Lecture I: Constrained Hamiltonian systems

(Courses in canonical gravity)

Yaser Tavakoli

December 15, 2014

1 Introduction

In canonical formulation of general relativity, geometry of space-time is given in terms of fields on spatial slices, Σ, whose geometry is encoded by a three metric h_{ab}, presenting the configuration variables. The space-time geometry is generally *covariant*, which is expressed by presence of constraint on the fields. These constraints are *diffeomorphism* and *Hamiltonian* constraints: More precisely, the diffeomorphism constraints generate deformation of spatial slices or coordinate changes. So, these constraints give the independence of the spatial geometry from the choice of space coordinates. The Hamiltonian constraints give the general covariance of space-time geometry for the time coordinates. It should be noted that, in this formulation, there is no absolute time, neither there is Hamiltonian generating evolution, just there is Hamiltonian constraint. From satisfying the Hamiltonian constraint, it is encoded the correlations between the physical fields of gravity and matter such that, the evolution in this framework is *relational*. The reproduction of a space-time metric in a coordinate-dependent way then requires one to choose a gauge and to compute the transformation in gauge parameters (including the coordinate) generated by the constraints.

It is quite often the case that theories of interest in modern physics are formulated as constrained systems. In particular, quantum gravity within canonical approach combines ideas from the constrained Hamiltonian systems for general relativity and Dirac's approach for quantization of these constrained systems. Dirac's theory of constrained Hamiltonian systems constitutes primary and secondary constraints, first-class and second-class constraints, and Dirac brackets. In quantum theory, the operator versions of first-class constraints become supplementary conditions on the wave function, provided these constraints are consistent with one another and with the Schrodinger equation. On the other hand, second-class constraints, become instead equations between quantum operators. Moreover, commutation relations are taken to correspond to Dirac-bracket relations, provided it is possible to find an irreducible representation of the Dirac-brackets algebra.

Thus we begin herein this chapter, by presenting Dirac's theory of constrained Hamiltonian systems, with emphasis on its application to canonical theory of gravity.

2 Lagrangian systems

A gauge theory may be thought of as one in which the dynamical variables are specified with respect to a "reference frame" whose choice is arbitrary at every instant of time. The physically important variables are those that are independent of the choice of the local reference frame. A transformation of the variables induced by a change in the arbitrary reference frame is called a *gauge transformation*. Physical variables (say "observables") are then said to be gauge invariant. We will see, in this chapter, that a gauge system is always a constrained Hamiltonian system. Therefore, we intend to study very briefly herein this chapter, the dynamics of constrained Hamiltonian system.

2.1 Phase space

A *phase space* is a space in which all possible states of a system are represented, with each possible state of the system corresponding to one unique point in the phase space. In a phase space, every degree of freedom or parameter of the system is represented as an axis of a multidimensional space; a one-dimensional system is called a phase line, while a two-dimensional system is called a phase plane. Furthermore, a phase space may contain very many dimensions. For instance, a gas containing many molecules may require a separate dimension for each particle's x, y and z positions and momenta as well as any number of other properties.

In classical mechanics the phase space coordinates are the generalized coordinates q^i and their conjugate generalized momenta p_i. The motion of an ensemble of systems in this space is studied by classical statistical mechanics. The local density of points in such systems obeys Liouville's Theorem, and so can be taken as constant. Within the context of a model system in classical mechanics, the phase space coordinates of the system at any given time are composed of all of the system's dynamical variables. Because of this, it is possible to calculate the state of the system at any given time in the future or the past, through integration of Hamilton's or Lagrange's equations of motion.

In quantum mechanics, the coordinates p and q of phase space normally become hermitian operators in a Hilbert space. But they may alternatively retain their classical interpretation, provided functions of them compose in novel algebraic ways (through Groenewold's star product), consistent with the uncertainty principle of quantum mechanics. Every quantum mechanical observable corresponds to a unique function or distribution on phase space, and vice versa. Expectation values in phase-space quantization are obtained isomorphically to tracing operator observables with the density matrix in Hilbert space: they are obtained by phase-space integrals of observables, with the Wigner quasi-probability distribution effectively serving as a measure.

Thus, by expressing quantum mechanics in phase space (the same ambit as for classical mechanics), the Weyl map facilitates recognition of quantum mechanics as a *deformation* (generalization) of classical mechanics, with deformation parameter $\hbar/S$, where S is the action of the relevant process. (Other familiar deformations in physics involve the deformation of classical Newtonian into relativistic mechanics, with deformation parameter v/c; or the deformation of Newtonian gravity into General Relativity, with deformation parameter Scwarzschild-radius/characteristic-dimension.)

Classical expressions, observables, and operations (such as Poisson brackets) are modified by $\hbar$-dependent quantum corrections, as the conventional commutative multiplication applying in classical mechanics is generalized to the noncommutative star-multiplication characterizing quantum mechanics and underlying its uncertainty principle.

2.2 The Lagrangian formalism

The usual procedure to determine the dynamics of a physical system with the Lagrangian $L(q^i, \dot{q}^i)$ follows an action principle. For a system of (finite) n configuration degrees of freedom q^i, where $i = 1, 2, ..., n$, the generic first-order action reads

$$S[q^i(t)] \;=\; \int L(q^i, \dot{q}^i) dt \; . \tag{2.1}$$

By using the usual variational techniques for the (stationary) action (2.1), the Euler–Lagrange equations is derived:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} \;=\; 0. \tag{2.2}$$

By expanding Eq. (2.2) we obtain a second order equation of motion:

$$W_{ij}\ddot{q}^j + \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j}\dot{q}^j - \frac{\partial L}{\partial q^i} \;=\; 0, \tag{2.3}$$

in which we have defined the matrix W_{ij} as

$$W_{ij} \;:=\; \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \; . \tag{2.4}$$

In general, a complete set of second-order equations of motion, coupled for all the n variables q^i, exists only if the matrix W_{ij} is non-degenerate. Then, at a given time, $\ddot{q}^j$ are uniquely determined by the positions and the velocities at that time; in other words, we can invert the matrix W_{ij} and obtain an explicit form for the equation of motion (2.3) as

$$\ddot{q}^j \; = \; (W_{ij})^{-1} \left(\frac{\partial L}{\partial q^i} - \frac{\partial^2 L}{\partial \dot{q}^i \partial q^k} \dot{q}^k \right) \; . \tag{2.5}$$

In this case, $\det(W_{ij})$ does not vanish. If, on the other hand, the matrix W_{ij} is not invertible, i.e. $\det(W_{ij}) = 0$, then, $\ddot{q}^j$ will not be uniquely determined by the positions and velocities, and the solution of (2.3) can contain arbitrary functions of time; such a system is said to be *gauge invariant*. Therefore, this is a key property of a gauge theory that the general solution of the equations of motion contains arbitrary functions of time, where the Lagrangian of the system is singular:

$$\det \left(\frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right) \; = \; 0 \; . \tag{2.6}$$

3 The Hamiltonian formulation

In Hamiltonian formulation, constraints can arise similar to the one we discussed previously in Lagrangian formulation. The departing point for the Hamiltonian formalism is to define the canonical momenta $p_i(q^j, \dot{q}^k)$ by

$$p_i \; = \; \frac{\partial L}{\partial \dot{q}^i} \; . \tag{3.1}$$

Using this relation in Eq. (2.4) we obtain $W_{ij} = \partial p_i / \partial \dot{q}^j$. Therefore, the relation (3.1) indicates n independent variables for p_is only if W_{ij} is invertible, such that one can at least locally solve for the $\dot{q}^i$.

If the matrix W_{ij} is not invertible, then there is no unique solution of Hamilton's equations of motion expressing the velocities in terms of the canonical coordinates q^i and conjugate momenta p_j (i.e. $\det(W_{ij}) = 0$). In that case, there exists certain relations connecting the momentum variables, of the type

$$\psi_s(q^i, p_j) \; = \; 0 \; . \tag{3.2}$$

The q's and p's are the dynamical variables of the Hamiltonian theory which are connected by the so-called *primary constraints* (3.2) of the Hamiltonian formalism. This indicates that, the map $(q^i, \dot{q}^i) \mapsto \left(q^i, p_j\right)$ (which is a one-to-one transformation of variables on the phase space, in the unconstrained case) maps the unconstrained phase space with coordinates $(q^i, \dot{q}^i)$ to the primary constraint surface $r : \psi_s(q^i, p_j) = 0$ of all points obtained as $\left(q^i, p_j(q,\dot{q})\right)$. Notice that, the unconstrained phase space of all $(q^i, \dot{q}^i)$, locally, has a complete set of coordinates given by $\left(q^i, p_j, \psi_s\right)$, however, globally, it is not straightforward to find the explicit expressions for the ψ_s.

Let us consider the Legendre transformation

$$H \;=\; \dot{q}^i p_i(q,\dot{q}) - L(q,\dot{q}) \;, \tag{3.3}$$

on the constrained manifold given by Eq. (3.2). It is seen that the relation for H in Eq. (3.3) refers to the time derivatives $\dot{q}^i$ rather than only to momenta p_j. Nevertheless, in order to H to be the *Hamiltonian* of a constrained system, the function H must be only a function of q^i and p_j; since in the constrained case, relation (3.1) is not invertible, one cannot replace all $\dot{q}^i$ in (3.3) by p_i, hence, no phase-space Hamiltonian as $H(q,p)$ would exist.

On the other hand, the Hamiltonian H must be always a well-defined functional of q^i and p_j, that is, by varying the $\dot{q}^i$ while p_j is fixed, the right-hand-side of Eq. (3.3) does not change. Therefore, the function $H(q^i, p_j)$ is well-defined. The variation of right-hand-side of Eq. (3.3) reads

$$\delta H \;=\; \dot{q}^i \delta p_i - \frac{\partial L}{\partial q^i}\delta q^i \;. \tag{3.4}$$

This relation shows that, variation of H involves only the variation of the q's and that of the p's; it does not involve the variation of the velocities $\dot{q}$'s. For a general variation of $H(q^i, p_j)$ on the momentum phase space, we obtain

$$\delta H \;=\; \frac{\partial H}{\partial q^i}\delta q^i + \frac{\partial H}{\partial p_i}\delta p_i \;. \tag{3.5}$$

By combining Eqs. (3.4) and (3.5) we find the equation

$$\left(\frac{\partial H}{\partial q^i} + \frac{\partial L}{\partial q^i}\right)\delta q^i + \left(\frac{\partial H}{\partial p_i} - \dot{q}^i\right)\delta p_i = 0 \;, \tag{3.6}$$

for $H(q^i, p_i)$ for any variation $(\delta q^i, \delta p_i)$ tangent to the primary constraint surface. Eq. (3.6) shows that the vector

$$V \;:=\; \left(\frac{\partial H}{\partial q^i} + \frac{\partial L}{\partial q^i} \;,\; \frac{\partial H}{\partial p_i} - \dot{q}^i\right) \;, \tag{3.7}$$

which satisfies the condition

$$V \;\cdot\; \begin{bmatrix} \delta q^i \\ \delta p_i \end{bmatrix} \;=\; 0 \;, \tag{3.8}$$

is normal to the constrained surface (3.2). Let us assume the surface (3.2) as $\psi_s = 0$ where $s = 1, ..., m$ (indicating to m independent constrained relations of this type), a basis of its normal space is given by

$$v_s \;:=\; \left(\frac{\partial \psi_s}{\partial q^i} \;,\; \frac{\partial \psi_s}{\partial p_j}\right), \tag{3.9}$$

which is the gradients of all the primary constraint functions. Thus, for some coefficients λ^s (which might be functionals on phase space) we have $V = \sum_s \lambda^s v_s$. Then, using this relation, together with Eqs. (3.7) and (3.9) we can derive the Hamiltonian equations of motion

$$\dot{q}^i = \frac{\partial H}{\partial p_i} - \lambda^s \frac{\partial \psi_s}{\partial p_i} , \tag{3.10}$$

$$\dot{p}_i = \frac{\partial L}{\partial q^i} = -\frac{\partial H}{\partial q^i} + \lambda^s \frac{\partial \psi_s}{\partial q^i} . \tag{3.11}$$

By a comparison to the Hamiltonian equations of motion of unconstrained systems, we can define a more general definition for the Hamiltonian of the system as

$$H_{\text{total}} = H - \lambda^s \psi_s . \tag{3.12}$$

In this case, we can rewrite the Hamiltonian equations of motion, in terms of the total Hamiltonian of the system, as

$$\dot{q}^i \approx \frac{\partial H_{\text{total}}}{\partial p_i} , \qquad \dot{p}_i \approx -\frac{\partial H_{\text{total}}}{\partial q^i} , \tag{3.13}$$

describing how the variables q^i and p_j vary in time, but, this equations involve unknown coefficients λ^s. Notice that, the 'weak equality' in equations above denotes an identity up to terms that vanish on the constraint surface.

Since terms $\lambda^s \psi_s$ in the total Hamiltonian vanish on the constraint surface, the value of the Hamiltonian does not change in the presence of the primary constraints, and is independent of λ^s. However, the evolution that Hamiltonian generates, which depends on derivatives of ψ_s, may not being independent of coefficients λ^s, if derivatives of ψ_s do not vanish. In order to investigate the role of λ^s on the evolution equations, it is convenient to introduce a certain formalism, namely the Poisson bracket formalism, which enables us to rewrite our equations briefly.

3.1 Poisson brackets

The Poisson bracket is an important binary operation in Hamiltonian mechanics, playing a central role in Hamilton's equations of motion, which govern the time-evolution of a Hamiltonian dynamical system. In a more general sense, the Poisson bracket is used to define a *Poisson algebra*, of which the algebra of functions on a *Poisson manifold* is a special case.

Definition. In canonical coordinates (q^i, p_j), on the phase space, given two functions $f(p,q)$, and $g(p,q)$, the Poisson bracket is defined as

$$\{f, g\} \;:=\; \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} \right) , \tag{3.14}$$

for a system with finitely many degrees of freedom.

The Poisson brackets have certain properties as following:

(i) They are antisymmetric in f and g:

$$\{f, g\} \;=\; -\{g, f\} \;.$$

(ii) They are linear in both entries:

$$\{f_1 + f_2,\; g\} \;=\; \{f_1,\; g\} + \{f_2,\; g\} \;.$$

(iii) They have the (Leibniz') product law

$$\{f_1 f_2,\; g\} \;=\; f_1 \{f_2,\; g\} + \{f_1,\; g\} f_2 \;.$$

also known as the "Poisson property".

(iv) They follow the *Jacobi identity*:

$$\{f,\; \{g, h\}\} + \{g,\; \{h, f\}\} + \{h,\; \{f, g\}\} \;=\; 0 \;.$$

If the Poisson bracket of f and g vanishes ($\{f, g\} = 0$), then f and g are said to be in mutual involution, and the operations of taking the Poisson bracket with respect to f and with respect to g commute.

With the help of Poisson brackets, we can now rewrite the equations of motion (3.13) as

$$\dot{q}^i \;\approx\; \{q^i,\; H_{\text{tot}}\} \;, \qquad\qquad \dot{p}_i \;\approx\; \{p_i,\; H_{\text{tot}}\} \;, \tag{3.15}$$

and for any arbitrary phase-space function $F(q^i, p_j)$, we can write

$$\dot{F} \;\approx\; \{F,\; H_{\text{tot}}\} \;. \tag{3.16}$$

This equation indicates that the total Hamiltonian H_{total} generates the dynamical flow of the phase-space variables in time.

From the conditions (ii) and (iii), the Poisson bracket can be always expressed as

$$\{f, g\} \;=\; \mathcal{P}^{ij} (\partial_i f)(\partial_j g) \;, \tag{3.17}$$

indicating that the Poisson structure can be captured by a contravariant 2-tensor (the so-called *Poisson tensor*), or a bivector $\mathcal{P}^{ij}$. Moreover, from condition (iv), the Jacobi identity, we have that $\mathcal{P}^{ij}$ must satisfy

$$\epsilon_{ikl}\mathcal{P}^{ij}\partial_j\mathcal{P}^{kl} \;=\; 0 \;. \tag{3.18}$$

In other words, it can be said that, any 2-tensor satisfying the condition (3.18) defines a Poisson structure (3.17).

A Poisson tensor $\mathcal{P}^{ij}$ may have an inverse which can be donated by $\Omega_{ij} := (\mathcal{P}^{-1})_{ij}$. The antisymmetry of tensors in Poisson geometry introduces an ambiguity, in lowering or raising the indices, which does not exist for Riemannian geometry; in Riemannian geometry, taking the inverse metric agrees with raising the indices of the metric by its inverse, however, this is not the case in Poisson geometry; a non-degenerate Poisson tensor defines a bijection (known as 'musical isomorphism'):

$$\begin{aligned} \mathcal{P}^\sharp :\; T^*M &\longrightarrow TM \;, \\ \alpha_i &\longmapsto \mathcal{P}^{ij}\alpha_j \;, \end{aligned} \tag{3.19}$$

which maps the co-tangent space to the tangent space of a manifold M, *raising* indices by contraction with $\mathcal{P}^{ij}$. Moreover, the inverse of the bijection $\mathcal{P}^\sharp$ defines a map $\mathbf{\Omega}^\flat = {\mathcal{P}^\sharp}^{-1}$ such that

$$\begin{aligned} \mathbf{\Omega}^\flat :\; TM &\longrightarrow T^*M \;, \\ v^i &\longmapsto (\mathcal{P})^{-1}_{ij}v^j \;=\; \Omega_{ij}v^j \;, \end{aligned} \tag{3.20}$$

lowering indices. Now, using tensor products one can lower or raise the indices of tensors of arbitrary degree; for example, we can lower the indices of $\mathcal{P}^{ij}$ using ${\mathcal{P}^\sharp}^{-1}$:

$$({\mathcal{P}^\sharp}^{-1})_{ik}({\mathcal{P}^\sharp}^{-1})_{jl}\mathcal{P}^{kl} \;=\; \delta^l_i({\mathcal{P}^\sharp}^{-1})_{jl} \;=\; -(\mathcal{P}^{-1})_{ij} \;=\; -\Omega_{ij} \;. \tag{3.21}$$

This indicates that, *in Poisson geometry, taking the inverse tensor agrees with raising the indices of the tensor by its inverse, only up to an opposite sign.* Therefore, we do not follow the convention of using the same letter for the tensor and its inverse, as we would do for a metric, but rather, keep separate symbols $\mathcal{P}^{ij}$ and Ω_{ij}.

Invertibility is not a general property of Poisson tensors. Nevertheless, the non-degenerate case with an existing inverse occurs often for Poisson tensors. This leads to several special properties, for example, if the inverse Ω_{ij} of $\mathcal{P}^{ij}$ exists, providing an antisymmetric covariant 2-tensor, it is called a *symplectic form* (see next section).

3.2 Symplectic geometry

Symplectic manifolds arise naturally in abstract formulations of classical mechanics as the cotangent bundles of manifolds, e.g., in the Hamiltonian formulation of classical mechanics, which provides one of the major motivations for the field: The set of all possible configurations of a system is modeled as a manifold, and this manifold's cotangent bundle describes the *phase space* of the system. The study of symplectic manifolds is called *symplectic geometry* or *symplectic topology*.

Definition. A symplectic manifold is a smooth manifold M, equipped with a *closed non-degenerate* differential 2-form Ω, called the symplectic form; this manifold is donated by (M, Ω). In other words, assigning a symplectic form Ω to a manifold M is referred to as giving M a *symplectic structure*.

The 'non-degeneracy' condition means that for all $p \in M$ we have the property that there does not exist non-zero $X \in T_pM$ such that $\Omega(X, Y) = 0$ for all $Y \in T_pM$. The 'skew-symmetric' condition means that for all $p \in M$ we have $\Omega(X, Y) = -\Omega(Y, X)$ for all $X, Y \in T_pM$. Recall that in odd dimensions antisymmetric matrices are not invertible. Since Ω is a differential two-form, the skew-symmetric condition implies that M has *even dimension*. The 'closed condition' means that the exterior derivative of Ω, namely $d\Omega$, is identically zero.

Definition. A diffeomorphism between two symplectic manifolds $f : (M, \Omega) \to (N, \Omega')$ is called *symplectomorphism*, if

$$f^*\Omega' \;=\; \Omega \;,$$

where f^* is the pullback of f. The symplectic diffeomorphisms from M to M are a (pseudo-)group, called the symplectomorphism group (see below).

The infinitesimal version of symplectomorphisms give the symplectic vector fields. A vector field $X \in \Gamma^\infty(TM)$ is called symplectic, if

$$\mathcal{L}_X \Omega \;=\; 0 \;.$$

Also, X is symplectic, iff the flow $\phi_t : M \to M$ of X is symplectic for every t. These vector fields build a Lie-subalgebra of $\Gamma^\infty(TM)$. Examples of symplectomorphisms include the canonical transformations of classical mechanics and theoretical physics, the flow associated to any Hamiltonian function, the map on cotangent bundles induced by any diffeomorphism of manifolds, and the coadjoint action of an element of a Lie group on a coadjoint orbit.

For a Poisson tensor $\mathcal{P}^{ij}$ the Jacobi identity (3.18) can be written as

$$\mathcal{P}^{ij}\partial_j\mathcal{P}^{kl} + \mathcal{P}^{kj}\partial_j\mathcal{P}^{li} + \mathcal{P}^{lj}\partial_j\mathcal{P}^{ik} = 0 \ . \tag{3.22}$$

Moreover, due to the inverse relationship between two tensors $\mathcal{P}^{ij}$ and Ω_{ij}, we have that $\Omega_{ij}\partial_k\mathcal{P}^{jl} = -\mathcal{P}^{jl}\partial_k\Omega_{ij}$; Using this in Eq. (3.22) we obtain

$$\mathcal{P}^{lk}\left(\partial_m\Omega_{nk} - \partial_n\Omega_{mk} + \partial_k\Omega_{mn}\right) \ = \ \mathcal{P}^{lk}\left(d\boldsymbol{\Omega}\right)_{mnk} \ = \ 0 \ , \tag{3.23}$$

which implies that the inverse of an invertible Poisson tensor is always a closed 2-form. Because of invertibility it is also non-degenerate. Therefore, the 2-form Ω_{ij} is a symplectic form, and the Poisson structure $(M, \mathcal{P})$ on a manifold M with an invertible Poisson tensor is equivalent to a symplectic structure $(M, \boldsymbol{\Omega})$.

Definition. Given a Poisson tensor, we associate the *Hamiltonian vector field* X_f to any function f on the Poisson manifold by

$$X_f : \ \mathcal{P}^\sharp df \ = \ -\mathcal{P}^{ij}(\partial_i f)\partial_j \ .$$

In terms of the Poisson bracket, one can write $X_f = \{\cdot, f\}$ as the action of the vector field on functions, to be inserted for the 'dot'.

(In Riemannian geometry, an analogous construction provides the normal vector to the surface given by $f = const.$) The Poisson bracket itself can be written in terms of Hamiltonian vector fields and the symplectic form:

$$\boldsymbol{\Omega}(X_f, X_g) \ = \ \Omega_{ij}\mathcal{P}^{ik}\mathcal{P}^{jl}\partial_k f\partial_l g \ = \ -\mathcal{P}^{ki}\partial_k f\partial_i g \ = \ -\{f, g\} \ . \tag{3.24}$$

The Hamiltonian vector field can be interpreted as the phase-space direction of change corresponding to the function f; for $f = p$ we have $X_p = \partial/\partial_q$. When f is one of the canonical coordinates, its Hamiltonian vector field is along its canonical momentum. In this sense, the Hamiltonian vector field generalizes the notion of momentum to arbitrary phase-space functions. Integrating the Hamiltonian vector field X_f to a 1-parameter family of diffeomorphisms, we obtain the Hamiltonian flow $\Phi_t^{(f)}$ generated by f. The dynamical flow of a canonical system is generated by the Hamiltonian function on phase space.

Definition. A *non-invertible* closed 2-form is called a *presymplectic form*; this provides the manifold on which it is defined with presymplectic geometry. Notice that, the presymplectic geometry can constitute of non-invertible Poisson tensors, providing a Poisson geometry, but, such geometry does not have an equivalent symplectic formulation.

Consider a 2-form Ω_{ij} defined as

$$\begin{aligned} \mathbf{\Omega}: \; TM &\longrightarrow T^*M \\ v^i &\longmapsto \Omega_{ij} v^j \; . \end{aligned}$$

Now, the kernel $Ker(\Omega)$ can be defined as,

$$Ker(\mathbf{\Omega}) \; = \; \{v^i \in TM : \quad \Omega_{ij} v^j = 0\} \; . \tag{3.25}$$

Therefore, these definitions imply that, a presymplectic form Ω_{ij} has a kernel $C \in TM$ of vector fields v^i satisfying $\Omega_{ij} v^j = 0$.

The vector fields v^j define a flow on phase space, which can be factored out by identifying all points on orbits of the flow. The resulting factor space is symplectic: every vector field in the kernel of Ω_{ij} is factored out. In this way, a reduced symplectic geometry can be associated with any presymplectic geometry.

Definition. Given a Poisson tensor $\mathcal{P}$, a *Casimir function* is defined as an element $C^I \in \mathcal{P}$ such that

$$\{C^I, f\} \; = \; \mathcal{P}^{ij} (\partial_i C^I)(\partial_j f) \; = 0 \; . \tag{3.26}$$

Therefore, using (3.26) we can show that for any Casimir function C^I there exists a one-form $dC^I := (\partial_i C^I)\partial_i$ such that $\mathcal{P}^\sharp(dC^I) = \mathcal{P}^{ij}(\partial_i C^I)\partial_j = 0$ which is a zero vector field. Thus, the one-forms dC^I are in the kernel of $\mathcal{P}^{ij}$.

3.3 Constraints on symplectic manifolds

Consider a symplectic manifold M, with symplectic form Ω_{ij} and Poisson tensor $\mathcal{P}^{ij}$ with a smooth Hamiltonian over it (for field theories, M would be infinite-dimensional). If we constrain the symplectic structure $(M, \mathbf{\Omega})$ to a subset $\mathcal{C}$ defined by the vanishing of constraint functions C^I:

$$C^I \; \approx \; 0 \; ,$$

there are different possibilities for symplectic properties of the subset $\mathcal{C}$. These properties are mainly determined by the Hamiltonian vector fields of the constraints (which are assumed to be non-vanishing in the neighborhood of the constraint surface):

Definition. We call a constraint C^I, *first class* with respect to all constraints if its Hamiltonian vector field is everywhere tangent to the constraint surface $\mathcal{C}$. We call it *second class*, if its Hamiltonian vector field is nowhere tangent to the constraint surface. Following the definition of the Hamiltonian vector field, this is equivalent to saying that, for all constraints C^J on the constraint surface

$$\{C^I, C^J\} \;=\; 0 \;, \tag{3.27}$$

if the constraint C^I is first class; it vanishes nowhere on the constraint surface if C^I is second class. (No condition is posed for the behavior of these Poisson brackets off the constraint surface.)

In summary, in a constrained Hamiltonian system, a dynamical quantity is called a first class constraint if its Poisson bracket with all the other constraints vanishes on the constraint surface (the surface implicitly defined by the simultaneous vanishing of all the constraints). The surface is called a *first-class constraint surface* if all constraints defining it are first class, and a *second-class constraint surface* if all constraints are second class.

We can equip the constraint surface with a presymplectic form $\bar{\mathbf{\Omega}}$ by pulling back the symplectic form to it. If

$$\begin{aligned} \iota :\; \mathcal{C} &\longrightarrow M \\ y^\alpha &\longmapsto x^i \;, \end{aligned} \tag{3.28}$$

is the embedding of the constraint surface in M, we write

$$\bar{\mathbf{\Omega}} \;=\; \iota^* \mathbf{\Omega} \;, \tag{3.29}$$

so that

$$\bar{\Omega}_{\alpha\beta} \;=\; \Omega_{ij} \frac{\partial x^i}{\partial y^\alpha} \frac{\partial x^j}{\partial y^\beta} \;. \tag{3.30}$$

Notice that we cannot directly equip constraint surface with a Poisson structure, since Poisson tensor, being contravariant, cannot be pulled back. If one of the constraints is first class, say C, the presymplectic form is degenerate: its Hamiltonian vector field X^i_C is tangent to the constraint surface and thus defines a vector field X^α on it by simple restriction. So that, by using the definition of Hamiltonian vector fields, we have

$$\bar{\Omega}_{\alpha\beta} X^\beta \;=\; \iota^*(\Omega_{ij} X^j_C) \;=\; \iota^*(\partial_i C) \;=\; 0. \tag{3.31}$$

Only if all constraints are second class, does a symplectic structure result on the constraint surface. In this case, the constraint surface can be used directly, as the phase space of the reduced system where the constraints are solved. If there are first class constraints, their Hamiltonian flow must be factored out to obtain the *reduced phase space* as the factor space of the presymplectic constraint surface by the Hamiltonian flow. Physically, this flow is the gauge flow generated by the constraints.

3.3.1 Dirac brackets

If one does not solve all the constraints, when using the Hamiltonian flows of phase space functions, most importantly the dynamical flow generated by the Hamiltonian, one must be careful that they do not leave the constraint surface. The *Dirac bracket* provides a modification to the Poisson brackets so as to ensure that the Hamiltonian flow generated by the old constraints with respect to the new Poisson structure is tangent to the constraint surface. More precisely, the two-form implied from the Dirac bracket is the restriction of the symplectic form to the constraint surface in phase space.

Definition. The *Dirac bracket* is defined as

$$\{f,g\}_{\mathrm{D}} \;:=\; \{f,g\} - \sum_{IJ}\{f,C^I\}\left(\{C^I,C^J\}\right)^{-1}\{C^J,g\}\,, \tag{3.32}$$

where the double sum is taken for all second-class constraints, for which the matrix inverse of $\{C^I,C^J\}$ is guaranteed to exist.

For any second class constraint C^K, using the definition of the Dirac bracket (3.32), we find that

$$\{f,C^K\}_{\mathrm{D}} \;:=\; \{f,C^K\} - \sum_{IJ}\{f,C^I\}\left(\{C^I,C^J\}\right)^{-1}\{C^J,C^K\} = 0, \tag{3.33}$$

which indicates that, the flow generated by the second class constraint vanishes and hence, it does not leave the constraint surface. Therefore, in using the Dirac bracket, only the flow of first class constraints need to be considered. Indeed, since the matrix $\{C^I,C^J\}$ is not invertible in the presence of the first class constraints, Dirac bracket cannot be removed in a similar way as mentioned for the second class constraints. In this case, to obtain the phase space with symplectic structure, the flows generated by first-class constraints must be factored out.

3.3.2 Constraint algebras

In addition to the primary constraints, involved in the total Hamiltonian (3.12), there are further equations to be satisfied by initial values: Consistency conditions are required because the time derivative of the primary constraints ψ_s, like the constraints themselves, must vanish at all time:

therefore, using (3.16) we can write

$$\dot{\psi}_s \ \approx\ \{\psi_s, H\} - \lambda^t\{\psi_s, \psi_t\} \ =:\ \{\psi_s, H\} - \lambda^t C_{st} \ =\ 0. \tag{3.34}$$

In this way, the structure of the constraint system is determined by the matrix $C_{st} := \{\psi_s, \psi_t\}$ introduced here: if $\det(C_{st}) \neq 0$, no further constraints result and we can fulfill the consistency condition (3.34) by solving them for all λ^t. If $\det(C_{st}) = 0$, not all λ^t can be determined to solve the consistency conditions completely. In this case, (3.34) implies *secondary constraints* which follow from the equations of motion, rather than from the basic definition of momenta as the primary constraints.

Definition. For any zero-eigenvector Z^s_r of the primary constraint matrix $Z^s_r C_{st} = 0$, *secondary constraints* in general take the form

$$Z^s_r\{\psi_s, H\} \ =\ 0\ . \tag{3.35}$$

If a non-trivial zero-eigenvector Z^s_r exists, $\psi_s = 0$ is preserved in time only if (3.35) is satisfied.

3.3.3 Gauge transformation

First class constraints generate *gauge transformations*. For any phase space function F, defined for any first class constraint C_s, the infinitesimal mapping

$$F(q,p) \ \longmapsto\ F(q,p) + \delta^{(s)}_\epsilon F(q,p) \ :=\ F(q,p) + \{F, \epsilon C_s\} \tag{3.36}$$

maps solutions to the constraints and equations of motion into other solutions: under this mapping, $\delta^{(s)}_\epsilon C_t \approx 0 \approx \delta^{(s)}_\epsilon H$. Interpreting transformations (3.36) as *gauge* means that we do not consider solutions mapped to each other by the Hamiltonian flow of first class constraints as physically distinct. Gauge transformations map different mathematical solutions into each other, but they are interpreted merely as different representations of the same physical solution.

Definition. For any phase space function f, if f has a complete vanishing Poisson bracket with all first-class constraints, it is called a complete *(or Dirac) observable.*

In order to have a well-defined theory, with unambiguous physical predictions, the infinitesimal flow $\{f, C_S\}$ only changes the mathematical representation but does not change the physics of the observable information. The change of f is just a gauge transformation without affecting the physical state.

In summary, in totally constrained systems, constraints play several roles as following:

- they constrain allowed field values to reside on the constraint surface;
- they generate gauge transformations as identifications of physically equivalent field configurations on the constraint surface
- they provide the total Hamiltonian of the system; this means that a particular combination of the constraints, generates Hamiltonian equations of motion in a time coordinate.

Any particular choice for the total Hamiltonian will result in equations of motion written in specific gauge. But since the theory is invariant under gauge transformations generated by constraints, the choice of a total Hamiltonian does not matter, and all sets of equations of motion obtained for different gauges are equivalent.

References

[1] P. A. M. Dirac, *Lecture in quantum mechanics*, (Dover, 2000)

[2] Marc Henneaux and Claudio Teitelboim, *Quantization of Gauge Systems*, (Princeton University Press, 1991).

[3] Martin Bojowald, *Canonical gravity and applications*, (Cambridge University Press, 2010).