Francisco Gómez H.

Introductory notes on

Mechanics and Symmetry

Notes from a mathematical approach

Contents

1 Hamiltonian Systems on Linear Symplectic Spaces **5**

1.1 Introduction 5

1.2 Symplectic and Poisson structures 7

Chapter 1

Hamiltonian Systems on Linear Symplectic Spaces

1.1 Introduction

Newton's second law for a particle moving in Euclidean three-dimensional space $\mathbb{R}^3$, under the influence of a *potential energy* $V(q)$, is

$$F = ma,$$

where $q \in \mathbb{R}^3$, $F(q) = -\nabla V(q)$ is the *force*, m is the mass of the particle, and

$$a = \frac{d^2 q}{dt^2} = \ddot{q}$$

is the acceleration (assuming we start in a postulated privileged coordinate frame called an *inertial frame*. The potential energy V is introduced through the notion of work and the assumption that the force field is conservative. The introduction of the *kinetic energy*

$$K = \frac{1}{2} m \left\| \dot{q} \right\|^2$$

is through the *power*, or *rate of work equation*:

$$\frac{dK}{dt} = m\langle \dot{q}, \ddot{q} \rangle = \langle \dot{q}, F \rangle,$$

where $\langle \cdot, \cdot \rangle$ denotes the innter product on $\mathbb{R}^3$.

The *Lagrangian* associated to this system is defined by

$$L(q^i, \dot{q}^i) = \frac{m}{2} \|\dot{q}\|^2 - V(q)$$

and one can check that the Newton's second law is equivalent to the *Euler-Lagrange equations*:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = 0,$$

In fact, it follows from the definition of the Lagrangian that

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = m\ddot{q}^i - \frac{\partial V}{\partial q^i} = (ma + \nabla V)_i.$$

The Euler-Lagrange equation is a second-order differential equation in q. Those equations are worthy of independent study for a general L, they are the equations for stationary values of the *action integral*:

$$\delta \int_{t_1}^{t_2} L(q^i, \dot{q}^i) dt = 0,$$

as will be detailed later. These *varitational principles* play a fundamental role throughout mechanics (both in particle mechanics and field theory).

A simple computation shows that $dE/dt = 0$, where E is the *total energy*

$$E = \frac{1}{2}m\|\dot{q}\|^2 + V(q).$$

Lagrange and Hamilton observed that it is convenient to introduce the momentum $p_i = m\dot{q}^i$ and rewrite E as a function of p_i and q^i by letting

$$H(p, q) = \frac{\|p\|^2}{2m} + V(q).$$

Newton's second law is equivalent to *Hamilton's canonical equations*

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

which is a first-order differential equations system in (p, q)-space, of *phase space.* In fact,

$$\dot{q}^i - \frac{\partial H}{\partial p_i} = \dot{q}^i - \frac{p_i}{m} = \dot{q}^i - \frac{m\dot{q}^i}{m}, \qquad \dot{p}_i + \frac{\partial H}{\partial q^i} = \frac{d}{dt}(m\dot{q}^i) + \frac{\partial V}{\partial q^i} = m\ddot{q}^i + \frac{\partial V}{\partial q^i}.$$

The first equation is equivalent to the definition of momentum and the second one to the Newton's second law. For a deeper understanding of Hamilton's equations we recall some matrix notation.

Let E be a real linear space and E^* its dual space. Let $e_1, ..., e_n$ be a basis of E with the associated dual basis for E^* denoted by $e^1, ..., e^n$. That is, e^i is defined by the equations

$$\langle e^i, e_j \rangle = e^i(e_j) = \delta_{ij},$$

where δ_{ij} is the Kronocker delta. Throughout these notes we will use Einstein summation convention. That is, when an index variable appears twice in a single term, it implies summation over all the values of the index. In other words, if there is no risk of confusion, summation symbol (Σ) will be omitted. For example, a vector $v \in E$ can be written as $v = v^i e_i$ and a covector $\alpha \in E^*$ as $\alpha = \alpha_i e^i$, where v^i and α_i are the*components* of v and α, respectively.

Let E and F be linear spaces. If $A : E \to F$ is a linear transformation, its *matrix* relative to bases $e_1, ..., e_n$ of E and $f_1, ..., f_m$ of F is denoted by A_i^j, and is defined by

$$A(e_i) = A_i^j f_j.$$

Thus, the columns of the matrix A are $A(e_1), ..., A(e_n)$. If $B : E \times F \to \mathbb{R}$ is a bilinear form, its matrix B_{ij} is defined by $B_{ij} = B(e_i, f_j)$. Define the *associated* linear map $B^\flat : E \to F^*$ by

$$B^\flat(v)(w) = B(v, w)$$

for all $v \in E$ and $w \in F$, and observe that $B^\flat(e_i) = B_{ij} f^j$. Since $B^\flat(e_i)$ is the i-th column of the matrix representing the linear map $B^\flat$, it follows that the matrix of $B^\flat$ in the bases $e_1, ..., e_n$ and $f^1, ..., f^m$ is

$$[B^\flat]_{ij} = B_{ji},$$

that is, it is the transpose of the matrix B_{ij}.

Let Z denote the linear space of pairs (p, q) and write $z = (p, q)$. Let the coordinates q^i, p_i be collectively denoted by z^I, with $I = 1, 2, ..., 2n$. One reason for the notation z is that if one thinks of z as a "complex variable" $z = q + ip$, then Hamilton's equations can be written in a simpler way. Let

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial q} - i\frac{\partial}{\partial p}\right).$$

Then, it follows from the previous notation that

$$2\frac{\partial H}{\partial \bar{z}} = \frac{\partial H}{\partial q} - i\frac{\partial H}{\partial p} = -\dot{p} + i\dot{q} = i(\dot{q} + i\dot{p}) = iz.$$

Then, the complex version of Hamilton's equations can be simply written as $\dot{z} = -2i\partial H/\partial \bar{z}$.

Note that the linear space Z inherites an inner product from $\mathbb{R}^{2n}$. Recall that, if $f : Z \to \mathbb{R}$ is a differentiable function, then $df_z : T_zZ \to \mathbb{R}$ is the linear map defined by

$$df_z(v) = \frac{d}{dt}f(z + tv)\bigg|_{t=0},$$

where the tangent vector $v \in T_zZ$ can be identified as an element of the linear space Z. The gradient of f at z, denoted by $\nabla f(z)$, is the unique tangent vector in T_zZ satisfying

$$df_z v = \langle \nabla f(z), v \rangle.$$

In other words, $\nabla f(z)$ is the *Riesz representation* of df_z under the inner product. In local coordinates,

$$df_z = \frac{\partial f}{\partial q^i}dq^i + \frac{\partial f}{\partial p_i}dp_i,$$

where $z = (q, p) \in Z$. Similarly, the gradient of f at z can be written as

$$\nabla f(z) = \frac{\partial f}{\partial q^i}\frac{\partial}{\partial q^i} + \frac{\partial f}{\partial p_i}\frac{\partial}{\partial p_i}.$$

One can identify vectors and covectors via the standard Euclidean inner product. Thus, df_z can be regarded as a row vector with entries $\partial f/\partial z_i$ and $\nabla f(z)$ as a column vector with the same entries.

1.2 Symplectic and Poisson structures

We can view Hamilton's equations as follows. Consider the operation

$$\nabla H(z) = \left(\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p}\right) \mapsto \left(\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q}\right) := X_H(z),$$

which forms a vector field X_H. This vector field is called the *Hamiltonian vector field*, from the differential of H, and is defined as the composition of certain linear map

$$R : T_zZ \to T_zZ$$

and the gradient $\nabla H(z)$ of H at z. In fact, the matrix of R with respect to the coordinates z^I is

$$[R] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \mathbb{J},$$

where we write $\mathbb{J}$ for that specific matriz which is sometimes called the *symplectic matrix*. Thus,

$$X_H(z) = R \cdot \nabla H(z),$$

or, if the components of X_H are denoted X^I, with $I = 1, ..., 2n$,

$$X^I = R^{IJ}\frac{\partial H}{\partial z^J}.$$

Alternatively, the equation above can be written as $X_H = \mathbb{J}\nabla H$, where ∇H is the gradient of H.

Note: Let $\mathbb{J} \in M_{2n\times 2n}(\mathbb{F})$ be the symplectic matrix. If $v \in \mathbb{F}^n$, then

$$v^\intercal \mathbb{J} = (\mathbb{J}v)^\intercal.$$

Previously, it was established a natural identification of 1-forms and tangent vectors. To be more precise, using the natural coordinates, df_z and $\nabla f(z)$ can be regarded as vectors with the same entries. Thus, the map R can be re-interpreted as a map $R : Z^* \to Z$ as follows

$$R(a_i dq^i + b_i dp_i) = b_i dq^i - a_i dp_i,$$

where $a_i, b_i \in \mathbb{R}$ and the covectors dp^i and dp_i represent the dual basis associated with the coordinates (q, p). Let $B(\alpha, \beta) = \langle \alpha, R(\beta)\rangle$ be the bilinear form associated to R, where $\langle\cdot,\cdot\rangle$ denotes the canonical pairing between Z^* and Z. One calls either the bilinear form B or its associated bilinear form R, the *Poisson structure*. The classical *Poisson bracket* is defined by

$$\{F, G\} = B(dF, dG) = (\nabla F)^\intercal \mathbb{J}\nabla G = \nabla F \cdot \mathbb{J}\nabla G,$$

where F and G are smooth functions from the linear space Z to $\mathbb{R}$. Recall that,

$$\nabla(f \cdot g)(z) = (Df_z)^\intercal g + (Dg_z)^\intercal f,$$

where $f, g : \mathbb{R}^n \to \mathbb{R}^k$ are differentiable functions. Moreover, Df_z and Dg_z are the $k \times n$ matrices of partial derivatives. Since $D(\nabla F) = \mathrm{Hess}(F)$, the Hessian of the function $F : Z \to \mathbb{R}$, it follows that, if $F, G : Z \to \mathbb{R}$ are differentiable functions and A is a constant matrix, then

$$\nabla(\nabla F \cdot A\nabla G) = \nabla(\nabla F \cdot \nabla(AG)) = \mathrm{Hess}(F)\nabla(AG) + \mathrm{Hess}(AG)\nabla F.$$

Siince Hessian matrix is symmetry, there is no transposes on the left hand side. Now, note that,

$$\begin{aligned}
\{F, \{G, H\}\} &= (\nabla F)^\intercal \mathbb{J}\nabla(\nabla G \cdot \mathbb{J}\nabla H) \\
&= (\nabla F)^\intercal \mathbb{J}(\mathrm{Hess}(G)\mathbb{J}\nabla H + (\mathbb{J}\,\mathrm{Hess}(H))^\intercal\nabla G), \\
&= (\nabla F)^\intercal \mathbb{J}(\mathrm{Hess}(G)\mathbb{J}\nabla H + \mathrm{Hess}(H)\mathbb{J}^\intercal\nabla G), \\
&= (\mathbb{J}\nabla F)^\intercal\,\mathrm{Hess}(G)(\mathbb{J}\nabla H) - (\mathbb{J}\nabla F)^\intercal\,\mathrm{Hess}(H)(\mathbb{J}\nabla G).
\end{aligned}$$

Similarly, one can easily find the formulas corresponding to rotate F, G and H:

$$\begin{aligned}
\{G, \{H, F\}\} &= (\mathbb{J}\nabla G)^\intercal\,\mathrm{Hess}(H)(\mathbb{J}\nabla F) - (\mathbb{J}\nabla G)^\intercal\,\mathrm{Hess}(F)(\mathbb{J}\nabla H), \\
\{H, \{F, G\}\} &= (\mathbb{J}\nabla H)^\intercal\,\mathrm{Hess}(F)(\mathbb{J}\nabla G) - (\mathbb{J}\nabla H)^\intercal\,\mathrm{Hess}(G)(\mathbb{J}\nabla F).
\end{aligned}$$

Finally, since each term $(\mathbb{J}\nabla F)^\intercal\,\mathrm{Hess}(G)(\mathbb{J}\nabla H)$ is one-dimensional, it follows that

$$[(\mathbb{J}\nabla F)^\intercal\,\mathrm{Hess}(G)(\mathbb{J}\nabla H)]^\intercal = (\mathbb{J}\nabla H)^\intercal\,\mathrm{Hess}(G)(\mathbb{J}\nabla F).$$

Thus, if F, G and H are differentiable functions from Z to $\mathbb{R}$, then

$$\{F, \{G, H\}\} + \{G, \{H, F\}\} + \{H, \{F, G\}\} = 0.$$

The above property is called the *Poisson identity*.

The *symplectic structure* Ω is the bilinear form associated to

$$R^{-1} : Z \to Z^*,$$

that is, $\Omega(v, w) = \langle R^{-1}v, w \rangle$, or equivalently, $\Omega^\flat = R^{-1}$. The matrix of Ω is $\mathbb{J}$, in the sense that

$$\Omega(v, w) = v^\mathsf{T} \mathbb{J} w.$$

Summarizing, there are four important maps involving the symplectic matrix:

- The symplectic form $\Omega : Z \times Z \to \mathbb{R}$, with matrix $\mathbb{J}$.
- The associated linear map $\Omega^\flat : Z \to Z^*$, with matrix $\mathbb{J}^\mathsf{T}$.
- The inverse map $\Omega^\sharp = (\Omega^\flat)^{-1} = R : Z^* \to Z$, with matrix $\mathbb{J}$.
- The Poisson form, $B : Z^* \times Z^* \to \mathbb{R}$, with matrix $\mathbb{J}$.

Thus, Hamilton's equations may be written as follows:

$$\dot{z} = X_H(z) = \Omega^\sharp dH(z),$$

under the identification between $dH(z)$ and $\nabla H(z)$. By multiplying both sides by $\Omega^\flat$, we get

$$\Omega^\flat X_H(z) = dH(z).$$

Thus, in terms of the symplectic form, one can write

$$\Omega(X_H(z), v) = dH(z) \cdot v$$

for all $z, v \in Z$. Problems such as rigid body dynamics, quantum mechanics as Hamiltonian system, and the motion of a particle in a rotating reference frace motivate the need to generalize these concepts.