total Hamiltonian*** is defined by $H_T = H + \lambda^i \phi_i$, $i = 1,\ldots,a'$ where the functions $\lambda^i \in C^\infty(P)$, $i = 1,\ldots,a'$ must satisfy, by definition, $\{H_T,\phi_j\}(x) = 0$, $j = 1,\ldots,a$, $x \in S$ or, equivalently,

$$\{H,\phi_j\}(x) + \lambda^i \{\phi_i,\phi_j\}(x) = 0,\ x \in S,\ i = 1,\ldots,a',\ j = 1,\ldots,a, \tag{28}$$

sum over $i=1,\ldots,a'$. We assume that the solutions $(\lambda^1,\ldots,\lambda^{a'})$ form a nonempty affine bundle $\Lambda \to S$.

For each section of Λ one has a Hamiltonian $H_T(x)=H(x)+\lambda^i(x)\phi_i(x)$, $x\in P$, and an equation of motion on S,

$$X_{H_T}=(dH_T)^{\#}.$$

The equations of motion can be described nicely using the Dirac bracket as we will see in a moment. Choose first class and second class constraints

$$(\psi_1,\ldots,\psi_{a-2s},\chi_1,\ldots,\chi_{2s})$$

as in lemma 3.19, then since $\{H_T,\chi_i\}(x)=0$, for any function $F\in \mathrm{F}(P)$ and any $x\in S$ we obtain

$$\{H_T,F\}_{(\chi)}(x)=\{H_T,F\}(x)-\{H_T,\chi_i\}c_\chi^{ij}(x)\{\chi_j,F\}(x)=\{H_T,F\}(x),\ x\in S.$$

Then we can write the equations of motion in terms of the Dirac bracket as

$$X_{H_T}(x)=X_{(\chi),H_T}(x),\ x\in S. \tag{29}$$

We want a more precise description of the equations of motion. The total Hamiltonian has the equivalent expression

$$H_T=H+\lambda'^i\psi_i'+\mu'^j\chi_j', \tag{30}$$

where $\psi_i'\in I_{(S',P)}\cap R^{(S,P)}$, $i=1,\ldots,a'-s'$, are such that $(d\psi_1'(x),\ldots,d\psi_{a'-s'}'(x))$ form a basis of

$$\{d\psi(x)\,|\,\psi\in I_{(S',P)}\cap R^{(S,P)}\},$$

for all $x\in S$, while $\chi_i'\in I_{(S',P)}$, $i=1,\ldots,s'$, are such that

$$\left(d\psi_1'(x),\ldots,d\psi_{a'-s'}'(x),d\chi_1'(x),\ldots,d\chi_{s'}'(x)\right)$$

form a basis of

$$\{d\psi(x)\,|\,\psi\in I_{(S',P)}\}$$

for all $x\in S$. Then $\left(\psi_1',\ldots,\psi_{a'-s'}',\chi_1',\ldots,\chi_{s'}'\right)$ can be chosen as a set of primary constraints which justifies the expression (30) for the total Hamiltonian. One can see that the rank of the matrix $\{\chi_j',\chi_i\}(x)$, $i=1,\ldots,2s$, $j=1,\ldots,s'$, $x\in S$ is s'. Now the conditions (28) are equivalent to $\{H_T,\psi_j\}(x)=0$, $j=1,\ldots,a-2s$, which gives $\{H,\psi_j\}(x)=0$, $j=1,\ldots,a-2s$ and $\{H_T,\chi_i\}(x)=0$, $i=1,\ldots,2s$, for all $x\in S$ which gives

$$\{H,\chi_i\}(x)+\mu'^j(x)\{\chi_j',\chi_i\}(x)=0,$$

$i=1,\ldots,2s$, $x\in S$, from which we obtain μ'^j, $j=1,\ldots,s'$, as well-defined functions (on a neighborhood of S, then we can extend them arbitrarily to P). One can write $\chi'_j = a_j^k\psi_k + b_j^l\chi_l$ with $a_j^k, b_j^l \in \mathrm{F}\,(P)$ uniquely defined on S, for $j=1,\ldots,s'$, $k=1,\ldots,a-2s$, $l=1,\ldots,2s$. Then the total Hamiltonian can be written

$$\begin{aligned} H_T &= H + \lambda''^i\psi'_i + \mu'^j a_j^k\psi_k + \mu'^j b_j^l\chi_l \\ &= H + \psi_{(S,S',H)} + \lambda''^i\psi'_i + \mu'^j b_j^l\chi_l, \end{aligned}$$

where $\psi_{(S,S',H)} = \mu'^j a_j^k\psi_k$ is a first class constraint, $\psi_{(S,S',H)} \in I^{(S,P)}$.

We can conclude, using the fact that $\{\chi_k, F\}_{(\chi)}(x)=0$, for all $x\in S$ and $k=1,\ldots,2s$ and (29), that

$$\{H_T,F\}(x) = \{H_T,F\}_{(\chi)}(x) = \{H+\psi_{(S,S',H)}+\lambda''^i\psi'_i, F\}_{(\chi)}(x)$$

for any function $F\in\mathrm{F}\,(P)$ and any $x\in S$.

We shall call

$$H_{(\chi,S,S')} := H + \psi_{(S,S',H)} + \lambda''^i\psi'_{(S,S',i)}$$

the Hamiltonian of the system with respect to the Dirac bracket $\{,\}_{(\chi)}$, where we have denoted χ'_i by $\psi'_{(S,S',i)}$ to emphasize that these functions depend on S and S'.

We have proven the following theorem.

Theorem 3.21. *Let (P,Ω) be a symplectic manifold and $S\subseteq S'\subseteq P$ given regularly defined submanifolds. In the situation described above, equations of motion on S can be written in the following equivalent ways:*

(a)

$$\begin{aligned} X_{H_T}(x) &= X_{(\chi),H_{(\chi,S,S')}}(x) = X_{(\chi),H+\psi_{(S,S',H)}+\lambda''^i\psi'_i}(x) \\ &= X_{(\chi),H}(x) + X_{(\chi),\psi_{(S,S',H)}}(x) + \lambda''^i X_{(\chi),\psi'_i}(x) \end{aligned} \quad (31)$$

for all $x\in S$.

Here $H\in\mathrm{F}\,(P)$ satisfies, by definition, $\{H,\psi\}(x)=0$, for all $x\in S$ and all primary first class constraints ψ; $\psi_{(S,S',H)}$ is a first class constraint depending on S, S' and H; $\{\psi'_1,\ldots,\psi'_{a'-s'}\}\subset I_{(S'P)}\cap R^{(S,P)}$ is a maximal independent (that is $\{d\psi'_1(x),\ldots,d\psi'_{a'-s'}(x)\}$ is linearly independent for each $x\in S$) set of primary first class constraints and λ''^i, $i=1,\ldots,a'-s'$, are arbitrary parameters. There is uniqueness of solution if and only if there are no primary first class constraints, that is, $a'-s'=0$.

(b)

$$\dot F = \{F, H_{(\chi,S,S')}\}_{(\chi)}.$$

(c)

$$\Omega(x)(\dot{x},\delta x) = dH(x)(\delta x),$$

where $(x,\dot{x}) \in T_x S$ *, for all* $\delta x \in T_x S'$.

Remark. **(a)** The Hamiltonian vector field (31) depends on a finite number of arbitrary parameters $\lambda''^{i} \in \mathbb{R}$, $i = 1,\ldots,a' - s'$. It is tangent to S for all values of the parameters and it generates an affine distribution. This should be compared with the affine bundle in equation (7). Any vector field, even time-dependent, X_t on P whose restriction to S is a section of that distribution gives equations of motion. Note that, since $\psi'_i \,|\, S = 0$, then for any choice of functions, even time-dependent, $\lambda''^{i}_t \in \mathcal{F}(P)$, $i = 1,\ldots,a' - s'$ we have

$$\lambda''^{i}_t X_{(\chi),\psi'_i}(x) = X_{(\chi),\lambda''^{i}_t \psi'_i}(x),$$

for all $x \in S$.

(b) The equations of motion can be globalized, using the bracket $\{,\}^V$, as far as one can find a global $\psi_{(S,S',H)}$.

4. Future work

The present work should be followed immediately by an extension of the Dirac theory, and also the dual Gotay-Nester theory, for the case of a Dirac dynamical system $(x,\dot{x}) \oplus d\mathcal{E}(x) \in D_x$. This will expand the field of applications, for instance, one will have a theory of constraints for LC circuits and holonomic systems, if the Dirac structure D is integrable.

The reduction theory for the constraint algorithm will also be the purpose of future work. This can be approached using the category of Dirac anchored vector bundles. This will extend part of the results in Cendra et al. [8].

Singular cases, where the strong regularity assumptions made in the present paper are weakened in several ways are also important and will be the purpose of future work.

Appendix

Lemma A.1. *There is a canonical inclusion* $\varphi : TQ \oplus T^*Q \to T^*TQ$. *In addition, consider the canonical two-forms* ω_{T^*Q} *and* ω_{T^*TQ} *on* T^*Q *and* T^*TQ *respectively, the canonical projection* $\mathrm{pr}_{T^*Q} : TQ \oplus T^*Q \to T^*Q$, *and define the presymplectic two-form* $\omega = \mathrm{pr}^*_{T^*Q}\,\omega_{T^*Q}$ *on* $TQ \oplus T^*Q$. *Then the inclusion preserves the corresponding two-forms, that is,* $\omega = \varphi^*\omega_{T^*TQ}$.

Proof. If $\tau_Q : TQ \to Q$ and $\tau_{TQ} : TTQ \to TQ$ are the tangent projections, we can consider the dual tangent rhombic

$$\begin{tikzcd}
& TTQ \arrow[dl, "\tau_{TQ}"'] \arrow[dr, "T\tau_{Q}"] & \\
TQ \arrow[dr, "\tau_Q"'] & & TQ \arrow[dl, "\tau_Q"] \\
& Q &
\end{tikzcd}$$

Define $\varphi : TQ \oplus T^*Q \to T^*TQ$ by $\varphi(v_q \oplus \alpha_q) \in T^*_{v_q}TQ$,

$$\varphi(v_q \oplus \alpha_q) \cdot w_{v_q} = \alpha_q \cdot T\tau_Q(w_{v_q}),$$

for $w_{v_q} \in T_{v_q}TQ$. Here $v_q \oplus \alpha_q$ denotes an element in the Pontryagin bundle over the point $q \in Q$. Note that the following diagram commutes.

$$\begin{tikzcd}
TQ \oplus T^*Q \arrow[r, "\varphi"] \arrow[d, "\mathrm{pr}_{TQ}"'] & T^*TQ \arrow[d, "\pi_{TQ}"] \\
TQ \arrow[r, "\mathrm{Id}_{TQ}"] & TQ
\end{tikzcd}$$

Let us see that φ is an injective vector bundle map from the bundle $\mathrm{pr}_{TQ} : TQ \oplus T^*Q \to TQ$ to the cotangent bundle $\pi_{TQ} : T^*TQ \to TQ$, over the identity of TQ. The last part of this assertion follows from the commutative diagram above.

First, if $\varphi(v_q \oplus \alpha_q) = \varphi(v'_{q'} \oplus \alpha'_{q'})$ then both sides are in the same fiber $T^*_{v_q}TQ = T^*_{v'_{q'}}TQ$, so $v_q = v'_{q'}$. Also, for all $w_{v_q} \in T^*_{v_q}TQ$ we have

$$\varphi(v_q \oplus \alpha_q) \cdot w_{v_q} = \varphi(v_q \oplus \alpha'_q) \cdot w_{v_q}$$

or

$$\alpha_q \cdot T\tau_Q(w_{v_q}) = \alpha'_q \cdot T\tau_Q(w_{v_q}).$$

Since $\tau_Q : TQ \to Q$ is a submersion, we have $\alpha_q = \alpha'_q$ and φ is injective.

Second, φ is linear on each fiber, since

$$\varphi(v_q \oplus (\alpha_q + \lambda\alpha'_q)) \cdot w_{v_q} = (\alpha_q + \lambda\alpha'_q) \cdot T\tau_Q(w_{v_q}) = \varphi(v_q \oplus \alpha_q) \cdot w_{v_q} + \lambda\varphi(v_q \oplus \alpha'_q) \cdot w_{v_q}$$

For the second part of the lemma, let us recall the definition of the canonical one-form on $\theta_{T^*Q} \in \Omega^1(T^*Q)$. For $\alpha_q \in T^*Q$, $\theta_{T^*Q}(\alpha_q)$ is an element of $T^*_{\alpha_q}T^*Q$ such that for $w_{\alpha_q} \in T_{\alpha_q}T^*Q$,

$$\theta_{T^*Q}(\alpha_q) \cdot w_{\alpha_q} = \alpha_q(T\pi_Q(w_{\alpha_q})),$$

where $\pi_Q : T^*Q \to Q$ is the cotangent bundle projection. With a similar notation, the canonical one-form $\theta_{T^*TQ} \in \Omega^1(T^*TQ)$ is given by

$$\theta_{T^*TQ}(\alpha_{v_q}) \cdot w_{\alpha_{v_q}} = \alpha_{v_q}(T\pi_{TQ}(w_{\alpha_{v_q}})).$$

Pulling back these forms to the Pontryagin bundle by φ and the projection $\mathrm{pr}_{T^*Q} : TQ \oplus T^*Q \to T^*Q$, we obtain the same one-form. Indeed, for $w \in T_{v_q \oplus \alpha_q}(TQ \oplus T^*Q)$, we get on one hand

$$(\mathrm{pr}^*_{T^*Q}\theta_{T^*Q})(v_q \oplus \alpha_q) \cdot w = \theta_{T^*Q}(\alpha_q) \cdot T\mathrm{pr}_{T^*Q}(w) = \alpha_q \cdot T(\pi_Q \circ \mathrm{pr}_{T^*Q})(w),$$

and on the other hand

$$\begin{aligned}(\varphi^*\theta_{T^*TQ})(v_q \oplus \alpha_q) \cdot w &= \theta_{T^*TQ}(\varphi(v_q \oplus \alpha_q)) \cdot T\varphi(w) = \\ \varphi(v_q \oplus \alpha_q) \cdot T\pi_{TQ}(T\varphi(w)) &= \alpha_q \cdot T(\tau_Q \circ \pi_{TQ} \circ \varphi)(w).\end{aligned}$$

However, the following diagram commutes

$$\begin{tikzcd}
TQ \oplus T^*Q \arrow[rr, "\varphi"] \arrow[d, "\mathrm{pr}_{T^*Q}"'] \arrow[drr, "\mathrm{pr}_{TQ}"] & & T^*TQ \arrow[d, "\pi_{TQ}"] \\
T^*Q \arrow[dr, "\pi_Q"'] & & TQ \arrow[dl, "\tau_Q"] \\
& Q &
\end{tikzcd}$$

so $\pi_Q \circ \mathrm{pr}_{T^*Q} = \tau_Q \circ \pi_{TQ} \circ \varphi$ and therefore $\mathrm{pr}^*_{T^*Q}\theta_{T^*Q} = \varphi^*\theta_{T^*TQ}$. Taking (minus) the differential of this identity, we obtain $\omega = \varphi^*\omega_{T^*TQ}$. □

In local coordinates, if we denote the elements of $TQ \oplus T^*Q$ and T^*TQ by (q,v,p) and (q,v,p,v) respectively, then $\varphi(q,v,p) = (q,v,p,0)$.

Acknowledgements

– We thank the following institutions for providing us with means to work on this paper: Universidad Nacional del Sur (projects PGI 24/L075 and PGI 24/ZL06); Universidad Nacional de la Plata (project X553 Ecuaciones Diferenciales Implícitas); Agencia Nacional de Promoción Científica y Tecnológica, Argentina (projects PICT 2006- 2219 and PICT 2010-2746); CONICET, Argentina (project PIP 2010-2012 11220090101018); European Community, FP7 (project IRSES "GEOMECH" 246981).

References

[1] H. Bursztyn and M. Crainic. Dirac geometry, quasi-Poisson actions and D/G-valued moment maps. *J. Differential Geom.*, 82(**3**), 501 (2009).

[2] F. Cantrijn, J. F. Cariñena, M. Crampin, and L. A. Ibort. Reduction of degenerate Lagrangian systems. *J. Geom. Phys.*, 3(**3**) 353, (1986).

[3] J. F. Cariñena, J. Gomis, L. A. Ibort, and N. Román. Canonical transformations theory for presymplectic systems. *J. Math. Phys.*, 26(**8**), 1961 (1985).

[4] J. F. Cariñena. Theory of singular Lagrangians. *Fortschr. Phys.*, 38(**9**), 641 (1990).

[5] J. F. Cariñena and M. F. Rañada. Blow-up regularization of singular Lagrangians. *J. Math. Phys.*, 25(**8**), 2430 (1984).

[6] J. F. Cariñena, C. López, and N. Román-Roy. Origin of the Lagrangian constraints and their relation with the Hamiltonian formulation. *J. Math. Phys.*, 29(**5**), 1143 (1988).

[7] H. Cendra and M. Etchechoury. Desingularization of implicit analytic differential equations. *J. Phys. A*, 39(**35**), 10975 (2006).

[8] H. Cendra, J. E. Marsden, and T. S. Ratiu. Lagrangian reduction by stages. *Mem. Amer. Math. Soc.*, 152(**722**) (2001).

[9] H. Cendra, J. E. Marsden, S. Pekarsky, and T. S. Ratiu. Variational principles for Lie-Poisson and Hamilton-Poincaré equations. *Mosc. Math. J.*, 3(**3**), 833, 1197 (2003).

[10] T. Courant and A. Weinstein. Beyond Poisson structures. In *Action hamiltoniennes de groupes. Troisième théorème de Lie (Lyon, 1986)*, *Travaux en Cours*, **27**, 39 (1988).

[11] T. Courant. Dirac manifolds. *Trans. Amer. Math. Soc.*, 319(**2**), 631 (1990).

[12] M. de León and D. Martín de Diego. On the geometry of non-holonomic Lagrangian systems. *J. Math. Phys.*, 37(**7**), 3389 (1996).

[13] M. de León, D. Martín de Diego, and P. Pitanga. A new look at degenerate Lagrangian dynamics from the viewpoint of almost-product structures. *J. Phys. A*, 28(**17**), 4951 (1995).

[14] M. Delgado-Téllez and A. Ibort. On the geometry and topology of singular optimal control problems and their solutions. *Discrete Contin. Dyn. Syst.*, (suppl.) 223, (2003). Dynamical systems and differential equations (Wilmington, NC, 2002).

[15] P. A. M. Dirac. Generalized Hamiltonian dynamics. *Canadian J. Math.*, **2**, 129 (1950).

[16] P. A. M. Dirac. Generalized Hamiltonian dynamics. *Proc. Roy. Soc. London. Ser. A*, **246**, 326 (1958).

[17] P. A. M. Dirac. *Lectures on quantum mechanics*. Belfer Graduate School of Science Monographs Series. Belfer Graduate School of Science, New York (1964).

[18] M. J. Gotay and J. M. Nester. Presymplectic Lagrangian systems. I. The constraint algorithm and the equivalence theorem. *Ann. Inst. H. Poincaré Sect. A (N.S.)*, 30(**2**), 129 (1979).

[19] M. J. Gotay and J. M. Nester. Generalized constraint algorithm and special presymplectic manifolds. In *Geometric methods in mathematical physics (Proc. NSF-CBMS Conf., Univ. Lowell, Lowell, Mass., 1979)*, *Lecture Notes in Math.*, **775**, 78 (1980).

[20] M. J. Gotay and J. Sniatycki. On the quantization of presymplectic dynamical systems via coisotropic imbeddings. *Comm. Math. Phys.*, 82(**3**), 377 (1981/82).

[21] M. J. Gotay, J. M. Nester, and G. Hinds. Presymplectic manifolds and the Dirac -Bergmann theory of constraints. *J. Math. Phys.*, 19(**11**), 2388, (1978).

[22] X. Gràcia and J. M. Pons. Constrained systems: a unified geometric approach. *Internat. J. Theoret. Phys.*, 30(**4**), 511 (1991).

[23] X. Gràcia and J. M. Pons. A generalized geometric framework for constrained systems. *Differential Geom. Appl.*, 2(**3**), 223 (1992).

[24] M. Henneaux and C. Teitelboim. *Quantization of gauge systems*. Princeton University Press, Princeton, NJ (1992).

[25] A. Ibort, M. de León, G. Marmo, and D. Martín de Diego. Non-holonomic constrained systems as implicit differential equations. *Rend. Sem. Mat. Univ. Politec. Torino*, 54(**3**),

295 (1996). Geometrical structures for physical theories, I (Vietri, 1996).
[26] A. Ibort, M. de León, J. C. Marrero, and D. Martín de Diego. Dirac brackets in constrained dynamics. *Fortschr. Phys.*, 47(**5**), 459 (1999).
[27] O. Krupková. A geometric setting for higher-order Dirac-Bergmann theory of constraints. *J. Math. Phys.*, 35(**12**), 6557 (1994).
[28] O. Krupková. A new look at Dirac's theory of constrained systems. In *Gravity, particles and space-time*, p. 507. World Sci. Publ., River Edge, NJ (1996).
[29] G. H. Livens. On Hamilton's principle and the modified function in analytical dynamics. *Proc. Roy. Soc. Edinburgh*, **39**, 113 (1919).
[30] G. Marmo, G. Mendella, and W. M. Tulczyjew. Constrained Hamiltonian systems as implicit differential equations. *J. Phys. A*, 30(**1**), 277, (1997).
[31] G. Mendella, G. Marmo, and W. M. Tulczyjew. Integrability of implicit differential equations. *J. Phys. A*, 28(**1**), 149 (1995).
[32] N. Mukunda. Time-dependent constraints in classical dynamics. *Phys. Scripta*, 21(**6**), 801 (1980).
[33] N. Mukunda. The life and work of P. A. M. Dirac. In *Recent developments in theoretical physics (Kottayam, 1986)*, p. 260. World Sci. Publishing, Singapore (1987).
[34] F. Leon Pritchard. On implicit systems of differential equations. *J. Differential Equations*, 194(**2**), 328 (2003).
[35] P. J. Rabier and W. C. Rheinboldt. A geometric treatment of implicit differential-algebraic equations. *J. Differential Equations*, 109(**1**), 110 (1994).
[36] R. Skinner. First-order equations of motion for classical mechanics. *J. Math. Phys.*, 24(**11**), 2581 (1983).
[37] R. Skinner and R. Rusk. Generalized Hamiltonian dynamics. I. Formulation on $T^*Q \oplus TQ$ *J. Math. Phys.*, 24(**11**), 2589 (1983).
[38] R. Skinner and R. Rusk. Generalized Hamiltonian dynamics. II. Gauge transformations. *J. Math. Phys.*, 24(**11**), 2595 (1983).
[39] J. Sniatycki. Dirac brackets in geometric dynamics. *Ann. Inst. H. Poincaré Sect. A (N.S.)*, **2**, 365 (1974).
[40] E. C. G. Sudarshan and N. Mukunda. *Classical dynamics: a modern perspective*. Wiley -Interscience, John Wiley & Sons, New York (1974).
[41] A. J. van der Schaft. Equations of motion for Hamiltonian systems with constraints. *J. Phys. A*, 20(**11**), 3271 (1987).
[42] A. Weinstein. The local structure of Poisson manifolds. *J. Differential Geom.*, 18(**3**), 523 (1983).
[43] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics. I. Implicit Lagrangian systems. *J. Geom. Phys.*, 57(**1**), 133 (2006).
[44] H. Yoshimura and J. E. Marsden. Dirac structures in Lagrangian mechanics. II. Variational structures. *J. Geom. Phys.*, 57(**1**), 209 (2006).