GEOMETRY AND DYNAMICS OF NONHOLONOMIC SYSTEMS

LUIS C. GARCÍA-NARANJO

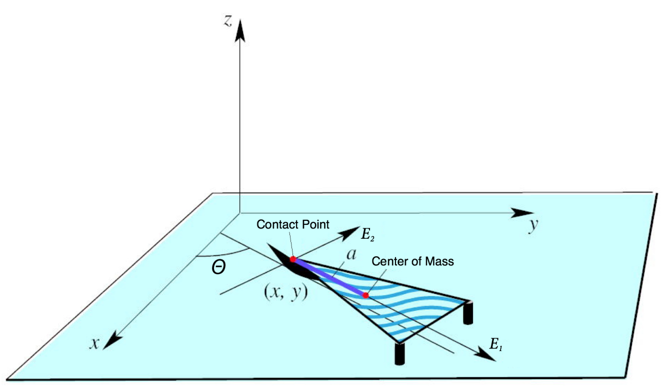
Abstract. This set of notes was prepared for the mini-course *Geometry and Dynamics of Nonholonomic Systems* given by the author at the 10^{th} ICMAT International Summer School on Geometry, Mechanics and Control organized by the GMC network from 20-24 June 2016 at La Cristalera, Miraflores de la Sierra, Madrid, Spain.

1. Examples

Mechanical systems with constraints on the velocities that are not derivatives of constraints in positions are termed nonholonomic. We introduce a series of examples that will be revisited throughout the course. Note that in all of the examples the constraints are linear and homogeneous on the velocities.

1.1. **Chaplygin sleigh.** Constraint:

$$\dot{y}\cos\theta = \dot{x}\sin\theta$$



Lagrangian:

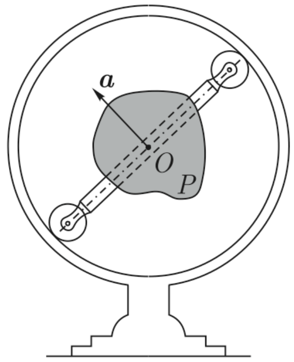
$$L(x,y,\theta,\dot{x},\dot{y},\dot{\theta}) = \frac{1}{2}\left((J+ma^2)\dot{\theta}^2 + m(\dot{x}^2+\dot{y}^2) + 2ma\dot{\theta}(\cos\theta\dot{y} - \sin\theta\dot{x})\right).$$

1.2. **Suslov problem.** Constraint $\langle a, \Omega\rangle = 0$. Ω is the angular velocity written in the body frame.

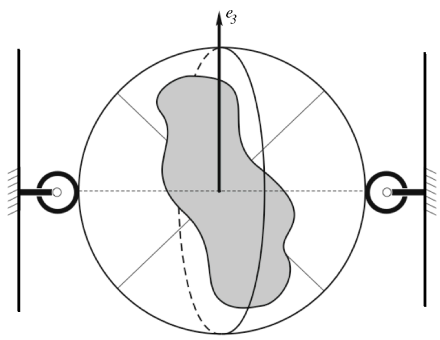
Lagrangian

$$L = \frac{1}{2}\langle \mathbb{I}\Omega, \Omega\rangle$$

($\mathbb{I}$ is the inertia tensor).



1.3. **Veselova problem.** Constraint $\langle e_3, \omega \rangle = \omega_3 = 0$. Here ω is the angular velocity written in the space frame.



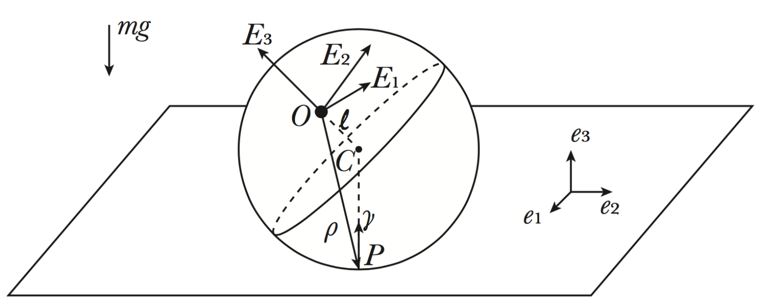
Lagrangian

$$L = \frac{1}{2} \langle \mathbb{I}\Omega, \Omega \rangle$$

1.4. **Chaplygin top (inhomogeneous sphere rolling on the plane).** Constraint $\dot{u} = B(\rho \times \Omega)$.

Here B is attitude matrix and ρ is the vector from O to P written in the body frame. C is geometric center of the sphere. O is center of mass. P is contact point.

u is the position of the center of mass O expressed in space coordinates.

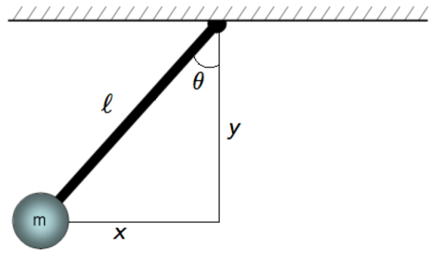


Lagrangian

$$L = \frac{1}{2}\langle \mathbb{I}\Omega, \Omega\rangle + \frac{m}{2}||\dot{u}||^2.$$

1.5. **Pendulum.** (Not really nonholonomic but we include it for comparison).

Constraint $x^2 + y^2 = \ell^2$ or $x\dot{x} + y\dot{y} = 0$.



Constraint: $x^2 + y^2 = \ell^2$

Lagrangian

$$L(x, y, \dot{x}, \dot{y}) = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - mgy.$$

2. Geometry of constraints that are linear in the velocities

The configuration space of a mechanical system is a smooth manifold Q^n whose coordinates $q^1, \ldots, q^n$ specify the configuration (position) of the system. The tangent bundle TQ is the space of positions and velocities. For the examples introduced above:

(i) Chaplygin sleigh: $Q = \mathbb{R}^2 \times S^1$. It is also convenient to think of $Q = SE(2)$ (not only a manifold but a Lie group).
(ii) Suslov problem: $Q = SO(3)$. This is the usual configuration space for rigid body dynamics.
(iii) Veselova problem: $Q = SO(3)$.
(iv) Chaplygin top: $Q = SO(3) \times \mathbb{R}^2$.
(v) Pendulum: $Q = \mathbb{R}^2$ (really S^1 because the constraint is integrable).

2.1. **Vector fields.** Recall that a vector field X on Q assigns a vector $X(q) \in T_qQ$ to each $q \in Q$ in a smooth manner. In local coordinates

$$X(q) = X^1(q)\frac{\partial}{\partial q^1} + \cdots + X^n(q)\frac{\partial}{\partial q^n} = X^j(q)\frac{\partial}{\partial q^j}$$

for some smooth functions $X^j : Q \to \mathbb{R}$. Here $\{\frac{\partial}{\partial q^1}, \ldots, \frac{\partial}{\partial q^n}\}$ is a basis of T_qQ induced by the coordinates.

A vector field defines an autonomous ordinary differential equation on Q. In local coordinates it is given by

$$\dot{q}^i = X^i(q^1, \ldots, q^n), \qquad i = 1, \ldots, n.$$

By the theorem of existence and uniqueness it defines a smooth flow $\phi_t : Q \to Q$ (existing at least for small $|t|$) such that $c(t) = \phi_t(q)$ is the unique curve that satisfies

$$c'(t) = X(c(t)), \qquad c(0) = q.$$

We call $c(t)$ an integral curve of X. (Note $(\phi_t)^{-1} = \phi_{-t}$ so ϕ_t is a diffeomorphism if t is small enough).

Example 2.1. *Take $Q = \mathbb{R}^2 \times S^1$. We can use coordinates (θ, x, y). Let $X_1 = \frac{\partial}{\partial \theta}$, $X_2 = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}$, $X_3 = -\sin\theta \frac{\partial}{\partial x} + \cos\theta \frac{\partial}{\partial y}$*

For X_1 we have

$$\phi_t(\theta, x, y) = (\theta + t, x, y).$$

For X_2 we have

$$\phi_t(\theta, x, y) = (\theta, x + t\cos\theta, y + t\sin\theta).$$

For X_3 we have

$$\phi_t(\theta, x, y) = (\theta, x - t\sin\theta, y + t\cos\theta).$$

2.2. **Vector fields act on functions.** For $f : Q \to \mathbb{R}$ we define the new function $X[f] = \pounds_X(f)$ by the rule

$$X[f](q) = \left.\frac{d}{dt}\right|_{t=0} f(\phi_t(q)).$$

We can also write

$$X[f](q) = \langle df(q), X(q) \rangle$$

or, in local coordinates,

$$X[f](q) = \frac{\partial f}{\partial q^i}(q) X^i(q).$$

2.3. **Lie bracket of vector fields.** If X and Y are vector fields on Q, and ϕ_t is the flow of X, we define the pull-back of Y by ϕ_t as the vector field on Q given by

$$\phi_t^*(Y)(q) = T_{\phi_t(q)}\phi_{-t}(Y(\phi_t(q))).$$

Example 2.2. *Recall*

$$X_1 = \frac{\partial}{\partial \theta}, \qquad X_2 = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}.$$

The pull-back of X_2 by the flow ϕ_t of X_1 is

$$(\phi_t)^*(X_2)(x, y, \theta) = \cos(\theta + t)\frac{\partial}{\partial x} + \sin(\theta + t)\frac{\partial}{\partial y}.$$

The Lie bracket $[X, Y]$ of the vector fields X, Y is the vector field defined by

$$[X, Y](q) = \lim_{t \to 0} \frac{\phi_t^* Y(q) - Y(q)}{t}$$

Example 2.3. *Example: $X_1 = \frac{\partial}{\partial \theta}$, $X_2 = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}$. The Lie bracket of X_1 and X_2 is*

$$\begin{aligned}[X_1, X_2](x, y, \theta) &= \lim_{t \to 0} \frac{1}{t}\left(\cos(\theta + t)\frac{\partial}{\partial x} + \sin(\theta + t)\frac{\partial}{\partial y} - \left(\cos(\theta)\frac{\partial}{\partial x} + \sin(\theta)\frac{\partial}{\partial y}\right)\right) \\ &= -\sin\theta \frac{\partial}{\partial x} + \cos\theta \frac{\partial}{\partial y} = X_3.\end{aligned}$$

In practice one can rarely apply this definition since we do not have a formula for the flow of a vector field. Brackets are efficiently computed using the following result.

Theorem. If X, Y are vector fields on Q and $f : Q \to \mathbb{R}$ then

$$[X, Y][f] = X[Y[f]] - Y[X[f]].$$

It follows that if

$$X = X^i \frac{\partial}{\partial q^i}, \qquad Y = Y^i \frac{\partial}{\partial q^i}$$

then

$$[X, Y] = DY \cdot X - DX \cdot Y.$$

A fundamental geometric object in nonholonomic systems is a *distribution*.

Definition 2.4. *A distribution $\mathcal{D}$ on Q is an assignment $q \mapsto \mathcal{D}_q \subset T_qQ$, where $\mathcal{D}_q$ is a subspace of T_qQ, such that around every point $q_0 \in Q$ there is a neighbourhood U of q_0 such that for all $q \in U$*

$$\mathcal{D}_q = \operatorname{span}\{X_1(q), \dots, X_k(q)\}$$

for smooth vector fields $X_1, \dots, X_k$.

The dimension of $\mathcal{D}_q$ is called the rank of $\mathcal{D}$ at $q \in Q$. If the rank of the distribution is a constant independent of $q \in Q$ we call such distribution *regular*. Then $\mathcal{D}$ is a subbundle of TQ that can be interpreted as a submanifold $D \subset TQ$.

We will consider regular distributions that are defined as the annihilator of a set of independent one-forms.

Example 2.5. *Take again $Q = S^1 \times \mathbb{R}^2$ with coordinates (θ, x, y) and let $\mathcal{D}$ be the annihilator of the one-form $\beta = \dot{y} \cos\theta - \dot{x} \sin\theta$. Then*

$$\mathcal{D} = span\{X_1, X_2\}$$

is a regular distribution of rank 2.

Definition 2.6 (Involutive distribution)**.** *The distribution $\mathcal{D}$ is called* involutive *if for any vector fields X, Y such that $X(q), Y(q) \in \mathcal{D}_q$ for $q \in U \subset Q$ open, we have*

$$[X, Y](q) \in \mathcal{D}_q.$$

Definition 2.7 (Integral manifold - local theory)**.** *An integral manifold through a point $q_0 \in Q$ is an immersed submanifold P of a neighbourhood $q_0 \in U \subset Q$ with the property that $T_qP = \mathcal{D}_q$ for all $q \in U$.*

Definition 2.8 (Integrable distribution)**.** *A distribution is* integrable *if there exists an integral manifold passing through each $q \in Q$.*

It follows that if a distribution is regular and integrable, then around every point $q_0 \in Q$ there exist local coordinates $(q^1, \dots q^k, q^{k+1}, \dots, q^n)$ such that the sets $q^{k+1} = c^{k+1}, \dots, q^n = c^n$ are integral manifolds. Using this, it is straightforward to show that an integrable distribution is necessarily involutive. The converse implication is the content of Frobenius Theorem.

Theorem 2.9 (Frobenius 1877)**.** *If $\mathcal{D}$ is constant rank then $\mathcal{D}$ is involutive if and only if $\mathcal{D}$ is integrable.*

Example 2.10. *The distribution $\mathcal{D}$ spanned by X_1 and X_2 is not involutive and therefore not integrable.*

Example 2.11. *Any rank one distribution is integrable. In particular, the distribution on $\mathbb{R}^2 \setminus \{(0,0)\}$ defined as the null-space of*

$$\beta = x\,dx + y\,dy.$$

In polar coordinates (r,θ), the integral manifolds are $r = a$; which are circles of radius a.

The constraints on the velocities of a mechanical system are nonholonomic if they define a non-integrable distribution on the configuration space Q. If such distribution is integrable, then the constraints are holonomic.

3. The Lagrange-D'Alembert principle

Our approach is based on the Lagrangian formulation of mechanics (Lagrange 1788). We assume that $q = (q^1, \dots, q^n)$ are any set of coordinates on a manifold Q that specify the *configuration* of the system and that the motion of the system is subjected to constraints that are linear and homogeneous on the velocities, say

$$\beta^a_k(q)\dot{q}^k = 0, \qquad a = 1, \dots, n-r, \tag{3.1}$$

where the vectors $\beta^a(q)$ are linearly independent. The constraints define a regular distribution $\mathcal{D}$ on Q and a *constraint submanifold* $D \subset TQ$.

We assume that the only forces acting on the system are conservative forces arising from a potential energy function $V = V(q)$ and the reaction forces. The equations of motion are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = R_i, \qquad i = 1, \dots n. \tag{3.2}$$

Here R_i are the components of the reaction force in this coordinate system and the Lagrangian $L = L(q,\dot{q}) = \frac{1}{2}\dot{q}^T A(q)\dot{q} - V(q)$ where the kinetic energy matrix $A(q)$ is symmetric and positive definite.

In order to obtain a closed set of equations of motion, we need to invoke a physical principle to specify the reaction force R. This is the *Lagrange-D'Alembert principle* of ideal constraints. It assumes that the reaction force R annihilates any possible displacement of the system. Namely, if $\dot{q}$ satisfies (3.1) then $R_i\dot{q}^i = 0$. Consequently, the reaction force R performs no work during the motion. It neither adds or takes away energy from the system.

If we define $E(q,\dot{q}) = \dot{q}^i \frac{\partial L}{\partial \dot{q}^i} - L$ then, *if the constraints* (3.1) *are satisfied*, a direct calculation *that uses* $R_i\dot{q}^i = 0$ shows that E is constant along the motion.

3.1. **Determination of the reaction force.** Our approach follows [1].

Denote by $\beta(q)$ the $(n-r)\times n$ matrix whose a^{th} row has entries $\beta^a_i(q)$. Under our assumptions, the matrix $\beta(q)$ has rank $n-r$.

The Lagrange-D'Alembert principle implies that R must be a linear combination of the rows of $\beta(q)$. Hence $R = \beta(q)^T\lambda$ for a vector $\lambda \in \mathbb{R}^{n-k}$. The entries λ_a of λ are sometimes called multipliers or Lagrange multipliers (even if they are *not necessarily* Lagrange multipliers arising from the theory of finding extrema with constraints). We shall see that λ is a function of $(q,\dot{q})$.

Differentiating (3.1) yields

$$\beta(q)\ddot{q} + \gamma(q, \dot{q}) = 0, \tag{3.3}$$

where $\gamma(q, \dot{q}) \in \mathbb{R}^{n-r}$ has components

$$\gamma^a(q, \dot{q}) = \frac{\partial \beta^a_k}{\partial q^j} \dot{q}^j \dot{q}^k.$$

Using that $L = \frac{1}{2}\dot{q}^T A(q)\dot{q} - V(q)$ we have

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = A(q)\ddot{q} + \eta(q, \dot{q}) + V'(q) \tag{3.4}$$

where the components of the vectors $\eta(q, \dot{q})$, $V'(q) \in \mathbb{R}^n$ are

$$\eta_i(q, \dot{q}) = \left(\frac{\partial A_{ij}}{\partial q^k}(q) - \frac{1}{2}\frac{\partial A_{jk}}{\partial q^i}(q)\right)\dot{q}^j\dot{q}^k, \qquad V'_i(q) = \frac{\partial V}{\partial q^i}(q).$$

Therefore (3.2) becomes

$$A(q)\ddot{q} + \eta(q, \dot{q}) + V'(q) = \beta(q)^T \lambda$$

Multiplying both sides of the equation by $\beta(q)A^{-1}(q)$ and using (3.3) yields

$$(\beta(q)A^{-1}(q)\beta(q)^T)\lambda = \beta(q)A^{-1}(q)\eta(q, \dot{q}) + \beta(q)A^{-1}(q)V'(q) - \gamma(q, \dot{q}).$$

The matrix $\beta(q)A^{-1}(q)\beta(q)^T$ is invertible since β has full rank. Therefore

$$\lambda(q, \dot{q}) = (\beta(q)A^{-1}(q)\beta(q)^T)^{-1}(\beta(q)A^{-1}(q)\eta(q, \dot{q}) + \beta(q)A^{-1}(q)V'(q) - \gamma(q, \dot{q})). \tag{3.5}$$

This choice of λ guarantees that the constraint functions

$$\phi^a(q, \dot{q}) = \beta^a_i(q)\dot{q}^i$$

are first integrals of (3.2). So, if (3.1) are satisfied at time $t = 0$, they are satisfied at all time.

Remark 3.1. The above formula for R depends on the basis β^a for the annihilator of $\mathcal{D}$. A different choice of basis, i.e. of matrix $\beta(q)$, say $\epsilon(q) = \psi(q)\beta(q)$, where $\psi(q)$ is an invertible $(n-r) \times (n-r)$ matrix, yields a different form of the reaction force, say $\hat{R}$. This choice of reaction force guarantees that instead the constraint functions $\tilde{\phi}^a(q, \dot{q}) = \psi^b_a(q)\beta^b_i(q)\dot{q}^i$ are first integrals of (3.2). However, the zero locus of ϕ^a and $\tilde{\phi}^a$ coincide, and equals D, and we have $R|_D = \hat{R}|_D$. Stated otherwise, the constraint force R is well-defined on the constraint space D but not on TQ. $\diamond$

Remark 3.2. Note that the dependence of η and γ on $\dot{q}$ is *quadratic*. As a consequence, the system is *reversible*. If $q(t)$ is a solution, then so is $\tilde{q}(t) := q(-t)$ (note also that the constraints (3.1) are satisfied by $\tilde{q}(t)$). This observation has implications in the dynamics and will be important in our study of measure preservation. $\diamond$

Remark 3.3. In practice one computes the value of the constraint forces in a case by case basis. However, this general approach illustrates an important point: while the constraints are of kinematical nature, the actual value of the constraint forces enforced by the Lagrange-D'Alembert principle involve the dynamical features of the problem, like its mass distribution and the potential. This follows from the dependence of λ on A and V. A concrete example is a ball rolling without slipping on the plane. As we shall see, if the ball is homogenous the constraint forces vanish. This is not the case if the mass distribution on the ball is not homogeneous. $\diamond$

3.2. **Example. The pendulum.** Our approach is somewhat inefficient and non-standard but it will illustrate details of systems with holonomic constraints further along the road.

We have $Q = \mathbb{R}^2$, the Lagrangian is $L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) - mgy$ and the constraint is $x^2 + y^2 = \ell^2$ where ℓ is the length of the rod of the pendulum. To apply Lagrange-D'Alembert's principle we differentiate the constraint to get

$$x\dot{x} + y\dot{y} = 0, \tag{3.6}$$

that has the form of the constraint equations (3.1). The matrix β is the row vector $\beta = (x, y)$ and λ is a scalar. The equations (3.2) can be written as

$$m\begin{pmatrix} \ddot{x} \\ \ddot{y} \end{pmatrix} = \begin{pmatrix} 0 \\ -mg \end{pmatrix} + \lambda \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ -mg \end{pmatrix} - \frac{T}{\sqrt{x^2+y^2}}\begin{pmatrix} x \\ y \end{pmatrix} \tag{3.7}$$

where $T = -\lambda\sqrt{x^2+y^2}$. We do not use equation (3.5) to determine λ. Instead, and just like we did in the general case, we differentiate the constraints and then use the equations of motion to determine T. Differentiating (3.6) yields

$$x\ddot{x} + y\ddot{y} = -\dot{x}^2 - \dot{y}^2.$$

Taking the scalar product of the equation of motion with (x, y) and using the above relation gives

$$-m(\dot{x}^2 + \dot{y}^2) = -mgy - T\sqrt{x^2+y^2}$$

so

$$T = \frac{m(\dot{x}^2 + \dot{y}^2) - mgy}{\sqrt{x^2+y^2}}.$$

For this value of T, the constraint function $\phi(x, y, \dot{x}, \dot{y}) = x\dot{x} + y\dot{y}$ is a first integral of (3.7). Along the zero level set of ϕ, $f(x, y) = x^2 + y^2$ is another first integral. We are interested in the flow of (3.7) restricted to the set where

$$f(x, y) = \ell^2, \qquad \phi(x, y, \dot{x}, \dot{y}) = 0.$$

Along this set

$$T = \frac{m(\dot{x}^2 + \dot{y}^2) - mgy}{\ell},$$

which is the tension on the rod of the pendulum that can be derived from a free-body diagram.

3.3. **Holonomic constraints.** If the distribution $\mathcal{D}$ is integrable then we can find local coordinates

$$q^1, \dots q^r, q^{r+1}, \dots, q^n,$$

such that the constraints take the form

$$\dot{q}^{r+1} = 0, \dots, \dot{q}^n = 0.$$

The equations of motion are

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} &= 0, \qquad i = 1, \dots r, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} &= \lambda_i, \qquad i = r+1, \dots n, \end{aligned} \tag{3.8}$$

where the multipliers λ_i guarantee that the constraints are satisfied.

Now fix an integral manifold P by the relations $q^{r+1} = c^{r+1}, \ldots, q^n = c^n$ and let $\tilde{L} = L|_{TP}$. Namely,

$$\tilde{L}(q^1, \ldots, q^r, \dot{q}^1, \ldots, \dot{q}^r) = L(q^1, \ldots, q^r, c^{r+1}, \ldots, c^n, \dot{q}^1, \ldots, \dot{q}^r, 0, \ldots, 0).$$

For $i = 1, \ldots, r$ we have

$$\left.\frac{\partial L}{\partial \dot{q}^i}\right|_{TP} = \frac{\partial \tilde{L}}{\partial \dot{q}^i}, \qquad \left.\frac{\partial L}{\partial q^i}\right|_{TP} = \frac{\partial \tilde{L}}{\partial q^i},$$

and therefore, the first set of equations in (3.8) becomes

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{q}^i}\right) - \frac{\partial \tilde{L}}{\partial q^i} = 0, \qquad i = 1, \ldots r.$$

In other words, when the constraints are holonomic, the equations of motion can be obtained by substituting them into the Lagrangian and computing the usual Euler-Lagrange equations.

Example 3.4. *For the case of the pendulum, the Lagrangian L in polar coordinates*

$$x = r\sin\theta, \qquad y = -r\cos\theta$$

is

$$L(\theta, r, \dot{\theta}, \dot{r}) = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) + mgr\cos\theta.$$

Taking P as the integral manifold $r = \ell$ leads to

$$\tilde{L}(\theta, \dot{\theta}) = \frac{m\ell^2}{2}\dot{\theta}^2 + mg\ell\cos\theta.$$

The corresponding Euler-Lagrange equations are

$$\ddot{\theta} + \frac{g}{\ell}\sin\theta = 0.$$

3.4. **Example 2. The Chaplygin sleigh.** This time $Q = S^1 \times \mathbb{R}^2$ with coordinates (θ, x, y). The Lagrangian is

$$L = \frac{1}{2}\Big((J + ma^2)\dot{\theta}^2 + m(\dot{x}^2 + \dot{y}^2) + 2ma\dot{\theta}(\cos\theta\dot{y} - \sin\theta\dot{x})\Big)$$

and the constraint is

$$-\dot{x}\sin\theta + \dot{y}\cos\theta = 0.$$

The equations of motion (3.2) become

$$\begin{aligned}
&(J + ma^2)\ddot{\theta} + \frac{d}{dt}(\cos\theta\dot{y} - \sin\theta\dot{x}) + ma\dot{\theta}(\sin\theta\dot{y} + \cos\theta\dot{x}) = 0, \\
&m\ddot{x} - \frac{d}{dt}(ma\dot{\theta}\sin\theta) = -\lambda\sin\theta, \\
&m\ddot{y} - \frac{d}{dt}(ma\dot{\theta}\sin\theta) = \lambda\cos\theta.
\end{aligned}$$

4. Rigid body dynamics

To describe the configuration of a rigid body we use two frames of reference. A *body frame* that is attached to the body and rotates with it and an inertial *space frame*. These frames are related by a matrix $B \in SO(3)$ called the *attitude matrix*. We will always assume that the body frame has its origin at the center of mass.

Take a vector whose coordinates with respect to the space frame are the entries of $q \in \mathbb{R}^3$ and whose coordinates with respect to the body frame are the entries of $Q \in \mathbb{R}^3$. The matrix B is such that $q = BQ$. For the most part, we will follow the convention of writing vectors written in the space frame with lower case letters and vectors written in the body frame with upper case letters.

If one performs a rotation of the space frame by $A \in SO(3)$, the new coordinates of our vector become $\tilde{q} = Aq$. And the relationship with Q becomes $\tilde{q} = ABQ$.

On the other hand, if we perform a rotation of the body frame by $A \in SO(3)$, the body coordinates of our vector are $\tilde{Q} = AQ$. The relationship with q becomes $q = BA^{-1}\tilde{Q}$.

Therefore, *rotations of the space axes correspond to left multiplication of the attitude matrix B* while *rotations of the body axes correspond to right multiplication of the attitude matrix B*. This simple fact is essential to understand the symmetries of a system that involves rigid bodies. For instance, the free rigid body is *left* invariant due to the isotropy of the ambient space.

A motion of the rigid body corresponds to a curve $B(t) \in SO(3)$.

Consider a vector Q with constant body coordinates (e.g. a material point in the body). Its space coordinates during the motion satisfy $q(t) = B(t)Q$ and hence

$$\dot{q} = \dot{B}Q = \dot{B}B^{-1}q = \hat{\omega}q = \omega \times q.$$

Here $\hat{\omega}$ is the skew-symmetric matrix (element of the Lie algebra $\mathfrak{so}(3)$) whose entries define the components $\omega_1, \omega_2, \omega_3$ of the vector ω. The convention that makes the formulae work is

$$\dot{B}B^{-1} = \hat{\omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

Analogously, if q is a vector with constant space coordinates (e.g. one of the axes of the space frame), its body coordinates throughout the motion are given by $Q(t) = B(t)^{-1}q$ and hence

$$\dot{Q} = -B^{-1}\dot{B}B^{-1}q = -B^{-1}\dot{B}Q = -\hat{\Omega}Q = Q \times \Omega.$$

where $\hat{\Omega}$ is the skew-symmetric matrix (element of the Lie algebra $\mathfrak{so}(3)$) whose entries define the components $\Omega_1, \Omega_2, \Omega_3$ of the vector Ω. The convention that makes the formulae work is

$$B^{-1}\dot{B} = \hat{\Omega} = \begin{pmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{pmatrix}.$$

The entries of ω and Ω are respectively the space and body coordinates of the *angular velocity vector* of the body. It is important to remember that they represent the same vector, just written with respect to different set of axes. They satisfy $\omega = B\Omega$.

The angular momentum vector depends linearly on the angular velocity. On the body frame, such dependence is $M = \mathbb{I}\Omega$ where $\mathbb{I}$ is the *inertia tensor of the body*. It is a symmetric, positive definite, constant (independent of B), 3×3 matrix. The kinetic energy of the body is the scalar product $\langle M, \Omega \rangle$.

The space representation of the angular momentum is the vector $m = BM$. In the absence of external forces it is constant. Therefore, from our calculation above, M satisfies

$$\dot{M} = M \times \Omega.$$

Equivalently,

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega.$$

These are Euler's equations for the motion of a free rigid body. They are complemented with the reconstruction equation

$$\dot{B} = B\hat{\Omega}.$$

The decoupling of the equations is a consequence of the symmetries of the problem mentioned above. The system is left invariant because the equations do not care in which way we choose the orientation of the space frame.

4.1. **Suslov problem.** Consider the motion of a rigid body with the nonholonomic constraint $\langle a, \Omega \rangle = 0$.

In the absence of constraints, the system evolves according to Euler's equations

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega.$$

If we enforce the constraint, we should bring in a reaction force[1] R

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega + R.$$

Lagrange-D'Alembert principle tells us that $\langle R, \Omega \rangle = 0$ if Ω satisfies the constraints. Therefore we take $R = \lambda a$ and determine λ. Differentiate the constraint to obtain $\langle a, \dot{\Omega} \rangle = 0$ and write

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega + \lambda a. \tag{4.1}$$

Taking the scalar product on both sides with $\mathbb{I}^{-1}a$ and performing the algebra we get

$$\lambda = -\frac{\langle (\mathbb{I}\Omega) \times \Omega, \mathbb{I}^{-1}a \rangle}{\langle a, \mathbb{I}^{-1}a \rangle} \tag{4.2}$$

This choice of λ guarantees that $\langle a, \Omega \rangle$ is a first integral of (4.1). We are interested in the zero level set.

Choose the body frame of the body such that the third axis E_3 is parallel to the vector a. The constraint becomes $\Omega_3 = 0$. The inertia tensor cannot be assumed to be diagonal. But by an appropriate rotation about the E_3 axis, we can assume that it has the form

$$\mathbb{I} = \begin{pmatrix} I_{11} & 0 & I_{13} \\ 0 & I_{22} & I_{23} \\ I_{13} & I_{23} & I_{33} \end{pmatrix}.$$

Exercise: Convince yourself of this statement.

[1]It would be more appropriate to speak of a constraint *torque*

A direct calculation shows that substituting $a = E_3$ and the above form of $\mathbb{I}$ and (4.2) into (4.1) yields

$$\begin{aligned} \dot{\Omega}_1 &= -\frac{1}{I_{11}}\left((I_{13}\Omega_1 + I_{23}\Omega_2)\Omega_2\right), \\ \dot{\Omega}_2 &= \frac{1}{I_{22}}\left((I_{13}\Omega_1 + I_{23}\Omega_2)\Omega_1\right), \end{aligned} \tag{4.3}$$

together with $\dot{\Omega}_3 = 0$.

Note that we get a closed system on $\mathbb{R}^2$. If I_{13} and I_{23} do not vanish simultaneously (the vector a is not a principal axis of inertia of the body), the phase portrait on the plane (Ω_1, Ω_2) looks like

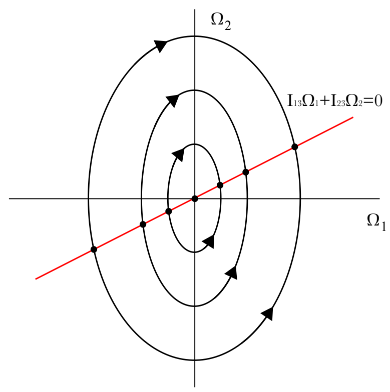


FIGURE 1. Phase portrait of the system (4.3).

The trajectories are contained on the energy level surfaces. Putting $\Omega_3 = 0$ in $\langle \mathbb{I}\Omega, \Omega\rangle$ gives

$$E = \frac{1}{2}\left(I_{11}\Omega_1^2 + I_{22}\Omega_2^2\right).$$

The level sets are hence ellipses. On the other hand the line $I_{13}\Omega_1 + I_{23}\Omega_2 = 0$ consists of equilibrium points.

The full dynamics of the system are complemented with the equation $\dot{B} = B\hat{\Omega}$. In the 5 dimensional phase space D we have families of attracting and repelling periodic orbits. This type of behavior cannot occur in Hamiltonian systems.

4.2. **Veselova problem.** The constraint can be written in body coordinates as $\langle \Omega, \gamma\rangle = 0$ where $\gamma = B^{-1}e_3$ is the *Poisson vector*. It is the expression of the 3rd axis of the body frame written in body coordinates. It evolves according to the kinematic equation

$$\dot{\gamma} = \gamma \times \Omega. \tag{4.4}$$

Note that $||\gamma||^2$ is a first integral of the above equation. We are interested in the level set $||\gamma|| = 1$ and we think of $\gamma \in S^2$.

In the absence of constraints, the system evolves according to Euler's equations

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega.$$

If we enforce the constraint, we should bring in a reaction force[2] R

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega + R.$$

Lagrange-D'Alembert principle tells us that $\langle R, \Omega \rangle = 0$ if Ω satisfies the constraints. Therefore we take $R = \lambda\gamma$ and determine λ. Differentiate the constraint to obtain $\langle \gamma, \dot{\Omega} \rangle = 0$ and write

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega + \lambda\gamma. \tag{4.5}$$

Taking the scalar product on both sides with $\mathbb{I}^{-1}\gamma$ and performing the algebra we get

$$\lambda = -\frac{\langle (\mathbb{I}\Omega) \times \Omega, \mathbb{I}^{-1}\gamma \rangle}{\langle \gamma, \mathbb{I}^{-1}\gamma \rangle} \tag{4.6}$$

This choice of λ guarantees that $\langle \gamma, \Omega \rangle$ is a first integral of the system (4.5), (4.4). We are interested in the level set where

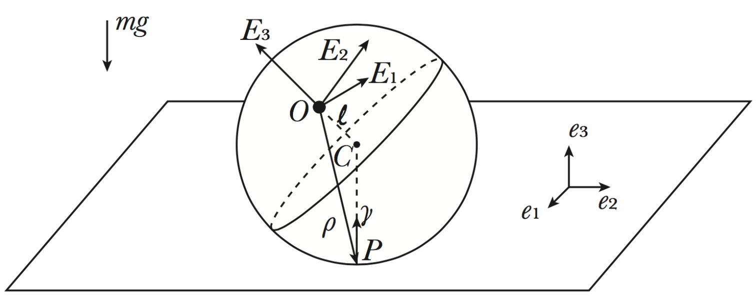
$$||\gamma|| = 1, \qquad \langle \Omega, \gamma \rangle = 0.$$

Such space is isomorphic to the tangent bundle of the sphere TS^2.

4.3. **Chaplygin top.** We denote by

- C the geometric center of the sphere.
- O the center of mass.
- P the contact point.
- $u = (x, y, z)$ the space coordinates of the point O. The space frame has the third vector e_3 perpendicular to the rolling plane and pointing upwards (against gravity).

The origin of the body axis is the center of mass O.



In the absence of constraints the equations of motion are

$$m\ddot{u} = -mge_3, \qquad \mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega,$$

where $\mathbb{I}$ is the inertia tensor, m is the mass of the sphere, g is the gravitational constant.

Now add the rolling without slipping reaction forces:

$$m\ddot{u} = -mge_3 + R_1, \qquad \mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega + R_2. \tag{4.7}$$

We now find the expression for the rolling constraint. A material point on the sphere with constant body coordinates Q has space coordinates

$$q(t) = u(t) + B(t)Q,$$

[2] It would be more appropriate to speak of a constraint *torque*

so

$$\dot{q} = \dot{u} + \dot{B}Q.$$

To obtain the rolling constraint we put $Q = \rho$ (the vector $\vec{OP}$ written in body coordinates) and enforce $\dot{q} = 0$ (the velocity of the contact point is zero). We get

$$\dot{u} = -\dot{B}\rho = -BB^{-1}\dot{B}\rho = -B(\hat{\Omega}\rho) = B(\rho \times \Omega).$$

So the nonholonomic constraint is

$$\dot{u} = B(\rho \times \Omega). \tag{4.8}$$

This constraint has one holonomic component: the ball cannot leave the table! Introduce the Poisson vector $\gamma = B^{-1}e_3$ and assume that the body frame $\{E_1, E_2, E_3\}$ is selected in such way that the geometric center of the sphere has body coordinates $-\ell E_3$. We can then write

$$\rho = -R\gamma - \ell E_3 \tag{4.9}$$

where R is the radius of the sphere.

The z coordinate of u satisfies

$$z = -\langle \rho, \gamma \rangle = R + \ell\gamma_3.$$

This is a holonomic constraint. No velocities are involved. Only a relationship between the coordinate z of u and one entries of B (actually, just the $(3,3)$-entry). Differentiating the above equation we get

$$\dot{z} = \ell\dot{\gamma}_3.$$

Exercise. Show that this equation is equivalent with the third component of (4.8).

Now we use the Lagrange-D'Alembert principle to obtain expressions for R_1, R_2 in (4.7). The principle implies that if $(\dot{u}, \Omega)$ satisfy (4.8), then $\langle R_1, \dot{u}\rangle + \langle R_2, \Omega\rangle = 0$. For such $(\dot{u}, \Omega)$ we have

$$\begin{aligned} 0 &= \langle R_1, B(\rho \times \Omega)\rangle + \langle R_2, \Omega\rangle \\ &= \langle (B^{-1}R_1) \times \rho + R_2, \Omega\rangle. \end{aligned}$$

Since Ω can be taken arbitrary we conclude that

$$R_2 = \rho \times (B^{-1}R_1).$$

Physically: R_2 is the moment of the force R_1 about the contact point.

Differentiating the constraints we find that

$$R_1 = m\ddot{u} + mge_3 = m\left(\dot{B}(\rho \times \Omega) + B(\dot{\rho} \times \Omega) + B(\rho \times \dot{\Omega}) + ge_3\right).$$

Therefore,

$$\begin{aligned} \mathbb{I}\dot{\Omega} &= (\mathbb{I}\Omega) \times \Omega + R_2 = (\mathbb{I}\Omega) \times \Omega + \rho \times (B^{-1}R_1) \\ &= (\mathbb{I}\Omega) \times \Omega + m\rho \times (\Omega \times (\rho \times \Omega)) + m\rho \times (\dot{\rho} \times \Omega) + m\rho \times (\rho \times \dot{\Omega}) + mg\rho \times \gamma. \end{aligned}$$

Complementing this equation with the kinematical condition

$$\dot{\gamma} = \gamma \times \Omega$$

and the relation (4.9), yields a closed system of equations for the evolution of γ, Ω. As with the Veselova system, the system possesses the geometric integral $||\gamma|| = 1$. We interpret $\gamma \in S^2$ so we obtain a system of equations on $S^2 \times \mathbb{R}^3$.

Introducing the angular momentum about the contact point

$$K = \mathbb{I}\Omega + m\rho \times (\Omega \times \rho),$$

the equations become

$$\dot{K} = K \times \Omega + m\dot{\rho} \times (\Omega \times \rho) + mg\rho \times \gamma, \qquad \dot{\gamma} = \gamma \times \Omega,$$

and the energy (along D) may be written as

$$E = \frac{1}{2}\langle K, \Omega\rangle + mg\ell\gamma_3.$$

Exercise Use the vector identity

$$a \times (b \times c) = \langle a, c\rangle b - \langle a, b\rangle c$$

to verify this statement.

5. Quasi-velocities

It is not clear how the equations of motion for the Suslov problem, the Veselov problem and the Chaplygin top, derived in section 4 relate to the general form of the equations of motion derived in section 3. In order to explain this relationship we introduce the concept of quasi-velocities.

The most common way to represent a velocity vector $V \in T_qQ$ is to write

$$V = (\dot{q}^1, \dots, \dot{q}^n)$$

where $(q^1, \dots, q^n)$ are local coordinates on the configuration manifold Q. What we are really thinking is that

$$V = \dot{q}^i \frac{\partial}{\partial q^i}$$

where $\{\frac{\partial}{\partial q^i}\}$ is the basis of T_qQ induced by the coordinates $(q^1, \dots, q^n)$. This gives a way to construct coordinates $(q^1, \dots q^n, \dot{q}^1, \dots, \dot{q}^n)$ for the tangent bundle TQ. This choice of coordinates is used in Lagrange's equations and leads to the canonical coordinates $(q^1, \dots q^n, p_1, \dots, p_n)$ where $p_i = \frac{\partial L}{\partial \dot{q}^i}$ that appear in the Hamiltonian formulation.

This is not the only way to construct coordinates for TQ and other choices are often more convenient. The geometric idea of quasi-velocities is to (locally) consider n linearly independent vector fields $X_1, \dots X_n$ and use $\{X_k(q)\}$ as a basis for the tangent space T_qQ. We can then write any velocity vector $V \in T_qQ$ as a linear combination

$$V = v^k X_k(q).$$

We can then use $(q^1, \dots q^n, v^1, \dots, v^n)$ as coordinates for the tangent bundle TQ. The collection of vector fields $\{X_k(q)\}$ are said to form a *moving frame*. In practice they are chosen in accordance with the symmetries or other important characteristics of the problem at hand.

Of course that velocities are a particular case of quasi-velocities when the moving frame is $\{\frac{\partial}{\partial q^k}\}$. However, the converse statement is not true. Given a moving frame $\{X_k(q)\}$ there need not exist coordinates $(\tilde{q}^1, \dots \tilde{q}^n)$ such that $X_k = \frac{\partial}{\partial \tilde{q}^k}$ for all k. This is easily seen since the vector fields X_k may not commute! This is crucial for the sequel so we introduce the (local) functions $C_{ij}^k : Q \to \mathbb{R}$ by the relations

$$[X_i, X_j] = C_{ij}^k X_k.$$

We shall refer to them as the structure coefficients associated to the frame $\{X_k(q)\}$.

Our first task is to rewrite the Euler-Lagrange equations (3.2) in quasi-velocities. The resulting equations receive the name of *Hamel's* equations.

Before we compute Hamel's equations note that the moving frame $\{X_k(q)\}$ induces a *moving co-frame* $\{\mu^k(q)\}$ of T^*Q. It is just the dual basis. Namely,

$$\langle \mu^i(q), X_j(q)\rangle = \delta^i_j,$$

for all $q \in Q$ where the frames are defined. If the frame is the standard $\{\frac{\partial}{\partial q^k}\}$ then the corresponding co-frame is $\{dq^k\}$ that we are used to. In order to obtain Hamel's equations we need to relate both the moving frame and the moving co-frame to the standard frames through appropriate linear combinations. Suppose that we have

$$X_j = \rho^k_j \frac{\partial}{\partial q^k}, \qquad \mu^i = \sigma^i_k \, dq^k \tag{5.1}$$

for locally defined functions $\rho^i_j, \sigma^j_i : Q \to \mathbb{R}$. For consistency we need $\rho^k_j \sigma^i_k = \delta^i_j$. Namely, the $n \times n$ matrices ρ^i_j and σ^j_i are inverses. So we also have

$$\frac{\partial}{\partial q^k} = \sigma^i_k X_i, \qquad dq^k = \rho^k_j \mu^j. \tag{5.2}$$

In terms of the induced coordinates for the tangent spaces we have

$$\dot q^j = \rho^j_k v^k, \tag{5.3}$$

or in matrix notation, $\dot q = \rho v$ where the matrix ρ is $n \times n$ and has entries ρ^j_k (in the k^{th} row and j^{th} column).

5.1. **Hamel's equations.** Now, to derive Hamel's equations, we write the Lagrangian $L : TQ \to \mathbb{R}$ in terms of the quasi-velocities. In order to apply the chain rule carefully we denote

$$\tilde L(q, v) = L(q, \dot q),$$

where we have abbreviated $v = (v^1, \dots v^n)$ for the quasi-velocities. Note that

$$\tilde L(q, v) = \frac{1}{2} v^T \rho(q)^T A(q) \rho(q) v - U(q).$$

Using the chain rule we find

$$\frac{\partial \tilde L}{\partial v^i} = \frac{\partial L}{\partial \dot q^k} \rho^k_i, \qquad \frac{\partial \tilde L}{\partial q^k} = \frac{\partial L}{\partial q^k} + \frac{\partial L}{\partial \dot q^j}\frac{\partial \rho^j_l}{\partial q^k} v^l, \tag{5.4}$$

where we have used (5.3). Using the first of these expressions, the Euler-Lagrange equations (3.2), and (5.3), we find

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial \tilde L}{\partial v^i}\right) &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot q^k}\right)\rho^k_i + \frac{\partial L}{\partial \dot q^k}\frac{\partial \rho^k_i}{\partial q^j}\dot q^j \\ &= \rho^k_i \frac{\partial L}{\partial q^k} + \frac{\partial L}{\partial \dot q^j}\frac{\partial \rho^j_i}{\partial q^k}\rho^k_l v^l.\end{aligned}$$

where the indices of summation j and k have been swapped in the second identity. Using now the second relation in (5.4) the above equation becomes

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial v^i}\right) = \rho_i^k \frac{\partial \tilde{L}}{\partial q^k} + \frac{\partial L}{\partial \dot{q}^j}\left(\frac{\partial \rho_i^j}{\partial q^k}\rho_l^k - \frac{\partial \rho_l^j}{\partial q^k}\rho_i^k\right) v^l. \tag{5.5}$$

On the other hand notice that

$$[X_i, X_l] = -\left(\frac{\partial \rho_i^j}{\partial q^k}\rho_l^k - \frac{\partial \rho_l^j}{\partial q^k}\rho_i^k\right)\frac{\partial}{\partial q^j} = -\sigma_j^m\left(\frac{\partial \rho_i^j}{\partial q^k}\rho_l^k - \frac{\partial \rho_l^j}{\partial q^k}\rho_i^k\right) X_m.$$

Therefore,

$$C_{il}^m = -\sigma_j^m\left(\frac{\partial \rho_i^j}{\partial q^k}\rho_l^k - \frac{\partial \rho_l^j}{\partial q^k}\rho_i^k\right),$$

or equivalently

$$\left(\frac{\partial \rho_i^j}{\partial q^k}\rho_l^k - \frac{\partial \rho_l^j}{\partial q^k}\rho_i^k\right) = -\rho_m^j C_{il}^m. \tag{5.6}$$

Substituting this expression into (5.7) and using the second expression in (5.4) we finally get

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial v^i}\right) = \rho_i^k \frac{\partial \tilde{L}}{\partial q^k} - C_{il}^m v^l \frac{\partial \tilde{L}}{\partial v^m}. \tag{5.7}$$

These are Hamel's equations that clearly hold for $i = 1, \dots, n$. When complemented with (5.3) they give a full system for the determination of the motion of the mechanical system.

Example. Euler top. The definition of the angular velocity in body coordinates implies

$$\dot{B} = B\hat{\Omega}.$$

Therefore, the time derivatives of any set of coordinates on $SO(3)$ are linear combinations of the entries of Ω with configuration dependent coefficients. In other words, the entries of Ω are quasi-velocities. We have

$$\tilde{L}(B, \Omega) = \frac{1}{2}\langle \mathbb{I}\Omega, \Omega\rangle.$$

The corresponding moving frame are the left invariant vector fields on $SO(3)$ that are obtained by left translation of the canonical basis of the Lie algebra $\mathfrak{so}(3)$. The structure coefficients C_{ij}^k are constant and equal to the structure constants of the Lie algebra $\mathfrak{so}(3)$. Hamel's equations are Euler's equations

$$\mathbb{I}\dot{\Omega} = (\mathbb{I}\Omega) \times \Omega.$$

5.2. **The equations of motion for a nonholnomic system in quasi-velocities.** The Hamel formalism will allow us to write the restriction of the Lagrange-D'Alembert equations (3.2) to the constraint subbundle $D \subset TQ$ defined by the constraint distribution $\mathcal{D}$. Moreover, the resulting equations will not involve multipliers.

We select the moving frame in such a way that the first r vector fields $X_1, \dots, X_r$ form a (local) basis of the constraint distribution $\mathcal{D}$ and the remaining $Y_1, \dots, Y_{n-r}$ span a complement of $\mathcal{D}_q$ at each point in Q. We denote our moving co-frame by $\{X_\alpha, Y_a\}$ where the index α runs from 1 to r

and the latin index a runs from 1 to $n-r$. The corresponding quasi-velocities are denoted v^α, u^a and the dual basis is $\{\mu^\alpha, \nu^a\}$.

Before writing the Lagrange-D'Alembert equations in quasi-velocities we write

$$\begin{aligned} X_\alpha &= \rho^k_\alpha \frac{\partial}{\partial q^k}, \qquad \mu^\alpha = \sigma^\alpha_k \, dq^k, \\ Y_a &= \rho^k_a \frac{\partial}{\partial q^k} \qquad \nu^a = \sigma^a_k \, dq^k, \end{aligned} \tag{5.8}$$

that are analogous to (5.1). And imply

$$\dot{q}^k = \rho^k_\alpha v^\alpha + \rho^k_a u^a.$$

Similar to (5.2) we have

$$\frac{\partial}{\partial q^k} = \sigma^\alpha_k X_\alpha + \sigma^a_k Y_a.$$

Since X_α is tangent to the constraints we have

$$\beta^a_k \rho^k_\alpha = 0.$$

The nonholonomic constraints can be written as $u^a = 0$ $a = 1, \dots, n-r$. Note that (q^i, v^α) serve as coordinates for the constraint subbundle $D \subset TQ$ (as a manifold, D has dimension $n+r$).

Proceeding as in the previous section we denote $\tilde{L}(q^i, v^\alpha, u^a) = L(q^i, \dot{q}^i)$. Using the chain rule, we find, in analogy with (5.4)

$$\frac{\partial \tilde{L}}{\partial v^\alpha} = \frac{\partial L}{\partial \dot{q}^k}\rho^k_\alpha, \qquad \frac{\partial \tilde{L}}{\partial u^a} = \frac{\partial L}{\partial \dot{q}^k}\rho^k_a, \qquad \frac{\partial \tilde{L}}{\partial q^k} = \frac{\partial L}{\partial q^k} + \frac{\partial L}{\partial \dot{q}^j}\left(\frac{\partial \rho^j_\beta}{\partial q^k} v^\beta + \frac{\partial \rho^j_a}{\partial q^k} u^a \right). \tag{5.9}$$

Using the Lagrange-D'Alembert equations (3.2) and the fact that $R_i \rho^i_\alpha = 0$, similar to (5.10) we obtain, for $\alpha = 1, \dots, r$,

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial v^\alpha} \right) = \rho^k_\alpha \frac{\partial \tilde{L}}{\partial q^k} + \frac{\partial L}{\partial \dot{q}^j}\left(\frac{\partial \rho^j_\alpha}{\partial q^k}\rho^k_\beta - \frac{\partial \rho^j_\beta}{\partial q^k}\rho^k_\alpha \right) v^\beta + \frac{\partial L}{\partial \dot{q}^j}\left(\frac{\partial \rho^j_\alpha}{\partial q^k}\rho^k_a - \frac{\partial \rho^j_a}{\partial q^k}\rho^k_\alpha \right) u^a. \tag{5.10}$$

Now, similar to (5.6) we have

$$\left(\frac{\partial \rho^j_\alpha}{\partial q^k}\rho^k_\beta - \frac{\partial \rho^j_\beta}{\partial q^k}\rho^k_\alpha \right) = -\rho^j_\gamma C^\gamma_{\alpha\beta} - \rho^j_a C^a_{\alpha\beta}.$$

where

$$[X_\alpha, X_\beta] = C^\gamma_{\alpha\beta} X_\gamma + C^a_{\alpha\beta} Y_a.$$

Substitution into (5.10) yields

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial v^\alpha} \right) = \rho^k_\alpha \frac{\partial \tilde{L}}{\partial q^k} - C^\gamma_{\alpha\beta} v^\beta \frac{\partial \tilde{L}}{\partial v^\gamma} - C^a_{\alpha\beta} v^\beta \frac{\partial \tilde{L}}{\partial u^a} + \frac{\partial L}{\partial \dot{q}^j}\left(\frac{\partial \rho^j_\alpha}{\partial q^k}\rho^k_a - \frac{\partial \rho^j_a}{\partial q^k}\rho^k_\alpha \right) u^a. \tag{5.11}$$

Now we restrict these equations to D. This means substituting $u^a = 0$ (and hence $\dot{u}^a = 0$ everywhere). This gives

$$\frac{d}{dt}\left(\left.\frac{\partial \tilde{L}}{\partial v^\alpha}\right|_{u^a=0}\right) = \rho^k_\alpha \left.\frac{\partial \tilde{L}}{\partial q^k}\right|_{u^a=0} - C^\gamma_{\alpha\beta} v^\beta \left.\frac{\partial \tilde{L}}{\partial v^\gamma}\right|_{u^a=0} - C^a_{\alpha\beta} v^\beta \left.\frac{\partial \tilde{L}}{\partial u^a}\right|_{u^a=0}. \tag{5.12}$$

If we introduce the *constrained Lagrangian* L_c as the restriction of L to D, i.e. $L_c = L|_D$, then we can write $L_c(q^i, v^\alpha) = \tilde{L}(q^i, v^\alpha, 0)$ and we have the identities

$$\left.\frac{\partial \tilde{L}}{\partial v^\alpha}\right|_{u^a=0} = \frac{\partial L_c}{\partial v^\alpha}, \qquad \left.\frac{\partial \tilde{L}}{\partial q^k}\right|_{u^a=0} = \frac{\partial L_c}{\partial q^k},$$

and (5.12) becomes

$$\frac{d}{dt}\left(\frac{\partial L_c}{\partial v^\alpha}\right) = \rho^k_\alpha \frac{\partial L_c}{\partial q^k} - C^\gamma_{\alpha\beta} v^\beta \frac{\partial L_c}{\partial v^\gamma} - C^a_{\alpha\beta} v^\beta \left.\frac{\partial \tilde{L}}{\partial u^a}\right|_{u^a=0}. \tag{5.13}$$

A final simplification is possible if we choose the vector fields Y_a in such a way that they span $\mathcal{D}^\perp$, which denotes the orthogonal complement of $\mathcal{D}$ with respect to the kinetic energy metric. In this case the kinetic energy matrix block diagonalizes

$$\rho(q)^T A(q) \rho(q) = \begin{pmatrix} A_D(q) & 0 \\ 0 & A_{D^\perp}(q) \end{pmatrix}.$$

Then, $\frac{\partial \tilde{L}}{\partial u^a}(q^i, v^\alpha, 0) = 0$ and equations (5.13) simplify to

$$\frac{d}{dt}\left(\frac{\partial L_c}{\partial v^\alpha}\right) = \rho^k_\alpha \frac{\partial L_c}{\partial q^k} - C^\gamma_{\alpha\beta} v^\beta \frac{\partial L_c}{\partial v^\gamma}, \tag{5.14}$$

and the functions $C^\gamma_{\alpha\beta}$ that appear in the equations can be characterised by the relation

$$\mathcal{P}([X_\alpha, X_\beta]) = C^\gamma_{\alpha\beta} X_\gamma$$

where $\mathcal{P} : TQ \to D$ is the bundle projector corresponding to the orthogonal decomposition $TQ = D \oplus D^\perp$.

Equations (5.14) are complemented by

$$\dot{q}^k = \rho^k_\alpha v^\alpha, \tag{5.15}$$

and define a system of first order differential equations for the variables (q^i, v^α). To see this note that

$$L_c(q^i, v^\alpha) = \frac{1}{2} v^T A_D(q) v - U(q).$$

Here $A_D(q)$ is an $r \times r$ positive definite matrix that defines a fibered metric on D. It is given by

$$A_D(q) = \rho_D(q)^T A(q) \rho_D(q)$$

where $\rho_D(q)$ is the $n \times r$ matrix with entries ρ^k_α (in the k^{th} row and α^{th} column). So $\frac{\partial L_c}{\partial v^\alpha}$ is linear in v^β.

Equations (5.14) and (5.15) define a vector field X_{nh} on D that describes the motion of the system for arbitrary admissible initial conditions. The vector field X_{nh} is not just an arbitrary vector field on D. It has the additional structure of being *second order* with respect to the bundle structure of D. Namely, if $\tau : D \to Q$ is the projection, then

$$T\tau \circ X_{nh} = id_D.$$

Another way to say this is that if $c(t)$ is an integral curve of X_{nh} then $T\tau(\dot{c}(t)) \in \mathcal{D}_{\tau(c(t))}$. Alternatively, there is a rank r distribution $\mathcal{E}$ on D

$$\mathcal{E}(q,v) = \operatorname{span}\left\{ X_\alpha(q) + 0 \cdot \frac{\partial}{\partial v^\alpha}, \frac{\partial}{\partial v^\alpha} \right\}$$

with the property that $X_{nh}(q,v) \in \mathcal{E}_{(q,v)}$ for all $(q,v) \in D$. Note that $\mathcal{E}$ has rank $2r$ and can be defined as the annihilator of the basic one-forms $\tau^*\beta^a$ on D. Also note that $\mathcal{E}$ is integrable if and only if $\mathcal{D}$ is integrable, i.e. if and only if the constraints are holonomic.

The conservation of energy can now be made transparent. Let $E_c = E|_D$. Then

$$E_c(q,v) = v^\alpha \frac{\partial L_c}{\partial v^\alpha}(q,v) - L_c(q,v).$$

A direct calculation that uses the skew-symmetry of $C_{\alpha\beta}^\gamma$ with respect to its lower indices shows that E_c is a first integral of X_{nh}.

A more thorough discussion of the geometric aspects of this approach can be found in e.g. [12].

5.3. **Example: Chaplygin sleigh.** We take

$$X_1 = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y}, \qquad X_2 = \frac{\partial}{\partial \theta}$$

as the generators of $\mathcal{D}$. The corresponding quasi-velocities u, ω satisfy

$$\dot{x} = u\cos\theta, \qquad \dot{y} = u\sin\theta, \qquad \dot{\theta} = \omega. \tag{5.16}$$

Physically, u is the velocity of the body in the direction of the blade and ω is the angular velocity.

Recall that the Lagrangian is

$$L(x,y,\theta,\dot{x},\dot{y},\dot{\theta}) = \frac{1}{2}\left((J+ma^2)\dot{\theta}^2 + m(\dot{x}^2+\dot{y}^2) + 2ma\dot{\theta}(\cos\theta\dot{y} - \sin\theta\dot{x}) \right).$$

A straightforward calculation shows that $\mathcal{D}^\perp = \operatorname{span}\{Y\}$ where

$$Y = \frac{ma}{J+ma^2}\frac{\partial}{\partial\theta} + \sin\theta\frac{\partial}{\partial x} - \cos\theta\frac{\partial}{\partial y}.$$

We have

$$\begin{aligned}
[X_1, X_2] &= \sin\theta\frac{\partial}{\partial x} - \cos\theta\frac{\partial}{\partial y} \\
&= Y - \frac{ma}{J+ma^2}X_2.
\end{aligned}$$

It follows that

$$\mathcal{P}([X_1,X_2]) = -\frac{ma}{J+ma^2}X_2,$$

and hence

$$C_{12}^1 = 0, \qquad C_{12}^2 = -\frac{ma}{J+ma^2}.$$

The constrained Lagrangian is

$$L_c = \frac{1}{2}\left((J + ma^2)\omega^2 + mu^2\right).$$

The equations of motion are

$$m\dot{u} = ma\omega^2$$
$$(J + ma^2)\dot{\omega} = -mau\omega,$$

and (5.16). Note that the above equations define a decoupled closed system on $\mathbb{R}^2$. The rank 4 distribution $\mathcal{E}$ is given by

$$\mathcal{E}_{(x,y,\theta,u,\omega)} = \operatorname{span}\left\{\cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial u}, \frac{\partial}{\partial \omega}\right\}$$

6. Almost Hamiltonian Formulation

Introduce the (quasi)-momenta

$$p_\alpha = \frac{\partial L_c}{\partial v^\alpha}(q, v), \qquad \alpha = 1, \ldots, k.$$

This is a q-dependent invertible linear change of variables between v and p. In matrix form, $p = A_D(q)v$, $v = A_D(q)^{-1}p$.

We can give a geometric interpretation of p_α analogous to the usual Hamiltonian formulation of mechanics. The mapping

$$(q, v) \xrightarrow{\psi} (q, p) = (q, A_D(q)v)$$

is a bundle isomorphism between D and D^* defined by the bundle metric on D. $p = (p_1, \ldots, p_r)$ are linear coordinates on the fibres of D^*. If we denote (q, v) and (q, p) respectively by v_q and p_q, and $p_q = \psi(v_q)$, then we have

$$p_q(\cdot) = \langle\langle v_q, \cdot\rangle\rangle_q,$$

where $\langle\langle\cdot,\cdot\rangle\rangle$ denotes the kinetic energy metric. In particular, if we recall that $v = (v^1, \ldots, v^r)$ are coordinates with respect to the basis $\{X_1, \ldots, X_r\}$, we can interpret the coordinate p_α as the linear function on the fibres of D given by

$$(p_\alpha)_q = \langle\langle X_\alpha(q), \cdot\rangle\rangle_q.$$

The energy function written in terms of (q, p) is the (constrained) Hamiltonian. It is given by

$$H_c(q, p) = p_\alpha v^\alpha - L_c = \frac{1}{2}p^T(A_D(q))^{-1}p + U(q),$$

and satisfies

$$\frac{\partial H_c}{\partial q^i} = -\frac{\partial L_c}{\partial q^i} \qquad \frac{\partial H_c}{\partial p_\alpha} = v^\alpha,$$

where it is understood that one should substitute $v = A_D(q)^{-1}p$ on the right hand side of the above equations.

The equations of motion can be written as the following first order system in D^*:

$$\dot{q}^k = \rho^k_\alpha\frac{\partial H_c}{\partial p_\alpha}, \qquad \dot{p}_\alpha = -\rho^k_\alpha\frac{\partial H_c}{\partial q^k} - C^\gamma_{\alpha\beta}p_\gamma\frac{\partial H}{\partial p_\beta}.$$

They define a vector field on D^* which is the push-forward of X_{nh} by ψ that we continue to denote by X_{nh}. Note that if $c(t)$ is a trajectory of the vector field X_{nh} and $\tau : D^* \to Q$ is the bundle projection, then

$$\tau(c(t)) \in \mathcal{D}_{\tau(c(t))}.$$

Another way to say the same thing is that X_{nh} is tangent to the distribution $\mathcal{F} = \psi_*(\mathcal{E})$ defined as the annihilator of $\tau^*\beta^a$.

The equations of motion can be written in vector form as

$$\begin{pmatrix} \dot q \\ \dot p \end{pmatrix} = \begin{pmatrix} 0 & \rho_D(q) \\ -\rho_D^T(q) & \mathcal{C}(q,p) \end{pmatrix} \begin{pmatrix} \frac{\partial H_c}{\partial q} \\ \frac{\partial H_c}{\partial p} \end{pmatrix}.$$

The size of the matrices above are

$$\begin{array}{rcccrcc} 0 & \text{is} & n \times n, & \qquad & \rho_D(q) & \text{is} & n \times r, \\ -\rho_D(q)^T & \text{is} & r \times n, & \qquad & \mathcal{C}(q,p) & \text{is} & r \times r. \end{array}$$

The matrix $\mathcal{C}(q,p)$ has components

$$\mathcal{C}(q,p)_{\alpha\beta} = -C^{\gamma}_{\alpha\beta}(q) p_\gamma.$$

The matrix

$$\pi_{nh}(q,p) = \begin{pmatrix} 0 & \rho_D(q) \\ -\rho_D^T(q) & \mathcal{C}(q,p) \end{pmatrix}$$

is $(n+r)\times(n+r)$ and skew-symmetric. Its kernel is spanned by the vectors $(\beta^a, 0)$ with $a = 1, \dots n-r$.

One can now define a bracket of functions on D^* by

$$\{f,g\}_{nh}(q,p) = \left(\frac{\partial f}{\partial q}, \frac{\partial f}{\partial p} \right) \pi_{nh}(q,p) \begin{pmatrix} \frac{\partial g}{\partial q} \\ \frac{\partial g}{\partial p} \end{pmatrix}.$$

Equivalently, the bracket can be defined in terms of the coordinate functions q^i, p_α as

$$\{q^i, q^k\}_{nh} = 0, \qquad \{q^i, p_\alpha\}_{nh} = \rho^i_\alpha = X_\alpha[q^i], \qquad \{p_\alpha, p_\beta\}_{nh} = -C^{\gamma}_{\alpha\beta} p_\gamma.$$

The above definition implies

$$\{f, gh\}_{nh} = h\{f,g\}_{nh} + g\{f,h\}_{nh}. \tag{6.1}$$

Moreover, using the skew-symmetry of $\pi_{nh}(q,p)$ we have

$$\{f,g\}_{nh} = -\{g,f\}_{nh}. \tag{6.2}$$

And that

$$X_{nh}[f] = \{f, H_c\}_{nh}$$

for any $f \in C^\infty(D^*)$. The last equation is reminiscent of the Poisson formulation of Hamiltonian systems. In fact, the identities (6.1) and (6.2) two of the properties that a Poisson bracket must satisfy. We shall see that the third property, the so-called *Jacobi identity*, only holds if the constraints are holonomic.

6.1. **Intrinsic description of the nonholonomic bracket and its relation with Dirac Brackets.** So far, we have thought of D^* as an abstract space. We can interpret it as a subbundle of T^*Q in the following way. We have the usual isomorphism (Legendre transformation)

$$\Psi : TQ \to T^*Q, \qquad \Psi(q,\dot{q}) = (q, A(q)\dot{q})$$

as usual: $\Psi(v_q)(\cdot) = \langle\langle v_q, \cdot\rangle\rangle$. If we denote $M = \Psi(D) \subset T^*Q$ then

$$\begin{tikzcd} D \arrow[r, "\Psi"] \arrow[d, "\psi"'] & M \\ D^* \arrow[ur, "\mathcal{P}^*"'] & \end{tikzcd}$$

From now on, we think of D^* as a subbundle of T^*Q.

Recall that D^* is equipped with a rank $2r$-distribution $\mathcal{F}$, and that X_{nh} is tangent to $\mathcal{F}$. The distribution $\mathcal{F}$ has the following properties:

(i) $\mathcal{F}$ is integrable if and only if the constraints are holonomic.

(ii) $\mathcal{F}$ is *symplectic*. We now explain what this means.

Let $(q,p) = m \in D^* \subset T^*Q$, and denote by Ω the canonical symplectic structure on T^*Q. Then Ω_m is a skew-symmetric non-degenerate bilinear form on $T_m(T^*Q)$. The distribution $\mathcal{F}$ defines a $2r$-dimensional subspace $\mathcal{F}_m \subset T_m(T^*Q)$. The condition that $\mathcal{F}$ is symplectic means that the the restriction of Ω_m to $\mathcal{F}_m$, denoted $\Omega|_{\mathcal{F}_m}$ is non-degenerate for all $m \in D^*$.

Apparently, this property was first noticed by Weber [16].

(iii) The vector field X_{nh} can be characterized by the following relation holding at all points $m \in D^*$:

$$i_{X_{nh}(m)}\left(\Omega_m|_{\mathcal{F}_m}\right) = dH_c(m)|_{\mathcal{F}_m}.$$

The validity of this characterization relies on the symplectic property of $\mathcal{F}$ and is reminiscent of the symplectic formulation of Hamiltonian mechanics. However, it is important to remember that $\Omega|_{\mathcal{F}_m}$ is *not* a two-form on D^*. This approach to the formulation of the equations of motion is sometimes called *distributional Hamiltonian* (see [6] and references therein).

(iv) Since $\mathcal{F}$ is symplectic, then at each $m \in D^*$ there is a decomposition

$$T_m(T^*Q) = \mathcal{F}_m \oplus \mathcal{F}_m^{\Omega}$$

where $\mathcal{F}_m^{\Omega}$ denotes the symplectic orthogonal of $\mathcal{F}_m$. Let $\mathcal{Q}_m : T_m(T^*Q) \to \mathcal{F}_m$ be the projector associated with this decomposition.

It is shown in [14] that the bracket $\{\cdot,\cdot\}_{nh}$ admits the following intrinsic definition

$$\{f,g\}_{nh}(m) = \Omega_m(\mathcal{Q}_m X_{\tilde{f}}(m), \mathcal{Q}_m X_{\tilde{g}}(m)).$$

Here $f, g \in C^\infty(D^*)$. $\tilde{f}, \tilde{g} \in C^\infty(T^*Q)$ are arbitrary smooth extensions of f, g, and $X_{\tilde{f}}$, $X_{\tilde{g}}$ are the Hamiltonian vector fields of $\tilde{f}, \tilde{g}$. ($i_{X_{\tilde{f}}}\Omega = d\tilde{f}$ and similarly for $X_{\tilde{g}}$).

Exercise Show that if $f \in C^\infty(D^*)$, the *almost Hamiltonian* vector field X_f^{nh} on D^* defined by

$$X_f^{nh}[g] = \{f,g\}_{nh}$$

for all $g \in C^\infty(D^*)$ satisfies $X_f^{nh} = \mathcal{Q}X_{\tilde{f}}$ where $X_{\tilde{f}}$ is the Hamiltonian vector field associated to $\tilde{f}$ with respect to the symplectic structure Ω on T^*Q where $\tilde{f}$ is an arbitrary smooth extension of f to T^*Q.

From the above exercise we draw the conclusion that the almost Hamiltonian vector fields X_f^{nh} with $f \in C^\infty(D^*)$ are tangent to $\mathcal{F}$. In fact one can show that the span of all such vector fields is precisely $\mathcal{F}$.

For a Poisson bracket the distribution spanned by the Hamiltonian vector fields is integrable and defines a *symplectic foliation*. If the constraints are nonholonomic, then $\mathcal{F}$ is non-integrable, and the bracket $\{\cdot,\cdot\}_{nh}$ cannot satisfy the Jacobi identity.

If the constraints are holonomic, then $\{\cdot,\cdot\}_{nh}$ coincides with Dirac's construction of Poisson brackets for systems with constraints (see [14]).

In conclusion, $\{\cdot,\cdot\}_{nh}$ satisfies the Jacobi identity if and only if the constraints are holonomic.

6.2. **Example.** Recall that for the Chaplygin sleigh we found the following equations of motion

$$\begin{aligned} m\dot{u} &= ma\omega^2 \\ (J+ma^2)\dot{\omega} &= -mau\omega, \\ \dot{x} &= u\cos\theta, \\ \dot{y} &= u\sin\theta, \\ \dot{\theta} &= \omega. \end{aligned}$$

The constrained Lagrangian

$$L_c = \frac{1}{2}\left((J+ma^2)\omega^2 + mu^2\right).$$

and

$$C_{12}^1 = 0, \qquad C_{12}^2 = -\frac{ma}{J+ma^2}.$$

We have

$$p_u = \frac{\partial L_c}{\partial u} = mu, \qquad p_\omega = \frac{\partial L_c}{\partial \omega} = (J+ma^2)\omega.$$

The Hamiltonian is

$$H = \frac{1}{2}\left(\frac{p_\omega^2}{J+ma^2} + \frac{p_u^2}{m}\right).$$

The equations of motion can be rewritten as

$$\begin{aligned} &\dot{x} = \frac{p_u}{m}\cos\theta, \qquad \dot{y} = \frac{p_u}{m}\sin\theta, \qquad \dot{\theta} = \frac{p_\omega}{J+ma^2} \\ &\dot{p}_u = \frac{map_\omega^2}{(J+ma^2)^2}, \qquad \dot{p}_\omega = -\frac{ap_u p_\omega}{J+ma^2}, \end{aligned}$$

or, equivalently,

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ \theta \\ p_u \\ p_\theta \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & \sin\theta & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -\cos\theta & -\sin\theta & 0 & 0 & \frac{ma}{J+ma^2}p_\omega \\ 0 & 0 & -1 & -\frac{ma}{J+ma^2}p_\omega & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial H_c}{\partial x} \\ \frac{\partial H_c}{\partial y} \\ \frac{\partial H_c}{\partial \theta} \\ \frac{\partial H_c}{\partial p_u} \\ \frac{\partial H_c}{\partial p_\omega} \end{pmatrix}.$$

The above 5×5 matrix defines the almost Poisson bracket $\{\cdot,\cdot\}_{nh}$. It has rank 4 and its null space is spanned by the vector $(-\sin\theta, \cos\theta, 0, 0, 0)$ that should be interpreted as the one-form $\sin\theta\, dx + \cos\theta\, dy$.

On the other hand, the vectors

$$\cos\theta\frac{\partial}{\partial x}+\sin\theta\frac{\partial}{\partial y},\qquad \frac{\partial}{\partial\theta},\qquad \frac{\partial}{\partial p_u},\qquad \frac{\partial}{\partial p_\omega},$$

span the range of the matrix that should be interpreted as the distribution $\mathcal{F}$.

7. Invariant measures

Recall that the generic solutions of the Suslov problem asymptotically approach periodic orbits as $t\to\pm\infty$. This behavior cannot occur in Hamiltonian mechanical systems. An obstruction for this is the existence of the *invariant Liouville measure*.

An invariant measure is a very important property for a dynamical system. We will investigate the existence of such an invariant for nonholonomic systems with symmetry. It will turn out that only exceptional systems possessing many symmetries have an invariant measure.

7.1. **Invariant measures for ODE's.** To fix ideas consider a vector field f defined on $\mathbb{R}^n$ and consider the autonomous system of differential equations

$$\dot{x}=f(x). \tag{7.1}$$

The flow ϕ_t of this equation satisfies

$$\frac{d\phi_t(x)}{dt}=f(\phi_t(x)),\qquad \phi_0(x)=x.$$

A smooth volume form on $\mathbb{R}^n$ is an expression like $\mu(x)\,dx$ where μ is a smooth real valued function on $\mathbb{R}^n$ such that $\mu(x)>0$ for all $x\in\mathbb{R}^n$ and dx is the euclidean volume form $dx=dx^1\wedge\cdots\wedge dx^n$ where $(x^1,\dots,x^n)$ are linear coordinates on $\mathbb{R}^n$. The function μ is called the density of the measure with respect to dx.

The measure $\mu(x)\,dx$ is *invariant* under the flow ϕ_t of (7.1) if

$$\int_A\mu(x)\,dx=\int_{\phi_t(A)}\mu(x)\,dx$$

for any region $A\subset\mathbb{R}^n$.

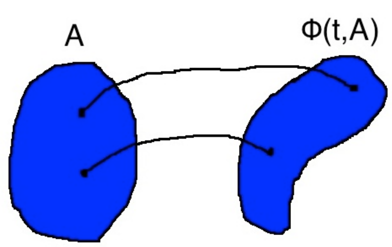


FIGURE 2. If $\mu(x)\,dx$ is an invariant measure then the volume of A with respect to $\mu(x)\,dx$ coincides with the volume of $\phi_t(A)$ with respect to $\mu(x)\,dx$ for all t and any $A\subset\mathbb{R}^n$.

Exercise Show that

$$\left.\frac{d}{dt}\right|_{t=0}\int_{\phi_t(A)}\mu(x)\,dx=\int_A\operatorname{div}(\mu(x)f(x))\,dx.$$

The condition for the existence of an invariant measure becomes the following linear PDE for μ

$$\operatorname{div}(\mu(x)f(x)) = 0.$$

The existence of a *global* solution to this equation is equivalent to the existence of an invariant measure. We stress that one needs a global solution. Locally, away from equilibrium points, the flow can be linearized and an invariant measure exists.

Asymptotic equilibria and asymptotic periodic orbits are obstructions to the existence of an invariant measure.

Symplectic Hamiltonian systems always preserve the Liouville measure. A direct calculation using the equality of mixed partials shows that the divergence of

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q^i}$$

equals zero (an intrinsic proof is also not hard to give using that the Hamiltonian flow preserves the symplectic form).

Our main tool to study the preservation of volumes for nonholonomic systems with symmetries is the following.

Proposition 7.1 (Kozlov [15]). *If the vector field f is homogeneous of degree 2, namely, if*

$$f(\lambda x) = \lambda^2 f(x)$$

for any $\lambda \in \mathbb{R}$ and $x \in \mathbb{R}^n$; then f preserves the measure $\mu(x)\,dx$ if and only if it preserves the measure dx and μ is a first integral.

Proof. Suppose that $\operatorname{div}(\mu f) = 0$. We shall prove that $\operatorname{div}(f) = 0$ and $\langle \nabla \mu, f\rangle = 0$. Since $\mu > 0$ we can put $\mu = e^\sigma$ for a smooth function σ on $\mathbb{R}^n$. We have

$$0 = \operatorname{div}(\mu f) = \operatorname{div}(e^\sigma f) = e^\sigma(\langle \nabla\sigma, f\rangle + \operatorname{div}(f)).$$

It follows that

$$\operatorname{div}(f) = -\langle \nabla\sigma, f\rangle.$$

Since f is homogeneous of degree 2, the left member in the last equation is homogeneous of degree 1. On the other hand, given that σ is smooth, the right hand side cannot be homogeneous of degree 1. Therefore, the above equation can only hold if both sides of the equality vanish.

The other implication is trivial. □

Example For the Suslov problem we found the set of decoupled equations on $\mathbb{R}^2$:

$$\dot{\Omega}_1 = -\frac{1}{I_{11}}\left((I_{13}\Omega_1 + I_{23}\Omega_2)\Omega_2\right),$$
$$\dot{\Omega}_2 = \frac{1}{I_{22}}\left((I_{13}\Omega_1 + I_{23}\Omega_2)\Omega_1\right).$$

The right hand side of these equations defines a homogeneous vector field of degree 2 on $\mathbb{R}^2$. By the proposition, our candidate for invariant measure is $d\Omega_1 \wedge d\Omega_2$. Taking divergence with respect to this measure we get:

$$\operatorname{div}(\,) = -\frac{I_{13}\Omega_2}{I_{11}} + \frac{I_{23}\Omega_1}{I_{22}}.$$

This quantity vanishes for arbitrary (Ω_1, Ω_2) if and only if $I_{13} = I_{23} = 0$.

Physically, the condition for measure preservation is that the axis of forbidden rotations is an eigenvector of the inertia tensor $\mathbb{I}$. In other words, the axis of forbidden rotations is a principal axis of inertia of the body.

For a general system, the equations of motion on the Hamiltonian side are given by

$$\dot{q}^i = \rho^i_\alpha \frac{\partial H}{\partial p_\alpha},$$
$$\dot{p}_\alpha = -\rho^i_\alpha \frac{\partial H}{\partial q^i} - C^\gamma_{\alpha\beta} p_\gamma \frac{\partial H}{\partial p_\gamma}$$

where we have written $H_c = H$ to simplify notation.

In the absence of a potential, the Hamiltonian is $H = \frac{1}{2} p^T A_D(q)^{-1} p$. Notice that under this condition the equations of motion for p_α are homogeneous quadratic. In this case, Proposition 7.1 can be generalized [10] to show that the only candidates for an invariant measure are basic measures $\mu(q)\, dq \wedge dp$ (the density does not depend on p).

8. Nonholonomic systems with symmetry

Consider, as before, a nonholonomic system with Lagrangian $L : TQ \to \mathbb{R}$, $L(q, \dot{q}) = \frac{1}{2}\dot{q}^T A(q)\dot{q} - U(q)$ and constraints $\beta(q)\dot{q} = 0$ that define a nonintegrable distribution $\mathcal{D}$ of rank r and a subbundle $D \subset TQ$.

Let G be a Lie group and suppose that it acts freely and properly on Q,

$$\Psi : G \times Q \to Q.$$

Ψ_g is a diffeomorphism on Q for every $g \in G$ and $\Psi_h \circ \Psi_g = \Psi_{hg}$ for all $g, h \in G$.

The tangent lift of Ψ is the action

$$\hat{\Psi} : G \times TQ \to TQ$$

defined by $\hat{\Psi}_g(v_q) = T_q\Psi_g(v_q)$ for $v_q \in T_qQ$.

We are interested in actions that *preserve the Lagrangian*:

$$L \circ \hat{\Psi}_g = \hat{\Psi}_g \qquad \text{for all} \qquad g \in G,$$

and also *preserve the constraints*:

$$\hat{\Psi}(\mathcal{D}_q) = \mathcal{D}_{\Psi_g(q)} \qquad \text{for all} \qquad g \in G,\ q \in Q.$$

In this case, $\hat{\Psi}$ can be restricted to define an action on D

$$\Phi : G \times D \to D, \qquad \Phi_g = \hat{\Psi}_g\Big|_D \quad \text{for all} \quad g \in G.$$

With the above hypothesis, the vector field X_{nh} is equivariant, namely,

$$X_{nh}(\Phi_g(q, v)) = T_{(q,v)}\Phi_g(X_{nh}(q, v)),$$

for all $g \in G$, $(q, v) \in D$, and the dynamics drop to the quotient space D/G that is rank r vector bundle over Q/G.

Example. Chaplygin sleigh. The configuration space $Q = \mathbb{R}^2 \times S^1$ has coordinates $q = (x, y, \theta)$.

The system is invariant under the action of the group $SE(2)$ of euclidean motions on the plane. An element of such space can be represented as $g = (R_\varphi, (A, B))$ where

$$R_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \qquad (A, B) \in \mathbb{R}^2.$$

The action Ψ is defined by

$$\Psi_g(x, y, \theta) = \left(R_\varphi \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} A \\ B \end{pmatrix}, \theta + \varphi \right).$$

It corresponds to our freedom to select the origin and orientation of the space frame without altering the dynamical equations. In fact, *any* problem of a body rolling on a homogenous plane will possess the same symmetry.

In this case Q/G is one point since the action is transitive, and D/G is isomorphic to a two dimensional vector space. The reduced equations of motion are

$$\begin{aligned} m\dot{u} &= ma\omega^2 \\ (J + ma^2)\dot{\omega} &= -mau\omega, \end{aligned}$$

that we had found before.

We note that $\mathcal{D}_q \subset T_q\mathrm{Orb}_G(q) = T_qQ$. In fact, we can take $Q = G$. With this interpretation of the configuration space, we have that it is a Lie group and both the constraints and the Lagrangian are invariant under left multiplication. This approach makes the Chaplygin sleigh an example of an *LL system*.

It is important to understand the geometric that allowed us to decouple the equations of motion and obtain the reduced system. The key is that the vector fields $X_1 = \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}$ and $X_2 = \frac{\partial}{\partial\theta}$ that define the moving frame for the quasi-velocities u, ω are equivariant with respect to the $SE(2)$ action on Q.

Example. Suslov Problem. The configuration space is $Q = SO(3)$. Both the constraint and the Lagrangian are written in terms of the angular velocity written in the body frame Ω without involving the attitude matrix B.

The symmetry group is $G = SO(3)$ acting by left multiplication $\Psi_g(B) = gB$. The Suslov problem is also an LL system. Once again the reduced system is isomorphic to $\mathbb{R}^2$ and the reduced equations are

$$\begin{aligned} \dot{\Omega}_1 &= -\frac{1}{I_{11}}\left((I_{13}\Omega_1 + I_{23}\Omega_2)\Omega_2\right), \\ \dot{\Omega}_2 &= \frac{1}{I_{22}}\left((I_{13}\Omega_1 + I_{23}\Omega_2)\Omega_1\right). \end{aligned}$$

The moving frame for Ω_1 Ω_2 is equivariant with respect to left multiplication on $SO(3)$.

Note that once again we have $\mathcal{D}_q \subset T_q\mathrm{Orb}_G(q) = T_qQ$.

Example. Veselova problem. The configuration space is $Q = SO(3)$, the Lagrangian is $L = \frac{1}{2}\langle \mathbb{I}\Omega, \Omega\rangle$. The constraint is $\omega_3 = \langle \gamma, \Omega\rangle = 0$ where $\gamma = B^{-1}e_3$.

The constraint involves the entries of B so we do not have an LL system.

The system has a $G = SO(2)$ symmetry. Denote elements in G by

$$g = \hat{R}_\varphi = \begin{pmatrix} R_\varphi & 0 \\ 0 & 1 \end{pmatrix}$$

where R_φ is the 2×2 rotation matrix defined above. For such g the action is

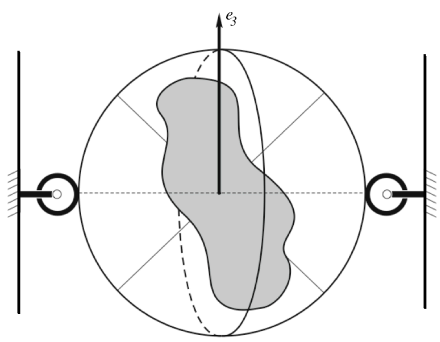
$$\Psi_g(B) = gB.$$

Note that the Poisson vector γ is invariant under the action of $SO(2)$:

$$\gamma \mapsto (gB)^{-1}e_3 = B^{-1}g^{-1}e_3 = B^{-1}e_3 = \gamma.$$

This is a mathematical verification of a simple fact. The action rotates the space axis about the third axis e_3. Hence e_3 and its avatar γ in the body frame remain invariant.

The existence of this symmetry is clear from the physical realization of the Veselova problem:



The evolution of the system does not care about the orientation that we choose for the e_1, e_2 space axes.

In this case we have $Q/G = SO(3)/SO(2) = S^2$. In fact γ can be used as a coordinate on S^2.

Also $TQ = \mathcal{D}_q \oplus T_q\mathrm{Orb}_G(q)$ for all $q \in Q$. Systems with this property are called *Chaplygin systems*. For this family of systems the reduced space D/G is isomorphic to $T(Q/G)$ (see e.g. [13]).

Hence, the reduced phase space for the Veselova system is TS^2. The reduced equations were given in section 4.2.

Example. Chaplygin top. Since the rolling takes place on a homogeneous plane, the system possesses an $SE(2)$ symmetry.

The configuration space is $Q = SO(3) \times \mathbb{R}^2$ with coordinates $q = (B, (x, y))$. The action of $g = (R_\varphi, (A, B)) \in SE(2)$ on Q is given by

$$\Psi_g(q) = \left(\hat{R}_\varphi B, R_\varphi \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} A \\ B \end{pmatrix} \right).$$

We have $Q/SE(2) = S^2$. Once again, the Poisson vector $\gamma \in S^2$ can be used as a coordinate on the orbit space. The reduced space is $D/SE(2)$ is a rank 3 vector bundle over S^2. The entries of Ω can be interpreted as linear coordinates on the fibers. The reduced equations of motion were given in section 4.3.

For this system we have $T_qQ = \mathcal{D}_q + T_q\mathrm{Orb}_G(q)$ with

$$\dim(\mathcal{D}_q \cap T_q\mathrm{Orb}_G(q)) = 1$$

for all $q \in Q$. The intersection is spanned by the action of the subgroup of $SE(2)$ that corresponds to rotations about the contact point.

8.1. **Reduced equations of motion.** Assume that the system has a symmetry group G. The action Φ on D can be transferred to D^*. We continue to use the symbol Φ for this new action. The bracket $\{\cdot,\cdot\}_{nh}$ is invariant. Namely,

$$\{f\circ\Phi_g, g\circ\Phi_g\}_{nh} = \{f,g\}_{nh}\circ\Phi_g.$$

If $\pi : D^* \to D^*/G$ is the orbit projection then we can define a bracket on D^*/G by the rule

$$\{F,G\}_{D^*/G} = \{F\circ\pi, G\circ\pi\}_{nh}$$

for functions $F, G \in C^\infty(D^*/G)$. The Hamiltonian function is invariant under the action and therefore there exists a function $h : D^*/G \to \mathbb{R}$ with the property that $h\circ\Phi_g = H$. (h is the *reduced* Hamiltonian). The reduced dynamics are described by the vector field X on D^*/G determined by the rule

$$X[F] = \{F,h\}_{D^*/G}$$

for all $F \in C^\infty(D^*/G)$.

We now proceed to write the equations of motion in local coordinates for D^*/G.

If the vector fields X_α that define the quasi-velocities v^α are G-equivariant, the corresponding momenta p_α are invariant functions on D^* and pass to the quotient D^*/G. We shall make this assumption from now on.

Denote by s^j some local coordinates on Q/G. Then we can take (s,p) as coordinates for D^*/G. The reduced bracket is defined in this coordinates by the relations

$$\{p_\alpha,p_\beta\}_{D^*/G} = -C_{\alpha\beta}^\gamma(s)p_\gamma, \qquad \{s^i,s^j\}_{D^*/G} = 0, \qquad \{s^i,p_\alpha\} = (\pi_*X_\alpha)[s^i].$$

Recall that the structure functions $C_{\alpha\beta}^\gamma(s)$ were defined by the relations

$$\mathcal{P}[X_\alpha,X_\beta] = C_{\alpha\beta}^\gamma X_\gamma,$$

where $\mathcal{P} : TQ \to D$ is the kinetic energy-orthogonal projection. These functions are G-invariant since the vector fields X_α are G-equivariant and by invariance of the kinetic energy. It follows that they can be expressed in terms of the reduced coordinates s^i.

Further simplifications can be made if the G-equivariant vector fields X_α are chosen as $\{X_\alpha\} = \{Z_a, Y_A\}$ such that

$$\operatorname{span}\{Z_a(q)\} = \mathcal{D}_q \cap T_q\mathrm{Orb}_G(q), \qquad \text{for all} \quad q\in Q$$

then $\pi_*(Z_a) = 0$.

Suppose that $\pi_*(Y_A) = \rho_A^i(s)\frac{\partial}{\partial s^i}$. Then

$$\{q^i,p_a\}_{D^*/G} = 0, \qquad \{q^i,p_A\}_{D^*/G} = \rho_A^i(s).$$

Moreover, if the $Y_A(q)$ are chosen to be orthogonal (with respect to the kinetic energy metric) to $\operatorname{span}\{Z_a(q)\}$. the kinetic energy matrix block diagonalizes. In particular

$$h(s,p_a,p_A) = \frac{1}{2}K^{ab}(s)p_ap_b + \frac{1}{2}T^{AB}(s)p_Ap_B$$

for certain symmetric positive definite matrices $K(s)$ and $T(s)$.

With this choice of moving frame, the reduced equations take the form:

$$\begin{aligned}
\dot{s}^i &= \rho^i_A(s)\frac{\partial h}{\partial p_A},\\
\dot{p}_A &= -\rho^i_A(s)\frac{\partial h}{\partial q^i} - C^{\beta}_{A\alpha}(s)p_\beta\frac{\partial h}{\partial p_\alpha},\\
\dot{p}_a &= -C^{\gamma}_{a\beta}(s)p_\gamma\frac{\partial h}{\partial p_\beta}.
\end{aligned}$$

This is the convenient form of the equations to investigate the existence of a smooth invariant measure. Multiplying the above vector field with $\mu = e^{\sigma(s)}$ and taking divergence with respect to $ds \wedge dp_A \wedge dp_a$ and setting the result equal to zero yields the following PDE for σ:

$$\frac{\partial}{\partial s^i}\left(e^\sigma \rho^i_A(s)\frac{\partial h}{\partial p_A}\right) + e^\sigma\frac{\partial}{\partial p_A}\left(-\rho^i_A(s)\frac{\partial h}{\partial q^i} - C^{\beta}_{A\alpha}(s)p_\beta\frac{\partial h}{\partial p_\alpha}\right) + e^\sigma\frac{\partial}{\partial p_a}\left(-C^{\gamma}_{a\beta}(s)p_\gamma\frac{\partial h}{\partial p_\beta}\right) = 0.$$

After some cancellations and rearrangement of terms we get

$$\left(\rho^i_A\frac{\partial \sigma}{\partial s^i} + \frac{\partial \rho^i_A}{\partial s^i} + C^{\alpha}_{A\alpha}\right)\frac{\partial h}{\partial p_A} + \left(C^{\alpha}_{a\alpha}(s)\right)\frac{\partial h}{\partial p_a} = 0. \tag{8.1}$$

Using that

$$\frac{\partial h}{\partial p_A} = T^{AB}(s)p_B, \qquad \frac{\partial h}{\partial p_\alpha} = K^{ab}(s)p_b$$

and that the matrices T and K are non-degenerate, it follows that (8.1) can only hold if

$$\begin{aligned}
&\rho^i_A\frac{\partial \sigma}{\partial s^i} + \frac{\partial \rho^i_A}{\partial s^i} + C^{\alpha}_{A\alpha}(s) = 0, \qquad \text{for all} \qquad A,\\
&\text{and} \qquad C^{\alpha}_{a\alpha}(s) = 0 \qquad \text{for all} \qquad a.
\end{aligned} \tag{8.2}$$

While the first set of the above conditions involve the unknown function σ, the second set does not. This can be very valuable to prove that there is no smooth invariant measure. If for some value of the coordinates s, we have $C^{\alpha}_{a\alpha}(s) \neq 0$, then there cannot exist a smooth invariant measure.

Example. Chaplygin sleigh. We have $\mathcal{D}_q \subset T_q\mathrm{Orb}_G(q)$ and

$$Z_1 = \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}, \qquad Z_2 = \frac{\partial}{\partial \theta}.$$

The set of vector fields $\{Y_A\}$ is empty. We have computed

$$C^1_{12} = 0, \qquad C^2_{12} = -\frac{ma}{J+ma^2}.$$

A necessary condition for the existence of an invariant measure is that

$$C^2_{12} = 0,$$

that can only happen if $a = 0$. Namely, if the contact point of the sleigh coincides with the center of mass.

Example. Chaplygin top. As we mentioned above, for this system the intersection $\mathcal{D}_q \cap T_q\mathrm{Orb}_G(q)$ has dimension 1 for all $q \in Q$. Associated to it there is a necessary condition for the existence of a smooth invariant measure $C^{\alpha}_{1\alpha}(s^1, s^2) = 0$ where (s^1, s^2) are coordinates for the 2-sphere S^2.

As it is shown in [10] the condition is satisfied if the following conditions on the parameters are satisfied

(i) $\ell = 0$. In this case the center of mass coincides with the geometric center. This system is called the *Chaplygin sphere* (it is a special case of the Chaplygin top). An invariant measure in this case was discovered by Chaplygin 1903.
(ii) $I_{11} = I_{22}$ and $I_{13} = I_{23} = 0$. The mass distribution on the sphere is axially symmetric. This system is called the Routh sphere. An invariant measure in this case was discovered by Routh 1884.

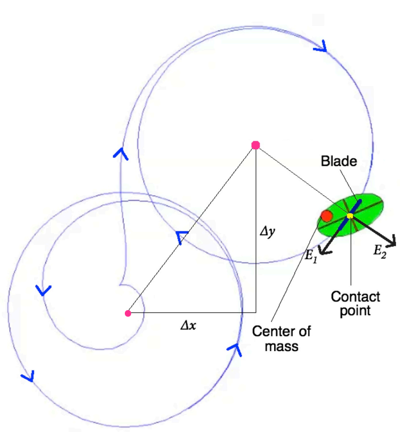
This analysis shows that, *in the absence of gravity*, there can only exist a smooth invariant measure in these two situations.

A similar approach can be applied to the rolling of a homogeneous ellipsoid on the plane [11]. The conclusion in this case is that there is an invariant measure if and only if two of the semi-axis of the ellipsoid are equal and it is a body of revolution.

This analysis suggests that invariant measures are *rare* in nonholonomic mechanics and leads to the following question: should one expect to find attractors (and hence, by reversibility of the dynamics, repellers) for generic nonholonomic systems? What happens to systems whose parameters are close to those for which the system has an invariant measure (for instance take $\ell/R << 1$ in the Chaplygin top)? At present time there is no satisfactory answer to these questions.

We mention that reversal (rattle-back like) phenomena have been shown to exist for the Chaplygin top [3].

Another system where the reversal phenomena (asymptotic motion approaching periodic orbits) can be explicitly seen is the hydrodynamic Chaplygin sleigh [8].



When circulation is added to the system, one observes two different behavior regimes. For small energies the motion is driven by the fluid's circulation and the trajectories are generically quasi-periodic on two-tori. On the other hand, once a threshold value of the energy is attained, the system exhibits reversal phenomena [9]. This energy dependent behavior is also observed in the motion of articulated vehicles (e.g. the case $n = 2$, $a = 0$ treated in [4]).

8.2. **Invariant measures for Chaplygin systems.** Suppose that we have a G-Chaplygin system. Namely, the action of the Lie group G satisfies

$$T_qQ = \mathcal{D}_q \oplus T_q\mathrm{Orb}_G(q)$$

for all $q \in Q$.

In this case the set of vector fields $\{Z_a\}$ is empty and only the first of the conditions (8.2) should be analysed. Moreover, for Chaplygin systems the indices A, B, C and i, j, k run over the same range and the matrix ρ_A^j is invertible. The condition for σ becomes

$$\frac{\partial \sigma}{\partial s^i} = g_i(s)$$

where the functions

$$g_i(s) = -\rho_i^B(s)\frac{\partial \rho_B^j}{\partial s^j}(s) - \rho_i^B(s)C_{B\alpha}^{\alpha}(s).$$

The existence of a function σ satisfying the above conditions is equivalent to the closeness of the one-form

$$g_i(s)\, ds^i.$$

This condition was derived in intrinsic terms in [5] (see also [17]).

An inhomogenous ball rolling without slipping on a cylinder is an example of a Chaplygin system. The system only possesses an invariant measure if the ball is homogeneous [10].

9. Nonholonomic Noether's Theorem

We recall the classical Noether Theorem for a Lagrangian holonomic system.

Let $\Psi : G \times Q \to Q$ be an action such that the lifted action $\hat{\Psi} : G \times TQ \to TQ$ preserves the Lagrangian function $L : TQ \to \mathbb{R}$, namely, $L \circ \hat{\Psi}_g = L$ for all $g \in G$.

Let $\xi \in \mathfrak{g}$, the Lie algebra of G, and denote by ξ_Q the *infinitesimal generator* of Ψ. Then ξ_Q is a vector field on Q defined by

$$\xi_Q(q) = \left.\frac{d}{dt}\right|_{t=0} \Psi_{\exp(t\xi)}(q).$$

Let ξ_Q be given in local coordinates by

$$\xi_Q(q) = \xi^i(q)\frac{\partial}{\partial q^i}.$$

The *tangent lift* of ξ_Q is the vector field $\hat{\xi}_Q$ on TQ given in bundle coordinates by

$$\hat{\xi}_Q(q, \dot{q}) = \xi^i(q)\frac{\partial}{\partial q^i} + \dot{q}^j\frac{\partial \xi^i}{\partial q^j}(q)\frac{\partial}{\partial \dot{q}^i}.$$

This vector field coincides with the infinitesimal generator ξ_{TQ} of $\hat{\Psi}$. Invariance of the Lagrangian implies

$$\hat{\xi}_Q[L] = \xi_{TQ}[L] = 0.$$

Now define the momentum p_ξ associated to the vector field ξ_Q. $p_\xi : TQ \to \mathbb{R}$ is the linear function on the fibers of TQ (on the velocities) given by

$$p_\xi = \langle\langle \xi_Q, \cdot \rangle\rangle_Q = \xi^i(q) \frac{\partial L}{\partial \dot{q}^i}.$$

A direct calculation, that uses the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0, \qquad i = 1, \dots, n, \tag{9.1}$$

gives

$$\begin{aligned} \dot{p}_\xi &= \frac{\partial \xi^i}{\partial q^j}\dot{q}^j \frac{\partial L}{\partial \dot{q}^i} + \xi^i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) \\ &= \frac{\partial \xi^i}{\partial q^j}\dot{q}^j \frac{\partial L}{\partial \dot{q}^i} + \xi^i \frac{\partial L}{\partial q^i} \\ &= \hat{\xi}_Q[L] = 0, \end{aligned}$$

which shows that p_ξ is a first integral.

In the nonholonomic setting, the Euler-Lagrange equations (9.1) get replaced by

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = R_i, \qquad i = 1, \dots, n, \tag{9.2}$$

where the reaction forces R_i obey the Lagrange-D'Alembert principle. In this case we get

$$\dot{p}_\xi = \hat{\xi}_Q[L] + \langle R, \xi_Q \rangle.$$

This straightforward calculation immediately leads to

Theorem 9.1. *If the action leaves L invariant and suppose that for a fixed Lie algebra element $\xi \in \mathfrak{g}$ we have $\xi_Q(q) \in \mathcal{D}_q$ for all $q \in Q$. Then p_ξ is a first integral of* (9.2).

Proof. By the Lagrange-D'Alembert principle $\langle R, \xi_Q \rangle = 0$ since $\xi_Q(q) \in \mathcal{D}_q$. □

The above result is sometimes referred to as *Nonholonomic Noether's Theorem*. The symmetry generated by the Lie algebra element ξ is called a *horizontal symmetry*.

Note that the theorem does not require D to be invariant under the action! In fact, in many cases this will not be the case.

Example. Chaplygin sleigh Consider $SO(2)$-action

$$\varphi \cdot (\theta, x, y).$$

The infinitesimal generator is $\xi_Q = \frac{\partial}{\partial \theta}$. Its tangent lift is $\hat{\xi}_Q = \frac{\partial}{\partial \theta}$. Recall that

$$L(x, y, \theta, \dot{x}, \dot{y}, \dot{\theta}) = \frac{1}{2}\left((J + ma^2)\dot{\theta}^2 + m(\dot{x}^2 + \dot{y}^2) + 2ma\dot{\theta}(\cos\theta \dot{y} - \sin\theta \dot{x})\right).$$

Then

$$\hat{\xi}_Q[L] = -ma\dot{\theta}(\cos\theta \dot{x} + \sin\theta \dot{y}) = -mau\omega.$$

The above vanishes for any value of u, ω only if $a = 0$. Conclusion $p_\xi = (J + ma^2)\dot{\theta}$ is constant if $a = 0$.

Example. Chaplygin sphere.

Take $G = SO(2)$. The action on $Q = SO(3) \times \mathbb{R}^2$ is

$$\varphi \cdot (B, (x, y)) = \left(\hat{R}_\varphi B, (x, y) \right).$$

The lift of this action certainly preserves the Lagrangian. On the other hand, *if the center of mass coincides with the geometric center* then the infinitesimal generator $\xi_Q(q) \in \mathcal{D}_q$. This can be understood as follows. The action is a rotation of the ball about the vertical axis passing through the center of mass. This does not violate the constraint as long as the axis passes through the contact point. But this situation only occurs if the center of mass coincides with the geometric center.

The corresponding first integral for the Chaplygin sphere is the vertical component of angular momentum $\langle \mathbb{I}\Omega, \gamma \rangle = \langle K, \gamma \rangle$.

9.1. **Gauge momenta.** There are a great number of examples of nonholonomic systems that possess first integrals linear in the velocities that do not correspond to horizontal symmetries.

To explain how they arise, note that in the presence of the kinetic energy metric, a linear function p_ξ on TQ is associated to a vector field ξ on Q.

$$p_\xi = \langle\langle \xi(q), \cdot \rangle\rangle_q = \xi(q)^T A(q) \dot{q} = \xi^j(q) \frac{\partial L}{\partial \dot{q}^j}(q, \dot{q}).$$

The vector field ξ is called the generator of p_ξ. We are interested in the restriction $p_\xi|_D$. It is not difficult to see that a linear function on D has a unique generator that is a section of D. Assume that $\xi(q) \in \mathcal{D}_q$ for all $q \in Q$. Proceeding as before we find

$$\dot{p}_\xi = \hat{\xi}[L] + \langle R, \xi \rangle.$$

Now restrict to D and use the fact that ξ takes values on D to conclude that $\langle R, \xi \rangle = 0|_D$. Hence we can write

$$\frac{d}{dt} \left(p_\xi|_D \right) = \hat{\xi}(L) \Big|_D.$$

Therefore, $p_\xi|_D$ *is a first integral of* (9.2) *if and only if its unique generator* ξ *(taking values on* D*) satisfies* $\hat{\xi}(L)\Big|_D = 0$.

The above result is elementary and does not assume the existence of symmetries. However, in a series of examples possessing linear first integrals, there is a Lie group G acting on the configuration space Q and whose tangent lift preserves both L and D. Moreover, the generator ξ of the first integral satisfies

$$\xi(q) \in \mathcal{D}_q \cap T_q \mathrm{Orb}_G(q)$$

for all $q \in Q$. First integrals whose generator satisfies this property are termed *gauge momenta* [2]. In order to determine these integrals for a given system, one finds a basis of sections $\{Z_a\}$ of $D \cap T\mathrm{Orb}_G$ and makes the following ansatz for the generator ξ:

$$\xi(q) = \sum_a f^a(q) Z_a(q).$$

One then attempts to solve the equation $\hat{\xi}(L)\Big|_D = 0$ for the functions f^a. At present, it is not clear what are the conditions that allow existence of solutions to these equations.

A very important property of gauge momenta is that they remain first integrals if a G-invariant potential is added to the system. The converse implication of this statement is also true [7].

Nonholonomic systems possessing gauge momenta include bodies of revolution rolling without slipping on the plane and the rolling of a homogeneous sphere on a surface of revolution.

References

[1] Agostinelli, C.,, Nuova forma sintetica delle equazioni del moto di un sistema anolonomo ed esistenza di un integrale lineare nelle velocità, *Boll. Un. Mat. Ital.* **11** (1956), 1–9.

[2] Bates L., Graumann H. and MacDonnell C., Examples of gauge conservation laws in nonholonomic systems, *Rep. Math. Phys.* **37** (1996), 295–308.

[3] Borisov, A.V., Kazakov, A. O. and Sataev, I.R., Reversal and chaotic attractor in the nonholonomic model of Chaplygin's top *Reg. and Chaot. Dyn.* **19**, (2014), 718–733.

[4] Bravo-Doddoli A. and García-Naranjo L. C., The motion of an articulated n-trailer vehicle. *Reg. and Chaot. Dyn.* **20** (2015), 497–517.

[5] Cantrijn, F, Cortés. J., de León, M., Martín de Diego, D., On the geometry of generalized Chaplygin systems. *Math. Proc. Camb. Phil. Soc.* (2002) **132**.

[6] R. Cushman, J.J. Duistermaat and J. Śnyaticki, *Geometry of Nonholonomically Constrained Systems.* Advanced Series in Nonlinear Dynamics **26** (World Scientific, Singapore, 2010).

[7] Fassò F., Giacobbe A. and Sansonetto, N., Linear weakly Noetherian constants of motion are horizontal gauge momenta. *J. Geom. Mech.* **4** (2012), 129–136.

[8] Fedorov Y. N. and García-Naranjo, L. C., The hydrodynamic Chaplygin sleigh, *J. Phys. A: Math. Theor.* **43** (2010) 434013 (18pp).

[9] Fedorov Y. N., García-Naranjo L. C. and Vankerschaver J., The motion of the 2D hydrodynamic Chaplygin sleigh in the presence of circulation, *Disc. and Cont. Dyn. Syst. Series A* **33** (2013) no. 9, 4017–4040.

[10] Fedorov Y. N., García-Naranjo L. C. and Marrero J. C., Unimodularity and preservation of volumes in nonholonomic mechanics. *J. Nonlinear Sci.* **25** (2015), 203–246

[11] García-Naranjo L. C. and Marrero J. C., Non-existence of an invariant measure for a homogeneous ellipsoid rolling on the plane. *Reg. and Chaot. Dyn.* **18** (2013), 372–379.

[12] Grabowski, J., de León, M., Marrero, J. C. and Martín de Diego, D. Nonholonomic constraints: a new viewpoint. *J. Math. Phys.* **50** (2009), 013520, 17 pp.

[13] Hochgerner, S., García-Naranjo, L.C., G-Chaplygin systems with internal symmetries, truncation, and an (almost) symplectic view of Chaplygin's ball. *J. Geom. Mech.* **1**, (2009), 35–53.

[14] Ibort, A., de León, M., Marrero, J.,C. and Martín de Diego, D., Dirac Brackets in Constrained Dynamics, *Fortschr. Phys.* **47**, (1999), 459–492.

[15] Kozlov V. V., Invariant measures of the Euler-Poincaré equations on Lie algebras *Funkt. Anal. Prilozh.* **22** 69–70 (Russian); English trans.: *Funct. Anal. Appl.* **22**, (1988), 58–59.

[16] Weber R. W., Hamiltonian Systems with Constraints and their Meaning in Mechanics *Arch. Rat. Mech. Anal.*, **91**, (1986), 309–335.

[17] Zenkov DV and Bloch AM, Invariant measures of nonholonomic flows with internal degrees of freedom *Nonlinearity* **16**, (2003), 1793–1807.

L.C. García-Naranjo: Departamento de Matemáticas y Mecánica, IIMAS-UNAM, Apdo. Postal: 20-726, Mexico City, 01000, Mexico.

E-mail address: luis@mym.iimas.unam.mx