# EULER-LAGRANGE EQUATIONS

EUGENE LERMAN

Contents

1. Classical system of $N$ particles in $\mathbb{R}^3$ 1
2. Variational formulation 2
3. Constrained systems and d'Alembert principle. 4
4. Legendre transform 6

## 1. Classical system of $N$ particles in $\mathbb{R}^3$

.

Consider a mechanical system consisting of $N$ particles in $\mathbb{R}^3$ subject to some forces. Let $x_i \in \mathbb{R}^3$ denote the position vector of the $i$th particle. Then all possible positions of the system are described by $N$-tuples $(x_1, \ldots, x_N) \in (\mathbb{R}^3)^N$. The space $(\mathbb{R}^3)^N$ is an example of a **configuration space**. The time evolution of the system is described by a curve $(x_1(t), \ldots, x_N(t))$ in $(\mathbb{R}^3)^N$ and is governed by Newton's second law:

$$m_i \frac{d^2 x_i}{dt^2} = F_i(x_1, \ldots, x_N, \dot{x}_1, \ldots, \dot{x}_N, t)$$

Here

- $F_i$ denotes the force on $i$th particle (which depends on the positions and velocities of all $N$ particles and on time),
- $\dot{x_i} = \frac{dx_i}{dt}$, and
- $m_i$ denotes the mass of the $i$th particle.

We now re-lable the variables. Let $q_{3i}, q_{3i+1}, q_{3i+2}$ be respectively the first, the second and the third coordinate of the vector $x_i$, $i = 1, \ldots, N$:

$$(q_{3i}, q_{3i+1}, q_{3i+2}) = x_i.$$

The configuration space of our system is then $\mathbb{R}^n$, where now $n = 3N$. The equations of motion take the form

$$m_\alpha \frac{d^2 q_\alpha}{dt^2} = F_\alpha(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t), \quad 1 \leq \alpha \leq n. \tag{1.1}$$

We now suppose that the forces are time-independent and conservative/ That is, we assume that there exists a function $V : \mathbb{R}^n \longrightarrow \mathbb{R}$ (a potential) such that

$$F_\alpha(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t) = F_\alpha(q_1, \ldots, q_n) = -\frac{\partial V}{\partial q_\alpha}(q_1, \ldots, q_n).$$

**Example 1.1.** For example if $N$ particles interact by gravitational attraction, then the potential is

$$V(x_1, \ldots, x_N) = -\gamma \sum_{i \neq j} \frac{m_i m_j}{\|x_i - x_j\|},$$

where $\gamma$ is a universal constant.

---

typeset February 26, 2013.

Under the assumption of time independent conservative forces the system of equations (1.1) takes the form

$$m_\alpha \frac{d^2 q_\alpha}{dt^2} = -\frac{\partial V}{\partial q_\alpha}(q_1, \ldots, q_n), \quad 1 \le \alpha \le n. \tag{1.2}$$

We rewrite equation (1.2) as a first order ODE by doubling the number of variables; this is a standard trick. Call the new variables, the velocities, $v_\alpha$:

$$\begin{cases} m_\alpha \frac{dv_\alpha}{dt} = -\frac{\partial V}{\partial q_\alpha}(q_1, \ldots, q_n) \\ \frac{dq_\alpha}{dt} = v_\alpha \end{cases} \tag{1.3}$$

A solution $(q(t), v(t))$ of the above system of equations is a curve in the tangent bundle $T\mathbb{R}^n$ with $\frac{d}{dt}q(t) = v(t)$. The tangent bundle $T\mathbb{R}^n$ is an example of a phase space.

It is standard to introduce a function

$$L(q, v) = \frac{1}{2}\sum_\alpha m_\alpha v_\alpha^2 - V(q)$$

on the phase space $T\mathbb{R}^n$. The function is called the Lagrangian of the system. It is the difference of the kinetic and the potential energies. We will see shortly that we can re-write (1.3) as

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_\alpha}\right) - \frac{\partial L}{\partial q_\alpha} = 0, \quad 1 \le \alpha \le n \tag{1.4}$$

The system of equations (1.4), which we will see are equivalent to Newton's law of motion, is an example of the Euler-Lagrange equations.

We now check that equations (1.3) and (1.4) are, indeed, the same. Since

$$\frac{\partial}{\partial v_\alpha}\left(\frac{1}{2}\sum_\beta m_\beta v_\beta^2\right) = m_\alpha v_\alpha$$

and

$$\frac{\partial}{\partial q_\alpha}\left(\frac{1}{2}\sum m_\beta v_\beta^2 - V\right) = -\frac{\partial V}{\partial q_\alpha} = F_\alpha$$

we get

$$0 = \frac{d}{dt}\left(\frac{\partial}{\partial v_\alpha}L\right) - \frac{\partial L}{\partial q_\alpha} = \frac{d}{dt}(m_\alpha v_\alpha) + \frac{\partial V}{\partial q_\alpha}.$$

Hence

$$m_\alpha \frac{dv_\alpha}{dt} = -\frac{\partial V}{\partial q_\alpha}.$$

So far introducing the Lagrangian did not give us anything new. We now show that it does indeed allow us to look at Newton's law from another point of view, and that the new point of view has interesting consequences.

## 2. Variational formulation

Let $L : T\mathbb{R}^n \to \mathbb{R}$ be a Lagrangian (i.e., a smooth function). Let $q^{(0)}, q^{(1)}$ be two points in $\mathbb{R}^n$. Consider the collection $\mathcal{P} = \mathcal{P}(q^{(0)}, q^{(1)})$ of all possible twice continuously differentiable ($C^2$) paths $\gamma : [a, b] \longrightarrow \mathbb{R}^n$ with $\gamma(a) = q^{(0)}, \gamma(b) = q^{(1)}$. That is, set

$$\mathcal{P} := \{\gamma : [a, b] \longrightarrow \mathbb{R}^n \mid \gamma(a) = q^{(0)}, \gamma(b) = q^{(1)}\}.$$

The Lagrangian $L$ defines a map

$$A_L : \mathcal{P} \to \mathbb{R}, \quad A_L(\gamma) := \int_a^b L(\gamma(t), \dot\gamma(t))\, dt,$$

called an action.

We are now in position to state:

**Hamilton's principle:** physical trajectories between two points $q^{(0)}, q^{(1)}$ of the system governed by the Lagrangian $L$ are critical points of the action functional $A_L : \mathcal{P}(q^{(0)}, q^{(1)}) \to \mathbb{R}$.

**Proposition 2.1.** *Hamilton's principle implies Euler-Lagrange equations and hence Newton's law of motion.*

*Proof.* The proof is well-known. Let $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ be a critical point (path, trajectory) of an action functional $A_L$; let $\gamma(t,\epsilon)$ be a family of paths depending on $\epsilon \in \mathbb{R}$ such that $\gamma(t,0) = \gamma(t)$ and such that for all $\epsilon$ we have $\gamma(a,\epsilon) = \gamma(a)$ and $\gamma(b,\epsilon) = \gamma(b)$ (i.e., we fix the end points).

Let $y(t) = \frac{\partial}{\partial \epsilon}\big|_{\epsilon=0}\gamma(t,\epsilon)$. Note that $y(a) = 0$ and $y(b) = 0$ since the end points are fixed. Also

$$\frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0}\left(\frac{\partial}{\partial t}\gamma(t,\epsilon)\right) = \frac{\partial}{\partial t}\left(\frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0}\gamma(t,\epsilon)\right) = \dot{y}(t).$$

Conversely, given a curve $y : [a,b] \to \mathbb{R}^n$ with $y(a) = y(b) = 0$, we can find a family of paths $\gamma(t,\epsilon)$ with fixed end points such that

$$\gamma(t,0) = \gamma(t) \quad \text{and} \quad \frac{\partial}{\partial \epsilon}\Big|_{\epsilon=0}\left(\frac{\partial}{\partial t}\gamma(t,\epsilon)\right) = \dot{y}(t).$$

For example, we may take

$$\gamma(t,\epsilon) = \gamma(t) + \epsilon y(t).$$

Now

$$\begin{aligned}
0 &= \frac{d}{d\epsilon}\bigg|_{\epsilon=0} A_L(\gamma(t,\epsilon)) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \int_a^b L(\gamma(t,\epsilon), \dot{\gamma}(t,\epsilon))\, dt \\
&= \int_a^b \frac{d}{d\epsilon}\Big|_0 L\, dt = \sum_\alpha \int_a^b \left(\frac{\partial L}{\partial q_\alpha}\frac{\partial q_\alpha}{\partial \epsilon}\Big|_{\epsilon=0} + \frac{\partial L}{\partial v_\alpha}\frac{\partial v_\alpha}{\partial \epsilon}\Big|_{\epsilon=0}\right) dt \\
&= \sum_\alpha \int_a^b \left(\frac{\partial L}{\partial q_\alpha} y_\alpha + \frac{\partial L}{\partial v_\alpha}\dot{y}_\alpha\right) dt \qquad (2.1)\\
&= \sum_\alpha \left\{\int_a^b \frac{\partial L}{\partial q_\alpha} y_\alpha\, dt + \frac{\partial L}{\partial v_\alpha} y_\alpha\Big|_a^b - \int_a^b \frac{d}{dt}\left(\frac{\partial L}{\partial v_\alpha}\right) y_\alpha\, dt\right\} \quad \text{(integration by parts)} \\
&= \sum_\alpha \int \left(\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt}\left(\frac{\partial L}{\partial v_\alpha}\right)\right) y_\alpha\, dt \quad (\text{since } y_\alpha(a) = y_\alpha(b) = 0) \text{ for all } \alpha).
\end{aligned}$$

We now recall without proof:

**Lemma 2.2.** *If $f : [a,b] \to \mathbb{R}$ is continuously differentiable function with the property that for any continuously differentiable function $y : [a,b] \to \mathbb{R}$ with $y(a) = y(b) = 0$, we have*

$$\int_a^b f(t) y(t)\, dt = 0,$$

*then $f$ has to be identically zero.*

We conclude that for any index $\alpha$ we must have

$$\frac{\partial L}{\partial q_\alpha} - \frac{d}{dt}\left(\frac{\partial L}{\partial v_\alpha}\right) = 0.$$

That is to say, Hamilton's principle implies the Euler-Lagrange equations. ☐

Note that we have proved that given a Lagrangian there is a vector field on $T\mathbb{R}^n$ whose integral curves are the critical curves of the corresponding action.

## 3. Constrained systems and d'Alembert principle.

We start by listing examples of constrained systems (all constraints are time-independent, that is, scleronomic).

**Example 3.1** (Spherical pendulum)**.** The system consists of a massive particle in $\mathbb{R}^3$ connected by a very light rod of length $\ell$ to a universal joint. The configuration space of this system is a sphere $S^2$ of radius $\ell$. The phase space is $TS^2$. This is an example of a holonomic constraint.

**Example 3.2** (Free rigid body)**.** The system consists of $N$ point masses in $\mathbb{R}^3$ maintaining fixed distances between each other:

$$\|x_i - x_j\| = \text{const}_{ij}.$$

We will see later that the configuration space is $E(3)$, the Euclidean group of distance preserving transformations of $\mathbb{R}^3$. It is not hard to show that $E(3)$ consists of rotations and translations. In fact we can represent $E(3)$ as a certain collection of matrices:

$$E(3) \simeq \{\begin{pmatrix} A & v \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{4^2} \mid A^T A = I, \det A = \pm 1, v \in \mathbb{R}^3\}.$$

The group $E(3)$ is a 6-dimensional manifold, hence the phase space for a free rigid body is $TE(3)$. The free rigid body is also an example of a system with a holonomic constraint.

**Example 3.3** (A quater rolling on a rough plane in an upright position (without slipping))**.** The configuration space is $(\mathbb{R}^2 \times S^1) \times S^1$, where the elements of $\mathbb{R}^2$ keep track of the point of contact of the quarter with the rough plane, the points in the first $S^1$ keeps track of the orientation of the plane of the quarter and points of the second $S^1$ keep track of the orientation of the design on the quarter. The phase space of the system is smaller than

$$T(\mathbb{R}^2 \times S^1 \times S^1)$$

because the point of contact of the quarter with the plane has to be stationary. This is an example of non-holonomic constraints, since the constraints on position do not determine the constraints on velocity: the roll-no-slip condition is extra.

This leads us to a definition.

**Definition 3.4.** Time-independent constraints are **holonomic** if the constraint on possible velocities are determined by the constraints on the configurations of the system. In other words if the constraints confine the configurations of the system to a submanifold $M$ of $\mathbb{R}^n$ and the corresponding phase space is $TM$, then the constraints are holonomic.

We will study only holonomic systems with an added assumption: constraint forces do no work.

We are now in position to formulate:

> **d'Alembert's principle:** If constraint forces do no work, then the true physical trajectory of the system are extremals of the action functional of the free system restricted to the paths lying in the constraint submanifold.

This principle is very powerful: we no longer need to know anything about the constraining forces except for the fact that they limit the possible configurations to a constraint submanifold. We now investigate the equations of motion that d'Alembert's principle produces.

Let $M \subseteq \mathbb{R}^n$ be a submanifold and let

$$L : T\mathbb{R}^n \supseteq TM \longrightarrow \mathbb{R}$$

be a Lagrangian for an unconstrained( "free") system. By d'Alembert's principle our system evolves along a path

$$\gamma : [a, b] \longrightarrow M$$

such that $\gamma$ is critical for

$$A_L : \{\sigma : [a,b] \longrightarrow M \mid \sigma(a) = q^{(0)}, \sigma(b) = q^{(1)}\} \to \mathbb{R}$$
$$A_L(\sigma) = \int_a^b L(\sigma, \dot{\sigma})\, dt$$

Suppose the end points $q^{(0)}$ and $q^{(1)}$ lie in some coordinate patch on $M$. Let $(q_1, \ldots, q_n)$ be the coordinates on the patch and let $(q_1, \ldots, q_n, v_1, \ldots, v_n)$ be the corresponding coordinates on the corresponding patch in $TM$. The same argument as before (cf. Proposition 2.1) gives us Euler-Lagrange equations

$$\frac{d}{dt}(\frac{\partial L}{\partial v_\alpha}) - \frac{\partial L}{\partial q_\alpha} = 0!$$

Note that these equations represent a vector field on a coordinate patch in the tangent bundle $TM$.

**Example 3.5** (Planar pendulum)**.** The system consists of a heavy particle in $\mathbb{R}^3$connected by a very light rod of length $\ell$ to a fixed point. Unlike the spherical pendulum the rod is only allowed to pivot in a fixed vertical plane. The configuration space $M$ is the circle $S^1 \subset \mathbb{R}^2$ of radius $\ell$. The Lagrangian of the free system is

$$L(x, v) = \frac{1}{2}m(v_1^2 + v_2^2) - mgx_2,$$

where

- $m$ is the mass of the particle,
- $x_1, x_2$ are coordinates on $\mathbb{R}^2$,
- $x_1, x_2, v_1, v_2$ are the corresponding coordinates on $T\mathbb{R}^2$ and
- $g$ is the gravitational acceleration (9.8 $m/s^2$).

We compute the equations of motion for the constraint system as follows Consider the embedding

$$S^1 \longrightarrow \mathbb{R}^2, \quad \varphi \mapsto (\ell \sin\varphi, -\ell\cos\varphi).$$

The corresponding embedding

$$TS^1 \longrightarrow T\mathbb{R}^2$$

is given by

$$(\varphi, v_\varphi) \mapsto (\ell\sin\varphi, -\ell\cos\varphi, \ell\cos\varphi\, v_\varphi, \ell\sin\varphi\, v_\varphi).$$

Therefore the constraint Lagrangian is given by

$$L(\varphi, v_\varphi) = \frac{1}{2}m(\ell^2\cos^2\varphi v_\varphi^2 + \ell^2\sin^2\varphi v_\varphi^2) + mg\ell\cos\varphi = \frac{1}{2}m\ell^2 v_\varphi^2 + mg\ell\cos\varphi.$$

The corresponding Euler-Lagrange equation is

$$\frac{d}{dt}(\frac{\partial L}{\partial v_\varphi}) - \frac{\partial L}{\partial \varphi} = 0,$$

i.e.,

$$m\ell^2\frac{dv_\varphi}{dt} + mg\ell\sin\varphi = 0.$$

Therefore

$$\ddot{\varphi} = -\frac{g}{\ell}\sin\varphi$$

is the equation of motion.

## 4. Legendre transform

We start with a brief digression which will serve as a toy example of a Legendre transform. Let

$$L : V \to \mathbb{R}$$

be a smooth function on a vector space $V$ and let

$$v_1, \dots v_n : V \to \mathbb{R}$$

be coordinates on $V$. For example, $\{v_i\}_{i=1}^n$ could be linear functionals on $V$ that form a basis of the dual space $V^*$. For each $v \in V$ consider the matrix

$$(\frac{\partial^2 L}{\partial v_i \partial v_j}(v))$$

of second order partials. It can be interpreted as a quadratic form $d^2L(v)$ on $V$ as follows: for $u, w \in V$ with coordinates $(u_1, \dots, u_n)$ and $(w_1, \dots, w_n)$ respectively we set

$$d^2L(v)(u,w) = \sum_{i,j} \frac{\partial^2 L}{\partial v_i \partial v_j}(v)\ u_i w_j \qquad \left(= u^T \left(\frac{\partial^2 L}{\partial v_i \partial v_j}(v)\right) w.\right)$$

The quadratic form $d^2L(v)$ also has a coordinate-free definition. By the chain rule, for any $u, w \in V$,

$$d^2L(v)(u,w) = \left.\frac{\partial^2}{\partial s \partial t} L(v + su + tw)\right|_{(0,0)}.$$

We note that the matrix $(\frac{\partial^2 L}{\partial v_i \partial v_j}(v))$ is invertible if and only if the quadratic form $d^2L(v)$ is nondegenerate. This ends a digression and we go back to studying Lagrangians on phase spaces.

Recall that given a Lagrangian

$$L : TM \to \mathbb{R}$$

and two points $m_1, m_2 \in M$, the corresponding action

$$A_L : \{\ C^1 \text{ paths connecting } m_1 \text{ to } m_2\} \to \mathbb{R}$$

is defined by

$$A_L(\sigma) = \int_a^b L(\sigma(t), \dot\sigma(t))\, dt.$$

Suppose that the points $m_1$, $m_2$ lie in a coordinate patch $U$ with coordinates $x_1, \dots, x_n : U \to \mathbb{R}$. Let $x_1, \dots, x_n, v_1, \dots, v_n$ be the corresponding coordinates on $TU \subset TM$. We saw that a path $\gamma : [a,b] \to U$, $\gamma(a) = m_1$, $\gamma(b) = m_2$, is critical for the action $A_L$ if and only if the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v_i}(\gamma(t), \dot\gamma(t))\right) - \frac{\partial L}{\partial x_i}(\gamma, \dot\gamma) = 0, \quad 1 \le i \le n$$

hold. Now

$$\frac{d}{dt}(\frac{\partial L}{\partial v_i}(\gamma, \dot\gamma)) = \sum_j (\frac{\partial^2 L}{\partial x_j \partial v_i}\dot\gamma_j + \frac{\partial^2 L}{\partial v_j \partial v_i}\ddot\gamma_j).$$

Hence the Euler-Lagrange equations read

$$\sum_j \frac{\partial^2 L}{\partial v_i \partial v_j}\ddot\gamma_j = \frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i}\dot\gamma_j \quad 1 \le i \le n.$$

We now make an important assumption: $L$ is a **regular** Lagrangian. That is. we assume that the matrix

$$(\frac{\partial^2 L}{\partial v_i \partial v_j}(x,v))$$

is invertible for all $(x, v) \in TU$. Equivalently we assume that for all $x \in U$ the form

$$\left. d^2 L\right|_{T_x M}(v)$$

is nondegenerate for all $v \in T_xM$. Then there exists an inverse matrix

$$(M_{ki}) = (M_{ki}(x,v)) = \left(\frac{\partial^2 L}{\partial v_i \partial v_j}(x,v)\right)^{-1}$$

It is defined by

$$\sum_i M_{ki} \frac{\partial^2 L}{\partial v_i \partial v_j} = \delta_{kj},$$

where, as usual, $\delta_{kj}$ is the Kronecker delta function. Under this assumption

$$\sum_{i,j} \underbrace{M_{ki} \frac{\partial^2 L}{\partial v_i \partial v_j}}_{\delta_{kj}} \ddot{\gamma}_j = \sum_i M_{ki} \left(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i} \dot{\gamma}_j\right),$$

hence

$$\ddot{\gamma}_k = \sum_i M_{ki} \left(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i} \dot{\gamma}_j\right) \tag{4.1}$$

**Exercise 4.1.** Let $g$ be a Riemannian metric on $M$ and let $L(x,v) = \frac{1}{2} g(x)(v,v)$. Check that this Lagrangian is regular. What does (4.1) look like for this $L$?

We can rewrite (4.1) as first order system in $2n$ variables.

$$\begin{aligned} \dot{x_j} &= v_j \\ (\ddot{x}_k =)\dot{v}_k &= \sum_i M_{ki} \left(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i} v_j\right) \end{aligned} \tag{4.2}$$

Note that in physics literature the coordinates $x_i$'s are usually called $q_i$'s and the corresponding coordinates $v_i$'s are usually called $\dot{q}_i$'s. The confusing point here is that the dot above $q_i$ does not stand for anything; $\dot{q}_i$ is simply a name of a coordinate.

Equation (4.2) means that we have a vector field $X_L$ on $TU$:

$$X_L(x,v) = \sum_j v_j \frac{\partial}{\partial x_j} + \sum_{k,i} M_{ki} \left(\frac{\partial L}{\partial x_i} - \sum_j \frac{\partial^2 L}{\partial x_j \partial v_i} v_j\right) \frac{\partial}{\partial v_k}$$

The vector field $X_L$ is called the **Euler-Lagrange vector field**. One can show that

**Proposition 4.1.** *$X_L$ is a well-defined vector field on the tangent bundle $TM$, i.e. it transforms correctly under the change of variables.*