Variational Calculus - Week 2 - The Euler Lagrange Equations

Antonio León Villares

September 2022

Contents

Based on the notes by Jelle Hartong, Chapter 3

1 The Euler-Lagrange Equations (with endpoint-fixed variations)

1.1 The Action Functional

- **What is the Lagrangian?**
  - a function of the form.
  $$L : \mathbb{R}^{2n+1} \to \mathbb{R}$$
  such that:
  $$(x, v, t) \mapsto L(x, v, t)$$
  where:
  $$x, v \in \mathbb{R}^n \quad t \in \mathbb{R}$$
  - in practice, we can think of x as a position, $v = \dot{x}(t)$ as the velocity of x, and t as a time parameter, which parametrises $x, \dot{x}$; in this way, the **Lagrangian** is a **composition** of functions:
  $$t \mapsto L(x(t), \dot{x}(t), t)$$
  - the **Lagrangian** generalises to higher order derivatives:
  $$L(x, \dot{x}, \ddot{x}, \ldots, t)$$
- **What is the action?**
  - let $\underline{P}, \underline{Q} \in \mathbb{R}^n$, and let $\mathcal{C}_{\underline{P},\underline{Q}}$ be the space of C^1 (continuously differentiable) curves:
  $$x : [0, 1] \to \mathbb{R}^n$$
  such that:
  $$x(0) = \underline{P} \qquad x(1) = \underline{q}$$
  - the **action** is a **functional**:
  $$I : \mathcal{C}_{\underline{P},\underline{Q}}$$
  involving the **Lagrangian**:
  $$I[x] = \int_0^1 L(x(t), \dot{x}(t), t)dt$$

1.2 The Euler-Lagrange Equation

The Euler-Lagrange equation allows us to derive an ODE which optimises the action.

> *The path $x = (x^1, x^2, \ldots, x^n) \in \mathcal{C}_{\underline{P},\underline{Q}}$ optimising the **action functional**:*
> $$I[x] = \int_0^1 L(x(t), \dot{x}(t), t)dt$$
> *is the solution to the **set of differential equations**:*
> $$\frac{\partial L}{\partial x^i} = \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i}, \qquad \forall i \in [1, n]$$

Proof. Recall, a path x will be a **critical point** of $I[x]$ if for **all** endpoint-fixed variations ε:

$$\frac{d}{ds} \left. I[x+s\varepsilon]\right|_{s=0} = 0$$

Assuming the Lagrangian is well-behaved, we can differentiate under the integral:

$$\frac{d}{ds} I[x+s\varepsilon] = \int_0^1 \frac{d}{ds}(L(x+s\varepsilon, \dot{x}+s\dot{\varepsilon}, t))dt$$

Now, keep in mind that $x+s\varepsilon, \dot{x}+s\dot{\varepsilon}$ actually represent a set of $2n$ variables. Let $q = x+s\varepsilon$. Thus, if we apply the chain rule:

$$\begin{aligned}
\frac{d}{ds} I[x+s\varepsilon] &= \int_0^1 \frac{d}{ds}(L(q,u,t))dt \\
&= \int_0^1 \left(\frac{\partial L}{\partial q^1}\frac{dq^1}{ds} + \ldots + \frac{\partial L}{\partial q^n}\frac{dq^n}{ds} + \frac{\partial L}{\partial \dot{q}^1}\frac{d\dot{q}}{ds} + \ldots + \frac{\partial L}{\partial \dot{q}^n}\frac{d\dot{q}}{ds} \right) dt \\
&= \int_0^1 \sum_{i=1}^{n} \left(\frac{\partial L}{\partial q^i}\varepsilon^i + \frac{\partial L}{\partial \dot{q}^i}\dot{\varepsilon}^i \right) dt \\
&= \sum_{i=1}^{n} \int_0^1 \left(\frac{\partial L}{\partial q^i}\varepsilon^i + \frac{\partial L}{\partial \dot{q}^i}\dot{\varepsilon}^i \right) dt
\end{aligned}$$

But then:

$$q^i(s=0) = x^i$$

so:

$$0 = \frac{d}{ds} \left. I[x+s\varepsilon]\right|_{s=0} = \sum_{i=1}^{n} \int_0^1 \left(\frac{\partial L}{\partial x^i}\varepsilon^i + \frac{\partial L}{\partial \dot{x}^i}\dot{\varepsilon}^i \right) dt$$

We would like to get rid of the $\dot{\varepsilon}^i$. As we did last week we apply integration by parts for the dot product. In particular, notice that:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\varepsilon^i \right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i} \right)\varepsilon^i + \frac{\partial L}{\partial \dot{x}^i}\dot{\varepsilon}^i$$

so we have that:

$$\begin{aligned}
0 &= \sum_{i=1}^{n} \int_0^1 \left(\frac{\partial L}{\partial x^i}\varepsilon^i + \frac{\partial L}{\partial \dot{x}^i}\dot{\varepsilon}^i \right) dt \\
&= \sum_{i=1}^{n} \int_0^1 \left(\frac{\partial L}{\partial x^i}\varepsilon^i + \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\varepsilon^i \right) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i} \right)\varepsilon^i \right) dt \\
&= \sum_{i=1}^{n} \left[\frac{\partial L}{\partial \dot{x}^i}\varepsilon^i \right]_0^1 + \int_0^1 \left(\frac{\partial L}{\partial x^i}\varepsilon^i - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i} \right)\varepsilon^i \right) dt \\
&= \sum_{i=1}^{n} \int_0^1 \left(\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i} \right) \right)\varepsilon^i dt
\end{aligned}$$

where we have used the fact that:

$$\varepsilon^i(0) = \varepsilon^i(1) = 0, \qquad \forall i \in [1,n]$$

But now, by the **Fundamental Lemma of Variational Calculus**, it must be the case that:

$$\frac{\partial L}{\partial x^i} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i} \right), \qquad \forall i \in [1,n]$$

□

If we let the action be the arclength functional, then:

$$L(x, \dot{x}, t) = \|\dot{x}\|$$

so by Euler-Lagrange:

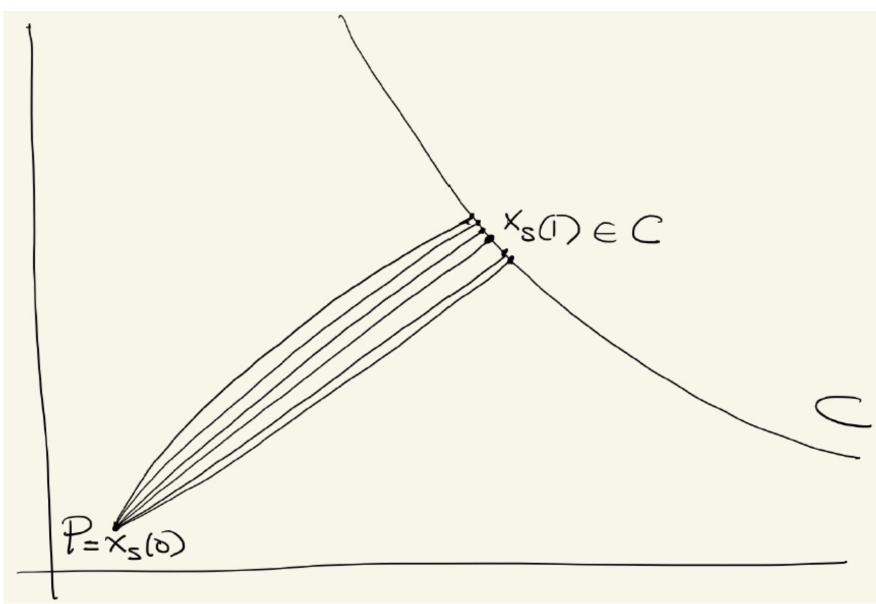
$$\frac{\partial L}{\partial x^i} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) \implies 0 = \frac{d}{dt}\left(\frac{\dot{x}^i}{\|\dot{x}^i\|}\right)$$

as before.

2 The Euler-Lagrange Equations (with variations of free endpoints)

2.1 One Endpoint-Fixed Variation: Geodesic From Point to Curve

- **What is a geodesic from a point to a curve?**
  - we consider the problem of finding the **shortest path** between a point $\underline{P} \in \mathbb{R}^2$ and a curve C defined in $\mathbb{R}^2$
  - C can be described implicitly by:
  $$g : \mathbb{R}^2 \to \mathbb{R}$$
  $$g(\underline{x}) = 0$$
  - the **geodesic** will be a **regular curve** of the form:
  $$x : [0, 1] \to \mathbb{R}$$
  satisfying:
  $$x(0) = \underline{P} \qquad x(1) \in C \iff g(x(1)) = 0$$



- **What condition must variations satisfy at the endpoint 1?**

- recall, we had a **family** of curves:

$$x(s,t) \qquad x(s,0) = \underline{P} \qquad x(s,1) \in C \iff g(x(s,1)) = 0$$

- we define the **variation** as:

$$\varepsilon(t) = \frac{\partial}{\partial s}(x(s,t))\Big|_{s=0}$$

which satisfies:

$$x(s,t) = x(t) + s\varepsilon(t) + sE(s,t)$$

- to know how the variation behaves at the endpoint, we can look at the derivative of $g(x(s,1))$ at $s=0$ (since when $s=0$, the family of curves collapses onto the "ideal" curve)
- thus, since $g(x(s,1)) = 0$ for any s, we must have:

$$\frac{d}{ds}g(x(s,1))\Big|_{s=0} = 0$$

- but recall the definition of the total derivative at a point $\underline{a}$ in the direction $\underline{y}$:

$$Dg(\underline{a})(\underline{y}) = \frac{d}{ds}g(\underline{a} + s\underline{y})$$

so:

$$0 = \frac{d}{ds}g(x(s,1))\Big|_{s=0} = Dg(x(1))(\varepsilon(1))|_{s=0} = \langle \nabla g(x(1)), \varepsilon(1)\rangle$$

- in other words either:
  * $\varepsilon(1) = 0$
  * $\varepsilon(1)$ is perpendicular to ∇g at $x(1)$, so $\varepsilon(1)$ will be **tangent** to C at $x(1)$

- **How does optimising the arclength change by including a variable endpoint?**
  - recall, when deriving the critical point for the arclength, we reached:

$$\frac{d}{ds}\left(S[x+s\varepsilon]\right)\Big|_{s=0} = \left[\left\langle \frac{\dot{x}}{\|\dot{x}\|}, \varepsilon\right\rangle\right]_0^1 - \int_0^1 \left\langle \frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right), \varepsilon\right\rangle) dt$$

  - with fixed endpoints $\varepsilon(0) = \varepsilon(1) = 0$, the term:

$$\left[\left\langle \frac{\dot{x}}{\|\dot{x}\|}, \varepsilon\right\rangle\right]_0^1$$

  disappeared
  - now, all **admissible** variations are either 0 at the endpoints, or **tangent** to C at $x(1)$
  - if $\varepsilon(1) = 0$, then we recover that the solution is:

$$\frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right) = 0$$

  that is, we expect the critical point to be a straight line
  - since $\varepsilon(1) = 0$, we thus have:

$$\frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right) = 0$$

  so our expression gets reduced to:

$$\frac{d}{ds}\left(S[x+s\varepsilon]\right)\Big|_{s=0} = \left[\left\langle \frac{\dot{x}}{\|\dot{x}\|}, \varepsilon\right\rangle\right]_0^1$$

- however, the **admissible** ε must also satisfy $\varepsilon(1)$ being tangent to C so:

$$\left[\left\langle \frac{\dot{x}}{\|\dot{x}\|}, \varepsilon \right\rangle\right]_0^1 = 0 \implies \left\langle \frac{\dot{x}(1)}{\|\dot{x}(1)\|}, \varepsilon(1) \right\rangle = 0$$

  We thus require that the line at $x(1)$ have a velocity which is **orthogonal** to $\varepsilon(1)$; in other words, at $t = 1$, the geodesic x is perpendicular to C

- thus, the geodesic joining a point to a curve will be the orthogonal line between the point and the curve

2.2 Variable Endpoint Variations: Geodesic From Curve to Curve

- **Can the variation at the endpoints also vary?**
  - we can consider the shortest path between 2 non-intersecting curves C_0, C_1
- **What are the admissible variations when the endpoints can vary?**
  - following similar work as before, we get that:
    - $\varepsilon(0) = 0$ or $\varepsilon(0)$ is tangent to C_0
    - $\varepsilon(1) = 0$ or $\varepsilon(1)$ is tangent to C_1
- **How does the critical point for the arclength functional change with variable endpoints?**
  - we still get that x must be a straight line
  - however, we also have the boundary conditions:
    - $x(0) \in C_0$ and $x(0)$ is orthogonal to C_0
    - $x(1) \in C_1$ and $x(1)$ is orthogonal to C_1
- **What happens if the variable endpoints are defined by a hypersurface?**
  - consider 2 hypersurfaces in $\mathbb{R}^n$ (up to now the analysis had just been curves in $\mathbb{R}^2$)
  - the same variational follows, and tells us that the shortest path in $\mathbb{R}^n$ between 2 hypersufaces are **straight** lines which are normal to the hypersurface at their intersection

2.3 The Action Functional for Variable Endpoint Variations

- **How do variable endpoints affect the critical point of the action functional?**
  - consider 2 **regular hypersurfaces** C_0, C_1 defined by:

$$g_0 : \mathbb{R}^n \to \mathbb{R} \qquad g_0(\underline{x}) = 0$$

$$g_1 : \mathbb{R}^n \to \mathbb{R} \qquad g_1(\underline{x}) = 0$$

  - recall, during our analysis for fixed endpoints we obtained:

$$\sum_{i=1}^{n} \left[\frac{\partial L}{\partial \dot{x}^i}\varepsilon^i\right]_0^1 + \int_0^1 \left(\frac{\partial L}{\partial x^i}\varepsilon^i - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\varepsilon^i\right) dt = 0$$

  - with variations which vanish at the endpoints, we get that:

$$\int_0^1 \left(\frac{\partial L}{\partial x^i}\varepsilon^i - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\varepsilon^i\right) dt = 0$$

    so the curve must satisfy the **Euler-Lagrange** equations

- for the remaining **admissible variations** (with $\varepsilon(0)$ tangent to C_0 and $\varepsilon(1)$ tangent to C_1) we thus need to satisfy:

$$\sum_{i=1}^{n}\left[\frac{\partial L}{\partial \dot{x}^i}\varepsilon^i\right]_0^1 = 0 \implies \sum_{i=1}^{n}\frac{\partial L}{\partial \dot{x}^i(1)}\varepsilon^i(1) - \sum_{i=1}^{n}\frac{\partial L}{\partial \dot{x}^i(0)}\varepsilon^i(0) = 0$$

- but if $\varepsilon(0) = 0$, we must have:

$$\sum_{i=1}^{n}\frac{\partial L}{\partial \dot{x}^i(1)}\varepsilon^i(1) = 0$$

  so the vector:

$$\begin{pmatrix} \frac{\partial L}{\partial \dot{x}^1(1)} \\ \vdots \\ \frac{\partial L}{\partial \dot{x}^n(1)} \end{pmatrix}$$

  must be **orthogonal** to $\varepsilon(1)$, and thus, must be orthogonal to C_1 at $t = 1$

- similarly, if $\varepsilon(1) = 0$, we must have:

$$\sum_{i=1}^{n}\frac{\partial L}{\partial \dot{x}^i(0)}\varepsilon^i(0) = 0$$

  then the vector:

$$\begin{pmatrix} \frac{\partial L}{\partial \dot{x}^1(0)} \\ \vdots \\ \frac{\partial L}{\partial \dot{x}^n(0)} \end{pmatrix}$$

  must be **orthogonal** to C_0 at $t = 0$

- if we use a different Lagrangian like $L(x, \dot{x}, \dots, t)$, then we will have the same Euler-Lagrange Equation, but we will have to cancel the boundary terms, which will depend on the boundary conditions of the problem itself

- **Will the critical point curve pass through the hypersurfaces orthogonally?**
  - notice the above analysis requires that the vectors with terms $\frac{\partial L}{\partial \dot{x}^i}$ must be **orthogonal** to the hypersurfaces at the endpoints $t = 0, 1$
  - however, it is not necessarily the case that $\frac{\partial L}{\partial \dot{x}^i}$ will be **collinear** with $\dot{x}^i$
  - hence, the critical point might not incide normally on the hypersurface
  - for instance, if:

$$L = \frac{1}{2}\|\dot{x}\|^2 - \langle x, \dot{x}\rangle$$

    then:

$$\frac{\partial L}{\partial \dot{x}} = \frac{\langle \dot{x}, \ddot{x}\rangle}{\|\dot{x}\|} - \langle \dot{x}, \dot{x}\rangle - \langle x, \ddot{x}\rangle$$

    so $\frac{\partial L}{\partial \dot{x}}$ isn't collinear to $\dot{x}$; hence, if $\frac{\partial L}{\partial \dot{x}}$ is orthogonal to the surface at the endpoints, the critical point, which has derivative $\dot{x}$ won't necessarily be orthogonal to the surface

3 Lemma: Lagrangian Invariant Given a Time Derivative

Let:

$$F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$$

be a ***twice continuously differentiable*** *function of* $t \in \mathbb{R}, x \in \mathbb{R}^n$.
Further, assume we are given 2 Lagrangians:

$$L(x, \dot{x}, t) \qquad L'(x, \dot{x}, t)$$

related by:

$$L'(x, \dot{x}, t) = L(x, \dot{x}, t) + \frac{d}{dt}F(x, t)$$

Then, the ***Euler-Lagrange*** *equations for the 2 Lagrangians will be the same:*

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{\partial L'}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{x}^i}\right) = 0$$

Proof. There are 2 ways of showing this. The first one will be more direct, but the second one will reveal the true effect of adding the extra term to the Lagrangian.

(1)

We first note that by the chain rule:

$$\frac{d}{dt}F(x, t) = \frac{\partial F}{\partial t} + \sum_{j=1}^{n} \dot{x}^j \frac{\partial F}{\partial x^j}$$

Thus:

$$\begin{aligned}
\frac{\partial L'}{\partial x^i} &= \frac{\partial}{\partial x^i}\left(L(x, \dot{x}, t) + \frac{d}{dt}F(x, t)\right) \\
&= \frac{\partial}{\partial x^i}\left(L(x, \dot{x}, t) + \frac{\partial F}{\partial t} + \sum_{j=1}^{n} \dot{x}^j \frac{\partial F}{\partial x^j}\right) \\
&= \frac{\partial L}{\partial x^i} + \frac{\partial^2 F}{\partial x^i \partial t} + \sum_{j=1}^{n} \dot{x}^j \frac{\partial^2 F}{\partial x^i \partial x^j} \\
&= \frac{\partial L}{\partial x^i} + \frac{d}{dt}\left(\frac{\partial F}{\partial x^i}\right)
\end{aligned}$$

Moreover:

$$\begin{aligned}
\frac{\partial L'}{\partial \dot{x}^i} &= \frac{\partial L}{\partial \dot{x}^i} + \frac{\partial}{\partial \dot{x}^i}\left(\sum_{j=1}^{n} \dot{x}^j \frac{\partial F}{\partial x^j}\right) \\
&= \frac{\partial L}{\partial \dot{x}^i} + \frac{\partial F}{\partial \dot{x}^i}
\end{aligned}$$

Thus:

$$\frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{x}^i}\right) = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) + \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^i}\right)$$

Now, the Euler-Lagrange Equations for L state that:

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = 0$$

But by the above, this implies that:

$$\left(\frac{\partial L'}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^i}\right)\right) - \left(\frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{x}^i}\right) - \frac{d}{dt}\left(\frac{\partial F}{\partial \dot{x}^i}\right)\right) = 0 \implies \frac{\partial L'}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{x}^i}\right) = 0$$

as required.

(2)

An alternative is to see how the different Lagrangians affect the action functionals used to derive the Euler-Lagrange equations:

$$I[x] = \int_0^1 L(x, \dot{x}, t)dt \qquad I'[x] = \int_0^1 L'(x, \dot{x}, t)dt$$

Notice that:

$$I'[x] = \int_0^1 L(x, \dot{x}, t) + \frac{d}{dt}F(x, t)dt = I[x] + [F(x,t)]_0^1$$

We consider the path which satisfies each Lagrangian, by applying variations:

$$\frac{d}{ds}I[x + s\varepsilon]\bigg|_{s=0} = 0$$

$$\frac{d}{ds}I'[x + s\varepsilon]\bigg|_{s=0} = \frac{d}{ds}I[x + s\varepsilon]\bigg|_{s=0} + \left[\sum_{i=1}^n \frac{\partial F}{\partial x^i}\varepsilon^i\right]_0^1 = 0$$

But now, recall that the variations for I lead to the equation:

$$\sum_{i=1}^n \left(\left[\frac{\partial L}{\partial \dot{x}^i}\varepsilon^i\right]_0^1 + \int_0^1 \left(\frac{\partial L}{\partial x^i}\varepsilon^i - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\varepsilon^i\right)dt\right) = 0$$

Thus, for I', we must have that:

$$\sum_{i=1}^n \left(\left[\left(\frac{\partial L}{\partial \dot{x}^i} + \frac{\partial F}{\partial x^i}\right)\varepsilon^i\right]_0^1 + \int_0^1 \left(\frac{\partial L}{\partial x^i}\varepsilon^i - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\varepsilon^i\right)dt\right) = 0$$

and this must be true for **all** admissible variations.
For the variations vanishing at the endpoints, we get that:

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = 0$$

so the Euler-Lagrange equations are the same for I and I', as expected.
But then, notice that for the remaining variations, we must also have that:

$$\sum_{i=1}^n \left[\left(\frac{\partial L}{\partial \dot{x}^i} + \frac{\partial F}{\partial x^i}\right)\varepsilon^i\right]_0^1 = 0$$

In other words, adding the F term to the Lagrangin will **not** alter the "shape" of the solution, but it **can** influence how the solution behaves at the endpoints, since it alters the boundary conditions. If the endpoints are fixed, then F won't affect the boundary conditions.

□

4 Previewing Noether's Theorem: Symmetries and Conservation

4.1 Physics Terminology

4.1.1 The Momentum

Thinking of x as position and $\dot{x}$ as velocity of a particle, then:

$$\frac{\partial L}{\partial \dot{x}^i}$$

is the ***momentum*** *of the particle,* ***conjugate*** *to the variable x^i. We often denote:*

$$p_i = \frac{\partial L}{\partial \dot{x}^i}$$

4.1.2 The Force

Recall, the ***impulse*** *is:*

$$I = F\Delta t$$

but also:

$$I = \Delta p = m\Delta v$$

where p is the momentum of a particle. In other words, if over a period of time the momentum changes, it must be due to there being a non-zero force being applied on the particle. In other words, since the Euler-Lagrange equations say:

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = 0$$

and if we have:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = \frac{dp}{dt} \neq 0$$

then:

$$\frac{\partial L}{\partial x^i} \neq 0$$

so we can think of:

$$\frac{\partial L}{\partial x^i}$$

as the ***force*** *being applied on a particle in the Euler-Lagrange equations.*

4.1.3 The Energy (Beltrami's Identity)

> ***Beltrami's Identity*** *gives us an expression for the* ***energy*** *of a system:*
>
> $$E = \sum_{i=1}^{n} \left(\dot{x}^i \frac{\partial L}{\partial \dot{x}^i} \right) - L$$

4.2 Linear Momentum is Conserved Under Translation

> *Let L be a* ***Lagrangian*** *independent of x:*
>
> $$\frac{\partial L}{\partial x} = 0$$
>
> *In other words, L is invariant under translation:*
>
> $$L(x, \dot{x}, t) = L(x + a, \dot{x}, t)$$
>
> *Then, the* ***momentum****:*
>
> $$\frac{\partial L}{\partial \dot{x}^i}$$
>
> *is* ***conserved*** *along curves which solve the* ***Euler-Lagrange Equations****; in other words, moving along the solution path won't change the momentum.*

Proof. This is direct from the Euler-Lagrange equations:

$$\frac{\partial L}{\partial x^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = 0 \implies \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) = 0$$

so momentum is **constant** as time changes.

□

4.3 Energy Conservation

4.3.1 Energy is Conserved Under Time Variations

> *If the Lagrangian is not* ***explicitly*** *dependent on time:*
>
> $$\frac{\partial L}{\partial \dot{x}^i}$$
>
> *then the* ***energy****:*
>
> $$E = \sum_{i=1}^{n} \left(\dot{x}^i \frac{\partial L}{\partial \dot{x}^i} \right) - L$$
>
> *is conserved along curves solving the* ***Euler-Lagrange*** *equations.*

Proof. We differentiate Beltrami's Identity:

$$\begin{aligned}
\frac{dE}{dt} &= \sum_{i=1}^{n} \frac{d}{dt}\left(\dot{x}^i \frac{\partial L}{\partial \dot{x}^i}\right) - \frac{dL}{dt} \\
&= \left(\sum_{i=1}^{n} \ddot{x}^i \frac{\partial L}{\partial \dot{x}^i} + \dot{x}^i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right)\right) - \left(\frac{\partial L}{\partial t} + \sum_{j=1}^{n} \frac{\partial L}{\partial x^j}\dot{x}^j + \frac{\partial L}{\partial \dot{x}^j}\ddot{x}^j\right) \\
&= \sum_{i=1}^{n} \ddot{x}^i \frac{\partial L}{\partial \dot{x}^i} + \dot{x}^i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) - \frac{\partial L}{\partial x^i}\dot{x}^i + \frac{\partial L}{\partial \dot{x}^i}\ddot{x}^i \\
&= \sum_{i=1}^{n} \dot{x}^i \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}^i}\right) - \frac{\partial L}{\partial x^i}\right) \\
&= 0
\end{aligned}$$

where the last step follows by the Euler-Lagrange Equations. □

4.3.2 Energy Conservation for Simplifying ODEs

Energy conservation is a very useful concept, for reducing complicated ODEs to more manageable ODEs.

For example, consider the Lagrangian:

$$L(y(x), y'(x), x) = \frac{\sqrt{1+(y')^2}}{y}$$

Then:

$$\begin{aligned}
\frac{\partial L}{\partial y} &= -\frac{\sqrt{1+(y')^2}}{y^2} \\
\frac{\partial L}{\partial y'} &= \frac{y'}{y\sqrt{1+(y')^2}}
\end{aligned}$$

So the Euler-Lagrange equations become:

$$\frac{d}{dx}\left(\frac{y'}{y\sqrt{1+(y')^2}}\right) = -\frac{\sqrt{1+(y')^2}}{y^2}$$

which is a very non-linear, second-order ODE.

However, notice that:

$$\frac{\partial L}{\partial x} = 0$$

so it follows that energy is conserved, so:

$$\frac{dE}{dx} = 0 \implies \frac{d}{dx}\left(y' \frac{y'}{y\sqrt{1+(y')^2}} - \frac{\sqrt{1+(y')^2}}{y}\right) = 0$$

We simplify:

$$\begin{aligned}
y' \frac{y'}{y\sqrt{1+(y')^2}} - \frac{\sqrt{1+(y')^2}}{y} &= \frac{(y')^2 - (1+(y')^2)}{y\sqrt{1+(y')^2}} \\
&= -\frac{1}{y\sqrt{1+(y')^2}}
\end{aligned}$$

Then:

$$\frac{d}{dx}\left(y'\frac{y'}{y\sqrt{1+(y')^2}} - \frac{\sqrt{1+(y')^2}}{y}\right) = 0 \implies -\frac{1}{y\sqrt{1+(y')^2}} = k$$

where $k \in \mathbb{R}$. Notice, this is now a first order ODE, with some pretty neat cancellations. Energy conservation has somehow "integrated" the original ODE to give us the above.

If we square both sides and rearrange:

$$\frac{1}{k^2y^2} - 1 = (y')^2 \implies y' = \pm\sqrt{\frac{1}{k^2y^2} - 1}$$

By separation of variables (using the positive square root):

$$\int \frac{1}{\sqrt{\frac{1}{k^2y^2} - 1}} dy = x + A$$

The integral on the LHS can be rewritten:

$$\int \frac{1}{\sqrt{\frac{1}{k^2y^2} - 1}} dy = \int \frac{1}{\sqrt{\frac{1-k^2y^2}{k^2y^2}}} dy = k\int \frac{y}{\sqrt{1-k^2y^2}} dy$$

Now, let:

$$u = 1 - k^2y^2 \qquad \frac{du}{dy} = -2k^2y$$

so:

$$\begin{aligned} k\int \frac{y}{\sqrt{1-k^2y^2}} dy &= k\int \frac{y}{\sqrt{u}} \times \frac{du}{-2k^2y} \\ &= -\frac{1}{2k}\int \frac{1}{\sqrt{u}} du \\ &= -\frac{\sqrt{u}}{k} \\ &= -\frac{\sqrt{1-k^2y^2}}{k} \end{aligned}$$

Thus, it follows that:

$$-\sqrt{1-k^2y^2} = kx + kA$$

Further simplifying:

$$\begin{aligned} &-\sqrt{1-k^2y^2} = kx + kA \\ \implies &1 - k^2y^2 = k^2x^2 - 2Akx + (kA)^2 \\ \implies &x^2 + y^2 = \frac{1 + 2Akx - k^2A^2}{k^2} \\ \implies &x^2 + y^2 = \frac{1}{k^2} + \frac{2Ax}{k} - A^2 \\ \implies &x^2 - \frac{2Ax}{k} + y^2 = \frac{1}{k^2} - A^2 \\ \implies &\left(x - \frac{A}{k}\right)^2 - \frac{A^2}{k^2} + y^2 = \frac{1}{k^2} - A^2 \\ \implies &\left(x - \frac{A}{k}\right)^2 + y^2 = \frac{1+A^2}{k^2} - A^2 \end{aligned}$$

So the critical points to the action functional are circles!

5 Exercises

1. **Let $x : [0,1] \to \mathbb{R}$ be a C^3 function. Let the Lagrangian L also depend on $\ddot{x}$. Derive the Euler-Lagrange equation arising from extremising the action:**

$$I[x] = \int_0^1 L(x, \dot{x}, \ddot{x}, t)dt$$

Generalise this further to Lagrangians depending on the first k derivatives of $x : [0,1] \to \mathbb{R}^n$, which should now be a C^{k+1} function.

2. **Find the shortest path in the plane starting at the origin and ending on the parabola:**

$$y = (x-3)^2$$

3. **What is the shortest distance between 2 non-overlapping circles on the plane?**

4. **Let C_1, C_2 be 2 closed, simple plane curves, with C_1 in the upper half-plane and C_2 in the lower half-plane. Show that there exists a straight line in the plane which intersects both C_1 and C_2 normally. The line might intersect each curve at more than one point, but the intersection is normal in at least one of the points.**

5. **Find the shortest path in the plane starting at $P = (2,1)$ and ending on the hyperbola $xy = 1$.**

6 Workshop

1. **Consider 2 points P, Q. Via translation and rotation, assume that:**

$$P = (0,0) \qquad Q = (\ell, 0)$$

We approximate the geodesic between P and Q by a piecewise linear path, made our of the straight line segments:

$$P \to (x,y) \to Q$$

The arclength of this piecewise lienar path is:

$$S(x,y) = \sqrt{x^2+y^2} + \sqrt{(\ell - x)^2 + y^2}$$

(a) **Find the critical points of S as a function of (x,y).**

We differentiate:

$$\begin{aligned}\frac{\partial S}{\partial x} &= \frac{1}{2\sqrt{x^2+y^2}} \times 2x + \frac{1}{2\sqrt{(\ell-x)^2+y^2}} \times 2(\ell - x) \times -1 \\ &= \frac{x}{\sqrt{x^2+y^2}} - \frac{\ell - x}{\sqrt{(\ell-x)^2+y^2}}\end{aligned}$$

similarly:

$$\frac{\partial S}{\partial y} = \frac{y}{\sqrt{x^2+y^2}} + \frac{y}{\sqrt{(\ell-x)^2+y^2}}$$

If (x,y) is a critical point, then:

$$\frac{\partial S}{\partial x}(x,y) = \frac{\partial S}{\partial y}(x,y) = 0$$

Notice, $\frac{\partial S}{\partial y} = 0$ if and only if $y = 0$. Hence, we require that:

$$\frac{x}{\sqrt{x^2}} - \frac{\ell - x}{\sqrt{(\ell - x)^2}} = 0 \implies \frac{x}{|x|} = \frac{\ell - x}{|\ell - x|}$$

This is always true, so long as $sgn(x) = sgn(\ell - x)$. If $x < 0$, the $\ell - x > 0$, so we must have $x > 0$. Since $x < \ell$, the critical points of S are:

$$\{(x, 0) \mid 0 < x < \ell\}$$

(b) **Calculate the Hessian at the critical points, and explain your answer.**

We compute explicitly:

$$\begin{aligned}\frac{\partial^2 S}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{x}{\sqrt{x^2+y^2}} - \frac{\ell - x}{\sqrt{(\ell-x)^2+y^2}}\right) \\ &= \frac{\sqrt{x^2+y^2} - \frac{x^2}{\sqrt{x^2+y^2}}}{x^2+y^2} - \frac{\sqrt{(\ell-x)^2+y^2} + \frac{(\ell-x)^2}{\sqrt{(\ell-x)^2+y^2}}}{(\ell-x)^2+y^2} \\ &= \frac{y^2}{(x^2+y^2)^{3/2}} + \frac{y^2}{((\ell-x)^2+y^2)^{3/2}}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 S}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{y}{\sqrt{x^2+y^2}} + \frac{y}{\sqrt{(\ell-x)^2+y^2}}\right) \\ &= -\frac{yx}{(x^2+y^2)^{3/2}} + \frac{y(\ell-x)}{((\ell-x)^2+y^2)^{3/2}}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 S}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{y}{\sqrt{x^2+y^2}} + \frac{y}{\sqrt{(\ell-x)^2+y^2}}\right) \\ &= \frac{\partial}{\partial y}\left(\frac{\sqrt{x^2+y^2} - \frac{y^2}{\sqrt{x^2+y^2}}}{x^2+y^2} + \frac{\sqrt{(\ell-x)^2+y^2} - \frac{y^2}{\sqrt{(\ell-x)^2+y^2}}}{(\ell-x)^2+y^2}\right) \\ &= \frac{x^2}{(x^2+y^2)^{3/2}} + \frac{(\ell-x)^2}{((\ell-x)^2+y^2)^{3/2}}\end{aligned}$$

Hence, setting $y = 0$ and $0 < x < \ell$:

$$\left.\frac{\partial^2 S}{\partial x^2}\right|_{y=0} = \left.\frac{\partial^2 S}{\partial x \partial y}\right|_{y=0} = 0$$

and:

$$\left.\frac{\partial^2 S}{\partial y^2}\right|_{y=0} = \frac{1}{x} + \frac{1}{\ell - x} = \frac{\ell}{x(\ell - x)} > 0$$

Thus, the Hessian is:

$$H(x, 0) = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\ell}{x(\ell - x)} \end{pmatrix}$$

This has non-negative eigenvalues, so H is positive semi-definite. Thus, $(x, 0)$ corresponds to minima, so long as $0 < x < \ell$.

(c) **Argue, as Euler would, that minimising the arclength from P to Q is the same a s minimising the arclength of the piecewise regular paths from P to $(x, 0)$, and from $(x, 0)$ to Q, for any $0 < x < \ell$. Hw does this show that the geodesic is the straight line segment between P and Q?**

We see that to minimise the distance between P and any intermediate point, we need to go along the line joining P and Q. In particular, if we then continue placing intermediate points, these will all lie along the same straight line, so the shortest path between P and Q must necesarily lie along this line.

2. **Consider the arclength functional:**

$$S[x] = \int_0^1 \|\dot{x}(t)\| dt$$

of regular curves $x(t) \in \mathbb{R}^2$ from $x(0) = \underline{P}$ to $x(1) = \underline{Q}$, where $P, Q \in \mathbb{R}^2$ are 2 distinct points. A reparametrisation is a continuously differentiable map:

$$t \mapsto \bar{t}(t), \qquad \frac{d\bar{t}}{dt} \neq 0, \qquad \forall t \in [0, 1], \bar{t} \in [a, b]$$

(so $\bar{t}$ is our new parameter, which depends on the original parameter t)

(a) **Show that the arclength functional is reparametrisation invariant:**

$$S[x] = \int_0^1 \|\dot{x}(t)\| dt = \int_a^b \|\dot{x}(\bar{t})\| d\bar{t}$$

where $\dot{x}(\bar{t})$ denotes the derivative of x with respect to $\bar{t}$.

Consider a parametrisation $\gamma(t), t \in [0, 1]$. Notice, since 2 parametrisations trace out the same curve, the only different between 2 parametrisations is how **quickly** they traverse the curve. Hence, define $\tau(t)$ such that, for some other parametrisation $x(t)$:

$$\gamma(t) = x(\tau(t)), \qquad \forall t \in [0, 1]$$

Then:

$$\begin{aligned} \int_0^1 \|\gamma'(t)\| dt &= \int_0^1 \left\| \frac{d}{dt}(x(\tau(t)) \right\| dt \\ &= \int_0^1 \left\| x'(\tau) \frac{d\tau}{dt} \right\| dt \\ &= \int_{\tau(0)}^{\tau(1)} \|x'(\tau)\| d\tau \end{aligned}$$

so the parametrisation doesn't affect the value of the arc length.

Without loss of generality we have assume that $\frac{d\tau}{dt} > 0$; otherwise, we can repeat the argument with the $-\tau$ instead.

(b) **Show that we can always choose a parametrisation such that:**

$$\|\dot{x}(\bar{t})\| = 1, \qquad \forall \bar{t}$$

This is called the *arclength parametrisation.*

Notice we want:

$$\|\gamma'(t)\| = 1 \implies \|x'(\tau)\| \frac{d\tau}{dt} = 1$$

hence, we have a differential equation to solve:

$$\frac{d\tau}{dt} = \frac{1}{\|x'(\tau)\|}$$

We have a boundary condition $\tau(0) = a$. Hence, this ODe has a unique solution, and so, we can always find our desired parametrisation τ.

(c) **The Euler-Lagrange equation for the arclength functional is:**

$$\frac{d}{dt}\left(\frac{\dot{x}}{\|\dot{x}\|}\right) = 0$$

which is the equation of a straight line. If we subsequently choose an arclength parametrisation (such that $\|\dot{x}\| = 1, \forall t$), then the equation for a straight line becomes $\ddot{x} = 0$. Let $x : [0,1] \to \mathbb{R}^2$ be a continuously differentiable curve in the plane with:

$$x(0) = (0,0) \qquad x(1) = (l,0)$$

Consider the functional:

$$S[x] = \int_0^1 \frac{1}{2}\|\dot{x}(t)\|^2 dt$$

Show that the Euler-Lagrange equation for this functional is precisely the equation $\ddot{x} = 0$.

We have that:

$$L = \frac{1}{2}\|\dot{x}(t)\|^2$$

so:

$$\frac{\partial L}{\partial x^i} = 0$$

$$\frac{\partial L}{\partial \dot{x}^i} = \frac{\dot{x}^i}{\|\dot{x}\|} = \dot{x}^i \implies \frac{d}{dt}\frac{\partial L}{\partial \dot{x}^i} = \ddot{x}^i$$

as required.

3. **Let $\rho : [0,1] \to \mathbb{R}^+, \psi : [0,1] \to \mathbb{R}$ be continuously differentiable functions with:**

$$\rho(0) = r_0 \qquad \rho(1) = r_1 \qquad \psi(0) = \theta_0 \qquad \psi(1) = \theta_1$$

We will assume that:

$$(r_0, \theta_0) \neq (r_1, \theta_1)$$

Consider the functional:

$$S[\rho, \psi] = \int_0^1 \frac{1}{2}\left(\dot{\rho}^2(t) + \rho^2(t)\dot{\psi}^2(t)\right) dt$$

Recall that the critical points of this functional satisfy the Euler-Lagrange Equations:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\psi}} = \frac{\partial L}{\partial \psi}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\rho}} = \frac{\partial L}{\partial \rho}$$

(a) **Show that one of the Euler-Lagrange equations imply that:**

$$\ell := \rho^2(t)\dot{\psi}(t)$$

is a constant.
Using:

$$L = \frac{1}{2}\left(\dot{\rho}^2 + \rho^2\dot{\psi}^2\right)$$

we have that:

$$\frac{\partial L}{\partial \rho} = \rho\dot{\psi}^2$$

$$\frac{\partial L}{\partial \dot{\rho}} = \dot{\rho}$$

$$\frac{\partial L}{\partial \psi} = 0$$

$$\frac{\partial L}{\partial \dot{\psi}} = \rho^2\dot{\psi}$$

Notice, the EL equation for Ψ says that:

$$\frac{d}{dt}\left(\rho^2\dot{\psi}\right) = 0 \implies \exists \ell \in \mathbb{R} \;:\; \rho^2\dot{\psi} = \ell$$

(b) **Show that the other Euler-Lagrange equation implies that:**

$$E := \dot{\rho}^2(t) + \frac{\ell^2}{\rho(t)}$$

is a positive constant. Why can't $E = 0$?
There are 2 ways of doing this:

(1) **Beltrami's Identity**

By Beltrami's Identity, we know that:

$$E = \dot{\psi}\frac{\partial L}{\partial \dot{\psi}} + \dot{\rho}\frac{\partial L}{\partial \dot{\rho}} - L$$

is constant. Hence:

$$\begin{aligned} E &= \dot{\psi}\frac{\partial L}{\partial \dot{\psi}} + \dot{\rho}\frac{\partial L}{\partial \dot{\rho}} - L \\ &= \rho^2\dot{\psi}^2 + \dot{\rho}^2 - \frac{1}{2}\left(\dot{\rho}^2 + \rho^2\dot{\psi}^2\right) \\ &= \frac{1}{2}\left(\dot{\rho}^2 + \rho^2\dot{\psi}^2\right) \\ &= \frac{1}{2}\left(\dot{\rho}^2 + \rho^2\left(\frac{\ell}{\rho^2}\right)^2\right) \qquad = \frac{1}{2}\left(\dot{\rho}^2 + \frac{\ell^2}{\rho^2}\right) \end{aligned}$$

Hence, redefining E:

$$E = \dot{\rho}^2 + \frac{\ell^2}{\rho^2}$$

must be constant.

---

② **Using EL**

The other EL gives us:

$$\ddot{\rho} = \rho\dot{\psi}^2 = \frac{\ell}{\rho^3}$$

If we multiply both sides by $\dot{\rho}$, we get:

$$\ddot{\rho}\dot{\rho} = \frac{\ell}{\rho^3}\dot{\rho}$$

Now, notice that:

$$\frac{d}{dt}\left(\dot{\rho}^2\right) = \dot{\rho}\ddot{\rho}$$

$$\frac{d}{dt}\left(-\frac{\ell^2}{2\rho^2}\right) = \frac{\ell}{\rho^3}\dot{\rho}$$

Thus:

$$\frac{d}{dt}\left(\dot{\rho}^2 + \frac{\ell^2}{2\rho^2}\right) = 0$$

as required.

Notice, this is clearly non-negative (it is a sum of squares). However, we can't have $E = 0$. If we did, then:

$$\dot{\rho} = 0 \qquad \ell = 0$$

If $\ell = 0$, then:

$$\rho\dot{\psi}^2 = 0$$

We know that $\rho \neq 0$ (otherwise E would be undefined), so it follows that $\dot{\psi} = 0$. Thus, both ρ, ψ must be constant. But this contradicts the fact that:

$$(r_0, \theta_0) \neq (r_1, \theta_1)$$

Hence, $E > 0$.

(c) **Show that if $l = 0$ then:**

$$\psi(t) = \theta_0 = \theta_1$$

$$\rho(t) = r_0 + t(r_1 - r_0)$$

Find E in terms of r_0, r_1.

If $\ell = 0$, as discussed above we must have that:

$$\dot{\psi} = 0$$

so ψ is constant:

$$\psi = \theta_0 = \theta_1$$

Moreover, we have that:

$$E = \dot{\rho}^2$$

Hence, $\dot{\rho}$ is a positive constant:

$$\rho = at + b$$

Using initial conditions:

$$\rho(0) = r_0 \implies b = r_0$$

$$\rho(1) = r_1 \implies a + r_0 = r_1$$

Hence:

$$\rho(t) = (r_1 - r_0)t + r_0$$

and so:

$$E = \dot{\rho}^2 = (r_1 - r_0)^2$$

(d) **Let $l > 0$. Solve the Euler-Lagrange equations to find $\rho(t), \psi(t)$. You may leave the answer in terms of integration constants, which can be determined in terms of the boundary values $r_0, r_1, \theta_0, \theta_1$.**

We solve the equations by using:

$$\ell = \rho^2 \dot{\psi}$$

$$E = \dot{\rho}^2 + \frac{\ell^2}{\rho^2}$$

Rearranging:

$$\dot{\rho}^2 = E - \frac{\ell^2}{\rho^2} \implies \frac{d\rho}{dt} = \frac{\sqrt{E\rho^2 - \ell^2}}{\rho}$$

Thus:

$$t + C = \int \frac{\rho}{\sqrt{E\rho^2 - \ell^2}} d\rho$$

If we let $u = E\rho^2 - \ell^2$ then:

$$\frac{du}{d\rho} = 2E\rho$$

so:

$$\begin{aligned} t + C &= \int \frac{\rho}{\sqrt{u}} \frac{1}{2E\rho} du \\ &= \frac{1}{2E} \int \frac{1}{\sqrt{u}} du \\ &= \frac{1}{E}\sqrt{u} \\ &= \frac{1}{E}\sqrt{E\rho^2 - \ell^2} \end{aligned}$$

Solving for ρ:

$$E\rho^2 - \ell^2 = E^2(t + C)^2 \implies \rho = \sqrt{E(t + C)^2 + \frac{\ell^2}{E}}$$

Using this, we have that:

$$\ell = \left(E(t + C)^2 + \frac{\ell^2}{E}\right) \dot{\psi}$$

so:

$$\dot{\psi} = \frac{E\ell}{\ell^2 + (E(t+c))^2} = \frac{E}{\ell}\frac{1}{1+\left(\frac{E(t+c)}{\ell}\right)^2}$$

This looks like the derivative of arctan, so applying the chain rule:

$$\psi = \arctan\left(\frac{E(t+C)}{\ell}\right) + D$$

(e) **Make the variable substitution**

$$\rho(t) = \sqrt{x^2(t) + y^2(t)}$$
$$\psi(t) = \arctan(y(t)/x(t))$$

and rewrite the functional $S[\rho, \psi]$ as a functional $S_1[x, y]$, and hence interpret the solutions found in part d) geometrically.

We compute:

$$\dot{\rho} = \frac{x\dot{x} + y\dot{y}}{\sqrt{x^2+y^2}} \implies \dot{\rho}^2 = \frac{(x\dot{x}+y\dot{y})^2}{x^2+y^2}$$
$$\dot{\psi} = \frac{1}{1+\frac{y^2}{x^2}} \times \left(\frac{\dot{y}x - \dot{x}y}{x^2}\right) = \frac{\dot{y}x - \dot{x}y}{x^2+y^2}$$

Then:

$$\rho^2\dot{\psi}^2 = \frac{(\dot{y}x - \dot{x}y)^2}{x^2+y^2}$$

So:

$$\begin{aligned}
\dot{\rho}^2 + \rho^2\dot{\psi}^2 &= \frac{(x\dot{x}+y\dot{y})^2 + (\dot{y}x - \dot{x}y)^2}{x^2+y^2} \\
&= \frac{(x\dot{x})^2 + 2x\dot{x}y\dot{y} + (y\dot{y})^2 + (\dot{y}x)^2 - 2\dot{y}x\dot{x}y + (\dot{x}y)^2}{x^2+y^2} \\
&= \frac{x^2(\dot{x}^2+\dot{y}^2) + y^2(\dot{x}^2+\dot{y}^2)}{x^2+y^2} \\
&= \dot{x}^2 + \dot{y}^2
\end{aligned}$$

Hence:

$$S[x,y] = \int \frac{1}{2}(\dot{x}^2 + \dot{y}^2)$$

The extremals of this Lagrangian must still be straight line paths. We can see that ρ, ψ represent angular coordinates for these paths.

4. **Let $x : [0,1] \to \mathbb{R}^3$ be a regular curve, such that:**

$$\|x(t)\| = 1, \qquad \forall t$$

In other words, x is a regular curve on the unit sphere in $\mathbb{R}^3$. Moreover, assume that:

$$x(0) = \underline{P} \qquad x(1) = Q$$

are 2 distinct points on the sphere. Using spherical polar coordinates, we can write:

$$x(t) = \begin{pmatrix} \sin\theta(t)\sin\phi(t) \\ \sin\theta(t)\cos\phi(t) \\ \cos\theta(t) \end{pmatrix}$$

for some continuously differentiable functions $\theta(t), \phi(t)$.

(a) (3 marks) **Show that the arclength:**

$$S[x] = \int_0^1 \|\dot{x}(t)\| dt$$

defines a functional:

$$S[\theta, \phi] = \int_0^1 \sqrt{\dot{\theta}^2(t) + \sin^2\theta(t)\dot{\phi}^2(t)} dt$$

We shall use:

$$\theta = \theta(t) \qquad \phi = \phi(t)$$

throughout for readability.

We compute $\dot{x}$ by using the chain rule:

$$\begin{aligned} \frac{d}{dt}(\sin\theta\sin\phi) &= \frac{d}{d\theta}(\sin\theta\sin\phi)\dot{\theta} + \frac{d}{d\phi}(\sin\theta\sin\phi)\dot{\phi} \\ &= \dot{\theta}\cos\theta\sin\phi + \dot{\phi}\sin\theta\cos\phi \end{aligned}$$

$$\begin{aligned} \frac{d}{dt}(\sin\theta\cos\phi) &= \frac{d}{d\theta}(\sin\theta\cos\phi)\dot{\theta} + \frac{d}{d\phi}(\sin\theta\cos\phi)\dot{\phi} \\ &= \dot{\theta}\cos\theta\cos\phi - \dot{\phi}\sin\theta\sin\phi \end{aligned}$$

$$\begin{aligned} \frac{d}{dt}(\cos\theta) &= \frac{d}{d\theta}(\cos\theta)\dot{\theta} \\ &= -\dot{\theta}\sin\theta \end{aligned}$$

Hence, we have that:

$$\dot{x} = \begin{pmatrix} \dot{\theta}\cos\theta\sin\phi + \dot{\phi}\sin\theta\cos\phi \\ \dot{\theta}\cos\theta\cos\phi - \dot{\phi}\sin\theta\sin\phi \\ -\dot{\theta}\sin\theta \end{pmatrix}$$

Thus:

$$\begin{aligned} \|\dot{x}\|^2 &= (\dot{\theta}\cos\theta\sin\phi + \dot{\phi}\sin\theta\cos\phi)^2 + (\dot{\theta}\cos\theta\cos\phi - \dot{\phi}\sin\theta\sin\phi)^2 + (-\dot{\theta}\sin\theta)^2 \\ &= \dot{\theta}^2\cos^2\theta\sin^2\phi + 2\dot{\theta}\dot{\phi}\cos\theta\sin\phi\sin\theta\cos\phi + \dot{\phi}^2\sin^2\theta\cos^2\phi \\ &+ \dot{\theta}^2\cos^2\theta\cos^2\phi - 2\dot{\theta}\dot{\phi}\cos\theta\sin\phi\sin\theta\cos\phi + \dot{\phi}^2\sin^2\theta\sin^2\phi + \dot{\theta}^2\sin^2\theta \\ &= \dot{\theta}^2\cos^2\theta(\sin^2\phi + \cos^2\phi) + \dot{\phi}^2\sin^2\theta(\sin^2\phi + \cos^2\phi) + \dot{\theta}^2\sin^2\theta \\ &= \dot{\theta}^2(\cos^2\theta + \sin^2\theta) + \dot{\phi}^2\sin^2\theta \\ &= \dot{\theta}^2 + \dot{\phi}^2\sin^2\theta \end{aligned}$$

Finally, it follows that:

$$S[x] = \int_0^1 \|\dot{x}(t)\| dt = \int_0^1 \sqrt{\dot{\theta}^2(t) + \sin^2\theta(t)\dot{\phi}^2(t)} dt$$

as required.

(b) (4 marks) **Write down the corresponding Euler-Lagrange equations and show that one of the equations implies that:**

$$\ell := \frac{\sin^2\theta(t)\dot{\phi}(t)}{\sqrt{\dot{\theta}^2(t)+\sin^2\theta(t)\dot{\phi}^2(t)}}$$

is a constant, and that $|\ell| \leq 1$.

Let:

$$L(\theta,\phi,\dot{\theta},\dot{\phi},t) = \sqrt{\dot{\theta}^2(t)+\sin^2\theta(t)\dot{\phi}^2(t)}$$

The Euler-Lagrange equations for the arclength $S[\theta,\phi]$ are:

$$\frac{d}{dt}\frac{\partial L}{\partial\dot{\theta}} = \frac{\partial L}{\partial\theta}$$
$$\frac{d}{dt}\frac{\partial L}{\partial\dot{\phi}} = \frac{\partial L}{\partial\phi}$$

We compute these quantities:

$$\frac{\partial L}{\partial\theta} = \frac{1}{2L}\times\frac{\partial}{\partial\theta}(\dot{\theta}^2+\sin^2\theta\dot{\phi}^2) = \frac{\dot{\phi}^2\sin\theta\cos\theta}{L}$$
$$\frac{\partial L}{\partial\dot{\theta}} = \frac{1}{2L}\times\frac{\partial}{\partial\dot{\theta}}(\dot{\theta}^2+\sin^2\theta\dot{\phi}^2) = \frac{\dot{\theta}}{L}$$
$$\frac{\partial L}{\partial\phi} = \frac{1}{2L}\times\frac{\partial}{\partial\phi}(\dot{\theta}^2+\sin^2\theta\dot{\phi}^2) = 0$$
$$\frac{\partial L}{\partial\dot{\phi}} = \frac{1}{2L}\times\frac{\partial}{\partial\dot{\phi}}(\dot{\theta}^2+\sin^2\theta\dot{\phi}^2) = \frac{\dot{\phi}\sin^2\theta}{L}$$

Hence, the Euler-Lagrange equations for the arclength are:

$$\frac{d}{dt}\left(\frac{\dot{\theta}}{L}\right) = \frac{\dot{\phi}^2\sin\theta\cos\theta}{L}$$

$$\frac{d}{dt}\left(\frac{\dot{\phi}\sin^2\theta}{L}\right) = 0$$

The second equation implies that:

$$\frac{\dot{\phi}\sin^2\theta}{L} = \frac{\dot{\phi}\sin^2\theta}{\sqrt{\dot{\theta}^2+\dot{\phi}^2\sin^2\theta}} = \ell \in \mathbb{R}$$

Moreover:

$$\begin{aligned}
|\ell| &= \frac{|\dot{\phi}|\sin^2\theta}{\sqrt{\dot{\theta}^2+\dot{\phi}^2\sin^2\theta}} \\
&\leq \frac{|\dot{\phi}|\sin^2\theta}{\sqrt{|\dot{\phi}|^2\sin^2\theta}}, \qquad \textit{(since } \dot{\theta}^2 \geq 0, \forall t \textit{ and } \dot{\phi}^2 = |\dot{\phi}|^2\textit{)} \\
&= \sqrt{\frac{|\dot{\phi}|^2\sin^4\theta}{|\dot{\phi}|^2\sin^2\theta}} \\
&= |\sin\theta| \\
&\leq 1
\end{aligned}$$

so it follows that:

$$|\ell| \leq 1$$

as required.

(c) **Argue that the geodesics on the sphere are great circles as follows. First notice that there is a rotation of the sphere about the origin which takes $\underline{P}$ to the North pole, and takes Q to the points with coordinates:**

$$\theta(1) = \theta_1 > 0 \qquad \phi(1) = 0$$

i. (4 marks) **Show that with these boundary conditions:**

$$S[\theta, \phi] \geq \theta_1$$

Without loss of generality, we can assume that $\theta(t), \phi(t)$ are (strictly) non-decreasing functions of t. This means that:

$$\dot{\theta} \geq 0 \qquad \dot{\phi} \geq 0$$

They must be monotonic, since they range over angles, and if they are (strictly) non-increasing, we can, just reparametrise via:

$$\theta := \theta(1-t) \qquad \phi := \phi(1-t)$$

to make them (stricly) non-decreasing.

Notice, $\forall t \in \mathbb{R}$:

$$\sin^2\theta\dot{\phi}^2 \geq 0$$

so:

$$\begin{aligned}
S[\theta, \phi] &= \int_0^1 \sqrt{\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2} dt \\
&\geq \int_0^1 \sqrt{\dot{\theta}^2} dt \\
&\geq \int_0^1 \dot{\theta} dt, \quad \textit{(since } \dot{\theta} \geq 0\textit{)} \\
&= \theta(1) - \theta(0) \\
&= \theta_1 - \theta(0)
\end{aligned}$$

But now, since $\underline{P}$ is at the North Pole, we have that:

$$\underline{P} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin\theta(0)\sin\phi(0) \\ \sin\theta(0)\cos\phi(0) \\ \cos\theta(0) \end{pmatrix}$$

which implies that:

$$\theta(0) = 0$$

Hence, we have shown that:

$$S[\theta, \phi] \geq \theta_1 - \theta(0) = \theta_1$$

as required.

ii. (3 marks) **Show that $S[\theta, \phi] = \theta_1$ precisely when $\phi(t) = 0$ for all t, so that $l = 0$. In this case, the curve is a segment of a meridian, which is part of a great circle.**

Firstly, if $\forall t \in [0,1], \phi(t) = 0$, then:

$$S[\theta, \phi] = \int_0^1 \sqrt{\dot{\theta}^2} dt = \theta_1$$

by the part above.

Notice, by the work above we have that:

$$\begin{aligned} S[\theta, \phi] &= \int_0^1 \sqrt{\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2} dt = \theta_1 \\ \iff \forall t \in [0,1], &\quad \sqrt{\dot{\theta}^2 + \sin^2\theta\dot{\phi}^2} = \sqrt{\dot{\theta}^2} \\ \iff \forall t \in [0,1], &\quad \sin^2\theta\dot{\phi}^2 = 0 \end{aligned}$$

Since $\theta(1) = \theta_1 > 0$, we know that $\theta(t)$ isn't always 0, so by the continuity of θ we must also have that $\sin^2\theta$ isn't always 0. Thus:

$$S[\theta, \phi] = \theta_1 \iff \forall t \in [0,1], \quad \dot{\phi}^2(t) = 0 \iff \forall t \in [0,1], \quad \dot{\phi}(t) = 0$$

However, this is true **if and only if**:

$$\forall t \in [0,1], \quad \phi(t) = C \in \mathbb{R}$$

But by the initial condition, $\phi(1) = 0$, so ϕ must be identically 0 for any $t \in [0,1]$, and so we have:

$$S[\theta, \phi] = \theta_1 \iff \forall t \in [0,1], \quad \phi(t) = 0$$

as required.

iii. (1 mark) **Argue that the original geodesic from $\underline{P}$ to $\underline{Q}$, which is obtained from this one by rotating the sphere back to the original position, is again part of a great circle.**

This geodesic is a great circle, which means there exists a plane Π in $\mathbb{R}^3$, going through the origin, such that the intersection of plane and sphere gives this geodesic. If we rotate the sphere back to its original position, since the sphere is rotated about the origin, the sphere remains centered at the origin and in the same location in $\mathbb{R}^3$. Thus, the plane Π will remain intersecting the sphere, albeit through the original geodesic. Hence, the original geodesic must be part of a great circle.