**H. Hergert**
Facility for Rare Isotope Beams
and Department of Physics & Astronomy

# PHY422/820: Classical Mechanics

FS 2020

Worksheet #4 (Sep 21 – 25)

September 28, 2020

## 1 Preparation

- Lemos, Section 2.5–2.7
- Goldstein, Sections 2.5–2.7

## 2 Lagrange Multipliers

### 2.1 Premise

Practical applications of variational principles often involve the optimization of a function in the presence of constraints: For instance, the examples below describe how we can find the largest possible rectangle that we can inscribe into an ellipse, or the minima of a two-dimensional function along a specific curve in the two-dimensional plane. Another example would be to optimize a probability distribution $p(x_i)$ of discrete events $x_i$ while ensuring that $\sum_i p(x_i) = 1$

To account for such constraints, we couple them to the function that is to be optimized using **Lagrange multipliers** and solve the optimization problem with these multipliers as additional variables. For instance, if we have a function $f(x_1, \ldots, x_n)$ and constraints $g_i(x_1, \ldots, x_n)$ linking the variables $x_i$, we define

$$\Lambda(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(x_1, \ldots, x_n) + \sum_i \lambda_i g_i(x_1, \ldots, x_n)\,, \tag{1}$$

and determine the extrema of $\Lambda$ as a function of the $x_i$ and $\lambda_i$ (see below). You can view the additional terms in the optimization problem as penalties that we impose for violating the constraint conditions.

The Lagrange multiplier techniques can be readily applied in variational calculus as well. In that case, the constraints $g_i(y, y', x), y = (y_1, \ldots, y_n)$ are coupled to the Lagrangian $f$ of our functional,

$$\widetilde{f}(y, y', x, \lambda) = f(y, y', x) + \sum_i \lambda_i g_i(y, x)\,, \tag{2}$$

and we again solve the variational problem as if $\lambda = (\lambda_1, \ldots, \lambda_m)$ were additional variables. Note that we had to require that the constraints are holonomic, i.e., integrable, in this expression.

## 2.2 The Second-Derivative Test

In the following examples, we are dealing with the optimization of functions of multiple variables, and we would like to determine whether an extremum that we found is a maximum, minimum, or a saddle point. For functions of one variable, this determination is readily made using the second derivative, but in multiple dimensions, the problem becomes more complex, and the presence of constraints adds additional complications.

Without constraints, we know that the extrema of a function $f(x_1, \ldots, x_n)$ can are identified by a vanishing gradient,

$$\vec{\nabla} f = \left( \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)^T = 0\,. \tag{3}$$

To determine the nature of the extrema, we can look at the **Hessian**, the matrix of second partial derivatives, which is defined by

$$\mathcal{H}[f] = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}. \tag{4}$$

If the Hessian is **positive definite**, then the extremum is an isolated minimum, if it is **negative definite**, an isolated maximum, otherwise we will have a saddle point (which may be degenerate, i.e., we could have curves, planes etc. of saddle points, depending on the dimension of $\vec{x}$). Otherwise, the test is inconclusive.

A symmetric matrix like the Hessian is positive definite if for all vectors $\vec{x}$ the product $\vec{x}^T \mathcal{H} \vec{x}$ is strictly positive. This is equivalent to stating that all eigenvalues $\mu_i$ of the matrix are positive. The Hessian is **positive semi-definite** if $\vec{x}^T \mathcal{H} \vec{x} \geq 0$ or $\mu_i \geq 0$. Negative definite and negative semi-definite matrices are defined analogously. A matrix that has positive and negative eigenvalues is called **indefinite**.

In problems with constraints, we use what is called the **bordered Hessian**, which is the Hessian of the Lagrangian $\Lambda$ (Eq. (1)):

$$\begin{aligned}\mathcal{H}[\Lambda] &= \begin{pmatrix} \frac{\partial^2 \Lambda}{\partial \lambda_1^2} & \cdots & \frac{\partial^2 \Lambda}{\partial \lambda_1 \lambda_m} & \frac{\partial^2 \Lambda}{\partial \lambda_1 \partial x_1} & \cdots & \frac{\partial^2 \Lambda}{\partial \lambda_1 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \Lambda}{\partial \lambda_m \lambda_1} & \cdots & \frac{\partial^2 \Lambda}{\partial \lambda_m^2} & \frac{\partial^2 \Lambda}{\partial \lambda_m \partial x_1} & \cdots & \frac{\partial^2 \Lambda}{\partial \lambda_m \partial x_n} \\ \frac{\partial^2 \Lambda}{\partial x_1 \lambda_1} & \cdots & \frac{\partial^2 \Lambda}{\partial x_1 \partial \lambda_m} & \frac{\partial^2 \Lambda}{\partial x_1^2} & \cdots & \frac{\partial^2 \Lambda}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \Lambda}{\partial x_n \lambda_1} & \cdots & \frac{\partial^2 \Lambda}{\partial x_n \partial \lambda_m} & \frac{\partial^2 \Lambda}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 \Lambda}{\partial x_n^2} \end{pmatrix} \\ &= \begin{pmatrix} 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} & \frac{\partial^2 \Lambda}{\partial x_1^2} & \cdots & \frac{\partial^2 \Lambda}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \Lambda}{\partial x_n \lambda_1} & \cdots & \frac{\partial^2 \Lambda}{\partial x_n \partial \lambda_m} & \frac{\partial^2 \Lambda}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 \Lambda}{\partial x_n^2} \end{pmatrix}.\end{aligned} \tag{5}$$

Here, we have used that $\Lambda$ is linear in the $\lambda_i$, and that $\frac{\partial \Lambda}{\partial \lambda_i} = g_i$.

We cannot apply the criterion for definiteness for the unconstrained Hessian: $\mathcal{H}[\Lambda]$ acts on $m+n$-dimensional vectors of the form $(\lambda_1, \ldots, \lambda_m, x_1, \ldots, x_n)^T$, and any vectors that only has components in the first $m$ entries will be mapped to 0. (Equivalently, the bordered Hessian is guaranteed to have at least $m$ eigenvalues that are zero.) Instead, the second-derivative test relies on sign conditions on the sequence of **leading principal minors**.

The **principal matrices** of an $n \times n$ matrix are obtained by deleting $k$ rows and columns, which we can do in $\binom{n}{k}$ ways in general. The **leading principal matrix** $P_k$ of dimension $k$ is obtained by deleting the *last* $n-k$ rows and columns. Let us denote the determinant of this matrix as $D_k$. For the bordered Hessian, we will then have

$$D_1 = \det P_1 = \det \begin{pmatrix} \frac{\partial^2 \Lambda}{\partial \lambda_1^2} \end{pmatrix} = 0\,, \tag{6}$$

$$D_2 = \det P_2 = \det \begin{pmatrix} \frac{\partial^2 \Lambda}{\partial \lambda_1^2} & \frac{\partial^2 \Lambda}{\partial \lambda_1 \lambda_2} \\ \frac{\partial^2 \Lambda}{\partial \lambda_2 \lambda_1} & \frac{\partial^2 \Lambda}{\partial \lambda_2^2} \end{pmatrix} = 0\,,$$

$$\vdots$$

$$D_{m+1} = \det P_{m+1} = \det \begin{pmatrix} 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \frac{\partial g_c}{\partial x_1} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_c}{\partial x_1} & \frac{\partial^2 \Lambda}{\partial x_1^2} \end{pmatrix}\,, \tag{7}$$

$$\vdots$$

$$D_{m+n} = \det P_{m+n} = \det \mathcal{H}[\Lambda]\,. \tag{8}$$

The test consists of evaluating the sequence of leading principal minors $D_k$ for $k = \min(2m+1, m+n), \ldots, m+n$, that is, the first $2m$ minors are neglected, and if $2m+1 > m+n$, then we only consider the determinant of the bordered Hessian itself. This means we consider $n-m$ minors for a system with $n-m$ degrees of freedom ($n$ coordinates, $m$ constraints). The *sufficient conditions* for local maxima and minima can then be stated as follows:

- An extremum is a local maximum if the signs of the $D_k$, evalutated at the point of interest, alternate and the first $D_k$ in the sequence has the sign $(-1)^{m+1}$.
- An extremum is a local minimum if all $D_k$ have the sign $(-1)^m$.
- Other outcomes of the test are inconclusive.

(In the unconstrained case $m = 0$, these conditions reduce to another criterion for the positive or negative definiteness of the unconstrained Hessian.)

### 2.3 Example: Rectangle Inscribed in an Ellipse

Consider the problem of inscribing the largest possible rectangle into an ellipse with semi-major axes $a$ and $b$, which is defined by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\,. \tag{9}$$

The area of the rectangle with corners $(-x,-y), (-x,y), (x,-y), (x,y)$ is given by the function

$$A(x,y) = (2x)(2y) = 4xy\,, \quad x, y \geq 0\,. \tag{10}$$

We introduce a Lagrange multiplier and couple the constraint (9) to this function,

$$\widetilde{A}(x,y,\lambda) = A(x,y) + \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1\right), \tag{11}$$

and proceed to maximize $\widetilde{A}$. The partial derivatives are

$$\frac{\partial\widetilde{A}}{\partial x} = 4y + 2\lambda\frac{x}{a^2}, \tag{12}$$

$$\frac{\partial\widetilde{A}}{\partial y} = 4x + 2\lambda\frac{y}{b^2}, \tag{13}$$

$$\frac{\partial\widetilde{A}}{\partial \lambda} = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1. \tag{14}$$

Let us first consider the boundaries of our domain, which are given by $(x=0,y)$ and $(x,y=0)$, respectively. The cases corresponds to the limit in which the rectangle turns into a line. While the are vanishes, the constraint can still be satisfied. On the first boundary, for example, we have

$$\frac{\partial\widetilde{A}}{\partial \lambda} = \frac{0^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \quad \Rightarrow \quad y = \pm b. \tag{15}$$

Our system of equations now becomes

$$\left.\frac{\partial\widetilde{A}}{\partial x}\right|_{x=0} = 0 \stackrel{!}{=} 0, \tag{16}$$

$$\left.\frac{\partial\widetilde{A}}{\partial y}\right|_{x=0} = 2\lambda\frac{y}{b^2} \stackrel{!}{=} 0, \tag{17}$$

$$\left.\frac{\partial\widetilde{A}}{\partial \lambda}\right|_{x=0} = \frac{y^2}{b^2} - 1 \stackrel{!}{=} 0. \tag{18}$$

The first equation is trivially true, while the second and third equations are solved by

$$x_1 = 0, \quad y_1 = b, \quad \lambda_1 = 0. \tag{19}$$

Analogously, we obtain

$$x_2 = a, \quad y_2 = 0, \quad \lambda_2 = 0. \tag{20}$$

on the other boundary. Mathematically, these are global minima of $\widetilde{A}(x,y,\lambda)$ and $A(x,y)$.

Now consider the interior of the domain, where $x>0, y>0$. We start with Eq. (13) and obtain

$$4x + 2\lambda\frac{y}{b^2} = 0 \quad \Rightarrow \quad x = -\lambda\frac{y}{2b^2}. \tag{21}$$

Note that this implies $\lambda < 0$, since $y$ must be positive. Plugging this solution into Eq. (12), we obtain

$$4y - \lambda^2\frac{y}{a^2b^2} = 0 \quad \Rightarrow \quad y\left(4 - \frac{\lambda^2}{a^2b^2}\right) = 0. \tag{22}$$

The solutions to this equation are $y=0$, which we considered separately above, and $\lambda = -2ab$ ($\lambda = 2ab$ is ruled out).

Let's proceed with $\lambda = -2ab$, which yields

$$x = -\lambda \frac{y}{2b^2} = y \frac{a}{b^2}\,. \tag{23}$$

Using this relation in Eq. (14), we have

$$\frac{y^2 a^2}{b^2 a^2} + \frac{y^2}{b^2} - 1 = 0 \quad \Rightarrow \quad y = \frac{b}{\sqrt{2}}\,. \tag{24}$$

Thus, the extremum in the domain's interior is

$$E_3: \quad x_3 = \frac{a}{\sqrt{2}}\,, \quad y_3 = \frac{b}{\sqrt{2}}\,, \quad \lambda_3 = -2ab\,. \tag{25}$$

The area of the resulting rectangle is

$$A\left(x_3, y_3\right) = 2ab\,, \tag{26}$$

and the constrained stationary point is obviously a maximum.

Although the nature of the extremum is clear, let us still perform the second-derivative test for practice. The bordered Hessian reads

$$\mathcal{H}(x,y,\lambda) = \begin{pmatrix} \frac{\partial^2 \widetilde{A}}{\partial \lambda^2} & \frac{\partial^2 \widetilde{A}}{\partial \lambda \partial x} & \frac{\partial^2 \widetilde{A}}{\partial \lambda \partial y} \\ \frac{\partial^2 \widetilde{A}}{\partial x \partial \lambda} & \frac{\partial^2 \widetilde{A}}{\partial x^2} & \frac{\partial^2 \widetilde{A}}{\partial x \partial y} \\ \frac{\partial^2 \widetilde{A}}{\partial y \partial \lambda} & \frac{\partial^2 \widetilde{A}}{\partial x \partial y} & \frac{\partial^2 \widetilde{A}}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 0 & \frac{2x}{a^2} & \frac{2y}{b^2} \\ \frac{2x}{a^2} & \frac{2\lambda}{a^2} & 4 \\ \frac{2y}{b^2} & 4 & \frac{2\lambda}{b^2} \end{pmatrix}. \tag{27}$$

We have $n = 2$ variables and $m = 1$ constraint, so we need to look at the leading principal minors of the bordered Hessian for $k = \min(2m+1, m+n), \ldots, m+n$. Here, this means $k = 3$, and the principal minor is the determinant of $\mathcal{H}$ itself. At the extremum $E_3$, we have

$$\det \mathcal{H}\left(\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}, -2ab\right) = \det \begin{pmatrix} 0 & \frac{\sqrt{2}}{a} & \frac{\sqrt{2}}{b} \\ \frac{\sqrt{2}}{a} & -\frac{4b}{a} & 4 \\ \frac{\sqrt{2}}{b} & 4 & \frac{-4a}{b} \end{pmatrix} = \frac{32}{ab} > 0\,, \tag{28}$$

matching the requirement for a constrained local maximum that at the sign should be $(-1)^{m+1} = +1$.

## 2.4 Example: Extrema of the Mexican Hat Potential

### Extrema Without Constraints

Let us consider the function

$$V(x,y) = -40(x^2+y^2) + (x^2+y^2)^2\,, \tag{29}$$

an example of a so-called *mexican hat* potential that is frequently used to discuss symmetry breaking phenomena in physics (see Fig. 1).

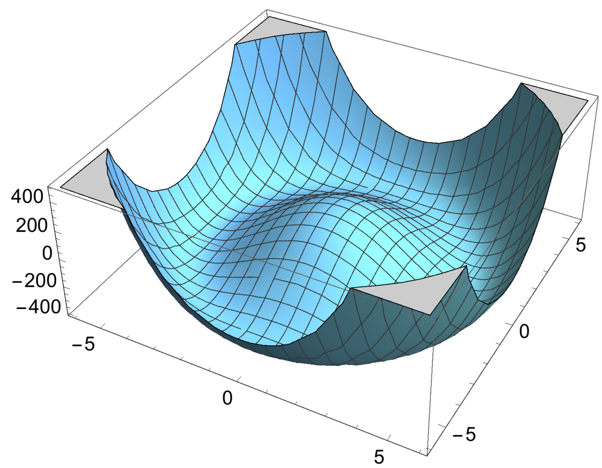

Figure 1: Mexican hat potential.

Let us first compute the extrema of $V(x,y)$. The partial derivatives with respect to the variables are

$$\frac{\partial V}{\partial x} = -80x + 4(x^2+y^2)x\,, \tag{30}$$

$$\frac{\partial V}{\partial y} = -80y + 4(x^2+y^2)y\,. \tag{31}$$

The extrema are obtained by solving

$$0 = -80x + 4(x^2+y^2)x\,, \tag{32}$$

$$0 = -80y + 4(x^2+y^2)y\,. \tag{33}$$

The solutions of Eq. (32) are

$$x_1 = 0 \tag{34}$$

and all points on a circle in the $xy$-plane with radius $r=\sqrt{20}$.

$$4(x^2+y^2) - 80 = 0 \quad \Rightarrow \quad x^2+y^2 = 20\,. \tag{35}$$

From Eq. (33), we again obtain the definition of the circle, as well as

$$y_1 = 0\,. \tag{36}$$

By inspecting the figure, we see that $(x_1,y_1)=(0,0)$ is a **local maximum**, while points on the circle are **degenerate global minima** in radial direction, and **saddle points** in any direction tangential to the circle. If we do not have a figure at hand, or a more complicated function, we can compute the Hessian and check its definiteness:

$$\mathcal{H}(x,y) = \begin{pmatrix} \frac{\partial^2 V}{\partial x^2} & \frac{\partial^2 V}{\partial x \partial y} \\ \frac{\partial^2 V}{\partial x \partial y} & \frac{\partial^2 V}{\partial y^2} \end{pmatrix} = \begin{pmatrix} -80 + 4(x^2+y^2) + 8x^2 & 8xy \\ 8xy & -80 + 4(x^2+y^2) + 8y^2 \end{pmatrix}\,. \tag{37}$$

For $(x_1,y_1)=(0,0)$, $\mathcal{H}(0,0)$ is diagonal and we can read off the doubly degenerate eigenvalue $h_{1/2}=-80$. Since all the eigenvalues are negative, the $\mathcal{H}(0,0)$ is negative definite, the point is a local maximum.

For any point on the ring, the Hessian becomes

$$\mathcal{H}(x,y)|_C = \begin{pmatrix} -80 + 4\cdot 20 + 8x^2 & 8xy \\ 8xy & -80 + 4\cdot 20 + 8y^2 \end{pmatrix} = \begin{pmatrix} 8x^2 & 8xy \\ 8xy & 8y^2 \end{pmatrix}\,. \tag{38}$$

The eigenvalues are $h_3=0$ and $h_4 = 8(x^2+y^2)=160$, so the matrix is indefinite. Parameterizing the circle by $(r,\phi)=(\sqrt{20},\phi)$, we find the expected result that $\vec{e}_3=\vec{e}_\phi=(-\sin\phi,\cos\phi)^T$ are eigenvectors associated with $h_3$, and $\vec{e}_4=\vec{e}_r=(\cos\phi,\sin\phi)^T$ are eigenvectors associated with $h_4$.

### Extrema Under Constraints

Let us now determine the extrema of the mexican hat potential under the constraint

$$y = x \quad \Rightarrow \quad f(x,y) \equiv y - x = 0\,. \tag{39}$$

We couple the constraint to the function $V(x,y)$ with a Langrange multiplier $\lambda$:

$$\widetilde{V}(x,y,\lambda) = V(x,y) + \lambda f(x,y)\,. \tag{40}$$

The partial derivatives of $\widetilde{V}$ and the resulting conditions for extrema are

$$\frac{\partial \widetilde{V}}{\partial x} = -80x + 4(x^2+y^2)x - \lambda = 0\,, \tag{41}$$

$$\frac{\partial \widetilde{V}}{\partial y} = -80y + 4(x^2+y^2)y + \lambda = 0\,, \tag{42}$$

$$\frac{\partial \widetilde{V}}{\partial \lambda} = y - x = 0\,. \tag{43}$$

Plugging the solution of (43) obviously into either of the other equations, e.g., Eq. (41), we can determine $\lambda$:

$$-80x + 8x^3 - \lambda = 0 \quad \Rightarrow \quad \lambda = 8x^3 - 80x\,. \tag{44}$$

The remaining equation now becomes

$$-80x + 8x^3 + 8x^3 - 80x = 16x(x^2-10) = 0\,, \tag{45}$$

with the solutions

$$x_1 = 0\,, x_{2/3} = \pm\sqrt{10}\,, \tag{46}$$

so our extrema are formally given by

$$(x_1,y_1,\lambda_1) = (0,0,0), \quad (x_2,y_2,\lambda_2) = (\sqrt{10},\sqrt{10},0), \quad (x_3,y_3,\lambda_3) = (-\sqrt{10},-\sqrt{10},0)\,. \tag{47}$$

(**Note: in general, the Lagrange multiplier(s) need not vanish at the stationary points; see App. 2.3.**)

The bordered Hessian of $\widetilde{V}$ now reads

$$\mathcal{H}(x,y,\lambda) = \begin{pmatrix} 0 & -1 & 1 \\ -1 & -80 + 8x^2 + 4(x^2+y^2) & 8xy \\ 1 & 8xy & -80 + 8x^2 + 4(x^2+y^2)\,, \end{pmatrix} \tag{48}$$

and it must be evaluated at the different stationary points. We have $n = 2$ variables and $m = 1$ constraints, so the second-derivative test requires us to look at the principal minors of $\mathcal{H}$ for $k = \min(2m+1, m+n), \ldots, m+n$. In the present case, the range gives us $k = 3$ and the principal minor is again $\mathcal{H}$ itself, just like in the previous example.

At $(x,y,\lambda) = (0,0,0)$, we have

$$\det \mathcal{H}(0,0,0) = \det \begin{pmatrix} 0 & -1 & 1 \\ -1 & -80 & 0 \\ 1 & 0 & -80 \end{pmatrix} = 160 > 0\,, \tag{49}$$

so the sign matches $(-1)^{m+1} = +1$ and we have a constrained maximum . At $(\sqrt{10},\sqrt{10},0)$ and $(-\sqrt{10},-\sqrt{10},0)$ we obtain the same bordered Hessian, and its determinant is

$$\det \mathcal{H}(\pm\sqrt{10},\pm\sqrt{10},0) = \det \begin{pmatrix} 0 & -1 & 1 \\ -1 & 80 & 80 \\ 1 & 80 & 80 \end{pmatrix} = -320 < 0\,, \tag{50}$$

implying that these points are constrained minima. This makes sense, of course: We precisely get the two global minima and the local maxima of $V(x,y)$ that are compatible with the constraint $y - x = 0$.

# 3 Constraints Revisited: Lagrange Equations of the First and Second Kind

## 3.1 Lagrangians with Explicit Constraints

One of the motivations for introducing the Lagrangian formalism was to achieve a description of the dynamics of a mechanical system that automatically incorporates constraints on the motion, so that we can directly work with the true degrees of freedom. This is certainly advantageous if we want so determine the solutions for a specific system as efficiently as possible, maybe even with pen and paper. However, there are good reasons why might want to have the option to treat some (or all) of the constraints explicitly:

- We might want to *know* the constraint forces in order to design a mechanical apparatus that can withstand the stresses that occur during its motion.
- We might want to implement general-purpose software that is capable of dealing with arbitrary user-defined dynamical systems.

In these cases, we can use the Lagrange multiplier technique described above to couple explicit constraints to the Lagrangian of our system.

The Lagrange formalism with explicit constraint is sometimes referred to as the **Lagrange formalism of the first kind**, as opposed to the **Lagrange formalism of the second kind**, which uses all constraints to eliminate as many degrees of freedom as possible.

### Holonomic Constraints

Any holonomic constraints can be directly coupled to the Lagrangian by defining

$$\widetilde{L}(q,\dot{q},t,\lambda) = L(q,\dot{q},t) + \sum_{a=1}^{c} \lambda_a f_a(q,t)\,, \tag{51}$$

with $q = (q_1, \ldots, q_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_c)$. Since the constraint equations are independent of the generalized velocity, the modified Lagrangian is compatible with d'Alembert's principle and the Principle of Least Action. Both approaches will yield the Lagrange equations

$$\frac{d}{dt}\frac{\partial \widetilde{L}}{\partial \dot{q}_j} - \frac{\partial \widetilde{L}}{\partial q_j} = 0 \quad \Leftrightarrow \quad \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \sum_{a=1}^{c} \lambda_a \frac{\partial f_a}{\partial q_j}\,, \tag{52}$$

and

$$\frac{d}{dt}\underbrace{\frac{\partial \widetilde{L}}{\partial \dot{\lambda}_a}}_{=0} - \frac{\partial \widetilde{L}}{\partial \lambda_a} = 0 \quad \Leftrightarrow \quad \frac{\partial \widetilde{L}}{\partial \lambda_a} = f_a(\vec{q},t) = 0\,. \tag{53}$$

Effectively, we are treating the Lagrange multipliers as additional generalized coordinates in the holonomic case. Note that

$$\lambda_a \frac{\partial f_a}{\partial q_j} = \lambda_a \sum_{i=1}^{A} \frac{\partial \vec{r}_i}{\partial q_j}\frac{\partial f}{\partial \vec{r}_i}\,, \tag{54}$$

where $\frac{\partial f}{\partial \vec{r}_i}$ is another way of writing $\vec{\nabla}_i f_a$ that brings out more clearly that we are dealing with an application of the chain rule. Here we recognize the constraint forces in the form we discussed previously, $\vec{F}_{C,a} = \lambda_a \vec{\nabla}_i f_a$, and how they get projected onto the configuration manifold to yield the generalized constraint force. We also see that the $\lambda$ factors we previously introduced are the Lagrange multipliers we use to couple the constraints to the Lagrangian of our system.

#### Nonholonomic Constraints

Unfortunately, nonholonomic constraints cannot be simply coupled to the Lagrangian directly. The treatment of such constraints is a long-standing problem that has led to several false starts and continuing misconceptions[1] — a discussion of the issues can be found, for instance, in a series of papers by M. Flannery [2, 3, 4]. Interest in the development of an efficient treatment for nonholonomic constraints has been revived in part by modern robotics research, where it could significantly aid in the construction of a robot's control software.

In the context of D'Alembert's principle, nonholonomic constraints produce generalized constraint forces that need to be added explicitly to the right-hand side of the Lagrange equations,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \sum_{b=1}^{s} \mu_b \frac{\partial g_b}{\partial \dot{q}_j} \,. \tag{55}$$

To distinguish them from explicit holonomic constraints, we denote the Lagrange multipliers of such forces by $\mu$ instead of $\lambda$. If we tried to add $\sum_b \mu_b g_b(q, \dot{q})$ to the Lagrangian, we would obtain

$$\mu_b \left[\frac{d}{dt}\left(\frac{\partial g_b}{\partial \dot{q}_j}\right) - \frac{\partial g_b}{\partial q_j}\right] , \tag{56}$$

which does not match the correct form in Eq. (55). This inability to couple the nonholonomic constraints to $L(q, \dot{q}, t)$ also prevent us from defining an action functional that is consistent with the equations of motion (55).

**Box 3.1: Lagrange Equations with Explicit Constraints**

The Lagrangre equations of the first kind for $r$ holonomic and $s$ nonholonomic, velocity-dependent constraints (no inequalities!) are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = \sum_{a=1}^{r} \lambda_a \frac{\partial f_a}{\partial q_j} + \sum_{b=1}^{s} \mu_b \frac{\partial g_b}{\partial \dot{q}_j} \,. \tag{I3.1-1}$$

## 3.2 Examples

### 3.2.1 Spinning a Mass on a String

Consider a mass that is being spun around on a string of length $l$ with a constant angular velocity $\omega$, parallel to the ground. In polar coordinates, the trajectory of the mass is

$$\vec{r} = r\vec{e}_r \,, \tag{57}$$

and therefore

$$\dot{\vec{r}} = \dot{r}\vec{e}_r + r\dot{\vec{e}}_\phi = \dot{r}\vec{e}_r + r\dot{\phi}\vec{e}_\phi \,. \tag{58}$$

The kinetic energy is

$$T = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\phi}^2\right) \,. \tag{59}$$

[1] Note that the discussion of nonholonomic constraints in the context of the principle of least action in [1] is erroneous as well - the specific example turns out to be correct if one foregoes the attempt to include the variational principle, but instead directly builds the constraint into the Lagrange equations of the first kind, Eqs. (I3.1-1), relying on d'Alembert's principle.

Since the motion occurs in a plane that is parallel to the ground, the potential is constant, and we can choose the origin of our coordinate system such that $V = 0$. The constraints of the motion are

$$r - l = 0\,, \quad \phi - \omega t = 0\,, \tag{60}$$

so we can define the modified Lagrangian

$$\widetilde{L} = \frac{1}{2} m \left(\dot{r}^2 + r^2\dot{\phi}^2\right) + \lambda_r\,(r - l) + \lambda_\phi\,(\phi - \omega t)\ . \tag{61}$$

The Lagrange equations of the first kind are now given by

$$\frac{d}{dt}\frac{\partial \widetilde{L}}{\partial \dot{r}} - \frac{\partial \widetilde{L}}{\partial r} = m\ddot{r} - mr\dot{\phi}^2 - \lambda_r = 0\,, \tag{62}$$

$$\frac{d}{dt}\frac{\partial \widetilde{L}}{\partial \dot{\phi}} - \frac{\partial \widetilde{L}}{\partial \phi} = \frac{d}{dt}\left(mr^2\dot{\phi}\right) - \lambda_\phi = 0\,, \tag{63}$$

$$-\frac{\partial \widetilde{L}}{\partial \lambda_r} = r - l = 0\,, \tag{64}$$

$$-\frac{\partial \widetilde{L}}{\partial \lambda_\phi} = \phi - \omega t = 0\,. \tag{65}$$

From the first two equations, we obtain the Lagrange multipliers

$$\lambda_r = m\ddot{r} - mr\dot{\phi}^2 \tag{66}$$

and

$$\lambda_\phi = \frac{d}{dt} mr^2\dot{\phi}\,. \tag{67}$$

The right-hand side of this equation shows that $\lambda_\phi$ is the time derivative of an angular momentum, i.e., a torque. We can now plug in the constraint equations for $\phi$ and $r$, and obtain

$$\lambda_\phi = \frac{d}{dt} ml^2\omega = 0\,, \tag{68}$$

i.e., the torque vanishes. This is consistent: $\omega =$ const. implies conservation of the angular momentum around the $z$ axis, so there cannot be an external torque acting on the system.

From the Lagrange equation (66) we obtain

$$\lambda_r = \underbrace{m\ddot{r}}_{=0} - ml\dot{\phi}^2 = -ml\omega^2 \ = \text{const.} \tag{69}$$

We see that $\lambda_r$ has the dimensions of a force, and it points in negative $r$ direction, i.e., toward the hub of the circle — this means that $\lambda_r$ is the **centripetal force** for the circular motion, which is provided by the string tension in the present case.

### 3.2.2 Cylinder Rolling Down an Inclinded Plane

Let us consider a cylinder of mass $M$ and radius $R$ that is rolling down an inclined plane (see figure). Its kinetic energy can be written as the sum of the center-of-mass translation and the cylinder's rotation around the symmetry axis through the center of mass (we will discuss this later in the course):

$$T = \frac{1}{2}M\dot{s}^2 + \frac{1}{2}I\dot{\phi}^2\,, \tag{70}$$

where $s$ is the distance the cylinder has rolled down the plane, $\phi$ is the rotation angle, and $I = \frac{1}{2}MR^2$.

We now assume that the cylinder is **rolling without slipping**, which means that

$$ds = Rd\phi\,. \tag{71}$$

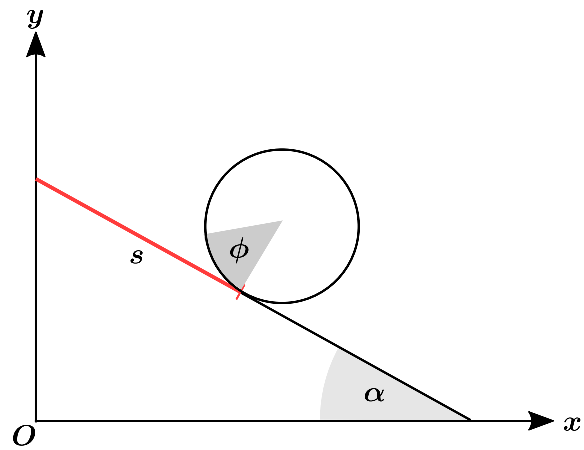


Figure 2: Cylinder rolling without slipping on an inclined plane.

This is a nonholonomic constraint that connects the generalized velocities,

$$g(\dot{s}, \dot{\phi}) = \dot{s} - R\dot{\phi} = 0\,. \tag{72}$$

The potential energy of the cylinder is given by

$$V = V_0 - Mgs\sin\alpha\,, \tag{73}$$

where $\alpha$ is the inclination angle. (Note that we have not explicitly considered the radius of the cylinder in the potential energy, which would just cause another constant offset that has no impact on the dynamics.) Thus, the Lagrangian is given by

$$L = \frac{1}{2}M\dot{s}^2 + \frac{1}{4}M\dot{\phi}^2 - V_0 + Mgs\sin\alpha\,. \tag{74}$$

Since the constraint is nonholonomic, we cannot couple it to the Lagrangian but add it to the Lagrange equation (see Eq. (I3.1-1)):

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{s}} - \frac{\partial L}{\partial s} = \mu\frac{\partial g}{\partial \dot{s}} \quad\Rightarrow\quad M\ddot{s} - Mg\sin\alpha = \mu\,, \tag{75}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = \mu\frac{\partial g}{\partial \dot{\phi}} \quad\Rightarrow\quad \frac{1}{2}MR^2\ddot{\phi} = -\mu R\,. \tag{76}$$

From the second equation and the time derivative of the constraint (72), we obtain

$$\mu = -\frac{1}{2}MR\ddot{\phi} = -\frac{1}{2}M\ddot{s}\,. \tag{77}$$

We plug this into the first equation of motion, which now becomes

$$M\ddot{s} - Mg\sin\alpha = -\frac{1}{2}M\ddot{s}\,, \tag{78}$$

and rearranging, we have

$$\frac{3}{2}\ddot{s} - g\sin\alpha = 0 \quad\Rightarrow\quad \ddot{s} = \frac{2}{3}g\sin\alpha\,. \tag{79}$$

(Note that the acceleration is in growing $s$ direction, i.e., downward along the plane, so the signs are correct.) Thus, the general solution to the equation of motion is given by

$$s(t) = s_0 + v_0 t + \frac{1}{2}(\frac{2}{3} g \sin\alpha) t^2 = s_0 + v_0 t + \frac{1}{3}(g \sin\alpha) t^2 \,, \tag{80}$$

and assuming the cylinder starts rolling from rest at the top of the plane, we have

$$s(t) = \frac{1}{3}(g \sin\alpha) t^2 \,. \tag{81}$$

We conclude the discussion of this example by noting that the nonholonomic constraint discussed here is actually **integrable**, i.e., a holonomic constraint in disguise. Since we only have two variables $s$ and $\phi$, we can integrate the constraint equation Eq. (72),

$$ds = R d\phi \quad \Rightarrow \quad s = R\phi + s_0 \,, \tag{82}$$

and use it to eliminate the $s$ or $\phi$ in favor of the other coordinate. In the next example, we will also consider rolling without slipping, but we will *not* be able to integrate the nonholonomic constraint.

#### 3.2.3 Disk Rolling on a Plane

We consider a disk that is rolling without slipping in a horizontal plane while remaining in an upright position, so that the rotational axis remains parallel to the plane (see figure). As generalized coordinates, we can choose

- the $x$ and $y$ coordinates of the disk's center of mass, which also correspond to the support point in the plane,
- the angle $\theta$ between the rotational axis and the $x$ axis, and
- the angle $\phi$ characterizing the disk's rotation around its axis.

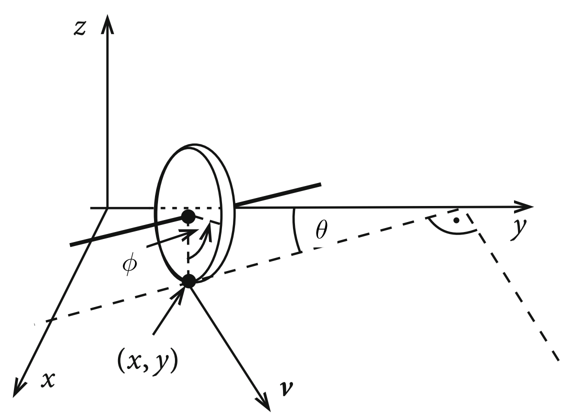


Figure 3: Disk rolling without slipping on a plane.

The condition for rolling without slipping relates the change of the center of mass's position to the change in the angle $\phi$ due to the rotation, just like in the previous example:

$$ds = R d\phi \,, \tag{83}$$

or in terms of velocities

$$\dot{s} = |\vec{v}| = R\dot{\phi} \,. \tag{84}$$

We also constrain $\vec{v}$ to be perpendicular to the rotational axis, which implies

$$\dot{x} = \dot{s}\cos\theta \,, \tag{85}$$

$$\dot{y} = \dot{s}\sin\theta \,. \tag{86}$$

The directional constraints can be combined with the rolling condition to yield

$$g_1(\vec{q}, \dot{\vec{q}}) \equiv \dot{x} - R\dot{\phi}\cos\theta = 0 \,, \tag{87}$$

$$g_2(\vec{q}, \dot{\vec{q}}) \equiv \dot{y} - R\dot{\phi}\sin\theta = 0\,. \tag{88}$$

The Lagrangian for the disk's unconstrained motion is the sum of the translational and rotational terms (see discussion of rigid bodies later in the course),

$$L = \frac{1}{2}M(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I_\phi\dot{\phi}^2 + \frac{1}{2}I_\theta\dot{\theta}^2\,, \tag{89}$$

where $I_\phi$ is the moment of inertia for the rotation around the disk's horizontal axis, and $I_\theta$ the moment of inertia for rotation around the vertical axis through the disk's center of mass and support point in the plane.

Using the Lagrangian (89) and the constraints (87), (88), we obtain the following Lagrange equations of the first kind:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = \sum_{b=1}^{2}\mu_b\frac{\partial g_b}{\partial \dot{x}} \qquad \Rightarrow \qquad M\ddot{x} = \mu_1\,, \tag{90}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{y}} - \frac{\partial L}{\partial y} = \sum_{b=1}^{2}\mu_b\frac{\partial g_b}{\partial \dot{y}} \qquad \Rightarrow \qquad M\ddot{y} = \mu_2\,, \tag{91}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} - \frac{\partial L}{\partial \phi} = \sum_{b=1}^{2}\mu_b\frac{\partial g_b}{\partial \dot{\phi}} \qquad \Rightarrow \qquad I_\phi\ddot{\phi} = -\mu_1 R\cos\theta - \mu_2 R\sin\theta\,, \tag{92}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} - \frac{\partial L}{\partial \theta} = \sum_{b=1}^{2}\mu_b\frac{\partial g_b}{\partial \dot{\theta}} \qquad \Rightarrow \qquad I_\theta\ddot{\theta} = 0\,. \tag{93}$$

Together with Eqs. (87) and (88) we now have six equations for six unknowns. From the equation of motion for $\theta$, we obtain

$$\theta(t) = \omega t + \theta_0 = \omega t\,, \tag{94}$$

where we have assumed the initial condition $\theta_0 = 0$ for simplicity.

Next, we differentiate the constraint equations again with respect to time, obtaining

$$\ddot{x} = R\ddot{\phi}\cos\omega t - R\omega\dot{\phi}\sin\omega t\,, \tag{95}$$

$$\ddot{y} = R\ddot{\phi}\sin\omega t + R\omega\dot{\phi}\cos\omega t\,. \tag{96}$$

These expressions can be used in the equations of motion for $x$ and $y$ to determine the Lagrange multipliers,

$$\mu_1 = MR\left(\ddot{\phi}\cos\omega t - \omega\dot{\phi}\sin\omega t\right)\,, \tag{97}$$

$$\mu_2 = MR\left(\ddot{\phi}\sin\omega t + \omega\dot{\phi}\cos\omega t\right)\,, \tag{98}$$

and plugging these expressions into the equation of motion for $\phi$, we finally obtain

$$\begin{aligned} I_\phi\ddot{\phi} &= -MR^2\cos\omega t\left(\ddot{\phi}\cos\omega t - \omega\dot{\phi}\sin\omega t\right) - MR^2\sin\omega t\left(\ddot{\phi}\sin\omega t + \omega\dot{\phi}\cos\omega t\right) \\ &= -MR^2\ddot{\phi} \end{aligned} \tag{99}$$

i.e.,

$$(I_\phi + MR^2)\ddot{\phi} = 0 \quad \Rightarrow \quad \dot{\phi}_0 = \text{const.}\,, \tag{100}$$

regardless of the moment of inertia of the disk. We still have to integrate the equations of motion for $x, y$, which read

$$\ddot{x} = -R\omega\dot{\phi}_0 \sin\omega t\,, \tag{101}$$

$$\ddot{y} = R\omega\dot{\phi}_0 \cos\omega t\,. \tag{102}$$

This means

$$\dot{x} = R\dot{\phi}_0 \cos\omega t + \dot{x}_0\,, \tag{103}$$

$$\dot{y} = R\dot{\phi}_0 \sin\omega t + \dot{y}_0\,, \tag{104}$$

and integrating once more, we have

$$x(t) = \frac{R}{\omega}\dot{\phi}_0 \sin\omega t + \dot{x}_0 t + x_0\,, \tag{105}$$

$$y(t) = -\frac{R}{\omega}\dot{\phi}_0 \cos\omega t + \dot{y}_0 t + y_0\,, \tag{106}$$

We can even determine the constraint forces which ensure that the disk rolls without slipping and remains upright.

$$\mu_1 = -MR\omega\dot{\phi} \sin\omega t\,, \tag{107}$$

$$\mu_2 = MR\omega\dot{\phi} \cos\omega t\,, \tag{108}$$

Finally, let us consider that the disk rolls in a straight line, so that $\omega$, the rotation around the disk's vertical axis, vanishes. Then

$$x(t) = \dot{x}_0 t + x_0\,, \tag{109}$$

$$y(t) = \dot{y}_0 t + y_0\,, \tag{110}$$

i.e., uniform linear motion of the center-of-mass with fixed velocity, and

$$\mu_1 = \mu_2 = 0\,. \tag{111}$$

## References


[1] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison-Wesley, 2002).

[2] M. R. Flannery, American Journal of Physics **73**, 265 (2005).

[3] M. R. Flannery, American Journal of Physics **79**, 932 (2011).

[4] M. R. Flannery, Journal of Mathematical Physics **52**, 032705 (2011).


# 4 Symmetries and Invariances

## 4.1 Cyclic Coordinates

As we have seen in several applications of the Lagrange formalism, the structure of the Lagrange equations implies the existence of conserved quantities whenver our system's Lagrangian depends on a generalized velocity $\dot{q}_i$ but not on the associated $q_i$:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} - \underbrace{\frac{\partial L}{\partial q_i}}_{=0} = 0 \quad \Rightarrow \quad p_i \equiv \frac{\partial L}{\partial \dot{q}_i} = \text{const.}\,. \tag{112}$$

We refer to such $q_i$'s as **cyclic coordinates**. The **generalized momentum** $p_i$ associated with each cyclic coordinate is **conserved**, i.e., constant in time. For mechanical systems, such quantities are also referred to as **constants of the motion**.

In general, we allow for Lagrangians $L(q, \dot{q}, t)$, but what happens if $L$ does not depend explicitly on time? Can we find a constant of the motion although there is no explicit Lagrange equation associated with $t$? To answer this question, consider

$$\begin{aligned} \frac{d}{dt} L(q, \dot{q}, t) &= \sum_{i=1}^{n} \left( \frac{\partial L}{\partial q_i} \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i \right) + \frac{\partial L}{\partial t} \\ &= \sum_{i=1}^{n} \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \frac{\partial L}{\partial \dot{q}_i} \frac{d}{dt} \dot{q}_i \right) + \frac{\partial L}{\partial t} \\ &= \frac{d}{dt} \left( \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) + \frac{\partial L}{\partial t} \,, \end{aligned} \tag{113}$$

where we have used the Lagrange equations in the second line. Collecting the total time derivatives on the left-hand side, we have

$$\frac{d}{dt} \left( \sum_{i=1}^{n} \left( \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - L \right) = - \frac{\partial L}{\partial t} \,. \tag{114}$$

If the Lagrangian does not explicitly depend on time, i.e., $\frac{\partial L}{\partial t} = 0$, we obtain another constant of motion [2] :

$$h(q, \dot{q}) = \sum_{i=1}^{n} \left( \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i \right) - L = \text{const.} \,, \tag{115}$$

This is sometimes referred to as the **Jacobi integral** of the system, which will eventually *turn into* the **Hamiltonian** when we make a change from the generalized coordinates and velocities to the proper variables, as discussed later.

## 4.2 Noether's Theorem

The existence of cyclic coordinates is fundamentally related to invariances of the Lagrangian — and therefore the action — under transformations of the generalized coordinates as first proven by Emmy Noether[3].

Let us consider a general infinitesimal transformation of the coordinates and time that changes the Lagrangian of a holonomic system at most by a total time derivative (cf. worksheet #3 and homework problem H4),

$$t \to \quad t' = t + \epsilon \tau(q(t), t) \,, \tag{116}$$

$$q_i \to \quad q_i'(t') = q_i(t) + \epsilon \eta_i(q(t), t) \,, \tag{117}$$

$$L(q, \dot{q}, t) \to \quad L(q'(t'), \dot{q}'(t'), t') = L(q'(t'), \dot{q}'(t'), t') + \epsilon \frac{d}{dt} F(q(t), t) \,, \tag{118}$$

[2]Note that this is equivalent to the Beltrami identity, which we used in our discussion of the brachistochrone problem.

[3]A translated version is available as Ref. [1, 2], or as an updated preprint in `https://arxiv.org/abs/physics/0503066`arXiv:physics/0503066.

where the time derivative of $q'$ is with respect to $t'$. Then the action of the system is invariant under the transformation (116)–(118) , and

$$J = \sum_i \frac{\partial L}{\partial \dot{q}_i} \left( \dot{q}_i \tau - \eta_i \right) - L\tau + F \tag{119}$$

is a **conserved quantity**[4]. The appearance of conserved quantities associated with the invariances of the Lagrangian and the action — i.e., the **symmetries** of the system — is the central statement of **Noether's theorem**, which is of fundamental importance in many domains of physics.

The proof of Noether's theorem is straightforward. First, we note that

$$\frac{dt'}{dt} = 1 + \epsilon\dot{\tau}\,, \quad \frac{dt}{dt'} = (1+\epsilon\dot{\tau})^{-1} = 1 - \epsilon\dot{\tau}\,, \tag{120}$$

where we have dropped higher order terms in the infinitesimal quantity $\epsilon$ in the second expression. Then

$$\frac{dq'}{dt'} = \frac{dt}{dt'}\frac{dq'}{dt} = (1-\epsilon\dot{\tau})\left(\dot{q} + \epsilon\dot{\eta}\right)\,, \tag{121}$$

where the time derivatives indicated by dots refer to $t$. Expanding and dropping terms beyond the linear order, we have

$$\frac{dq'}{dt'} = \dot{q} + \epsilon(\dot{\eta} - \dot{q}\dot{\tau})\,. \tag{122}$$

Now we are ready to consider the action. If $S$ is indeed invariant under the transformation (116)–(118), we must have

$$\Delta S = \int_{t_1'}^{t_2'} dt' L(q'(t'), \dot{q}'(t'), t') - \int_{t_1}^{t_2} dt \left[ L(q(t), \dot{q}(t), t) + \epsilon \frac{dF(q(t),t)}{dt} \right] \stackrel{!}{=} 0\,. \tag{123}$$

Expanding the coordinates in the first term, we have

$$\begin{aligned}
\Delta S &= \int_{t_1}^{t_2} dt(1+\epsilon\dot{\tau}) L(q+\epsilon\eta, \dot{q} + \epsilon(\dot{\eta}-\dot{q}\dot{\tau}), t+\epsilon\tau) - \int_{t_1}^{t_2} dt \left[ L(q(t), \dot{q}(t), t) + \epsilon\frac{dF(q(t),t)}{dt} \right] \\
&= \int_{t_1}^{t_2} dt(1+\epsilon\dot{\tau}) \left[ L(q,\dot{q},t) + \sum_i \left( \frac{\partial L}{\partial q_i}\epsilon\eta_i + \frac{\partial L}{\partial \dot{q}_i}\epsilon(\dot{\eta}_i - \dot{q}_i\dot{\tau}) \right) + \frac{\partial L}{\partial t}\epsilon\tau \right] \\
&\quad - \int_{t_1}^{t_2} dt \left[ L(q,\dot{q},t) + \epsilon\frac{dF}{dt} \right] \\
&= \epsilon \int_{t_1}^{t_2} dt \left[ \sum_i \left( \frac{\partial L}{\partial q_i}\eta_i + \frac{\partial L}{\partial \dot{q}_i}(\dot{\eta}_i - \dot{q}_i\dot{\tau}) \right) + \frac{\partial L}{\partial t}\tau + L\dot{\tau} - \frac{dF}{dt} \right] \\
&= \epsilon \int_{t_1}^{t_2} dt \left[ \sum_i \left( \left( \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right)\eta_i + \frac{\partial L}{\partial \dot{q}_i}\dot{\eta}_i \right) - \left( \sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L \right)\dot{\tau} + \frac{\partial L}{\partial t}\tau - \frac{dF}{dt} \right]\,,
\end{aligned} \tag{124}$$

where we have dropped terms of $\mathcal{O}(\epsilon^2)$, and used the Lagrange equations in the last step. Now we note that the first sum is

$$\sum_i \left( \left( \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right)\eta_i + \frac{\partial L}{\partial \dot{q}_i}\dot{\eta}_i \right) = \frac{d}{dt}\sum_i \left( \frac{\partial L}{\partial \dot{q}_i}\eta_i \right)\,, \tag{125}$$

[4]In classical and quantum field theory, the conserved quantity is usually referred to as the **Noether current** $J_\mu$ ($\mu = 0, \ldots, 3$) and the volume integral $\int d^3r\, J_0(\vec{r})$ defines the **Noether charge**. This is why we use the symbol $J$ for the conserved quantity.

and that we can use the Jacobi integral (115) to rewrite the second and third terms:

$$-\left(\sum_i \frac{\partial L}{\partial \dot{q}_i}\dot{q}_i - L\right)\dot{\tau} + \frac{\partial L}{\partial t}\tau = -h\dot{\tau} - \frac{dh}{dt}\tau = -\frac{d}{dt}(h\tau)\,. \tag{126}$$

Thus,

$$\Delta S = \epsilon \int_{t_1}^{t_2} dt \frac{d}{dt}\left[\sum_i \left(\frac{\partial L}{\partial \dot{q}_i}\eta_i\right) - h\tau - F\right] \stackrel{!}{=} 0\,, \tag{127}$$

and since this must hold for arbitrary $\epsilon$ and paths, we have

$$\frac{d}{dt}\left[\sum_i \left(\frac{\partial L}{\partial \dot{q}_i}\eta_i\right) - h\tau - F\right] = 0\,, \tag{128}$$

i.e.,

$$J = -\sum_i \left(\frac{\partial L}{\partial \dot{q}_i}\eta_i\right) + h\tau + F = \sum_i \frac{\partial L}{\partial \dot{q}_i}\left(\dot{q}_i\tau - \eta_i\right) - L\tau + F = \text{const.}, \tag{129}$$

which completes the proof. The sign change is mere convention here — we will see why it is aesthetically useful below. Let us now use Noether's theorem to study various invariances.

## 4.3 Spatial Translations

Consider a Lagrangian that is invariant under a translation of the coordinate system i.e., a change of coordinates

$$\vec{r}_i(t) \longrightarrow \vec{r}_i'(t) = \vec{r}_i(t) + \epsilon\vec{e}\,, \tag{130}$$

where $\epsilon$ is time independent and $\vec{e}$ is a constant unit vector in $\mathbb{R}^3$. For simplicity, we consider an $N$-particle system without constraints, but the conclusions are readily generalized to holonomic systems.

The kinetic energy is obviously invariant under a translation because $\dot{\vec{r}}_i'(t) = \dot{\vec{r}}_i$, and the potential energy is invariant if it only depends on the particles' relative coordinates since

$$\vec{r}_i' - \vec{r}_j' = \vec{r}_i + \epsilon\vec{e} - \vec{r}_j - \epsilon\vec{e} = \vec{r}_i - \vec{r}_j\,. \tag{131}$$

Thus, a translationally invariant Lagrangian can be written as

$$L(\vec{r},\dot{\vec{r}}) = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\vec{r}}_i^{\,2} - V(\vec{r}_1 - \vec{r}_2, \ldots, \vec{r}_i - \vec{r}_j, \ldots)\,, \tag{132}$$

In the notation introduced above, we have

$$\vec{\eta}_i = \vec{e}\,, \quad \tau = 0\,, \quad F = 0\,, \tag{133}$$

and therefore Noether's theorem (119) implies that

$$\text{const.} = \sum_{i=1}^{N}\sum_{k=1}^{3} \frac{\partial L}{\partial \dot{x}_{ik}}\eta_{ik} = \sum_{k=1}^{3}\underbrace{\sum_{i=1}^{N} \frac{\partial L}{\partial \dot{x}_{ik}}}_{\equiv P_k} e_k = \vec{P}\cdot\vec{e}\,, \tag{134}$$

i.e., the component of the total momentum $\vec{P}$ in the direction of $\vec{e}$ is conserved (think of the block sliding down a wedge discussed on worksheet #3, where we found conservation of the total momentum in $x$ direction).

If the translational invariance holds for an *arbitrary* unit vector $\vec{e}$ in $\mathbb{R}^3$, we have a **spatially homogenous** system in which no single point is preferred. In that case, each component of $\vec{P}$ must be conserved, and we obtain three constants of motion (see Box 4.2). Thus, we find a deep connection between the fundamental structure of space and a conservation law — we will come back to this.

## 4.4 Spatial Rotations

Next, we consider a Lagrangian that remains invariant under rotations by an angle $\epsilon$ around a spatially fixed axis $\vec{e}$. We can express such a rotation in vectorial fashion as

$$\vec{r}_i(t) \longrightarrow \vec{r}_i'(t) = \vec{r}_i(t)\cos\epsilon + \vec{e}(\vec{e}\cdot\vec{r}_i)(1-\cos\epsilon) + (\vec{e}\times\vec{r}_i)\sin\epsilon\,. \tag{135}$$

To see this, we choose a spherical coordinate system such that the $z$ axis is aligned with the rotational axis, $\vec{e} = \vec{e}$. The rotation by an angle $\epsilon$ around this axis can then be expressed in matrix form as

$$\vec{r}(\epsilon) = \begin{pmatrix} \cos\epsilon & -\sin\epsilon & 0 \\ \sin\epsilon & \cos\epsilon & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x\cos\epsilon - y\sin\epsilon \\ x\sin\epsilon + y\cos\epsilon \\ z \end{pmatrix}\,. \tag{136}$$

With $\vec{e} = \vec{e}_z$, Eq. (135) becomes

$$\vec{r}(\epsilon) = \begin{pmatrix} x\cos\epsilon \\ y\cos\epsilon \\ z\cos\epsilon \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z(1-\cos\epsilon) \end{pmatrix} + \begin{pmatrix} -y\sin\epsilon \\ x\sin\epsilon \\ 0 \end{pmatrix} = \begin{pmatrix} x\cos\epsilon - y\sin\epsilon \\ x\sin\epsilon + y\cos\epsilon \\ z \end{pmatrix}\,. \tag{137}$$

Noting that

$$\eta_i = \vec{e}\times\vec{r}_i(t)\,, \quad \tau = 0\,, \quad F = 0\,, \tag{138}$$

Noether's theorem (119) yields

$$\text{const.} = \sum_{i=1}^{N} m_i\dot{\vec{r}}_i\cdot(\vec{e}\times\vec{r}_i) = \vec{e}\cdot\underbrace{\sum_{i=1}^{N} m_i(\vec{r}_i\times\dot{\vec{r}}_i)}_{\equiv\vec{L}} = \vec{e}\cdot\vec{L}\,, \tag{139}$$

where we have introduced the total angular momentum $\vec{L}$ and used

$$\vec{a}\cdot(\vec{b}\times\vec{c}) = \vec{b}\cdot(\vec{c}\times\vec{a}) = \vec{c}\cdot(\vec{a}\times\vec{b})\,. \tag{140}$$

We see that the total angular momentum along the rotational axis $\vec{e}$ is conserved.

If the Lagrangian is invariant under rotations around an *arbitrary* axis $\vec{e}$ in $\mathbb{R}^3$, no direction is preferred and we call a system **spatially isotropic**. An example would be a Lagrangian containing a potential that only depends on the relative *distances* $|\vec{r}_i - \vec{r}_j|$ of the particles. In this case all three components of the total angular momentum $\vec{L}$ are conserved, and we have three constants of motion (see Box 4.2).

### 4.5 Galilean Boosts

According to the special principle of relativity, the laws of physics must be the same in any inertial frame, i.e., any frame moving with a constant velocity with respect to the observer's frame. In non-relativistic mechanics, transformations between such coordinate frames are referred to as **Galilean boosts**, and they have the form

$$\vec{r}_i'(t) = \vec{r}_i(t) + \epsilon \vec{u}_0\, t \tag{141}$$

with a fixed velocity $\vec{u}_0$.

It is easy to see that a potential of the form $V(\vec{r}_1 - \vec{r}_2, \ldots, \vec{r}_i - \vec{r}_j, \ldots)$ will be invariant under Galilean boosts while the kinetic energy will *not* be invariant under such transformations, because

$$T(\epsilon = 0) = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{\vec{r}}_i^2 \quad \longrightarrow \quad T(\epsilon) = \sum_{i=1}^{N} \frac{1}{2} m_i (\dot{\vec{r}}_i + \epsilon \vec{u}_0)^2 \,. \tag{142}$$

However, we can show that the kinetic energy difference between the two frames is a total time derivative:

$$\begin{aligned} L(\vec{r}_i', \dot{\vec{r}}_i', t) &\\ &= \frac{1}{2} \sum_{i=1}^{N} m_i (\dot{\vec{r}}_i + \epsilon \vec{u}_0)^2 - V(\vec{r}_1 + \epsilon \vec{u}_0 t - \vec{r}_2 - \epsilon \vec{u}_0 t, \ldots, \vec{r}_{N-1} + \epsilon \vec{u}_0 t - \vec{r}_N - \epsilon \vec{u}_0 t) \\ &= \frac{1}{2} \sum_{i=1}^{N} m_i \dot{\vec{r}}_i^2 + \sum_{i=1}^{N} m_i \left( \epsilon \dot{\vec{r}}_i \cdot \vec{u}_0 + \frac{1}{2} \epsilon^2 \vec{u}_0^2 \right) - V(\vec{r}_1 - \vec{r}_2, \ldots, \vec{r}_{N-1} - \vec{r}_N) \\ &= L(\vec{r}_i, \dot{\vec{r}}_i, t) + \frac{d}{dt} \underbrace{\left( \sum_{i=1}^{N} m_i \left( \epsilon \vec{u}_0 \cdot \vec{r}_i + \frac{1}{2} \epsilon^2 \vec{u}_0^2 t \right) \right)}_{\equiv \epsilon F + \mathcal{O}(\epsilon^2)} . \end{aligned} \tag{143}$$

Thus, we have

$$\vec{\eta}_i = \vec{u}_0\, t \,, \quad \tau = 0 \,, \quad F = \sum_{i=1}^{N} m_i \vec{u}_0 \cdot \vec{r}_i \,, \tag{144}$$

and Eq. (119) implies

$$\sum_{i=1}^{N} m_i \dot{\vec{r}}_i \cdot \vec{u}_0 t - \sum_{i=1}^{N} m_i \vec{r}_i \cdot \vec{u}_0 = \vec{u}_0 (\vec{P} t - M \vec{R}) = \text{const.} \,, \tag{145}$$

where we have introduced the total momentum and the center-of-mass position vector

$$\vec{R} = \frac{1}{M} \sum_{i=1}^{N} m_i \vec{r}_i \,. \tag{146}$$

For arbitray $\vec{u}_0$, we obtain three constants of motion:

$$\vec{P} t - M \vec{R} = \text{const.} \tag{147}$$

(see Box 4.2). If we write the constant as $-M\vec{R}_0$, we can solve for $\vec{R}$ and find

$$\vec{R}(t) = \frac{1}{M} \vec{P} t + \vec{R}_0 \,, \tag{148}$$

which is the trajectory of the center of mass undergoing uniform linear motion. If the direction of $\vec{u}_0$ is fixed, we only obtain one constant of motion, and the motion of the center of mass is only uniform along the direction of $\vec{u}_0$.

## 4.6 Translations in Time

A system whose properties are invariant under temporal translations

$$t \longrightarrow t + \epsilon \tag{149}$$

is called **homogenous in time**. This means that the results of any measurement are independent of the specific time at which it is conducted. Applying Noether's theorem with

$$\eta_i = 0\,, \quad \tau = 1\,, \quad F = 0\,, \tag{150}$$

we see that

$$J = -\sum_i \left(\frac{\partial L}{\partial \dot{q}_i}\eta_i\right) + h\tau + F = h = \text{const.}\,, \tag{151}$$

i.e., the Jacobi integral (115) is conserved. To understand its physical meaning, we need to analyze the terms $\frac{\partial L}{\partial \dot{q}_i}\dot{q}_i$ in Eq. (115). For now, we restrict the discussion to systems without velocity-dependent potentials or dissipation — we will consider such systems later.

### Scleronomic Constraints

In a holonomic system with scleronomic constraints, we have $\vec{r}_i = \vec{r}_i(q)$, which implies

$$\dot{\vec{r}}_i = \sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j \tag{152}$$

and

$$T = \frac{1}{2}\sum_{i=1}^{N} m_i \dot{\vec{r}}_i \cdot \dot{\vec{r}}_i = \frac{1}{2}\sum_{j,k=1}^{n}\left(\sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial q_j}\cdot\frac{\partial \vec{r}_i}{\partial q_k}\right)\dot{q}_j\dot{q}_k \equiv \frac{1}{2}\sum_{j,k=1}^{n} M_{jk}\dot{q}_j\dot{q}_k\,. \tag{153}$$

Here, we have introduced the so-called **mass tensor M**. The kinetic energy's partial derivative with respect to $\dot{q}_j$ is given by

$$\begin{aligned}\frac{\partial T}{\partial \dot{q}_j} &= \frac{1}{2}\sum_{kl} M_{kl}\frac{\partial}{\partial \dot{q}_j}\left(\dot{q}_k\dot{q}_l\right) = \frac{1}{2}\sum_{kl} M_{kl}\left(\frac{\partial \dot{q}_k}{\partial \dot{q}_j}\dot{q}_l + \dot{q}_k\frac{\partial \dot{q}_l}{\partial \dot{q}_j}\right)\\ &= \frac{1}{2}\sum_{kl} M_{kl}\left(\delta_{jk}\dot{q}_l + \delta_{jl}\dot{q}_k\right) = \sum_k M_{jk}\dot{q}_k\,,\end{aligned} \tag{154}$$

where we have used the symmetry of the mass tensor ($M_{jk} = M_{kj}$) and the freedom to rename summation variables. This means that

$$\sum_j \frac{\partial T}{\partial \dot{q}_j}\dot{q}_j = \sum_{jk} M_{jk}\dot{q}_j\dot{q}_k = 2T\,, \tag{155}$$

and since we only consider velocity-independent potentials, the Jacobi integral becomes

$$h(q,\dot{q}) = \sum_{j=1}^{n}\frac{\partial L}{\partial \dot{q}_j}\dot{q}_j - L = \sum_{j=1}^{n}\frac{\partial T}{\partial \dot{q}_j}\dot{q}_j - L = 2T - (T - V) = T + V\,. \tag{156}$$

Thus, invariance with respect to translations in time implies the **conservation of energy** in holonomic systems with scleronomic constraints, and the Jacobi integral is the *total energy* of such a system (see Box 4.2).

**Box 4.1: Euler's Homogenous Function Theorem**

A homogenous function $F(x_1, \ldots, x_n)$ of degree $k$ has the property

$$F(\lambda x_1, \ldots, \lambda x_n) = \lambda^k F(x_1, \ldots, x_n)\,, \tag{I4.1-1}$$

which means that

$$\sum_{i=1}^{n} x_i \frac{\partial F}{\partial x_i} = kF\,, \tag{I4.1-2}$$

Thus, for holonomic systems with scleronomic constraints, the kinetic energy is a homogenous function of degree 2 in the generalized velocities.

### Rheonomic Systems

Let us now consider holonomic systems with rheonomic constraints. In this case, we have $\vec{r}_i = \vec{r}_i(q,t)$, and

$$\dot{\vec{r}}_i = \sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial \vec{r}_i}{\partial t}\,. \tag{157}$$

The kinetic energy now becomes

$$T = \frac{1}{2}\sum_{i=1}^{N} m_i \left(\sum_{k=1}^{n} \frac{\partial \vec{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \vec{r}_i}{\partial t}\right) \cdot \left(\sum_{l=1}^{n} \frac{\partial \vec{r}}{\partial q_l}\dot{q}_l + \frac{\partial \vec{r}_i}{\partial t}\right)\,. \tag{158}$$

Computing the first term in the Jacobi integral, we find

$$\begin{aligned}
\sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j}\dot{q}_j &= \sum_{j=1}^{n} \frac{\partial T}{\partial \dot{q}_j}\dot{q}_j = \sum_{i=1}^{N}\sum_{j,k=1}^{n} m_i \left(\frac{\partial \vec{r}_i}{\partial q_j}\underbrace{\frac{\partial \dot{q}_k}{\partial \dot{q}_j}}_{=\delta_{jk}}\right) \cdot \underbrace{\left(\sum_{l=1}^{n} \frac{\partial \vec{r}_i}{\partial q_l}\dot{q}_l + \frac{\partial \vec{r}_i}{\partial t}\right)}_{=\dot{\vec{r}}_i} \dot{q}_j \\
&= \sum_{i=1}^{N} m_i \left(\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j\right) \cdot \dot{\vec{r}}_i = \sum_{i=1}^{N} m_i \underbrace{\left(\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial \vec{r}_i}{\partial t}\right)}_{=\dot{\vec{r}}_i} \cdot \dot{\vec{r}}_i - \sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial t} \cdot \dot{\vec{r}}_i \\
&= 2T - \sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial t} \cdot \dot{\vec{r}}_i\,,
\end{aligned} \tag{159}$$

and in total

$$h(q,\dot{q}) = \sum_{j=1}^{n} \frac{\partial L}{\partial \dot{q}_j}\dot{q}_j - L = T + V - \sum_{i=1}^{N} m_i \dot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial t}\,. \tag{160}$$

Thus, the Jacobi integral does *not* correspond to the total energy in this case unless the third term vanishes. This term is the projection of the momentum on the tangent vector $\frac{\partial \vec{r}_i}{\partial t}$, which results from the time dependence of the constraint in the rheonomic case (compare Eqs. (152) and (157)). Thus, if the change of the constraint is orthogonal to the direction of motion, $h$ would still be the total energy.

**Example: Bead on a Rotating Wire**

As an example for a rheonomic system, we consider a bead on a rotating wire, as previously discussed. For simplicity, we choose $\alpha = \pi/2$, so that the coordinates are

$$x = q\cos\omega t\,, \quad y = q\sin\omega t\,, \quad z = 0\,, \tag{161}$$

and the Lagrangian becomes

$$\begin{aligned} L &= \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) \\ &= \frac{1}{2}m\left(\dot{q}^2(\cos^2\omega t + \sin^2\omega t) + q^2\omega^2(\sin^2\omega t + \cos^2\omega t)\right) \\ &= \frac{1}{2}m\left(\dot{q}^2 + \omega^2 q^2\right)\,. \end{aligned} \tag{162}$$

Now

$$\frac{\partial L}{\partial \dot{q}} = m\dot{q}\,, \quad \frac{\partial L}{\partial q} = mq\omega^2\,, \tag{163}$$

and the Lagrange equation yields the equation of motion

$$\ddot{q} - \omega^2 q = 0\,. \tag{164}$$

The Jacobi integral becomes

$$h(q,\dot{q}) = \frac{\partial L}{\partial \dot{q}}\dot{q} - L(q,\dot{q}) = \frac{1}{2}m\left(\dot{q}^2 - q^2\omega^2\right)\,, \tag{165}$$

which is conserved because

$$\frac{dh}{dt} = -\frac{\partial L}{\partial t} = 0 \tag{166}$$

(see Eq. (114)). However, it is *not* the total energy, which we can see by considering the relation (160). Since $V = 0$, the total energy is identical to the kinetic energy,

$$T = \frac{1}{2}m\left(\dot{q}^2 + \omega^2 q^2\right)\,. \tag{167}$$

The momentum is

$$m\dot{\vec{r}} = m\left(\dot{q}\cos\omega t - q\omega\sin\omega t,\ \dot{q}\sin\omega t + q\omega\cos\omega t,\ 0\right)^T \tag{168}$$

and the explicit time dependence of the coordinate due to the rheonomic constraint imposed by the rotating wire yields

$$\frac{\partial \vec{r}}{\partial t} = \left(-q\omega\sin\omega t,\ q\omega\cos\omega t,\ 0\right)^T\,, \tag{169}$$

so we obtain

$$m\dot{\vec{r}}\cdot\frac{\partial \vec{r}}{\partial t} = mq^2\omega^2\,. \tag{170}$$

The time derivative of the total energy is given by

$$\frac{dE}{dt} = \frac{d(T+V)}{dt} = \frac{d}{dt}\left(\frac{1}{2}m\left(\dot{q}^2 + \omega^2 q^2\right)\right) = m\dot{q}\left(\ddot{q} + \omega^2 q\right)\,, \tag{171}$$

and plugging in Eq. (164), we obtain

$$\frac{dE}{dt} = \frac{d}{dt}m\omega^2 q^2 = 2m\omega^2 q\dot{q}\,. \tag{172}$$

This is the change in the energy the motor needs to provide to keep the wire spinning at a constant angular velocity as the pearl slides to different positions $q$ along the wire.

**Box 4.2: Constants of Motion**

For a closed system of $N$ (non-relativistic) particles that only interact through conservative forces which depend on $\vec{r}_i - \vec{r}_j$, Noether's theorem implies the existence of **10 constants of motion**:

| Symmetry Transformation | | Conserved Quantity | |
|---|---|---|---|
| temporal translation | $t' = t + \tau$ | $h = \text{const.}$ | Jacobi integral |
| | | $h = E = \text{const.}$ | total energy (scleronomic systems) |
| spatial translation | $\vec{r}' = \vec{r} + \vec{a}$ | $\vec{P} = \text{const.}$ | total momentum |
| rotation | $\vec{r}' = R\vec{r}$ | $\vec{L} = \text{const.}$ | total angular momentum |
| Galilean boost | $\vec{r}' = \vec{r} + \vec{u}\,t$ | $M\vec{R} - \vec{P}t = \text{const.}$ | |

## 4.7 Additional Remarks and Summary

- The constants of motion that are associated with the fundamental symmetries of (nonrelativistic) spacetime are summarized in Box 4.2.
- In general, a system with $N$ continuous degrees of freedom will admit at most $N$ independent constants of the motion. In the extreme case, all generalized coordinates are either immediately found to be cyclic, or we can perform a symmetry transformation to new cyclic coordinates (see problem G11).
- It is important to recognize that the symmetries are conceptually more fundamental than the constants of motion. Noether's theorem merely provides us with the machinery to construct *a* conserved quantitiy associated with a particular continuous symmetry of the action or the Lagrangian. It is easy to see that arbitrary linear combinations or even products of the constants of motion would be conserved in time as well.
- It is worth stressing that Noether's theorem applies to **continuous** symmetries of the Lagrangian and the action. Its proof relies on the existence on the continuity and differentiability of the Lagrangian, which we Taylor expand for small changes in the generalized coordinates as well as time.
- Examples of discrete symmetries for which Noether's theorem does not apply are
  1. the reflection symmetry $\vec{r} \to -\vec{r}$ (cf. homework problem H8),
  2. the discrete translational symmetry in a crystal,
  $$\frac{d}{dt}(\vec{r} + n\vec{a}_i) = \dot{\vec{r}}, V(\vec{r} + n\vec{a}_i) = V(\vec{r})\,, \quad \vec{a}_i = \text{const.}, n \in \mathbb{Z}\,, \tag{173}$$
  where $\vec{a}_i$ are the directions along which we shift between lattice sites,
  3. or the invariance of the pendulum Lagrangian
  $$L = \frac{1}{2}ml^2\dot{\theta}^2 - mgl\cos\theta \tag{174}$$
  under the discrete shift $\theta \to \theta + n \cdot 2\pi\,, n \in \mathbb{Z}$. Unlike a continuous shift in $\theta$, this invariance is not associated with angular momentum conservation because
  $$\frac{d}{dt}\frac{\partial L}{\partial\dot{\theta}} = \frac{\partial L}{\partial\theta} \neq 0 \tag{175}$$

unless $g = 0$.

## References

[1] E. Noether, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1918**, 235 (1918).

[2] E. Noether, Transport Theory and Statistical Physics **1**, 186 (1971).

# 5 Group Exercises

## Problem G9 – Lagrange Equations of the First and Second Kind

We consider gravity acting on a particle of mass $m$ that glides without friction on the interior of the rotational paraboloid

$$az = x^2 + y^2 \,, \quad a = \text{const.} \tag{176}$$

1. Construct the Lagrangian $L(\rho, \phi, z)$ in cylindrical coordinates of a particle that is moving **without constraints** under the influence of gravity. Determine the Lagrange equations, and compare them to the equations resulting from Newton's Second Law in cylindrical coordinates.

2. Implement the constraint that the particle moves on the paraboloid (in cylindrical coordinates) by adding it to the Lagrangian with a Lagrange multiplier, and derive the Lagrange equations for the modified Lagrangian $\widetilde{L}(r, \phi, z, \lambda)$.

3. Determine $\lambda$ and use it to decouple the equations of motion.

Let us now exploit that the constraint is holonomic, and immediately use the Lagrange formalism of the second kind.

4. Use the constraint to identify suitable generalized coordinates $q_i$, and construct the Lagrangian $L(q_i, \dot{q}_i)$.

5. Derive the Lagrange equations for the generalized coordinates, and compare them to your results from step 3.

6. Show that the particle will move with an angular velocity $\omega = \sqrt{2g/a}$ if we restrict its trajectory to a circle at fixed height $h$.

## Problem G10 – Noether's Theorem for an Oscillator

[**cf. Lemos, 2.30**] Consider a harmonic oscillator with mass $m$ and frequency $\omega$.

1. Show that

   $$C(x, \dot{x}, t) = \dot{x} \cos \omega t + \omega x \sin \omega t \tag{177}$$

   is a conserved quantity.

2. Prove that this quantity is associated with the invariance of the action under the continuous transformation

$$t' = t \,, \tag{178}$$

$$x'(t') = x(t) - \frac{\epsilon}{m} \cos \omega t \,, \tag{179}$$

$$F(t) = \omega x \sin \omega t \,. \tag{180}$$

## Problem G11 – A System with Two Degrees of Freedom

**[cf. Lemos, problem 2.17]** The Lagrangian for a system with two degrees of freedom is

$$L = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) - (\alpha x + \beta y) \,, \tag{181}$$

where $\alpha, \beta \neq 0$.

1. Prove that the Lagrangian (and therefore the action) are invariant under the coordinate transformation

   $$x \to x' = x + \epsilon\beta \,, \quad y \to y' = y - \epsilon\alpha \,. \tag{182}$$

   Using Noether's theorem, show that the quantity

   $$A = \beta\dot{x} - \alpha\dot{y} \tag{183}$$

   is conserved.

2. Introduce new generalized coordinates

   $$\bar{x} = \alpha x + \beta y \,, \quad \bar{y} = \beta x - \alpha y \tag{184}$$

   and show that one of the new coordinates is cyclic in the new Lagrangian $L(\bar{x}, \bar{y}, \dot{\bar{x}}, \dot{\bar{y}}$. Prove that the quantity $A$ from the previous part of the problem is proportional to the momentum that is conjugate to the cyclic coordinate.