

MATH-F-204 Analytical mechanics: Lagrangian and Hamiltonian formalism

Glenn Barnich

Physique Théorique et Mathématique
Université libre de Bruxelles and International Solvay Institutes
Campus Plaine C.P. 231, B-1050 Bruxelles, Belgique
Bureau 2.O6. 217, E-mail: Glenn.Barnich@ulb.be, Tel: 02 650 58 01

February 22, 2026

Abstract

The aim of the course is an introduction to Lagrangian and Hamiltonian mechanics. The fundamental equations as well as standard applications of Newtonian mechanics are derived from variational principles.

From a mathematical perspective, the course and exercises (i) constitute an application of differential and integral calculus (primitives, graphs and functions); (ii) provide examples and explicit solutions of first and second order systems of differential equations; (iii) give an implicit introduction to differential manifolds and to realizations of Lie groups and algebras (change of coordinates, parametrization of a rotation, Galilean group, canonical transformations, one-parameter subgroups, infinitesimal transformations).

From a physical viewpoint, explicit and elegant solutions to idealized, but nevertheless concrete problems are provided. The more formal aspects of the course, such as (i) the covariance of the Lagrangian formalism under changes of Lagrange coordinates, (ii) symmetries and conservation laws in the form of Noether's theorem as applied to Galilean invariance, (iii) canonical transformations in the Hamiltonian formalism, pave the way for a thorough understanding of the structure of quantum mechanics, both in its operatorial and path integral formulations, and of field theories with Poincaré invariance or general relativity.

The course of 30 hours goes hand in hand with 30 hours of exercises. The focus of the course is the general theoretical setting and abstract developments. Simple examples and concrete applications of these general concepts are treated during the exercise sessions. A full understanding of the material requires both of these complementary parts.

Contents

1	Extremisation, constraints and Lagrange multipliers	4
1.1	Unconstrained extremisation	4
1.2	Constraints and regularity conditions	4
1.3	Constrained extremisation	7
2	Constrained systems and d'Alembert principle	8
2.1	Holonomic constraints	8
2.2	d'Alembert's theorem	9
2.3	Principle of virtual work	10
2.3.1	Worked Exercise 1: The Lever	10
2.3.2	Worked Exercise 2: Block and Tackle (Palan)	11
2.4	Non-holonomic constraints	12
2.5	d'Alembert's principle	14
3	Hamilton and least action principles	16
3.1	Natural and virtual trajectories	16
3.2	Hamilton's principle	16
3.3	Least action principle	17
4	Lagrangian mechanics	19
4.1	Lagrange coordinates	19
4.2	Lagrange equations	19
4.3	Dissipative systems	21
4.4	Euler-Lagrange equations	21
4.4.1	Worked Exercise 3: A Lagrangian of sufficient generality	23
4.5	Points of equilibrium	24
5	Applications of the Lagrangian formalism	25
5.1	The simple pendulum	25
5.1.1	Small oscillations	26
5.1.2	Conservation of energy	26
5.1.3	Points of equilibrium	27
5.1.4	Vanishing velocity	27
5.1.5	Motion if $\alpha = \pi$	28
5.1.6	Oscillations	28
5.1.7	Revolutions	29
5.2	The brachistochrone	31
5.2.1	Statement of the problem	31
5.2.2	Main equations	31
5.2.3	Finding the curve	32

5.2.4	Properties of the motion	34
6	Hamiltonian formalism	35
6.1	Legendre transform	35
6.2	Legendre transform in mechanics and Hamilton's equations	37
6.3	Poisson brackets	38
7	Covariance of the Lagrangian formalism	41
7.1	Change of dependent variables	41
7.2	Change of the independent variable	42
7.3	Combined change of variables	43
8	Canonical transformations	45
8.1	Generating functionals	45
8.2	Invariance of canonical Poisson brackets	48
8.3	Infinitesimal canonical transformations	49
8.4	One-parameter group of canonical transformations	49
8.5	Liouville theorem	50
9	Noether's theorem and Galilean relativity	53
9.1	Noether's first theorem	53
9.2	Applications	54
9.2.1	Explicit time independence and conservation of energy	54
9.2.2	Cyclic variables and conservation of canonical momentum	54
9.2.3	One-parameter groups of transformations	55
9.2.4	Parametrization of a rotation	57
9.2.5	One-parameter groups of transformations involving time	60
9.2.6	Time-translation symmetry and conservation of energy	62
9.2.7	Galilean boosts	62
9.3	Hamiltonian version of Noether's theorem	63
9.4	More applications	64
9.4.1	Hamiltonian formulation of the Galilean algebra	64
9.4.2	Worked Exercise 4: Geodesic motion and Killing vectors	64
9.4.3	Worked Exercise 5: Geodesic motion and Killing tensors	65
9.4.4	Worked Exercise 6: Laplace-Runge-Lenz vector	65
9.4.5	Worked Exercise 7: Harmonic oscillator conserved quantities	65
10	Reduced and alternative variational principles	66
10.1	Routhian procedure for cyclic variables	66
10.2	Parametrized systems	67
10.3	Maupertius' / Jacobi's principle	68
11	Dynamics in non-inertial frames	70
12	Relativistic particles	71
13	Charged particle in electromagnetic field	74
13.1	Newtonian formulation	74
13.2	Lagrangian formulation	74
13.3	Hamiltonian formulation	75

14 To be added	77
References	77

Chapter 1

Extremisation, constraints and Lagrange multipliers

The purpose of this chapter is to re-derive in the finite-dimensional case standard results on constrained extremisation so as to motivate and clarify analogous results needed in the functional case. It is based on chapters II.1, II.2, II.5 of [8]. The discussion on constraints and regularity conditions is taken from chapter 1 of [11].

1.1 Unconstrained extremisation

The problem consists of finding the extrema (=stationary points) of a function $F(x^1, \dots, x^n)$ of n variables x^a , $a = 1, \dots, n$ in the interior of a domain. We write the variation of such a function under variations δx^a of the variables x^a as

$$\delta F = \delta x^a \frac{\partial F}{\partial x^a}.$$

Here and throughout, we use the summation convention over repeated indices, $\delta x^a \frac{\partial F}{\partial x^a} = \sum_{a=1}^n \delta x^a \frac{\partial F}{\partial x^a}$. Since for a collection of functions $f_a(x^b)$,

$$f_a \delta x^a = 0, \forall \delta x^a \implies f_a = 0, \quad (1.1)$$

it follows that the extrema $\bar{x}^a$ are determined by the vanishing of the gradient of F ,

$$\delta F = 0 \iff \left. \frac{\partial F}{\partial x^a} \right|_{x=\bar{x}} = 0. \quad (1.2)$$

Recall further that the nature (maximum, minimum, saddle point) of each extremum is determined by the eigenvalues of the matrix

$$\left. \frac{\partial^2 F}{\partial x^a \partial x^b} \right|_{x=\bar{x}}. \quad (1.3)$$

Remark: The geometric interpretation of the LHS of (1.1) is that the vectors $\vec{f}$ and $\vec{\delta x}$ are perpendicular (in Euclidean space $\mathbb{R}^n$).

1.2 Constraints and regularity conditions

Very often, one is interested in a problem of extremisation, where the variables x^a are subjected to constraints, that is to say where r independent relations

$$G_m(x) = 0, \quad m = 1, \dots, r, \quad (1.4)$$

are supposed to hold. Standard regularity on the functions G_m are that the rank of the matrix $\frac{\partial G_m}{\partial x^a}$ is r (locally, in a neighborhood of $G_m = 0$). Without loss of generality (by renaming if necessary some of the variables), we can assume that

$$\left| \frac{\partial G_m}{\partial x^a} \right| \neq 0, \text{ for } a = n - r + 1, \dots, n. \quad (1.5)$$

One can prove¹ that this implies that one may locally solve the constraints in terms of the r last coordinates,

$$G_m = 0 \iff x^\Delta = X^\Delta(x^\alpha), \quad \Delta = n - r + 1, \dots, n, \quad \alpha = 1, \dots, n - r. \quad (1.6)$$

In this case, one refers to the x^Δ as dependent and to the x^α as independent variables.

This also means that one may choose locally a new system of coordinates such that G_m are r of the new coordinates, $x^a \longleftrightarrow q^\alpha, G_m$, with $\alpha = 1, \dots, n - r$.

Example: Consider a circle of radius R in $\mathbb{R}^2$: $x^1, x^2 \in \mathbb{R}^2$,

$$G = \sqrt{(x^1)^2 + (x^2)^2} - R = 0 \quad (1.7)$$

One may solve the constraint to express x^2 , the dependent variable, in terms of x^1 , the independent variable: $x^2 = \pm \sqrt{R^2 - (x^1)^2}$. Alternatively, if one uses polar coordinates $x^1 = r \cos \theta$, $x^2 = r \sin \theta$, the constraint becomes $G = r - R$. Hence, if one uses $\theta, \rho = r - R$ as new coordinates, $q = \theta, G = \rho$.

Lemma 1. *A function f that vanishes when the constraints hold can be written as a combination of the functions defining the constraints:*

$$\boxed{f|_{G_m=0} = 0 \iff f = g^m G_m}, \quad (1.8)$$

for some coefficients $g^m(x)$.

That the condition is sufficient ($\Leftarrow$) is obvious. That it is necessary ($\Rightarrow$) is shown in the new coordinate system, where $f(q^\alpha, G_m = 0) = 0$ and thus

$$f(q^\alpha, G_m) = \int_0^1 d\tau \frac{d}{d\tau} f(q^\alpha, \tau G_m) = G_m \int_0^1 d\tau \frac{\partial f}{\partial G_m}(q^\alpha, \tau G_m), \quad (1.9)$$

so that $g^m = \int_0^1 d\tau \frac{\partial f}{\partial G_m}(q^\alpha, \tau G_m)$.

In the following, we will use the notation $f \approx 0$ for a function that vanishes when the constraints hold, so that the lemma becomes $f \approx 0 \iff f = g^m G_m$.

Lemma 2. *Suppose that the constraints hold and that one restricts oneself to variations δx^a tangent to the constraint surface,*

$$G_m = 0, \quad \frac{\partial G_m}{\partial x^a} \delta x^a \approx 0. \quad (1.10)$$

It follows that

$$\boxed{f_a \delta x^a \approx 0, \quad \forall \delta x^a \text{ satisfying (1.10)} \iff f_a \approx \mu^m \frac{\partial G_m}{\partial x^a}}, \quad (1.11)$$

for some $\mu^m(x^a)$.

¹The explicit proof, which is not needed here, goes under the name of “implicit function theorem”.

NB: In this lemma, the use of $\approx$ means that these equalities in (1.10) and (1.11) are understood as holding when $G_m = 0$.

That the condition is sufficient ($\Leftarrow$) is again direct when contracting the expression for f_a in the RHS of (1.11) with δx^a and using the second of (1.10).

In the adapted coordinate system (q^α, G_m) , variations tangent to the constraints correspond to δq^α arbitrary, and $\delta G_m = 0$. The LHS of (1.11) becomes $f_\alpha(q^\beta, 0)\delta q^\alpha = 0$, which is equivalent to $f_\alpha(q^\beta, 0) = 0$. Since $\mu^m \frac{\partial G_m}{\partial q^\alpha} = 0$, the RHS of (1.11) is satisfied for the α -components. It is also satisfied for the n components by choosing $\mu^m = \delta_n^m f^n$.

In a more general coordinate system, that the condition is necessary ($\Rightarrow$) follows by first subtracting from the LHS of (1.11) the combination $\lambda^m \frac{\partial G_m}{\partial x^a} \delta x^a \approx 0$, with arbitrary λ^m , which gives

$$(f_a - \lambda^m \frac{\partial G_m}{\partial x^a}) \delta x^a \approx 0. \quad (1.12)$$

The regularity conditions imply that, locally, there exists an invertible matrix $M^{m\Delta}$ such that

$$M^{m\Delta} \frac{\partial G_{m'}}{\partial x^\Delta} \approx \delta_{m'}, \quad \frac{\partial G_m}{\partial x^\Delta} M^{m\Gamma} \approx \delta_\Delta^\Gamma. \quad (1.13)$$

Indeed, according Lemma 1, there exists functions $M^{\Gamma m}$ such that

$$x^\Delta - X^\Delta = M^{\Delta m} G_m, \quad (1.14)$$

Conversely, when describing the surface by $x^\Gamma - X^\Gamma$, there exists functions $M_{m\Delta}$ such that

$$G_m = M_{m\Gamma} (x^\Gamma - X^\Gamma). \quad (1.15)$$

Substituting this in the previous equation and differentiating with respect to x^Δ , it follows that

$$M^{\Delta m} M_{m\Gamma} \approx \delta_\Gamma^\Delta. \quad (1.16)$$

The second of equation (1.13) then follows by differentiating (1.14) with respect to x^Δ ,

$$\delta_\Gamma^\Delta = \frac{\partial M^{\Delta m}}{\partial x^\Gamma} G_m + M^{\Delta m} \frac{\partial G_m}{\partial x^\Gamma}, \quad (1.17)$$

and the first because the right and the left inverse are the same.

In terms of the various types of variables, (1.12) may be decomposed as

$$(f_\Delta - \lambda^m \frac{\partial G_m}{\partial x^\Delta}) \delta x^\Delta + (f_\alpha - \lambda^m \frac{\partial G_m}{\partial x^\alpha}) \delta x^\alpha \approx 0. \quad (1.18)$$

Since the λ^m are arbitrary, we are free to fix

$$\lambda^m = M^{m\Gamma} f_\Gamma := \mu^m \iff f_\Delta \approx \mu^m \frac{\partial G_m}{\partial x^\Delta}, \quad (1.19)$$

so that the first term (1.18) vanishes. We then remain with

$$(f_\alpha - \mu^m \frac{\partial G_m}{\partial x^\alpha}) \delta x^\alpha \approx 0. \quad (1.20)$$

Since the variables x^α are unconstrained and the variations δx^α independent, this implies as in (1.1) that

$$f_\alpha \approx \mu^m \frac{\partial G_m}{\partial x^\alpha}. \quad (1.21)$$

□

1.3 Constrained extremisation

Extremisation of a function $F(x^a)$ under r constraints $G_m = 0$ can then be done in two equivalent ways:

Method 1: The first is to solve the constraints as in (1.6) and to extremise the function

$$\boxed{F^G(x^\alpha) = F(x^\alpha, X^\Delta(x^\beta))}. \quad (1.22)$$

In terms of the unconstrained variables x^α , we are back to an unconstrained extremisation problem, whose extrema are determined by

$$\boxed{\frac{\partial F^G}{\partial x^\alpha} = 0}. \quad (1.23)$$

Indeed, on the one hand,

$$\frac{\partial F^G}{\partial x^\alpha} = \frac{\partial F}{\partial x^\alpha} \Big|_{x^\Delta = X^\Delta} + \frac{\partial F}{\partial x^\Gamma} \Big|_{x^\Delta = X^\Delta} \frac{\partial X^\Gamma}{\partial x^\alpha}. \quad (1.24)$$

On the other hand, differentiating (1.14) with respect to x^α ,

$$\frac{\partial X^\Gamma}{\partial x^\alpha} = \frac{\partial (X^\Gamma - x^\Gamma)}{\partial x^\alpha} = -\frac{\partial M^{\Gamma m}}{\partial x^\alpha} G_m - M^{\Gamma m} \frac{\partial G_m}{\partial x^\alpha}, \quad (1.25)$$

so that, if $G_m = 0$, or equivalently, if one substitutes x^Δ by X^Δ , (1.23) may be written as

$$\frac{\partial F}{\partial x^\alpha} \approx \mu^m \frac{\partial G_m}{\partial x^\alpha}, \quad \mu^m \approx \frac{\partial F}{\partial x^\Gamma} M^{\Gamma m}. \quad (1.26)$$

Furthermore, if $G_m = 0$, one may also write

$$\frac{\partial F}{\partial x^\Delta} = \frac{\partial F}{\partial x^\Gamma} \delta_\Delta^\Gamma \approx \frac{\partial F}{\partial x^\Gamma} M^{\Gamma m} \frac{\partial G_m}{\partial x^\Delta} \approx \mu^m \frac{\partial G_m}{\partial x^\Delta}. \quad (1.27)$$

by using (1.17) and the definition of μ^m in (1.26).

Method 2: Instead of restricting oneself to the space of independent variables x^a , one extends the space of variables x^a by additional variables λ^m , called Lagrange multipliers, and one considers the unconstrained extremisation of the function,

$$\boxed{F^\lambda(x^a, \lambda^m) = F(x^a) - \lambda^m G_m, \quad \delta F^\lambda = 0}. \quad (1.28)$$

By unconstrained extremisation, we understand that the variables x^a, λ^m are considered as unconstrained at the outset, with independent variations $\delta x^a, \delta \lambda^m$.

That this gives the same extrema can be seen as follows:

$$0 = \delta F^\lambda = \left(\frac{\partial F}{\partial x^a} - \lambda^m \frac{\partial G_m}{\partial x^a} \right) \delta x^a - \delta \lambda^m G_m, \quad (1.29)$$

implies

$$\frac{\partial F}{\partial x^a} = \lambda^m \frac{\partial G_m}{\partial x^a}, \quad G_m = 0. \quad (1.30)$$

A variation of the second equation then implies that $\frac{\partial G_m}{\partial x^a} \delta x^a = 0$, while contraction of the first equation with δx^a gives us (1.12) of the proof of lemma 2. We can then continue as in the proof of this lemma to conclude that $\frac{\partial F}{\partial x^a} \approx \mu^m \frac{\partial G_m}{\partial x^a}$ for some μ^m , and thus that both methods define the same extrema.

Chapter 2

Constrained systems and d'Alembert principle

The chapter starts with a description of holonomic constraints on the trajectories of point particles. Real displacements are distinguished from virtual (instantaneous) displacements. D'Alembert's theorem as an alternative equivalent formulation of Newtonian mechanics is stated and proved. For non-holonomic constraints, the same is achieved through d'Alembert's theorem. A criterion to distinguish holonomic from non-holonomic constraints is provided. This chapter is based on [4], chapters 1.1, 1.2, 1.3.

2.1 Holonomic constraints

Consider $\mathcal{N}$ particles moving in space $\mathbb{R}^3$, with coordinates are denoted by x_i^α , $i = 1, \dots, \mathcal{N}$, $\alpha = 1, 2, 3$. Alternatively, we can label them by x^a , with $a = 1, \dots, 3\mathcal{N}$. We assume that they are submitted to r constraints of the form

$$G_m(x_i^\alpha, t) = 0. \quad (2.1)$$

Such constraints that only involve the positions but not the velocities of the particles are called *holonomic*. As before, we assume regularity conditions: the rank of matrix $\frac{\partial G_m}{\partial x_i^\alpha}$ is maximal for all values of t . When the constraints do not depend explicitly on time, they are called *scleronomic*, if they do, they are called *rheonomic*.

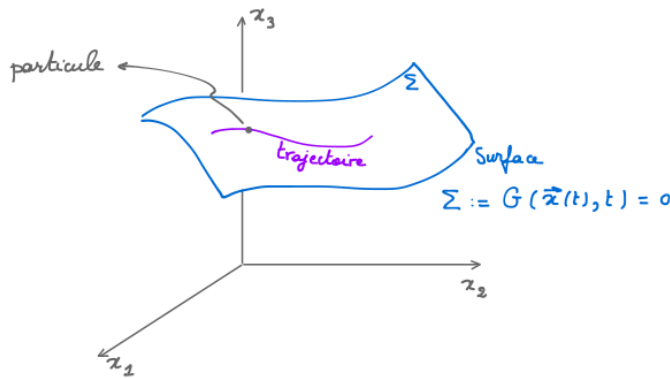


Figure 2.1: A particle in $\mathbb{R}^3$ constrained to a 2-dimensional surface Σ . Source: Script by N. Boulanger

During a motion, the particles are assumed to remain on the surface defined by the constraints. If this

motion dx_i^α takes place during an infinitesimally small time interval dt , this means that

$$dG_m = dx_i^\alpha \frac{\partial G_m}{\partial x_i^\alpha} + dt \frac{\partial G_m}{\partial t} = 0. \quad (2.2)$$

In other words, the constraints are constraints on the trajectories of the particles and give rise by differentiation with respect to time to constraints on the velocities,

$$\dot{x}_i^\alpha \frac{\partial G_m}{\partial x_i^\alpha} + \frac{\partial G_m}{\partial t} = 0, \quad (2.3)$$

where, for a trajectory $x_i^\alpha(t)$, the notation $\dot{x}_i^\alpha = \frac{dx_i^\alpha}{dt}$ and $\ddot{x}_i^\alpha = \frac{d^2 x_i^\alpha}{(dt)^2}$ is used.

Consider two trajectories $x_i^\alpha(t)$ and $\tilde{x}_i^\alpha(t)$. What we will be interested in below is (instantaneous) *virtual displacements* i.e., the difference of these trajectories $\Delta x_i^\alpha(t) = \tilde{x}_i^\alpha(t) - x_i^\alpha(t)$ at fixed times t . It follows that

$$\begin{aligned} \Delta \frac{dx_i^\alpha(t)}{dt} &= \Delta \lim_{\epsilon \rightarrow 0} \frac{x_i^\alpha(t + \epsilon) - x_i^\alpha(t)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{\tilde{x}_i^\alpha(t + \epsilon) - x_i^\alpha(t + \epsilon) - (\tilde{x}_i^\alpha(t) - x_i^\alpha(t))}{\epsilon} \\ &= \frac{d\tilde{x}_i^\alpha(t)}{dt} - \frac{dx_i^\alpha(t)}{dt} = \frac{d}{dt} \Delta x_i^\alpha(t). \end{aligned} \quad (2.4)$$

More specifically we will be interested in infinitesimal displacements δx_i^α of this type that are compatible with the constraints,

$$\delta x_i^\alpha \frac{\partial G_m}{\partial x_i^\alpha} \approx 0. \quad (2.5)$$

Lemma 3. *The elementary work δW of the reaction forces due to the constraints vanishes for all infinitesimal virtual displacements that are compatible with the constraints if and only if the reaction forces are a combination of the gradients of the constraints,*

$$\delta W = R_\alpha^i \delta x_i^\alpha \approx 0 \iff R_\alpha^i \approx \mu^m \frac{\partial G_m}{\partial x_i^\alpha}, \quad (2.6)$$

for some $\mu^m(x_i^\alpha, t)$.

Indeed, except for the interpretation in terms of work and forces, this follows directly from the previous lemma. The notation δW means that this elementary work is not necessarily the variation of a function. The geometric interpretation is that the reaction forces are orthogonal to the constraints. (For each m , the gradient $\frac{\partial G_m}{\partial x_i^\alpha}$ is a vector that is orthogonal to the tangent vectors of the surface defined by $G_m = 0$.)

In what follows, we limit ourselves (almost) exclusively to such constraints, i.e., to constraints whose reaction forces do not do any work during a virtual displacement. In particular this means that there can be no friction due to these constraints. *Such constraints are called ideal.*

2.2 d'Alembert's theorem

The dynamics of the $\mathcal{N}$ particles is governed by Newton's equations,

$$m_{(i)} \ddot{x}_i^\alpha = F_i^\alpha + R_i^\alpha, \quad (2.7)$$

where the parenthesis around the index i means that there is no summation, and where F_i^α are the components of the applied forces acting on the i th particles, while R_i^α are the reaction forces due to the constraints, assumed to be ideal. The latter keep the particles confined to the constraint surface.

Theorem 1 (d'Alembert). *Newton's equation for a system of point particles subjected to ideal constraints are equivalent to the condition that the elementary work of the sum of the applied forces and of the forces of inertia ($-m_{(i)}\ddot{x}_i^\alpha$) vanishes for all infinitesimal virtual displacements compatible with the constraints,*

$$\boxed{(F_i^\alpha - m_{(i)}\ddot{x}_i^\alpha)\delta x_i^\alpha \approx 0}. \quad (2.8)$$

The theorem states that Newton's equations (2.7), together with the fact that the reaction forces associated to ideal constraints are of the form $R_\alpha^i \approx \mu^m \frac{\partial G_m}{\partial x_i^\alpha}$ on account of equation (2.6) of Lemma 3 are equivalent to (2.8) when the constraints (2.1) hold, and when limiting oneself to compatible infinitesimal virtual displacements, (2.5). That the condition is sufficient ($\implies$) follows by contracting (2.7) with δx_i^α and using (2.6) together with (2.5) to conclude that (2.8) holds. That the condition is necessary ($\impliedby$) follows because (2.1), (2.5) together with (2.8) imply, as in lemma 3, that $F_i^\alpha - m_{(i)}\ddot{x}_i^\alpha \approx -\mu^m \frac{\partial G_m}{\partial x_i^\alpha}$, for some μ^m , which are Newton's equation (2.7), together with (2.6).

2.3 Principle of virtual work

If one uses d'Alembert's theorem in the particular case of a system at equilibrium, the accelerations vanish, $\ddot{x}_i^\alpha = 0$, and the theorem reduces to the

Theorem 2 (Principle of virtual work). *The virtual work done by the applied forces that keep a system in equilibrium are zero.*

2.3.1 Worked Exercise 1: The Lever

Using the principal of virtual work, find the equilibrium position of a horizontal lever of fulcrum O , with a downward vertical force of modulus A at a distance a of O to the left and a downward vertical force of modulus B at a distance b to the right of O . Hint: The equilibrium condition links A, B, a, b . Relate the vertical displacements of the extremities δs_A and δs_B to the rotation angle $\delta\varphi$ (see figure).

Solution: the condition that the virtual work of the applied forces should vanish is

$$\vec{A} \cdot \vec{\delta s}_A + \vec{B} \cdot \vec{\delta s}_B = 0 \iff A\delta s_A + B\delta s_B = 0, \quad (2.9)$$

Furthermore $\delta s_A = -a\delta\varphi$, $\delta s_B = b\delta\varphi$. Hence

$$(-Aa + Bb)\delta\varphi = 0, \forall \delta\varphi \implies \boxed{Aa = Bb}. \quad (2.10)$$

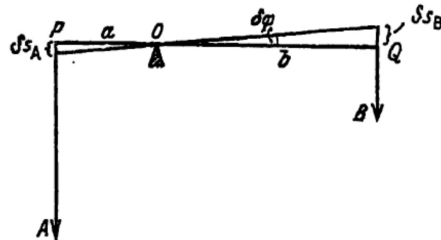
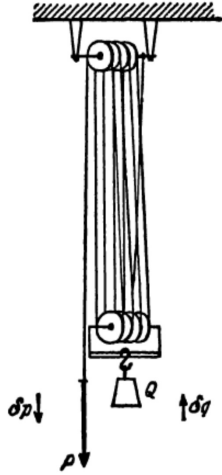
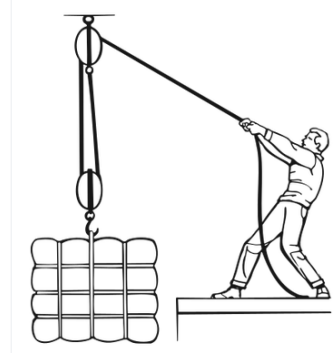


FIG. 10a. Lever with arms a and b under vertical loads A and B ,

Figure 2.2: Lever. Source: Sommerfeld Mechanics, chapter II.9



(a) Block and tackle. Source: Sommerfeld, Mechanics, chapter II.9



(b) Block and tackle. Source: Wikipedia

2.3.2 Worked Exercise 2: Block and Tackle (Palan)

In a tackle with n pulleys at the fixed upper end and n pulleys at the moving lower end lower, let Q be the modulus of the weight force that one wishes to lift, and P the modulus of the lifting force applied at the to the other end of the rope. In a virtual displacement, let δp be the displacement of the end of the rope and $-\delta q$ the displacement of the load (which is negative if δp is positive, see figure).

1. When taking the tackle as an ideal system without friction, what does the principle of virtual work imply ?
2. If you lift the load by δq , each of the $2n$ lengths of rope between the pulleys is shortened by δq .
3. What is the relation between δp and δq ?
4. Find the link between Q , P and the number of pulleys n at the upper and lower end of the tackle to maintain the load suspended.

Solution:

Application of the principle:

$$P\delta p - Q\delta q = 0. \quad (2.11)$$

For 1 pulley at both the upper and the lower end, if one lifts by δq , each of the 2 pieces of rope is shortened by δq so that $\delta p = 2\delta q$.

For 2 pulleys at both the upper and the lower end, if one lifts by δq , each of the 4 pieces of rope is shortened by δq so that $\delta p = 4\delta q$.

For n pulleys at both the upper and the lower end, $\delta p = 2n\delta q$ and

$$(2nP - Q)\delta q = 0, \forall \delta q \implies \boxed{P = \frac{Q}{2n}}. \quad (2.12)$$

2.4 Non-holonomic constraints

Non-holonomic constraints are constraints on the velocities of point particles as in (2.3), but that do not originate from holonomic constraints, i.e., from constraints on positions, through differentiation with respect to time.

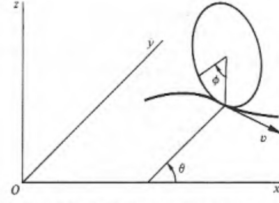


FIGURE 1.5 Vertical disk rolling on a horizontal plane.

Figure 2.4: Source: Goldstein Classical Mechanics, chapter 1.3

As an example, consider a vertical disk or coin rolling without sliding on a horizontal plane, $z = 0$ in standard Cartesian coordinates. Let (x_C, y_C, z_C) the coordinates of the center of the disk. Only x_C, y_C are relevant since $z_C = a$, with a the radius of the disk. Let ϕ be the angle of rotation of the disk starting from a fixed initial direction and θ the angle made by $\vec{1}_x$ and the axes of rotation of the disk. Since the velocity $\vec{v}$ of the center of the disk is normal to this axes, and lies in a plane of constant z , it follows that the angle between $\vec{1}_x$ and $\vec{v}$ is $\theta - \frac{\pi}{2}$, so that

$$x_C = v \cos(\theta - \frac{\pi}{2}) = v \sin \theta, \quad y_C = v \sin(\theta - \frac{\pi}{2}) = -v \cos \theta. \quad (2.13)$$

Futhermore, the velocity of a point on the edge of the disk is $a\dot{\phi}$. This is the module of the velocity of the point of contact of the disk with the plane, and since there is no sliding this is also the velocity of the center of the disk, $v = a\dot{\phi}$. By injecting into the previous relations, we thus get the following relations between velocities for this motion

$$\begin{cases} \dot{x}_C - a \sin \theta \dot{\phi} = 0, \\ \dot{y}_C + a \cos \theta \dot{\phi} = 0, \end{cases} \quad (2.14)$$

and the question is whether these constraints on velocities come from constraints on positions by differentiation with respect to time.

More generally, consider r constraints that are at most linear in velocities $\dot{x}^a$,

$$K_a^m \dot{x}^a + K_0^m = 0, \quad (2.15)$$

where $K_a^m(x^b, t), K_0^m(x^b, t)$ are functions that depend on the coordinates x^a , $a = 1, \dots, n$ (which may or may not simply be the positions x_i^α in Cartesian space of the $\mathcal{N}$ particles), and possibly an explicit dependence on time t . The question then is whether the relation on the velocities (2.15), or any equivalent relation on the velocities

$$\Lambda^m_{m'} (K_a^{m'} \dot{x}^a + K_0^{m'}) = 0, \quad (2.16)$$

with $\Lambda^m_{m'}(x^b, t)$ an invertible matrix, can be written as

$$\Lambda^m_{m'} (K_a^{m'} \dot{x}^a + K_0^{m'}) = \frac{\partial G^m}{\partial x^a} \dot{x}^a + \frac{\partial G^m}{\partial t}. \quad (2.17)$$

for r functions $G^m(x^a, t)$. Since the velocities are arbitrary, this relation can be decomposed as

$$\begin{cases} \frac{\partial G^m}{\partial x^a} = \Lambda^m_{m'} K_a^{m'}, \\ \frac{\partial G^m}{\partial t} = \Lambda^m_{m'} K_0^{m'}, \end{cases} \quad (2.18)$$

In order to answer this question, it is useful to introduce $x^A = (t, x^a)$, $A = 0, \dots, n$, so that the above $n + 1$ equations can be written in a unified way as

$$\frac{\partial G^m}{\partial x^A} = \Lambda^m_{m'} K_A^{m'}. \quad (2.19)$$

Necessary conditions¹ for the existence of such functions G^m is that mixed second derivatives of the right hand sides vanish,

$$\frac{\partial^2 G_m}{\partial x^A \partial x^B} = \frac{\partial^2 G_m}{\partial x^B \partial x^A} \implies \frac{\partial}{\partial x^A} (\Lambda^m_{m'} K_B^{m'}) = \frac{\partial}{\partial x^B} (\Lambda^m_{m'} K_A^{m'}). \quad (2.20)$$

This is however not the condition that we need because it still contains the unknown functions $\Lambda^m_{m'}$. What we need is a condition on the functions K_A^m alone. A standard technical assumption on these functions, that we assume to hold is that they are linearly independent in the following sense: if there exist some functions $\lambda_m(x^B)$ such that $\lambda_m K_A^m = 0$ then necessarily $\lambda_m = 0$. The condition we need is then a consequence of the following theorem.

Theorem 3 (Frobenius). *For linearly independent functions K_A^m , the set of conditions*

$$\left(\frac{\partial}{\partial x^{[A_1}} K_{A_2}^m \right) K_{A_3}^1 \dots K_{A_{r+2}}^r = 0, \quad (2.21)$$

is necessary and sufficient for the existence of functions $\Lambda^m_{m'}(x^B)$ (invertible) and $G^m(x^B)$ such that

$$\frac{\partial G^m}{\partial x^A} = \Lambda^m_{m'} K_A^{m'}. \quad (2.22)$$

In equation (2.21), the square bracket denotes complete antisymmetrization of the indices $A_1, \dots, A_{r+2}$, divided by $(r + 2)!$, the number of permutations of $r + 2$ objects. More explicitly,

$$\left(\frac{\partial}{\partial x^{[A_1}} K_{A_2}^m \right) K_{A_3}^1 \dots K_{A_{r+2}}^r = \frac{1}{(r + 2)!} \sum_{\sigma \in P_{r+2}} (1)^{|\sigma|} \left(\frac{\partial}{\partial x^{A_{\sigma(1)}}} K_{A_{\sigma(2)}}^m \right) K_{A_{\sigma(3)}}^1 \dots K_{A_{\sigma(r+2)}}^r, \quad (2.23)$$

where P_{r+2} denotes a permutations of $r + 2$ elements and $|\sigma| = 0$ if the permutation is even, while $|\sigma| = 1$ if the permutation is odd.

Let us prove that these conditions are indeed necessary ($\Leftarrow$). We start from

$$0 = \left(\frac{\partial}{\partial x^{[A_1}} \frac{\partial G^m}{\partial x^{A_2}} \right) K_{A_3}^1 \dots K_{A_{r+2}}^r, \quad (2.24)$$

which holds because the second order derivatives $\frac{\partial^2 G^m}{\partial x^{A_1} \partial x^{A_2}}$ are symmetric in A_1, A_2 , so that, when completely antisymmetrizing in the right hand side of equation (2.24), one finds always 0. Since we assume that (2.22) is true, we find by substitution

$$\begin{aligned} 0 &= \left(\frac{\partial}{\partial x^{[A_1}} (\Lambda^m_{m'} K_{A_2}^{m'}) \right) K_{A_3}^1 \dots K_{A_{r+2}}^r \\ &= \left(\frac{\partial}{\partial x^{[A_1}} \Lambda^m_{m'} \right) K_{A_2}^{m'} K_{A_3}^1 \dots K_{A_{r+2}}^r + \Lambda^m_{m'} \left(\frac{\partial}{\partial x^{[A_1}} K_{A_2}^{m'} \right) K_{A_3}^1 \dots K_{A_{r+2}}^r \end{aligned} \quad (2.25)$$

¹It can in fact be shown, as done later in the course, that these conditions are also locally sufficient.

The first term on the last line vanishes because the antisymmetry in $A_2, \dots, A_{r+2}$ implies the antisymmetry in $m', 1, \dots, r$, but m' takes a value between 1 and r so that one of the values arises twice and a completely antisymmetric expression then vanishes. Equation (2.21) then follows because $\Lambda^m_{m'}$ is invertible.

The proof that the condition is sufficient ($\implies$) can be found in [3] or [10]. $\square$

In the particular case of the disk, we have

$$x^A = (t, x_C, y_C, \phi, \theta), \quad (2.26)$$

and we have two constraints, $r = 2$. More explicitly,

$$\begin{aligned} K_0^1 = 0, \quad K_1^1 = 1, \quad K_2^1 = 0, \quad K_3^1 = -a \sin \theta = -a \sin x^4, \quad K_4^1 = 0, \\ K_0^2 = 0, \quad K_1^2 = 0, \quad K_2^2 = 1, \quad K_3^2 = a \cos \theta = a \cos x^4, \quad K_4^2 = 0. \end{aligned} \quad (2.27)$$

Let us show that the constraints are non-holonomic, i.e., that there is at least one of the conditions in (2.21) that is not satisfied. If we take $m = 1$, they become

$$\left(\frac{\partial}{\partial x^{[A_1}} K_{A_2]}^1 \right) K_{A_3}^1 K_{A_4]}^2 = 0. \quad (2.28)$$

Choosing $A_1 = 4, A_2 = 3, A_3 = 1, A_4 = 2$, one can check that only the first term in the complete antisymmetrization is non-vanishing, so that the result of the left hand side is given by $\frac{1}{4!}(-a \cos x^4)$ which does not vanish.

2.5 d'Alembert's principle

In the proof of d'Alembert theorem, nothing crucially depended on the fact that the constraints were holonomic. It can thus be extended to non-holonomic ones (by the replacement $\frac{\partial G^m}{\partial x^a} \rightarrow K_a^m$ in the proof). Suppose then that there are r non holonomic constraints

$$K_a^m \dot{x}^a + K_0^m = 0, \quad (2.29)$$

and consider infinitesimal virtual displacements δx^a that are compatible,

$$K_a^m \delta x^a = 0. \quad (2.30)$$

d'Alembert's principle then states that Newton's equation

$$m_{(a)} \ddot{x}^a = F^a + R^a, \quad (2.31)$$

together with the condition that the forces of reaction are of the form

$$R_a \approx \lambda_m K_a^m, \quad (2.32)$$

are equivalent to the condition that the sum of the applied forces and of the forces of inertia produce no work during any compatible infinitesimal virtual displacement δx^a ,

$$(F^a - m_{(a)} \ddot{x}^a) \delta x^a \approx 0. \quad (2.33)$$

Systems covered by d'Alembert's theorem and principle are

- systems of point particles that move on polished surfaces that are fixed or moving (holonomic constraints),

- motions of solids with one point having a prescribed trajectory
- sliding of solids on polished surfaces
- rolling without sliding of solids on rough surfaces
- fixed links between solids

What is not covered by the theorem and principle are constraints that involve friction or constraints that are expressed by inequalities.

Chapter 3

Hamilton and least action principles

In this chapter, Hamilton's principle and the least action principle are derived from d'Alembert's theorem. It is based on [4], chapter 2.1.

3.1 Natural and virtual trajectories

Consider trajectories $x^a = x^a(t)$ that pass through a fixed point at some initial time t_i and another fixed point at a final time t_f , $x^a(t_i) = x_i^a$, $x^a(t_f) = x_f^a$. A *natural* trajectory is a trajectory that satisfies Newton's equations together with the constraints. Virtual trajectories are trajectories that originate from a virtual displacement of a natural trajectory. As before, we consider infinitesimal virtual variations $\delta x^a(t)$.

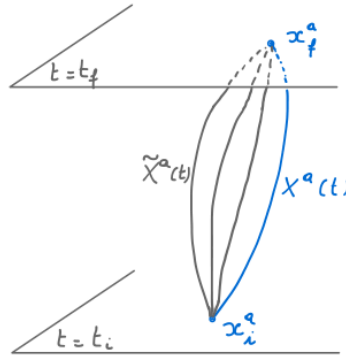


Figure 3.1: Natural and virtual trajectories. Source: Script by N. Boulanger

In particular, infinitesimal means that we compute to first order in δx^a . Furthermore, recall that such displacements commute with the time derivative, $\delta \dot{x}^a(t) = \frac{d}{dt} \delta x^a$.

3.2 Hamilton's principle

According to d'Alembert's theorem or principle, natural trajectories satisfy

$$(m_{(a)} \ddot{x}_a - F_a) \delta x^a \approx 0. \quad (3.1)$$

By an integrations by parts, (which means here using Leibniz rule in the way that is done when doing integrations by parts), we have

$$\frac{d}{dt}(m_{(a)}\dot{x}_a\delta x^a) - m_{(a)}\dot{x}_a\delta\dot{x}^a - F_a\delta x^a \approx 0. \quad (3.2)$$

In terms of the kinetic energy and its variation, we have

$$T = \frac{1}{2}m_{(a)}\dot{x}_a\dot{x}^a, \quad \delta T = m_{(a)}\dot{x}_a\delta\dot{x}^a. \quad (3.3)$$

The last term on the left hand side of (3.2) is the infinitesimal work done by the applied forces,

$$\delta W = F_a\delta x^a. \quad (3.4)$$

The use of the symbol δ is to indicate that in general $F_a\delta x^a$ is not necessarily integrable and the variation of a single function.

In these terms, D'Alembert's principle can be written as

$$\frac{d}{dt}(m_{(a)}\dot{x}_a\delta x^a) \approx \delta T + \delta W. \quad (3.5)$$

When integrating this relation along a natural trajectory, one gets

$$\left[m_{(a)}\dot{x}_a\delta x^a \right]_{t_i}^{t_f} \approx \int_{t_i}^{t_f} dt [\delta T + \delta W]. \quad (3.6)$$

Limiting oneself to virtual instantaneous variations that do not change the end points, $\delta x^a(t_i) = 0 = \delta x^a(t_f)$, we get:

Theorem 4 (Hamilton's principle). *A natural trajectory satisfies*

$$\int_{t_i}^{t_f} dt [\delta T + \delta W] \approx 0, \quad (3.7)$$

for all infinitesimal virtual variations compatible with the constraints and that keep the end points fixed.

3.3 Least action principle

In the particular case when the forces derive from a potential,

$$F_a = -\frac{\partial V}{\partial x^a} \iff F_a\delta x^a = -\delta V. \quad (3.8)$$

and $\delta W = -\delta V$. In this case, suitable integrability conditions

$$\frac{\partial}{\partial x^a}F_b - \frac{\partial}{\partial x^b}F_a = 0, \quad (3.9)$$

have to be satisfied. The force is conservative and the work does not depend on the path taken between the initial and the end points. We can write (3.7) as

$$\int_{t_i}^{t_f} dt \delta(T - V) = \delta \int_{t_i}^{t_f} dt (T - V) \approx 0, \quad (3.10)$$

where the variation commutes with integration since it is done at a fixed time. Defining the action, respectively the Lagrangian, as

$$S[x^a(t)] = \int_{t_i}^{t_f} dt L, \quad L = T - V, \quad (3.11)$$

Theorem 5 (Least action principle). *A natural trajectory is an extremal of the action,*

$$\boxed{\delta S \approx 0}, \quad (3.12)$$

for all infinitesimal virtual variations compatible with the constraints and that keep the end points fixed.

Remarks:

1. As in the finite dimensional case considered in the first chapter, where the problem of finding the extrema of a function of several variables, we thus have again an extremisation problem. The analogy is as follows: finding the extrema can be done by finding the points where the gradient vanishes. In the present case this corresponds to solving Newton's equations. Alternatively, one is looking for the points where the variation of the function vanishes. In the present case, this corresponds to finding the trajectories that extremize the variation of the action. The object whose extremals one is searching for is no longer a function, but a functional, that is to say, the integral of a function L of several functions $x^a(t)$ of time. Suppose for instance that $x^a(t_i) = x^a(t_f)$. One may then develop the trajectories in terms of Fourier series,

$$x^a(t) = \frac{c_0}{2} + \sum_{n=0}^{\infty} c_n^a \cos \frac{2\pi(t-t_i)}{t_f-t_i} + s_n^a \sin \frac{2\pi(t-t_i)}{t_f-t_i}, \quad (3.13)$$

so that, after substitution, the action becomes a function of the Fourier coefficients, $S[x^a(t)] = S[c_0^a, c_n^a, s_n^a]$ and the extremisation problem becomes that of a function of an infinite, countable number of variables.

2. How should one take the constraints into account ? Again there are two options. The first is to explicitly solve the constraints in terms of a number of independent variables $q^\alpha(t)$ for the trajectories, and reduce the extremisation problem to one in those variables. This is what we will do in detail in the next chapter. The second approach consists in working with Lagrange multipliers, that now also become trajectories, and consider the unconstrained extremisation problem for the extended action

$$S[x^a(t), \lambda^m(t)] = \int_{t_i}^{t_f} dt \left[L(x^a(t), \dot{x}^a(t)) - \lambda^m(t) G_m(x^a(t), t) \right]. \quad (3.14)$$

Chapter 4

Lagrangian mechanics

4.1 Lagrange coordinates

In the case of r independent holonomic constraints,

$$G_m(x^a, t) = 0, \quad m = 1, \dots, r, \quad (4.1)$$

that satisfy the regularity assumptions, there exists locally a change of coordinates

$$x^a = x^a(q^b, t), \quad q^a = q^a(x^b, t), \quad a, b = 1, \dots, 3\mathcal{N} = n, \quad (4.2)$$

such that

$$\begin{cases} q^a = q^{n-r+m} = G_m, & a = n - r + 1, \dots, n, \\ q^a = q^\alpha, & a, \alpha = 1, \dots, n - r. \end{cases} \quad (4.3)$$

The q^α are the independent coordinates, also called Lagrange coordinates. Satisfying the constraints in this coordinate system means to put the last r coordinates to 0, $q^{n-r+m} = 0$. Furthermore, virtual displacements that are compatible with the constraints satisfy $\delta q^{n-r+m} = 0$ while the δq^α are arbitrary and unconstrained. In terms of the old variables, we thus have that, when the constraints hold, variations compatible with the constraints are given by

$$\delta x^a \approx \left. \frac{\partial x^a}{\partial q^\alpha} \right|_{q^{n-r+m}=0} \delta q^\alpha. \quad (4.4)$$

4.2 Lagrange equations

In terms of Lagrange coordinates, we have

$$\delta W \approx F_a \delta x^a \approx Q_\alpha \delta q^\alpha, \quad Q_\alpha = F_a \left. \frac{\partial x^a}{\partial q^\alpha} \right|_{q^{n-r+m}=0}, \quad (4.5)$$

where Q_α is called the generalized force vector of Lagrange. For notational simplicity, we omit the bar with the indication that one should impose $q^{n-r+m} = 0$. In order to work out the kinetic energy $T = \frac{1}{2} m_{(a)} \dot{x}_a \dot{x}^a$, in terms of the Lagrange coordinates, one uses

$$\dot{x}^a = \frac{\partial x^a}{\partial q^\beta} \dot{q}^\beta + \frac{\partial x^a}{\partial t}, \quad (4.6)$$

so that the kinetic energy becomes explicitly

$$T(q^\alpha, \dot{q}^\alpha, t) = \frac{1}{2} m_{(a)} \left(\frac{\partial x^a}{\partial q^\alpha} \dot{q}^\alpha + \frac{\partial x^a}{\partial t} \right)^2 = m_{(a)} \left[\frac{1}{2} \frac{\partial x_a}{\partial q^\alpha} \frac{\partial x^a}{\partial q^\beta} \dot{q}^\alpha \dot{q}^\beta + \frac{\partial x_a}{\partial q^\alpha} \frac{\partial x^a}{\partial t} \dot{q}^\alpha + \frac{1}{2} \frac{\partial x_a}{\partial t} \frac{\partial x^a}{\partial t} \right], \quad (4.7)$$

while its variation is

$$\delta T = \frac{\partial T}{\partial q^\alpha} \delta q^\alpha + \frac{\partial T}{\partial \dot{q}^\alpha} \delta \dot{q}^\alpha = \left(\frac{\partial T}{\partial q^\alpha} - \frac{d}{dt} \frac{\partial T}{\partial \dot{q}^\alpha} \right) \delta q^\alpha + \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}^\alpha} \delta q^\alpha \right). \quad (4.8)$$

For any function $F(q^\alpha, \dot{q}^\alpha, t)$, the Euler-Lagrange derivative is defined by

$$\frac{\delta F}{\delta q^\alpha} = \frac{\partial F}{\partial q^\alpha} - \frac{d}{dt} \frac{\partial F}{\partial \dot{q}^\alpha}. \quad (4.9)$$

Again, in the context of Hamilton's or the least action principle, the last term of (4.8) does not contribute when restricting oneself to variations vanishing at the end points. Hamilton's principle then becomes

$$\int_{t_i}^{t_f} dt \left[\frac{\delta T}{\delta q^\alpha} + Q_\alpha \right] \delta q^\alpha(t) = 0, \quad (4.10)$$

for all *unconstrained* $\delta q^\alpha(t)$ (that vanish at the end points). The du Bois-Reymond lemma then implies the Lagrange equations

$$\boxed{\frac{\delta T}{\delta q^\alpha} = -Q_\alpha}. \quad (4.11)$$

That these differential equations are equivalent to Newton's equations for ideal holonomic constraints can also be seen directly as follows. One starts from

$$\begin{cases} m_{(a)} \ddot{x}_a = F_a + \mu^m \frac{\partial G_m}{\partial x^a} \\ G_m = 0 \end{cases}. \quad (4.12)$$

The first set of equations are contracted with $\frac{\partial x^a}{\partial q^\alpha}$, and then the constrained are imposed $q^{n-r+m} = 0$. The constraints are then trivially fulfilled. When using that

$$\frac{\partial G_m}{\partial x^a} \frac{\partial x^a}{\partial q^\alpha} = 0, \quad (4.13)$$

and the definition of the Q_α , the first set of equations reduces to

$$m_{(a)} \ddot{x}_a \frac{\partial x^a}{\partial q^\alpha} = Q_\alpha, \quad (4.14)$$

with the understanding that $q^{n-r+m} = 0$. It remains to be shown that

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}^\alpha} - \frac{\partial T}{\partial q^\alpha} = m_{(a)} \ddot{x}_a \frac{\partial x^a}{\partial q^\alpha} \iff \frac{d}{dt} \left(\dot{x}_a \frac{\partial \dot{x}^a}{\partial \dot{q}^\alpha} \right) - \dot{x}_a \frac{\partial \dot{x}^a}{\partial q^\alpha} = \frac{d}{dt} \left(\dot{x}_a \frac{\partial x^a}{\partial q^\alpha} \right) - \dot{x}_a \frac{d}{dt} \frac{\partial x^a}{\partial q^\alpha}. \quad (4.15)$$

This is satisfied since it follows directly from (4.6) that

$$\frac{\partial \dot{x}^a}{\partial \dot{q}^\alpha} = \frac{\partial x^a}{\partial q^\alpha}, \quad \frac{\partial \dot{x}^a}{\partial q^\alpha} = \frac{\partial^2 x^a}{\partial q^\alpha \partial q^\beta} \dot{q}^\beta + \frac{\partial^2 x^a}{\partial q^\alpha \partial t} = \frac{d}{dt} \frac{\partial x^a}{\partial q^\alpha}. \quad (4.16)$$

4.3 Dissipative systems

In some dissipative systems there exists, besides the applied and the reaction forces, also friction forces proportional to the velocities. In this case,

$$\begin{cases} m_{(a)}\ddot{x}_a = F_a + \mu^m \frac{\partial G_m}{\partial x^a} - \zeta_{ab}\dot{x}^b \\ G_m = 0 \end{cases}, \quad (4.17)$$

When contracting with $\frac{\partial x^a}{\partial q^\alpha}$, the additional term on the right hand side is

$$-\zeta_{ab}\dot{x}^b \frac{\partial x^a}{\partial q^\alpha}. \quad (4.18)$$

If the constant friction coefficients are symmetric, $\zeta_{ab} = \zeta_{ba}$, one may write the associated Lagrange equations in terms of the Rayleigh's dissipation function

$$D = \frac{1}{2}\zeta_{ab}\dot{x}^a\dot{x}^b, \quad (4.19)$$

as

$$\boxed{\frac{\delta T}{\delta q^\alpha} = -Q_\alpha - \frac{\partial D}{\partial \dot{q}^\alpha}}, \quad (4.20)$$

because

$$\frac{\partial D}{\partial \dot{q}^\alpha} = \zeta_{ab} \frac{\partial \dot{x}^a}{\partial \dot{q}^\alpha} \dot{x}^b = \zeta_{ab}\dot{x}^b \frac{\partial x^a}{\partial q^\alpha}. \quad (4.21)$$

4.4 Euler-Lagrange equations

In the absence of such friction forces and if all applied forces derive from a potential $V(x^a)$, natural trajectories have to satisfy the Euler-Lagrange equations,

$$\boxed{\frac{\delta L}{\delta q^\alpha} = 0}. \quad (4.22)$$

These equations can either be obtained from the Lagrange equations (4.11) by using that $L = T - V$ and

$$Q_\alpha = F_a \frac{\partial x^a}{\partial q^\alpha} = -\frac{\partial V}{\partial x^a} \frac{\partial x^a}{\partial q^\alpha} = -\frac{\partial V}{\partial q^\alpha}, \quad \frac{\delta V}{\delta q^\alpha} = \frac{\partial V}{\partial q^\alpha}, \quad (4.23)$$

where $V(q^\alpha, t) = V(x^a(q^\alpha, t))$. Alternatively, one derive the Euler-Lagrange equations directly from the least action principle:

$$0 = \delta S = \int_{t_i}^{t_f} dt \left(\frac{\partial L}{\partial q^\alpha} \delta q^\alpha + \frac{\partial L}{\partial \dot{q}^\alpha} \delta \dot{q}^\alpha \right) = \left[\frac{\partial L}{\partial \dot{q}^\alpha} \delta q^\alpha \right]_{t_i}^{t_f} + \int_{t_i}^{t_f} dt \frac{\delta L}{\delta q^\alpha} \delta q^\alpha. \quad (4.24)$$

The passage from the first to the second equality involves an integrations by parts, the boundary term vanishes because of the assumption that the variations vanish at the end points, and the Euler-Lagrange equations then follow by using the fact that the variations $\delta q^\alpha(t)$ are arbitrary and by the du Bois-Reymond lemma.

Remarks

- In some problems, the potential may depend on velocities, $V = V(q^\alpha, \dot{q}^\alpha, t)$ or, more generally, the Lagrangian may be of form $L(q^\alpha, \dot{q}^\alpha, t)$, without a direct interpretation in terms of T and V . In this case, the least action principle for natural trajectories still gives rise to the Euler-Lagrange equations for these trajectories.
- When making explicit the second order time derivatives in the left hand side of the Euler-Lagrange derivatives,

$$-\frac{\delta L}{\delta q^\alpha} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} - \frac{\partial L}{\partial q^\alpha} = \frac{\partial^2 L}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} \ddot{q}^\beta - f_\alpha(q^\beta, \dot{q}^\beta, t), \quad (4.25)$$

it follows that the Euler-Lagrange equations are equivalent to differential equations in normal form if and only if

$$\left| \frac{\partial^2 L}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} \right| \neq 0. \quad (4.26)$$

We call this the regular case. In this case, the inverse matrix $M^{\gamma\alpha} \frac{\partial^2 L}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} = \delta_\beta^\gamma$ where $M^{\gamma\alpha} = M^{\gamma\alpha}(q^\beta, \dot{q}^\beta, t)$, exists and the normal form of the Euler-Lagrange equations is

$$\ddot{q}^\gamma = F^\gamma, \quad F^\gamma(q^\alpha, \dot{q}^\alpha, t) = M^{\gamma\alpha} f_\alpha. \quad (4.27)$$

Unless otherwise specified, we will always be in the regular case in the rest of the course.

- It follows from (4.24) that one may define the functional derivative of the action with respect to a trajectory $q^\alpha(t)$ by

$$\frac{\delta S}{\delta q^\alpha(t)} = \frac{\delta L}{\delta q^\alpha} \Big|_{q^\alpha=q^\alpha(t), \dot{q}^\alpha=\frac{d}{dt}q^\alpha(t)}. \quad (4.28)$$

- If $L' = cL$, $c \in \mathbb{R}$, the solutions of the Euler-Lagrange equations are unchanged.
- Suppose that $L' = L + \frac{df}{dt}$ for some $f = f(q^\alpha, t)$. It follows that $\frac{\delta L'}{\delta q^\alpha} = \frac{\delta L}{\delta q^\alpha}$. Indeed,

$$\frac{\delta}{\delta q^\alpha} \left(\frac{df}{dt} \right) = \frac{\delta}{\delta q^\alpha} \left(\frac{\partial f}{\partial q^\beta} \dot{q}^\beta + \frac{\partial f}{\partial t} \right) = \frac{\partial^2 f}{\partial q^\alpha \partial q^\beta} \dot{q}^\beta + \frac{\partial^2 f}{\partial q^\alpha \partial t} - \frac{d}{dt} \frac{\partial f}{\partial q^\alpha} = 0. \quad (4.29)$$

Conversely, suppose that $\frac{\delta L'}{\delta q^\alpha} = \frac{\delta L}{\delta q^\alpha}$, or if $M(q^\alpha, \dot{q}^\alpha, t) = L' - L$, that¹ $\frac{\delta M}{\delta q^\alpha} = 0$. It follows that $M = \frac{df}{dt}$ with $f = f(q^\alpha, t)$. Indeed, in $\frac{\delta M}{\delta q^\alpha}$ the only terms that involve second order derivatives are given by $-\frac{\partial^2 M}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} \ddot{q}^\beta$ and have to vanish separately. It thus follows that $\frac{\partial^2 M}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} = 0$. When writing

$$\begin{aligned} M(q^\alpha, \dot{q}^\alpha, t) &= M(0, 0, t) + \int_0^1 d\tau \frac{d}{d\tau} M(\tau q^\alpha, \tau \dot{q}^\alpha, t) \\ &= \frac{d}{dt} \int_0^t dt' M(0, 0, t') + \int_0^1 \frac{d\tau}{\tau} (q^\alpha \frac{\partial M}{\partial q^\alpha} + \dot{q}^\alpha \frac{\partial M}{\partial \dot{q}^\alpha})(\tau q^\alpha, \tau \dot{q}^\alpha, t). \end{aligned} \quad (4.30)$$

By the usual manipulation, we may write

$$q^\alpha \frac{\partial M}{\partial q^\alpha} + \dot{q}^\alpha \frac{\partial M}{\partial \dot{q}^\alpha} = q^\alpha \frac{\delta M}{\delta q^\alpha} + \frac{d}{dt} (q^\alpha \frac{\partial M}{\partial \dot{q}^\alpha}), \quad (4.31)$$

¹Note that in this case, we consider this equation to be an identity in the variables $q^\alpha, \dot{q}^\alpha, t$ and not a system of differential equations for trajectories.

where the first term vanishes by assumption and $N = q^\alpha \frac{\partial M}{\partial \dot{q}^\alpha}$ does not depend on $\dot{q}^\beta$ since we have already shown that $\frac{\partial^2 M}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} = 0$. Hence $N = N(q^\alpha, t)$. Furthermore

$$\begin{aligned} \left(\frac{d}{dt}N\right)(\tau q^\alpha, \tau \dot{q}^\alpha, t) &= (\dot{q}^\alpha \frac{\partial N}{\partial q^\alpha} + \frac{\partial N}{\partial t})(\tau q^\alpha, \tau \dot{q}^\alpha, t) = \tau \dot{q}^\alpha \frac{\partial N}{\partial q^\alpha}(\tau q^\alpha, t) + \frac{\partial N}{\partial t}(\tau q^\alpha, t) \\ &= \frac{d}{dt}N(\tau q^\alpha, t). \end{aligned} \quad (4.32)$$

Now $\frac{d}{dt}$ can be taken out of the integral and

$$M = \frac{df}{dt}, \quad f(q^\alpha, t) = \int^t dt' M(0, 0, t') + \int_0^1 \frac{d\tau}{\tau} N(\tau q^\alpha, t). \quad (4.33)$$

4.4.1 Worked Exercise 3: A Lagrangian of sufficient generality

Consider a Lagrangian of the form

$$L = \frac{1}{2} g_{\beta\gamma} \dot{q}^\beta \dot{q}^\gamma + A_\beta \dot{q}^\beta - V, \quad (4.34)$$

where $g_{\beta\gamma} = g_{\gamma\beta}$, A_β , V are functions of q^δ and t only.

Show that in the regular case, the “metric” $g_{\beta\gamma}$ is invertible. Show that the left hand sides of the Euler-Lagrange equations may be written as

$$-\frac{\delta L}{\delta q^\alpha} = g_{\alpha\beta} \ddot{q}^\beta + \Gamma_{\alpha\beta\gamma} \dot{q}^\beta \dot{q}^\gamma - F_{\alpha\beta} \dot{q}^\beta + \frac{\partial V}{\partial q^\alpha} + \frac{\partial g_{\alpha\beta}}{\partial t} \dot{q}^\beta + \frac{\partial A_\alpha}{\partial t}, \quad (4.35)$$

where the “Christoffel symbols of the first kind” $\Gamma_{\alpha\beta\gamma}$ and the “field strengths” $F_{\alpha\beta}$ are given by

$$\Gamma_{\alpha\beta\gamma} = \frac{1}{2} \left(\frac{\partial g_{\alpha\beta}}{\partial q^\gamma} + \frac{\partial g_{\alpha\gamma}}{\partial q^\beta} - \frac{\partial g_{\beta\gamma}}{\partial q^\alpha} \right), \quad F_{\alpha\beta} = \frac{\partial A_\beta}{\partial q^\alpha} - \frac{\partial A_\alpha}{\partial q^\beta}. \quad (4.36)$$

Solution: The first part follows from

$$\frac{\partial^2 L}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} = g_{\alpha\beta}. \quad (4.37)$$

For the second, we compute

$$-\frac{\partial L}{\partial q^\alpha} = -\left(\frac{1}{2} \frac{\partial g_{\beta\gamma}}{\partial q^\alpha} \dot{q}^\beta \dot{q}^\gamma + \frac{\partial A_\beta}{\partial q^\alpha} \dot{q}^\beta - \frac{\partial V}{\partial q^\alpha} \right), \quad (4.38)$$

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} = g_{\alpha\beta} \ddot{q}^\beta + \frac{\partial g_{\alpha\beta}}{\partial q^\gamma} \dot{q}^\gamma \dot{q}^\beta + \frac{\partial A_\alpha}{\partial q^\beta} \dot{q}^\beta + \frac{\partial g_{\alpha\beta}}{\partial t} \dot{q}^\beta + \frac{\partial A_\alpha}{\partial t}. \quad (4.39)$$

The result follows by adding the terms on the right hand sides of these two equations and using

$$\frac{\partial g_{\alpha\beta}}{\partial q^\gamma} \dot{q}^\gamma \dot{q}^\beta = \frac{1}{2} \left(\frac{\partial g_{\alpha\beta}}{\partial q^\gamma} + \frac{\partial g_{\alpha\gamma}}{\partial q^\beta} \right) \dot{q}^\gamma \dot{q}^\beta. \quad (4.40)$$

For later use, let us note that a geometric representation of the kinetic term $T = \frac{1}{2} g_{\beta\gamma} \dot{q}^\beta \dot{q}^\gamma$ is obtained by considering the “line element”

$$\overline{ds} = \sqrt{|g_{\beta\gamma} dq^\beta dq^\gamma|}, \quad (4.41)$$

which defines a distance in configuration space and where, as in the example of the work of applied forces, the bar in the notation $\overline{ds}$ indicates that this is in general not an exact differential, and writes

$$T = \frac{1}{2} \sigma \left(\frac{\overline{ds}}{dt} \right)^2, \quad (4.42)$$

where σ denotes the sign of $g_{\beta\gamma} \dot{q}^\beta \dot{q}^\gamma$.

4.5 Points of equilibrium

Consider a Lagrangian of the form of the previous section, and assume that there is no explicit time dependence in the coefficients A_α . Equilibrium points are points $\bar{q}^\beta$ such that

$$\frac{\partial V}{\partial q^\alpha}(\bar{q}^\beta) = 0. \quad (4.43)$$

In this case, the trajectories $q^\beta(t) = \bar{q}^\beta$, with $\bar{q}^\beta$ constant in time, are natural trajectories. Indeed, when evaluated at this trajectories, which satisfy $\dot{q}^\beta = 0 = \ddot{q}^\beta$, the left hand side of the equations of motion (4.35) vanish since in addition the term $\frac{\partial A_\alpha}{\partial t}$ is absent and (4.43) holds.

If at the point of equilibrium, we have a local minimum of the potential,

$$\det \frac{\partial^2 V}{\partial q^\alpha \partial q^\beta}(\bar{q}) > 0 \quad (4.44)$$

the equilibrium is stable, if not it is unstable. Note that the potential that is relevant here is the potential in terms of the Lagrange coordinates.

Chapter 5

Applications of the Lagrangian formalism

5.1 The simple pendulum

Consider a point particle of mass m that is attached by means of a rigid bar of negligible mass to a fixed point O . It is submitted to the gravitational force. The reaction force of the bar on the point particle is perpendicular to the motion and thus produces no work.

We assume that the motion takes place in a vertical plane. For instance, this plane is determined by two straight lines, the first from O to the point particle when the pendulum is at equilibrium, and the second from O to the point where the particle is released with zero velocity.

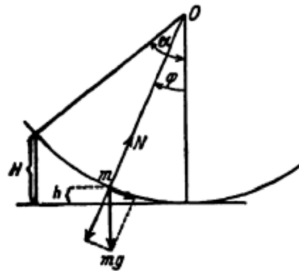


FIG. 24. Simple pendulum. Component of gravity along the direction of motion.

Figure 5.1: Simple pendulum. Source: Sommerfeld, Mechanics, chapter III.15

Let us choose coordinates on $\mathbb{R}^3$ where this plane is the plane (x, z) with O sitting at the origin, so that the motion is constrained to take place on the circle S of radius l centered at the origin.

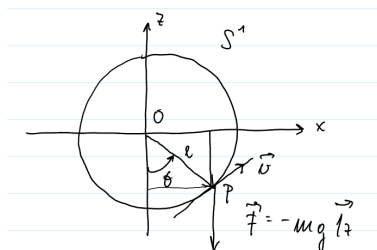


Figure 5.2: Coordinate system for planar motion

The equation for the constraint is

$$x^2 + z^2 = l^2. \quad (5.1)$$

Going to polar coordinates, with angle θ chosen as the deviation from the negative z axis,

$$x = r \sin \theta, \quad z = -r \cos \theta, \quad (5.2)$$

the constraint becomes

$$r - l = 0. \quad (5.3)$$

The Lagrange coordinate in this case is θ , and velocities are given by

$$\dot{x} = l \cos \theta \dot{\theta}, \quad \dot{z} = l \sin \theta \dot{\theta}. \quad (5.4)$$

In terms of the Lagrange coordinate, the kinetic energy is

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{z}^2) = \frac{1}{2}m(\dot{x}^2 + \dot{z}^2) = \frac{1}{2}ml^2\dot{\theta}^2 \quad (5.5)$$

while the potential energy is

$$V = mgz = -mgl \cos \theta. \quad (5.6)$$

The Lagrangian is

$$L = T - V = m\left(\frac{1}{2}l^2\dot{\theta}^2 + gl \cos \theta\right). \quad (5.7)$$

The Euler-Lagrange derivative is

$$\frac{\partial L}{\partial \theta} - \frac{d}{dt} \frac{\partial L}{\partial \dot{\theta}} = -mgl \sin \theta - ml^2\ddot{\theta}, \quad (5.8)$$

and the equation of motion is equivalent to

$$\boxed{\ddot{\theta} + \omega^2 \sin \theta = 0, \quad \omega^2 = \frac{g}{l}}. \quad (5.9)$$

5.1.1 Small oscillations

For small oscillations, $\theta \ll 1$, $\sin \theta \sim \theta$, the equation describes “harmonic” oscillations, with general solution given by

$$\theta(t) = A \sin(\omega t + \varphi), \quad \omega = \sqrt{\frac{g}{l}}. \quad (5.10)$$

5.1.2 Conservation of energy

In the general case, like in many systems, the energy which is given by $E = T + V$, or explicitly

$$E = m\left(\frac{1}{2}l^2\dot{\theta}^2 - gl \cos \theta\right), \quad (5.11)$$

in the current case, is a conserved quantity. By this one means that the energy is constant on all solutions to the equations of motion. Indeed,

$$\frac{d}{dt}E = ml^2\dot{\theta}(\ddot{\theta} + \frac{g}{l} \sin \theta) \quad (5.12)$$

vanishes for all solutions of (5.9).

Instead of solving the second order equation, whose general solution contains two integration constants, one may also solve the first order equation provided by (5.11), which is equivalent to

$$\dot{\theta}^2 = 2\omega^2 \left[\frac{E}{mgl} + \cos \theta \right], \quad (5.13)$$

in terms of E and an additional integration constant. Since the left hand side of this equation is always positive, it follows in particular that E must be such that

$$\frac{E}{mgl} \geq -\cos \theta \geq -1. \quad (5.14)$$

5.1.3 Points of equilibrium

The gradient of the potential

$$\frac{\partial V}{\partial \theta} = mgl \sin \theta \quad (5.15)$$

vanishes if and only if

$$\theta = 0 \quad \text{or} \quad \theta = \pi. \quad (5.16)$$

These two possibilities define the points of equilibrium where the pendulum is at rest, $\theta(t) = 0$ or $\theta(t) = \pi$. Since

$$\frac{\partial^2 V}{(\partial \theta)^2} = mgl \cos \theta, \quad (5.17)$$

the former is a stable point of equilibrium whereas the latter is unstable.

5.1.4 Vanishing velocity

On account of relation (5.13), the velocity vanishes if

$$\cos \theta = -\frac{E}{mgl} \quad (5.18)$$

which is possible only if

$$-1 \leq \frac{E}{mgl} \leq 1. \quad (5.19)$$

Setting

$$\cos \alpha = -\frac{E}{mgl}, \quad 0 \leq \alpha \leq \pi, \quad (5.20)$$

we may write

$$\dot{\theta}^2 = 2\omega^2 (\cos \theta - \cos \alpha). \quad (5.21)$$

The velocity vanishes if $\theta = \pm \alpha$.

5.1.5 Motion if $\alpha = \pi$

In the case $\alpha = \pi$, the relation becomes

$$\dot{\theta}^2 = \frac{2g}{l}(\cos \theta + 1). \quad (5.22)$$

Since

$$\cos^2 \frac{\theta}{2} = \left(\frac{e^{i\frac{\theta}{2}} + e^{-i\frac{\theta}{2}}}{2} \right)^2 = \frac{e^{i\theta} + e^{-i\theta} + 2}{4} = \frac{1}{2}(1 + \cos \theta), \quad (5.23)$$

this becomes

$$\dot{\theta}^2 = \frac{4g}{l} \cos^2 \frac{\theta}{2}, \quad (5.24)$$

so that

$$\dot{\theta} = \pm 2\omega \cos \frac{\theta}{2} \iff \pm \omega dt = \frac{d\frac{\theta}{2}}{\cos \frac{\theta}{2}}. \quad (5.25)$$

By integration, it follows that

$$\pm \omega(t - t_0) = \left[\ln \tan\left(\frac{x}{4} + \frac{\pi}{4}\right) \right]_{\theta_0}^{\theta}. \quad (5.26)$$

Indeed,

$$\frac{d}{dx} \ln \tan\left(\frac{x}{4} + \frac{\pi}{4}\right) = \frac{\cos\left(\frac{x}{4} + \frac{\pi}{4}\right)}{4 \sin\left(\frac{x}{4} + \frac{\pi}{4}\right) \cos^2\left(\frac{x}{4} + \frac{\pi}{4}\right)} = \frac{1}{2 \sin\left(\frac{x}{2} + \frac{\pi}{2}\right)} = \frac{1}{2 \cos \frac{x}{2}}. \quad (5.27)$$

Let us assume that $t_0 = 0$, $\theta_0 = 0$, and $\dot{\theta}(t_0) > 0$. This means that one has to choose the plus sign in equations (5.25) and (5.26) and that the pendulum starts at the stable point of equilibrium with an initial velocity such that one may reach the unstable point of equilibrium. Equation (5.26) simplifies to

$$\omega t = \ln \tan\left(\frac{\theta}{4} + \frac{\pi}{4}\right), \quad (5.28)$$

and the time to reach $\theta = \pi$ diverges,

$$\lim_{\theta \rightarrow \pi} \ln \tan\left(\frac{\theta}{4} + \frac{\pi}{4}\right) = \infty. \quad (5.29)$$

5.1.6 Oscillations

Consider now the case where $0 < \alpha < \pi$, or equivalently, on account of (5.20), the case where $-1 < \frac{E}{mgl} < 1$.

From $\ddot{\theta} = -\frac{g}{l} \sin \theta$ and the fact that $\sin \theta > 0$ for $\theta = \alpha > 0$, it follows that the acceleration is negative while we have already seen that the velocity vanishes. Hence, if θ was increasing towards α , it will start decreasing after reaching α . For $\theta = -\alpha < 0$, $\sin \theta < 0$, the acceleration is positive and if θ was decreasing towards $-\alpha$, it will start increasing after having reaching $-\alpha$. We thus have an oscillation between $-\alpha$ and α .

Let us define

$$k^2 = \frac{1}{2}\left(1 + \frac{E}{mgl}\right) \iff k^2 = \frac{1}{2}(1 - \cos \alpha) \iff k = \sin \frac{\alpha}{2}, \quad 0 < k < 1. \quad (5.30)$$

The equation of motion (5.13) becomes

$$\dot{\theta}^2 = 2\omega^2[2k^2 + \cos \theta - 1] = 4\omega^2[k^2 - \sin^2 \frac{\theta}{2}]. \quad (5.31)$$

The integral to be computed is

$$\pm\omega(t - t_0) = \frac{1}{2} \int_{\theta_0}^{\theta} \frac{d\xi}{\sqrt{k^2 - \sin^2 \frac{\xi}{2}}}. \quad (5.32)$$

The change of variables $\sin \frac{\theta}{2} = k \sin \varphi$, $\sin \frac{\xi}{2} = k \sin \eta$, which implies that $\frac{1}{2} \cos \frac{\xi}{2} d\xi = k \cos \eta d\eta$, brings this integral to the form

$$\pm\omega(t - t_0) = \frac{1}{2} \int_{\varphi_0}^{\varphi} \frac{2k \cos \eta d\eta}{\cos \frac{\xi}{2} \sqrt{k^2 - k^2 \sin^2 \eta}} = \int_{\varphi_0}^{\varphi} \frac{d\eta}{\cos \frac{\xi}{2}} = \int_{\varphi_0}^{\varphi} \frac{d\eta}{\sqrt{1 - k^2 \sin^2 \eta}}. \quad (5.33)$$

In terms of the elliptic integral of first kind defined by

$$F(\varphi, k) = \int_0^{\varphi} \frac{d\eta}{\sqrt{1 - k^2 \sin^2 \eta}} = -F(-\varphi, k), \quad (5.34)$$

we have

$$\pm\omega(t - t_0) = F(\varphi, k) - F(\varphi_0, k). \quad (5.35)$$

The period T of the motion is determined as follows. Because $k = \sin \frac{\alpha}{2}$ and $\sin \frac{\theta}{2} = k \sin \varphi$, if $\theta = \pm\alpha$, $\varphi = \pm\frac{\pi}{2}$, while $\theta = 0$ corresponds to $\varphi = 0$. T is twice the time needed to go from $-\alpha$ to α . Indeed, for θ going from $-\alpha$ to $+\alpha$, $\dot{\theta} \geq 0$, and one uses the + sign in (5.35), so that the time needed is $F(\frac{\pi}{2}, k) - F(-\frac{\pi}{2}, k)$. For θ going from α to $-\alpha$, $\dot{\theta} \leq 0$, and one uses the minus sign in (5.35), so that the time needed is $-[F(-\frac{\pi}{2}, k) - F(\frac{\pi}{2}, k)]$.

$$\omega T = 2[F(\frac{\pi}{2}, k) - F(-\frac{\pi}{2}, k)] = 4F(\frac{\pi}{2}, k) = 4K(k), \quad (5.36)$$

where the complete elliptic integral is defined by

$$K(k) = F(\frac{\pi}{2}, k). \quad (5.37)$$

Suppose now that $k^2 = \sin^2 \frac{\alpha}{2} \ll 1$, which means that oscillations are small, and let us develop in $\sin^2 \frac{\alpha}{2}$. To first order, the period becomes

$$\begin{aligned} T &= \frac{4}{\omega} \int_0^{\frac{\pi}{2}} \frac{d\eta}{\sqrt{1 - \sin^2 \frac{\alpha}{2} \sin^2 \eta}} = \frac{4}{\omega} \int_0^{\frac{\pi}{2}} d\eta (1 + \frac{1}{2} \sin^2 \frac{\alpha}{2} \sin^2 \eta) + O(\sin^4 \frac{\alpha}{2}) \\ &= \frac{4}{\omega} (\frac{\pi}{2} + \frac{1}{4} \sin^2 \frac{\alpha}{2} \int_0^{\frac{\pi}{2}} d\eta (1 - \cos 2\eta)) + O(\sin^4 \frac{\alpha}{2}) = \frac{2\pi}{\omega} (1 + \frac{1}{4} \sin^2 \frac{\alpha}{2}) + O(\sin^4 \frac{\alpha}{2}). \end{aligned} \quad (5.38)$$

The leading term is $\frac{2\pi}{\omega}$, the period of the harmonic oscillator.

5.1.7 Revolutions

The second generic case is when $\frac{E}{mgl} > 1$, or equivalently $k > 1$. The motion becomes a revolution since $\dot{\theta}$ never vanishes, so it is either always strictly positive or strictly negative. θ will thus become bigger than π or smaller than $-\pi$. When setting $k_1 = \frac{1}{k}$, $k_1 < 1$ and equation (5.31) becomes

$$\dot{\theta}^2 = \frac{4\omega^2}{k_1^2} (1 - k_1^2 \sin^2 \frac{\theta}{2}). \quad (5.39)$$

Integrating, it now follows directly from the definition (5.34) that

$$\pm \frac{\omega}{k_1}(t - t_0) = \frac{1}{2} \int_{\theta_0}^{\theta} \frac{d\eta}{\sqrt{1 - k_1^2 \sin^2 \frac{\eta}{2}}} = F\left(\frac{\theta}{2}, k_1\right) - F\left(\frac{\theta_0}{2}, k_1\right). \quad (5.40)$$

In this case, a period of one revolution corresponds to the time needed to go from $\theta = -\pi$ to $\theta = \pi$,

$$T = 2 \frac{k_1}{\omega} F\left(\frac{\pi}{2}\right) = 2 \frac{k_1}{\omega} K(k_1). \quad (5.41)$$

5.2 The brachystochrone

The word means either “shortest time” or “the curve of quickest descent” and concerns a famous prize problem formulated by Bernoulli in the 17th century, and solved by himself, Newton and Leibniz. In the way it was solved by Bernoulli, it is famous in the history of mathematics because it has lead to the calculus of variations, which is at heart of the Lagrangian method.

5.2.1 Statement of the problem

Consider 2 points P and Q in a vertical plane, find the ideal constraint relating these two points that minimizes the time needed for a massive point particle sliding with no friction under the influence of gravity to reach Q when being released with zero velocity at P .

5.2.2 Main equations

Without loss of generality, we may choose a coordinate system such that the vertical plane is (x, z) with the arrival point Q at the origin, $Q : (0, 0)$, and the initial point somewhere above, to the left, $P : (x^*, z^*)$ with $x^* < 0$ and $z^* > 0$.

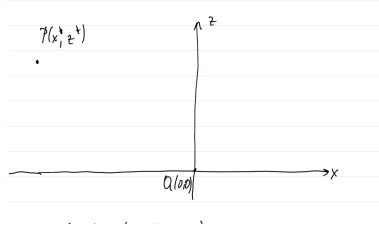


Figure 5.3: Coordinate system for sliding particle

The time needed to fall by a distance

$$\overline{ds} = \sqrt{dx^2 + dz^2} \quad (5.42)$$

is

$$|dt| = \frac{|\overline{ds}|}{v}, \quad (5.43)$$

where v is the magnitude of the velocity along the curve.

One may describe the curve that one is looking for by the graph of the function $x(z)$, with $x(0) = 0$ and $x(z^*) = x^*$, so that

$$dt = -\frac{\sqrt{1 + \left(\frac{dx}{dz}\right)^2}}{v} dz, \quad (5.44)$$

where the minus sign is due to the fact that if t increases, z decreases.

Conservation of energy gives

$$\frac{1}{2}mv_1^2 + mgz_1 = \frac{1}{2}mv_2^2 + mgz_2. \quad (5.45)$$

When applied to the current case with $P: (x_2, z_2) = (x^*, z^*)$ and $v_2 = 0$, while (x_1, z_1) is a generic point $(x(z), z)$ along the curve where the velocity is v , it follows that

$$\frac{1}{2}mv^2 = mg(z^* - z) \implies v = \sqrt{2g(z^* - z)}, \quad (5.46)$$

so that (5.44) becomes

$$dt = -\sqrt{\frac{1 + \left(\frac{dx}{dz}\right)^2}{2g(z^* - z)}} dz. \quad (5.47)$$

Integrating gives the time needed for the fall as

$$T = t_0 - t_* = -\int_{z^*}^0 \sqrt{\frac{1 + \left(\frac{dx}{dz}\right)^2}{2g(z^* - z)}} dz. \quad (5.48)$$

This time of fall is a functional of the function $x(z)$, $T = T[x(z)]$. But we know how to find extremals of such functionals. By the analogy $T \leftrightarrow S$, $x(z) \leftrightarrow q(\tau)$, the extremal with fixed end points is given functions for which the Euler-Lagrange derivative

$$\frac{\delta L}{\delta z} = \frac{\partial L}{\partial x} - \frac{d}{dz} \frac{\partial L}{\partial x'}, \quad x' = \frac{dx}{dz}, \quad (5.49)$$

vanishes with

$$L(x, x', z) = \sqrt{\frac{1 + \left(\frac{dx}{dz}\right)^2}{2g(z^* - z)}}. \quad (5.50)$$

Since there is no explicit dependence on x , the equation takes the form

$$\frac{d}{dz} \frac{\partial L}{\partial x'} = 0. \quad (5.51)$$

5.2.3 Finding the curve

This equation may be integrated once to get

$$\frac{\partial L}{\partial x'} = c, \quad (5.52)$$

for some integration constant c , or more explicitly,

$$\frac{1}{2} \sqrt{\frac{2g(z^* - z)}{1 + (x')^2}} \frac{2x'}{2g(z^* - z)} = c \iff \frac{x'}{\sqrt{(1 + (x')^2)2g(z^* - z)}} = c. \quad (5.53)$$

Taking the square in order to solve for x' gives

$$(x')^2 = c^2(1 + (x')^2)(2g(z^* - z)) \iff (x')^2(1 - 2gc^2(z^* - z)) = 2gc^2(z^* - z), \quad (5.54)$$

and then

$$x' = \pm \frac{c\sqrt{2g(z^* - z)}}{\sqrt{(1 - 2gc^2(z^* - z))}} = \pm \frac{\sqrt{z^* - z}}{\sqrt{2R - (z^* - z)}}, \quad 2R = \frac{1}{2gc^2}. \quad (5.55)$$

By integration, one gets

$$x - x^* = \pm \int_{z^*}^z \frac{\sqrt{z^* - z'}}{\sqrt{2R - (z^* - z')}} dz'. \quad (5.56)$$

Setting

$$Z = z^* - z = 2R \cos^2 \frac{\theta}{2} = R(1 + \cos \theta), \quad (5.57)$$

we have that z decreases from z^* to 0, $z^* \geq z \geq 0$, while Z increases from 0 to z^* , $0 \leq Z \leq z^* (\leq 2R)$, and then θ decreases from π to some θ_0 , $\pi \geq \theta \geq \theta_0 (\geq 0)$. The integral becomes

$$\begin{aligned} x - x^* &= \mp \int_0^{z^*-z} dZ \frac{\sqrt{Z}}{\sqrt{2R-Z}} = \pm \int_\pi^\theta d\xi 2R \cos \frac{\xi}{2} \sin \frac{\xi}{2} \frac{\sqrt{2R} \cos \frac{\xi}{2}}{\sqrt{2R} \sin \frac{\xi}{2}} = \pm 2R \int_\pi^\theta d\xi \cos^2 \frac{\xi}{2} \\ &= \pm R \int_\pi^\theta d\xi (1 + \cos \xi) = \pm R(\theta - \pi + \sin \theta), \quad (5.58) \end{aligned}$$

In order to fix the sign, starting from $\theta = \pi$, $z = z^*$, $Z = 0$, $x = x^*$, let us now take $x = x^* < 0$ and $\theta = \pi - \alpha$ for small positive α . We then have

$$z - z^* = -R(1 + \cos(\pi - \alpha)) = -R(1 - \cos \alpha) = -R(1 - (1 - \frac{\alpha^2}{2} + O(\alpha^4))) \sim -R \frac{\alpha^2}{2} < 0, \quad (5.59)$$

so that z indeed decreases when α increases. Furthermore,

$$x - x^* = \pm R(-\alpha + \sin(\pi - \alpha)) = \pm R(-\alpha + \sin \alpha) = \pm R(-\alpha + \alpha - \frac{\alpha^3}{3!} + O(\alpha^5)) \sim \mp R \frac{\alpha^3}{3!}, \quad (5.60)$$

Since we want x to increase when α increases, we need the lower (+) sign in the final expression, and thus the lower (-) sign in the final expression for $x - x^*$ in (5.58).

The parametric equations for the curve we are looking for are thus

$$\boxed{z - z^* = -R(1 + \cos \theta), \quad x - x^* = R(\pi - \theta - \sin \theta)}. \quad (5.61)$$

These are the equations for a cycloid, which is the curve followed by a point on the circumference of a disc that rolls without sliding on a line. More precisely, in the current case they describe an inverted cycloid.

Indeed, the equations for a standard cycloid are

$$x = a(t - \sin t), \quad y = a(1 - \cos t), \quad (5.62)$$

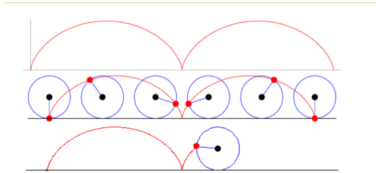


Figure 5.4: Standard cycloid. Source Wolfram MathWorld

The identification in our case relies on $t \rightarrow \pi - \theta$, $a \rightarrow R$, $x \rightarrow x - x^*$, $y \rightarrow -(z - z^*)$.

For a disc that rolls without sliding, one should have that the velocity of the center of the disc should be $v = R\dot{\theta} = \sqrt{(\frac{dx_C}{dt})^2 + (\frac{dz_C}{dt})^2}$, and thus, by multiplying by $\frac{dt}{d\theta}$, $R = \sqrt{(\frac{dx_C}{d\theta})^2 + (\frac{dz_C}{d\theta})^2}$. Here we have that the motion of the center of the disc is

$$x_C = R(\pi - \theta), \quad z_C = -R, \quad (5.63)$$

so that $\frac{dx_C}{d\theta} = -R$, $\frac{dz_C}{d\theta} = 0$, which indeed satisfies the required relation.

Each point of the circumference of the disc is animated by the composition of this translation with a rotation around the center of the disc at $(-R, 0)$,

$$x_{\text{circ}} = -R \sin \theta, \quad z_{\text{circ}} = -R \cos \theta. \quad (5.64)$$

5.2.4 Properties of the motion

We start at $\theta = \pi$, which corresponds to (x^*, z^*) , then θ decreases to θ_0 which corresponds to $(0, 0)$. This angle is determined by

$$z^* = R(1 + \cos \theta_0) \iff \frac{z^* - R}{R} = \cos \theta_0, \quad x^* = R(\pi - \theta_0 + \sin \theta_0). \quad (5.65)$$

It follows from (5.56) that $z^* - z \geq 0$ and $z^* - z \leq 2R$, so that, if $z = 0$, $0 \leq z^* \leq 2R$. From the first equation, it then follows that there exists $\pi \geq \theta_0 \geq 0$. The last equation then fixes the integration constant R in terms of the initial positions x^*, z^* .

The extremum for z is reached when $\frac{dz}{d\theta} = R \sin \theta = 0$, which implies that $\theta = 0$. It is a minimum since $\frac{d^2z}{(d\theta)^2} = R \cos \theta$, which is positive at $\theta = 0$.

Since in addition $\frac{dx}{d\theta} = -R(1 + \cos \theta)$, the velocity vanishes at $\theta = \pi$ and $\theta = -\pi$, and there is thus an oscillation between these two extreme values.

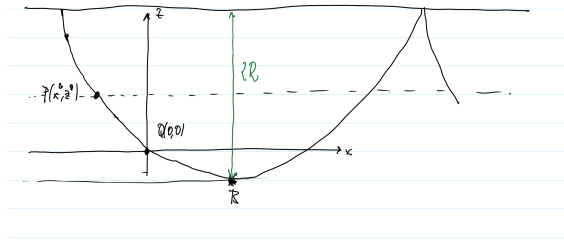


Figure 5.5: Cycloid oscillation

If $z^* = 2R$, $z - 2R = -R(1 + \cos \theta)$, the extremum is at $(0, 0)$ and the motion is along the whole branch of the cycloid.

Finally, let us study the temporal dependence starting from (5.47),

$$dt = -\sqrt{\frac{1 + \left(\frac{dx}{dz}\right)^2}{2g(z^* - z)}} dz \implies \frac{dt}{d\theta} = -\sqrt{\frac{\left(\frac{dz}{d\theta}\right)^2 + \left(\frac{dx}{d\theta}\right)^2}{2g(z^* - z)}}. \quad (5.66)$$

Injecting the values for the cycloid,

$$\frac{dt}{d\theta} = -\sqrt{\frac{(R \sin \theta)^2 + (R(1 + \cos \theta))^2}{2gR(1 + \cos \theta)}} = -\sqrt{\frac{R}{g}}. \quad (5.67)$$

The duration of the motion for a generic point on the cycloid is

$$t - t^* = -\sqrt{\frac{R}{g}}(\theta - \pi) = \sqrt{\frac{R}{g}}(\pi - \theta). \quad (5.68)$$

It follows that the duration of the fall from (x^*, z^*) to $(0, 0)$ is

$$t(\theta_0) - t^* = \sqrt{\frac{R}{g}}(\pi - \theta_0). \quad (5.69)$$

More importantly, a half period corresponds to the motion from π to $-\pi$,

$$t(-\pi) - t^* = \sqrt{\frac{R}{g}}2\pi. \quad (5.70)$$

Since this is true for all z^* , the period is independent of the amplitude, that is to say of the initial height, and the oscillation is isochronous.

A video of this state of affairs can be found here,

<https://www.facebook.com/TrustMyScience/videos/982002581956755/>.

Chapter 6

Hamiltonian formalism

In the regular case, the Euler-Lagrange equations are equivalent to a system of n second order equations

$$\ddot{q}^\alpha = f^\alpha(q^\beta, \dot{q}^\beta, t). \quad (6.1)$$

These equations may be turned into an equivalent system of $2n$ first order differential equations by introducing additional variables v^α and writing,

$$\dot{q}^\alpha = v^\alpha, \quad \dot{v}^\alpha = f^\alpha(q^\beta, v^\beta, t). \quad (6.2)$$

This system does not in general derive from a variational principle. One of the aims of the Hamiltonian formulation is to construct an equivalent system of $2n$ first order differential equations that does follow from a variational principle. This will be done through a suitable Legendre transform below.

6.1 Legendre transform

Take a function of one variable,

$$y = y(x), \quad (6.3)$$

and suppose that one is interested in the extrema of this function, i.e., the points $\bar{x}$ such that $y'(\bar{x}) = 0$. The problem that one would like to address is to represent this function by using the derivative

$$p = \frac{dy}{dx}, \quad (6.4)$$

as the dependent variable without losing any information. Note that at each point x , the derivative represents the slope of the tangent to the curve at that point, $\frac{dy}{dx} = \tan \theta$, with θ the angle of the tangent and the horizontal axis.

One possibility would be to invert the last equation to get $x = x(p)$, which is possible locally when $\frac{d^2y}{dx^2} \neq 0$, and then to substitute into (6.3). If the original function was $y = x^2$, one gets $p = 2x$, $x = \frac{1}{2}p$ and then $y = \frac{1}{4}p^2$. The problem is however that the relation

$$y = \frac{1}{4}\left(\frac{dy}{dx}\right)^2, \quad (6.5)$$

does not allow one to uniquely reconstruct the parabola, with its extremum at $\bar{x} = 0$. Indeed, separation of variables gives

$$\frac{dy}{2\sqrt{y}} = dx \longrightarrow \sqrt{y} = x - c \longrightarrow y = (x - c)^2, \quad (6.6)$$

and the original parabola is not uniquely reconstructed.

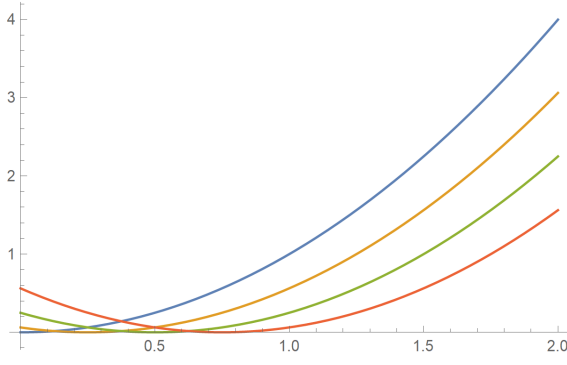


Figure 6.1: $y = (x - c)^2$

A better solution consists in representing the original curve $y = y(x)$ by the family of tangent lines at each point. In order to do so, one represents each member of this family by two numbers (p, Ψ) , the slope of the tangent p and the $-\Psi$, the intersection of the tangent line to the curve at the point x with the y axis. One then replaces the relation $y = y(x)$ by $\psi = \psi(p)$. In terms of equations, this means that

$$p = \frac{\Delta y}{\Delta x} = \frac{y + \Psi}{x - 0} \iff \Psi = px - y. \quad (6.7)$$

More precisely, given $y = y(x)$, define

$$p = \frac{dy}{dx}. \quad (6.8)$$

This relation can be inverted if $\frac{d^2y}{dx^2} \neq 0$ which we assume to be the case below. Inverting gives $x = x(p)$, and in particular also that

$$\left. \frac{dy}{dx} \right|_{x(p)} = p. \quad (6.9)$$

The Legendre transform $\Psi(p)$ is defined as

$$\Psi(p) = \left[px - y(x) \right]_{x=x(p)} = px(p) - y(x(p)). \quad (6.10)$$

In the particular case where $y = x^2$, $p = 2x$, $x = \frac{1}{2}p$, $\Psi(p) = \frac{1}{4}p^2$.

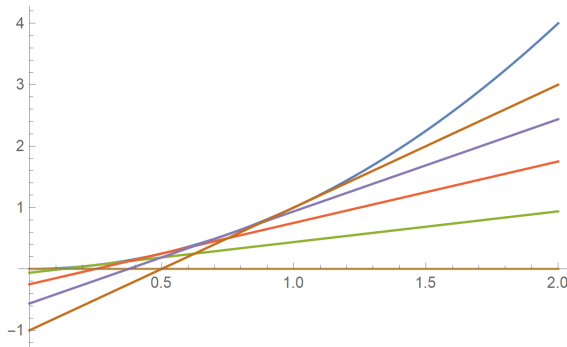


Figure 6.2: Family of tangent lines to $y = x^2$

Theorem 6. *The Legendre transform is invertible. More precisely, the Legendre transform of the Legendre transform is the starting point function.*

Indeed, we have to define $\tilde{x} = \frac{d\Psi}{dp}$ and invert this relation to get $p = p(\tilde{x})$ and then define $\tilde{y}(\tilde{x}) = [\tilde{x}p - \Psi(p)]|_{p(\tilde{x})}$. When using the definition of Ψ ,

$$\tilde{x}(p) = \frac{d}{dp}[px(p) - y(x(p))] = x(p) + \cancel{p \frac{dx'}{dp}} - \cancel{\frac{dy}{dx}} \bigg|_{x(p)} \frac{dx}{dp}. \quad (6.11)$$

It follows that $\tilde{x} = x$ and that

$$x = \frac{d\Psi}{dp} \quad (6.12)$$

is the inverse relation to $p = \frac{dy}{dx}$. Furthermore,

$$\tilde{y}(x) = \left[xp - [px - y(x)]|_{x(p)} \right] \bigg|_{p(x)} = \cancel{xp(x)} - \cancel{p(x)x} + y(x). \quad (6.13)$$

In the particular case where $\Psi(p) = \frac{1}{4}p^2$, $x = \frac{\partial\Psi}{\partial p} = \frac{1}{2}p$ so that $p = 2x$ and $y = [xp - \frac{1}{4}p^2]|_{p=2x} = x^2$.

The Legendre transform may be generalized to a function of several variables where some variables are transformed and others remain passive. It has important applications for instance in the context of thermodynamics.

6.2 Legendre transform in mechanics and Hamilton's equations

Consider a Lagrangian, $L(q^\alpha, \dot{q}^\alpha, t)$, $\alpha = 1, \dots, n$. By definition, the Hamiltonian is the Legendre transform of the Lagrangian with respect to all the generalized velocities $\dot{q}^\alpha$. One thus defines the canonical momenta by

$$p_\alpha = \frac{\partial L}{\partial \dot{q}^\alpha}. \quad (6.14)$$

In order to be able to invert this relation,

$$\dot{q}^\alpha = \dot{q}^\alpha(q^\beta, p_\beta, t), \quad (6.15)$$

one needs

$$\left| \frac{\partial^2 L}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} \right| \neq 0, \quad (6.16)$$

and the Hamiltonian is

$$H(q^\beta, p_\beta, t) = \left[p_\alpha \dot{q}^\alpha - L \right] \bigg|_{\dot{q}^\alpha(q^\beta, p_\beta, t)}. \quad (6.17)$$

By the theorem on the inverse Legendre transform, we thus have that

$$\dot{q}^\alpha = \frac{\partial H}{\partial p_\alpha} \quad (6.18)$$

is the inverse of relation (6.14) and that

$$L(q^\alpha, \dot{q}^\alpha, t) = \left[p_\alpha \dot{q}^\alpha - H(q^\beta, p_\beta, t) \right] \bigg|_{p(q, \dot{q}, t)}. \quad (6.19)$$

This suggests to study the least action principle based on

$$S_H[q, p] = \int_{t_i}^{t_f} dt L_H, \quad L_H = p_\alpha \dot{q}^\alpha - H(q^\beta, p_\beta, t), \quad (6.20)$$

which is first order in time derivatives. For infinitesimal virtual instantaneous variations $\delta q^\alpha(t), \delta p_\alpha(t)$ such that $\delta q^\alpha(t_f) = 0 = \delta q^\alpha(t_i)$, one finds

$$\delta S_H = \int_{t_i}^{t_f} dt \left[\delta p_\alpha (\dot{q}^\alpha - \frac{\partial H}{\partial p_\alpha}) + \frac{d}{dt} (p_\alpha \delta q^\alpha) + (-\dot{p}_\alpha - \frac{\partial H}{\partial q^\alpha}) \delta q^\alpha \right], \quad (6.21)$$

so that the equations of motion that one obtains when requiring that the first order action principle be an extremum for all such variations are Hamilton's equations

$$\boxed{\dot{q}^\alpha = \frac{\partial H}{\partial p_\alpha}, \quad \dot{p}_\alpha = -\frac{\partial H}{\partial q^\alpha}}. \quad (6.22)$$

These $2n$ first order differential equations appear here as the Euler-Lagrange equations of L_H with respect to p_α, q^α .

Theorem 7. *Hamilton's equations are equivalent to the Euler-Lagrange equations of motions associated to the Lagrangian L , $\frac{\delta L}{\delta q^\alpha} = 0$.*

The proof proceeds as follows. From the definition of the Legendre transform, a general variation of the Hamiltonian dH is given by

$$\begin{aligned} \frac{\partial H}{\partial q^\alpha} dq^\alpha + \frac{\partial H}{\partial p_\alpha} dp_\alpha + \frac{\partial H}{\partial t} dt &= d \left\{ \left[p_\alpha \dot{q}^\alpha - L(q, \dot{q}, t) \right] \Big|_{\dot{q}(q,p,t)} \right\} \\ &= dp_\alpha \dot{q}^\alpha \Big|_{\dot{q}(q,p,t)} + p_\alpha \left(\cancel{\frac{\partial \dot{q}^\alpha}{\partial q^\beta} dq^\beta} + \cancel{\frac{\partial \dot{q}^\alpha}{\partial p_\beta} dp_\beta} + \cancel{\frac{\partial \dot{q}^\alpha}{\partial t} dt} \right) - \frac{\partial L}{\partial q^\alpha} \Big|_{\dot{q}(q,p,t)} dq^\alpha \\ &\quad - \frac{\partial L}{\partial \dot{q}^\alpha} \Big|_{\dot{q}(q,p,t)} \left(\cancel{\frac{\partial \dot{q}^\alpha}{\partial q^\beta} dq^\beta} + \cancel{\frac{\partial \dot{q}^\alpha}{\partial p_\beta} dp_\beta} + \cancel{\frac{\partial \dot{q}^\alpha}{\partial t} dt} \right) - \frac{\partial L}{\partial t} \Big|_{\dot{q}(q,p,t)} dt. \end{aligned} \quad (6.23)$$

It follows that

$$\frac{\partial H}{\partial p_\alpha} = \dot{q}^\alpha \Big|_{\dot{q}(q,p,t)}, \quad \frac{\partial H}{\partial q^\alpha} = -\frac{\partial L}{\partial q^\alpha} \Big|_{\dot{q}(q,p,t)}, \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \Big|_{\dot{q}(q,p,t)}. \quad (6.24)$$

We already know that the first equation, which is equivalent to the first set of Hamilton's equations in (6.22) is the inverse to the definition of the canonical momenta in (6.14). When injecting this and the second relation in the second set of Hamilton's equations, and expressing the equations in terms of $q, \dot{q}, t$, one finds

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \right) = \frac{\partial L}{\partial q^\alpha} \iff \frac{\delta L}{\delta q^\alpha} = 0. \quad (6.25)$$

□

6.3 Poisson brackets

By definition, phase space is the $2n$ dimensional space with local coordinates given by the so-called canonical variables (q^α, p_α) , $\alpha = 1, \dots, n$. Consider then phase space functions with an additional explicit dependence on time, $F(q^\alpha, p_\alpha, t)$. For natural trajectories, that is to say, when Hamilton's equations (6.22) are satisfied, the time evolution of such a function is given by

$$\begin{aligned} \frac{dF}{dt} &= \frac{\partial F}{\partial q^\alpha} \dot{q}^\alpha + \frac{\partial F}{\partial p_\alpha} \dot{p}_\alpha + \frac{\partial F}{\partial t} = \frac{\partial F}{\partial q^\alpha} \frac{\partial H}{\partial p_\alpha} - \frac{\partial F}{\partial p_\alpha} \frac{\partial H}{\partial q^\alpha} + \frac{\partial F}{\partial t} \\ &= \{F, H\} + \frac{\partial F}{\partial t}. \end{aligned} \quad (6.26)$$

where the Poisson bracket of two phase space functions F_1, F_2 (that may depend explicitly on time) is defined by

$$\{F_1, F_2\} = \frac{\partial F_1}{\partial q^\alpha} \frac{\partial F_2}{\partial p_\alpha} - \frac{\partial F_1}{\partial p_\alpha} \frac{\partial F_2}{\partial q^\alpha}. \quad (6.27)$$

This definition implies in particular the “canonical Poisson brackets”

$$\{q^\alpha, q^\beta\} = 0 = \{p_\alpha, p_\beta\}, \quad \{q^\alpha, p_\beta\} = \delta_\beta^\alpha = -\{p_\beta, q^\alpha\}. \quad (6.28)$$

Furthermore, in these terms, Hamilton’s equations become

$$\dot{q}^\alpha = \{q^\alpha, H\}, \quad \dot{p}_\alpha = \{p_\alpha, H\}. \quad (6.29)$$

The Poisson bracket satisfies

- skew-symmetry

$$\{F, G\} = -\{G, F\}, \quad (6.30)$$

- right (and as a consequence of skew-symmetry, also left) linearity

$$\{F, c_1 G_1 + c_2 G_2\} = c_1 \{F, G_1\} + c_2 \{F, G_2\}, \quad c_1, c_2 \in \mathbb{R}. \quad (6.31)$$

- right (and as a consequence of skew-symmetry, also left) Leibniz rule

$$\{F, GK\} = \{F, G\}K + G\{F, K\}. \quad (6.32)$$

- Jacobi identity

$$\begin{aligned} \{F, \{G, K\}\} + \{G, \{K, F\}\} + \{K, \{F, G\}\} &= 0 \\ \iff \{F, \{G, K\}\} &= \{G, \{F, K\}\} + \{\{F, G\}, K\}, \end{aligned} \quad (6.33)$$

where the equivalence of the two expressions relies on the skew-symmetry of the Poisson bracket.

The proof of all of these properties, except for the Jacobi identity, follows directly from the definition. For the Jacobi identity, the proof may be streamlined by introducing a collective notation for the phase space coordinates,

$$z^A = (q^\alpha, p_\alpha), \quad A = 1, \dots, 2n, \quad (6.34)$$

together with the matrix,

$$\omega = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix}, \quad \omega^{AB} = \begin{pmatrix} 0 & \delta^{A, B-n} \\ -\delta^{A-n, B} & 0 \end{pmatrix}, \quad (6.35)$$

which is skew-symmetric, $\omega^{AB} = -\omega^{BA}$, and whose inverse is the transposed matrix with components denoted by ω_{AB} , $\omega^{AB}\omega_{BC} = \delta_C^A$, which is also skew-symmetric $\omega_{AB} = -\omega_{BA}$. In these terms, Poisson brackets can be written compactly as

$$\{F, G\} = \frac{\partial F}{\partial z^A} \omega^{AB} \frac{\partial G}{\partial z^B}, \quad \{z^A, z^B\} = \omega^{AB}, \quad (6.36)$$

while Hamilton’s equation become

$$\dot{z}^A = \{z^A, H\} = \omega^{AB} \frac{\partial H}{\partial z^B}. \quad (6.37)$$

To each phase space function F , one may associated the first order differential operator

$$v_F = v_F^B \frac{\partial}{\partial z^B}, \quad v_F^B = \frac{\partial F}{\partial z^A} \omega^{AB}, \quad (6.38)$$

which may also be written as

$$v_F = \{F, \cdot\}. \quad (6.39)$$

In particular, one may write Poisson brackets as

$$\{F, G\} = v_F G = -v_G F. \quad (6.40)$$

For the proof of the Jacobi identity, one starts by re-writing the second of (6.33) as

$$v_F(v_G K) - v_G(v_F K) = v_{\{F, G\}} K. \quad (6.41)$$

Developing the left-hand side, one gets

$$\begin{aligned} & v_F^A \frac{\partial}{\partial z^A} (v_G^B \frac{\partial K}{\partial z^B}) - v_G^B \frac{\partial}{\partial z^B} (v_F^A \frac{\partial K}{\partial z^A}) \\ &= v_F^A \frac{\partial v_G^B}{\partial z^A} \frac{\partial K}{\partial z^B} + \cancel{v_F^A v_G^B \frac{\partial^2 K}{\partial z^A \partial z^B}} - v_G^B \frac{\partial v_F^A}{\partial z^B} \frac{\partial K}{\partial z^A} - \cancel{v_G^B v_F^A \frac{\partial^2 K}{\partial z^B \partial z^A}} \\ &= [v_F^B \frac{\partial v_G^A}{\partial z^B} - v_G^B \frac{\partial v_F^A}{\partial z^B}] \frac{\partial K}{\partial z^A} \\ &= [\frac{\partial F}{\partial z^B} \omega^{BC} \frac{\partial}{\partial z^C} (\frac{\partial G}{\partial z^D}) \omega^{DA} - \frac{\partial G}{\partial z^B} \omega^{BC} \frac{\partial}{\partial z^C} (\frac{\partial F}{\partial z^D}) \omega^{DA}] \frac{\partial K}{\partial z^A}. \end{aligned} \quad (6.42)$$

For the right hand side, one finds

$$\begin{aligned} v_{\{F, G\}}^A \frac{\partial K}{\partial z^A} &= \frac{\partial}{\partial z^B} [\frac{\partial F}{\partial z^C} \omega^{CD} \frac{\partial G}{\partial z^D}] \omega^{BA} \frac{\partial K}{\partial z^A} \\ &= [\frac{\partial^2 F}{\partial z^B \partial z^C} \omega^{CD} \frac{\partial G}{\partial z^D} + \frac{\partial F}{\partial z^C} \omega^{CD} \frac{\partial^2 G}{\partial z^B \partial z^D}] \omega^{BA} \frac{\partial K}{\partial z^A} \\ &= [\frac{\partial^2 F}{\partial z^D \partial z^B} \omega^{BC} \frac{\partial G}{\partial z^C} \omega^{DA} + \frac{\partial F}{\partial z^C} \omega^{CB} \frac{\partial^2 G}{\partial z^D \partial z^B} \omega^{DA}] \frac{\partial K}{\partial z^A}, \end{aligned} \quad (6.43)$$

which agrees with the previous equation when using skew-symmetry of ω^{BC} in the first term of the right hand side.

Chapter 7

Covariance of the Lagrangian formalism

7.1 Change of dependent variables

Consider a non-singular change of variables of the generalized Lagrange coordinates,

$$q^\alpha = q^\alpha(q'^\beta, t), \quad \det\left(\frac{\partial q^\alpha}{\partial q'^\beta}\right) \neq 0, \quad (7.1)$$

with inverse transformation denoted by $q'^\beta = q'^\beta(q^\alpha, t)$. The Lagrangian in the new coordinates is

$$L'(q', \dot{q}', t) = L(q(q', t), \frac{d}{dt}q(q', t), t) := L| \quad (7.2)$$

where the bar means that one has done the change of variables inside the function, with $\frac{d}{dt}q^\alpha(q'^\beta, t) = \frac{\partial q^\alpha}{\partial q'^\gamma} \dot{q}'^\gamma + \frac{\partial q^\alpha}{\partial t}$.

By direct computation one finds

$$\begin{aligned} \frac{\delta L'}{\delta q'^\beta} &= \frac{\partial L'}{\partial q'^\beta} - \frac{d}{dt} \frac{\partial L'}{\partial \dot{q}'^\beta} \\ &= \frac{\partial L}{\partial q^\alpha} \Big| \frac{\partial q^\alpha}{\partial q'^\beta} + \frac{\partial L}{\partial \dot{q}^\alpha} \Big| \left(\frac{\partial^2 q^\alpha}{\partial q'^\beta \partial q'^\gamma} \dot{q}'^\gamma + \frac{\partial^2 q^\alpha}{\partial q'^\beta \partial t} \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \Big| \frac{\partial q^\alpha}{\partial q'^\beta} \right) \\ &= \frac{\partial q^\alpha}{\partial q'^\beta} \left[\frac{\partial L}{\partial q^\alpha} \Big| - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \Big| \right) \right]. \end{aligned} \quad (7.3)$$

Indeed, when $\frac{d}{dt}$ in the last term of the first line hits $\frac{\partial q^\alpha}{\partial q'^\beta}$ the previous term in the first line is cancelled. As a consequence, out of the second and third term of the first line, one only remains with $\frac{d}{dt}$ that hits $\frac{\partial L}{\partial \dot{q}^\alpha} \Big|$.

Furthermore,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \Big| \right) = \frac{\partial^2 L}{\partial q^\beta \partial \dot{q}^\alpha} \Big| \frac{d}{dt} q^\beta(q', t) + \frac{\partial^2 L}{\partial \dot{q}^\beta \partial \dot{q}^\alpha} \Big| \frac{d^2}{dt^2} q^\beta(q', t) + \frac{\partial^2 L}{\partial t \partial \dot{q}^\alpha} \Big| = \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} \Big| \right), \quad (7.4)$$

where $f(q, \dot{q}, \ddot{q}, t) \Big| = f(q(q', t), \frac{d}{dt}q(q', t), \frac{d^2}{dt^2}q(q', t), t)$. Hence,

$$\boxed{\frac{\delta L'}{\delta q'^\beta} = \frac{\partial q^\alpha}{\partial q'^\beta} \left(\frac{\delta L}{\delta q^\alpha} \Big| \right)}. \quad (7.5)$$

This equation captures the covariance of Euler-Lagrange equations under changes of Lagrange coordinates. The meaning is as follows. Suppose that $\bar{q}^\alpha(t)$ is a natural trajectory, i.e., a solution to the Euler-Lagrange equations of motion,

$$\frac{\delta L}{\delta q^\alpha}(\bar{q}^\alpha(t), \frac{d}{dt}\bar{q}^\alpha(t), t) = 0. \quad (7.6)$$

If one changes variables, $q'^\beta = q'^\beta(q^\alpha, t)$, with inverse given in (7.1), it follows that $\bar{q}'^\beta(t) = q'^\beta(\bar{q}^\alpha(t), t)$ is a solution for the variational pinciple defined by $L'(q', \dot{q}', t)$. Indeed, (7.5) is equivalent to

$$\frac{\partial q'^\beta}{\partial q^\alpha} \frac{\delta L'}{\delta q'^\beta} \left(q'(q, t), \frac{d}{dt} q'(q, t), \frac{d^2}{dt^2} q'(q, t), t \right) = \frac{\delta L}{\delta q^\alpha} (q, \dot{q}, \ddot{q}, t). \quad (7.7)$$

When evaluating at $\bar{q}^\alpha(t)$, the RHS vanishes. The result that $\frac{\delta L'}{\delta q'^\beta}$ evaluated at $\bar{q}'^\beta(t)$ vanishes then follows from invertibility of $\frac{\partial q'^\beta}{\partial q^\alpha}$.

7.2 Change of the independent variable

The action is a functional, it produces a number if one evaluates it on a trajectory,

$$S[q] = \int_{t_i}^{t_f} dt L \left(q(t), \frac{d}{dt} q(t), t \right) \in \mathbb{R}. \quad (7.8)$$

Under an arbitrary reparametrisation of time, $t = t(t')$, with inverse denoted by $t' = t'(t)$, one finds the same number provided that

$$S[q] = \int_{t'=t'(t_i)}^{t'=t'(t_f)} dt' \frac{dt}{dt'} L \left(q(t(t')), \frac{dt'}{dt} \frac{d}{dt'} q(t(t')), t(t') \right). \quad (7.9)$$

This suggests that the Euler-Lagrange equations are covariant under $t = t(t')$ and $q'^\alpha(t') = q^\alpha(t(t'))$ provided that

$$\boxed{L' \left(q', \frac{dq'}{dt'}, t' \right) = \frac{dt}{dt'} L \left(q(t(t')), \frac{dt'}{dt} \frac{d}{dt'} q(t(t')), t(t') \right)}. \quad (7.10)$$

with

$$\frac{dq'^\alpha}{dt'} = \frac{dq^\alpha}{dt} \Big|_{t=t(t')} \frac{dt}{dt'} = \dot{q}^\alpha \Big|_{t=t(t')} \frac{dt}{dt'}. \quad (7.11)$$

Consider first a simple example. If $L = \frac{1}{2} m \dot{q}^2$, $\frac{\delta L}{\delta q} = 0 \iff \ddot{q} = 0 \iff q(t) = vt + c$ with $v, c \in \mathbb{R}$. Let $t, t' > 0$ and take $t = \sqrt{t'} \iff t' = t^2$.

It follows that $\frac{dt'}{dt} = 2t = 2\sqrt{t'}$, and the solution becomes $q'(t') = v\sqrt{t'} + c$. The new Lagrangian is

$$L' = \frac{dt}{dt'} \frac{1}{2} m \left(\frac{dt'}{dt} \frac{dq'}{dt'} \right)^2 = m \sqrt{t'} \left(\frac{dq'}{dt'} \right)^2. \quad (7.12)$$

The new equations of motion are

$$\frac{\delta L'}{\delta q'} = \frac{\partial L'}{\partial q'} - \frac{d}{dt'} \frac{\partial L'}{\partial \frac{dq'}{dt'}} = -\frac{d}{dt'} \left(2m \sqrt{t'} \frac{dq'}{dt'} \right) = -m \frac{1}{\sqrt{t'}} \frac{dq'}{dt'} - 2m \sqrt{t'} \frac{d^2 q'}{dt'^2} = 0, \quad (7.13)$$

which is equivalent to

$$\frac{d^2 q'}{dt'^2} = -\frac{1}{2t'} \frac{dq'}{dt'}. \quad (7.14)$$

This equation holds since

$$\frac{dq'}{dt'} = \frac{1}{2} \frac{v}{\sqrt{t'}}, \quad \frac{d^2 q'}{dt'^2} = -\frac{1}{4} \frac{v}{(\sqrt{t'})^3}. \quad (7.15)$$

The proof in the general case proceeds as follows.

$$\frac{\delta L'}{\delta q'^\alpha} = \frac{\partial L'}{\partial q'^\alpha} - \frac{d}{dt'} \frac{\partial L'}{\partial \frac{dq'^\alpha}{dt'}} = \frac{dt}{dt'} \frac{\partial L}{\partial q^\alpha} - \frac{d}{dt'} \left(\frac{dt}{dt'} \frac{\partial L}{\partial \dot{q}^\alpha} \right) = \frac{dt}{dt'} \left[\frac{\partial L}{\partial q^\alpha} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \right) \right] = \frac{dt}{dt'} \left[\frac{\delta L}{\delta q^\alpha} \right], \quad (7.16)$$

where the last equality holds provided one can show that

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \right) = \left[\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} \right]. \quad (7.17)$$

As in the preceding section, $\frac{\delta L'}{\delta q'^\alpha} = \frac{dt}{dt'} \left(\frac{\delta L}{\delta q^\alpha} \right)$ means that, if $q^\alpha(t)$ is a natural trajectory for $L(q, \dot{q}, t)$, then so is $q'^\alpha(t') = q^\alpha(t(t'))$ for $L'(q', \dot{q}', t')$.

In order to show (7.17), we have

$$\begin{aligned} & \frac{d}{dt} \left[\frac{\partial L}{\partial \dot{q}^\alpha} \left(q'(t'), \frac{dt'}{dt} \frac{dq'}{dt'}, t(t') \right) \right] \\ &= \frac{\partial^2 L}{\partial q^\beta \partial \dot{q}^\alpha} \left| \frac{dt'}{dt} \frac{dq'^\beta}{dt'} + \frac{\partial^2 L}{\partial \dot{q}^\beta \partial \dot{q}^\alpha} \left(\frac{d^2 t'}{dt^2} \frac{dq'^\beta}{dt'} + \left(\frac{dt'}{dt} \right)^2 \frac{d^2 q'^\beta}{dt'^2} \right) + \frac{\partial^2 L}{\partial t \partial \dot{q}^\alpha} \left| \frac{dt'}{dt} \frac{dt'}{dt'} \right|, \end{aligned} \quad (7.18)$$

while

$$\left[\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} \right] = \frac{\partial^2 L}{\partial q^\beta \partial \dot{q}^\alpha} |\dot{q}^\beta| + \frac{\partial^2 L}{\partial \dot{q}^\beta \partial \dot{q}^\alpha} |\ddot{q}^\beta| + \frac{\partial^2 L}{\partial t \partial \dot{q}^\alpha}. \quad (7.19)$$

When taking into account (7.11), the first terms on the RHS do agree, since so do the last terms, we only have to show that middle terms do as well, which is the case if

$$\ddot{q}^\beta = \frac{d^2 t'}{dt^2} \frac{dq'^\beta}{dt'} + \left(\frac{dt'}{dt} \right)^2 \frac{d^2 q'^\beta}{dt'^2}. \quad (7.20)$$

Differentiating (7.11) with respect to t' gives

$$\frac{d^2 q'^\beta}{dt'^2} = \frac{d}{dt'} \left(\frac{dt}{dt'} \dot{q}^\beta \right) = \frac{d^2 t}{dt'^2} \dot{q}^\beta + \left(\frac{dt}{dt'} \right)^2 \ddot{q}^\beta, \quad (7.21)$$

which implies that

$$\ddot{q}^\beta = \left(\frac{dt'}{dt} \right)^2 \frac{d^2 q'^\beta}{dt'^2} - \left(\frac{dt'}{dt} \right)^2 \frac{d^2 t}{dt'^2} \dot{q}^\beta. \quad (7.22)$$

This gives the result when using (7.11) and

$$0 = \frac{d}{dt} \left(\frac{dt'}{dt} \frac{dt}{dt'} \right) = \frac{d^2 t'}{dt^2} \frac{dt}{dt'} + \left(\frac{dt'}{dt} \right)^2 \frac{d^2 t}{dt'^2} \implies \frac{d^2 t'}{dt^2} = - \left(\frac{dt'}{dt} \right)^3 \frac{d^2 t}{dt'^2}. \quad (7.23)$$

7.3 Combined change of variables

When combining the results of the two previous sections, Euler-Lagrange equations are covariant under

$$\begin{cases} q^\alpha = q'^\alpha(q'^\beta, t') \\ t = t(t') \end{cases} \quad (7.24)$$

with inverse

$$\begin{cases} q'^\alpha = q'^\alpha(q^\beta, t) \\ t' = t'(t) \end{cases} \quad (7.25)$$

provided that

$$\boxed{L'\left(q'^\beta(t'), \frac{d}{dt'}q'^\beta(t'), t'\right) = \frac{dt}{dt'}L\left(q^\alpha(q'^\beta(t'), t'), \frac{dt'}{dt}\frac{d}{dt'}q^\alpha(q'^\beta(t'), t'), t(t')\right)}, \quad (7.26)$$

with $\frac{d}{dt'}q^\alpha(q'^\beta(t'), t') = \frac{\partial q^\alpha}{\partial q'^\gamma} \frac{dq'^\gamma}{dt'} + \frac{\partial q^\alpha}{\partial t'}$. Indeed, showing that

$$\frac{\delta L'}{\delta q'^\beta} = \frac{dt}{dt'} \left(\frac{\delta L}{\delta q^\alpha} \right) \bigg| \frac{\partial q^\alpha}{\partial q'^\beta}. \quad (7.27)$$

can now be done in a straightforward way by combining the intermediate steps of the proofs of the previous two sections.

Chapter 8

Canonical transformations

8.1 Generating functionals

Under a general (invertible) transformation of the dependent variables,

$$q'^\alpha = q'^\alpha(q, p, t), \quad p'_\alpha = p'_\alpha(q, p, t) \iff z'^A = z^A(z, t), \quad (8.1)$$

there is no guarantee that the Hamiltonian equations of motion keep, the same form, that is to say, that there exists a suitable Hamiltonian $H'(q', p', t)$ such that the Hamiltonian equations of motion in the old variables (6.22) are equivalent to

$$\dot{q}'^\alpha = \frac{\partial H'}{\partial p'_\alpha}, \quad \dot{p}'_\alpha = -\frac{\partial H'}{\partial q'^\alpha}. \quad (8.2)$$

However, if one can find a function $G(z', t)$ such that

$$[p_\alpha \dot{q}^\alpha - H]|_{z(z', t)} = p'_\alpha \dot{q}'^\alpha - H'(z', t) + \frac{dG}{dt}, \quad (8.3)$$

then Hamilton's equations in the new variables (8.2) are equivalent to Hamilton's equations in the old variables (6.22).

The proof uses the fact that the addition of a total derivative does not change the equations of motion, so that the Euler-Lagrange equations of motion with respect to z'^A of the right hand side are given by (8.2). The result follows from the covariance of the Euler-Lagrange equations of motions, which guarantees that the Euler-Lagrange equations of motion of the left hand side with respect to z'^A are equivalent to the Euler Lagrange equations of L_H with respect to z^A , which coincide with Hamilton's equations (6.22).

More explicitly, in terms of equations, recall that $L_H(z, \dot{z}, t) = p_\alpha \dot{q}^\alpha - H(z, t)$. It follows that

$$L'_H(z', \dot{z}', t) = L_H(z(z', t), \frac{d}{dt}z(z', t), t) = p'_\alpha \dot{q}'^\alpha - H'(z', t) + \frac{d}{dt}G \quad (8.4)$$

where the first equality is the definition of the transformed first order Lagrangian, while the second equality is equivalent to the assumption (8.3).

When taking the Euler-Lagrange derivative of this equation with respect to z'^A , we get on the left hand side

$$\frac{\delta L'_H}{\delta z'^A} = \frac{\partial z^B}{\partial z'^A} \frac{\delta L_H}{\delta z^B} \Big|_{z=z(z', t)}, \quad (8.5)$$

when using covariance of the Euler-Lagrange derivatives (cf. equation (7.5)). Since Euler-Lagrange derivatives annihilate total derivatives, (4.29), the right hand side reduces to

$$\frac{\delta}{\delta z'^A} [p'_\alpha \dot{q}'^\alpha - H'(z', t)] \quad (8.6)$$

When requiring these Euler-Lagrange derivatives to vanish, one finds Hamilton's equations of motion in the new variables, (8.2), while requiring (8.5) to vanish is equivalent to Hamilton's equations in the old variables.

Suppose now that one may use as variables q^α, q'^α, t , or in other words that $p_\alpha = p_\alpha(q', p'(q, p, t), t)$ may be solved for p_α so as to provide $p_\alpha = p_\alpha(q, q', t)$. Let us define

$$\Gamma_1(q, q', t) = G(q', p'(q, p(q, q', t), t), t). \quad (8.7)$$

In these variables, (8.3) becomes

$$\begin{aligned} p_\alpha(q, q', t)\dot{q}^\alpha - H(q, p(q, q', t), t) \\ = p'_\alpha(q, p(q, q', t), t)\dot{q}'^\alpha - H'\left((q', p'(q, p(q, q', t), t))\right) + \dot{q}^\alpha \frac{\partial \Gamma_1}{\partial q^\alpha} + \dot{q}'^\alpha \frac{\partial \Gamma_1}{\partial q'^\alpha} + \frac{\partial \Gamma_1}{\partial t}. \end{aligned} \quad (8.8)$$

By identifying the coefficients of the time derivatives, this is equivalent to

$$\boxed{p_\alpha = \frac{\partial \Gamma_1}{\partial q^\alpha}, \quad p'_\alpha = -\frac{\partial \Gamma_1}{\partial q'^\alpha}, \quad H' = H + \frac{\partial \Gamma_1}{\partial t}}. \quad (8.9)$$

The information on the canonical transformation defined by (8.3) is thus contained in the “generating function” (of the first kind) $\Gamma_1(q, q', t)$, where one assumes that

$$\left| \frac{\partial^2 \Gamma_1}{\partial q^\alpha \partial q'^\beta} \right| \neq 0 \quad (8.10)$$

so that one may solve $p_\alpha = p_\alpha(q, q', t)$ in terms of $q'^\alpha = q'^\alpha(q, p, t)$.

In particular for instance, if $\Gamma_1 = q^\alpha \delta_{\alpha\beta} q'^\beta$, the transformation becomes $p_\alpha = q'_\alpha, p'_\alpha = -q_\alpha$, so that this transformation exchanges the role played by the generalized coordinates and the momenta,

$$(q^\alpha, p_\alpha) \rightarrow (p_\alpha, -q^\alpha). \quad (8.11)$$

In the considerations below, let us focus for notational simplicity on a single degree of freedom and omit the explicit time dependence. Consider then the (opposite of the) Legendre transform of $\Gamma_1(q, q')$ for which $p' = -\frac{\partial \Gamma_1}{\partial q'}$ with respect to q' ,

$$\Gamma_2(q, p') = \left[p' q' + \Gamma_1(q, q') \right] \Big|_{q'=q'(q, p')}. \quad (8.12)$$

This is the generating function of the second kind, for which

$$p = \frac{\partial \Gamma_2}{\partial q}, \quad q' = \frac{\partial \Gamma_2}{\partial p'}. \quad (8.13)$$

In particular, if $\Gamma_2 = qp'$, the transformation is the identity, $p = p', q' = q$.

The Legendre transform of Γ_1 with respect to q , where $p = \frac{\partial \Gamma_1}{\partial q}$ is the generating function of the third kind,

$$\Gamma_3(p, q') = \left[-pq + \Gamma_1(q, q') \right] \Big|_{q=q(p, q')}, \quad (8.14)$$

for which

$$p' = -\frac{\partial \Gamma_3}{\partial q'}, \quad q = -\frac{\partial \Gamma_3}{\partial p}. \quad (8.15)$$

In particular, if $\Gamma_3 = pq'$, the transformation is $p = -p', q' = -q$.

Finally, the generating function of the fourth kind is the Legendre transform of $\Gamma_2(q, p')$, for which $p = \frac{\partial \Gamma_2}{\partial q}$ with respect to q ,

$$\Gamma_4(p, p') = \left[-pq + \Gamma_2(q, p') \right] \Big|_{q=q(p, p')}, \quad (8.16)$$

for which

$$q' = \frac{\partial \Gamma_4}{\partial p'}, \quad q = -\frac{\partial \Gamma_4}{\partial p}. \quad (8.17)$$

In particular, if $\Gamma_4 = pp'$, $q' = p$, $q = -p'$.

Remark: For several canonical pairs, one may have generating functions of different type in the different canonical pairs.

How can one remember these formulas ?

$$q, q' : pdq = p'dq' + d\Gamma_1(q, q'), \quad \boxed{p = \frac{\partial \Gamma_1}{\partial q}, \quad p' = -\frac{\partial \Gamma_1}{\partial q'}}, \quad (8.18)$$

$$\begin{aligned} q, p' : p'dq' &= d(p'q') - dp'q', \quad \Gamma_2 = \Gamma_1 + p'q', \\ pdq &= -dp'q' + d\Gamma_2(q, p'), \quad \boxed{p = \frac{\partial \Gamma_2}{\partial q}, \quad q' = \frac{\partial \Gamma_2}{\partial p'}}, \end{aligned} \quad (8.19)$$

$$\begin{aligned} p, q' : pdq &= d(pq) - dpq, \quad \Gamma_3 = \Gamma_1 - pq, \\ -dpq &= p'dq' + d\Gamma_3(p, q'), \quad \boxed{q = -\frac{\partial \Gamma_3}{\partial p}, \quad p' = -\frac{\partial \Gamma_3}{\partial q'}}, \end{aligned} \quad (8.20)$$

$$\begin{aligned} p, p' : \Gamma_4 &= \Gamma_2 - pq = \Gamma_3 + p'q', \\ -dpq &= -dp'q' + d\Gamma_4(p, p'), \quad \boxed{q = -\frac{\partial \Gamma_4}{\partial p}, \quad q' = \frac{\partial \Gamma_4}{\partial p'}}. \end{aligned} \quad (8.21)$$

Remark: The transformation

$$q' = \mu q, \quad p' = \nu p, \quad \mu, \nu \in \mathbb{R}^*, \quad (8.22)$$

is a canonical transformation with

$$H'(q', p',) = \mu\nu H\left(\frac{q'}{\mu}, \frac{p'}{\nu}\right) \quad (8.23)$$

because

$$p'\dot{q}' - H' = \mu\nu(p\dot{q} - H) \Big|_{q=\frac{q'}{\mu}, p=\frac{p'}{\nu}}, \quad (8.24)$$

and the Euler-Lagrange equations are unchanged when multiplying the Lagrangian by a non-vanishing constant. There exists no generating function however. The same holds for the canonical transformation

$$q' = p, \quad p' = q, \quad H' = -H. \quad (8.25)$$

A canonical transformation that admits a generating function is called a restricted canonical transformation.

8.2 Invariance of canonical Poisson brackets

Theorem 8. *A canonical transformation is a restricted canonical transformation if and only if*

$$\boxed{\{q'^\alpha, q'^\beta\} = 0 = \{p'_\alpha, p'_\beta\}, \quad \{q'^\alpha, p'_\beta\} = \delta_\beta^\alpha, \quad \Longleftrightarrow \quad \{z'^A, z'^B\} = \omega^{AB}}. \quad (8.26)$$

Remarks:

(i) The last condition becomes explicitly

$$\frac{\partial z'^A}{\partial z^C} \omega^{CD} \frac{\partial z'^B}{\partial z^D} = \omega^{AB} \Longleftrightarrow \frac{\partial z'^A}{\partial z^C} \omega^{CD} = \omega^{AB} \frac{\partial z^D}{\partial z'^B}, \quad (8.27)$$

or

$$\begin{pmatrix} \frac{\partial q'^\alpha}{\partial q^\beta} & \frac{\partial q'^\alpha}{\partial p_\beta} \\ \frac{\partial p'_\alpha}{\partial q^\beta} & \frac{\partial p'_\alpha}{\partial p_\beta} \end{pmatrix} \begin{pmatrix} 0 & \delta_\gamma^\beta \\ -\delta_\beta^\gamma & 0 \end{pmatrix} = \begin{pmatrix} 0 & \delta_\beta^\alpha \\ -\delta_\alpha^\beta & 0 \end{pmatrix} \begin{pmatrix} \frac{\partial q^\gamma}{\partial q'^\beta} & \frac{\partial p_\gamma}{\partial q'^\beta} \\ \frac{\partial q^\gamma}{\partial p'_\beta} & \frac{\partial p_\gamma}{\partial p'_\beta} \end{pmatrix}, \quad (8.28)$$

which is equivalent to the following relations between mixed partial derivatives:

$$\begin{aligned} \frac{\partial q'^\alpha}{\partial p_\gamma} &= -\frac{\partial q^\gamma}{\partial p'_\alpha} (a), & \frac{\partial q'^\alpha}{\partial q^\gamma} &= -\frac{\partial p_\gamma}{\partial p'_\alpha} (b), \\ \frac{\partial p'_\alpha}{\partial p_\gamma} &= -\frac{\partial q^\gamma}{\partial q'^\alpha} (c), & \frac{\partial p'_\alpha}{\partial q^\gamma} &= -\frac{\partial p_\gamma}{\partial q'^\alpha} (d). \end{aligned} \quad (8.29)$$

(ii) As a consequence of the theorem, for a restricted canonical transformation, Poisson brackets can be computed in the old or in the new variables. Indeed, if $F'(z') = F(z(z'))$, $G'(z') = G(z(z'))$,

$$\begin{aligned} \{F', G'\}_{z'} &= \frac{\partial F'}{\partial z'^A} \omega^{AB} \frac{\partial G'}{\partial z'^B} = \frac{\partial F}{\partial z^C} \Big|_{z=z(z')} \frac{\partial z^C}{\partial z'^A} \omega^{AB} \frac{\partial z^D}{\partial z'^B} \frac{\partial G}{\partial z^D} \Big|_{z=z(z')} \\ &= \left(\frac{\partial F}{\partial z^C} \omega^{CD} \frac{\partial G}{\partial z^D} \right) \Big|_{z=z(z')} = \left(\{F, G\} \right) \Big|_{z=z(z')}, \end{aligned} \quad (8.30)$$

where we have used that, by multiplying by the Jacobian matrices of the inverse transformations, the first of (8.27) is equivalent to

$$\frac{\partial z^C}{\partial z'^A} \omega^{AB} \frac{\partial z^D}{\partial z'^B} = \omega^{CD}. \quad (8.31)$$

(iii) Equation (8.31) means that the inverse of a canonical transformation is a canonical transformation. Consider then also a second canonical transformations, $z'' = z''(z')$. The Jacobian matrices of the composed transformation $z'' = z''(z) = z''(z'(z))$ are given by

$$\frac{\partial z''^A}{\partial z^B} = \frac{\partial z''^A}{\partial z'^C} \Big|_{z'=z'(z)} \frac{\partial z'^C}{\partial z^B}, \quad (8.32)$$

which allows to prove directly that the composition of two canonical transformations is a canonical transformation. This implies that canonical transformations form a (sub-) group (of the group of invertible transformations).

The detailed proof of the theorem is omitted here and can be found in the handwritten “french” notes.

In the following, we only consider restricted canonical transformations but omit the adjective “restricted”.

8.3 Infinitesimal canonical transformations

In order to study an infinitesimal canonical transformation close to the identity, we start from a generating function of the second kind of the form

$$\Gamma_2 = p'_\beta q^\beta + \epsilon G(q, p', t), \quad (8.33)$$

because we have already seen that for $\epsilon = 0$, such a canonical transformation reduces to the identity. For $\epsilon \neq 0$, the associated canonical transformation is defined by

$$p_\alpha = \frac{\partial \Gamma_2}{\partial q^\alpha} = p'_\alpha + \epsilon \frac{\partial G}{\partial q^\alpha}, \quad q'^\alpha = \frac{\partial \Gamma_2}{\partial p'_\alpha} = q^\alpha + \epsilon \frac{\partial G}{\partial p'_\alpha}. \quad (8.34)$$

In the second terms of these relations one may replace p'_β by p_β in G , since one is already at first order on ϵ , so that $G = G(q, p, t)$. One may then write,

$$\delta q^\alpha = q'^\alpha - q^\alpha = \epsilon \frac{\partial G}{\partial p_\alpha} = \epsilon \{q^\alpha, G\}, \quad \delta p_\alpha = p'_\alpha - p_\alpha = -\epsilon \frac{\partial G}{\partial q^\alpha} = \epsilon \{p_\alpha, G\}, \quad (8.35)$$

and thus also

$$\boxed{\delta z^A = \epsilon \{z^A, G\}}. \quad (8.36)$$

The phase space function $G(q, p, t)$ is called the generator of the infinitesimal canonical transformation.

8.4 One-parameter group of canonical transformations

A family of canonical transformation depending on a parameter λ ,

$$z_\lambda^A = Z^A(z, \lambda), \quad (8.37)$$

such that

$$z_{\lambda_1 + \lambda_2}^A = Z^A(z, \lambda_1 + \lambda_2) = Z^A(Z(z, \lambda_1), \lambda_2), \quad z_0^A = Z^A(z, 0) = z^A, \quad (8.38)$$

is called a one-parameter group of canonical transformations.

The transformation $z_{\lambda+\epsilon}^A$ with ϵ infinitesimal is an infinitesimal canonical transformation that reduces to z_λ^A for $\epsilon = 0$. It follows from (8.36) that there exists a generator $G(z_\lambda)$ such that

$$z_{\lambda+\epsilon}^A - z_\lambda^A = \epsilon \{z_\lambda^A, G(z_\lambda)\}_{z_\lambda}. \quad (8.39)$$

This can also be written as

$$\boxed{\frac{d}{d\lambda} z_\lambda^A = -\{G(z_\lambda), z_\lambda^A\}_{z_\lambda}}, \quad (8.40)$$

or, when using the property that Poisson brackets can be evaluated before or after a canonical transformation,

$$\frac{d}{d\lambda} Z^A(z, \lambda) = -\{G(Z(z, \lambda)), Z^A(z, \lambda)\}_z. \quad (8.41)$$

Conversely, this equation determines uniquely a one-parameter group of canonical transformations close to the identity.

For all functions $F(z_\lambda)$ that do not depend explicitly on λ ,

$$\frac{dF}{d\lambda} = \frac{\partial F}{\partial z_\lambda^A} \frac{dz_\lambda^A}{d\lambda} = -\{G(z_\lambda), z_\lambda^A\} \frac{\partial F}{\partial z_\lambda^A} = -\{G(z_\lambda), F(z_\lambda)\}_{z_\lambda}. \quad (8.42)$$

This implies in particular that $\frac{dG(z_\lambda)}{d\lambda} = 0$ if $\frac{\partial G}{\partial \lambda} = 0$. Assuming that $G(z_\lambda)$ has no explicit λ dependence thus implies that

$$\frac{d^2 z_\lambda^A}{(d\lambda)^2} = \{G(z_\lambda), \{G(z_\lambda), z_\lambda^A\}_{z_\lambda}\}_{z_\lambda}, \quad (8.43)$$

and, by iteration, that

$$\frac{d^n z_\lambda^A}{(d\lambda)^n} = (-1)^n \{G(z_\lambda), \{G(z_\lambda), \dots, \{G(z_\lambda), z_\lambda^A\}_{z_\lambda}\}_{z_\lambda}\}_{z_\lambda}, \quad (8.44)$$

with n brackets. When using that $z_0^A = z^A$ and by putting λ to zero, this gives

$$\left. \frac{d^n z_\lambda^A}{(d\lambda)^n} \right|_{\lambda=0} = (-1)^n \{G(z), \{G(z), \dots, \{G(z), z^A\}\}\}. \quad (8.45)$$

One then may write the function in a neighborhood of the identity in terms of its Taylor series as

$$z_\lambda^A = z^A + \sum_{n=1} \frac{(-1)^n \lambda^n}{n!} \{G(z), \{G(z), \dots, \{G(z), z^A\}\}\}. \quad (8.46)$$

This Taylor series may also be written as

$$z_\lambda^A = e^{-\lambda \{G, \cdot\}} z^A = e^{-\lambda v_G} z^A. \quad (8.47)$$

Remarks:

(i) When the generator of a one-parameter group of canonical transformations depends explicitly on the parameter λ , it remains true that the one-parameter group is uniquely determined by the generator in a neighborhood of the identity. The explicit formula is however more involved than the exponential series above.

(ii) Given initial conditions $z_0^A = z^A$, the solution to the equations of motion $z_t^A = z^A(z_0, t)$ is a one-parameter group of canonical transformations whose generator is the Hamiltonian. This follows from the fact that, when $\lambda = t$ and $G = H$, the equations (8.40) coincide with Hamilton's equations of motion. Furthermore, if H does not explicitly depend on time, $\frac{\partial H}{\partial t} = 0$, the solution of the equations of motion may be written as

$$z_t = e^{-t \{H, \cdot\}} z, \quad (8.48)$$

in a neighborhood of the identity.

(iii) If $K(z)$ is a first integral that does not depend explicitly on time, $\{K, H\} = 0$, it follows that $K(z_t) = K(z_0)$. Indeed, $\frac{dK(z_t)}{dt} = \{K(z_t), H(z_t)\}_{z_t} = 0$.

(iv) Conversely, when considering the one-parameter group of canonical transformations z_λ^A generated by the first integral $K(z)$ which does not depend explicitly on time, for a Hamiltonian that does not explicitly on λ , it follows that $H(z_\lambda, t) = H(z_0, t)$ since $\frac{dH(z_\lambda, t)}{d\lambda} = -\{K(z_\lambda), H(z_\lambda, t)\}_{z_\lambda} = 0$ so that $H(z_\lambda, t) = H(z_0, t)$.

The general case when the generator depends explicitly on the parameter can be found in the handwritten "french" notes.

8.5 Liouville theorem

For a system of first order differential equations,

$$\dot{x}^a = v^a(x^b, t), \quad a, b = 1, \dots, n, \quad (8.49)$$

there is existence and uniqueness for the solution of the initial value problem,

$$x_t^a = X^a(t; x_0^b, t_0), \quad X^a(t_0; x_0^b, t_0) = x_0^a, \quad (8.50)$$

under suitable conditions which we assume to be satisfied.

One may also interpret this equation as (time dependent) transformations that map the points with coordinates x_0^a to the points with coordinates x_t^a . This group of transformation is also called a “flow”, in analogy with fluid mechanics.

Consider then a function $\rho(x^a, t)$ and its integral over a domain $\mathcal{D}(t)$ of $\mathbb{R}^n$,

$$I(t) = \int_{\mathcal{D}(t)} \rho(x, t) dx^1 \dots dx^n. \quad (8.51)$$

When using the flow, to each point of the domain $\mathcal{D}(t)$ corresponds a single point at a time t' . This defines the domain $\mathcal{D}(t')$.

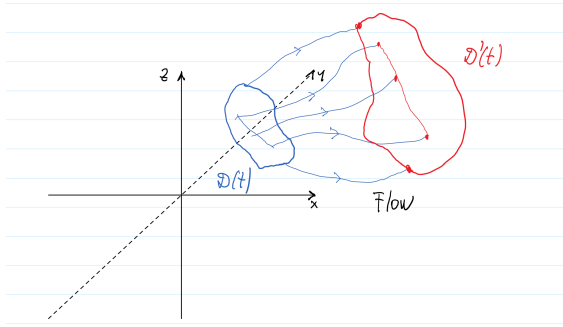


Figure 8.1: Flow of the points of a domain

One wants to understand how the integral changes in time,

$$I(t') - I(t) = \int_{\mathcal{D}(t')} \rho(x, t') dx^1 \dots dx^n - \int_{\mathcal{D}(t)} \rho(x, t) dx^1 \dots dx^n. \quad (8.52)$$

In order to do so, one uses that

$$\int_{\mathcal{D}(t')} \rho(x, t') dx^1 \dots dx^n = \int_{\mathcal{D}(t)} \rho(X(t'; x, t), t') \left| \frac{\partial X^a}{\partial x^b} \right| dx^1 \dots dx^n, \quad (8.53)$$

to bring back the integral over the new domain to an integral over the old domain.

To see that this is indeed the case, take an example in 1 dimension. Remember that for a definite integral,

$$I = \int_{x=a}^{x=b} f(x) dx, \quad (8.54)$$

if one makes a change of variables, $x = g(y)$, the integral becomes

$$I = \int_{y=g^{-1}(a)}^{y=g^{-1}(b)} f(g(y)) \frac{dg}{dy} dy. \quad (8.55)$$

Take the domain at time $\mathcal{D}(t)$ to be the set of points that belong to the interval $[a, b]$, $x_t \in [a, b]$ and

$$I(t) = \int_{x=a}^{x=b} \rho(x, t) dx. \quad (8.56)$$

Under the map $x_{t'} = X(t'; x, t)$, the end-point of the new domain $\mathcal{D}(t')$ are $X(t'; a, t)$ and $X(t'; b, t)$, the images of $x_t = a$ and $x_t = b$. It follows that

$$I(t') = \int_{x=X(t';a,t)}^{x=X(t';b,t)} \rho(x, t') dx. \quad (8.57)$$

When doing the change of variables $x = X(t'; y, t)$ it follows that

$$I(t') = \int_{y=X^{-1}(t';X(t';a,t),t)=a}^{y=X^{-1}(t';X(t';b,t),t)=b} \rho(X(t'; y, t), t') \frac{\partial X}{\partial y} dy = \int_{x=a}^{x=b} \rho(X(t'; x, t), t') \frac{\partial X}{\partial x} dx, \quad (8.58)$$

and thus, by generalizing to multiple variables and integrals,

$$\boxed{I(t') - I(t) = \int_{\mathcal{D}(t)} \left[\rho(X^a(t'; x, t), t') \left| \frac{\partial X^a}{\partial x^b} \right| - \rho(x, t) \right] dx^1 \dots dx^n.} \quad (8.59)$$

In particular, it follows that $I(t') = I(t)$ if ρ is constant and the Jacobian unity, $|\partial X / \partial x| = 1$. If $\rho = 1$, $I(t)$ represents the volume of $\mathcal{D}(t)$ which remains constant for a flow with unit Jacobian.

Consider transformations on phase space, so that the coordinates x^a are now the coordinates z^A . For a canonical transformation, we have

$$\frac{\partial z'^A}{\partial z^C} \omega_{CD} \frac{\partial z'^B}{\partial z^D} = \omega^{AB} \iff J\omega J^T = \omega, \quad J^A{}_B = \frac{\partial z'^A}{\partial z^B}, \quad (8.60)$$

it follows that $\det J = \pm 1$ since $\det \omega = 1 \neq 0$. For all canonical transformations connected to the identity, one thus has by continuity that $\det J = 1$. (One may show further that this is the case for all canonical transformations.) Since the Hamiltonian evolution is a one-parameter group of canonical transformations, with unit Jacobian, we have:

Theorem 9 (Liouville). *The volume of any domain of phase space is preserved under Hamiltonian time evolution.*

Let us now analyse the consequence of (8.59) for an infinitesimal flow, $t' = t + \epsilon$. In this case,

$$X^a(t + \epsilon; x^b, t) = x^a + \epsilon v^a(x^b, t) + \mathcal{O}(\epsilon^2), \quad (8.61)$$

while

$$\left| \frac{\partial X^a}{\partial x^b} \right| = \left| \delta_b^a + \epsilon \frac{\partial v^a}{\partial x^b} + \mathcal{O}(\epsilon^2) \right| = 1 + \epsilon \frac{\partial v^a}{\partial x^a} + \mathcal{O}(\epsilon^2). \quad (8.62)$$

As a consequence,

$$\frac{dI(t)}{dt} = \lim_{\epsilon \rightarrow 0} \frac{I(t + \epsilon) - I(t)}{\epsilon} = \int_{\mathcal{D}(t)} \left[\frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial x^a} v^a + \rho \frac{\partial v^a}{\partial x^a} \right] dx^1 \dots dx^n. \quad (8.63)$$

It follows that

$$\frac{dI(t)}{dt} = 0, \quad \forall \mathcal{D}(t) \iff \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x^a} (\rho v^a) = 0. \quad (8.64)$$

The equation for ρ on the right hand side is called a “continuity equation”.

For a Hamiltonian system, we have $\rho = \rho(z, t)$, while $v^a \rightarrow \omega^{AB} \frac{\partial H}{\partial z^B}$, and the continuity equation becomes

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial z^A} (\rho \omega^{AB} \frac{\partial H}{\partial z^B}) = 0 \iff \boxed{\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0}. \quad (8.65)$$

This is the Liouville equation in statistical mechanics, for which $\rho = 1$ is a particular solution.

Chapter 9

Noether's theorem and Galilean relativity

9.1 Noether's first theorem

A *first integral* or *conserved quantity* is a function $K(q, \dot{q}, t)$ that is constant on all natural trajectories,

$$\frac{dK}{dt}|_{\bar{q}} = 0, \quad \text{with } \frac{\delta L}{\delta q^\alpha}|_{\bar{q}} = 0. \quad (9.1)$$

An *infinitesimal symmetry* is an infinitesimal virtual variation that leaves the Lagrangian invariant up to a total derivative,

$$\delta_Q q^\alpha = Q^\alpha(q, \dot{q}, t) \text{ such that } \delta_Q L = \frac{dF(q, \dot{q}, t)}{dt}. \quad (9.2)$$

Noether's first theorem states that:

Theorem 10 (Noether's 1st). *Every infinitesimal symmetry gives rise to a conserved quantity.*

Indeed,

$$\delta_Q L = \delta_Q q^\alpha \frac{\partial L}{\partial q^\alpha} + \delta_Q \dot{q}^\alpha \frac{\partial L}{\partial \dot{q}^\alpha} = Q^\alpha \frac{\partial L}{\partial q^\alpha} + \frac{dQ^\alpha}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} = \frac{dF}{dt}, \quad (9.3)$$

where the second equality follows because infinitesimal virtual variations commute with the total derivative, while the last equality follows by the assumption that the variation is a symmetry. Using

$$\frac{dQ^\alpha}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} = \frac{d}{dt} \left(Q^\alpha \frac{\partial L}{\partial \dot{q}^\alpha} \right) - Q^\alpha \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\alpha}, \quad (9.4)$$

the latter combines with the first term to produce the Euler-Lagrange derivatives, while the former combines with the total derivative on the RHS,

$$Q^\alpha \frac{\delta L}{\delta q^\alpha} = \frac{d}{dt} (F - \frac{\partial L}{\partial \dot{q}^\alpha} Q^\alpha). \quad (9.5)$$

Since the LHS vanishes when evaluated at any natural trajectory, it follows that associated the conserved quantity is given by

$$K = p_\alpha Q^\alpha - F, \quad p_\alpha = \frac{\partial L}{\partial \dot{q}^\alpha}. \quad (9.6)$$

The overall minus sign is conventional and amounts to writing Noether's first theorem as

$$Q^\alpha \frac{\delta L}{\delta q^\alpha} + \frac{d}{dt} K = 0. \quad (9.7)$$

Remark: Neither the theorem nor the expression depends on the choice of representative for the action: if $L' = L + \frac{df(q,t)}{dt}$, it follows that δ_Q is also an infinitesimal symmetry of L' ,

$$\delta_Q L' = \delta_Q L + \frac{d}{dt}(\delta_Q f) = \frac{d}{dt}(F + \delta_Q f), \quad (9.8)$$

with associated conserved quantity

$$K' = F + \cancel{\frac{\partial f}{\partial q^\alpha} Q^\alpha} - \left(\frac{\partial L}{\partial \dot{q}^\alpha} + \cancel{\frac{\partial f}{\partial \dot{q}^\alpha}} \right) Q^\alpha = K. \quad (9.9)$$

9.2 Applications

9.2.1 Explicit time independence and conservation of energy

Consider a Lagrangian that does not depend explicitly on time, $L = L(q, \dot{q})$, $\frac{\partial L}{\partial t} = 0$. It follows that

$$\frac{dL}{dt} = \frac{\partial L}{\partial q^\alpha} \dot{q}^\alpha + \frac{\partial L}{\partial \dot{q}^\alpha} \ddot{q}^\alpha = \frac{\delta L}{\delta q^\alpha} \dot{q}^\alpha + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^\alpha} \dot{q}^\alpha \right), \quad (9.10)$$

In other words,

$$E = p_\alpha \dot{q}^\alpha - L, \quad (9.11)$$

is the conserved quantity associated to the infinitesimal symmetry

$$\delta_Q q^\alpha = \dot{q}^\alpha. \quad (9.12)$$

In particular, if

$$L = \frac{1}{2} g_{\alpha\beta}(q) \dot{q}^\alpha \dot{q}^\beta + A_\alpha(q) \dot{q}^\alpha - V(q), \quad (9.13)$$

the associated conserved quantity

$$E = g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta + \cancel{A_\alpha \dot{q}^\alpha} - \frac{1}{2} g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta - \cancel{A_\alpha \dot{q}^\alpha} + V = \frac{1}{2} g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta + V. \quad (9.14)$$

corresponds to the mechanical energy.

9.2.2 Cyclic variables and conservation of canonical momentum

Suppose that a given variable q^α is “cyclic” which means that L does not depend on q^α but may depend on the associated velocity $\dot{q}^\alpha$. To fix ideas, suppose that the last of n variables, q^n is cyclic, $\frac{\partial L}{\partial q^n} = 0$. It follows that $\delta_Q q^\alpha = \delta_n^\alpha$ is an infinitesimal symmetry,

$$\delta_Q L = \delta_n^\alpha \frac{\partial L}{\partial q^\alpha} + \frac{d(\delta_n^\alpha)}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} = \frac{\partial L}{\partial q^n} + \frac{d(1)}{dt} \frac{\partial L}{\partial \dot{q}^n} = 0. \quad (9.15)$$

Since $F = 0$, the associated conserved quantity turns out to be the canonical momentum associated to q^n ,

$$\frac{\partial L}{\partial \dot{q}^\alpha} (\delta_n^\alpha) = \frac{\partial L}{\partial \dot{q}^n} = p_n. \quad (9.16)$$

9.2.3 One-parameter groups of transformations

Suppose that the Lagrangian is invariant under a set of invertible transformations $q'^\alpha = q'^\alpha(q^\beta, t; s)$ that depend continuously on a parameters s , and reduce to the identity at $s = 0$, $q'^\alpha(q^\beta, t; 0) = q^\alpha$. By invariance we mean that the transformed Lagrangian in terms of the new variables is equal to the old Lagrangian in terms of the same variables,

$$L'(q', \dot{q}', t) = L(q(q', t; s), \frac{d}{dt}q(q', t; s), t) = L(q', \dot{q}', t), \quad (9.17)$$

where the first equality merely expresses how the new Lagrangian in terms of the new coordinates is obtained in terms of the old one, while the second equality expresses the invariance of the Lagrangian. When using the inverse transformation, the latter is equivalent to

$$L(q, \dot{q}, t) = L(q'(q, t; s), \frac{d}{dt}q'(q, t; s), t). \quad (9.18)$$

In this case, the transformation

$$\boxed{\delta_R q^\alpha = R^\alpha(q, t) := \frac{\partial q'^\alpha}{\partial s} \Big|_{s=0}}, \quad (9.19)$$

defines an infinitesimal symmetry.

Indeed, $q'^\alpha = q^\alpha + sR^\alpha(q, t) + O(s^2)$, while $\frac{dq'^\alpha}{dt} = \dot{q}^\alpha + s\frac{dR^\alpha}{dt} + O(s^2)$. If $\delta_R \dot{q}^\alpha = \frac{\partial}{\partial s} \frac{dq'^\alpha}{dt} \Big|_{s=0}$, it follows that $\delta_R \dot{q}^\alpha = \frac{dR^\alpha}{dt}$, as it should for an infinitesimal virtual variation. In these terms, the invariance condition (9.18) becomes

$$L(q, \dot{q}, t) = L(q + sR + O(s^2), \dot{q} + s\frac{dR}{dt} + O(s^2), t), \quad \forall s. \quad (9.20)$$

In particular, for $s = 0$, this is an identity, while the terms linear in s on the RHS have to vanish. The latter can be obtained by taking $\frac{\partial}{\partial s}$ at $s = 0$, and yield

$$R^\alpha \frac{\partial L}{\partial q^\alpha} + \frac{dR^\alpha}{dt} \frac{\partial L}{\partial \dot{q}^\alpha} = 0, \quad (9.21)$$

which can also be written as $\delta_R L = 0$. According to Noether's theorem, the associated conserved quantity is given by

$$\boxed{K = p_\alpha \delta_R q^\alpha}. \quad (9.22)$$

Note that a cyclic variable can be interpreted as a particular case. Indeed, the transformation that leaves the Lagrangian invariant is given by $q'^\alpha = q^\alpha + s\delta_1^\alpha$.

Another particular case is a system of particles in three-dimensional Euclidean space, with a Lagrangian of the type

$$L = T - V, \quad T = \frac{1}{2} \sum_{i=1}^N m_{(i)} \dot{x}_{(i)}^\alpha \dot{x}_{\alpha(i)}, \quad V = V(r_{ij}). \quad (9.23)$$

In this expression, the summation in the kinetic term over the different particles has been written out explicitly, while the summation over the 3 components of the position of a single particle is implicit through the summation convention. Note also that we have defined $x_{\alpha(i)} = \delta_{\alpha\beta} x_{(i)}^\beta$, and that the potential $V(r_{ij})$ with $i < j$ is assumed to depend only on the relative distances between the different particles,

$$r_{ij} = \sqrt{(x_{(i)}^\alpha - x_{(j)}^\alpha)(x_{\alpha(i)} - x_{\alpha(j)})}. \quad (9.24)$$

If one now considers translations in Euclidean space,

$$\boxed{x_i'^\alpha = x_i^\alpha + a^\alpha}, \quad (9.25)$$

for constant a^α (which correspond to three different parameters s), invariance of the Lagrangian in the form (9.18) is manifest since $\frac{dx_i'^\alpha}{dt} = \frac{dx_i^\alpha}{dt}$ and $r_{ij}' = r_{ij}$. The associated infinitesimal transformations are

$$R_{\gamma i}^\alpha = \frac{\partial x_i'^\alpha}{\partial a^\gamma} \Big|_{a=0} = \delta_\gamma^\alpha, \quad (9.26)$$

with associated conserved quantities given by the components of total linear momenta,

$$\boxed{P_\gamma = \sum_i \frac{\partial L}{\partial \dot{x}_{(i)}^\gamma} = \sum_i m_{(i)} \dot{x}_{(i)}^\gamma}. \quad (9.27)$$

For rotations, we have

$$\boxed{x_i'^\alpha = A^\alpha_\beta x_i^\beta}, \quad A^\alpha_\beta \delta_{\alpha\gamma} A^\gamma_\delta = \delta_{\beta\delta} \iff A^T A = \mathbb{1}, \quad (9.28)$$

with A a constant 3×3 orthogonal matrix. Again, the invariance of the Lagrangian is manifest since orthogonality implies that $\frac{dx_{(i)}'^\alpha}{dt} \delta_{\alpha\gamma} \frac{dx_{(i)}'^\gamma}{dt} = \frac{dx_{(i)}^\alpha}{dt} \delta_{\alpha\gamma} \frac{dx_{(i)}^\gamma}{dt}$ while distances are also left invariant. In order to find the parameters, we note that if $A^\alpha_\beta = \delta^\alpha_\beta + \omega^\alpha_\beta + O(\omega^2)$, with ω^α_β a “small” matrix, the orthogonality condition

$$(\delta^\alpha_\beta + \omega^\alpha_\beta + O(\omega^2)) \delta_{\alpha\gamma} (\delta^\gamma_\delta + \omega^\gamma_\delta + O(\omega^2)) = \delta_{\beta\delta}, \quad (9.29)$$

implies to first order in ω that

$$\omega_{\delta\beta} + \omega_{\beta\delta} = 0 \iff \omega + \omega^T = 0. \quad (9.30)$$

This implies that ω can be written in terms of three independent rotation parameters θ^γ as

$$\omega^\alpha_\beta = \theta^\gamma \epsilon^\alpha_{\gamma\beta} \iff \omega = \begin{pmatrix} 0 & -\theta^3 & \theta^2 \\ \theta^3 & 0 & -\theta^1 \\ -\theta^2 & \theta^1 & 0 \end{pmatrix}, \quad (9.31)$$

and that the associated infinitesimal transformations are

$$R_{\gamma i}^\alpha = \frac{\partial x_i'^\alpha}{\partial \theta^\gamma} \Big|_{\theta=0} = \epsilon^\alpha_{\gamma\beta} x_i^\beta, \quad (9.32)$$

with associated conserved quantities given by the components of the total angular momentum,

$$\boxed{J_\gamma = \epsilon_{\gamma\alpha\beta} \sum_i x_{(i)}^\alpha m_{(i)} \dot{x}_{(i)}^\beta}. \quad (9.33)$$

We will show below that a rotation matrix that is connected to the identity¹ can be written as a rotation of an angle $0 \leq \theta < 2\pi$ around a fixed axis. The rotation may be determined a vector

$$\vec{\theta} = \theta \vec{n} = \theta^\alpha \vec{e}_\alpha, \quad (9.34)$$

with $\vec{e}_1 = \vec{1}_x$, $\vec{e}_2 = \vec{1}_y$, $\vec{e}_3 = \vec{1}_z$, of norm

$$\theta = \sqrt{(\theta^1)^2 + (\theta^2)^2 + (\theta^3)^2}, \quad (9.35)$$

$$A^\alpha_\beta = \cos \theta \delta^\alpha_\beta + \left(\frac{1 - \cos \theta}{\theta^2} \right) \theta^\alpha \theta_\beta - \frac{\sin \theta}{\theta} \epsilon^\alpha_{\beta\gamma} \theta^\gamma. \quad (9.36)$$

As before, one then finds that

$$\frac{\partial A^\alpha_\beta}{\partial \theta^\gamma} \Big|_{\theta=0} = \epsilon^\alpha_{\gamma\beta}. \quad (9.37)$$

¹This means that one can assume that, if the different parameters are set to zero, the matrix that determines the rotation can be taken to be the unit matrix.

9.2.4 Parametrization of a rotation

Consider a rotation matrix A in 3 dimensions with entries A^α_β , $\alpha, \beta = 1, 2, 3$,

$$A^T A = \mathbb{1} \iff A_{\alpha\beta} A^\alpha_\gamma = \delta_{\beta\gamma}, \quad (9.38)$$

where the Kronecker symbols $\delta_{\alpha\beta}, \delta^{\alpha\beta}$ are used to lower and raise indices. By taking the determinant, it follows that $\det A = \pm 1$. In the following, we will concentrate on rotation matrices with unit determinant,

$$\det A = 1. \quad (9.39)$$

Invariant line

For such a matrix, the rotation, i.e., the transformation

$$x'^\alpha = A^\alpha_\beta x^\beta, \quad (9.40)$$

leaves a line invariant.

Indeed, such a line is defined by a non trivial solution to the equation

$$x^\alpha = A^\alpha_\beta x^\beta \iff (A - \mathbb{1})\vec{x} = 0. \quad (9.41)$$

The existence of a non-trivial solution $\vec{x} \neq \vec{0}$ requires that

$$\det(A - \mathbb{1}) = 0. \quad (9.42)$$

Let us suppose that $\det(A - \mathbb{1}) = k$. From the defining equation of a rotation matrix, we have

$$A^T(A - \mathbb{1}) = \mathbb{1} - A^T = (\mathbb{1} - A)^T = -(A - \mathbb{1})^T. \quad (9.43)$$

By taking the determinant of this equation, it follows that

$$1k = (-1)^3 k \implies k = -k \implies k = 0. \quad (9.44)$$

Let us denote by $\vec{\theta}$ a non trivial solution to $A\vec{\theta} = \vec{\theta}$. In other words, the line which is left invariant by the rotation is $\lambda\vec{\theta}$, $\lambda \in \mathbb{R}$. Furthermore $\vec{\theta} = \theta\vec{n}$, $\theta = \sqrt{(\theta^1)^2 + (\theta^2)^2 + (\theta^3)^2}$ is the norm of $\vec{\theta}$, while $\vec{n} = \frac{1}{\theta}\theta^\alpha \vec{e}_\alpha$, with components $\hat{\theta}^\alpha = \frac{\theta^\alpha}{\theta}$ is the unit vector along the axis of rotation, while $\vec{e}_\alpha$ are the unit vectors along the coordinate axis of $\mathbb{R}^3$.

Geometry of a rotation in 3 dimensions

If $\vec{OP} = x^\alpha \vec{e}_\alpha = \vec{x}$ denotes the vector relating the origin O with coordinates $x^\alpha = 0$ to an arbitrary point P with coordinates x^α , the image of this vector under the rotation (9.40) is the vector $\vec{OP}' = x'^\alpha \vec{e}_\alpha = \vec{x}'$, where the coordinates of P' are x'^α . Furthermore, by suitably choosing the orientation of $\vec{n}$, one may arrange that $(\vec{n}, \vec{OP}, \vec{OP}')$ is oriented positively.

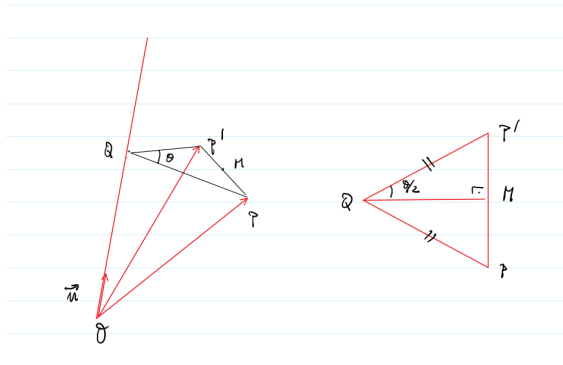


Figure 9.1: Geometry of a rotation

A rotation preserves the scalar products and thus the norms of vectors,

$$\vec{u}' \cdot \vec{v}' = u'_\alpha v'^\alpha = A_{\alpha\beta} u^\beta A^\alpha_\gamma v^\gamma = u^\beta \delta_{\beta\gamma} v^\gamma = u_\beta v^\beta. \quad (9.45)$$

Let us define the point Q to be the orthogonal projection of P on the line defined by $\vec{n}$, $\overrightarrow{OQ} = (\overrightarrow{OP} \cdot \vec{n})\vec{n}$, and show that Q, P, P' lie in the plane perpendicular to $\vec{n}$. Writing $\overrightarrow{OP} = \overrightarrow{OQ} + \overrightarrow{QP}$, it follows that

$$\overrightarrow{QP} \cdot \vec{n} = (\overrightarrow{OP} - \overrightarrow{OQ}) \cdot \vec{n} = \overrightarrow{OP} \cdot \vec{n} - (\overrightarrow{OP} \cdot \vec{n})\vec{n} \cdot \vec{n} = 0 \quad (9.46)$$

Furthermore,

$$\overrightarrow{QP'} \cdot \vec{n} = \overrightarrow{OP'} \cdot \vec{n} - \overrightarrow{OQ} \cdot \vec{n} = A(\overrightarrow{OP}) \cdot A(\vec{n}) - \overrightarrow{OQ} \cdot \vec{n} = \overrightarrow{OP} \cdot \vec{n} - \overrightarrow{OQ} \cdot \vec{n} = 0. \quad (9.47)$$

Furthermore

$$||\overrightarrow{QP}|| = ||\overrightarrow{QP'}||. \quad (9.48)$$

Indeed, starting from $||\overrightarrow{OP}||^2 = ||\overrightarrow{OP'}||^2$, by using that

$$||\vec{a} + \vec{b}||^2 = a^2 + b^2 + 2ab \cos \phi, \quad (9.49)$$

with a, b the norms of $\vec{a}, \vec{b}$ and ϕ the angle between them, one develops the left hand side as

$$||\overrightarrow{OP}||^2 = ||\overrightarrow{OQ} + \overrightarrow{QP}||^2 = ||\overrightarrow{OQ}||^2 + ||\overrightarrow{QP}||^2 + 2\overrightarrow{OQ} \cdot \overrightarrow{QP}. \quad (9.50)$$

The last term vanishes since both vectors are orthogonal. The result follows by applying the same reasoning to the right hand side, that is to say decomposing $\overrightarrow{OP'} = \overrightarrow{OQ} + \overrightarrow{QP'}$, and after having computed the norm, cancelling the common term $||\overrightarrow{OQ}||^2$ on both sides of the equation.

The triangle QPP' thus has two sides of equal length, QP and QP' so that, if M is the point in the middle of PP' , QM bisects the angle at Q and $\overrightarrow{QM}$ is orthogonal to $\overrightarrow{PP'}$. In the plane orthogonal to the axis of rotation, the angle at Q , which will be called θ , is the angle of rotation that brings you from P to P' .

By using that

$$\vec{a} \times \vec{b} = ab \sin \phi \vec{n}, \quad (9.51)$$

where $\vec{n}$ is the normal to the plane formed by $\vec{a}, \vec{b}$, it follows that

$$\overrightarrow{QM} \times \overrightarrow{PP'} = QM PP' \vec{n}, \quad (9.52)$$

since the angle between the two vectors is $\frac{\pi}{2}$ radians. Furthermore,

$$QM = QP' \cos \frac{\theta}{2}, \quad \frac{PP'}{2} = MP' = QP' \sin \frac{\theta}{2}. \quad (9.53)$$

Taking the quotient of these two equations yields

$$2 \tan \frac{\theta}{2} = \frac{PP'}{QM}. \quad (9.54)$$

Since $(\vec{n}, \overrightarrow{OP}, \overrightarrow{OP'})$ is positively oriented, it follows for the unit vectors that $\vec{n} \times \overrightarrow{1_{QM}} = \overrightarrow{1_{PP'}}$. Multiplying by the previous equation yields

$$2 \tan \frac{\theta}{2} \vec{n} \times \overrightarrow{1_{QM}} = \frac{PP'}{QM} \overrightarrow{1_{PP'}}. \quad (9.55)$$

Defining

$$\vec{\Omega} = \tan \frac{\theta}{2} \vec{n}, \quad \Omega = \tan \frac{\theta}{2}, \quad (9.56)$$

one may rewrite the former equation as

$$2\vec{\Omega} \times \overrightarrow{QM} = \overrightarrow{PP'}. \quad (9.57)$$

Analytic expression of the rotation matrix

We want an expression of $\overrightarrow{OP'}$ in terms of $\overrightarrow{OP}$.

By using that $\overrightarrow{QM} = \frac{1}{2}(\overrightarrow{QP} + \overrightarrow{QP'})$, it follows from the previous equation that

$$\overrightarrow{PP'} = \vec{\Omega} \times (\overrightarrow{QP} + \overrightarrow{QP'}) = \vec{\Omega} \times (\overrightarrow{OP} + \overrightarrow{OP'} - 2\overrightarrow{OQ}), \quad (9.58)$$

where the last term does not contribute since it is parallel to $\vec{\Omega}$. When using that

$$\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}, \quad (9.59)$$

it follows by applying $\vec{\Omega} \times$ to (9.58) that

$$\vec{\Omega} \times \overrightarrow{PP'} = (\vec{\Omega} \cdot (\overrightarrow{OP} + \overrightarrow{OP'}))\vec{\Omega} - (\vec{\Omega} \cdot \vec{\Omega})(\overrightarrow{OP} + \overrightarrow{OP'}). \quad (9.60)$$

When using on the left hand side that

$$\overrightarrow{PP'} = \overrightarrow{QP'} - \overrightarrow{QP} = \overrightarrow{OP'} - \overrightarrow{OP}, \quad (9.61)$$

and on the right hand side that

$$\vec{\Omega} \cdot \overrightarrow{OP} = \vec{\Omega} \cdot \overrightarrow{OQ} = \vec{\Omega} \cdot \overrightarrow{OP'} \quad (9.62)$$

equation (9.60) becomes

$$\vec{\Omega} \times \overrightarrow{OP'} - \vec{\Omega} \times \overrightarrow{OP} = 2(\vec{\Omega} \cdot \overrightarrow{OP})\vec{\Omega} - \Omega^2(\overrightarrow{OP} + \overrightarrow{OP'}). \quad (9.63)$$

Using (9.61) in the left hand side of (9.58) and (9.63) in the right hand side yields

$$\overrightarrow{OP'} - \overrightarrow{OP} = 2\vec{\Omega} \times \overrightarrow{OP} + 2(\vec{\Omega} \cdot \overrightarrow{OP})\vec{\Omega} - \Omega^2(\overrightarrow{OP} + \overrightarrow{OP'}), \quad (9.64)$$

or equivalently

$$\overrightarrow{OP'} = \frac{1}{1 + \Omega^2} \left[(1 - \Omega^2)\overrightarrow{OP} + 2\vec{\Omega} \times \overrightarrow{OP} + 2(\vec{\Omega} \cdot \overrightarrow{OP})\vec{\Omega} \right]. \quad (9.65)$$

The trigonometric relations

$$\frac{1 - \Omega^2}{1 + \Omega^2} = \frac{1 - \tan^2 \frac{\theta}{2}}{1 + \tan^2 \frac{\theta}{2}} = \frac{\frac{\cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2}}{\cos^2 \frac{\theta}{2}}}{\frac{\cos^2 \frac{\theta}{2} + \sin^2 \frac{\theta}{2}}{\cos^2 \frac{\theta}{2}}} = \cos^2 \frac{\theta}{2} - \sin^2 \frac{\theta}{2} = \cos \theta, \quad (9.66)$$

$$\frac{2\Omega}{1 + \Omega^2} = 2 \cos^2 \frac{\theta}{2} \tan \frac{\theta}{2} = 2 \sin \frac{\theta}{2} \cos \frac{\theta}{2} = \sin \theta, \quad (9.67)$$

$$\frac{2\Omega^2}{1 + \Omega^2} = \frac{2 \tan^2 \frac{\theta}{2}}{1 + \tan^2 \frac{\theta}{2}} = 2 \sin^2 \frac{\theta}{2} = 1 - \cos \theta, \quad (9.68)$$

then allow one to rewrite

$$\vec{x}' = A\vec{x} = \cos \theta \vec{x} + \sin \theta \vec{n} \times \vec{x} + (1 - \cos \theta)(\vec{n} \cdot \vec{x})\vec{n} = \cos \theta \vec{x} + \frac{\sin \theta}{\theta} \vec{\theta} \times \vec{x} + \frac{1 - \cos \theta}{\theta^2} (\vec{\theta} \cdot \vec{x})\vec{\theta}. \quad (9.69)$$

When expressed in components, this gives

$$x'^{\alpha} = A^{\alpha}_{\beta} x^{\beta} = \cos \theta x^{\alpha} + \frac{\sin \theta}{\theta} \epsilon^{\alpha}_{\beta\gamma} \theta^{\beta} x^{\gamma} + \left(\frac{1 - \cos \theta}{\theta^2} \right) \theta^{\alpha} \theta_{\beta} x^{\beta}. \quad (9.70)$$

When identifying the coefficient of x^β on the right hand side, this gives the rotation matrix,

$$A^\alpha{}_\beta = \cos \theta \delta^\alpha_\beta + \left(\frac{1 - \cos \theta}{\theta^2} \right) \theta^\alpha \theta_\beta - \frac{\sin \theta}{\theta} \epsilon^\alpha{}_{\beta\gamma} \theta^\gamma, \quad (9.71)$$

or, equivalently, in terms of

$$\vec{n} = \hat{\theta}^\alpha e_\alpha, \quad \hat{\theta}^\alpha = \frac{\theta^\alpha}{\theta}, \quad (9.72)$$

as

$$A^\alpha{}_\beta = \cos \theta (\delta^\alpha_\beta - \hat{\theta}^\alpha \hat{\theta}_\beta) + \hat{\theta}^\alpha \hat{\theta}_\beta + \sin \theta \hat{\theta}^\gamma \epsilon^\alpha{}_{\gamma\beta}. \quad (9.73)$$

Worked exercise: Rotation around the $\vec{1}_z$ axis

In order to find the analytic expression for a rotation matrix around the z axis by an angle θ in the (x, y) plane from the expression (9.70), we note that $\theta^1 = \theta^2 = 0$, while $\theta^3 = \theta$. It follows that $-\frac{\sin \theta}{\theta} \epsilon^\alpha{}_{\beta\gamma} \theta^\gamma = -\frac{\sin \theta}{\theta} \epsilon^\alpha{}_{\beta 3} \theta$, so that

$$\begin{aligned} A_{\vec{1}_z} &= \begin{pmatrix} \cos \theta & 0 & 0 \\ 0 & \cos \theta & 0 \\ 0 & 0 & \cos \theta \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{\theta^2(1-\cos \theta)}{\theta^2} \end{pmatrix} + \begin{pmatrix} 0 & -\sin \theta & 0 \\ \sin \theta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \quad (9.74)$$

which is the expected result from the viewpoint of a rotation matrix in the (x, y) plane.

9.2.5 One-parameter groups of transformations involving time

Consider now a one-parameters group of invertible transformations that affects time as well,

$$\begin{cases} q'^\alpha = q'^\alpha(q^\beta, t; s), \\ t' = t'(t; s) \end{cases} \implies q'^\alpha(t') = q'^\alpha(q^\beta(t(t'; s)), t(t'; s); s), \quad (9.75)$$

with $q'^\alpha = q'^\alpha(q^\beta, t; 0) = q^\alpha$, $t'(t; 0) = t$. According to the equation (7.26) of the previous section, the transformed Lagrangian is

$$L'(q'(t'), \frac{dq'(t')}{dt'}, t') = \frac{dt}{dt'} L(q(q'(t'), t'), \frac{dt'}{dt} \frac{d}{dt'} q(q'(t'), t'), t(t')), \quad (9.76)$$

where we have temporarily suppressed the s dependence. Let us assume that the transformed Lagrangian is equal to the starting point Lagrangian (in terms of the new variables), up to a total derivative,

$$L'(q'(t'), \frac{dq'(t')}{dt'}, t') = L(q'(t'), \frac{d}{dt'} q'(t'), t') - \frac{d}{dt'} G(q'(t'), t'), \quad (9.77)$$

with $G'|_{s=0} = 0$. When using the inverse transformation and multiplying by $\frac{dt'}{dt}$, the generalized invariance condition (9.77), where the left hand side has been expressed in terms of L by using (9.76), can also be written as

$$L(q(t), \frac{d}{dt} q(t), t) = \frac{dt'}{dt} L(q'(q, t), \frac{dt'}{dt} \frac{d}{dt} q'(q, t), t'(t)) - \frac{d}{dt} G(q(t), t), \quad (9.78)$$

with $G(q(t), t) = G'(q'(q, t), t'(t))$.

In terms of the action, this generalized invariance condition means that as a functional of the new trajectories and the transformed end-points, the

$$S[q'] = \int_{t'=t'(t_i)}^{t'=t'(t_f)} dt L(q'(t), \frac{d}{dt}q'(t), t), \quad (9.79)$$

equals the action as a functional of the old trajectories, up to a boundary term. Indeed, by first changing the name of the dummy integration variable, and then doing the change of variables from primed to unprimed ones, we get

$$S[q] = \int_{t'=t'(t_i)}^{t'=t'(t_f)} dt' L(q'(t'), \frac{d}{dt'}q'(t'), t') = \int_{t=t_i}^{t=t_f} dt \frac{dt'}{dt} L(q'(q(t), t), \frac{dt}{dt'} \frac{d}{dt}q'(q(t), t), t'(t)). \quad (9.80)$$

When using the generalized invariance condition (9.78), this gives

$$S[q] = \int_{t_i}^{t_f} dt L(q(t), \frac{d}{dt}q(t), t) + \left[G(q(t), t) \right]_{t_i}^{t_f}. \quad (9.81)$$

Note that the boundary term does not contribute when computing the variation of trajectories with fixed end-points.

Expanding the transformed trajectories to first order in s gives

$$\begin{cases} q'^\alpha(t') = q^\alpha(t) + sR^\alpha(q^\beta, t) + O(s^2), \\ t' = t + s\xi(t) + O(s^2), \end{cases}, \quad G = sG_1(q, t) + O(s^2), \quad (9.82)$$

the associated equal-time infinitesimal virtual variation is

$$\begin{aligned} \boxed{\delta_S q^\alpha(t)} &= \lim_{s \rightarrow 0} \frac{1}{s} (q'^\alpha(t) - q^\alpha(t)) = \lim_{s \rightarrow 0} \frac{1}{s} (q'^\alpha(t' - s\xi(t) + O(s^2)) - q^\alpha(t)) \\ &= \lim_{s \rightarrow 0} \frac{1}{s} (\cancel{q^\alpha(t)} + sR^\alpha(q^\beta(t), t) - s \frac{dq'^\alpha}{dt'}(t')\xi(t) + O(s^2) - \cancel{q^\alpha(t)}) = \boxed{R^\alpha - \dot{q}^\alpha \xi}. \end{aligned} \quad (9.83)$$

Such a variation leaves the Lagrangian invariant up to a total derivative. Indeed, the generalized invariance condition (9.78) can be written as

$$L(q, \dot{q}, t) = (1 + s\dot{\xi}) \left(L(q + sR, (1 - s\dot{\xi}) \frac{d}{dt}(q + sR), t + s\xi) - s \frac{d}{dt} G_1 + O(s^2) \right). \quad (9.84)$$

To order 0 in s , this is an identity, while the terms of order 1 yield

$$\dot{\xi} L(q, \dot{q}, t) + \frac{\partial L}{\partial q^\alpha} R^\alpha + \frac{\partial L}{\partial \dot{q}^\alpha} \left(\frac{d}{dt} R^\alpha - \dot{\xi} \dot{q}^\alpha \right) + \cancel{\frac{\partial L}{\partial t}(q, \dot{q}, t) \xi} - \frac{d}{dt} G_1 = 0. \quad (9.85)$$

The first term can be integrated by parts

$$\dot{\xi} L = \frac{d}{dt}(\xi L) - \xi \left(\frac{\partial L}{\partial q^\alpha} \dot{q}^\alpha + \frac{\partial L}{\partial \dot{q}^\alpha} \ddot{q}^\alpha + \cancel{\frac{\partial L}{\partial t}} \right), \quad (9.86)$$

and, when cancelling (as indicated) and combining like terms, we get

$$\frac{\partial L}{\partial q^\alpha} (R^\alpha - \dot{q}^\alpha \xi) + \frac{\partial L}{\partial \dot{q}^\alpha} \frac{d}{dt} (R^\alpha - \dot{q}^\alpha \xi) = \frac{d}{dt} (G_1 - \xi L) \iff \delta_S L = \frac{d}{dt} (G_1 - \xi L). \quad (9.87)$$

The associated conserved quantity is

$$\boxed{K = p_\alpha \delta_S q^\alpha + \xi L - G_1}. \quad (9.88)$$

9.2.6 Time-translation symmetry and conservation of energy

In particular, a Lagrangian does not depend explicit on time, $\frac{\partial L}{\partial t} = 0$ iff it possesses the symmetry $q'^\alpha = q^\alpha$, $t' = t - s$. In this case $G = 0$, $\xi = -1$, $R^\alpha = 0$, $\delta_S q^\alpha = \dot{q}^\alpha$. The associated conserved quantity is

$$E = p_\alpha \dot{q}^\alpha - L, \quad (9.89)$$

discussed previously, which now appears from the viewpoint of the symmetry of the system under constant time translations.

9.2.7 Galilean boosts

Galilean transformations on a system of particles are described by

$$x_i'^\alpha = A^\alpha_\beta x_i^\beta + v^\alpha t + a^\alpha, t' = t + b, \quad A^T A = \mathbb{1}. \quad (9.90)$$

They consists of a combination of spatial translations, rotations, time-translations, described by the constant parameter b , and changes of observers with a constant velocity in a given direction, sometimes called “Galilean boosts” and described by the 3 parameters v^α .

Sometimes it is useful to provide the following 5×5 matrix form of these transformations (suppressing the particle label)

$$\begin{pmatrix} x'^\alpha \\ t' \\ 1 \end{pmatrix} = \begin{pmatrix} A^\alpha_\beta & v^\alpha & a^\alpha \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^\beta \\ t \\ 1 \end{pmatrix}. \quad (9.91)$$

We have already analyzed invariance of the particle Lagrangian (9.23) under spatial translations (parameters a^α) and rotations (parameters ω^α of the rotation matrix A) and the associated conserved quantities P_α , J_α . For the time-translation, explicit time-independence implies that the associated conserved quantity is the energy, $E = \sum_i \frac{\partial L}{\partial \dot{x}_i^\alpha} \dot{x}_i^\alpha - L$. So only the Galilean boosts,

$$x_i'^\alpha = x_i^\alpha + v^\alpha t, \quad t' = t, \quad (9.92)$$

remain to be analyzed explicitly.

More generally, for the whole set of Galilean transformations, we have $\frac{dt'}{dt} = 1$, $\frac{dx_i'^\alpha(dt')}{dt} = A^\alpha_\beta \frac{dx_i^\beta(t)}{dt} + v^\alpha$. For the kinetic energy of the particle Lagrangian (9.23), this gives

$$\begin{aligned} T\left(\frac{dx_i'^\alpha(t')}{dt'}\right) &= \frac{1}{2} \sum_i m_{(i)} \left(A^\alpha_\beta \frac{dx_{(i)}^\beta}{dt} + v^\alpha \right) \left(A_{\alpha\gamma} \frac{dx_{(i)}^\gamma}{dt} + v_\alpha \right) \\ &= \frac{1}{2} \sum_i m_{(i)} \frac{dx_{(i)}^\alpha}{dt} \frac{dx_{\alpha(i)}}{dt} + \sum_i m_{(i)} v_\alpha A^\alpha_\beta \frac{dx_{(i)}^\beta}{dt} + \frac{1}{2} \sum_i m_{(i)} v^\alpha v_\alpha \\ &= T\left(\frac{dx_i^\alpha}{dt}\right) + \frac{d}{dt} \left(\sum_i m_{(i)} v_\alpha (A^\alpha_\beta x_{(i)}^\beta + \frac{1}{2} t v^\alpha) \right). \end{aligned} \quad (9.93)$$

Since relative distances are invariant even if one translates the individual coordinates linearly in time (by the same amount), the potential is also invariant. It follows that

$$G = \sum_i m_{(i)} v_\alpha \left(A^\alpha_\beta x_{(i)}^\beta + \frac{1}{2} t v^\alpha \right), \quad (9.94)$$

and, for Galilean boosts parametrized by v^γ , $R_{\gamma i}^\alpha = \delta_\gamma^\alpha t$, $\xi = 0$, $\frac{\partial G}{\partial v^\gamma}|_{a,b,\theta,v=0} = \sum_i m_{(i)} x_{\gamma(i)}$. The associated conserved quantities are thus

$$K_\gamma = \sum_i m_{(i)} (\dot{x}_{\gamma(i)} t - x_{\gamma(i)}). \quad (9.95)$$

9.3 Hamiltonian version of Noether's theorem

As a consequence of (6.26), in the Hamiltonian formalism, a first integral is a (possibly time dependent) phase space function $K(z^A, t)$ that satisfies

$$\boxed{\{K, H\} + \frac{\partial K}{\partial t} = 0}. \quad (9.96)$$

It follows directly from the properties of the Poisson bracket and the partial derivative that first integrals form a vector space. Furthermore,

Theorem 11. *The Poisson bracket $\{K_1, K_2\}$ of two first integrals K_1, K_2 is a first integral.*

Indeed,

$$\begin{aligned} & \{\{K_1, K_2\}, H\} + \frac{\partial}{\partial t}\{K_1, K_2\} \\ &= \{\{K_1, H\}, K_2\} + \{K_1, \{K_2, H\}\} + \frac{\partial}{\partial t} \frac{\partial K_1}{\partial z^A} \omega^{AB} \frac{\partial K_2}{\partial z^B} + \frac{\partial K_1}{\partial z^A} \omega^{AB} \frac{\partial}{\partial t} \frac{\partial K_2}{\partial z^B} \\ &= \{\{K_1, H\} + \frac{\partial K_1}{\partial t}, K_2\} + \{K_1, \{K_2, H\} + \frac{\partial K_2}{\partial t}\} = 0, \end{aligned} \quad (9.97)$$

where we have used the Jacobi identity.

By definition, a Lie algebra is a vector space equipped with a bi-linear, skew-symmetric operation that satisfies the Jacobi identity. As a consequence, in the Hamiltonian formalism, first integrals equipped with the Poisson bracket form a Lie algebra.

In the Hamiltonian formalism, infinitesimal symmetries are variations

$$\delta_Z z^A = Z^A(z, t) \iff \begin{cases} \delta_Z q^\alpha = Q^\alpha(q, p, t), \\ \delta_Z p_\alpha = P_\alpha(q, p, t), \end{cases} \quad (9.98)$$

that leave the first order Hamiltonian Lagrangian $L_H = \dot{q}^\alpha p_\alpha - H(z, t)$ invariant up to a total time derivative,

$$\delta_Z L_H = \frac{d}{dt} F(z, t). \quad (9.99)$$

Theorem 12. *There is a 1-1 correspondence between symmetries and first integrals, up to constants. Furthermore, the first integral is the canonical generator of the symmetry in the sense that*

$$Z^A = \{z^A, K\} = -v_K^A. \quad (9.100)$$

Writing out explicitly (9.99), we get

$$\begin{aligned} & P_\alpha \dot{q}^\alpha + p_\alpha \frac{d}{dt} Q^\alpha - Q^\alpha \frac{\partial H}{\partial q^\alpha} - P_\alpha \frac{\partial H}{\partial p_\alpha} = \frac{d}{dt} F \\ & \iff P_\alpha \dot{q}^\alpha - Q^\alpha \dot{p}_\alpha - Q^\alpha \frac{\partial H}{\partial q^\alpha} - P_\alpha \frac{\partial H}{\partial p_\alpha} = \frac{d}{dt} (F - Q^\alpha p_\alpha). \end{aligned} \quad (9.101)$$

Defining

$$\boxed{K = Q^\alpha p_\alpha - F}, \quad (9.102)$$

this equation may be re-written as

$$P_\alpha \dot{q}^\alpha - Q^\alpha \dot{p}_\alpha - Q^\alpha \frac{\partial H}{\partial q^\alpha} - P_\alpha \frac{\partial H}{\partial p_\alpha} + \dot{q}^\alpha \frac{\partial K}{\partial q^\alpha} + \dot{p}_\alpha \frac{\partial K}{\partial p_\alpha} + \frac{\partial K}{\partial t} = 0. \quad (9.103)$$

Because this is an identity that has to hold for all $q^\alpha, p_\alpha, \dot{q}^\alpha, \dot{p}_\alpha, t$, one may identify the coefficients of $\dot{q}^\alpha, \dot{p}_\alpha$ to find

$$P_\alpha = -\frac{\partial K}{\partial \dot{q}^\alpha} = \{p_\alpha, K\}, \quad Q^\alpha = \frac{\partial K}{\partial p_\alpha} = \{q^\alpha, K\} \iff Z^A = \{z^A, K\} = -v_K^A. \quad (9.104)$$

Inserting these expressions back into (9.103) yields

$$-\frac{\partial K}{\partial p_\alpha} \frac{\partial H}{\partial q^\alpha} + \frac{\partial K}{\partial q^\alpha} \frac{\partial H}{\partial p_\alpha} + \frac{\partial K}{\partial t} = 0 \iff \{K, H\} + \frac{\partial K}{\partial t} = 0, \quad (9.105)$$

which shows that K is a first integral. Note that, for a given symmetry, F , and thus also K , are defined only up to a constant by this procedure.

Conversely, for a given first integral K , one may follow the proof in the reverse direction to show that $Z^A = \{z^A, K\}$ defines a symmetry. Finally, if two first integrals define the same symmetry, $\{z^A, K_1\} = Z^A = \{z^A, K_2\}$, it follows that $\{z^A, K_1 - K_2\} = 0$, or $\omega^{AB} \frac{\partial(K_1 - K_2)}{\partial z^B} = 0$, and thus that $\frac{\partial(K_1 - K_2)}{\partial z^B} = 0$. As a consequence, $K_1 - K_2 = c(t)$. But since $K_1 - K_2$ is a first integral, $\{K_1 - K_2, H\} + \frac{\partial(K_1 - K_2)}{\partial t} = 0$, it follows that c is a constant.

Theorem 13. *Infinitesimal symmetries equipped with the commutator of symmetries constitute a Lie algebra isomorphic to the Lie algebra of first integrals equipped with the Poisson bracket, up to constants.*

Proof :

9.4 More applications

9.4.1 Hamiltonian formulation of the Galilean algebra

9.4.2 Worked Exercise 4: Geodesic motion and Killing vectors

Consider the generalized coordinates q^α and the Lagrangian

$$L = \frac{1}{2} g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta, \quad (9.106)$$

where the “metric” $g_{\alpha\beta} = g_{\alpha\beta}(q)$ is a given function of the coordinates.

Defining the Christoffel symbols of the second kind by

$$\Gamma_{\beta\gamma}^\delta = g^{\delta\alpha} \Gamma_{\alpha\beta\gamma}, \quad (9.107)$$

where $g^{\delta\alpha}$ denotes the inverse metric, $g^{\delta\alpha} g_{\alpha\beta} = \delta_\beta^\delta$, show that the equations of motion are equivalent to the geodesic equation

$$\ddot{q}^\delta + \Gamma_{\beta\gamma}^\delta \dot{q}^\beta \dot{q}^\gamma = 0. \quad (9.108)$$

Solution: This follows directly by multiplying the two non-vanishing first terms in (4.35) by the inverse metric.

Let us introduce the notation $\partial_\alpha = \frac{\partial}{\partial q^\alpha}$. Suppose that there is a $\xi^\alpha(q)$ that satisfies the “Killing equation”

$$\xi^\gamma \partial_\gamma g_{\alpha\beta} + \partial_\alpha \xi^\gamma g_{\gamma\beta} + \partial_\beta \xi^\gamma g_{\alpha\gamma} = 0. \quad (9.109)$$

Show that the transformation $\delta_\xi^1 q^\alpha = \xi^\alpha(q)$ defines a symmetry with associated Noether charge

$$K_\xi^1 = g_{\alpha\beta} \dot{q}^\alpha \xi^\beta. \quad (9.110)$$

Solution: The variation of the Lagrangian vanishes on account of (9.109),

$$\delta_\xi^1 L = \frac{1}{2} \partial_\gamma g_{\alpha\beta} \xi^\gamma + \dot{q}^\alpha \dot{q}^\beta + g_{\alpha\gamma} \dot{q}^\alpha \partial_\beta \xi^\gamma \dot{q}^\beta = 0. \quad (9.111)$$

The result then follows from the formula for the Noether charge.

9.4.3 Worked Exercise 5: Geodesic motion and Killing tensors

This exercise requires an understanding of the covariant derivative D_μ defined in terms of the Levi-Civita connection $\Gamma_{\beta\gamma}^\delta$. It is metric compatible $D_\alpha g_{\beta\gamma} = 0$. The Killing equation (9.109) can then be shown to be equivalent to

$$D_{(\alpha_0} \xi_{\alpha_1)} = 0, \quad (9.112)$$

where indices are lowered and raised with the metric $g_{\mu\nu}$ and its inverse, and a round parenthesis denotes complete symmetrization of the enclosed indices,

$$M_{(\alpha_1 \dots \alpha_k)} = \frac{1}{k!} \sum_{\sigma \in S_k} M_{\alpha_{\sigma(1)} \dots \alpha_{\sigma(k)}}. \quad (9.113)$$

A Killing tensor is a symmetric tensor $\xi_{\alpha_1 \dots \alpha_k} = \xi_{(\alpha_1 \dots \alpha_k)}$ that satisfies

$$D_{(\alpha_0} \xi_{\alpha_1 \dots \alpha_k)} = 0. \quad (9.114)$$

In particular, a Killing vector corresponds to the particular case where $k = 1$.

Show that

$$K_\xi^k = \frac{1}{k!} \xi_{\alpha_1 \dots \alpha_k} \dot{q}^{\alpha_1} \dots \dot{q}^{\alpha_k} \quad (9.115)$$

is a conserved quantity and that the associated symmetry is

$$\delta_\xi^k q^\alpha = \frac{1}{(k-1)!} \xi_{\alpha_2 \dots \alpha_k}^{\alpha} \dot{q}^{\alpha_2} \dots \dot{q}^{\alpha_k}. \quad (9.116)$$

Solution:

$$\begin{aligned} \frac{d}{dt} K_\xi^k &= \frac{1}{k!} \partial_{\alpha_0} \xi_{\alpha_1 \dots \alpha_k} \dot{q}^{\alpha_0} \dot{q}^{\alpha_1} \dots \dot{q}^{\alpha_k} + \frac{1}{(k-1)!} \xi_{\alpha_1 \dots \alpha_k} \ddot{q}^{\alpha_1} \dot{q}^{\alpha_2} \dots \dot{q}^{\alpha_k} \\ &= \frac{1}{k!} \partial_{\alpha_0} \xi_{\alpha_1 \dots \alpha_k} \dot{q}^{\alpha_0} \dot{q}^{\alpha_1} \dots \dot{q}^{\alpha_k} + \frac{1}{(k-1)!} \xi_{\alpha_1 \dots \alpha_k} \left(-g^{\alpha_1 \beta} \frac{\delta L}{\delta q^\beta} \right) \dot{q}^{\alpha_2} \dots \dot{q}^{\alpha_k} \\ &\quad - \frac{1}{(k-1)!} \xi_{\beta \alpha_2 \dots \alpha_k} \Gamma_{\alpha_0 \alpha_1}^\beta \dot{q}^{\alpha_0} \dot{q}^{\alpha_1} \dots \dot{q}^{\alpha_k}. \end{aligned} \quad (9.117)$$

The first and the third term vanish on account of (9.115). Since the second term is proportional to the field equations, K_ξ^k is a conserved quantity, and one may read off the characteristic of the symmetry (9.116).

9.4.4 Worked Exercise 6: Laplace-Runge-Lenz vector

9.4.5 Worked Exercise 7: Harmonic oscillator conserved quantities

Chapter 10

Reduced and alternative variational principles

10.1 Routhian procedure for cyclic variables

In a system with n Lagrange coordinates q^α , suppose that q^n is a cyclic (ignorable, kinosthenic) variable, and let us denote by $q^{\bar{\alpha}}$ the first $n - 1$ Lagrange coordinates. Noether's implies that $p_n = \frac{\partial L}{\partial \dot{q}^n} = c_n$ is a constant of the motion. Suppose that this relation allows to express $\dot{q}^n$ as a function of the other variables

$$\dot{q}^n = U^n(q^{\bar{\alpha}}, \dot{q}^{\bar{\alpha}}, c_n, t). \quad (10.1)$$

Note that this implies that

$$\left. \frac{\partial L}{\partial \dot{q}^n} \right|_{\dot{q}^n = U^n} = c_n. \quad (10.2)$$

It follows that the Lagrangian equations of motion, i.e., a system of n differential equations of second order can be reduced to a system of $n - 1$ differential equations of second order.

The Routhian procedure provides a reduced Lagrangian $\bar{L}(q^{\bar{\alpha}}, \dot{q}^{\bar{\alpha}}, t)$ such that the reduced system are the Euler-Lagrange equations of the reduced Lagrangian. For the starting point variational principle, we have

$$\delta \int_{t_i}^{t_f} dt L = \int_{t_i}^{t_f} dt \frac{\delta L}{\delta q^\alpha} \delta q^\alpha + \left[\frac{\partial L}{\partial \dot{q}^\alpha} \delta q^\alpha \right]_{t_i}^{t_f} \quad (10.3)$$

On a natural trajectory, this reduces to

$$\delta \int_{t_i}^{t_f} dt L \approx [p_{\bar{\alpha}} \delta q^{\bar{\alpha}}]_{t_i}^{t_f} + c_n [\delta q^n]_{t_i}^{t_f}, \quad (10.4)$$

Since the last term can be written as $\delta \int_{t_i}^{t_f} c_n \dot{q}^1$, this suggest that the appropriate reduced action principle, that gives rise to the reduced system of equations for arbitrary variations $\delta q^{\bar{\alpha}}$ that vanish at the end points is

$$\boxed{\bar{L}(q^{\bar{\alpha}}, \dot{q}^{\bar{\alpha}}, t) = (L - p_n \dot{q}^n)|_{\dot{q}^n = U^n} = L|_{\dot{q}^n = U^n} - c_n U^n}. \quad (10.5)$$

It is applicable for trajectories that have the same constant p_n than the natural trajectory. Indeed, the reduced system of equations is

$$\left[\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^{\bar{\alpha}}} \right] \Big|_{\dot{q}^n = U^n, \ddot{q}^n = \frac{d}{dt} U^n} = \frac{\partial L}{\partial q^{\bar{\alpha}}} \Big|_{\dot{q}^n = U^n}, \quad \frac{\partial L}{\partial \dot{q}^n} = c_n, \quad (10.6)$$

while the Euler-Lagrange equations for the reduced Lagrangian are

$$-\frac{\delta \bar{L}}{\delta q^{\bar{\alpha}}} = \frac{d}{dt} \left[\frac{\partial L}{\partial \dot{q}^{\bar{\alpha}}} \Big|_{\dot{q}^n = U^n} + \frac{\partial L}{\partial \dot{q}^n} \Big|_{\dot{q}^n = U^n} \frac{\partial U^n}{\partial \dot{q}^{\bar{\alpha}}} - c_n \frac{\partial U^n}{\partial \dot{q}^{\bar{\alpha}}} \right] - \frac{\partial L}{\partial q^{\bar{\alpha}}} \Big|_{\dot{q}^n = U^n} - \frac{\partial L}{\partial \dot{q}^n} \Big|_{\dot{q}^n = U^n} \frac{\partial U^n}{\partial q^{\bar{\alpha}}} + c_n \frac{\partial U^n}{\partial q^{\bar{\alpha}}} = 0, \quad (10.7)$$

which indeed agree when using equation (10.2).

Finally, note that one may also arrive to the reduced action principle by using a Lagrange multiplier,

$$L^\lambda(q^\alpha, \dot{q}^\alpha, \dot{q}^n, t, \lambda) = L - p_n \dot{q}^n - \lambda(p_n - c_n). \quad (10.8)$$

10.2 Parametrized systems

The solution of the Euler-Lagrange equations of motion may be given in parametrized form. Rather than using $q^\alpha = q^\alpha(t)$, one introduces a parameter τ and describes the solution in parametrized form through $t = t(\tau)$, $q^\alpha = q^\alpha(\tau)$. If $t' = \frac{dt}{d\tau} \neq 0$, $q'^\alpha = \frac{dq^\alpha}{d\tau}$, the appropriate variation principle involves as independent variable τ and contains $n + 1$ Lagrange coordinates q^α, t . It is given by

$$S = \int_{\tau_i}^{\tau_f} d\tau L^P, \quad L^P(q^\alpha, t, q'^\alpha, t') = t' L(q^\alpha, \dot{q}^\alpha, t) \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}}. \quad (10.9)$$

Indeed,

$$\begin{aligned} -\frac{\delta L^P}{\delta q^\alpha} &= \frac{d}{d\tau} \left[t' \frac{\partial L}{\partial \dot{q}^\alpha} \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}} \right] - t' \frac{\partial L}{\partial q^\alpha} \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}}, \\ -\frac{\delta L^P}{\delta t} &= \frac{d}{d\tau} \left(L \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}} - \frac{\partial L}{\partial \dot{q}^\alpha} \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}} \frac{q'^\alpha}{t'} \right) - t' \frac{\partial L}{\partial t} \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}}. \end{aligned} \quad (10.10)$$

When using that

$$\frac{d^2}{dt^2} q^\alpha = \frac{d}{dt} \left(\frac{q'^\alpha}{t'} \right) = \frac{1}{t'} \frac{d}{d\tau} \left(\frac{q'^\alpha}{t'} \right) = \frac{q''^\alpha}{(t')^2} - \frac{q'^\alpha t''}{(t')^3}, \quad (10.11)$$

the first n Euler-Lagrange equations associated to (10.10) become

$$t' \left[\frac{\delta L}{\delta q^\alpha} \right] \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}, \ddot{q}^\alpha = \frac{q''^\alpha}{(t')^2} - \frac{q'^\alpha t''}{(t')^3}} = 0, \quad (10.12)$$

and are thus equivalent to the original Euler-Lagrange equations of motion. This is the same reasoning as in the previous discussion concerning a change of the independent variable.

When working out the derivative with respect to τ and using the previous discussion, one finds for the last expression of (10.10),

$$-\frac{\delta L^P}{\delta t} = -q'^\alpha \left[\frac{\delta L}{\delta q^\alpha} \right] \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}, \ddot{q}^\alpha = \frac{q''^\alpha}{(t')^2} - \frac{q'^\alpha t''}{(t')^3}} \quad (10.13)$$

which vanishes as a consequence of the previous n equations.

Note that L^P does not define a regular system. Indeed, when denoting the $n + 1$ Lagrange coordinates (t, q^α) by q^A , $L^P(q^A, q'^A)$ is homogeneous of degree 1 in the generalized velocities q'^A ,

$$L^P(q^A, \mu q'^A) = \mu L^P(q^A, q'^A), \quad (10.14)$$

which implies, by differentiation with respect to μ and setting $\mu = 0$ (Euler's theorem for homogeneous functions) that

$$q'^B \frac{\partial L^P}{\partial q'^B} = L^P. \quad (10.15)$$

Another derivation with respect to q'^A yields

$$\frac{\partial^2 L^P}{\partial q'^A \partial q'^B} q'^B + \frac{\partial L^P}{\partial q'^A} = \frac{\partial L^P}{\partial q'^A}, \quad (10.16)$$

which means that q'^B is an eigenvector with zero eigenvalue of the matrix $\frac{\partial^2 L^P}{\partial q'^A \partial q'^B}$, so that this matrix is degenerate.

10.3 Maupertius'/ Jacobi's principle

If L does not depend explicitly on time t , t is a cyclic variable of the variational principle defined by L^P . The associated conserved momentum, p_t is minus the energy of the original variational principle defined by L ,

$$p_t = \frac{\partial L^P}{\partial t'} = (L - p_\alpha \dot{q}^\alpha) \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}} = -E \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}} = -\mathbf{E}, \quad (10.17)$$

for some constant $\mathbf{E}$. If the original Lagrangian was regular, one may solve this equation for $t' = T'(q^\alpha, q'^\alpha)$ provided that not all generalized velocities vanish. Indeed,

$$\frac{\partial^2 L^P}{(\partial t')^2} = \left(\cancel{\frac{\partial L}{\partial \dot{q}^\alpha}} - \cancel{\frac{\partial L}{\partial \dot{q}^\alpha}} - \dot{q}^\alpha \frac{\partial^2 L}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} \right) \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}} \left(-\frac{q'^\beta}{(t')^2} \right) = \frac{1}{t'} \left(\dot{q}^\alpha \frac{\partial^2 L}{\partial \dot{q}^\alpha \partial \dot{q}^\beta} \dot{q}^\beta \right) \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}}. \quad (10.18)$$

In this case, the variational principle of Maupertius/Jacobi is obtained by applying the Routhian procedure to L^P . It gives rise to the reduced variational principle, applicable for trajectories with the same energy $\mathbf{E}$ than the natural trajectory,

$$\bar{L}^P = T' L \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{T'}} + E \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{T'}} T' = p_\alpha \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{T'}} q'^\alpha. \quad (10.19)$$

when taking into account (10.17).

For an explicitly time-independent L of the form (9.13), one finds from (10.17) that

$$\frac{1}{2} g_{\alpha\beta} q'^\alpha q'^\beta (T')^{-2} + V = \mathbf{E} \implies T' = \pm \sqrt{\frac{g_{\alpha\beta} q'^\alpha q'^\beta}{2(\mathbf{E} - V)}}, \quad (10.20)$$

where the sign follows that of t' . This implies on the one hand that

$$t - t_i = \pm \int_{\tau_i}^{\tau} d\tau' \sqrt{\frac{g_{\alpha\beta} q'^\alpha q'^\beta}{2(\mathbf{E} - V)}} = \pm \int_{\gamma} \frac{\bar{ds}}{\sqrt{2(\mathbf{E} - V)}}, \quad (10.21)$$

where γ is a path from some initial configuration q_i^α to a final configuration q^α , and on the other that

$$\bar{L}^P = \frac{1}{T'} g_{\alpha\beta} q'^\alpha q'^\beta + A_\alpha q'^\alpha = \pm \sigma \sqrt{2(\mathbf{E} - V)} \sqrt{|g_{\alpha\beta} q'^\alpha q'^\beta|} + A_\alpha q'^\alpha, \quad (10.22)$$

where σ denotes the sign of $g_{\alpha\beta} q'^\alpha q'^\beta$ and

$$\boxed{\bar{S}^P = \int \left[\pm \sigma \sqrt{2(\mathbf{E} - V)} \bar{ds} + A_\alpha dq^\alpha \right]}. \quad (10.23)$$

The Lagrange multiplier version of Maupertius'/Jacobi's variational principle is

$$S^{P,\lambda} = \int d\tau \left[p_\alpha \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}} q'^\alpha + \lambda [E \Big|_{\dot{q}^\alpha = \frac{q'^\alpha}{t'}} - \mathbf{E}] \right]. \quad (10.24)$$

Remarks:

(i) In particular, if $V = 0 = A_\alpha$, the variational principle of Maupertius/Jacobi for $S = \int dt \frac{1}{2} g_{\alpha\beta} \dot{q}^\alpha \dot{q}^\beta$ reduces to the reparametrization invariant variational principle associated to $\bar{S}^P = \pm \sigma \sqrt{2|\mathbf{E}|} \int \bar{ds}$.

(ii) In this context, a systematic derivation of the solution to the brachistochrone 5.2.1 starting from Hamilton's principle is possible. The appropriate action is

$$S[x, z] = \int_{t_*}^{t_0} dt \left[\frac{1}{2} m (\dot{x}^2 + \dot{z}^2) - mgz \right], \quad (10.25)$$

with initial and final conditions $x(t_*) = x^*, z(t_*) = z^*, x(t_0) = 0, z(t_0) = 0$. The objective is to determine the spatial path $x(z)$ such that the time of descent is minimal. In order to use conservation of energy, the use of Maupertius'/Jacobi's principle is appropriate. The initial conditions imply that $E = mgz^*$, so that equation (10.21) yields for the time of descent between (x^*, z^*) and $(0, 0)$

$$t_0 - t_* = - \int_{z_*}^0 dz \sqrt{\frac{1 + (x')^2}{2g(z^* - z)}}, \quad x' = \frac{dx}{dz}, \quad (10.26)$$

in agreement with equation (5.48).

The remaining equations are the Euler-Lagrange equations associated to

$$\bar{S}^P = -m \int dz \sqrt{2g(z^* - z)} \sqrt{1 + (x')^2}. \quad (10.27)$$

Since x is a cyclic variable,

$$p_x = \sqrt{2g(z^* - z)} \frac{x'}{\sqrt{1 + (x')^2}} = d, \quad (10.28)$$

for some constant $d \geq 0$. The sign is fixed by requiring that x increases when z increases. Taking the square yields

$$(x')^2 = \frac{d^2}{2g(z^* - z) - d^2} = \frac{2R'}{z^* - z - 2R'}, \quad 2R' = \frac{d^2}{2g} \quad (10.29)$$

Through integration, one then finds

$$x - x^* = \int_{z^*}^z dz' \frac{\sqrt{2R'}}{\sqrt{z^* - z - 2R'}} = - \int_0^{z-z^*} dZ \frac{\sqrt{2R'}}{\sqrt{Z - 2R'}} \quad (10.30)$$

Setting $Z = R'(1 - \cos \theta) = 2R' \sin^2 \frac{\theta}{2}$.

solve this to find $x = x(z; x^*, z^*, c')$, and then compute the time of fall on-shell, then extremize it

Chapter 11

Dynamics in non-inertial frames

Chapter 12

Relativistic particles

When one considers relativistic particles rather than non-relativistic ones, the action

$$S_P = \int dt \sum_{a=1}^N \frac{1}{2} m_{(a)} \dot{x}_a^i \dot{x}_{ai}, \quad (12.1)$$

for N free non-relativistic particles must be replaced by

$$S_P = - \sum_{a=1}^N m_{(a)} c \int d\tau \sqrt{-\frac{dx_a^\mu}{d\tau} \frac{dx_{a\mu}}{d\tau}}. \quad (12.2)$$

This action reduces to the non-relativistic one in the limit of small velocities as compared to the speed of light. Indeed, if we choose $\tau = x^0 = ct$,

$$\begin{aligned} & - m_{(a)} c \int d\tau \sqrt{-\frac{dx_a^\mu}{d\tau} \frac{dx_{a\mu}}{d\tau}} \\ &= - m_{(a)} c^2 \int dt \sqrt{1 - \frac{1}{c^2} \frac{d\vec{x}_a}{dt} \cdot \frac{d\vec{x}_a}{dt}} \approx - m_{(a)} c^2 \int dt \left(1 - \frac{1}{2} \frac{1}{c^2} \frac{d\vec{x}_a}{dt} \cdot \frac{d\vec{x}_a}{dt} \right). \end{aligned} \quad (12.3)$$

The first term is a constant which does not contribute to the Euler-Lagrange equations and thus can be dropped, while the second term reduces to the action for a collection of non-relativistic particles, (12.1).

Let us now set $c = 1$ and study in more details a single relativistic particle, whose worldline is described by $x^\mu(\tau)$ in Minkowski spacetime. The associated action is

$$S_P[x^\mu] = \int d\tau L_P, \quad L_P = -m \sqrt{-\dot{x}^\nu \dot{x}_\nu}, \quad (12.4)$$

with $\dot{x}^\nu = \frac{dx^\nu}{d\tau}$. For variations that vanish at the boundary of the τ interval, the variation of the Lagrangian is

$$\delta S_P = -m \int d\tau \frac{-\delta \dot{x}^\mu \dot{x}_\mu}{\sqrt{-\dot{x}^\nu \dot{x}_\nu}} = -m \int d\tau \delta x^\mu \frac{d}{d\tau} \left(\frac{\dot{x}_\mu}{\sqrt{-\dot{x}^\nu \dot{x}_\nu}} \right), \quad (12.5)$$

and the Euler-Lagrange equations are given by

$$\frac{\delta L_P}{\delta x^\mu} = -m \frac{d}{d\tau} \left(\frac{\dot{x}_\mu}{\sqrt{-\dot{x}^\nu \dot{x}_\nu}} \right) = 0. \quad (12.6)$$

It is not possible to pass to the Hamiltonian formulation through a Legendre transformation: the Hessian matrix

$$\frac{\partial^2 L_P}{\partial \dot{x}^\mu \partial \dot{x}^\lambda} \quad (12.7)$$

is not invertible. Indeed,

$$\frac{\partial^2 L_P}{\partial \dot{x}^\mu \partial \dot{x}^\lambda} = -m \frac{-\dot{x}^\nu \dot{x}_\nu \eta_{\mu\lambda} + \dot{x}_\mu \dot{x}_\lambda}{(\sqrt{-\dot{x}^\nu \dot{x}_\nu})^3}, \quad \frac{\partial^2 L_P}{\partial \dot{x}^\mu \partial \dot{x}^\lambda} \dot{x}^\lambda = 0, \quad (12.8)$$

so that $\dot{x}^\lambda$ is an eigenvector of eigenvalue 0.

The correct first order action principle whose equation of motion are equivalent to the Lagrangian ones in (12.7) is given by

$$S[x^\mu, p_\mu, \lambda] = \int d\tau [\dot{x}^\mu p_\mu - \lambda(p^\mu p_\mu + m^2)]. \quad (12.9)$$

The associated Euler-Lagrange equations with respect to x^μ, p_μ, λ are

$$\dot{p}_\mu = 0, \quad \dot{x}^\mu - 2\lambda p^\mu = 0, \quad p^\mu p_\mu + m^2 = 0. \quad (12.10)$$

It is a “constrained Hamiltonian system” in the sense that $\lambda(\tau)$ is a Lagrange multiplier for the constraint $p^\mu p_\mu + m^2$ that does not involve time derivatives of the phase space variables x^μ, p_μ .

The general solution to the equations of motion is

$$p_\mu(\tau) = p_\mu(0) \equiv p_\mu, \quad x^\mu(\tau) = x^\mu(0) + 2p^\mu(0) \int_0^\tau d\tau' \lambda(\tau'), \quad p^0 = \pm \sqrt{p_i p^i + m^2}. \quad (12.11)$$

where p_μ are τ -independent constants. Alternatively, if $\lambda > 0$, the equations for p_μ may be solved as

$$p_\mu = -\frac{1}{2\lambda} \dot{x}_\mu. \quad (12.12)$$

When substituting in the equation for λ , one finds

$$\frac{\dot{x}^\mu \dot{x}_\mu}{4\lambda^2} + m^2 = 0 \iff \lambda = \frac{\sqrt{-\dot{x}^\nu \dot{x}_\nu}}{2m}. \quad (12.13)$$

When substituting both of these equations into the equation for x^μ , i.e., the first equation in (12.10), one recovers the Lagrangian equations (12.6). It is in this sense that the Lagrangian equations (12.6) and the first order equations (12.10) are equivalent. Finally, note also that the first order action (12.9) reduces to the Lagrangian action (12.4) when substituting both (12.12) and (12.13).

The x^μ can be considered as coordinates on Minkowski spacetime with metric $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ that is used to lower indices. Its inverse is denoted by $\eta^{\mu\nu} = \text{diag}(-1, 1, 1, 1)$, $\eta_{\mu\nu} \eta^{\nu\rho} = \delta_\mu^\rho$ and is used to raise indices.

Poincaré transformations are defined by

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu + a^\mu, \quad (12.14)$$

where the matrix $\Lambda^\mu{}_\nu$ satisfies

$$\Lambda^\mu{}_\nu \eta_{\mu\sigma} \Lambda^\sigma{}_\rho = \eta_{\nu\rho} \iff \Lambda^T \eta \Lambda = \eta \iff \Lambda_\sigma{}^\nu \Lambda^\sigma{}_\rho = \delta_\nu^\rho \iff (\Lambda^\sigma{}_\rho)^{-1} = \Lambda_\sigma{}^\rho, \quad (12.15)$$

The transformations with $\Lambda^\mu{}_\nu = 0$, $x'^\mu = x^\mu + a^\mu$ are called translations, those with $a^\mu = 0$, $x'^\mu = \Lambda^\mu{}_\nu x^\nu$ Lorentz transformations.

For Poincaré transformations (that do not depend on τ), the Lagrangian L_P and thus the action S_P is invariant,

$$S_P[x'^\mu] = S_P[x^\mu]. \quad (12.16)$$

As in the discussion of Galilean invariance, infinitesimal Lorentz transformations are of the form

$$\Lambda^\mu{}_\nu = \delta_\nu^\mu + \omega^\mu{}_\nu + O(\omega^2), \quad \omega_{\mu\nu} = -\omega_{\nu\mu}, \quad (12.17)$$

and contain 6 parameters, while infinitesimal translations are characterized by $a^\mu = 0 + \alpha^\mu$ and contain 4 parameters. The associated infinitesimal Poincaré transformations on Minkowski spacetime are given by

$$\delta_{\omega,\alpha} x^\mu = \omega^\mu{}_\nu x^\nu + \alpha^\mu. \quad (12.18)$$

According to Noether's theorem, there are 10 conserved quantities given by

$$K_{\omega,\alpha} = \frac{\partial L_P}{\partial \dot{x}^\mu} \delta_{\omega,\alpha} x^\mu = m \frac{\dot{x}_\mu \delta_{\omega,\alpha} x^\mu}{\sqrt{-\dot{x}^\nu \dot{x}_\nu}}. \quad (12.19)$$

The individual conserved quantities are denoted by $P^\mu, L^{\mu\nu}$,

$$K_{\omega,\alpha} = \frac{1}{2} \omega_{\mu\nu} L^{\mu\nu} + \alpha_\mu P^\mu, \quad (12.20)$$

and are explicitly given by

$$P^\mu = m \frac{\dot{x}^\mu}{\sqrt{-\dot{x}^\sigma \dot{x}_\sigma}}, \quad L^{\mu\nu} = m \frac{\dot{x}^\mu x^\nu - \dot{x}^\nu x^\mu}{\sqrt{-\dot{x}^\sigma \dot{x}_\sigma}}. \quad (12.21)$$

For several particles, one finds instead,

$$P^\mu = \sum_{a=1}^N m_{(a)} \frac{\dot{x}_a^\mu}{\sqrt{-\dot{x}_a^\sigma \dot{x}_{\sigma,a}}}, \quad L^{\mu\nu} = \sum_{a=1}^N m_{(a)} \frac{\dot{x}_a^\mu x_a^\nu - \dot{x}_a^\nu x_a^\mu}{\sqrt{-\dot{x}_a^\sigma \dot{x}_{\sigma,a}}}. \quad (12.22)$$

relativistic invariance, to be continued

Chapter 13

Charged particle in electromagnetic field

A prerequisite for this chapter is the Lagrangian formulation of Maxwell's equation using gauge potentials.

13.1 Newtonian formulation

Empirical considerations have shown that the force that acts on a charged point particle under the influence of an electromagnetic field is given by the Lorentz force,

$$\vec{F} = q(\vec{E} + \dot{\vec{x}} \times \vec{B}). \quad (13.1)$$

For a charged particle that moves at a non relativistic velocity $v \ll c$, Newton's second law thus implies

$$m\ddot{\vec{x}} = q(\vec{E} + \dot{\vec{x}} \times \vec{B}). \quad (13.2)$$

13.2 Lagrangian formulation

In order to derive the equations of motion (13.2) from a variational principle, one needs to describe the electromagnetic fields in terms of potentials, $\phi(x^i, t)$, $\vec{A}(x^i, t)$.

$$\vec{B} = \vec{\nabla} \times \vec{A}, \quad \vec{E} = -\partial_t \vec{A} - \vec{\nabla} \phi, \quad (13.3)$$

the appropriate Lagrangian is

$$L = \frac{1}{2}m\dot{\vec{x}}^2 - q\phi + q\dot{\vec{x}} \cdot \vec{A}. \quad (13.4)$$

Indeed, the associated Euler-Lagrange derivatives are

$$\begin{aligned} -\frac{\delta L}{\delta x^i} &= \frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} - \frac{\partial L}{\partial x^i} = \frac{d}{dt}(m\dot{x}_i + qA_i) + q\left(\frac{\partial \phi}{\partial x^i} - \dot{x}^j \frac{\partial A_j}{\partial x^i}\right) \\ &= m\ddot{x}_i + q\left(\frac{\partial A_i}{\partial x^j} \dot{x}^j + \frac{\partial A_i}{\partial t}\right) + q\left(\frac{\partial \phi}{\partial x^i} - \dot{x}^j \frac{\partial A_j}{\partial x^i}\right). \end{aligned} \quad (13.5)$$

Re-organizing, the associated equations of motion are

$$m\ddot{x}_i = q(\partial_i A_j - \partial_j A_i) \dot{x}^j - q(\partial_i \phi + \partial_t A_i) = q\epsilon_{ijk} \epsilon^{klm} \partial_l A_m \dot{x}^j - q(\partial_i \phi + \partial_t A_i), \quad (13.6)$$

which coincide with (13.2).

13.3 Hamiltonian formulation

The canonical momenta are given by

$$p_i = \frac{\partial L}{\partial \dot{x}^i} = m\dot{x}_i + qA_i \iff \dot{x}^i = \frac{1}{m}(p^i - qA^i), \quad (13.7)$$

while the Hamiltonian is

$$H = [p_i \dot{x}^i - L]_{\dot{x}=\dot{x}(x,p)} = \frac{1}{m}(p^i - qA^i)p_i - \frac{1}{2m}(p^i - qA^i)(p_i - qA_i) + q\phi - q\frac{1}{m}(p^i - qA^i)A_i. \quad (13.8)$$

After regrouping terms,

$$H = \frac{1}{2m}(p^i - qA^i)(p_i - qA_i) + q\phi. \quad (13.9)$$

In other words, as compared to the Hamiltonian of a free particle, $H = \frac{1}{2m}p^i p_i$, the Hamiltonian of a charged particle in an electromagnetic field is obtained by adding the potential $V = q\phi$ and the so-called rule of “minimal substitution”, $p^i \rightarrow p^i - qA^i$.

Remarks:

(i) In the case of N particles with coordinates x_a^i , mass m_a and charge q_a , the Lagrangian becomes

$$L = \sum_{a=1}^N \left[\frac{1}{2}m_{(a)}\dot{x}_a^i \dot{x}_{ai} - q_{(a)}\phi(\vec{x}_a, t) + q_{(a)}\dot{x}_a^i A_i(\vec{x}_a, t) \right], \quad (13.10)$$

while the Hamiltonian is given by

$$H = \sum_{a=1}^N \left[\frac{1}{2m_{(a)}} [p_a^i - q_{(a)}A^i(\vec{x}_a, t)][p_{ai} - q_{(a)}A_i(\vec{x}_a, t)] + q_{(a)}\phi(\vec{x}_a, t) \right]. \quad (13.11)$$

One may rewrite the action associated to the Lagrangian (13.10) as $S = S_P + S_I$, where the action for N free non-relativistic particles is given in (12.1), while

$$S_I = \sum_{a=1}^N \left[-q_{(a)}\phi(\vec{x}_a, t) + q_{(a)}\dot{x}_a^i A_i(\vec{x}_a, t) \right], \quad (13.12)$$

which may be re-written as

$$S_I = \int d^4x A_\mu(x) J^\mu(x), \quad J^0 = \sum_{a=1}^N q_{(a)}\delta^{(3)}(\vec{x} - \vec{x}_a(t)), \quad J^i = \sum_{a=1}^N \frac{q_{(a)}}{c} \frac{dx_a^i}{dt} \delta^{(3)}(\vec{x} - \vec{x}_a(t)), \quad (13.13)$$

is the action that describes the interaction of the particles with the electromagnetic field.

(ii) In order to distinguish more clearly the coordinates on which the electromagnetic fields and potentials depend from the position of the particles, it is often useful to change notation and describe the latter by z_a^i rather than x_a^i .

(iii) In relativistic notation, the current may be written as

$$J^\mu = \sum_{a=1}^N \int d\tau \frac{q_{(a)}}{c} \frac{dx_a^\mu}{d\tau} \delta^{(4)}(x^\nu - x_a^\nu(\tau)), \quad (13.14)$$

That this expression reduces to that in (13.13) can be seen as follows:

$$\int d\tau \frac{q_{(a)}}{c} \frac{dx_a^\mu}{d\tau} \delta^{(4)}(x^\nu - x_a^\nu(\tau)) = \dots = \frac{q_{(a)}}{c} \frac{dx_a^\mu}{d\tau} \delta^{(3)}(\vec{x} - \vec{x}_a) \quad (13.15)$$

where, in order to show the last equality, one uses the following property of the δ function,

$$\int dx f(x) \delta(g(x)) = \frac{f(x_0)}{|g'(x_0)|}, \quad (13.16)$$

where $g(x)$ is a monotonic function that vanishes in x_0 , $g(x_0) = 0$.

In this case, one needs to replace the action for the particles S_P by that appropriate for relativistic ones.

complete by using Felsager exercise 1.4.1 pages 20,28 and equation 1.34 in reverse.

Chapter 14

To be added

geodesic motion on the surface of the sphere as introduction/application to the chapter “A Lagrangian of sufficient generality”

mostly track 2 chapters

- Hamilton-Jacobi theory
- normal modes
- full proof of Frobenius theorem (inverse)
- Lagrangian approach to two body problem
- reduction of order and reduced variational principle “Routhian”, Lagrangian and Hamiltonian version
- Maupertius, time-reparametrization invariance, apply to go from Lagrangian formulation of brachystochrone problem, to one where the action becomes the time of fall
- more details on Lagrangian field theories and Poincaré invariance

Bibliography

- [1] G. Mayné. “Cours de mécanique analytique. 2ème candidature Sciences ULB. Mathématique et Physique”. unpublished. 1988.
- [2] P. Gaspard. *ULB MATHF204: Mécanique analytique, Notes de cours*. ULB, 2014.
- [3] P. Spindel. *Mécanique. Volume 1: Mécanique newtonienne*. Gordon and Breach, 2004.
- [4] P. Spindel. *Mécanique. Volume 2: Mécanique analytique*. Gordon and Breach, 2002.
- [5] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics*. Addison Wesley, 2002. ISBN: 9780201657029. URL: <https://books.google.lu/books?id=tJCuQgAACAAJ>.
- [6] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Second. Graduate Texts in Mathematics. Springer, 1989.
- [7] E. Sudarshan and N. Mukunda. *Classical Dynamics: A Modern Perspective*. John Wiley & Sons, 1974.
- [8] C. Lanczos. *The Variational Principles of Mechanics*. Dover Books On Physics. Dover Publications, 1986. ISBN: 9780486650678. URL: <https://books.google.lu/books?id=ZW0YYr8wk2IC>.
- [9] H. B. Callen. *Thermodynamics and an introduction to thermostatistics*. Second. John Wiley & Sons, 1985.
- [10] B. Schutz. *Geometrical methods of mathematical physics*. Cambridge University Press, 1980.
- [11] M. Henneaux and C. Teitelboim. *Quantization of Gauge Systems*. Princeton University Press, 1992.
- [12] C. Bunster. “The Principle of Least Action, from the "Vis-viva" to Quantum Black Hole”. Sixteenth Arnold Sommerfeld Lecture Series, available online at <https://www.theorie.physik.uni-muenchen.de/activities> 2014. URL: https://www.theorie.physik.uni-muenchen.de/activities/lectures/sixteenth_series/bunster_public/video_bunster/index.html.