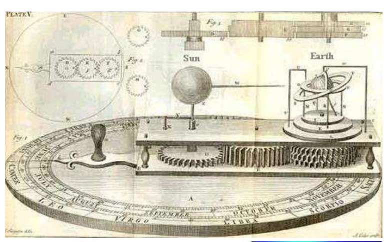
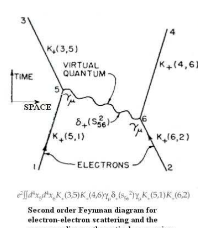
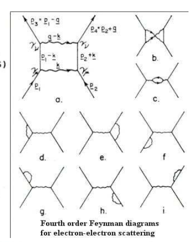
An introduction to Lagrangian and Hamiltonian mechanics

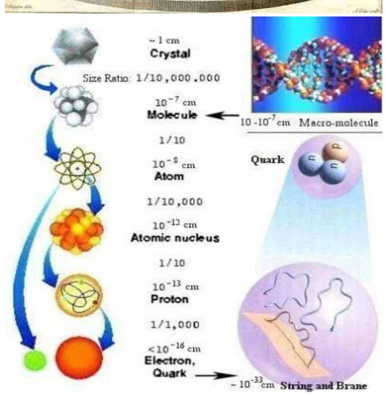
Lecture notes

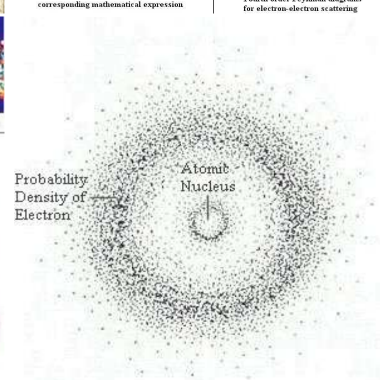
Simon J.A. Malham











1. Introduction

Newtonian mechanics took the Apollo astronauts to the moon. It also took the voyager spacecraft to the far reaches of the solar system. However Newtonian mechanics is a consequence of a more general scheme. One that brought us quantum mechanics, and thus the digital age. Indeed it has pointed us beyond that as well. The scheme is Lagrangian and Hamiltonian mechanics. Its original prescription rested on two principles. First that we should try to express the state of the mechanical system using the minimum representation possible and which reflects the fact that the physics of the problem is coordinate-invariant. Second, a mechanical system tries to optimize its *action* from one split second to the next; often this corresponds to minimizing its total energy as it evolves from one state to the next. These notes are intended as an elementary introduction into these ideas and the basic prescription of Lagrangian and Hamiltonian mechanics. A pre-requisite is the thorough understanding of the calculus of variations, which is where we begin.

2. Calculus of variations

Many physical problems involve the minimization (or maximization) of a quantity that is expressed as an integral.

2.1. Example (Euclidean geodesic). Consider the path that gives the shortest distance between two points in the plane, say (x_1, y_1) and (x_2, y_2). Suppose that the general curve joining these two points is given by $y = y(x)$. Then our goal is to find the function $y(x)$ that minimizes the arclength:

$$\begin{aligned} J(y) &= \int_{(x_1,y_1)}^{(x_2,y_2)} \mathrm{d}s \\ &= \int_{x_1}^{x_2} \sqrt{1 + (y_x)^2}\,\mathrm{d}x. \end{aligned}$$

Here we have used that for a curve $y = y(x)$, if we make a small increment in x, say Δx, and the corresponding change in y is Δy, then by Pythagoras' theorem the corresponding change in length along the curve is $\Delta s = +\sqrt{(\Delta x)^2 + (\Delta y)^2}$. Hence we see that

$$\Delta s = \frac{\Delta s}{\Delta x}\,\Delta x = \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2}\,\Delta x.$$

Note further that here, and hereafter, we use $y_x = y_x(x)$ to denote the derivative of y, i.e. $y_x(x) = y'(x)$ for each x for which the derivative is defined.

2.2. Example (Brachistochrome problem; John and James Bernoulli 1697). Suppose a particle/bead is allowed to slide freely along a wire under gravity (force $F = -gk$ where k is the unit upward vertical vector) from a point (x_1, y_1) to the origin $(0, 0)$. Find the curve $y = y(x)$ that minimizes the time of descent:

$$\begin{aligned} J(y) &= \int_{(x_1,y_1)}^{(0,0)} \frac{1}{v}\,\mathrm{d}s \\ &= \int_{x_1}^{0} \frac{\sqrt{1 + (y_x)^2}}{\sqrt{2g(y_1 - y)}}\mathrm{d}x. \end{aligned}$$

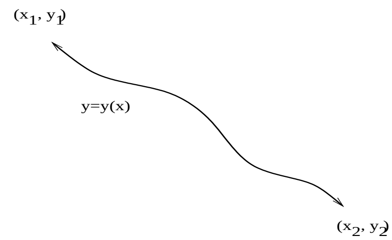


FIGURE 1. In the Euclidean geodesic problem, the goal is to find the path with minimum total length between points (x_1, y_1) and (x_2, y_2).

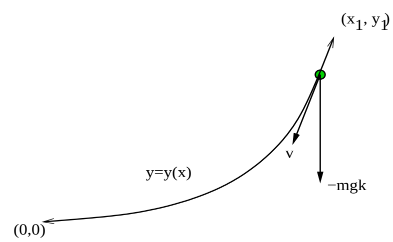


FIGURE 2. In the Brachistochrome problem, a bead can slide freely under gravity along the wire. The goal is to find the shape of the wire that minimizes the time of descent of the bead.

Here we have used that the total energy, which is the sum of the kinetic and potential energies,

$$E = \tfrac{1}{2}mv^2 + mgy,$$

is constant. Assume the initial condition is $v = 0$ when $y = y_1$, i.e. the bead starts with zero velocity at the top end of the wire. Since its total energy is constant, its energy at any time t later, when its height is y and its velocity is v, is equal to its initial energy. Hence we have

$$\tfrac{1}{2}mv^2 + mgy = 0 + mgy_1 \qquad \Leftrightarrow \qquad v = +\sqrt{2g(y_1 - y)}.$$

3. Euler–Lagrange equation

We can see that the two examples above are special cases of a more general problem scenario.

3.1. Classical variational problem. Suppose the given function $F(\cdot,\cdot,\cdot)$ is twice continuously differentiable with respect to all of its arguments. Among all functions/paths $y = y(x)$, which are twice continuously differentiable on the interval $[a, b]$ with $y(a)$ and $y(b)$ specified, find the one which extremizes the *functional* defined by

$$J(y) := \int_a^b F(x, y, y_x)\,\mathrm{d}x.$$

Theorem 1 (**Euler–Lagrange equations**). *The function $u = u(x)$ that extremizes the functional J necessarily satisfies the* Euler–Lagrange equation

$$\frac{\partial F}{\partial u} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial u_x}\right) = 0,$$

on $[a, b]$.

Proof. Consider the family of functions on $[a, b]$ given by

$$y(x;\epsilon) := u(x) + \epsilon\eta(x),$$

where the functions $\eta = \eta(x)$ are twice continuously differentiable and satisfy $\eta(a) = \eta(b) = 0$. Here ϵ is a small real parameter, and of course, the function $u = u(x)$ is our 'candidate' extremizing function. We set (see for example Evans [**4**, Section 3.3])

$$\varphi(\epsilon) := J(u + \epsilon\eta).$$

If the functional J has a local maximum or minimum at u, then u is a stationary function for J, and for all η we must have

$$\varphi'(0) = 0.$$

To evaluate this condition for the functional given in integral form above, we need to first determine $\varphi'(\epsilon)$. By direct calculation,

$$\begin{aligned}
\varphi'(\epsilon) &= \frac{\mathrm{d}}{\mathrm{d}\epsilon} J(u + \epsilon\eta) \\
&= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\int_a^b F(x, u + \epsilon\eta, u_x + \epsilon\eta_x)\,\mathrm{d}x \\
&= \int_a^b \frac{\partial}{\partial\epsilon} F(x, y(x;\epsilon), y_x(x;\epsilon))\,\mathrm{d}x \\
&= \int_a^b \left(\frac{\partial F}{\partial y}\frac{\partial y}{\partial\epsilon} + \frac{\partial F}{\partial y_x}\frac{\partial y_x}{\partial\epsilon}\right)\mathrm{d}x \\
&= \int_a^b \left(\frac{\partial F}{\partial y}\eta(x) + \frac{\partial F}{\partial y_x}\eta'(x)\right)\mathrm{d}x.
\end{aligned}$$

The integration by parts formula tells us that

$$\int_a^b \frac{\partial F}{\partial y_x}\eta'(x)\,\mathrm{d}x = \left[\frac{\partial F}{\partial y_x}\,\eta(x)\right]_{x=a}^{x=b} - \int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y_x}\right)\eta(x)\,\mathrm{d}x.$$

Recall that $\eta(a) = \eta(b) = 0$, so the boundary term (first term on the right) vanishes in this last formula, and we see that

$$\varphi'(\epsilon) = \int_a^b \left(\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y_x}\right)\right)\eta(x)\,\mathrm{d}x.$$

If we now set $\epsilon = 0$, the condition for u to be a critical point of J is

$$\int_a^b \left(\frac{\partial F}{\partial u} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial u_x}\right)\right)\eta(x)\,\mathrm{d}x = 0.$$

Now we note that the functions $\eta = \eta(x)$ are arbitrary, using Lemma 1 below, we can deduce that pointwise, i.e. for all $x \in [a, b]$, necessarily u must satisfy the Euler–Lagrange equation shown. □

3.2. Useful lemma.

LEMMA 1 (**Useful lemma**). *If* $\alpha(x)$ *is continuous in* $[a,b]$ *and*

$$\int_a^b \alpha(x)\,\eta(x)\,\mathrm{d}x = 0$$

for all continuously differentiable functions $\eta(x)$ *which satisfy* $\eta(a) = \eta(b) = 0$*, then* $\alpha(x) \equiv 0$ *in* $[a,b]$.

PROOF. Assume $\alpha(z) > 0$, say, at some point $a < z < b$. Then since α is continuous, we must have that $\alpha(x) > 0$ in some open neighbourhood of z, say in $a < \underline{z} < z < \overline{z} < b$. The choice

$$\eta(x) = \begin{cases} (x-\underline{z})^2(\overline{z}-x)^2, & \text{for } x \in [\underline{z},\overline{z}], \\ 0, & \text{otherwise}, \end{cases}$$

which is a continuously differentiable function, implies

$$\int_a^b \alpha(x)\,\eta(x)\,\mathrm{d}x = \int_{\underline{z}}^{\overline{z}} \alpha(x)\,(x-\underline{z})^2(\overline{z}-x)^2\,\mathrm{d}x > 0,$$

a contradiction. □

3.3. Remarks. Some important theoretical and practical points to keep in mind are as follows.

(1) The Euler–Lagrange equation is a *necessary* condition: if such a $u = u(x)$ exists that extremizes J, then u satisfies the Euler–Lagrange equation. Such a u is known as a *stationary function* of the *functional* J.
(2) The Euler–Lagrange equation above is an ordinary differential equation for u; this will be clearer once we consider some examples presently.
(3) Note that the extremal solution u is independent of the coordinate system you choose to represent it (see Arnold [**2**, Page 59]). For example, in the Euclidean geodesic problem, we could have used polar coordinates (r,θ), instead of Cartesian coordinates (x,y), to express the total arclength J. Formulating the Euler–Lagrange equations in these coordinates and then solving them will tell us that the extremizing solution is a straight line (only it will be expressed in polar coordinates).
(4) Let $\mathbb{Y}$ denote a function space; in the context above $\mathbb{Y}$ was the space of twice continuously differentiable functions on $[a,b]$ which are fixed at $x = a$ and $x = b$. A *functional* is a real-valued map and here $J\colon \mathbb{Y} \to \mathbb{R}$.
(5) We define the *first variation* $\delta J(u,\eta)$ of the functional J, at u in the direction η, to be $\delta J(u,\eta) := \varphi'(0)$.
(6) Is u a maximum, minimum or saddle point for J? The physical context should hint towards what to expect. Higher order variations will give you the appropriate mathematical determination.
(7) The functional J has a local minimum at u iff there is an open neighbourhood $U \subset \mathbb{Y}$ of u such that $J(y) \geq J(u)$ for all $y \in U$. The functional J has a local maximum at u when this inequality is reversed.
(8) We will generalize all these notions to multi-dimensions and systems later.

Solution (Euclidean geodesic). Recall, this variational problem concerns finding the shortest distance between the two points (x_1, y_1) and (x_2, y_2) in the plane. This is equivalent to minimizing the total arclength functional

$$J(y) = \int_{x_1}^{x_2} \sqrt{1 + (y_x)^2}\, \mathrm{d}x.$$

Hence in this case, the *integrand* is $F(x, y, y_x) = \sqrt{1 + (y_x)^2}$. From the general theory outlined above, we know that the extremizing solution satisfies the Euler–Lagrange equation

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y_x}\right) = 0.$$

We substitute the actual form for F we have in this case which gives

$$\begin{aligned}
& & -\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial}{\partial y_x}\left(\left(1+(y_x)^2\right)^{\frac{1}{2}}\right)\right) &= 0 \\
&\Leftrightarrow & \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{y_x}{\left(1+(y_x)^2\right)^{\frac{1}{2}}}\right) &= 0 \\
&\Leftrightarrow & \frac{y_{xx}}{\left(1+(y_x)^2\right)^{\frac{1}{2}}} - \frac{(y_x)^2 y_{xx}}{\left(1+(y_x)^2\right)^{\frac{3}{2}}} &= 0 \\
&\Leftrightarrow & \frac{\left(1+(y_x)^2\right) y_{xx}}{\left(1+(y_x)^2\right)^{\frac{3}{2}}} - \frac{(y_x)^2 y_{xx}}{\left(1+(y_x)^2\right)^{\frac{3}{2}}} &= 0 \\
&\Leftrightarrow & \frac{y_{xx}}{\left(1+(y_x)^2\right)^{\frac{3}{2}}} &= 0 \\
&\Leftrightarrow & y_{xx} &= 0.
\end{aligned}$$

Hence $y(x) = c_1 + c_2 x$ for some constants c_1 and c_2. Using the initial and starting point data we see that the solution is the straightline function (as we should expect)

$$y = \left(\frac{y_2 - y_1}{x_2 - x_1}\right)(x - x_1) + y_1.$$

Note that this calculation might have been a bit shorter if we had recognised that this example corresponds to the third special case in the next section.

4. Alternative form and special cases

4.1. Alternative form. The Euler–Lagrange equations given by

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y_x}\right) = 0$$

are equivalent to the following alternative formualtion of the Euler–Lagrange equations

$$\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}x}\left(F - y_x \frac{\partial F}{\partial y_x}\right) = 0.$$

4.2. Special cases. When the functional F does not explicitly depend on one or more variables, then the Euler–Lagrange equations simplify considerably. We have the following three notable cases:

(1) If $F = F(y, y_x)$ only, i.e. it does not *explicitly* depend on x, then the alternative form for the Euler–Lagrange equation implies

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(F - y_x \frac{\partial F}{\partial y_x}\right) = 0 \qquad \Leftrightarrow \qquad F - y_x \frac{\partial F}{\partial y_x} = c,$$

for some arbitrary constant c.

(2) If $F = F(x, y_x)$ only, i.e. it does not *explicitly* depend on y, then the Euler–Lagrange equation implies

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y_x}\right) = 0.$$

(3) If $F = F(y_x)$ only, then the Euler–Lagrange equation implies

$$\begin{aligned} 0 &= \frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y_x}\right) \\ &= \frac{\partial F}{\partial y} - \left(\frac{\partial^2 F}{\partial x \partial y_x} + y_x \frac{\partial^2 F}{\partial y \partial y_x} + y_{xx} \frac{\partial^2 F}{\partial y_x \partial y_x}\right) \\ &= -y_{xx} \frac{\partial^2 F}{\partial y_x \partial y_x}. \end{aligned}$$

Hence $y_{xx} = 0$, i.e. $y = y(x)$ is a linear function of x and has the form

$$y = c_1 + c_2 x,$$

for some constants c_1 and c_2.

4.3. Solution (Brachistochrome problem). Recall, this variational problem concerns a particle/bead which can freely slide along a wire under the force of gravity. The wire is represented by a curve $y = y(x)$ from (x_1, y_1) to the origin $(0, 0)$. The goal is to find the shape of the wire, i.e. $y = y(x)$, which minimizes the time of descent of the bead, which is given by the functional

$$J(y) = \int_{x_1}^{0} \sqrt{\frac{1 + (y_x)^2}{2g(y_1 - y)}}\, \mathrm{d}x = \frac{1}{\sqrt{2g}} \int_{x_1}^{0} \sqrt{\frac{1 + (y_x)^2}{(y_1 - y)}}\, \mathrm{d}x.$$

Hence in this case, the integrand is

$$F(x, y, y_x) = \sqrt{\frac{1 + (y_x)^2}{(y_1 - y)}}.$$

From the general theory, we know that the extremizing solution satisfies the Euler–Lagrange equation. Note that the multiplicative constant factor $1/\sqrt{2g}$ should not affect the extremizing solution path; indeed it divides out of the Euler–Lagrange equations. Noting that the integrand F does not explicitly depend on x, then the alternative form for the Euler–Lagrange equation may be easier:

$$\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}x}\left(F - y_x \frac{\partial F}{\partial y_x}\right) = 0.$$

This immediately implies that for some constant c, the Euler–Lagrange equation is

$$F - y_x \frac{\partial F}{\partial y_x} = c.$$

Now substituting the form for F into this gives

$$\begin{aligned}
& \frac{\big(1+(y_x)^2\big)^{\frac12}}{(y_1-y)^{\frac12}} - y_x \frac{\partial}{\partial y_x}\left(\frac{\big(1+(y_x)^2\big)^{\frac12}}{(y_1-y)^{\frac12}}\right) = c \\
\Leftrightarrow\quad & \frac{\big(1+(y_x)^2\big)^{\frac12}}{(y_1-y)^{\frac12}} - y_x \frac{\frac12\cdot 2\cdot y_x}{(y_1-y)^{\frac12}\big(1+(y_x)^2\big)^{\frac12}} = c \\
\Leftrightarrow\quad & \frac{\big(1+(y_x)^2\big)^{\frac12}}{(y_1-y)^{\frac12}} - \frac{(y_x)^2}{(y_1-y)^{\frac12}\big(1+(y_x)^2\big)^{\frac12}} = c \\
\Leftrightarrow\quad & \frac{1+(y_x)^2}{(y_1-y)^{\frac12}\big(1+(y_x)^2\big)^{\frac12}} - \frac{(y_x)^2}{(y_1-y)^{\frac12}\big(1+(y_x)^2\big)^{\frac12}} = c \\
\Leftrightarrow\quad & \frac{1}{(y_1-y)\big(1+(y_x)^2\big)} = c^2.
\end{aligned}$$

We can now rearrange this equation so that

$$(y_x)^2 = \frac{1}{c^2(y_1-y)} - 1 \qquad \Leftrightarrow \qquad (y_x)^2 = \frac{1-c^2y_1-y}{c^2y_1-c^2y}.$$

If we set $a = y_1$ and $b = \frac{1}{c^2} - y_1$, then this equation becomes

$$y_x = \left(\frac{b+y}{a-y}\right)^{\frac12}.$$

To find the solution to this ordinary differential equation we make the change of variable

$$y = \tfrac12(a-b) - \tfrac12(a+b)\cos\theta.$$

If we substitute this into the ordinary differential equation above and use the chain rule we get

$$\tfrac12(a+b)\sin\theta\,\frac{\mathrm{d}\theta}{\mathrm{d}x} = \left(\frac{1-\cos\theta}{1+\cos\theta}\right)^{\frac12}.$$

Now we use that $1/(\mathrm{d}\theta/\mathrm{d}x) = \mathrm{d}x/\mathrm{d}\theta$, and that $\sin\theta = \sqrt{1-\cos^2\theta}$, to get

$$\begin{aligned}
& \frac{\mathrm{d}x}{\mathrm{d}\theta} = \tfrac12(a+b)\sin\theta\cdot\left(\frac{1+\cos\theta}{1-\cos\theta}\right)^{\frac12} \\
\Leftrightarrow\quad & \frac{\mathrm{d}x}{\mathrm{d}\theta} = \tfrac12(a+b)\cdot(1-\cos^2\theta)^{\frac12}\cdot\left(\frac{1+\cos\theta}{1-\cos\theta}\right)^{\frac12} \\
\Leftrightarrow\quad & \frac{\mathrm{d}x}{\mathrm{d}\theta} = \tfrac12(a+b)\cdot(1+\cos\theta)^{\frac12}(1-\cos\theta)^{\frac12}\cdot\left(\frac{1+\cos\theta}{1-\cos\theta}\right)^{\frac12} \\
\Leftrightarrow\quad & \frac{\mathrm{d}x}{\mathrm{d}\theta} = \tfrac12(a+b)(1+\cos\theta).
\end{aligned}$$

We can directly integrate this last equation to find x as a function of θ. In other words we can find the solution to the ordinary differential equation for $y = y(x)$

above in parametric form, which with some minor rearrangment, can expressed as (here d is an arbitrary constant of integration)

$$\begin{aligned} x + d &= \tfrac{1}{2}(a+b)(\theta + \sin\theta), \\ y + b &= \tfrac{1}{2}(a+b)(1 - \cos\theta). \end{aligned}$$

This is the parametric representation of a *cycloid*.

5. Multivariable systems

We consider possible generalizations of functionals to be extremized. For more details see for example Keener [**5**, Chapter 5].

5.1. Higher derivatives. Suppose we are asked to find the curve $y = y(x) \in \mathbb{R}$ that extremizes the functional

$$J(y) := \int_a^b F(x, y, y_x, y_{xx})\,\mathrm{d}x,$$

subject to $y(a)$, $y(b)$, $y_x(a)$ and $y_x(b)$ being fixed. Here the functional quantity to be extremised depends on the curvature y_{xx} of the path. Necessarily the extremizing curve y satisfies the Euler–Lagrange equation (which is an ordinary differential equation)

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y_x}\right) + \frac{\mathrm{d}^2}{\mathrm{d}x^2}\left(\frac{\partial F}{\partial y_{xx}}\right) = 0.$$

Note this follows by analogous arguments to those used for the classical variational problem above.

5.2. Multiple dependent variables. Suppose we are asked to find the multi-dimensional curve $\boldsymbol{y} = \boldsymbol{y}(x) \in \mathbb{R}^N$ that extremises the functional

$$J(\boldsymbol{y}) := \int_a^b F(x, \boldsymbol{y}, \boldsymbol{y}_x)\,\mathrm{d}x,$$

subject to $\boldsymbol{y}(a)$ and $\boldsymbol{y}(b)$ being fixed. Note $x \in [a,b]$ but here $\boldsymbol{y} = \boldsymbol{y}(x)$ is a curve in N-dimensional space and is thus a vector so that (here we use the notation $' \equiv \mathrm{d}/\mathrm{d}x$)

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{y}_x = \begin{pmatrix} y_1' \\ \vdots \\ y_N' \end{pmatrix}.$$

Necessarily the extremizing curve $\boldsymbol{y}$ satisfies a set of Euler–Lagrange equations, which are equivalent to a system of ordinary differential equations, given for $i = 1, \ldots, N$ by:

$$\frac{\partial F}{\partial y_i} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y_i'}\right) = 0.$$

6. Lagrange multipliers

For the moment, let us temporarily put aside variational calculus and consider a problem in standard multivariable calculus.

6.1. Constrained optimization problem. Consider the following problem.

> Find the stationary points of the scalar function $f(\boldsymbol{x})$ where $\boldsymbol{x} = (x_1, \ldots, x_N)$ subject to the constraints $g_k(\boldsymbol{x}) = 0$, where $k = 1, \ldots, m$, with $m < N$.

Note that the graph $y = f(\boldsymbol{x})$ of the function f represents a hyper-surface in $(N + 1)$-dimensional space. The constraints are given implicitly; each one also represents a hyper-surface in $(N+1)$-dimensional space. In principle we could solve the system of m constraint equations, say for $x_1, \ldots, x_m$ in terms of the remaining variables $x_{m+1}, \ldots, x_N$. We could then substitute these into f, which would now be a function of $(x_{m+1}, \ldots, x_N)$ only. (We could solve the constraints for any subset of m variables x_i and substitute those in if we wished or this was easier, or avoided singularities, and so forth.) We would then proceed in the usual way to find the stationary points of f by considering the partial derivative of f with respect to all the remaining variables $x_{m+1}, \ldots, x_N$, setting those partial derivatives equal to zero, and then solving that system of equations. However solving the constraint equations may be very difficult, and the method of *Lagrange multipliers* provides an elegant alternative (see McCallum *et. al.* [**9**, Section 14.3]).

6.2. Method of Lagrange multipliers. The idea is to convert the constrained optimization problem to an 'unconstrained' one as follows. Form the *Lagrangian function*

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) := f(\boldsymbol{x}) + \sum_{k=1}^{m} \lambda_k \, g_k(\boldsymbol{x}),$$

where the parameter variables $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$ are known as the *Lagrange multipliers*. Note that $\mathcal{L}$ is a function of both $\boldsymbol{x}$ and $\boldsymbol{\lambda}$, i.e. of $N + m$ variables in total. The partial derivatives of $\mathcal{L}$ with respect to all of its dependent variables are:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x_j} &= \frac{\partial f}{\partial x_j} + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_j}, \\ \frac{\partial \mathcal{L}}{\partial \lambda_k} &= g_k, \end{aligned}$$

where $j = 1, \ldots, N$ and $k = 1, \ldots, m$. At the stationary points of $\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda})$, necessarily all these partial derivatives must be zero, and we must solve the following 'unconstrained' problem:

$$\begin{aligned} \frac{\partial f}{\partial x_j} + \sum_{k=1}^{m} \lambda_k \frac{\partial g_k}{\partial x_j} &= 0, \\ g_k &= 0, \end{aligned}$$

where $j = 1, \ldots, N$ and $k = 1, \ldots, m$. Note we have $N + m$ equations in $N + m$ unknowns. Assuming that we can solve this system to find a stationary point $(\boldsymbol{x}_*, \boldsymbol{\lambda}_*)$ of $\mathcal{L}$ (there could be none, one, or more) then $\boldsymbol{x}_*$ is *also* a stationary point of the original *constrained* problem. Recall: what is important about the formulation of the Lagrangian function $\mathcal{L}$ we introduced above, is that the given constraints mean that (on the constraint manifold) we have $\mathcal{L} = f + 0$ and therefore the stationary points of f and $\mathcal{L}$ coincide.

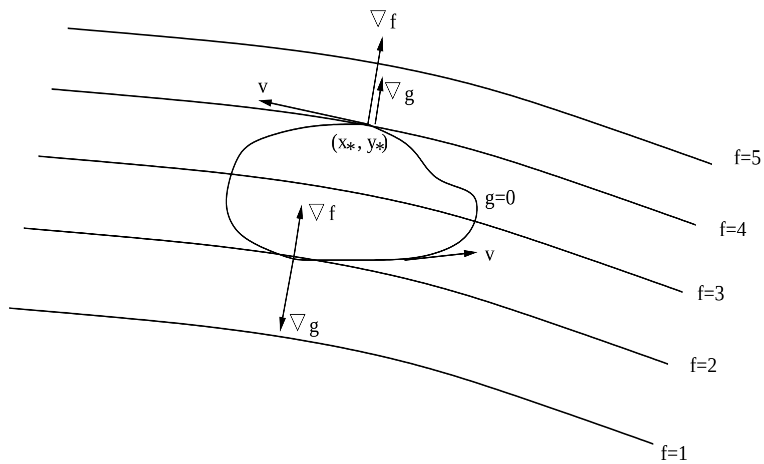


FIGURE 3. At the constrained extremum ∇f and ∇g are parallel. (This is a rough reproduction of the figure on page 199 of McCallum *et. al.* [**9**])

6.3. Geometric intuition. Suppose we wish to extremize (find a local maximum or minimum) the *objective function* $f(x, y)$ subject to the *constraint* $g(x, y) = 0$. We can think of this as follows. The graph $z = f(x, y)$ represents a surface in three dimensional (x, y, z) space, while the constraint represents a curve in the x-y plane to which our movements are restricted.

Constrained extrema occur at points where the contours of f are tangent to the contours of g (and can also occur at the endpoints of the constraint). This can be seen as follows. At any point (x, y) in the plane ∇f points in the direction of maximum increase of f and thus perpendicular to the level contours of f. Suppose that the vector v is tangent to the constraining curve $g(x, y) = 0$. If the directional derivative $f_v = \nabla f \cdot v$ is positive at some point, then moving in the direction of v means that f increases (if the directional derivative is negative then f decreases in that direction). Thus at the point (x_*, y_*) where f has a constrained extremum we must have $\nabla f \cdot v = 0$ and so both ∇f and ∇g are perpendicular to v and therefore parallel. Hence for some scalar parameter λ (the Lagrange multiplier) we have at the constrained extremum:

$$\nabla f = \lambda \nabla g \qquad \text{and} \qquad g = 0.$$

Notice that here we have three equations in three unknowns x, y, λ.

7. Constrained variational problems

A common optimization problem is to extremize a functional J with respect to paths y which are constrained in some way.

7.1. Constrained variational problem. We consider the following formulation.

Find the extrema of the functional

$$J(\boldsymbol{y}) := \int_a^b F(x, \boldsymbol{y}, \boldsymbol{y}_x)\, \mathrm{d}x,$$

where $\boldsymbol{y} = (y_1, \ldots, y_N) \in \mathbb{R}^N$ subject to the set of m constraints, for $k = 1, \ldots, m < N$:

$$G_k(x, \boldsymbol{y}) = 0.$$

To solve this constrained variational problem we use the method of Lagrange multipliers as follows. Note that the m constraint equations above imply

$$\int_a^b \lambda_k(x)\, G_k(x, \boldsymbol{y})\, \mathrm{d}x = 0,$$

for each $k = 1, \ldots, m$. Note that the λ_k's are the Lagrange multipliers, which with the constraints expressed in this integral form, can in general be functions of x. We now form the equivalent of the Lagrangian function which here is the functional

$$\begin{aligned} \tilde{J}(\boldsymbol{y}) &:= \int_a^b F(x, \boldsymbol{y}, \boldsymbol{y}_x)\, \mathrm{d}x + \sum_{k=1}^m \int_a^b \lambda_k(x)\, G_k(x, \boldsymbol{y})\Big)\, \mathrm{d}x \\ \Leftrightarrow \qquad \tilde{J}(\boldsymbol{y}) &= \int_a^b \Big(F(x, \boldsymbol{y}, \boldsymbol{y}_x) + \sum_{k=1}^m \lambda_k(x)\, G_k(x, \boldsymbol{y}) \Big)\, \mathrm{d}x. \end{aligned}$$

The integrand of this functional is

$$\tilde{F}(x, \boldsymbol{y}, \boldsymbol{y}_x, \boldsymbol{\lambda}) := F(x, \boldsymbol{y}, \boldsymbol{y}_x) + \sum_{k=1}^m \lambda_k(x)\, G_k(x, \boldsymbol{y}),$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$. We know from the classical variational problem that if $(\boldsymbol{y}, \boldsymbol{\lambda})$ extremize $\tilde{J}$ then necessarily they must satisfy the Euler–Lagrange equations:

$$\begin{aligned} \frac{\partial \tilde{F}}{\partial y_i} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial \tilde{F}}{\partial y_i'} \right) &= 0, \\ G_k(x, \boldsymbol{y}) &= 0, \end{aligned}$$

for $i = 1, \ldots, N$ and $k = 1, \ldots, m$. This is a system of *differential-algebraic equations*: the first set of relations are ordinary differential equations, while the constraints are algebraic relations.

7.2. Integral constraints. If the constraints are (already) in integral form so that we have

$$\int_a^b G_k(x, \boldsymbol{y})\, \mathrm{d}x = 0,$$

for $k = 1, \ldots, m$, then set

$$\begin{aligned} \tilde{J}(\boldsymbol{y}) &:= J(\boldsymbol{y}) + \sum_{k=1}^m \lambda_k \int_a^b G_k(x, \boldsymbol{y})\, \mathrm{d}x \\ &= \int_a^b \Big(F(x, \boldsymbol{y}, \boldsymbol{y}_x) + \sum_{k=1}^m \lambda_k\, G_k(x, \boldsymbol{y}) \Big)\, \mathrm{d}x, \end{aligned}$$

and proceed exactly as above. The Euler–Lagrange equations are the same.

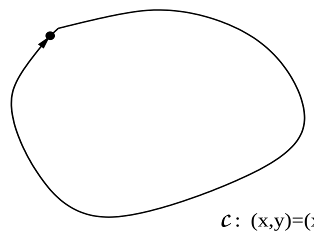


FIGURE 4. For the isoperimetrical problem, the closed curve $\mathcal{C}$ has a fixed length ℓ, and the goal is to choose the shape that maximizes the area it encloses.

7.3. Example (Dido's/isoperimetrical problem). The goal of this classical constrained variational problem is as follows. Find the shape of the closed curve, of a given fixed length ℓ, that encloses the maximum possible area.

Suppose that the curve is given in parametric coordinates $\big(x(\tau), y(\tau)\big)$ where the parameter $\tau \in [0, 2\pi]$. By Stokes' theorem, the area enclosed by a closed contour $\mathcal{C}$ is

$$\tfrac{1}{2}\oint_{\mathcal{C}} x\,\mathrm{d}y - y\,\mathrm{d}x.$$

Hence our goal is to extremize the functional

$$J(x,y) := \tfrac{1}{2}\int_0^{2\pi} (x\dot{y} - y\dot{x})\,\mathrm{d}\tau,$$

subject to the constraint

$$\int_0^{2\pi} \sqrt{\dot{x}^2 + \dot{y}^2}\,\mathrm{d}\tau = \ell.$$

Note that the constraint can be expressed in standard integral form as follows. Since ℓ is fixed we have

$$\int_0^{2\pi} \tfrac{1}{2\pi}\ell\,\mathrm{d}\tau = \ell.$$

Hence the constraint can be expressed in the form

$$\int_0^{2\pi} \sqrt{\dot{x}^2 + \dot{y}^2} - \tfrac{1}{2\pi}\ell\ \mathrm{d}\tau = 0.$$

To solve this variational constraint problem we use the method of Lagrange multipliers and form the functional

$$\begin{aligned}\tilde{J}(x,y) &:= J(x,y) + \lambda\bigg(\int_0^{2\pi} \sqrt{\dot{x}^2 + \dot{y}^2} - \tfrac{1}{2\pi}\ell\ \mathrm{d}\tau\bigg)\\ &= \int_0^{2\pi} \tilde{F}(x,y,\dot{x},\dot{y})\,\mathrm{d}\tau\,.\end{aligned}$$

where the integrand $\tilde{F}$ is given by

$$\tilde{F}(x,y,\dot{x},\dot{y}) = \tfrac{1}{2}(x\dot{y} - y\dot{x}) + \lambda(\dot{x}^2 + \dot{y}^2)^{\frac{1}{2}} - \lambda\tfrac{1}{2\pi}\ell.$$

Note there are two dependent variables x and y (here the parameter τ is the independent variable). By the theory above we know that the extremizing solution

(x, y) necessarily satisfies an Euler–Lagrange system of equations, which are the pair of ordinary differential equations

$$\frac{\partial \tilde{F}}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial \tilde{F}}{\partial \dot{x}}\right) = 0,$$
$$\frac{\partial \tilde{F}}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial \tilde{F}}{\partial \dot{y}}\right) = 0,$$

together with the integral constraint condition. Substituting the form for $\tilde{F}$ above, the pair of ordinary differential equations are

$$\tfrac{1}{2}\dot{y} - \frac{\mathrm{d}}{\mathrm{d}\tau}\left(-\tfrac{1}{2}y + \frac{\lambda\dot{x}}{(\dot{x}^2+\dot{y}^2)^{\frac{1}{2}}}\right) = 0,$$
$$-\tfrac{1}{2}\dot{x} - \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\tfrac{1}{2}x + \frac{\lambda\dot{y}}{(\dot{x}^2+\dot{y}^2)^{\frac{1}{2}}}\right) = 0.$$

Integrating both these equations with respect to τ we get

$$y - \frac{\lambda\dot{x}}{(\dot{x}^2+\dot{y}^2)^{\frac{1}{2}}} = c_2 \qquad \text{and} \qquad x - \frac{\lambda\dot{y}}{(\dot{x}^2+\dot{y}^2)^{\frac{1}{2}}} = c_1,$$

for arbitrary constants c_1 and c_2. Combining these last two equations reveals

$$(x-c_1)^2 + (y-c_2)^2 = \frac{\lambda^2\dot{y}^2}{\dot{x}^2+\dot{y}^2} + \frac{\lambda^2\dot{x}^2}{\dot{x}^2+\dot{y}^2} = \lambda^2.$$

Hence the solution curve is given by

$$(x-c_1)^2 + (y-c_2)^2 = \lambda^2,$$

which is the equation for a circle with radius λ and centre (c_1, c_2). The constraint condition implies $\lambda = \ell/2\pi$ and c_1 and c_2 can be determined from the initial or end points of the closed contour/path.

The isoperimetrical problem has quite a history. It was formulated in Virgil's poem the *Aeneid*, one account of the beginnings of Rome; see Wikipedia [**13**] or Montgomery [**11**]. Quoting from Wikipedia (Dido was also known as Elissa):

> Eventually Elissa and her followers arrived on the coast of North Africa where Elissa asked the local inhabitants for a small bit of land for a temporary refuge until she could continue her journeying, only as much land as could be encompassed by an oxhide. They agreed. Elissa cut the oxhide into fine strips so that she had enough to encircle an entire nearby hill, which was therefore afterwards named Byrsa "hide". (This event is commemorated in modern mathematics: The "isoperimetric problem" of enclosing the maximum area within a fixed boundary is often called the "Dido Problem" in modern calculus of variations.)

Dido found the solution—it her case a half-circle—and the semi-circular city of *Carthage* was founded.

8. Lagrangian dynamics

8.1. Newton's equations. Consider a system of N particles in three dimensional space, each with position vector $\boldsymbol{r}_i(t)$ for $i = 1, \ldots, N$. Note that each $\boldsymbol{r}_i(t) \in \mathbb{R}^3$ is a 3-vector. We thus need $3N$ coordinates to specify the system, this is the *configuration space*. Newton's 2nd law tells us that the equation of motion for the ith particle is

$$\dot{\boldsymbol{p}}_i = \boldsymbol{F}_i^{\text{ext}} + \boldsymbol{F}_i^{\text{con}},$$

for $i = 1, \ldots, N$. Here $\boldsymbol{p}_i = m_i \boldsymbol{v}_i$ is the linear momentum of the ith particle and $\boldsymbol{v}_i = \dot{\boldsymbol{r}}_i$ is its velocity. We decompose the total force on the ith particle into an external force $\boldsymbol{F}_i^{\text{ext}}$ and a *constraint* force $\boldsymbol{F}_i^{\text{con}}$. By external forces we imagine forces due to gravitational attraction or an electro-magnetic field, and so forth.

8.2. Holonomic constraints. By a constraint on a particles we imagine that the particle's motion is limited in some rigid way. For example the particle/bead may be constrained to move along a wire or its motion is constrained to a given surface. If the system of N particles constitute a rigid body, then the distances between all the particles are rigidly fixed and we have the constraint

$$|\boldsymbol{r}_i(t) - \boldsymbol{r}_j(t)| = c_{ij},$$

for some constants c_{ij}, for all $i, j = 1, \ldots, N$. All of these are examples of *holonomic constraints*.

Definition 1 (**Holonomic constraints**). *For a system of particles with positions given by $\boldsymbol{r}_i(t)$ for $i = 1, \ldots, N$, constraints that can be expressed in the form*

$$g(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N, t) = 0,$$

are said to be holonomic*. Note they only involve the configuration coordinates.*

We will *only* consider systems for which the constraints are holonomic. Systems with constraints that are non-holonomic are: gas molecules in a container (the constraint is only expressible as an inequality); or a sphere rolling on a rough surface without slipping (the constraint condition is one of matched velocities).

8.3. Degrees of freedom. Let us suppose that for the N particles there are m holonomic constraints given by

$$g_k(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N, t) = 0,$$

for $k = 1, \ldots, m$. The positions $\boldsymbol{r}_i(t)$ of all N particles are determined by $3N$ coordinates. However due to the constrants, the positions $\boldsymbol{r}_i(t)$ are not all independent. In principle, we can use the m holonomic constraints to eliminate m of the $3N$ coordinates and we would be left with $3N - m$ independent coordinates, i.e. the dimension of the configuration space is actually $3N - m$.

Definition 2 (**Degrees of freedom**). *The dimension of the configuration space is called the* number of degrees of freedom*, see Arnold* [**2**, Page 76].

Thus we can transform from the 'old' coordinates $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N$ to new *generalized coordinates* $q_1, \ldots, q_n$ where $n = 3N - m$:

$$\begin{aligned} \boldsymbol{r}_1 &= \boldsymbol{r}_1(q_1, \ldots, q_n, t), \\ &\vdots \\ \boldsymbol{r}_N &= \boldsymbol{r}_N(q_1, \ldots, q_n, t). \end{aligned}$$

8.4. D'Alembert's principle. We will restrict ourselves to systems for which the net work of the constraint forces is zero, i.e. we suppose

$$\sum_{i=1}^{N} \boldsymbol{F}_i^{\mathrm{con}} \cdot \mathrm{d}\boldsymbol{r}_i = 0,$$

for every small change $\mathrm{d}\boldsymbol{r}_i$ of the configuration of the system (for t fixed). If we combine this with Newton's 2nd law we see that we get

$$\sum_{i=1}^{N} (\dot{\boldsymbol{p}}_i - \boldsymbol{F}_i^{\mathrm{ext}}) \cdot \mathrm{d}\boldsymbol{r}_i = 0.$$

This is *D'Alembert's principle*. In particular, no forces of constraint are present.

The assumption that the constraint force does no net work is quite general. It is true in particular for holonomic constraints. For example, for the case of a rigid body, the internal forces of constraint do no work as the distances $|\boldsymbol{r}_i - \boldsymbol{r}_j|$ between particles is fixed, then $\mathrm{d}(\boldsymbol{r}_i - \boldsymbol{r}_j)$ is perpendicular to $\boldsymbol{r}_i - \boldsymbol{r}_j$ and hence perpendicular to the force between them which is parallel to $\boldsymbol{r}_i - \boldsymbol{r}_j$. Similarly for the case of the bead on a wire or particle constrained to move on a surface—the normal reaction forces are perpendicular to $\mathrm{d}\boldsymbol{r}_i$.

8.5. Lagrange's equations (from D'Alembert's principle). In his *Mécanique Analytique [1788]*, Lagrange sought a "coordinate-invariant expression for mass times acceleration", see Marsden and Ratiu [**10**, Page 231]. This lead to Lagrange's equations of motion. Consider the transformation to generalized coordinates

$$\boldsymbol{r}_i = \boldsymbol{r}_i(q_1, \ldots, q_n, t),$$

for $i = 1, \ldots, N$. If we consider a small increment in the displacements $\mathrm{d}\boldsymbol{r}_i$ then the corresponding increment in the work done by the external forces is

$$\sum_{i=1}^{N} \boldsymbol{F}_i^{\mathrm{ext}} \cdot \mathrm{d}\boldsymbol{r}_i = \sum_{i,j=1}^{N,n} \boldsymbol{F}_i^{\mathrm{ext}} \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j} \mathrm{d}q_j = \sum_{j=1}^{n} Q_j \, \mathrm{d}q_j.$$

Here we have set for $j = 1, \ldots, n$,

$$Q_j = \sum_{i=1}^{N} \boldsymbol{F}_i^{\mathrm{ext}} \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j},$$

and think of these as *generalized forces*.

Further, the change in kinetic energy mediated through the momentum (the first term in D'Alembert's principle), due to the increment in the displacements $\mathrm{d}\boldsymbol{r}_i$, is given by

$$\sum_{i=1}^{N} \dot{\boldsymbol{p}}_i \cdot \mathrm{d}\boldsymbol{r}_i = \sum_{i=1}^{N} m_i \dot{\boldsymbol{v}}_i \cdot \mathrm{d}\boldsymbol{r}_i = \sum_{i,j=1}^{N,n} m_i \dot{\boldsymbol{v}}_i \cdot \frac{\partial \boldsymbol{r}_i}{\partial q_j} \mathrm{d}q_j.$$

From the product rule we know that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\boldsymbol{v}_i\cdot\frac{\partial\boldsymbol{r}_i}{\partial q_j}\right) &\equiv \dot{\boldsymbol{v}}_i\cdot\frac{\partial\boldsymbol{r}_i}{\partial q_j}+\boldsymbol{v}_i\cdot\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\boldsymbol{r}_i}{\partial q_j}\right)\\ &\equiv \dot{\boldsymbol{v}}_i\cdot\frac{\partial\boldsymbol{r}_i}{\partial q_j}+\boldsymbol{v}_i\cdot\frac{\partial\boldsymbol{v}_i}{\partial q_j}.\end{aligned}$$

Also, by differentiating the transformation to generalized coordinates (keeping t fixed), we see that

$$\boldsymbol{v}_i\equiv\sum_{j=1}^{n}\frac{\partial\boldsymbol{r}_i}{\partial q_j}\,\dot{q}_j \qquad\Rightarrow\qquad \frac{\partial\boldsymbol{v}_i}{\partial\dot{q}_j}\equiv\frac{\partial\boldsymbol{r}_i}{\partial q_j}.$$

Using these last two identities we see that

$$\begin{aligned}\sum_{i=1}^{N}\dot{\boldsymbol{p}}_i\cdot\mathrm{d}\boldsymbol{r}_i &= \sum_{j=1}^{n}\left(\sum_{i=1}^{N}m_i\dot{\boldsymbol{v}}_i\cdot\frac{\partial\boldsymbol{r}_i}{\partial q_j}\right)\mathrm{d}q_j\\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{N}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(m_i\boldsymbol{v}_i\cdot\frac{\partial\boldsymbol{r}_i}{\partial q_j}\right)-m_i\boldsymbol{v}_i\cdot\frac{\partial\boldsymbol{v}_i}{\partial q_j}\right)\right)\mathrm{d}q_j\\ &= \sum_{j=1}^{n}\left(\sum_{i=1}^{N}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(m_i\boldsymbol{v}_i\cdot\frac{\partial\boldsymbol{v}_i}{\partial\dot{q}_j}\right)-m_i\boldsymbol{v}_i\cdot\frac{\partial\boldsymbol{v}_i}{\partial q_j}\right)\right)\mathrm{d}q_j\\ &= \sum_{j=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial}{\partial\dot{q}_j}\left(\sum_{i=1}^{N}\tfrac{1}{2}m_i|\boldsymbol{v}_i|^2\right)\right)-\frac{\partial}{\partial q_j}\left(\sum_{i=1}^{N}\tfrac{1}{2}m_i|\boldsymbol{v}_i|^2\right)\right)\mathrm{d}q_j.\end{aligned}$$

Hence if the kinetic energy is defined to be

$$T:=\sum_{i=1}^{N}\tfrac{1}{2}m_i|\boldsymbol{v}_i|^2,$$

then we see that D'Alembert's principle is equivalent to

$$\sum_{j=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial\dot{q}_j}\right)-\frac{\partial T}{\partial q_j}-Q_j\right)\mathrm{d}q_j=0.$$

Since the q_j for $j=1,\ldots,n$, where $n=3N-m$, are all independent, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial\dot{q}_j}\right)-\frac{\partial T}{\partial q_j}-Q_j=0,$$

for $j=1,\ldots,n$. If we now assume that the work done depends on the initial and final configurations only and not on the path between them, then there exists a potential function $V(q_1,\ldots,q_n)$ such that

$$Q_j=-\frac{\partial V}{\partial q_j}$$

for $j=1,\ldots,n$ (such forces are said to be *conservative*). If we define the *Lagrange function* or *Lagrangian* to be

$$L=T-V,$$

then we see that D'Alembert's principle is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0,$$

for $j = 1, \ldots, n$. These are known as *Lagrange's equations*. As already noted the n-dimensional subsurface of $3N$-dimensional space, on which the solutions to Lagrange's equations lie (where $n = 3N - m$), is called the *configuration space*. It is parameterized by the n generalized coordinates $q_1, \ldots, q_n$.

9. Hamilton's principle

9.1. Action. We consider mechanical systems with holonomic constraints and with all other forces conservative. Recall, we define the *Lagrange function* or *Lagrangian* to be

$$L = T - V,$$

where

$$T = \sum_{i=1}^{N} \tfrac{1}{2} m_i |\boldsymbol{v}_i|^2$$

is the total kinetic energy for the system, and V is its potential energy.

Definition 3 (**Action**). *If the Lagrangian L is the difference of the kinetic and potential energies for a system, i.e. $L = T - V$, we define the* action *A from time t_1 to t_2 to be the functional*

$$A(\boldsymbol{q}) := \int_{t_0}^{t_1} L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)\, \mathrm{d}t,$$

where $\boldsymbol{q} = (q_1, \ldots, q_n)$.

9.2. Hamilton's principle of least action. Hamilton [1834] realized that Lagrange's equations of motion were equivalent to a variational principle (see Marsden and Ratiu [**10**, Page 231]).

Theorem 2 (**Hamilton's principle of least action**). *The correct path of motion of a mechanical system with holonomic constraints and conservative external forces, from time t_1 to t_2, is a stationary solution of the action. Indeed, the correct path of motion $\boldsymbol{q} = \boldsymbol{q}(t)$, with $\boldsymbol{q} = (q_1, \ldots, q_n)$, necessarily and sufficiently satisfies Lagrange's equations of motion*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0,$$

for $j = 1, \ldots, n$.

Quoting from Arnold [**2**, Page 60], it is Hamilton's form of the principle of least action "because in many cases the action of $\boldsymbol{q} = \boldsymbol{q}(t)$ is not only an extremal but also a minimum value of the action functional".

9.3. Example (Kepler problem). Consider a particle of mass m moving in an inverse square law force field, $-\mu m/r^2$, such as a small planet or astroid in the gravitational field of a star or larger planet. The Lagrangian $L = T - V$ is given by

$$L(r, \dot{r}, \theta, \dot{\theta}) = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{\mu m}{r}.$$

From Hamilton's principle the equations of motion are given by Lagrange's equations, which here, taking the generalized coordinates to be $q_1 = r$ and $q_2 = \theta$, are the pair of ordinary differential equations

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} &= 0, \\ \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} &= 0, \end{aligned}$$

which on substituting the form for the Lagrangian above, become

$$\begin{aligned} m\ddot{r} - mr\dot{\theta}^2 + \frac{\mu m}{r^2} &= 0, \\ \frac{\mathrm{d}}{\mathrm{d}t}\big(mr^2\dot{\theta}\big) &= 0. \end{aligned}$$

9.4. Non-uniqueness of the Lagrangian. Two Lagrangian's L_1 and L_2 that differ by the total time derivative of any function of $\boldsymbol{q} = (q_1, \ldots, q_n)$ and t generate the same equations of motion. In fact if

$$L_2(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = L_1(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) + \frac{\mathrm{d}}{\mathrm{d}t}\big(f(\boldsymbol{q}, t)\big),$$

then for $j = 1, \ldots, n$ direct calculation reveals that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_2}{\partial \dot{q}_j}\right) - \frac{\partial L_2}{\partial q_j} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L_1}{\partial \dot{q}_j}\right) - \frac{\partial L_1}{\partial q_j}.$$

10. Constraints

Given a Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ for a system, suppose we realize the system has some constraints (so the q_j are not all independent).

10.1. Holonomic constraints. Suppose we have m holonomic constraints of the form

$$G_k(q_1, \ldots, q_n, t) = 0,$$

for $k = 1, \ldots, m < n$. We can now use the method of Lagrange multipliers with Hamilton's principle to deduce that the equations of motion are given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} &= \sum_{k=1}^{m} \lambda_k(t)\frac{\partial G_k}{\partial q_j}, \\ G_k(q_1, \ldots, q_n, t) &= 0, \end{aligned}$$

for $j = 1, \ldots, n$ and $k = 1, \ldots, m$. We call the quantities on the right above

$$\sum_{k=1}^{m} \lambda_k(t)\frac{\partial G_k}{\partial q_j},$$

the *generalized forces of constraint*.

10.2. Example (simple pendulum). Consider the motion of a simple pendulum bob of mass m that swings at the end of a light rod of length a. The other end is attached so that the rod and bob can swing freely in a plane. If g is the acceleration due to gravity, then the Lagrangian $L = T - V$ is given by

$$L(r, \dot{r}, \theta, \dot{\theta}) = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + mgr\cos\theta,$$

together with the constraint

$$r - a = 0.$$

We could just substitute $r = a$ into the Lagrangian, obtaining a system with one degree of freedom, and proceed from there. However, we will consider the system as one with two degrees of freedom, $q_1 = r$ and $q_2 = \theta$, together with a constraint $G(r) = 0$, where $G(r) = r - a$. Hamilton's principle and the method of Lagrange multipliers imply that the system evolves according to the pair of ordinary differential equations together with the algebraic constraint given by

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} &= \lambda \frac{\partial G}{\partial r},\\
\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} &= \lambda \frac{\partial G}{\partial \theta},\\
G &= 0.
\end{aligned}$$

Substituting the form for the Lagrangian above, the two ordinary differential equations together with the algebraic constraint become

$$\begin{aligned}
m\ddot{r} - mr\dot{\theta}^2 - mg\cos\theta &= \lambda,\\
\frac{\mathrm{d}}{\mathrm{d}t}\big(mr^2\dot{\theta}\big) + mgr\sin\theta &= 0,\\
r - a &= 0.
\end{aligned}$$

Note that the constraint is of course $r = a$, which implies $\dot{r} = 0$. Using this, the system of differential algebraic equations thus reduces to

$$ma^2\ddot{\theta} + mga\sin\theta = 0,$$

which comes from the second equation above. This first equation tells us that the Lagrange multiplier is given by

$$\lambda(t) = -ma\dot{\theta}^2 - mg\cos\theta.$$

The Lagrange multiplier has a physical interpretation, it is the normal reaction force, which here is the tension in the rod.

11. Hamiltonian dynamics

Hamiltonian. We consider mechanical systems that are holonomic and and conservative (or for which the applied forces have a generalized potential). For such a system we can construct a Lagrangian $L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$, where $\boldsymbol{q} = (q_1, \ldots, q_n)$, which is the difference of the total kinetic T and potential V energies. These mechanical systems evolve according to the n Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0,$$

for $j = 1, \ldots, n$. These are each second order ordinary differential equations and so the system is determined for all time once $2n$ initial conditions $\big(\boldsymbol{q}(t_0), \dot{\boldsymbol{q}}(t_0)\big)$

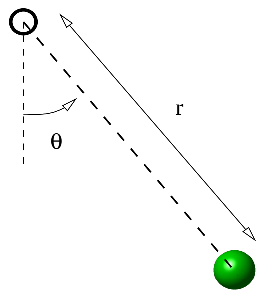


FIGURE 5. The mechanical problem for the simple pendulum, can be thought of as a particle of mass m moving in a vertical plane, that is constrained to always be a distance a from a fixed point. In polar coordinates, the position of the mass is (r, θ) and the constraint is $r = a$.

are specified (or n conditions at two different times). The state of the system is represented by a point $\boldsymbol{q} = (q_1, \ldots, q_n)$ in *configuration space*.

DEFINITION 4 (**Generalized momenta**). *We define the* generalized *momenta for a Lagrangian mechanical system to be*

$$p_j = \frac{\partial L}{\partial \dot{q}_j},$$

for $j = 1, \ldots, n$. *Note that we have* $p_j = p_j(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)$ *in general, where* $\boldsymbol{q} = (q_1, \ldots, q_n)$ *and* $\dot{\boldsymbol{q}} = (\dot{q}_1, \ldots, \dot{q}_n)$.

In terms of the generalized momenta, Lagrange's equations become, for $j = 1, \ldots, n$:

$$\dot{p}_j = \frac{\partial L}{\partial q_j}.$$

Further, *in principle*, we can solve the relations above which define the generalized momenta, to find functional expressions for the $\dot{q}_j$ in terms of q_i, p_i and t. In other words we can solve the relations defining the generalized momenta to find $\dot{q}_j = \dot{q}_j(\boldsymbol{q}, \boldsymbol{p}, t)$ where $\boldsymbol{q} = (q_1, \ldots, q_n)$ and $\boldsymbol{p} = (p_1, \ldots, p_n)$.

DEFINITION 5 (**Hamiltonian**). *We define the* Hamiltonian function *as the* Legendre transform *of the Lagrangian function, i.e. the Hamiltonian is defined by*

$$H(\boldsymbol{q}, \boldsymbol{p}, t) := \dot{\boldsymbol{q}} \cdot \boldsymbol{p} - L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t),$$

where $\boldsymbol{q} = (q_1, \ldots, q_n)$ *and* $\boldsymbol{p} = (p_1, \ldots, p_n)$ *and we suppose* $\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}}(\boldsymbol{q}, \boldsymbol{p}, t)$.

Note that in this definition we used the notation for the dot product

$$\dot{\boldsymbol{q}} \cdot \boldsymbol{p} = \sum_{j=1}^{n} \dot{q}_j \, p_j.$$

The *Legendre transform* is nicely explained in Arnold [**2**, Page 61] and Evans [**4**, Page 121]. In practice, as far as solving example problems herein, it simply involves the task of solving the equations for the generalized momenta above, to find $\dot{q}_j$ in terms of q_i, p_i and t.

Hamilton's equations of motion. From Lagrange's equation of motion, we can, using the definitions for the generalized momenta

$$p_j = \frac{\partial L}{\partial \dot{q}_j},$$

and Hamiltonian

$$H = \sum_{j=1}^{n} \dot{q}_j \, p_j - L,$$

above, deduce Hamilton's equations of motion.

THEOREM 3 (**Hamilton's equations of motion**). *With the Hamiltonian defined as the Legendre transform of the Lagrangian, Lagrange's equations of motion imply*

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i}, \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}, \end{aligned}$$

for $i = 1, \ldots, n$. These are Hamilton's canonical equations*, consisting of $2n$ first order equations of motion.*

PROOF. Using the chain rule and the definition for the generalized momenta we have

$$\frac{\partial H}{\partial p_i} = \sum_{j=1}^{n} \frac{\partial \dot{q}_j}{\partial p_i} p_j + \dot{q}_i - \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial p_i} \equiv \dot{q}_i,$$

and

$$\frac{\partial H}{\partial q_i} = \sum_{j=1}^{n} \frac{\partial \dot{q}_j}{\partial q_i} p_j - \frac{\partial L}{\partial q_i} - \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j} \frac{\partial \dot{q}_j}{\partial q_i} \equiv -\frac{\partial L}{\partial q_i}.$$

Now using Lagrange's equations, in the form $\dot{p}_j = \partial L / \partial q_j$, the last relation reveals

$$\dot{p}_i = -\frac{\partial H}{\partial q_i},$$

for $i = 1, \ldots, n$. Collecting these relations together, we see that Lagrange's equations of motion imply Hamilton's canonical equations as shown. □

Two further observations are also useful. Direct calculation reveals that

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t},$$

indicating that explicit dependence of H or L on t is present or absent together. Also, using the chain rule and Hamilton's equations we see that

$$\begin{aligned} & \frac{\mathrm{d}H}{\mathrm{d}t} = \sum_{i=1}^{n} \frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial p_i} \dot{p}_i + \frac{\partial H}{\partial t} \\ \Leftrightarrow \qquad & \frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t}. \end{aligned}$$

Hence if H does *not* explicitly depend on t (and therefore L also) then

$$H \text{ is a } \begin{cases} \text{constant of the motion,} \\ \text{conserved quantity,} \\ \text{integral of the motion.} \end{cases}$$

Hence the absense of explicit t dependence in the Hamiltonian H could serve as a more general definition of a conservative system, though in general H may not be the total energy. However for *simple mechanical systems* for which the kinetic energy $T = T(\boldsymbol{q}, \dot{\boldsymbol{q}})$ is a homogeneous quadratic function in $\dot{\boldsymbol{q}}$, and the potential $V = V(\boldsymbol{q})$, then the Hamiltonian H will be the total energy:

$$\sum_{i=1}^{n} \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2T \qquad \Rightarrow \qquad H = 2T - (T - V) = T + V.$$

12. Hamiltonian formulation

12.1. Summary. To construct Hamilton's canonical equations for a mechanical system proceed as follows:

(1) Choose your generalized coordinates $\boldsymbol{q} = (q_1, \ldots, q_n)$ and construct

$$L(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) = T - V.$$

(2) Define and compute the generalized momenta

$$p_i = \frac{\partial L}{\partial \dot{q}_i},$$

for $i = 1, \ldots, n$. Solve these relations to find $\dot{q}_i = \dot{q}_i(\boldsymbol{q}, \boldsymbol{p}, t)$.

(3) Construct and compute the Hamiltonian function

$$H = \sum_{j=1}^{n} \dot{q}_j \, p_j - L,$$

(4) Write down Hamilton's equations of motion

$$\begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i}, \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}, \end{aligned}$$

for $i = 1, \ldots, n$, and evaluate the partial derivatives of the Hamiltonian on the right.

12.2. Example (simple harmonic oscillator). The Lagrangian for the simple harmonic oscillator, which consists of a mass m moving in a quadatric potential field with characteristic coefficient k, is

$$L(x, \dot{x}) = \tfrac{1}{2} m \dot{x}^2 - \tfrac{1}{2} k x^2.$$

The corresponding generalized momentum is

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$$

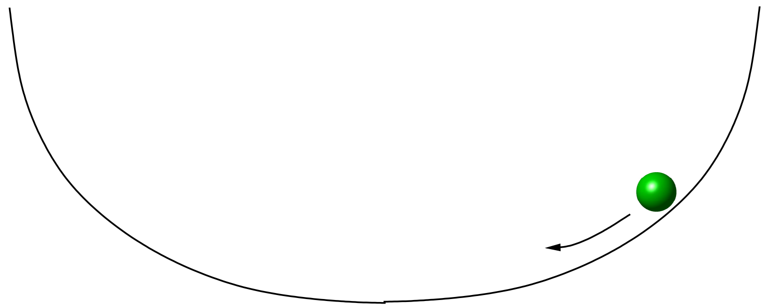


FIGURE 6. The mechanical problem for the simple harmonic oscillator consists of a particle moving in a quadratic potential field. As shown, we can think of a ball of mass m sliding freely back and forth in a vertical plane, without energy loss, inside a parabolic shaped bowl. The horizontal position $x(t)$ is its displacement.

which is the usual momentum. This implies $\dot{x} = p/m$ and so the Hamiltonian is given by

$$\begin{aligned} H(x,p) &= \dot{x}\,p - L(x,\dot{x}) \\ &= \frac{p}{m}\,p - \left(\tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}kx^2\right) \\ &= \frac{p^2}{m} - \left(\tfrac{1}{2}m\left(\frac{p}{m}\right)^2 - \tfrac{1}{2}kx^2\right) \\ &= \tfrac{1}{2}\frac{p^2}{m} + \tfrac{1}{2}kx^2. \end{aligned}$$

Note that is last expression is just the sum of the kinetic and potential energies and so H is the total energy. Hamilton's equations of motion are thus given by

$$\begin{aligned} \dot{x} &= \partial H/\partial p, \\ \dot{p} &= -\,\partial H/\partial x, \end{aligned} \qquad \Leftrightarrow \qquad \begin{aligned} \dot{x} &= p/m, \\ \dot{p} &= -\,kx. \end{aligned}$$

Note that combining these two equations, we get the usual equation for a harmonic oscillator: $m\ddot{x} = -kx$.

12.3. Example (Kepler problem). Recall the Kepler problem for a mass m moving in an inverse-square central force field with characteristic coefficient μ. The Lagrangian $L = T - V$ is

$$L(r,\dot{r},\theta,\dot{\theta}) = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{\mu m}{r}.$$

Hence the generalized momenta are

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \qquad \text{and} \qquad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}.$$

These imply $\dot{r} = p_r/m$ and $\dot{\theta} = p_\theta/mr^2$ and so the Hamiltonian is given by

$$\begin{aligned} H(r,\theta,p_r,p_\theta) &= \dot{r}\,p_r + \dot{\theta}p_\theta - L(r,\dot{r},\theta,\dot{\theta}) \\ &= \frac{1}{m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) - \left(\tfrac{1}{2}m\left(\frac{p_r^2}{m^2} + r^2\frac{p_\theta^2}{m^2r^4}\right) + \frac{\mu m}{r}\right) \\ &= \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) - \frac{\mu m}{r}, \end{aligned}$$

which is this case is also the total energy. Hamilton's equations of motion are

$$\begin{aligned} \dot{r} &= \partial H/\partial p_r, & & & \dot{r} &= p_r/m, \\ \dot{\theta} &= \partial H/\partial p_\theta, & & \Leftrightarrow & \dot{\theta} &= p_\theta/mr^2, \\ \dot{p_r} &= -\,\partial H/\partial r, & & & \dot{p_r} &= p_\theta^2/mr^3 - \mu m/r^2, \\ \dot{p_\theta} &= -\,\partial H/\partial \theta, & & & \dot{p_\theta} &= 0. \end{aligned}$$

Note that $\dot{p_\theta} = 0$, i.e. we have that p_θ is constant for the motion. This property corresponds to the *conservation of angular momentum*.

12.4. Remark. The Lagrangian $L = T - V$ may change its functional form if we use different variables $(\boldsymbol{Q}, \dot{\boldsymbol{Q}})$ instead of $(\boldsymbol{q}, \dot{\boldsymbol{q}})$, but its magnitude will not change. *However*, the functional form and magnitude of the Hamiltonian both depend on the generalized coordinates chosen. In particular, the Hamiltonian H may be conserved for one set of coordinates, but not for another.

12.5. Example (harmonic oscillator on a moving platform). Consider a mass-spring system, mass m and spring stiffness k, contained within a massless cart which is translating horizontally with a fixed velocity U—see Figure 7. The constant velocity U of the cart is maintained by an external agency. The Lagrangian $L = T - V$ for this system is

$$L(q,\dot{q},t) = \tfrac{1}{2}m\dot{q}^2 - \tfrac{1}{2}k(q - Ut)^2.$$

The resulting equation of motion of the mass is

$$m\,\ddot{q} = -k(q - Ut).$$

If we set $Q = q - Ut$, then the equation of motion is

$$m\,\ddot{Q} = -kQ.$$

Let us now consider the Hamiltonian formulation using two different sets of coordinates.

First, using the generalized coordinate q, the corresponding generalized momentum is $p = m\dot{q}$ and the Hamiltonian is

$$H(q,p,t) = \frac{p^2}{2m} + \frac{k}{2}(q - Ut)^2.$$

Here the Hamiltonian H *is* the total energy, but it is *not* conserved (there is an external energy input maintaining U constant).

Second, using the generalized coordinate Q, the Lagrangian $\tilde{L} = T - V$ is

$$\tilde{L}(Q,\dot{Q},t) = \tfrac{1}{2}m\dot{Q}^2 + m\dot{Q}U + \tfrac{1}{2}mU^2 - \tfrac{1}{2}kQ^2.$$

Here, the generalized momentum is $P = m\dot{Q} + mU$ and the Hamiltonian is

$$\tilde{H}(Q,P) = \frac{(P - mU)^2}{2m} + \frac{k}{2}Q^2 - \frac{m}{2}U^2.$$

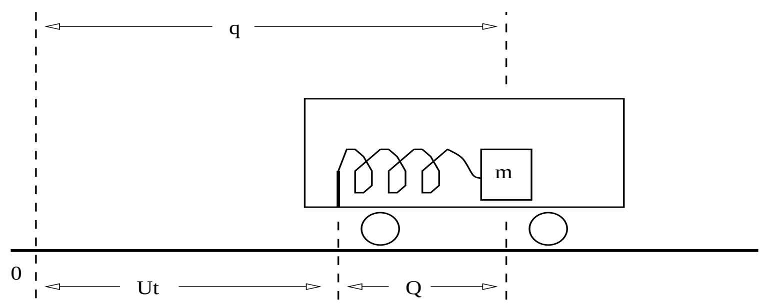


FIGURE 7. Mass-spring system on a massless cart.

Note that $\tilde{H}$ does not explicitly depend on t. Hence $\tilde{H}$ *is* conserved, but it is *not* the total energy.

13. Symmetries and conservation laws

13.1. Cyclic coordinates. We have already seen that

$$\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}.$$

Hence if the Hamiltonian does not depend explicitly on t, then it is a *constant* or *integral* of the motion; sometimes called Jacobi's integral. It *may* be the total energy. Further from the definition of the generalized momenta $p_i = \partial L/\partial \dot{q}_i$, Lagrange's equations, and Hamilton's equations for the generalized momenta, we have

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} = \frac{\partial L}{\partial q_i}.$$

From these relations we can see that if q_i is explicitly absent from the Lagrangian L, then it is explicitly absent from the Hamiltonian H, and

$$\dot{p}_i = 0.$$

Hence p_i is a conserved quantity, i.e. constant of the motion. Such a q_i is called *cyclic* or *ignorable*. Note that for such coordinates q_i, the transformation

$$t \to t + \Delta t,$$
$$q_i \to q_i + \Delta q_i,$$

leave the Lagrangian/Hamiltonian unchanged. This invariance signifies a fundamental *symmetry* of the system.

13.2. Example (Kepler problem). Recall, the Lagrangian $L = T - V$ for the Kepler problem is

$$L(r, \dot{r}, \theta, \dot{\theta}) = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + \frac{\mu m}{r},$$

and the Hamiltonian is

$$H(r, \theta, p_r, p_\theta) = \frac{1}{2m}\left(p_r^2 + \frac{p_\theta^2}{r^2}\right) - \frac{\mu m}{r}.$$

Note that L, H are independent of t and therefore H is conserved (and here it is the total energy). Further, L, H are independent of θ and therefore $p_\theta = mr^2\dot{\theta}$ is

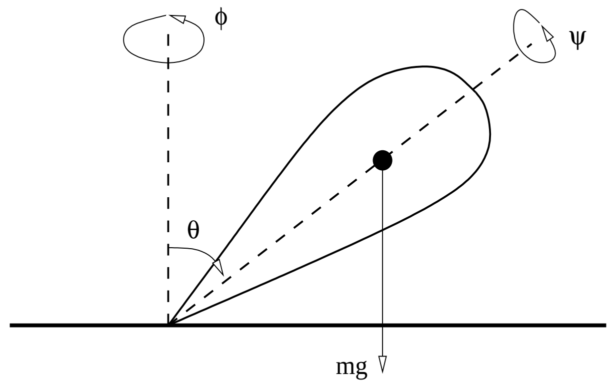


FIGURE 8. Simple axisymmetric top.

conserved also. The original problem has two degrees of freedom (r, θ). However, we have just established two integrals of the motion, namely H and p_θ. This system is said to be *integrable*.

13.3. Example (axisymmetric top). Consider a simple axisymmetric top of mass m with a fixed point on its axis of symmetry; see Figure 8. Suppose the centre of mass is a distance a from the fixed point, and the principle moments of inertia are $A = B \neq C$. We assume there are no torques about the symmetry or vertical axes. The configuration of the top is given in terms of the Euler angles (θ, ϕ, ψ) as shown in Figure 8. The Lagrangian $L = T - V$ is

$$L = \tfrac{1}{2}A(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta) + \tfrac{1}{2}C(\dot{\psi} + \dot{\phi}\cos\theta)^2 - mga\cos\theta.$$

The generalized momenta are

$$\begin{aligned} p_\theta &= A\dot{\theta}, \\ p_\phi &= A\dot{\phi}\sin^2\theta + C\cos\theta(\dot{\psi} + \dot{\phi}\cos\theta), \\ p_\psi &= C(\dot{\psi} + \dot{\phi}\cos\theta). \end{aligned}$$

Using these the Hamiltonian is given by

$$H = \frac{p_\theta^2}{2A} + \frac{(p_\phi - p_\psi\cos\theta)^2}{2A\sin^2\theta} + \frac{p_\psi^2}{2C} + mga\cos\theta.$$

We see that L, H are both independent of t, ψ and ϕ. Hence H, p_ψ and p_ϕ are conserved. Respectively, they represent the total energy, the angular momentum about the symmetry axis and the angular momentum about the vertical.

14. Exercises

Exercise (Euler–Lagrange alternative form). Show that the Euler–Lagrange equation is equivalent to

$$\frac{\partial F}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}x}\left(F - y_x\frac{\partial F}{\partial y_x}\right) = 0.$$

Exercise (soap film). A soap film is stretched between two rings of radius a which lie in parallel planes a distance $2x_0$ apart—the axis of symmetry of the two rings is coincident—see Figure 9.

(a) Explain why the surface area of the surface of revolution is given by

$$J(y) = 2\pi \int_{-x_0}^{x_0} y\sqrt{1+(y_x)^2}\,\mathrm{d}x,$$

where radius of the surface of revolution is given by $y = y(x)$ for $x \in [-x_0, x_0]$.

(b) Show that extremizing the surface area $J(y)$ in part (a) leads to the following ordinary differential equation for $y = y(x)$:

$$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2 = C^{-2}y^2 - 1$$

where C is an arbitrary constant.

(c) Use the substitution $y = C\cosh\theta$ and the identity $\cosh^2\theta - \sinh^2\theta = 1$ to show that the solution to the ordinary differential equation in part (b) is

$$y = C\cosh\big(C^{-1}(x+b)\big)$$

where b is another arbitrary constant. Explain why we can deduce that $b = 0$.

(d) Using the end-point conditions $y = a$ at $x = \pm x_0$, discuss the existence of solutions in relation to the ratio a/x_0.

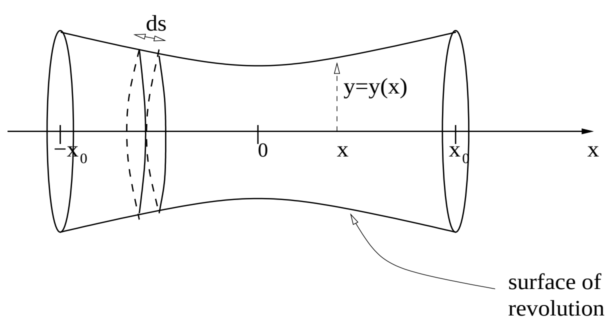


FIGURE 9. Soap film stretched between two concentric rings. The radius of the surface of revolution is given by $y = y(x)$ for $x \in [-x_0, x_0]$.

Exercise (hanging rope). We wish to compute the shape $y = y(x)$ of a uniform heavy hanging rope that is supported at the points $(-a, 0)$ and $(a, 0)$. The rope hangs so as to minimize its total potential energy which is given by the functional

$$\int_{-a}^{+a} \rho g y\sqrt{1+(y')^2}\,\mathrm{d}x,$$

where ρ is the mass density of the rope and g is the acceleration due to gravity. Suppose that the total length of the rope is fixed and given by ℓ, i.e. we have a constraint on the system given by

$$\int_{-a}^{+a} \sqrt{1+(y')^2}\,\mathrm{d}x = \ell.$$

(a) Use the method of Lagrange multipliers to show that the Euler–Lagrange equations in this case is

$$(y')^2 = c^2(y+\lambda)^2 - 1,$$

where c is an arbitrary constant and λ the Lagrange multiplier.

(b) Use the substitution $c(y+\lambda) = \cosh\theta$ to show that the solution to the ordinary differential equation in (a) is

$$y = c^{-1}\cosh\bigl(c(x+b)\bigr) - \lambda,$$

for some constant b. This problem is symmetric about the origin. Why does this imply $b = 0$?

(c) Use the boundary conditions that $y = 0$ at $x = \pm a$ and the constraint condition, respectively, to show that

$$\begin{aligned} c\,\lambda &= \cosh(ac), \\ c\,\ell &= 2\sinh(ac). \end{aligned}$$

Under what condition does the second equation have a real solution? What is the physical significance of this condition?

Exercise (spherical geodesic). This problem has some obvious applications to aviation: imagine flying from Paris to New York. The problem is to find the shortest distance between two points on a sphere, where the path between the two points must lie on the surface. It is natural to use spherical polar coordinates and to take the radial distance to be fixed, say $r = r_0$ with r_0 constant. The relationship between spherical and cartesian coordinates is

$$\begin{aligned} x &= r\,\sin\theta\,\cos\phi, \\ y &= r\,\sin\theta\,\sin\phi, \\ z &= r\,\cos\theta, \end{aligned}$$

where θ and ϕ are the latitude (measured from the north pole axis) and azimuthal angles, respectively. The goal is thus to minimize the total arclength from Paris, the point a on the sphere surface, to New York, the point b, i.e. to minimize the total arclength functional

$$\int_a^b \mathrm{d}s,$$

where s measures arclength on the surface of the sphere.

(a) Use spherical polar coordinates to show that we can express the total arclength functional in the form

$$J(\phi) := \int_{\theta_a}^{\theta_b} r_0 \sqrt{1 + \sin^2\theta \left(\frac{\mathrm{d}\phi}{\mathrm{d}\theta}\right)^2}\,\mathrm{d}\theta.$$

Here θ_a and θ_b are the latitude angles of the points a and b on the sphere surface. *Hint:* you will need to compute

$$\mathrm{d}s = \sqrt{(\mathrm{d}x)^2 + (\mathrm{d}y)^2 + (\mathrm{d}z)^2}.$$

(b) Show that the Euler–Lagrange equation is

$$\frac{r_0\sin^2\theta\cdot\phi'}{\bigl(1+\sin^2\theta\cdot(\phi')^2\bigr)^{\frac{1}{2}}} = c,$$

for some arbitrary constant c. Hence using the initial condition, deduce that the solution is an arc of a great circle.

Exercise (motion of relativistic particles). A particle with position $x(t) \in \mathbb{R}^3$ at time t prescribes a path that minimizes the functional

$$\int_{t_0}^{t_1} \left(-m_0 c^2 \sqrt{1 - \frac{\dot{x}^2}{c^2}} - U(x) \right) \mathrm{d}t,$$

subject to $x(t_0) = a$ and $x(t_1) = b$. Show that the equation of evolution of the particle is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(m \frac{\mathrm{d}x}{\mathrm{d}t} \right) = -\nabla U, \qquad \text{where} \qquad m = \frac{m_0}{\sqrt{1 - \frac{\dot{x}^2}{c^2}}}.$$

This is the equation of motion of a relativistic particle in an inertial system, under the influence of the force $-\nabla U$. Note that the *relativistic mass* m of the particle depends on its velocity. This mass goes to infinity if the particle approaches the velocity of light c.

Exercise (central force field). A particle of mass m moves under the central force $F = -\mathrm{d}V(r)/\mathrm{d}r$ in the spherical coordinate system such that

$$(x, y, z) = (r \cos\phi \sin\theta, r \sin\phi \sin\theta, r \cos\theta).$$

Using that the total kinetic energy of the system is given by

$$T = \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2),$$

show that the Lagrangian of the system in terms of r, θ and ϕ and their time derivatives is

$$L = \tfrac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2 + r^2 \sin^2\theta \dot{\phi}^2) - V(r).$$

From this Lagrangian, show that: (i) the quantity $m r^2 \dot{\phi} \sin^2\theta$ is a constant of the motion (call this h); (ii) the two remaining equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(r^2 \dot{\theta}\big) = \frac{h^2}{m^2 r^2} \cot\theta \,\mathrm{cosec}^2\theta,$$

$$\ddot{r} = r \dot{\theta}^2 + \frac{h^2}{m^2 r^3} \mathrm{cosec}^2\theta - \frac{1}{m} \frac{\mathrm{d}V}{\mathrm{d}r}.$$

Exercise (spherical pendulum). An inextensible string of length ℓ is fixed at one end and has a bob of mass m attached to the other. This bob swings freely under gravity, forming a spherical pendulum. Show that the Lagrangian for this system is given by

$$L(\theta, \phi, \dot{\theta}, \dot{\phi}) = \tfrac{1}{2} m \ell^2 (\dot{\theta}^2 + \sin^2\theta \dot{\phi}^2) + m g \ell \cos\theta,$$

where (θ, ϕ) are spherical angle coordinates centred at the fixed end of the pendulum where θ measures the angle to the vertical downward direction, and ϕ represents the azimuthal angle. Identify a conserved quantity. Write down the pair of Euler–Lagrange equations and use them to show, first that

$$m \ell^2 \sin^2\theta \cdot \dot{\phi} = \text{constant} = K,$$

and second, the equation of motion is given by

$$\ddot{\theta} = -\frac{g}{\ell} \sin\theta + \left(\frac{K}{m \ell^2} \right)^2 \cot\theta \mathrm{cosec}^2\theta.$$

Exercise (horizontal Atwood machine). A string of length ℓ was a mass m at each end passes through a hole in a horizontal frictionless plane. One mass moves horizontally on the plane, the other hangs vertically downwards.

(a) Explain why the Lagrangian for this system has the form

$$L(r,\theta,\dot{r},\dot{\theta}) = \tfrac{1}{2}m(2\dot{r}^2 + r^2\dot{\theta}^2) + mg(\ell - r),$$

where (r,θ) are the plane polar coordinates of the mass that moves on the plane.

(b) Show that the generalized momenta p_r and p_θ corresponding to the coordinates r and θ, respectively, are given by

$$p_r = 2m\dot{r} \qquad \text{and} \qquad p_\theta = mr^2\dot{\theta}.$$

(c) Using the results from part (b), show that the Hamiltonian for this system is given by

$$H(r,\theta,p_r,p_\theta) = \frac{p_r^2}{4m} + \frac{p_\theta^2}{2mr^2} - mg(\ell - r).$$

(d) Explain why the Hamiltonian H and generalized momentum p_θ are constants of the motion.

(e) In light of the information in part (d) above, we can express the Hamiltonian in the form

$$H(r,\theta,p_r,p_\theta) = \frac{p_r^2}{4m} + V(r),$$

where

$$V(r) = \frac{p_\theta^2}{2mr^2} - mg(\ell - r).$$

In other words, we can now think of the system as a particle moving in a potential given by $V(r)$.

Sketch V as a function of r. Describe qualitatively the different dynamics for the particle you might expect to see.

Exercise (pendulum with moving frictionless support). A pendulum system consists of a light rod, of length ℓ, with a mass M connected at one end that can slide freely along the x-axis, and a mass m at the other end that swings freely in the vertical plane containing the x-axis. If $\mu(t)$ represents the position of the mass M along the x-axis at time t, and $\theta(t)$ is the angle the rod makes with the vertical, show that the Lagrangian for this system is

$$L(\mu,\theta,\dot{\mu},\dot{\theta}) = \tfrac{1}{2}M\dot{\mu}^2 + \tfrac{1}{2}m(\ell^2\dot{\theta}^2 + \dot{\mu}^2 + 2\ell\dot{\mu}\dot{\theta}\cos\theta) + mg\ell\cos\theta.$$

Derive explicit expressions for the generalized momenta p_μ and p_θ conjugate to μ and θ, respectively. Explain why the Hamiltonian (no need to derive it) and p_μ are constants of the motion. Assume $p_\mu = 0$ (this corresponds to assuming that the centre of mass of the system is not uniformly translating in the x-direction) and show that

$$(M+m)\mu = -m\ell\sin\theta + A,$$

where A is an arbitrary constant.

Hence write down the Euler–Lagrange equation of motion for the angle $\theta = \theta(t)$.

Use your result/condition for $\mu = \mu(t)$ above to show that the position of the mass m at time t in Cartesian x and y coordinates is given by

$$x = \left(\frac{M\ell}{M+m}\right)\sin\theta + \frac{A}{M+m},$$
$$y = -\ell\cos\theta.$$

What is the shape of this curve with respect to the x and y coordinates?

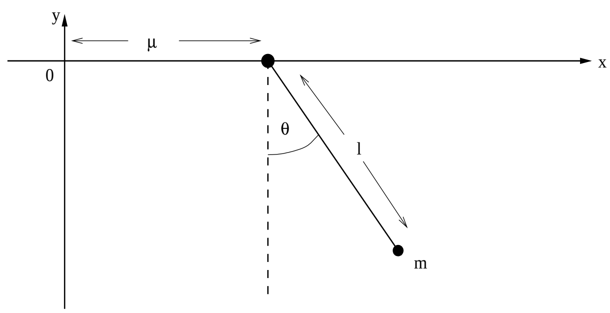


FIGURE 10. Pendulum with moving frictionless support.

Exercise (particle in a cone). A cone of semi-angle α has its axis vertical and vertex downwards, as in Figure 11. A point mass m slides without friction on the inside of the cone under the influence of gravity which acts along the negative z direction. The Lagrangian for the particle is

$$L(r,\theta,\dot r,\dot\theta) = \tfrac{1}{2}m\left(r^2\dot\theta^2 + \frac{\dot r^2}{\sin^2\alpha}\right) - \frac{mgr}{\tan\alpha},$$

where (r,θ) are plane polar coordinates as shown in Figure 11.

(a) Show that the generalized momenta p_r and p_θ corresponding to the coordinates r and θ, respectively, are given by

$$p_r = \frac{m\dot r}{\sin^2\alpha} \qquad \text{and} \qquad p_\theta = mr^2\dot\theta.$$

(b) Show that the Hamiltonian for this system is given by

$$H(r,\theta,p_r,p_\theta) = \frac{\sin^2\alpha}{2m}\,p_r^2 + \frac{p_\theta^2}{2mr^2} + \frac{mgr}{\tan\alpha}.$$

(c) Explain why the Hamiltonian H and generalized momentum p_θ are constants of the motion.

(d) In light of the information in part (c) above, we can express the Hamiltonian in the form

$$H(r,\theta,p_r,p_\theta) = \frac{\sin^2\alpha}{2m}\,p_r^2 + V(r),$$

where

$$V(r) = \frac{p_\theta^2}{2mr^2} + \frac{mgr}{\tan\alpha}.$$

In other words, we can now think of the system as a particle moving in a potential given by $V(r)$.

Sketch V as a function of r. Describe qualitatively the different dynamics for the particle you might expect to see.

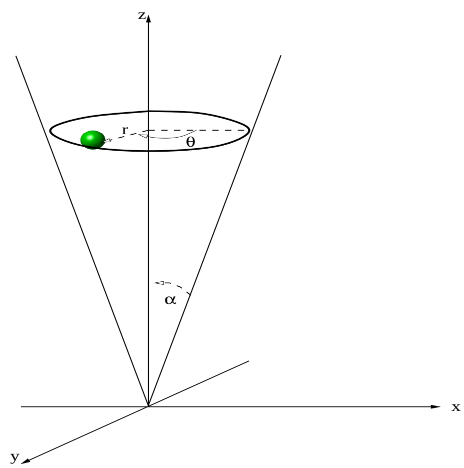


FIGURE 11. Particle sliding without friction inside a cone of semi-angle α, axis vertical and vertex downwards.

Exercise (spherical pendulum revisited). An inextensible string of length ℓ is fixed at one end and has a bob of mass m attached to the other. This bob swings freely under gravity, forming a spherical pendulum. Recall the form of the Lagrangian from the exercise above. Write down the Hamiltonian for this system, and identify a constant, J, of the motion (distinct from the Hamiltonian, which is also conserved). If θ is the angle the string makes with the vertical, show that the Hamiltonian can be written in the form

$$H = \frac{p_\theta^2}{2m\ell^2} + U(\theta).$$

where θ and p_θ are the canonical coordinates and $U(\theta)$ is the effective potential. Sketch $U(\theta)$ showing that it has a local minimum at θ_0, where θ_0 satisfies,

$$\left(\frac{J}{m\ell}\right)\cos\theta_0 = g\ell\sin^4\theta_0.$$

Briefly describe the possible behaviour of this system.

15. Notes

15.1. Multivariable variation. Suppose we are asked to find the surface $y = y(\boldsymbol{x})$ that, for $\boldsymbol{x} \in \Omega \subseteq \mathbb{R}^n$, extremises the functional

$$J(y) := \int_\Omega F(\boldsymbol{x}, y, \nabla y)\,\mathrm{d}\boldsymbol{x},$$

subject to y being fixed at the boundary $\partial\Omega$ of the domain Ω. Note here $\nabla y \equiv \nabla_{\boldsymbol{x}} y$ is simply the gradient of y, i.e. it is the vector of partial derivatives of y with respect to each of the components x_i $(i = 1, \dots, n)$ of $\boldsymbol{x}$:

$$\nabla y = \begin{pmatrix} \partial y/\partial x_1 \\ \vdots \\ \partial y/\partial x_n \end{pmatrix}.$$

Necessarily y satisfies an Euler–Lagrange equation which is a partial differential equation given by

$$\frac{\partial F}{\partial y} - \nabla \cdot \big(\nabla_{y_{\boldsymbol{x}}} F\big) = 0,$$

where '$\nabla\cdot$ ' is the usual divergence gradient operator with respect to $\boldsymbol{x}$, and

$$\nabla_{y_{\boldsymbol{x}}} F = \begin{pmatrix} \partial F/\partial y_{x_1} \\ \vdots \\ \partial F/\partial y_{x_n} \end{pmatrix},$$

where to keep the formula readable, with ∇y the usual gradient of y, we have set $y_{\boldsymbol{x}} \equiv \nabla y$ so that $y_{x_i} = (\nabla y)_i = \partial y/\partial x_i$ for $i = 1, \dots, n$.

Example (Laplace's equation). The variational problem here is to find the field $\psi(x_1, x_2, x_3)$, for $\boldsymbol{x} = (x_1, x_2, x_3) \in \Omega \subseteq \mathbb{R}^3$, that extremizes the mean-square gradient average

$$J(\psi) := \int_\Omega |\nabla \psi|^2 \,\mathrm{d}\boldsymbol{x}.$$

This is an example of case (3) above of multi-dimensional generalizations. In this case the integrand of the functional J is

$$F(\boldsymbol{x}, \psi, \nabla\psi) = |\nabla\psi|^2 \equiv (\psi_{x_1})^2 + (\psi_{x_2})^2 + (\psi_{x_3})^2.$$

Note that the integrand F depends on the partial derivatives of ψ only. Using the form for the Euler–Lagrange equation in case three above we get

$$\begin{aligned}
& & -\nabla_{\boldsymbol{x}} \cdot \big(\nabla_{y_{\boldsymbol{x}}} F\big) &= 0 \\
&\Leftrightarrow & \begin{pmatrix} \partial/\partial x_1 \\ \partial/\partial x_2 \\ \partial/\partial x_3 \end{pmatrix} \cdot \begin{pmatrix} \partial F/\partial \psi_{x_1} \\ \partial F/\partial \psi_{x_2} \\ \partial F/\partial \psi_{x_3} \end{pmatrix} &= 0 \\
&\Leftrightarrow & \frac{\partial}{\partial x_1}\big(2\psi_{x_1}\big) + \frac{\partial}{\partial x_2}\big(2\psi_{x_2}\big) + \frac{\partial}{\partial x_3}\big(2\psi_{x_3}\big) &= 0 \\
&\Leftrightarrow & \psi_{x_1x_1} + \psi_{x_2x_2} + \psi_{x_3x_3} &= 0 \\
&\Leftrightarrow & \nabla^2\psi &= 0.
\end{aligned}$$

This is *Laplace's equation* for ψ in the domain Ω; the solutions are called *harmonic functions*. Note that implicit in writing down the Euler–Lagrange partial differential equation above, we assumed that ψ was fixed at the boundary $\partial\Omega$, i.e. Dirichlet boundary conditions were specified.

Example (Helmholtz's equation). This is a constrained variational version of the problem that generated Laplace's equation. Find the field $\psi(x_1, x_2, x_3)$ that extremizes the functional

$$J(y) := \int_\Omega |\nabla\psi|^2 \,\mathrm{d}\boldsymbol{x},$$

subject to the constraint

$$\int_\Omega \psi^2 \,\mathrm{d}\boldsymbol{x} = 1.$$

This constraint corresponds to saying that the total energy is bounded and in fact renormalized to unity. Assuming zero boundary conditions, $\psi(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \partial\Omega$, the method of Lagrange multipliers implies ψ satisfies Helmholtz's equation in Ω:

$$-\nabla^2\psi = \lambda\psi.$$

Here λ is the Lagrange multiplier which also represents the eigenvalue parameter.

Example (stretched vibrating string). Suppose a string is tied between the two fixed points $x = 0$ and $x = \ell$. Let $y(x,t)$ be the small displacement of the string at position $x \in [0,\ell]$ and time $t > 0$ from the equilibrium position $y = 0$. If μ is the uniform mass per unit length of the string which is stretched to a tension K, the kinetic and potential energy of the string are given by

$$T = \tfrac{1}{2}\mu \int_0^\ell y_t^2 \,\mathrm{d}x, \qquad \text{and} \qquad V = K\left(\int_0^\ell \big(1 + (y_x)^2\big)^{1/2} \,\mathrm{d}x - \ell\right)$$

respectively, where subscripts indicate partial derivatives and the effect of gravity is neglected. If the oscillations of the string are quite small, we can replace the expression for V by

$$V = \frac{K}{2}\int_0^\ell y_x^2 \,\mathrm{d}x.$$

By applying Hamilton's principle to the action

$$A = \int_{t_1}^{t_2} (T - V)\,\mathrm{d}t$$

one can show that $y(x,t)$ satisfies the wave equation

$$y_{tt} = c^2\, y_{xx},$$

where $c^2 = K/\mu$.

15.2. Variation on manifolds. In general we may be faced with the problem of extremizing a functional which depends on several components which are themselves defined over a multi-dimensional domain $\Omega \subseteq \mathbb{R}^n$. In most applications though Ω will be the interval $\Omega = [a,b] \subset \mathbb{R}$ (there are some exceptions coming presently though). The functions $\boldsymbol{y}$ that are defined on Ω, which we vary to find extremizing solution, can in general be manifold valued, i.e. $\boldsymbol{y} \in \mathbb{M}$ where $\mathbb{M}$ is a smooth N-dimensional submanifold of a higher dimensional Euclidean space. For example, the motions or paths may be restricted or confined to lie on the surface of the sphere for example, in which case $\mathbb{M} = S^2$. Hence we have the maps

$$\begin{aligned} \boldsymbol{y}&\colon \Omega \to \mathbb{M}, \\ \boldsymbol{y}&\colon \boldsymbol{x} \mapsto \boldsymbol{y}(\boldsymbol{x}), \end{aligned} \qquad \text{and} \qquad \begin{aligned} F&\colon \Omega \times \mathrm{T}\mathbb{M} \to \mathbb{R}, \\ F&\colon \big(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{x}), \nabla\boldsymbol{y}(\boldsymbol{x})\big) \mapsto F\big(\boldsymbol{x}, \boldsymbol{y}(\boldsymbol{x}), \nabla\boldsymbol{y}(\boldsymbol{x})\big), \end{aligned}$$

where $\mathrm{T}\mathbb{M}$ is the tangent bundle of $\mathbb{M}$. The functional to be extremized is thus a map $J\colon \mathbb{Y} \to \mathbb{R}$, where if $C^2\big(\Omega;\mathbb{M}\big)$ is the space of twice continuously differentiable functions on Ω, then

$$\mathbb{Y} := \big\{\boldsymbol{y} \in C^2\big(\Omega;\mathbb{M}\big) \colon \boldsymbol{y} \in \partial\Omega \text{ given}\big\}.$$

Note that if $\Omega = [a,b]$ then the boundary conditions in $\mathbb{Y}$ are equivalent to $\boldsymbol{y}(a)$ and $\boldsymbol{y}(b)$ being specified.

15.3. Non-holonomic constraints. Mechanical systems with certain types of non-holonomic constraints can also be treated, in particular constraints of the form

$$\sum_{j=1}^{n} A(\boldsymbol{q},t)_{kj}\dot{q}_j + b_k(\boldsymbol{q},t) = 0,$$

for $k = 1, \ldots, m$, where $\boldsymbol{q} = (q_1, \ldots, q_n)$. Note the assumption is that these equations are not integrable, in particular not exact, otherwise the constraints would be holonomic.

15.4. Non-conservative forces. If the system has non-conservative forces it may still be possible to find a *generalized potential* function V such that

$$Q_j = -\frac{\partial V}{\partial q_j} + \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial V}{\partial \dot{q}_j}\right),$$

for $j = 1, \ldots, n$. From such potentials we can still deduce Lagrange's equations of motion. Examples of such generalized potentials are velocity dependent potentials due to electro-magnetic fields, for example the Lorentz force on a charged particle.

15.5. Lagrangian mechanics on manifolds. the configuration space is an n-dimensional submanifold $\mathbb{M}$ of $3N$-dimensional Euclidean space—it is the manifold consisting of the intersections of the holonomic constraint hypersurfaces in $\mathbb{R}^{3N}$. The Lagrangian is map $L\colon \mathrm{T}\mathbb{M} \times [a,b] \to \mathbb{R}$, i.e. from the tangent bundle of $\mathbb{M}$ (and the interval $[a,b]$ when it is non-autonomous) to $\mathbb{R}$. For more details see Arnold [**2**, Chapter 4] and Marsden and Ratiu [**10**, Chapter 8].

15.6. Noether's theorem. To accept only those symmetries which leave the Lagrangian unchanged is needlessly restrictive. When searching for conservation laws (integrals of the motion), we can in general consider transformations that leave the *action integral* 'invariant enough' so that we get the same equations of motion. This is the idea underlying (Emmy) Noether's theorem; see Arnold [**2**, Page 88].

Acknowledgements

These notes are dedicated to Dr. Frank Berkshire whose enthusiam and knowledge inspired me as a student. The lecture notes herein, are largely based on the first half of Frank's Dynamics course that I attended as a third year undergraduate at Imperial College in the Autumn term of 1989.

References

[1] Abraham, R. and Marsden, J.E. 1978 *Foundations of mechanics*, Second edition, Westview Press.
[2] Arnold, V.I. 1978 *Mathematical methods of classical mechanics*, Graduate Texts in Mathematics.
[3] Berkshire, F. 1989 *Lecture notes on Dynamics*, Imperial College Mathematics Department.
[4] Evans, L.C. 1998 *Partial differential equations* Graduate Studies in Mathematics, Volume 19, American Mathematical Society.
[5] Keener, J.P. 2000 *Principles of applied mathematics: transformation and approximation*, Perseus Books.
[6] Krantz, S.G. 1999 *How to teach mathematics*, Second Edition, American Mathematical Society.
[7] Kibble, T.W.B. and Berkshire, F.H. 1996 *Classical mechanics*, Fourth Edition, Longman.
[8] Goldstein, H. 1950 *Classical mechanics*, Addison–Wesley.
[9] McCallum, W.G. *et. al.* 1997 *Multivariable calculus*, Wiley.

[10] Marsden, J.E. and Ratiu, T.S. 1999 *Introduction to mechanics and symmetry*, Second edition, Springer.
[11] Montgomery, R. 2002 *A tour of subRiemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, Volume 91, American Mathematical Society.
[12] http://universe-review.ca/F15-particle.htm
[13] http://en.wikipedia.org/wiki/Dido_(Queen_of_Carthage)

Department of Mathematics, Heriot-Watt University; 11th January 2010