Dynamical systems with non-integrable constraints, vakonomic mechanics, sub-Riemannian geometry, and non-holonomic mechanics

A. V. Borisov, I. S. Mamaev, and I. A. Bizyaev

Abstract. This is a survey of the main forms of equations of dynamical systems with non-integrable constraints, divided into two large groups. The first group contains systems arising in vakonomic mechanics and optimal control theory, with the equations of motion obtained from the variational principle, and the second contains systems in classical non-holonomic mechanics, when the constraints are ideal and therefore the D’Alembert–Lagrange principle holds.

Bibliography: 134 titles.

Keywords: non-integrable constraints, vakonomic mechanics, optimal control theory, sub-Riemannian geometry, non-holonomic mechanics, invariant measure.

Contents

1. Introduction 784
2. Hamilton’s principle for systems with non-integrable constraints 790
  2.1. Equations of conditional extremals 790
  2.2. A Lagrangian for unconditional extremals 792
  2.3. An example: the Bottema–Hamel system 793
  2.4. A parametric representation 794
  2.5. The maximum principle 796
  2.6. The law of conservation of energy 798
  2.7. An example: the Bottema–Hamel system (continued) 798
1. Introduction

First of all, we recall some characteristic features of a mechanical system with holonomic constraints. Let $\mathbf{q} = (q^1, \ldots, q^n)$ be generalized coordinates in a configuration space $\mathcal{N}$ and suppose that certain constraints defined in this space impose restrictions on the coordinates. We say that these constraints are holonomic (and the system is also holonomic) and represent them in the form

$$f^\mu(\mathbf{q}) = 0, \quad \mu = 1, \ldots, k, \quad k < n. \tag{1}$$

Apart from the constraints, the system is also characterized by a Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}})$, which is determined by the kinetic and potential energy of the system and is a polynomial of degree at most two in the velocities $\dot{\mathbf{q}}$.

In deriving the equations of motion of a holonomic system, the authors mostly use the following two different axiomatic principles.

The d’Alembert–Lagrange principle, which in this case can be represented in the form

$$\left[ \left( \frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} \right] \delta q^i = Q_i \delta q^i, \quad \frac{\partial f^\mu}{\partial q^i} \delta q^i = 0, \quad \mu = 1, \ldots, k,$$

where $\mathbf{Q} = (Q_1, \ldots, Q_n)$ denotes the generalized forces. According to this principle, the work of the reaction forces along virtual (possible) displacements $\delta \mathbf{q} = (\delta q^1, \ldots, \delta q^n)$ satisfying the constraints is zero.
Hamilton’s variational principle, which says that the trajectories $q(t)$ of the system are extremals of the action functional

$$A = \int_{t_1}^{t_2} L(q(t), \dot{q}(t)) \, dt$$

in the class of curves satisfying the constraints (1).

As is well known, in this case these principles produce the same equations of motion. Here are the main properties of the equations:

- the equations of motion can be represented in a (canonical) Hamiltonian form, and thus preserve the phase volume (that is, possess the standard invariant measure) by Liouville’s theorem;
  
- the classical principle of determinacy holds, namely, there exists a unique trajectory corresponding to a prescribed initial position and a prescribed velocity of the system.

In applications the Lagrangian is usually reduced (possibly using the Maupertuis principle) to a homogeneous quadratic form in the velocities, which is in addition non-degenerate and positive definite:

$$L = \frac{1}{2} g_{ij}(q) \dot{q}^i \dot{q}^j.$$ 

In this case trajectories of the system are geodesics of a certain Riemannian metric $g_{ij}$. Recall that by definition a geodesic is a trajectory of minimum length that connects two points in the configuration space $\mathcal{N}$, that is, a ‘shortest’ curve. A further property holds in this case:

- a geodesic (shortest) trajectory is also a ‘straightest’ curve, that is, a curve whose geodesic curvature vanishes.

2. It is known that quite a few problems in various areas of mechanics and mathematics reduce to an analysis of systems with constraints represented in the differential form

$$f^\mu(q, \dot{q}) = \alpha^i_\mu(q) \dot{q}^i = 0, \quad \mu = 1, \ldots, k, \quad k < n.$$ 

The study of a possible reduction of such constraints to the form (1) reduces to studying the integrability of systems of Pfaffian equations,\footnote{In fact, after multiplication by $dt$ the constraint equations (2) take the standard form of Pfaffian equations.} and a criterion for reducibility can be found using the Frobenius theorem.

We call constraints (2) that cannot be reduced to the form (1) non-integrable (or non-holonomic). It turns out to be a special feature of such constraints that upon applying the d’Alembert–Lagrange principle (see §4) and Hamilton’s variational principle (see §2) to the same system, one obtains different equations of motion. This was first noted by Hertz and Poincaré and then was more thoroughly investigated by Hamel [76], [78].

Remark 1. Here we confine ourselves to constraints (2) which are linear and homogeneous in the velocities, though many of the results below can be directly generalized to affine (inhomogeneous) constraints [52].
3. The choice of a dynamical principle underlying the derivation of the equations of motion is determined by the method of realization of the constraints described by the formal relations (1) or (2). Here by realization we mean that some dynamical factors present in the unconstrained system result in constraints in the limit. A realization of non-holonomic constraints can involve either potential forces [116] or also non-potential forces [101]. To make this general approach clearer we divide systems with constraints (2) into two groups and consider them separately in what follows.

The first group contains systems whose equations of motion are obtained from the variational principle. These systems were considered by Kozlov [88]–[90] in the framework of vakonomic mechanics (derived from ‘variational axiomatic kind’), where the constraints (2) are realized by passing to a limit in the kinetic energy of the system as the inertial characteristics (mass, moments of inertia, and so on) tend to infinity. Such an anisotropy of the inertial properties of the system arises in the motion of a plate in an ideal fluid, due to added masses (see §3.6).

In this case the equations of motion are Hamiltonian and therefore (by Liouville’s theorem) preserve the phase volume. On the other hand, they have a number of special properties, in contrast to holonomic systems. Let us look more closely at such systems.

- They do not satisfy the principle of determinacy: the same initial position and initial (admissible) velocity correspond to a whole family of trajectories. As a result, the equations of motion involve ‘unobservable’ variables, in number equal to the number of constraints. This indeterminacy, foreign to classical mechanics, spurred a discussion, which is available in [93] and [84]. In vakonomic mechanics the initial values of the ‘unobservable’ variables are interpreted as deviations of the initial velocity in the free system (existing before passage to the limit) from the relations prescribed by the constraints [93].

- For the systems with non-integrable constraints under consideration the Hamiltonian is degenerate in the momenta. If it can be reduced to a homogeneous form which is quadratic in the momenta, then we obtain the problem of geodesics with a degenerate metric, which is considered in sub-Riemannian geometry [108].

- Trajectories of a system with non-integrable constraints that are obtained from the variational principle are not the ‘straightest’ curves. This was discussed in [2], Chap. 1, § 4, using as an example a non-integrable constraint defined in \( \mathbb{R}^3 \).

The second group contains systems in classical non-holonomic mechanics, when the constraints (2) are ideal (so that the corresponding reaction forces perform no work) and therefore the d’Alembert–Lagrange principle holds. The best known example here is the problem of a rigid body rolling on the plane, when non-integrable constraints express the absence of slipping at the point of contact between the body and the plane. A comprehensive historical survey of the development of non-holonomic mechanics can be found in [53].

The original motivation to study the problem of realizing the constraints (2) in mechanics came from Painlevé’s paradoxes [112]. It is a classical result going back to Carathéodory [60] (and proved in [81], [91], [110]) that non-integrable constraints appear as the viscous friction coefficient goes to infinity. Investigations of the passage to the limit and of the asymptotic formulae approximating the non-holonomic
motion on a finite interval of time are still being continued (for instance, see [79] and [87]).

A more general passage to the limit, when both the viscous friction coefficient and the parameters of the system in the kinetic energy tend to infinity, was considered in [93]. We also mention systems with non-integrable servoconstraints (introduced by Béghin [5]), which were discussed in [96] and [97]. In this case the constraints (2) are realized in an ‘active’ way, using adjustable generalized forces (arising via control, for instance) or by means of suitable changes in the inertial properties of the system.

In contrast to vaconomic mechanics, the principle of determinacy holds for systems in non-holonomic mechanics. Furthermore, the equations of motion preserve the energy integral (only when the constraints are homogeneous [52]) and therefore are also said to be conservative, which is not quite correct since these systems do not preserve the phase volume (that is, do not have a continuous invariant measure) in general [92]. This absence of an invariant measure is responsible for many non-trivial dynamical phenomena, such as the reversal of a rattleback (a Celtic stone).

4. In the general case the equations of motion in non-holonomic mechanics cannot be represented in Hamiltonian form. Moreover, the absence of a smooth invariant measure in the general case leads to phenomena typical of dissipative systems occurring in non-holonomic systems [74], [40] (for instance, there can be strange attractors in the phase space).

On the other hand, the case when a continuous invariant measure exists is also possible under certain restrictions on the parameters of a non-holonomic system. Moreover, in some examples it turns out that the equations of motion can be represented in Hamiltonian form, but only after rescaling time. Such systems are said to be conformally Hamiltonian. Consequently, methods in Hamiltonian mechanics developed for integrability theory, stability, topological analysis, perturbation theory (KAM-theory), and so on, can be used. A search for a conformally Hamiltonian representation leads to the problem of Hamiltonization and of describing various (topological) obstructions to Hamiltonization [20]. Various situations encountered in non-holonomic mechanics are schematically presented in Fig. 1.

One can say that the historical problem of Hamiltonization of non-holonomic systems goes back to [66] and [67], where it was conjectured that the equations of motion of a balanced but dynamically asymmetric ball (the so-called Chaplygin ball) on a plane cannot be reduced to a conformally Hamiltonian form. Moreover, this was even proved in [67] using geometric formalism. Nevertheless, this system was ‘Hamiltonized’ in [37] using an explicit Poisson structure, which turned out to be non-linear. (Note that we speak here about ‘Hamiltonizing’ the reduced system, whereas even for a homogeneous ball it is not always possible to reduce the full system of equations to Hamiltonian form [35]. Problems of Hamiltonization were discussed in [17] in relation to systems of hydrodynamical type.

The method of reducing multipliers, considered originally by Chaplygin and subsequently developed further in [22] and [45], proved to be very efficacious in reducing equations of non-holonomic mechanics to a conformally Hamiltonian form.

Systems arising in non-holonomic mechanics are intermediate between Hamiltonian and dissipative systems. In [39] and [54] this variety of types of behaviour
was called a hierarchy of dynamics, and it is brought about by the existence (or absence) of various tensor invariants (conservation laws), which significantly affects the dynamics of the system.

5. Thus, systems with constraints (2) whose trajectories are found using the d’Alembert–Lagrange principle do not deliver extremals of some functional. For all this, there is some confusion in the literature, going back to Griffith [75], who used the variational principle to describe the motion of non-holonomic systems. The first step in clearing up this question was made by Vershik and Gershkovich [127], and complete clarity was achieved by Kozlov in [93]. There he reviewed the main forms of equations of dynamical systems with non-integrable constraints developed in the cycle of papers [88]–[90]. Moreover, as illustrations of applications he presented several examples, some of which were new.

Geometric structures arising in non-holonomic mechanics have been investigated by Schouten, Vrancanu, Singh (see [118]), and others. We mention especially Wagner’s paper [124], where he introduced the notion of curvature for a non-holonomic manifold and gave examples when geometric constructions can be used to integrate the equations of motion. For instance, in the problem of motion for a Chaplygin sleigh\(^2\) it was shown that the corresponding manifold has curvature zero, and therefore one can introduce local coordinates in which the coefficients of the connection vanish, so that the equations of motion can be explicitly integrated. A geometric approach to non-holonomic systems was also considered in [126].

Non-holonomic mechanics provides many more examples than does vakonomic mechanics. For the latter we know of only a single interesting example: the problem of a plate in a fluid with a strong anisotropy of inertial properties. We consider this problem in §3.6. Nevertheless, variational problems with non-integrable constraints

---

\(^2\)Recall that a Chaplygin sleigh is a rigid body that moves with two perfectly smooth legs and a weightless sharp wheel (a disk or a knife edge) in contact with a supporting surface. The wheel prevents its point of contact from sliding in the direction orthogonal to the plane of the wheel.
arise in a natural way in optimal control theory. Let us now look more closely at these problems.

6. Solving the constraint equations (1) for the generalized velocities $\dot{q}$, we represent them in the parametric form

$$\dot{q}^i = b_{ji}(q)u^j, \quad j = 1, \ldots, n - k,$$

where $u = (u^1(t), \ldots, u^{n-k}(t))$ are unknown (control) functions.

The problem consists in choosing $u$ so as to move the system from a fixed initial position $(q_1, \dot{q}_1)$ to a terminal position $(q_2, \dot{q}_2)$ while minimizing a certain cost functional (efficiency index)

$$\mathcal{A} = \int_{t_1}^{t_2} L(q(t), \dot{q}(t), u(t)) \, dt.$$

In contrast to vakonomic mechanics, here $L$ is not necessarily the Lagrange function of the system, but is chosen based on the context of the problem, for instance, so as to minimize the length of the trajectory in the configuration space or the energy expended on the control (see [1], for example).

Using the Pontryagin maximum principle (see §2.5), we can obtain for the functions $u$ equations which are similar in form to equations in vakonomic mechanics. From the standpoint of solvability of boundary-value problems in control theory, ‘unobservable’ variables have a natural explanation: without this additional flexibility there would be no extremals with prescribed initial and terminal positions.

In this case the controllability test reduces to verifying the assumptions of the Rashevskii–Chow theorem (for instance, see [131] and [130]), which reduces in fact to verifying that the constraints (2) can be taken to holonomic form.

Among the best known examples of systems in optimal control theory with constraints (3) we mention the classical car parking problem (see [103], for instance): by adjusting the translational velocity and the turning angle of the front wheels, move a car from an initial position to another prescribed position so as to minimize the path length. In this case the non-integrable constraint is expressed by the condition that the projection of the car’s velocity on the direction orthogonal to the plane of a front wheel is zero.

In this survey we show that in optimal control problems non-integrable constraints arise not only due to kinematic restrictions as in the car parking problem; they can also be dictated by the common (usually, zero-) level set of first integrals that the system continues to have when a control is imposed. As an example, we consider the problem of applying a control to a body at rest in a fluid by means of rotors. In this case the non-integrable constraints (3) express the conservation of momentum and angular momentum of the system (see §3.4). A similar problem of controlling, by means of rotors, a Chaplygin ball rolling on a plane was recently considered in [19], [31], and [32], and we discuss this in §3.3 from a more general standpoint.

Remark 2. We remark that non-integrable constraints in the form (3) also arise in optics [6], [104], in describing a polarized light ray in a light guide. Understanding the nature of the corresponding equations of motion is an interesting problem, which has apparently not yet been solved.
2. Hamilton’s principle for systems with non-integrable constraints

2.1. Equations of conditional extremals. Consider a system whose configuration space is an \( n \)-dimensional manifold \( N \) locally parameterized by some generalized coordinates \( \mathbf{q} = (q^1, \ldots, q^n) \). We assume that the generalized velocities \( \dot{\mathbf{q}} = (\dot{q}^1, \ldots, \dot{q}^n) \) satisfy homogeneous linear equations, the constraints

\[
f^\mu(\mathbf{q}, \dot{\mathbf{q}}) = a_i^\mu(\mathbf{q}) \dot{q}^i = 0, \quad \mu = 1, \ldots, k, \tag{4}
\]

where we sum over repeated indices in this paper. We also assume that the constraints are non-integrable (non-holonomic), that is, cannot be represented as the time derivatives of some functions of coordinates.

For each point \( \mathbf{q} \) let \( \mathcal{D}_\mathbf{q} \) denote the subspace of the tangent space \( T\mathcal{N}_\mathbf{q} \) determined by the constraint equations. The union of these subspaces is the constraint distribution, which is a submanifold \( \mathcal{D} \subset T\mathcal{N} \):

\[
\mathcal{D} = \{(\mathbf{q}, \dot{\mathbf{q}}) \mid a_i^\mu(\mathbf{q}) \dot{q}^i = 0, \mu = 1, \ldots, k\}.
\]

This is the space of observable states of the system. Note that here \( \mathcal{D} \) is not the phase space: as will be clear from what follows, the complete state of the system is characterized also by certain ‘hidden’ parameters.

Remark 3. For simplicity the expression \( \dot{q}^i \) is used in two senses in mechanics. On the one hand, this is the time derivative of the function \( q^i(t) \), and on the other hand, \( \dot{q}^i \) will often denote the coordinates of a point in the tangent space \( T\mathcal{N}_\mathbf{q} \). The particular meaning is easy to see from the context.

Apart from the constraints, the system is characterized by its Lagrangian function \( L(\mathbf{q}, \dot{\mathbf{q}}) \), which, when restricted to \( \mathcal{D} \), satisfies the condition of non-degeneracy in the velocities.

Remark 4. In applications the Lagrangian \( L(\mathbf{q}, \dot{\mathbf{q}}) \) is usually a polynomial of degree at most two in the velocities.

In the case under consideration the equations of the dynamics of the system are derived from the natural generalization of Hamilton’s principle of stationary action:

find a path \( \mathbf{q}(t) \) that is an extremal of the action functional

\[
A = \int_{t_1}^{t_2} L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \, dt \tag{5}
\]

in the class of curves with fixed endpoints satisfying the constraint equations \( (4) \).

Before going further, we recall that for any vector field \( \xi(\mathbf{q}) = \xi^i(\mathbf{q}) \frac{\partial}{\partial q^i} \) on \( \mathcal{N} \) we can define its natural lift \( \check{\xi}(\mathbf{q}, \dot{\mathbf{q}}) \) to \( T\mathcal{N} \) by

\[
\check{\xi}(F(\mathbf{q}, \dot{\mathbf{q}})) = \xi^i \frac{\partial F}{\partial q^i} + \dot{\xi}^i \frac{\partial F}{\partial \dot{q}^i}, \quad \dot{\xi}^i = \frac{\partial \xi^i}{\partial \dot{q}^k} \dot{q}^k.
\]
If the field is multiplied by an arbitrary function of the coordinates \( g(q) \), then the lift becomes
\[
\hat{g}\xi(F) = g\hat{\xi}(F) + \dot{g}\xi^i \frac{\partial F}{\partial \dot{q}^i}, \quad \dot{g} = \frac{\partial g}{\partial \dot{q}^k} \dot{q}^k.
\]

In any basis of vector fields \( E_\alpha = E^i_\alpha(q) \partial / \partial q^i \), non-commuting in general, with \([E_\alpha, E_\beta] = c^\gamma_{\alpha\beta}(q) E_\gamma\), the velocity \( \dot{q} \) and the field \( \xi \) can be expressed in the form
\[
\dot{q} = \omega^\alpha E_\alpha \quad \text{and} \quad \xi = \xi^\alpha E_\alpha,
\]
and the lift \( \hat{\xi} \) has the expression
\[
\hat{\xi}(F(q, \omega)) = \xi^i \frac{\partial F}{\partial q^i} + c^\gamma_{\alpha\beta}(q) \omega^\alpha \xi^i \frac{\partial F}{\partial \omega^\gamma}.
\] (6)

Let \( q(t), t \in [t_1, t_2] \), be an admissible path in \( \mathcal{N} \), which means that \( q(t) \) satisfies the constraints (4). A variation of it can be represented as
\[
q_{\varepsilon}(t) = (q^1(t) + \varepsilon \xi^1(q), \ldots, q^n(t) + \varepsilon \xi^n(q)) + O(\varepsilon^2), \quad \varepsilon \in \mathbb{R},
\] (7)
where \( \varepsilon \) is a small quantity and \( \xi(q) \) is a vector field on \( \mathcal{N} \) such that
\[
\xi(q(t_1)) = \xi(q(t_2)) = 0.
\]

Moreover, the assumption that for each \( \varepsilon \) the curves \( q_{\varepsilon}(t) \) satisfy the constraints leads to the additional equations
\[
\hat{\xi}(f^\mu(q, \dot{q})) = a^\mu_i \dot{\xi}^i + \frac{\partial a^\mu_i}{\partial q^m} \xi^m \dot{q}^i = 0
\] (8)
for the vector field of variations.

For variations (7) satisfying (8), we can use the method of (Lagrange) undetermined multipliers to represent the extremality condition for the action (5) in the form
\[
\frac{dA}{d\varepsilon} \bigg|_{\varepsilon=0} = \int_{t_1}^{t_2} \left( \lambda_0 \hat{\xi}(L(q, \dot{q})) - \lambda_\mu \hat{\xi}(f^\mu(q, \dot{q})) \right) dt
\]
\[
= \int_{t_1}^{t_2} \left[ \xi^i \left( \lambda_0 \frac{\partial L}{\partial \dot{q}^i} - \lambda_\mu \frac{\partial a^\mu_i}{\partial q^m} \dot{q}^m \right) + \dot{\xi}^i \left( \lambda_0 \frac{\partial L}{\partial \dot{q}^i} - \lambda_\mu a^\mu_i \right) \right] dt = 0.
\] (9)

**Remark 5.** Geometrically, this is an analogue of a well-known fact in the theory of multivariate functions: the gradients of the function and the constraints are linearly dependent at a point of conditional extremum. Thus, upon discretization of the integral in (5) this gives us conditions for the extremum in the form (9). Nevertheless, a rigorous proof of such results requires considerable effort (for instance, see [4], [23]).

Hence, integrating by parts and taking (4) into account, we obtain a system of equations for the required trajectories in the form
\[
\frac{d}{dt} \left( \lambda_0 \frac{\partial L}{\partial \dot{q}^i} - \lambda_\mu a^\mu_i \right) - \lambda_0 \frac{\partial L}{\partial q^i} + \lambda_\mu \frac{\partial a^\mu_i}{\partial q^m} \dot{q}^m = 0, \quad a^\mu_i(q) \dot{q}^i = 0.
\] (10)
Here only the ratios of the quantities $\lambda_0, \lambda_1, \ldots, \lambda_k$ are determined, and not the quantities themselves. Thus, we can assume that $\lambda_0$ can take the values 0 and 1, and then the functions $q^i(t)$ and $\lambda_\mu(t), i = 1, \ldots, n, \mu = 1, \ldots, k,$ are the unknowns in (10). For $\lambda_0 = 1$ the extremals described by these equations are said to be normal, and for $\lambda_0 = 0$ they are said to be abnormal. In what follows, the set of undetermined multipliers will be denoted by $\lambda = (\lambda_1, \ldots, \lambda_k)$. Here we will only be interested in normal extremals, so we set $\lambda_0 = 1$.

We see that the phase space (state space) of our system is a $2n$-dimensional manifold involving the ‘hidden’ parameters $\lambda = (\lambda_1, \ldots, \lambda_k)$:

$$\mathcal{M} = \{(q, \dot{q}, \lambda) \mid q_\mu^i(q)\dot{q}^j = 0\}.$$ 

In particular, this shows that the principle of determinacy from classical mechanics fails for (10): in this case, for the same initial values of the coordinates $q(0) = q_0$ and the velocity $\dot{q}(0) = v_0$ there exists a $k$-parameter family of solutions that is parameterized by the initial conditions for the undetermined multipliers $\lambda(0) = \lambda_0$. For such systems we have the ‘weak’ principle of determinacy [90]: knowing the dynamics on a whole interval $(t_1, t_2)$ with $t_2 > t_1$ completely determines the motion of the system.

On the other hand, for the problem with given boundary conditions $q(t_1) = q_1$ and $q(t_2) = q_2$ to have a solution we must prescribe $2n$ initial conditions. Since the constraints (4) reduce the number of initial conditions on the distribution $D$ (by $k$), this deficiency is compensated for by the possibility of prescribing arbitrary initial conditions for the undetermined multipliers $\lambda_1, \ldots, \lambda_k$.

### 2.2. A Lagrangian for unconditional extremals.

It is well known that we can derive (10) using the new Lagrangian

$$\mathcal{L}(q, \dot{q}, \lambda) = L(q, \dot{q}) - \lambda_\mu f_\mu(q, \dot{q})$$

and the corresponding action functional

$$\mathcal{A} = \int_{t_1}^{t_2} \mathcal{L}(q, \dot{q}, \lambda) \, dt,$$  

which must be varied with respect to the variables $q$ and $\lambda$ without any restrictions on the variations of the latter.

**Theorem 1.** An admissible path $q(t), t \in [t_1, t_2]$, is a normal conditional extremal of the action functional (5) if and only if $q(t)$ and $\lambda(t)$ give an unconditional extremal of (11), that is, satisfy the Euler–Lagrange equations

$$\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^i}\right) - \frac{\partial \mathcal{L}}{\partial q^i} = 0, \quad \left(\frac{\partial \mathcal{L}}{\partial \lambda_\mu}\right) - \frac{\partial \mathcal{L}}{\partial \lambda_\mu} = 0, \quad i = 1, \ldots, n, \quad \mu = 1, \ldots, k.$$

If the Lagrangian is a homogeneous quadratic form in the velocities,

$$L = \frac{1}{2} g_{ij}(q)\dot{q}^i\dot{q}^j,$$
then the conditional extremals \( \gamma = \{ q_e(t), \ t \in [t_1, t_2] \} \) connecting the two given points \( q_e(t_1) \) and \( q_e(t_2) \) coincide with the curves admitted by the constraints and having an extremal length defined by

\[
l(\gamma) = \int_{t_1}^{t_2} \sqrt{g_{ij}(q_e(t)) \dot{q}_e^i(t) \dot{q}_e^j(t)} \, dt.
\]

Such curves \( \gamma \) are called sub-Riemannian geodesics \([108]\), because they minimize the distance between points locally in the class of curves admitted by the constraints.

2.3. An example: the Bottema–Hamel system. As an example, consider a system describing the motion of a rigid body with a flat bottom on a fixed flat support when one of the points on the support surface carries a ‘mechanism’ implementing a constraint which allows the point of the body that is in contact with this ‘mechanism’ to move only in a given direction on the support. Such a constraint was formally considered by Bottema \([58]\) and Hamel \([77]\) and is currently being widely discussed (\([13]\), \([18]\), \([119]\)).

We consider two coordinate systems (see Fig. 2):

- the first system \( Oxy \) is fixed, with origin at the point where the constraint is imposed, and the \( Oy \) axis is directed parallel to the admissible direction of motion;
- the second system \( Cx_1x_2 \) is rigidly attached to the moving body.

![Figure 2](image_url)

The position and orientation of the body are specified by the Cartesian coordinates \((x, y)\) of the point \( C \) with respect to the space-fixed coordinate system and the rotation angle \( \varphi \) of the body-fixed frame \( Cx_1x_2 \) relative to the space-fixed frame. Thus, the configuration space is the group of motions of the Euclidean plane

\[
\mathcal{N} = \{ q = (\varphi, x, y) \mid \varphi \mod 2\pi \} = \text{SE}(2),
\]

and the constraint equation takes the form

\[
\dot{y} - x\dot{\varphi} = 0.
\]
We give the Lagrangian function in the form
\[ L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + I_0\dot{\phi}^2) - U(q). \]

Writing out the equations of motion (10) and eliminating \( \dot{y} \) with the help of (12), we then obtain
\[
\begin{align*}
\dot{\phi} &= v \phi, \\
\dot{x} &= v_x, \\
\dot{y} &= x v \phi,
\end{align*}
\]
\[
\begin{align*}
\dot{v} \phi &= -(I_0 + x^2)^{-1}\left(v_x(xv\phi + \lambda) + x\frac{\partial U}{\partial \phi} + x\frac{\partial U}{\partial y}\right), \\
\dot{v}_x &= \lambda v \phi - \frac{\partial U}{\partial x}, \\
\dot{\lambda} &= (I_0 + x^2)^{-1}\left(v_x(I_0v\phi - \lambda x) - x\left(\frac{\partial U}{\partial \phi} + x\frac{\partial U}{\partial y}\right)\right).
\end{align*}
\]  

(13)

In §2.7 we look at this system from the point of view of Hamiltonian formalism and find an explicit solution.

2.4. A parametric representation. We see that the above method for finding extremals of a functional in the presence of constraints results in equations involving the unknown functions \( q(t) \) and \( \lambda(t) \) in an asymmetric fashion. Now we discuss a method for finding extremals that leads to equations which are more symmetric with respect to the unknowns.

First we solve the constraint equations (4) for \( \dot{q} \) in the parametric form
\[
\dot{q}^i = v^i(q, u) = \tau^i_\alpha(q)v^\alpha, \quad \alpha = 1, \ldots, n-k,
\]  

(14)

where \( u = (u^1, \ldots, u^{n-k}) \in \mathbb{R}^{n-k} \) denotes real parameters, and the vectors \( \tau^i_\alpha \) form a basis of the kernel of the \( n \times k \) matrix \( A(q) = ||a_i^\alpha(q)|| \) describing the constraints (that is, \( A\tau^i_\alpha \equiv 0 \)). Substituting \( \dot{q} \) into the Lagrangian, we obtain
\[ \tilde{L}(q, u) = L(q, v(q, u)). \]

The original problem (see §2.1) now reduces to finding the extremals of the functional
\[ A = \int_{t_1}^{t_2} \tilde{L}(q(t), u(t)) \, dt \]  

(15)

in the class of curves \( q(t), u(t), t \in [t_1, t_2] \), satisfying the constraint equations
\[
\frac{dq^i(t)}{dt} - v^i(q(t), u(t)) = 0, \quad i = 1, \ldots, n,
\]  

(16)

and having the projection \( q(t) \) with fixed endpoints
\[ q(t_1) = q_1 \quad \text{and} \quad q(t_2) = q_2. \]

We recall that in optimal control the Lagrangian \( L(q, u) \) is called the cost function.

By analogy with §2.2 we conclude that the conditional extremals in question coincide with the unconditional extremals of the functional
\[ \tilde{A} = \int_{t_1}^{t_2} \tilde{\mathcal{L}}(q(t), u(t), \dot{q}(t), p(t)) \, dt, \]
\[ \tilde{\mathcal{L}} = \tilde{L}(q, u) + p_i(q^i - v^i(q, u)), \quad i = 1, \ldots, n, \]
where \( p = (p_1, \ldots, p_n) \) are undetermined multipliers, and the variables \( q, u, \) and \( p \) are independently varied. After variation we obtain the following system of equations for the required extremals:

\[
\frac{\partial \mathcal{L}}{\partial u^\alpha} = \frac{\partial \mathcal{L}}{\partial u^\alpha} - p_i \frac{\partial v^i}{\partial u^\alpha} = 0, \quad \alpha = 1, \ldots, n - k,
\]

\[
\frac{\partial \mathcal{L}}{\partial p_i} = \dot{q}^i - v^i(q, u) = 0,
\]

\[
\left( \frac{\partial \mathcal{L}}{\partial q^i} - \frac{\partial \mathcal{L}}{\partial \dot{q}^i} \right)_{q = v(q, u)} = \dot{p}_i + p_m \frac{\partial v^m}{\partial q^i} - \frac{\partial \mathcal{L}}{\partial q^i} = 0, \quad i = 1, \ldots, n.
\]

(17)

Since, as noted above (in §2.1), the restriction of the Lagrangian to the constraints is non-degenerate, by using the first \( n - k \) equations we can express the quantities \( u^\alpha, \alpha = 1, \ldots, n - k, \) as functions of \( q \) and \( p \):

\[ u^\alpha = u^\alpha(q, p). \]

We now define the Hamiltonian of the system in the standard way by

\[ H(q, p) = p_i v^i(q, u_*(q, p)) - \mathcal{L}(q, u_*(q, p)). \]

(18)

Using the rules for differentiating composite functions and the relation

\[ \frac{\partial \mathcal{L}}{\partial u^\alpha} = \frac{\partial \mathcal{L}}{\partial u^\alpha}, \]

we obtain the canonical Hamiltonian equations describing the evolution of the variables \( q \) and \( p \):

\[ \dot{q}^i = \frac{\partial H(q, p)}{\partial p_i}, \quad \dot{p}_i = - \frac{\partial H(q, p)}{\partial q^i}. \]

(19)

Here the Hamiltonian \( H(q, p) \) turns out to be degenerate in the momenta (see §2.5 for details).

For an unconstrained system we have \( \dot{q}^i = u^i \), and in view of (17), \( u_*(q, p) \) and \( H(q, p) \) are given by the standard Legendre transformation

\[ p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}^i}, \quad H(q, p) = p_i u^i(q, p) - \mathcal{L}(q, u_*(q, p)). \]

(20)

**Remark 6.** The connection between this representation and the representation in terms of the variables \( q, \dot{q}, \) and \( \lambda \) considered in §§2.1 and 2.2 is given by a natural generalization of the Legendre transform

\[ p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}^i} = \frac{\partial \mathcal{L}}{\partial \dot{q}^i} - \lambda_\mu a^\mu_i(q) \]

(see [3]), which at each point \( q \) defines a map \( T_q: (\dot{q}, \lambda) \mapsto p \) that is invertible on the manifold \( D \) determined by the constraints (4). Then the Hamiltonian is given by

\[ H(q, p) = \left. \left( \frac{\partial \mathcal{L}}{\partial \dot{q}^i} \dot{q}^i - L \right) \right|_{(q, \lambda) \mapsto p}. \]

(21)
2.5. The maximum principle. The representation from the previous subsection is significantly more convenient in the case when the admissible velocities of the system are defined only on some subset of the constraint distribution \( \mathcal{D} \). In terms of the representation (14) this can be written as \( \mathbf{u} \in \mathcal{U} \subset \mathbb{R}^{n-k} \). As a rule, in applications \( \mathcal{U} \) is a bordered submanifold, whose boundary \( \partial \mathcal{U} \) is not necessarily smooth since \( \mathcal{U} \) is defined by a system of inequalities. In this situation we follow tradition by considering conditions for trajectories minimizing a functional, because this leads to the well-known Pontryagin maximum principle.

In this case the first \( n-k \) equations in (17) are replaced by the condition that \( \mathbf{u}_*(\mathbf{q}, \mathbf{p}) \) solves the equation

\[
\widetilde{\mathcal{L}} \bigg|_{\mathbf{u} = \mathbf{u}_*(\mathbf{q}, \mathbf{p})} = \min_{\mathbf{u} \in \mathcal{U}} \widetilde{\mathcal{L}} (\mathbf{q}, \dot{\mathbf{q}}, \mathbf{u}, \mathbf{p}).
\]

**Remark 7.** Of course, in the case of arbitrary extremals we must replace \( \min \) by the condition of being extremal.

This condition is usually replaced by the equivalent condition that \( \mathbf{u}_*(\mathbf{q}, \mathbf{p}) \) delivers a maximum of the Pontryagin function:

\[
\mathcal{H} \bigg|_{\mathbf{u} = \mathbf{u}_*(\mathbf{q}, \mathbf{p})} = \max_{\mathbf{u} \in \mathcal{U}} \mathcal{H} (\mathbf{q}, \mathbf{p}, \mathbf{u}),
\]

\[
\mathcal{H} (\mathbf{q}, \mathbf{p}, \mathbf{u}) = p_i v^i (\mathbf{q}, \mathbf{u}) - L (\mathbf{q}, \mathbf{u}).
\]

(22)

Since we will vary \( \mathbf{q} \) and \( \mathbf{p} \) as before, from (17) we obtain the equations for the trajectory \( \mathbf{q}(t), \mathbf{p}(t) \), which we can write in the form

\[
\dot{q}^i = \frac{\partial \mathcal{H}}{\partial p_i} \bigg|_{\mathbf{u} = \mathbf{u}_*(\mathbf{q}, \mathbf{p})}, \quad \dot{p}_i = - \frac{\partial \mathcal{H}}{\partial q^i} \bigg|_{\mathbf{u} = \mathbf{u}_*(\mathbf{q}, \mathbf{p})}.
\]

We now show that these equations can still be represented in Hamiltonian form (19) not only for \( \mathbf{u}_*(\mathbf{q}, \mathbf{p}) \) in the interior of \( \mathcal{U} \), but also when \( \mathbf{u}_*(\mathbf{q}, \mathbf{p}) \) is a non-singular boundary point on \( \partial \mathcal{U} \). In fact, for the Hamiltonian (18) we have

\[
\frac{\partial H}{\partial q^i} = \left( \frac{\partial \mathcal{H}}{\partial q^i} + \frac{\partial \mathcal{H}}{\partial u^\alpha} \frac{\partial u^\alpha (\mathbf{q}, \mathbf{p})}{\partial q^i} \right) \bigg|_{\mathbf{u} = \mathbf{u}_*(\mathbf{q}, \mathbf{p})}.
\]

If the point of maximum \( \mathbf{u}_* \) lies in the interior of \( \mathcal{U} \), then we have \( \frac{\partial \mathcal{H}}{\partial u^\alpha} = 0 \), \( \alpha = 1, \ldots, n-k \), at this point. If \( \mathbf{u}_* \in \partial \mathcal{U} \), then the vector

\[
\mathbf{N} = \left( \frac{\partial \mathcal{H}}{\partial u^1} \bigg|_{\mathbf{u} = \mathbf{u}_*}, \ldots, \frac{\partial \mathcal{H}}{\partial u^{n-k}} \bigg|_{\mathbf{u} = \mathbf{u}_*} \right)
\]

is orthogonal to \( \partial \mathcal{U} \) (with respect to the standard Euclidean scalar product in \( \mathbb{R}^{n-k} \)), while the vectors

\[
\mathbf{T}_i = \left( \frac{\partial u^1}{\partial q^i}, \ldots, \frac{\partial u^{n-k}}{\partial q^i} \right)
\]

are tangent to \( \partial \mathcal{U} \) (because we differentiate with respect to a parameter) and therefore \( (\mathbf{N}, \mathbf{T}_i) = 0 \). Using the same arguments for derivatives with respect to the \( p_i \),
we get that at non-singular points
\[
\frac{\partial H}{\partial q^i} = \frac{\partial H}{\partial q^j}
\bigg|_{u=u_*(q,p)} \quad \text{and} \quad \frac{\partial H}{\partial p_i} = \frac{\partial H}{\partial p_j}
\bigg|_{u=u_*(q,p)} = v^i(q, u_*(q, p)). \tag{23}
\]

As a result, we arrive at the following statement.

**Proposition 1.** Let \( u_*(q, p) \in U \subset \mathbb{R}^{n-k} \) be a maximum point of the Pontryagin function (22) for fixed values of \( q \) and \( p \). Then trajectories minimizing the functional (15) under the constraints (16) can be represented as \((q(t), u_*(q(t), p(t)))\), where \( q(t), p(t) \) is the solution of the canonical Hamiltonian system
\[
\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}
\]
with Hamiltonian (18) that satisfies the boundary conditions
\[
q(t_1) = q_1, \quad q(t_2) = q_2.
\]

Another important consequence of (23) is the degeneracy of \( H \) in the momenta.

**Proposition 2.** If \( u_*(q, p) \) is not a boundary point of \( U \), then
\[
\text{rank} \left| \frac{\partial^2 H}{\partial p_i \partial p_j} \right| = n - k.
\]

**Proof.** By (14) and (23),
\[
\frac{\partial H(q, p)}{\partial p_i} a^p_i(q) \equiv 0.
\]
Differentiating this identity with respect to the \( p_j \), we see that the Hessian matrix
\[
\left| \frac{\partial^2 H}{\partial p_i \partial p_j} \right|
\]
has a \( k \)-dimensional kernel. \( \square \)

In applications the set \( U \subset \mathbb{R}^{n-k} \) is usually given by a system of non-strict inequalities, and as a consequence \( U \) (together with the boundary \( \partial U \)) is a union of connected smooth submanifolds of different dimension:
\[
U = \bigcup_{\alpha} \mathcal{M}_\alpha^{n_\alpha}, \quad n_\alpha \in \mathbb{N}, \quad n_\alpha \leq n - k,
\]
where the \( n_\alpha \) are the corresponding dimensions. Repeating the arguments in the proof of Proposition 2, we can show that if \( u_*(q, p) \in \mathcal{M}_\alpha^{n_\alpha} \), then
\[
\text{rank} \left| \frac{\partial^2 H}{\partial p_i \partial p_j} \right| = n_\alpha.
\]

Thus, in the general case we can conclude that:

1) the equations describing the extremals of the functional (15) can be represented in the canonical Hamiltonian form in \( T^*\mathcal{N} = \{ (q, p) \} \), with Hamiltonian \( H(q, p) \) defined on the whole of \( T^*\mathcal{N} \);
2) the whole phase space is a union
\[ T^*N = \bigcup_{\alpha} \tilde{M}^{2n}_\alpha \]
of 2n-dimensional domains, each of which corresponds to a submanifold \( M^{n\alpha}_\alpha \subset U \), and moreover,
\[ \text{if } (q, p) \in \tilde{M}^{2n}_\alpha, \text{ then } \operatorname{rank} \frac{\partial^2 H(q, p)}{\partial p_i \partial p_j} = n_\alpha, \ u_*(q, p) \in \tilde{M}^{n\alpha}_\alpha. \]

2.6. The law of conservation of energy. If the Lagrange function and the constraints do not explicitly depend on the time, then the system has the well-known energy integral
\[ E = \frac{\partial L}{\partial \dot{q}_i} \dot{q}_i + \frac{\partial L}{\partial \dot{\lambda}_\mu} \dot{\lambda}_\mu - L = \frac{\partial L}{\partial q_i} \dot{q}_i - L. \]
Note that this integral does not explicitly involve the ‘hidden’ variables \( \lambda \).
In the Hamiltonian representation the energy integral coincides with the Hamiltonian
\[ E = H(q, p). \]

2.7. An example: the Bottema–Hamel system (continued). We represent the equations of motion (13) in the canonical Hamiltonian form. To do this we take \( u = (v_\varphi, v_x) \) to be parameters and represent the Pontryagin function (22) in the form
\[ H = \frac{1}{2} \left( p_\varphi^2 + \left( \frac{p_\varphi + x p_y}{I_0 + x^2} \right)^2 \right) + U(q) - \frac{1}{2} (p_x - v_x)^2 - \frac{1}{2} (I_0 + x^2) \left( v_\varphi - \frac{p_\varphi + x p_y}{I_0 + x^2} \right)^2. \]
We see that for fixed \( q \) and \( p \) it attains a maximum when the last two terms vanish, so that
\[ v_\varphi = \frac{p_\varphi + x p_y}{I_0 + x^2}, \quad v_x = p_x. \]

Accordingly, for the Hamiltonian (18) we have
\[ H = \frac{1}{2} \left( \frac{(p_\varphi + x p_y)^2}{I_0 + x^2} + p_x^2 \right) + U(q). \]
Furthermore, \( \operatorname{rank} \frac{\partial^2 H}{\partial p_i \partial p_j} = 2 \) everywhere.

Remark 8. We can obtain the same result using the generalized Legendre transformation (20). In this case
\[ p_\varphi = I_0 \dot{\varphi} + \lambda x, \quad p_x = \dot{x}, \quad \text{and} \quad p_y = \dot{y} - \lambda. \]
Inverting this transformation on the constraint \( \dot{y} - x\dot{\phi} = 0 \), we get that
\[
\dot{\phi} = \frac{p_\phi + xp_y}{I_0 + x^2}, \quad \dot{x} = p_x, \quad \dot{y} = \frac{x(p_\phi + xp_y)}{I_0 + x^2}, \quad \text{and} \quad \lambda = \frac{xp_\phi - I_0p_y}{I_0 + x^2}.
\]
Substituting into (21), we obtain the Hamiltonian (25).

Let us consider more closely the case of sub-Riemannian geodesics, that is, the case when \( U(q) = 0 \). Since the Hamiltonian does not explicitly depend on \( \varphi \) and \( y \), here we have the two first integrals
\[
p_\varphi = \text{const} \quad \text{and} \quad p_y = \text{const}.
\]

We now restrict the system to the common level set of the first integrals, setting
\[
p_y = p \neq 0, \quad p_\varphi = ap, \quad \text{and} \quad H = \frac{1}{2}p^2\mathcal{E},
\]
where \( a, b, \) and \( \mathcal{E} \) are some constants. Then
\[
\dot{x}^2 = \frac{p^2}{I_0 + x^2}(\mathcal{E}(I_0 + x^2) - (a + x)^2),
\]
\[
\dot{\phi} = \frac{p}{I_0 + x^2}(a + x), \quad \dot{y} = \frac{p}{I_0 + x^2}(a + x)x.
\]
If two fixed points \( q_1 \) and \( q_2 \) in \( \mathcal{N} \) must be connected by an admissible curve, then (26) shows that we must solve the two non-linear equations
\[
\Delta \varphi = \int_{x_1}^{x_2} \frac{a + x}{\sqrt{I_0 + x^2}} \frac{dx}{G(x)}, \quad \Delta y = \int_{x_1}^{x_2} \frac{x(a + x)}{\sqrt{I_0 + x^2}} \frac{dx}{G(x)},
\]
\[
G(x) = \mathcal{E}(I_0 + x^2) - (a + x)^2, \quad \Delta \varphi = \varphi_2 - \varphi_1, \quad \Delta y = y_2 - y_1,
\]
where the constant values of the first integrals \( a \) and \( \mathcal{E} \) are the unknowns.

For completeness we also consider the case when the range \( U \) of the parameters \( u \) is bounded by a rectangle (see Fig. 3, (a)):
\[
-\omega_0 \leq \varphi \leq \omega_0, \quad -v_0 \leq v_x \leq v_0.
\]
In this case \( U \) consists of nine smooth submanifolds: the open rectangle \( \mathcal{M}^2 \), the four line intervals \( \mathcal{M}^1_\alpha, \alpha = 1, \ldots, 4 \), that are its sides, and the four points \( \mathcal{M}^0_\beta, \beta = 1, \ldots, 4 \), that are its vertices. Correspondingly, the phase space of the variables \( (q, p) \) is also partitioned into nine regions, in each of which we have some particular representation for the Hamiltonian which is different from the others. This is more convenient to represent graphically; see Fig. 3, (b).

Using the Pontryagin function (24), we see that \( H_2(q, p) \) coincides with the Hamiltonian (25), and the other functions are defined as follows:
\[
H_1(\sigma_1) = \frac{1}{2}p_x^2 + (-1)^{\sigma_1}(p_\varphi + xp_y)\omega_0 + U(q) - \frac{1}{2}(I_0 + x^2)\omega_0^2,
\]
\[
\tilde{H}_1(\sigma_2) = \frac{1}{2} \left( \frac{p_\varphi + xp_y}{I_0 + x^2} \right)^2 + (-1)^{\sigma_2}p_xv_0 + U(q) - \frac{1}{2}v_0^2,
\]
\[
H_0(\sigma_1, \sigma_2) = (-1)^{\sigma_1}p_xv_0 + (-1)^{\sigma_2}(p_\varphi + xp_y)\omega_0 + U(q) - \frac{1}{2}(I_0 + x^2) - \frac{1}{2}v_0^2,
\]
where the term \(-v_0^2/2\) is not essential and can be dropped.
3. Equations of conditional extremals in a non-holonomic basis. Examples

3.1. Poincaré–Chetaev Hamiltonian formalism. It is well known (for instance, see [50] or [42]) that in many systems in classical mechanics (connected mostly with rigid-body dynamics) the equations of motion take a form that is most natural and most convenient for investigations and numerical experiments when one uses a non-holonomic (that is, coordinate-free) basis to represent generalized velocities. The Lagrangian form of these equations was found by Poincaré [114], and their Hamiltonian form was found by Chetaev [64]. We consider the use of a non-holonomic basis to present the equations of conditional extremals.

To parameterize the tangent planes $T\mathcal{N}_q$ we use a non-holonomic basis of vector fields $\mathbf{E}_s = E^i_s(q)\partial/\partial q^i, \ s = 1, \ldots, n$. Then

$$\dot{q}^i = E^i_s(q)\omega^s, \quad [\mathbf{E}_s, \mathbf{E}_r] = c^t_{sr}(q)\mathbf{E}_t, \quad (27)$$
where $\det \|E^i_s\| \neq 0$. In the new variables the constraints are expressed by

$$\tilde{f}^\mu(q, \omega) = \tilde{a}^\mu_s(q)\omega^s = 0, \quad \mu = 1, \ldots, k,$$

(28)

where $\tilde{a}^\mu_s(q) = a^i_i E^i_s$.

This change of basis in $TN$ induces a natural change of basis in the cotangent bundle $T^*N$, which has the form

$$M_s = E^i_s(q)p_i, \quad s = 1, \ldots, n.$$

In the new variables the Poisson bracket becomes

$$\{q^i, M_s\} = E^i_s(q), \quad \{M_s, M_r\} = -c^t_{sr}(q)M_t.$$

In fact, using that $\{q^i, p_j\} = \delta^i_j$ we find that

$$\{q^i, M_s\} = \{q^i, p_k\}E^k_s = E^i_s,$$

$$\{M_s, M_r\} = p_k E^i_s\{p_i, E^k_r\} + p_i E^k_s\{E^i_s, p_k\} = p_t[E_r, E_s]^i = c^t_{ts}M_t.$$

Equations (19) can now be written in the non-canonical Hamiltonian form as

$$\dot{q}^i = \{q^i, H\} = E^i_s \frac{\partial H}{\partial M_s}, \quad \dot{M}_s = \{M_s, H\} = -c^t_{ts}M_t \frac{\partial H}{\partial M_r} + E^i_s \frac{\partial H}{\partial q^i}.$$  

(29)

We now discuss the following problem:

describe the general form of a (degenerate) Hamiltonian that gives rise to the constraints (28).

To do this we take two systems of vector fields forming a basis in $TN$ such that the fields $\tau^s_\alpha = \tau^s_\alpha(q)E^i_s, \alpha = 1, \ldots, n - k$, are everywhere tangent to the subspaces $Dq$ determined by the constraints (28), while the fields $n^s_\mu = n^s_\mu(q)E^i_s, \mu = 1, \ldots, k$, are everywhere transversal to the $Dq$, so that

$$\tilde{a}^\mu_s \tau^s_\alpha \equiv 0, \quad \det \|\Gamma^\mu_\nu\| \neq 0, \quad \Gamma^\mu_\nu = \tilde{a}^\mu_s n^s_\nu.$$  

(30)

We define new momenta by the equalities

$$K_\alpha = \tau^s_\alpha M_s \quad \text{and} \quad \Lambda_\mu = n^s_\mu M_s.$$

**Proposition 3.** The trajectories of the Hamiltonian system (29) satisfy the constraint equations (28) if and only if the Hamiltonian depends only on the variables $K$:

$$H = H(K_1, \ldots, K_{n-k}).$$

**Proof.** In accordance with (30) the quasi-velocities $\omega$ satisfying equations (28) are linear combinations of the vectors $\tau^s_\alpha$:

$$\omega = \lambda^\alpha(q)\tau^s_\alpha(q).$$

On the other hand, since the substitution (27) is invertible, we find that

$$\omega^s = \frac{\partial H}{\partial M_s} = \lambda^\alpha \tau^s_\alpha.$$
At the same time,
\[
\frac{\partial K_\alpha}{\partial M_s} = \tau^s_\alpha,
\]
which shows that the functions \( H, K_1, \ldots, K_{n-k} \) are functionally dependent. \( \square \)

**Remark 9.** This result holds not only for a non-holonomic basis, but also for a coordinate basis. For example, if
\[
E_i = \frac{\partial}{\partial q^i},
\]
then
\[
K_\alpha = \tau^i_\alpha p_i, \quad \alpha = 1, \ldots, n-k,
\]
where the vectors \( \tau_\alpha = (\tau^1_\alpha, \ldots, \tau^n_\alpha) \), given in the coordinate basis, are tangent to the constraint distribution \( \mathcal{D} \). The Hamiltonian vector fields generated by the Hamiltonians \( K_\alpha \), which are linear in the momenta, project on \( TN \) in a natural way and coincide with the fields \( \tau_\alpha(q) \).

Thus, we can conclude finally that

*the extremals of the functional (5) that lie in the constraint distribution \( \mathcal{D} \) are given by a Hamiltonian system with degenerate Hamiltonian \( H = H(K) \) and with the Poisson bracket of the free (unconstrained) system.*

For example, in the case of a sub-Riemannian geodesic flow we have a Hamiltonian
\[
H = \frac{1}{2} G^{ij}(q) p_i p_j, \quad (31)
\]
where \( \det \|G^{ij}\| = 0 \). Assume that the kernel of this matrix tensor has the same dimension \( k \) at all points \( q \) and defines covector fields \( a^\mu(q) = (a^\mu_1(q), \ldots, a^\mu_n(q)) \) such that
\[
G^{ij} a^\mu_j = 0, \quad \mu = 1, \ldots, k. \quad (32)
\]
We complete these to form a covector basis by some fields
\[
b^\alpha(q) = (b^\alpha_1, \ldots, b^\alpha_n), \quad \alpha = 1, \ldots, n-k,
\]
and make the change of basis
\[
p_i = K_\alpha b^\alpha_i(q) + \Lambda_\mu a^\mu_i(q)
\]
in each cotangent space \( T^*N_q \). Substituting this into (31) and using (32), we see that \( \tilde{H} \) is independent of \( \Lambda \):
\[
\tilde{H}(q, K) = \frac{1}{2} G^{\alpha\beta}(q) K_\alpha K_\beta, \quad G^{\alpha\beta} = G^{ij} b^\alpha_i b^\beta_j,
\]
where \( G = \|G^{\alpha\beta}\| \) is a non-singular \((n-k)\times(n-k)\) matrix.

After the Legendre transformation with respect to the \( K \)-variables we find the Lagrangian function
\[
\tilde{L}(q, \omega) = \frac{1}{2} g_{\alpha\beta}(q) \omega^\alpha \omega^\beta, \quad \omega^\alpha = \frac{\partial H}{\partial K_\alpha} = G^{\alpha\beta} K_\beta,
\]
where \( g = \|g_{\alpha\beta}\| = G^{-1} \) is the inverse \((n-k)\times(n-k)\) matrix.
Finally, we give a simple formula allowing us to calculate the quasi-momenta $K_\alpha$, $\alpha = 1, \ldots, n - k$, in the case of linear constraints. We derive it by solving (28) in the parametric form $\omega^s = \tau^s_\alpha(q)u^\alpha$ and defining the function

$$
\Phi(M, u) = (M_s\omega^s)|_{\omega = \tau^s_\alpha u^\alpha}.
$$

Then

$$
K_\alpha = \frac{\partial \Phi}{\partial u^\alpha} = \tau^s_\alpha M_s.
$$

(33)

3.2. Formal variational equations for a convex body rolling on a horizontal plane. We consider the question of conditional extremals of the action in the classical problem of a rigid body moving without slipping on a horizontal plane. As already noted, we do not obtain the ‘proper’ equations of rolling dynamics in this way (for these we must use the d’Alembert–Lagrange principle), but we obtain equations which are useful for solving optimal control problems.

Fix two coordinate systems: a space-fixed system $Oxyz$ with origin $O$ on the plane of rolling and the $Oz$ axis orthogonal to this plane, and a body-fixed system $Cx_1x_2x_3$ with origin at the centre of mass of the body and axes coinciding with the principal axes of inertia (see Fig. 4).

![Figure 4](image)

We specify the position of the body in space by the coordinates of the centre of mass $R_C = (x, y, z)$ with respect to the space-fixed reference frame $Oxyz$, and we characterize its orientation by the orthogonal $3 \times 3$ matrix $Q$ whose columns are equal to the coordinates of the space-fixed basis vectors $e_x, e_y, e_z$ relative to the moving frame $Cx_1x_2x_3$. This matrix can be parameterized by the Euler angles $(\theta, \varphi, \psi)$:

$$
Q = \begin{pmatrix}
\cos \varphi \cos \psi - \cos \theta \sin \psi \sin \varphi & \cos \varphi \sin \psi + \cos \theta \cos \psi \sin \varphi & \sin \varphi \sin \theta \\
-\sin \varphi \cos \psi - \cos \theta \sin \psi \cos \varphi & -\sin \varphi \sin \psi + \cos \theta \cos \psi \cos \varphi & \cos \varphi \sin \theta \\
\sin \theta \sin \psi & -\sin \theta \cos \psi & \cos \theta
\end{pmatrix},
$$

(34)

so that the configuration space of the system is the group of motions of Euclidean 3-space:

$$
N = \{(R_C, Q)\} = SE(3).
$$

To parameterize the velocities we also use variables invariant under the group of motions, the components of the angular velocity vector $\omega = (\omega_1, \omega_2, \omega_3)$ of the
body and the velocity of its centre of mass $\mathbf{v} = (v_1, v_2, v_3)$ relative to the moving frame. Then
\[
\dot{\theta} = \omega_1 \cos \varphi - \omega_2 \sin \varphi, \quad \dot{\varphi} = \omega_3 - \frac{\cos \theta}{\sin \theta} (\omega_1 \sin \varphi + \omega_2 \cos \varphi),
\]
\[
\dot{\psi} = \frac{\omega_1 \sin \varphi + \omega_2 \cos \varphi}{\sin \theta},
\]
and
\[
\mathbf{R}_P = \mathbf{Q}(\mathbf{v} + \mathbf{\omega} \times \mathbf{r} + \dot{\mathbf{r}}),
\]
where $\mathbf{R}_P = (x, y, 0)$ is the radius vector of the point of contact $P$ in the fixed frame $Oxyz$, and $\mathbf{r}$ is the vector from the centre of mass to the point of contact in the moving frame $Cx_1x_2x_3$. For a convex body $\mathbf{r}$ is fully determined by the orientation of the body and is independent of the angle of precession, that is, $\mathbf{r} = \mathbf{r}(\theta, \varphi)$, so that
\[
\dot{\mathbf{r}} = \frac{\partial \mathbf{r}}{\partial \theta} \dot{\theta} + \frac{\partial \mathbf{r}}{\partial \varphi} \dot{\varphi}.
\]

The condition of rolling without slipping means that at the point of contact with the plane the velocity of the point of the rigid body vanishes; in our variables this is expressed by
\[
f = \mathbf{v} + \mathbf{\omega} \times \mathbf{r} = 0.
\]

We denote the quasi-momenta corresponding to the quasi-velocities $\mathbf{v}$ and $\mathbf{\omega}$ by $\mathbf{p}$ and $\mathbf{M}$, and we write the corresponding Poisson structure using superfluous variables:
\[
\begin{align*}
\{M_i, M_j\} &= -\varepsilon_{ijk}M_k, \quad \{M_i, p_j\} = -\varepsilon_{ijk}p_k, \\
\{M_i, \alpha_j\} &= -\varepsilon_{ijk}\alpha_k, \quad \{M_i, \beta_j\} = -\varepsilon_{ijk}\beta_k, \quad \{M_i, \gamma_j\} = -\varepsilon_{ijk}\gamma_k, \\
\{p_i, x\} &= -\alpha_i, \quad \{p_i, y\} = -\beta_i, \quad \{p_i, z\} = -\gamma_i,
\end{align*}
\]
where $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, $\beta = (\beta_1, \beta_2, \beta_3)$, and $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ are the column vectors of the matrix $\mathbf{Q}$, which satisfy the equalities
\[
\alpha^2 = \beta^2 = \gamma^2 = 1, \quad (\alpha, \beta) = (\beta, \gamma) = (\gamma, \alpha) = 0.
\]

Although this system of variables is quite superfluous, it has the very useful property that in these variables the Poisson structure (36) is linear and homogeneous. Furthermore, the equations of motion arising in applications usually have a simple algebraic form in these variables (that is, they do not involve functions such as sin, cos, and so on).

In the vector form the Hamiltonian equations of motion can be written as follows:
\[
\begin{align*}
\dot{\mathbf{M}} &= \mathbf{M} \times \frac{\partial H}{\partial \mathbf{M}} + \mathbf{p} \times \frac{\partial H}{\partial \mathbf{p}} + \alpha \times \frac{\partial H}{\partial \alpha} + \beta \times \frac{\partial H}{\partial \beta} + \gamma \times \frac{\partial H}{\partial \gamma}, \\
\dot{\mathbf{p}} &= \mathbf{p} \times \frac{\partial H}{\partial \mathbf{M}} - \frac{\partial H}{\partial x} \alpha - \frac{\partial H}{\partial y} \beta - \frac{\partial H}{\partial z} \gamma, \\
\dot{\alpha} &= \alpha \times \frac{\partial H}{\partial \mathbf{M}}, \quad \dot{\beta} = \beta \times \frac{\partial H}{\partial \mathbf{M}}, \quad \dot{\gamma} = \gamma \times \frac{\partial H}{\partial \mathbf{M}}, \\
\dot{x} &= \left(\alpha, \frac{\partial H}{\partial \mathbf{p}}\right), \quad \dot{y} = \left(\beta, \frac{\partial H}{\partial \mathbf{p}}\right), \quad \dot{z} = \left(\gamma, \frac{\partial H}{\partial \mathbf{p}}\right).
\end{align*}
\]
For this system we now indicate degenerate Hamiltonians that lead to the constraint (35). By Proposition 3, to do this we must take a system of quasi-momenta corresponding to solutions of the constraint equations. We parameterize a solution of (35) as follows:

\[
\omega = u, \quad v = r \times u.
\]

By (33) it corresponds to a set of quasi-momenta which can be represented in vector form as

\[
K = \frac{\partial}{\partial u} \left[ ((M, \omega) + (p, v)) \Big|_{v=r \times u, \omega=u} \right] = M + p \times r.
\]

Thus, for the degenerate Hamiltonians

\[
H = \tilde{H}(q, K), \quad q = (R_C, Q),
\]

we always have the constraint (35) corresponding to rolling without slipping.

In this case we do not discuss the question of whether the motions described by the system (37) with Hamiltonian (38) are physically realized. We underline again that a rolling motion of a body on an absolutely rough plane under the action of external fields is described by equations of motion in non-holonomic mechanics (see §4.3) which are derived from the d’Alembert–Lagrange principle and are not Hamiltonian in general.

### 3.3. A vakoball and the Jurdjevic ball.

3.3.1. Assume that a ball whose centre of mass \(C\) coincides with its geometric centre moves without slipping on a horizontal plane, and let the vector \(r = -R\gamma\), where \(\gamma\) is the vertical basis vector expressed in the moving frame, be equal to the third column vector of the matrix \(Q\) in (34). Then the constraint (35) can be expressed by

\[
v + R\gamma \times \omega = 0.
\]

The authors of [19] and [84] considered extremals of such a system with the Lagrangian function equal to the kinetic energy of the ball, which has the following expression in the moving frame \(Cx_1x_2x_3\) with axes equal to the principal axes of inertia:

\[
L = \frac{1}{2} \left( mv^2 + (\omega, I\omega) \right),
\]

where \(m\) and \(I = \text{diag}(I_1, I_2, I_3)\) is the inertia tensor of the ball. Using the previous subsection, §3.2, we can show that the Hamiltonian of this system is

\[
\tilde{H} = \frac{1}{2} (K, J^{-1}(\gamma)K),
\]

\[
K = M + R\gamma \times p, \quad J(\gamma) = (I + mR^2E) - mR^2\gamma \otimes \gamma,
\]

(39)

where \(R\) is the radius of the ball and \(E\) is the identity matrix. In this case

\[
\omega = \frac{\partial H}{\partial M} = \frac{\partial \tilde{H}}{\partial K} = J^{-1}(\gamma)K \quad \text{and} \quad v = \frac{\partial H}{\partial p} = R \frac{\partial \tilde{H}}{\partial K} \times \gamma = R(J^{-1}(\gamma)K) \times \gamma.
\]
By (37), in this case we can single out a system of nine equations that describes the evolution of the vectors $\mathbf{M}$, $\mathbf{p}$, and $\mathbf{\gamma}$:

\[
\begin{align*}
\dot{\mathbf{M}} &= \mathbf{M} \times \mathbf{\omega} + R \mathbf{p} \times (\mathbf{\omega} \times \mathbf{\gamma}) + \mathbf{\gamma} \times \left( \frac{\partial \tilde{H}}{\partial \mathbf{\gamma}} + R \mathbf{p} \times \mathbf{\omega} \right), \\
\dot{\mathbf{p}} &= \mathbf{p} \times \mathbf{\omega}, \\
\dot{\mathbf{\gamma}} &= \mathbf{\gamma} \times \mathbf{\omega}.
\end{align*}
\]

(40)

It follows from these equations that $\mathbf{\gamma}$ and $\mathbf{p}$ are constant vectors relative to the space-fixed frame $Oxyz$. For this reason the equations have the natural first integrals

\[
F_1 = \mathbf{\gamma}^2, \quad F_2 = \mathbf{p}^2, \quad \text{and} \quad F_3 = (\mathbf{\gamma}, \mathbf{p}).
\]

(41)

We choose the fixed frame so that the $Ox$ axis lies in the plane of the vectors $\mathbf{\gamma}$ and $\mathbf{p}$. Then the evolution of the centre of mass of the body and the angle of precession is described by the equations

\[
\begin{align*}
\dot{x} &= \frac{R}{\sqrt{P^2 - p_z^2}} (\mathbf{\gamma} \times \mathbf{p}, \mathbf{\omega}), \\
\dot{y} &= -\frac{R}{\sqrt{P^2 - p_z^2}} (\mathbf{p} - p_z \mathbf{\gamma}, \mathbf{\omega}), \\
\dot{z} &= 0, \\
\dot{\psi} &= \frac{2 \gamma_1 \gamma_2 (\mathbf{p}, \mathbf{\omega})}{\gamma_1^2 - \gamma_2^2},
\end{align*}
\]

where the constant first integrals (41) are specified by

\[
F_1 = 1, \quad F_2 = P^2, \quad F_3 = p_z.
\]

The simplest integrable case mentioned in [19] and [84] occurs when the inertia tensor is spherical:

\[
\mathbf{I} = I_z \mathbf{E}.
\]

In this case it is convenient to pass to the space-fixed coordinate system $Oxyz$, so that for the components

\[
M_x = (\mathbf{M}, \mathbf{\alpha}), \quad M_y = (\mathbf{M}, \mathbf{\beta}), \quad \text{and} \quad M_z = (\mathbf{M}, \mathbf{\gamma})
\]

of the moment vector and the components

\[
p_x = (\mathbf{p}, \mathbf{\alpha}), \quad p_y = (\mathbf{p}, \mathbf{\beta}), \quad p_z = (\mathbf{p}, \mathbf{\gamma})
\]

of the momentum we obtain a Lie algebra isomorphic to $\text{so}(3) \oplus \mathbb{R}^3$:

\[
\{M_x, M_y\} = M_z, \quad \{M_y, M_z\} = M_x, \quad \{M_z, M_x\} = M_y
\]

(42)

(the other brackets are zero). In the new variables the Hamiltonian (39) has the form

\[
\tilde{H} = \frac{1}{2} \left( \frac{(M_x - R p_y)^2 + (M_y + R p_x)^2}{I_x} + \frac{M_z^2}{I_z} \right), \quad I_x = I_z + mR^2.
\]

Hence, the momenta of the system are preserved:

\[
p_x = \text{const}, \quad p_y = \text{const}, \quad p_z = \text{const}.
\]
Consequently, the system describing the evolution of \( M_x, M_y, \) and \( M_z \) is equivalent to a special case of the Joukowski–Volterra system [42], which describes the dynamics of a rigid body with a fixed point that is equipped with a rotating rotor. On a level surface of the integrals

\[
p_x = -R^{-1}k_y, \quad p_y = R^{-1}k_x
\]

we obtain

\[
\begin{align*}
\dot{M}_x &= (I_x^{-1} - I_z^{-1})M_yM_z + I_x^{-1}k_yM_z, \\
\dot{M}_y &= -(I_x^{-1} - I_z^{-1})M_xM_z - I_x^{-1}k_xM_z, \\
\dot{M}_z &= I_x^{-1}(M_xk_y - M_yk_x).
\end{align*}
\]  

(43)

For known solutions of this system, the position and orientation of the ball can be found by quadratures:

\[
\dot{x} = -RI_x^{-1}(k_y - M_y), \quad \dot{y} = RI_x^{-1}(k_x - M_x), \quad \dot{z} = 0,
\]

\[
\dot{Q} = Q\Omega, \quad \Omega = \begin{pmatrix} 0 & I_x^{-1}M_z & -I_x^{-1}(M_y - k_y) \\
-I_x^{-1}M_z & 0 & I_x^{-1}(M_x - k_x) \\
I_x^{-1}(M_y - k_y) & -I_x^{-1}(M_x - k_x) & 0 \end{pmatrix}.
\]  

(44)

We can show that the trajectories of this system correspond to conditional extremals of the action functional with the Lagrangian

\[
L = \frac{1}{2}(\mu_1(v_x^2 + v_y^2) + \mu_2R^2\omega_z^2), \quad \mu_1 = m + \frac{I_z}{R^2}, \quad \mu_2 = \frac{I_z}{R^2}.
\]

3.3.2. We remark that in [19], [31], and [32] a formulation of the problem is discussed which is slightly different at first glance, namely, a Chaplygin ball (balanced but dynamically asymmetric) rolls on a plane with the constraint (35) and supplied with balanced rotors whose rotation axes are body-fixed (see Fig. 5). The equations of motion of this system (see §4), which are different from (40) in general, have a vector integral of motion, the angular momentum with respect to the point of contact:

\[
\mathcal{H} = I\omega + mR^2\gamma \times (\omega \times \gamma) + \sum_{k=1}^{N_r} J_k u_k n_k = \text{const},
\]

(45)

where the \( J_k, u_k, \) and \( n_k \) are the moments of inertia, the angular velocities, and the direction vectors of the axes of the rotors, and \( N_r \) is the number of rotors. For constant angular velocities \( (u_k = \text{const}, k = 1, \ldots, N_r) \) the system is integrable [105] and can be represented in a conformally Hamiltonian form [45]. Without time rescaling it is non-Hamiltonian even for \( u_k \equiv 0 \) [35].

However, if we do not treat the \( u_k \) as prescribed functions of time but rather as controls of the system (other ways to control the ball were considered in [12] and [83]), then the situation is different. For example, consider motions of the ball that begin and end in a state of rest, when \( \omega = 0 \) and \( u_k = 0 \). Then the first integral (45) vanishes. Assume in addition that the ball carries three rotors whose axes are not coplanar. Then the angular and linear velocity of the ball can be uniquely expressed in terms of the controls \( u_k \):

\[
\omega = J^{-1}(\gamma)\left(\sum_{k=1}^{3} J_k u_k n_k\right), \quad \mathbf{v} = R\omega \times \gamma,
\]

(46)
where $J(\gamma)$ is defined in (39). These relations describe precisely the distribution determined by the constraint (35). Furthermore, by taking a suitable functional to be minimized we obtain equation (40) for the optimal control problem.

The case when the ball carries just two controlled rotors instead of three [117] is also interesting. In this case the velocities of the ball can be expressed in the form

$$\begin{pmatrix} \omega \\ v \end{pmatrix} = u_1 \chi_1(\gamma) + u_2 \chi_2(\gamma),$$

where $\chi_1(\gamma)$ and $\chi_2(\gamma)$ are six-dimensional vectors which can be found from (46). Using equation (45), we see that the corresponding distribution is defined by the system of constraints

$$v + R\gamma \times \omega = 0 \quad \text{and} \quad (J(\gamma)\omega, a) = 0,$$

where $a$ is a body-fixed vector orthogonal to $n_1$ and $n_2$ (for example, $a = n_1 \times n_2$). In this case the Hamiltonian of the system can be expressed in terms of the components of the vector

$$K' = (J(\gamma)a) \times (M + R\gamma \times p).$$

3.3.3. Besides the model of rolling without slipping considered above, a model of rolling without slipping and spinning (see Fig. 6) is also used in mechanics and robotics; see the comprehensive surveys of the literature in [45] and [51]. In this case the constraint equations for the ball can be written as

$$v + R\gamma \times \omega = 0, \quad (\omega, e_z) = 0,$$

where $e_z$ is a normal vector to the plane of rolling. For this system, sub-Riemannian geodesics are described by Hamiltonian equations with the Poisson bracket (42) and a Hamiltonian $H$ depending on the variables

$$K_x = M_x - Rp_y \quad \text{and} \quad K_y = M_y + Rp_x.$$
For instance, Jurdjevic [80] considered the problem of minimizing the functional

\[ \int_{t_1}^{t_2} \frac{1}{2} (v_x^2 + v_y^2) \, dt \]

(the so-called plate-ball problem). The system in [80] is a limiting case of the system (43), (44), where we formally let \( I_z \to \infty \).

3.4. A rotor-controlled body in a fluid.

3.4.1. Consider a rigid body, a shell, with three axially-symmetric internal rotors, moving in an ideal fluid (see Fig. 7). We assume that the weight of the body with the rotors is equal to the weight of the displaced fluid, so that the system has zero buoyancy.
The configuration space of this system\(^3\) is known to be the group of motions of Euclidean space \(N = SE(3)\). We introduce two coordinate systems for its parameterization, a fixed one \(Oxyz\) and a moving one \(Cx_1x_2x_3\) rigidly attached to the shell.

In the general case the centre of mass of the whole system (consisting of the shell and the rotors) need not be at the point \(C\). The position of the body in space will be determined by the coordinates \(r\) of the centre of mass in the fixed frame \(Oxyz\), and its orientation will be determined by the orthogonal matrix \(Q\) with columns formed by the coordinates of the fixed basis vectors with respect to the moving frame \(Cx_1x_2x_3\). We parameterize \(Q\) in terms of the Euler angles \(\xi = (\psi, \theta, \phi)\):

\[
Q = \begin{pmatrix}
\cos \phi \cos \psi - \cos \theta \sin \psi \sin \phi & \cos \phi \sin \psi + \cos \theta \cos \psi \sin \phi & \sin \phi \sin \theta \\
- \sin \phi \cos \psi - \cos \theta \sin \psi \cos \phi & - \sin \phi \sin \psi + \cos \theta \cos \psi \cos \phi & \cos \phi \cos \theta \\
\sin \theta \sin \phi & - \sin \theta \cos \phi & \cos \theta 
\end{pmatrix}.
\]

The evolution of \(r\) and \(\xi\) is described by the following kinematic relations:

\[
\dot{r} = QTv, \quad \dot{\xi} = \Xi\omega,
\]

\[
\Xi = \begin{pmatrix}
\sin \phi & \cos \phi & 0 \\
\sin \theta & \sin \theta & 0 \\
\cos \phi & - \sin \phi & 0 \\
- \cot \theta \sin \phi & - \cot \theta \cos \phi & 1
\end{pmatrix},
\]

where \(v\) and \(\omega\) are the velocity vector of the centre of mass and the angular velocity vector of the body, respectively.

In the general case the kinematic energy of the full system is

\[
T = \frac{1}{2}(\omega, A\omega) + (v, B\omega) + \frac{1}{2}(v, Cv) + (\omega, \kappa),
\]

where \(A\), \(B\), and \(C\) are matrices determined by the shape of the shell and the distribution of mass in the body (see examples below) with \(A\) and \(C\) always non-singular, and \(\kappa\) is the vector of the total gyrostatic momentum of the rotors, which is expressed in terms of the angular velocities \(w_k\) of their rotation by

\[
\kappa = \sum_{k=1}^{3} J_k w_k n_k,
\]

with \(J_k\) and \(n_k\) the moments of inertia of the rotors and the direction vectors of their axes.

Remark 10. By specifying the origin and the directions of the axes of the moving coordinate system \(Cx_1x_2x_3\), we can eliminate six parameters in the formula for the kinetic energy (47). In particular, \(B\) can (if necessary) be made symmetric, and one of the matrices \(A\) and \(C\) can be made diagonal.

---

\(^3\)Here we consider the case when the rotational velocities of the rotors are either given functions of time or control parameters; the rotational angles of the rotors are not additional degrees of freedom.
We define the momentum $\mathcal{P}$ and the angular momentum $\mathcal{M}$ of the system by

$$\mathcal{P} = \frac{\partial T}{\partial v} = B\omega + Cv \quad \text{and} \quad \mathcal{M} = \frac{\partial T}{\partial \omega} = A\omega + B^T v + \kappa.$$  

If no external forces other than the hydrodynamical resistance act on the system, then the equations of motion can be represented in the form

$$\dot{\mathcal{P}} = \mathcal{P} \times \omega, \quad \dot{\mathcal{M}} = \mathcal{M} \times \omega + \mathcal{P} \times v,$$

and therefore the relations

$$B\omega + Cv = 0 \quad \text{and} \quad A\omega + B^T v + \kappa = 0$$  

are invariant.

Now let us consider the case when the angular velocities $w_k$ of the rotors and therefore the gyrostatic momentum $\kappa$ are control parameters. For example, by using (48) we can express the angular velocity and the velocity of the centre of mass in terms of $\kappa$:

$$\omega = (B^T C^{-1} B - A)^{-1} \kappa, \quad v = C^{-1} B (B^T C^{-1} B - A)^{-1} \kappa.$$  

Furthermore, it is the submanifold (48) which contains the trajectories corresponding to motions that start and end in a state of rest (of both the body and the rotors). Consequently, to analyze the control of the body by means of internal rotors in the ideal fluid approximation, we must investigate the system subject to constraint equations of the form

$$B\omega + Cv = 0.$$  

**Remark 11.** Another formulation of the optimal control problem for a body in a fluid that is an autonomous underwater vehicle (AUV) was considered in [7]. We recall that the AUV in [7] is stretched in shape in one direction, and it was thus assumed that it moves only in this direction: there is no motion in the two perpendicular directions because of viscous friction (that is, the non-integrable constraints $v_1 = 0$ and $v_2 = 0$ are imposed on the system).

3.4.2. In the general case the constraints (49) define a completely non-holonomic distribution. We consider cases when this distribution is holonomic or not completely non-holonomic (see [131] and [57] for details).

1) If $B = 0$, then $r = \text{const}$.

2) If $C = cE$ and $B$ is skew-symmetric,

$$B = \begin{pmatrix} 0 & b_3 & -b_2 \\ -b_3 & 0 & b_1 \\ b_2 & -b_1 & 0 \end{pmatrix},$$

then the holonomic constraints can be represented as

$$cr + Q^T b = \text{const}, \quad b = (b_1, b_2, b_3).$$
3) If \( C = \text{diag}(c_1, c_1, c_3) \) and

\[
B = \begin{pmatrix}
0 & b_3 & 0 \\
-b_3 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix},
\]

then

\[ c_1 r + b_3 e_3 = \text{const}, \quad e_3 = (0, 0, 1). \]

3.4.3. We showed above (in § 3.1) that optimal trajectories of a constrained system (that is, extremals of a certain functional) are defined by a Hamiltonian system whose Poisson structure corresponds to the free (unconstrained) Hamiltonian system, and the Hamiltonian is degenerate in the momenta: \( H = \tilde{H}(K, r, \xi) \).

In this case the equations have the form (37). To find the quasi-momenta \( K \) that are arguments of the Hamiltonian we parameterize the constraints (49) as follows:

\[
\omega = u, \quad v = -D^T u, \quad D^T = C^{-1} B.
\]

Using (33), we find that

\[
K = \frac{\partial}{\partial u} ((M, \omega) + (p, v)) \big|_{\omega = u, v = -D^T u} = M - Dp.
\]

We now look at the special case of sub-Riemannian geodesics given by a Hamiltonian of the form

\[
H = \frac{1}{2}(K, G K),
\]

(50)

where \( G \) is a constant matrix. The equations describing the evolution of \( M \) and \( p \) can be singled out in the form

\[
\dot{M} = M \times \frac{\partial H}{\partial M} + p \times \frac{\partial H}{\partial p}, \quad \dot{p} = p \times \frac{\partial H}{\partial M}.
\]

(51)

Equations (51) with the Hamiltonian (50) are Kirchhoff’s equations in rigid body dynamics. In was shown, for instance, in [42] that this is a Hamiltonian system with Poisson structure corresponding to the algebra \( e(3) \) [111]. Apart from the Hamiltonian \( H \), the system (51) has the additional integrals

\[
C_1 = (M, p) \quad \text{and} \quad C_2 = p^2,
\]

which are Casimir functions of \( e(3) \). We see that we need a further additional integral to integrate the system (51) using Liouville’s theorem.

Several integrable cases of Kirchhoff’s equations are known [42], some of which are realized for the degenerate Hamiltonian (50):

- Kirchhoff’s case, when

\[
D = \text{diag}(d_1, d_1, d_3), \quad G = \text{diag}(g_1, g_1, g_3),
\]

and the additional integral is

\[
F = M_3;
\]
• **Clebsch’s family**, when
  \[
  D = \text{diag}(d_1, d_2, d_3), \quad G = D^{-1},
  \]
  and the additional integral is
  \[
  F = M^2 - \det (p, D^{-1}p);
  \]

• the **Steklov–Lyapunov family**, for which
  \[
  D = \text{diag}(d_1, d_2, d_3), \quad G = \text{diag}(-d_1 + d_2 + d_3, d_1 - d_2 + d_3, d_1 + d_2 - d_3),
  \]
  and the additional integral has the form
  \[
  F = M^2 + 2(M, Dp) + (d_2 - d_3)^2 p_1^2 + (d_1 - d_3)^2 p_2^2 + (d_1 - d_2)^2 p_3^2.
  \]

**Remark 12.** We can obtain the last two cases by combining first integrals of the Clebsch and Steklov–Lyapunov families into a quadratic form and then reducing it to a Hamiltonian of the form (50). In the same way we can show that other integrable cases of Kirchhoff’s equations cannot be generalized for a Hamiltonian of the form (50).

### 3.4.4

In conclusion we discuss the conditions on the matrices \( A, B, \) and \( C \) in the expression for the kinetic energy (47) that are imposed by various symmetries of the shell and the distribution of mass.

1) **The shell and the distribution of mass of the system are symmetric relative to a rotation through the angle \( 2\pi/n, n > 2 \).** In this case, for a suitably chosen coordinate system \( Cx_1x_2x_3 \) (with \( C \) lying on the symmetry axis, which coincides with the \( Cx_3 \) axis) we can obtain the matrices in the form

\[
A = \text{diag}(a_1, a_2, a_3), \quad B = \text{diag}(b_1, b_2, b_3), \quad C = \text{diag}(c_1, c_2, c_3).
\]

2) **The shell and the distribution of mass have a plane of symmetry.** Choosing \( Cx_1x_2x_3 \) so that the plane \( Cx_2x_3 \) coincides with the plane of symmetry, we can obtain

\[
A = \begin{pmatrix}
a_{11} & 0 & 0 \\
0 & a_{22} & a_{23} \\
0 & a_{23} & a_{33}
\end{pmatrix}, \quad
B = \begin{pmatrix}
b_{12} & b_{13} & 0 \\
b_{12} & 0 & 0 \\
b_{13} & 0 & 0
\end{pmatrix}, \quad
C = \begin{pmatrix}
c_{11} & 0 & 0 \\
0 & c_{22} & c_{23} \\
0 & c_{23} & c_{33}
\end{pmatrix}.
\]

(Making a rotation about the \( Cx_1 \) axis, we can in addition diagonalize \( A \) or \( C \).)

3) **The shell has three mutually orthogonal planes of symmetry.** Choosing a system of coordinates \( Cx_1x_2x_3 \) so that \( C \) is the geometric centre of the shell and the axes are directed along the planes of symmetry, we can obtain

\[
A = \text{diag}(a_1, a_2, a_3), \quad B = \begin{pmatrix}
b_{12} & b_{13} & 0 \\
-b_{12} & 0 & b_{23} \\
b_{13} & -b_{23} & 0
\end{pmatrix}, \quad
C = \text{diag}(c_1, c_2, c_3).
\]
3.5. A vakonomic Chaplygin sleigh. Consider a system consisting of a platform sliding along a horizontal plane, with a small wheel (or knife edge) fixed at a point $P$ of the platform, which allows $P$ to move only in the prescribed direction parallel to the plane of the wheel (the knife edge); see Fig. 8. In addition, there is a ‘manipulator’ attached at a point $M$ of the platform and allowing one to prescribe the velocity of this point. This version of the problem is typical of optimal control, although it is far from a physical realization.

We choose two coordinate systems:

• a space-fixed system $Oxy$,
• a body-fixed system $Cx_1x_2$, with origin at the point where the manipulator is attached and with the $Cx_1$ axis parallel to the plane of the wheel (the knife edge).

![Figure 8](image)

The position and orientation of the body are specified by the Cartesian coordinates $(x, y)$ of the point $C$ in the fixed frame and the angle of rotation $\varphi$ of the moving frame $Cx_1x_2$ relative to the fixed frame. Thus, the configuration space is the group of motions of the Euclidean plane:

$$\mathcal{N} = \{q = (\varphi, x, y) \mid \varphi \text{ mod } 2\pi\} = SE(2).$$

Let $a$ be the $x$-coordinate of $M$ and let $b$ be the $y$-coordinate of $P$; then the components of the velocities of these points with respect to the moving frame $Cx_1x_2$ can be expressed as follows in terms of the generalized velocities of the system:

$$v_M = (\dot{x} \cos \varphi + \dot{y} \sin \varphi, -\dot{x} \sin \varphi + \dot{y} \cos \varphi + a \dot{\varphi}),$$

$$v_P = (\dot{x} \cos \varphi + \dot{y} \sin \varphi - b \dot{\varphi}, -\dot{x} \sin \varphi + \dot{y} \cos \varphi).$$

Since $P$ cannot move in the direction of the $Cx_2$ axis, we obtain the constraint equation

$$-\dot{x} \sin \varphi + \dot{y} \cos \varphi = 0. \tag{52}$$

Denoting the components of the velocity of the point $M$ (at which the manipulator is fixed) with respect to the fixed frame $Oxy$ by $u_x$ and $u_y$, we get that

$$v_M = (u_x \cos \varphi + u_y \sin \varphi, u_x \sin \varphi + u_y \cos \varphi).$$
Thus, the evolution of the position and orientation of the platform is described by the equations

\[
\dot{x} = (u_x \cos \varphi + u_y \sin \varphi) \cos \varphi, \quad \dot{y} = (u_x \cos \varphi + u_y \sin \varphi) \sin \varphi, \\
\dot{\varphi} = u_x \sin \varphi - u_y \cos \varphi.
\]

Let \( p_x, p_y, \) and \( p_\varphi \) be the canonical momenta corresponding to the generalized coordinates in the cotangent bundle \( T^*N \). Then the quasi-momenta \( K_1 \) and \( K_2 \) that are arguments of the degenerate Hamiltonian \( \tilde{H}(q, K_1, K_2) \) corresponding to the constraint (52) can be taken to be

\[
K_1 = p_x \cos \varphi + p_y \sin \varphi \quad \text{and} \quad K_2 = \frac{p_\varphi}{a}.
\]

The simplest quadratic Hamiltonian

\[
\tilde{H} = \frac{1}{2}(K_1^2 + K_2^2)
\]

(53)
corresponds to conditional extremals of the action with the Lagrangian

\[
L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + a^2 \dot{\varphi}^2) = \frac{1}{2}(u_x^2 + u_y^2).
\]

(54)

**Remark 13.** The coefficients of the terms in (53) can always be made equal by a simplest canonical transformation

\[
x \to k^{-1}x, \quad y \to k^{-1}y, \quad p_x \to kp_x, \quad p_y \to kp_y.
\]

**Remark 14.** The system with Lagrangian (54) and the constraint (52) arises in the car parking problem (for instance, see [103] and the references therein). In this case the control is implemented by adjusting the translational velocity and the angle of rotation of the front wheels.

Then the system has the obvious (cyclic) first integrals

\[
p_x = \text{const} \quad \text{and} \quad p_y = \text{const}.
\]

On a level surface of these integrals and the energy integral we let

\[
p_x = ap \cos \alpha, \quad p_y = ap \sin \alpha, \quad H = \frac{a^2 p^2}{2} h,
\]

where \( \alpha, p, \) and \( h \) are some constants. Then the equations can be represented as

\[
\dot{\varphi}^2 = p^2 (h - \cos^2(\varphi - \alpha)), \\
\frac{\dot{x}}{a} = p \cos(\varphi - \alpha) \cos \varphi, \quad \frac{\dot{y}}{a} = p \sin(\varphi - \alpha) \sin \varphi.
\]

For an extremal with fixed initial point \((x_1, y_1, \varphi_1)\) and terminal point \((x_2, y_2, \varphi_2)\) we can find the values of \( \alpha \) and \( h \) from the system of (non-linear) equations

\[
\Delta \varphi = \pm \int_{\varphi_1}^{\varphi_2} \frac{\cos(\varphi - \alpha) \cos \varphi}{\sqrt{h - \cos^2(\varphi - \alpha)}} \, d\varphi, \\
\Delta \varphi = \pm \int_{\varphi_1}^{\varphi_2} \frac{\cos(\varphi - \alpha) \sin \varphi}{\sqrt{h - \cos^2(\varphi - \alpha)}} \, d\varphi, \\
\Delta x = \frac{x_2 - x_1}{a}, \quad \Delta y = \frac{y_2 - y_1}{a}.
\]
3.6. A vakonomic ice skate on an inclined plane. Here we consider more closely a balanced Chaplygin sleigh, an ice skate. This problem is not only interesting from the standpoint of control theory, but it also has a hydrodynamic realization, which was proposed by Kozlov [90] in the framework of vakonomic mechanics.

In the previous subsection we set $a = 0$ and defined the quasi-momenta $K_1$ and $K_2$ by

$$K_1 = p_x \cos \varphi + p_y \sin \varphi \quad \text{and} \quad K_2 = p_\varphi.$$  

In [90] Kozlov analyzed extremals of this system when the Lagrange function corresponds to an ice skate on an inclined plane:

$$L = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2) + \frac{1}{2} I \dot{\varphi}^2 - \mu y,$$  

where $m$ and $I$ are the mass and moment of inertia of the skate and $\mu$ is its weight.

Using the previous subsection, we can show that the degenerate Hamiltonian of the system with Lagrangian (55) and constraint (52) has the representation

$$H = \frac{1}{2} m (p_x \cos \varphi + p_y \sin \varphi)^2 + \frac{1}{2} I p_\varphi^2 + \mu y.$$  

In this case the system has the cyclic first integral

$$p_x = \text{const},$$

and after integration of the equation we get for $p_y$ that

$$p_y = -c - \mu t, \quad c = \text{const}.$$  

The evolution of the other variables is described by the system of equations

$$\dot{p}_\varphi = mk_1 k_2, \quad \dot{\varphi} = I p_\varphi, \quad \dot{x} = mk_1 \cos \varphi, \quad \dot{y} = mk_1 \sin \varphi,$$  

$$k_1 = p_x \cos \varphi - (c + \mu t) \sin \varphi, \quad k_2 = p_x \sin \varphi + (\mu t + c) \cos \varphi.$$  

(56)

For $p_x = 0$ this system has the equilibrium state

$$p_\varphi = 0, \quad \varphi = 0,$$

in which the skate is at rest, with horizontally directed knife edge. On the other hand, the ‘hidden’ parameter

$$\lambda = p_x \sin \varphi - p_y \cos \varphi = \mu t + c$$

increases monotonically with time.

To understand the physical meaning of the initial value of $\lambda$, we consider the realization of this system in vakonomic mechanics, as indicated by Kozlov in [90].

We represent the kinetic energy of an elliptic plate in an ideal fluid as

$$T = \frac{1}{2} (a_1 v_1^2 + a_2 v_2^2 + b \omega^2),$$

where $\omega$ and $\mathbf{v} = (v_1, v_2)$ are the angular velocity and the translational velocity of the plate, respectively, and $a_1$, $a_2$, and $b$ are the added masses.
We consider the limit as one semiaxis of the plate tends to zero (so that \(a_2 \to \infty\)) while \(a_1\) and \(b\) remain finite (see [90] for details). Then in a neighbourhood of zero \(v_2\) varies with amplitude decreasing as \(a_2\) increases, and in the limit we obtain the constraint \(v_2 = 0\). That is, in the presence of gravity we get a system equivalent to the ice skate problem considered earlier.

Note that the kinetic energy remains finite if the velocity \(v_2\) has the order of \(a^{-1}\). The proportionality constant is equal to the initial value of the hidden parameter \(\lambda\) (see [93] for details).

Remark 15. In the general case the behaviour of an elliptic plate in an ideal fluid in a gravity field \((U = \mu y)\) is described by Kirchhoff’s equations

\[
\begin{align*}
& a_1 \dot{v}_1 = a_2 \omega v_2 - \mu \sin \varphi, \quad a_2 \dot{v}_2 = -a_1 \omega v_1 - \mu \cos \varphi, \quad b \dot{\omega} = (a_1 - a_2)v_1 v_2, \\
& \dot{\varphi} = \omega, \quad \dot{x} = v_1 \cos \varphi - v_2 \sin \varphi, \quad \dot{y} = v_1 \sin \varphi + v_2 \cos \varphi
\end{align*}
\]

(for instance, see [43]), where \(\varphi\) is the rotation angle, and \(x\) and \(y\) are the coordinates of the centre of mass. These equations have the additional integral

\[
p_x = a_1 v_1 \cos \varphi - a_2 v_2 \sin \varphi.
\]

Furthermore,

\[
p_y = a_1 v_1 \sin \varphi + a_2 v_2 \cos \varphi = -\mu t - c, \quad c = \text{const}.
\]

This yields the equations

\[
\begin{align*}
\ddot{\varphi} &= \frac{a_2 - a_1}{b a_1 a_2} k_1 k_2, \quad \dot{x} = \frac{k_1}{a_1} \cos \varphi + \frac{k_2}{a_2} \sin \varphi, \quad \dot{y} = \frac{k_1}{a_1} \sin \varphi - \frac{k_2}{a_2} \cos \varphi,
\end{align*}
\]

which in the limit as \(a_2 \to \infty\) reduce to (56) up to a change of variables. The asymptotic behaviour of this system was considered in [36].

Figure 9. The trajectory of the centre of mass of the knife edge for fixed \(m = 1, I = 1, \mu = 1, p_x = 0.4\) and initial conditions \(p_\varphi(0) = 0, \varphi(0) = 0, x(0) = 0,\) and \(y(0) = 0,\) versus different values of the constant \(c.\)
In connection with this realization, a special motion of the skate was considered in [65]: at the initial moment of time the knife edge is horizontal ($\varphi = 0$), $p_\varphi = 0$, and $p_x \neq 0$ (so that the initial velocity is also horizontally directed). It was shown that a ‘floating up’ effect is observed in this case: the body goes up above its initial position (see Fig. 9).

Remark 16. Note that in the non-holonomic setting the skate on an inclined plane does not slide down, but displays a horizontal drift. Furthermore, if the translational velocity vanishes at the initial moment of time ($v_1(0) = 0$), then the skate draws a cycloid [110].


4.1. Equations of motion. As in previous sections, we take a system whose configuration space is an $n$-dimensional manifold $\mathcal{N} = \{ q = (q_1, \ldots, q_n) \}$, when the generalized velocities $\dot{q} = (\dot{q}^1, \ldots, \dot{q}^n)$ satisfy the constraint equations (4), so that the trajectory in the tangent bundle lies in the distribution $\mathcal{D} \subset T\mathcal{N}$:

$$\mathcal{D} = \{(q, \dot{q}) \mid a_i^\mu(q)\dot{q}^i = 0, \mu = 1, \ldots, k \},$$

where here and below we sum over repeated indices.

In addition, the system is characterized by the Lagrangian function $L(q, \dot{q})$, which is calculated without taking the constraints into account (that is, it is defined everywhere in $T\mathcal{N}$) and satisfies the non-degeneracy condition

$$\det \left| \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right| \neq 0.$$

The equations of motion of the system that describe its dynamics (that is, a vector field on $\mathcal{M}$) are determined by the d’Alembert–Lagrange principle, which in its modern formulation can be written as follows [3], [48]:

for a system under the action of potential forces, at each moment of time an admissible smooth path $(q(t), \dot{q}(t))$ satisfies the system of equations

$$a_i^\mu(q)\dot{q}^i = 0,$$

$$\left[ \left( \frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} \right] \xi^i = 0, \quad a_i^\mu(q)\xi^i = 0, \quad \mu = 1, \ldots, k, \quad (57)$$

for all possible variations $\xi = (\xi^1, \ldots, \xi^n)$.

Remark 17. The addition of non-potential external forces $F = (F_1, \ldots, F_n)$ implies that the right-hand side of the first equation in (57) must be equal to $F_i \xi^i$.

Remark 18. Although the expression in square brackets in (57) is the variational derivative of the Lagrangian, when the constraints are non-holonomic, a solution $(q(t), \dot{q}(t))$ of these equations is not a conditional extremal of the functional
\[ f_{t_1}^{t_2} L dt. \] The reason is that, as already noted (see §2.1), a vector field of variations of curves \( \xi(q) \) which preserves the constraints must satisfy the equation (8) instead of the last equation in (57):

\[ \hat{\xi}(f^\mu(q, \dot{q})) = a_\mu^i \dot{\xi}^i + a_\mu^i \ddot{\xi}^i = 0, \quad \dot{\alpha}_i^\mu = \frac{\partial a_i^\mu}{\partial q^k} \dot{q}^k, \quad \xi^i = \frac{\partial \xi^i}{\partial \dot{q}^k} \dot{q}^k, \]

where \( \hat{\xi} \) is a lift of the vector field to \( T \mathcal{N} \).

The mechanical meaning of this principle is as follows: for all virtual displacements \( \delta q = \varepsilon \xi, \varepsilon \ll 1 \), admitted by the constraints the work of the reaction forces of the constraints must be equal to zero. In this connection the principle is also called the condition of ideal constraints. In physical applications the equations (57) can be obtained by passing to the limit in the system describing a rigid body rolling on a surface in the presence of viscous friction ([79], [81], [91], [110]) as the coefficient of friction goes to infinity. Such a surface is sometimes said to be absolutely rough, and the rolling model is said to be non-holonomic. We remark that such passages to the limit can result in a variety of different non-holonomic constraints, depending on the type of the physical surface [45].

To find explicit equations of motion of the system, we differentiate the constraint equations with respect to time and solve (57) using undetermined multipliers:

\[ \left( \frac{\partial L}{\partial \dot{q}^i} \right)^* - \frac{\partial L}{\partial q^i} = \lambda_\mu a_\mu^i, \quad (a_\mu^i \dot{q}^1)^* = 0. \]

These equations form a linear system in the unknowns \( \dot{q} = (\dot{q}^1, \ldots, \dot{q}^n) \) and \( \lambda = (\lambda_1, \ldots, \lambda_k) \). It can be represented in matrix form as

\[ B \dot{q} - A^T \lambda = b, \quad A \dot{q} = \beta, \]

\[ B = \left\| \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \right\|, \quad A = \| a_\mu^i \|, \]

where the components of the vectors \( b = (b_1, \ldots, b_n) \) and \( \beta = (\beta^1, \ldots, \beta^k) \) are functions of \( q \) and \( \dot{q} \) and can explicitly be expressed as

\[ b_i = \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i}, \quad \beta^\mu = -\frac{\partial a_i^\mu}{\partial q^j} \dot{q}^i \dot{q}^j. \]

It follows from (59) and the non-degeneracy that the undetermined multipliers can be found explicitly as functions of \( q \) and \( \dot{q} \):

\[ \lambda = (AB^{-1} A^T)^{-1}(\beta - AB^{-1} b). \]

Then we obtain the equations of motion explicitly in the form

\[ \ddot{q} = B^{-1} (b + A^T \lambda), \]

defined everywhere on \( T \mathcal{N} \). As a consequence of such a definition, we have the following natural result.
Proposition 4. The functions defining the constraint equations (4) are first integrals of the system (60):

\[ f^\mu(q, \dot{q}) = a^\mu_i(q)\dot{q}^i = \text{const.} \] (61)

Thus, the system (60) has a well-defined restriction to \( \mathcal{D} \) and defines the required vector field.

We see that in this case the undetermined multipliers have a meaning quite different from that of the undetermined multipliers introduced above (see §2). Here they are the reactions of constraints, and we need not specify any additional initial conditions to determine them. Hence, this mechanical system with (non-holonomic) constraints satisfies the classical principle of determinacy: a unique motion of the system corresponds to any initial position and velocity \((q, \dot{q}) \subset \mathcal{D}\). This also shows that the phase space of our non-holonomic system coincides with \( \mathcal{D} \).

4.2. The law of conservation of energy. It is straightforward to show that the system (60) restricted to \( \mathcal{D} \) has an energy integral:

\[ E = \left( \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i - L \right) \bigg|_\mathcal{D}. \] (62)

In fact, differentiating the expression in brackets with respect to time and taking (58) into account, we get that

\[ \left( \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i - L \right) = \left[ \left( \frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} \right] \dot{q}^i = \lambda_\mu c^\mu, \]

where the \( c^\mu \) are the constant values of the integrals (61). On \( \mathcal{D} \) we have \( c^\mu = 0 \), \( \mu = 1, \ldots, k \), which yields the required result.

Remark 19. We see that for inhomogeneous constraints the energy integral in the form (62) is not a constant, but nevertheless, in some cases the system has a certain generalization of the energy integral [52].

4.3. A non-holonomic basis and a pseudo-Hamiltonian representation. We choose new (local) coordinates parameterizing \( T\mathcal{N} \), the same as we used in §3.1. Recall that to do this we must define a special (non-holonomic) basis of vector fields on \( \mathcal{N} \) such that the fields

\[ \tau_\alpha = \tau^i_\alpha \frac{\partial}{\partial q^i}, \quad \alpha = 1, \ldots, n - k, \]

are everywhere tangent to the subspaces \( \mathcal{D}_q \) determined by the constraints (4), and the fields

\[ n_\mu = n^i_\mu \frac{\partial}{\partial q^i}, \quad \mu = 1, \ldots, k, \]

are everywhere transversal to \( \mathcal{D}_q \), so that

\[ a^\mu_i \tau^i_\alpha \equiv 0, \quad \det \|\Gamma^\mu_\nu\| \neq 0, \quad \Gamma^\mu_\nu = a^\mu_i n^i_\nu. \] (63)
In each tangent space $T \mathcal{N}_q$ we make the change of generalized velocities

$$\dot{q}^i = \omega^\alpha \tau^i_\alpha + w^\mu n^i_\mu$$

(in mechanics these new variables $\omega = (\omega^1, \ldots, \omega^{n-k})$, $w = (w^1, \ldots, w^k)$ are called quasi-velocities of the system). We parameterize $T \mathcal{N}$ by the variables $(q, \omega, w)$. Equations (57) and the constraints can now be represented in the form

$$\left[ \left( \frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} \right] \tau^i_\alpha = 0, \quad w^\mu = 0, \quad \alpha = 1, \ldots, n-k, \quad \mu = 1, \ldots, k. \quad (64)$$

We can show that for each vector field $\xi(q) = (\xi^1(q), \ldots, \xi^n(q))$ on $\mathcal{N}$ we have the natural relation

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}^i} \xi^i \right) - \dot{\xi}(L) = \left[ \left( \frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i} \right] \xi^i, \quad (65)$$

where $\dot{\xi}$ is a lift of the vector field $\xi$ to the tangent bundle $T \mathcal{N}$.

After the substitution (64) the function $F(q, \dot{q})$ on $T \mathcal{N}$ will be denoted by $F(q, \omega, w) = F(q, \dot{q})|_{q=\omega^\alpha \tau^i_\alpha + w^\mu n^i_\mu}$. By the rule for differentiating composite functions,

$$\frac{\partial \tilde{F}}{\partial q^i} = \frac{\partial F}{\partial \dot{q}^i} + \frac{\partial F}{\partial \dot{q}^k} \left( \omega^\alpha \frac{\partial \tau^k_\alpha}{\partial \dot{q}^i} + w^\mu \frac{\partial n^k_\mu}{\partial \dot{q}^i} \right),$$

$$\frac{\partial \tilde{F}}{\partial \omega^\alpha} = \frac{\partial F}{\partial \dot{q}^i} \tau^i_\alpha, \quad \frac{\partial \tilde{F}}{\partial w^\mu} = \frac{\partial F}{\partial \dot{q}^i} n^i_\mu. \quad (66)$$

Moreover, for each function $\tilde{F}(q, \omega, w)$ we denote its restriction to the constraints by $\tilde{F}^\ast(q, \omega) = \tilde{F}(q, \omega, w)|_{w=0}$.

Using (65) and the differentiation rules (66), we can now write the equations of dynamics for the system as

$$\frac{d}{dt} \left( \frac{\partial \tilde{L}^\ast}{\partial \omega^\alpha} \right) - (\tilde{\tau}^\ast_\alpha(\tilde{L}))^\ast = 0, \quad \alpha = 1, \ldots, n-k. \quad (67)$$

Note that this is a universal form of the equations of motion of a system governed by the d’Alembert–Lagrange principle, be it constrained or unconstrained. For instance, in the special case when there are no constraints, we obtain the well-known dynamics equations in quasi-coordinates (for instance, see [42]), as pointed out by Poincaré [114] in the case where the configuration space is a Lie group.

To write them explicitly we use the formula (6) for a lift of a vector field in a non-holonomic basis. We have

$$\tilde{\tau}^\alpha(F(q, \omega, w)) = \tau^i_\alpha \frac{\partial F}{\partial \dot{q}^i} + c^\gamma_\beta \omega^\beta \frac{\partial F}{\partial \omega^\gamma} + \tilde{c}^\mu_\beta \omega^\beta \frac{\partial F}{\partial \omega^\mu} + d^\beta_\mu w^\mu \frac{\partial F}{\partial \omega^\beta} + \tilde{d}^\mu_\alpha w^\mu \frac{\partial F}{\partial \omega^\mu},$$

$$[\tau^\alpha, \tau^\beta] = c^\gamma_\alpha(q) \tau^\gamma + \tilde{c}^\mu_\alpha(q) n^\mu, \quad [\tau^\alpha, n^\mu] = d^\beta_\alpha(q) \tau^\beta + \tilde{d}^\nu_\alpha(q) n^\nu,$$
where \([\cdot, \cdot]\) is the Lie bracket of vector fields. Finally, we obtain the equations of motion on \(D\) in the form
\[
\left(\frac{\partial \tilde{L}^*}{\partial \omega^\alpha}\right) - \tau^i_\alpha \frac{\partial \tilde{L}^*}{\partial q^i} = -c^\gamma_{\alpha\beta} \frac{\partial \tilde{L}^*}{\partial \omega^\gamma} \omega^\beta - \bar{c}^\mu_{\alpha\beta} \left(\frac{\partial \tilde{L}}{\partial w^\mu}\right) \omega^\beta, \quad \dot{q}^i = \tau^i_\alpha \omega^\alpha.
\] (67)

Thus, in contrast to extremals (§§2 and 3), trajectories of the non-holonomic system (67) depend not only on the restriction of the Lagrange function \(eL^*\) to \(D\), but also on the restrictions of the derivatives \((\partial eL/\partial w^\mu)^*\).

Now we show that equations (67) can be represented in pseudo-Hamiltonian form. To do this we use a Legendre transform of the form
\[
M^\alpha = \frac{\partial eL^*}{\partial \omega^\alpha}, \quad H(q, M) = (\omega^\alpha M^\alpha - \tilde{L}^*)|_{\omega=M}, \quad \alpha = 1, \ldots, n - k,
\]
where \(M = (M_1, \ldots, M_{n-k})\). In the variables \((q, M)\) the system (67) has the form
\[
\dot{M}^\alpha = -\tau^i_\alpha \frac{\partial H}{\partial q^i} - c^\gamma_{\alpha\beta} M^\gamma \frac{\partial H}{\partial M^\beta} - \bar{c}^\mu_{\alpha\beta} p^\mu \frac{\partial H}{\partial M^\beta}, \quad \dot{q}^i = \tau^i_\alpha \frac{\partial H}{\partial M^\alpha},
\]
where \(p_\mu(q, M) = \left(\frac{\partial \tilde{L}^*}{\partial w^\mu}\right)|_{\omega=M} \). These equations were also derived in [125] using another method.

Denoting the system of new variables \((q^1, \ldots, q^n, M_1, \ldots, M_{n-k})\) by \(x\), we can write these equations as
\[
\dot{x}_i = J_{ij}(x) \frac{\partial H}{\partial x_j}, \quad i = 1, \ldots, 2n - k,
\] (68)
where the \(J_{ij}(x)\) are the components of the skew-symmetric matrix
\[
J = \begin{pmatrix} 0 & T^T \\ -T & -\Pi \end{pmatrix}, \quad \Pi(q, M) = \|c^\gamma_{\alpha\beta}(q) M^\gamma + \bar{c}^\mu_{\alpha\beta}(q) p_\mu(q, M)\|, \quad T(q) = \|\tau^i_\alpha(q)\|.
\]
The matrix \(J\) satisfies the Jacobi identity if and only if the constraints are holonomic. Hence, although non-holonomic systems have the pseudo-Hamiltonian representation (68), their dynamics is essentially different from that of Hamiltonian systems. Corresponding examples can be found in [12], [33], and [51].

5. A rigid body rolling on a plane: non-holonomic equations

5.1. Equations of motion. We consider the problem of a rigid body rolling without slipping on a horizontal plane in a gravity field (see Fig. 4).

As in §3.2, rolling without slipping means that the velocity of the body vanishes at the point of contact with the plane:
\[
f = v + \omega \times r = 0,
\] (69)
where $\omega$ and $v$ are the angular velocity and the velocity of the centre of mass of the body with respect to the moving reference system, and $r$ is the radius vector from the centre of mass to the point of contact. For a convex body $r$ can be expressed in terms of the normal $\gamma$ at the point of contact by using the formula for the Gaussian map

$$\gamma = -\frac{\nabla f(r)}{|\nabla f(r)|},$$

where $f(r) = 0$ is an equation describing the surface of the body. In what follows we assume that $r = r(\gamma)$.

In view of the previous section, we obtain from the d’Alembert–Lagrange principle the equations of motion in the following form (see [49] for details):

$$\dot{M} = M \times \omega + m \dot{r} \times (\omega \times r) + m g r \times \gamma, \quad \dot{\gamma} = \gamma \times \omega, \quad (70)$$

$$\dot{\alpha} = \alpha \times \omega, \quad \dot{\beta} = \beta \times \omega, \quad \dot{x} = (\alpha, \dot{r}), \quad \dot{y} = (\beta, \dot{r}), \quad (71)$$

where $M$ is the angular momentum of the body with respect to the point of contact $P$, referred to the moving frame $C x_1 x_2 x_3$. The other variables (as in §3.2) are the basis vectors $\alpha, \beta, \gamma$ of the space-fixed coordinate system and the radius vector $R_P = (x, y, 0)$ of the point of contact.

The angular velocity $\omega$ can be expressed in terms of $M$ and $\gamma$ using the relation

$$M = \mathbf{I} \omega, \quad \mathbf{I} = \mathbf{I} + m r^2 - m r \otimes r,$$

where $\mathbf{I} = \text{diag}(I_1, I_2, I_3)$ is the inertia tensor of the body.

In this way we decouple the closed system of equations (70) describing the evolution of $M$ and $\gamma$. Since the constraints (69) are homogeneous and linear in the velocities, the system (70) possesses an energy integral (see §4.2) and a geometric integral:

$$E = \frac{1}{2} (M, \omega) - mg(r, \gamma), \quad \gamma^2 = 1.$$  

In order for the system (70) to be integrable by the Euler–Jacobi theorem, we need to find two additional first integrals and an invariant measure.

In contrast to equations (37) considered above, (70) does not have an invariant measure with a continuous density in the general case, and cannot be represented in Hamiltonian form (see [39] for details). In the next subsection we describe conditions for the existence of an invariant measure; for particular non-holonomic systems these conditions impose certain restrictions on their dynamical and geometric parameters.

5.2. Necessary conditions for the existence of a continuous invariant measure. A function $\rho$ is called the density of an invariant measure of the system (70) if it satisfies Liouville’s equation

$$\text{div}(\rho(x)u(x)) = 0, \quad (72)$$

where $x = (M, \gamma)$ and $u(x)$ is the vector field defined by (70). Here the function $\rho$ is assumed to be differentiable; more general questions about non-differentiable measures are discussed in [99].
If $\rho$ is a continuous function of the coordinates $x$, then for any domain $\Gamma_t$ in the phase space which is transformed by the flow of the system, its volume relative to the density $\rho$ is conserved:

$$\int_{\Gamma_t} \rho(x) \, dx = \text{const}.$$ 

Let us consider conditions under which a solution of (72) can be found explicitly. For this we represent the equations of motion as

$$\dot{M} = \tilde{M}(\gamma, M), \quad \dot{\gamma} = \tilde{\gamma}(\gamma, M).$$

Then Liouville’s equation has the form

$$\frac{1}{\rho} \left(\left(\frac{\partial \rho}{\partial \gamma}, \tilde{\gamma}\right) + \left(\frac{\partial \rho}{\partial M}, \tilde{M}\right)\right) = -\sum_{i=1}^{3} \frac{\partial \tilde{M}_i}{\partial M_i},$$

where we have taken into account that $\sum_i \frac{\partial \tilde{\gamma}_i}{\partial \gamma_i} = 0$. As follows from (70), the right-hand side of the above equation is a function which is linear and homogeneous in $M$, and it can be represented in the form

$$\sum_{i=1}^{3} \frac{\partial \tilde{M}_i}{\partial M_i} = (\chi(\gamma), \omega).$$

In view of this relation we can rewrite the equation for the invariant measure as

$$\left(\frac{\partial \log \rho}{\partial \gamma}, \tilde{\gamma}\right) + \left(\frac{\partial \log \rho}{\partial M}, \tilde{M}\right) + (\chi(\gamma), \omega) = 0. \quad (73)$$

When $\rho = \rho(\gamma)$, (73) reduces to

$$\left(\frac{\partial}{\partial \gamma} \log \rho(\gamma) \times \gamma + \chi(\gamma), \omega\right) = 0.$$

Since this must hold for arbitrary $\omega$, it follows that

$$\gamma \times \frac{\partial}{\partial \gamma} \log \rho(\gamma) = \chi(\gamma).$$

This relation has the following consequence.

**Proposition 5.** If a system of equations (70) has an invariant measure $\rho(\gamma) \, dM \, d\gamma$, then the function $\chi(\gamma)$ satisfies the relations

$$(\chi(\gamma), \gamma) = 0 \quad \text{and} \quad \sum_{i=1}^{3} \frac{\partial \chi_i}{\partial \gamma_i} = 0. \quad (74)$$
5.3. Rolling of a triaxial ellipsoid. Suppose that a body has the shape of a triaxial ellipsoid, with a surface given by the equation
\[(r, B^{-1}r) = 1, \quad B = \text{diag}(a_1^2, a_2^2, a_3^2),\]
where \(a_1, a_2,\) and \(a_3\) are the semiaxes of the ellipsoid. Inverting the Gauss map, we obtain the explicit expression
\[r = -\frac{B\gamma}{\sqrt{(B\gamma, \gamma)}}.\]

Using the criterion (74), we find conditions for the existence of a continuous invariant measure:
\[(a_2^2 - a_3^2)(ma_1^2I_1 - I_2I_3) + (a_3^2 - a_1^2)(ma_2^2I_2 - I_1I_3)\]
\[+ (a_1^2 - a_2^2)(ma_3^2I_3 - I_1I_2) = 0,\]
\[a_1^2a_2^2a_3^2(I_1(a_2^2 - a_3^2) + I_2(a_3^2 - a_1^2) + I_3(a_1^2 - a_2^2)) = 0.\]

Remark 20. In the absence of gravity, when \(g = 0,\) we can show (by analogy with [51]) that condition (75) is not only sufficient but also necessary for the existence of a continuous invariant measure.

We can parameterize the second relation in (75) as follows:
\[I_i = \mu + \nu a_i^2, \quad i = 1, 2, 3.\]

Then from the first relation we obtain
\[\nu(\nu - m)(a_1^2 - a_3^2)(a_2^2 - a_3^2)(a_2^2 - a_1^2) = 0.\]
This equality holds in the following cases.
- **An ellipsoid with spherical moment of inertia** (\(\nu = 0\)): in this case an invariant measure was found by Yaroshchuk [132].
- **An ellipsoid of revolution**: this case was considered by Chaplygin [61].
- **An ellipsoid with special moments of inertia** (\(\nu = m\)): in this case an invariant measure was found by Yaroshchuk [133].

Condition (76) coincides with the condition for the existence of vertical permanent rotations found in [82], which are fixed points of the system (70):
\[\gamma = \chi = \text{const}, \quad \omega = k\chi, \quad k^2 = \frac{mg}{(\nu - m)(\chi, B\chi)}, \quad \chi^2 = 1.\]

It was shown in [41] that at these equilibria (except for the three cases above) \(\text{Tr } A^{(L)} \neq 0,\) where \(A^{(L)}\) is the linearization matrix. Hence, in accordance with [92], the equilibria (77) are asymptotically stable in the general case, which explains the absence of a continuous invariant measure.

If condition (76) fails to hold (as for a homogeneous ellipsoid, for example), then the question remains open of possible ‘dissipation’ effects (for example, attractors such as fixed points, limit cycles, or strange attractors).
If the body has the shape of an elliptic disk \((a_3 = 0)\), then the second relation in (76) holds identically, while the first, in view of the equality \(I_3 = I_1 + I_2\), implies a restriction on the distribution of mass:

\[a_1^2 I_1^2 - a_2^2 I_2^2 + ma_1^2 a_2^2 (I_1 - I_2) = 0.\]

In this case Liouville’s equation has the solution

\[\rho(\gamma) = \left(\det \tilde{I}\right)^{-1/4}.\]

This case when there is an invariant measure was not known previously and is presented here for the first time. In addition, it follows from the Schwarzschild–Littlewood theorem (see [94] and [100] for details) that under almost all initial conditions the disk does not fall on the plane in this case.

6. Short sketch of the development of non-holonomic mechanics in recent decades

The paper [53] is a comprehensive survey of the development of non-holonomic mechanics covering a period up until the publication of the book [110] of Neimark and Fufaev. For this reason we consider in greater detail the main results obtained from the 1980s to the present day which are directly related to our survey.

Several surveys on non-holonomic mechanics (for instance, [106] and [113]) have been primarily devoted to the elaboration of specific questions in the theory and did not contain any new statements of problems. In our opinion the development of non-holonomic mechanics should mainly be encouraged by elaborating known approaches and working out new methods for the solution of concrete problems rather than by creating an abstract theory. This is why in this brief survey we are mostly concerned with various non-holonomic problems which in some sense can be regarded as models, and with results obtained in investigating them.

1. The problem of a balanced, dynamically asymmetric ball rolling without slipping on a plane plays a special role in analyzing specific properties of non-holonomic systems. It was first considered by Chaplygin [62], who derived the equations of motion and reduced them to quadratures. In addition, he gave a geometric interpretation of the motion of such a ball, which was subsequently called a ‘Chaplygin ball’.

From the modern point of view, the problem of a Chaplygin ball reduces to an investigation of a reduced system describing the evolution of the angular momentum \(M\) of the ball and the normal \(\gamma\) to the plane at the point of contact. In this case a non-singular level set of the first integrals is diffeomorphic to a 2-torus. For known solutions \(M(t)\) and \(\gamma(t)\), the orientation of the ball and the dynamics of the point of contact can be recovered by quadratures.

A topological analysis of the reduced system was carried out in [85]. It turned out that in fact the bifurcation diagram for this case coincides with the diagram for the Euler case. On the other hand, the reduced system can be represented in Hamiltonian form only after rescaling time (that is, it is conformally Hamiltonian). An explicit conformally Hamiltonian form was found by Borisov and Mamaev [37], and it was shown in [38] that the reduced system is not Hamiltonian if time is not rescaled: different periods of revolution on resonance tori are an obstruction to this.
Subsequently, Duistermaat [66], after studying Chaplygin’s paper [62] in detail and in a certain sense reformulating it in more modern terms, remarked that Chaplygin’s results on the behaviour of the point of contact were not accurate, or at least were lacking rigorous proof. In [33] methods of topological analysis and analytic results due to Bohl and Weyl were used to derive conditions for the trajectory of the point of contact to be bounded or unbounded.

Two integrable (in the Euler–Jacobi sense) generalizations are known in the problem of a rolling Chaplygin ball: the first involves the addition of the Brun potential field\(^4\) and was obtained by Kozlov [92] in 1985, and the second involves the addition of a gyrostatic term and was pointed out by Markeev [105] in 1986. In both cases the equations of motion can still be represented in a conformally Hamiltonian form. Nevertheless, these systems have not yet been explicitly integrated: after rescaling time we obtain a Hamiltonian system on a sphere with a quadratic integral that contains terms linear in the momenta. No general method has been developed for integrating such systems. In the particular case when the area constant is equal to zero and there exist no such terms, only the case involving the Brun field has been integrated [68]. A topological study of the case found by Markeev was carried out in [109]. We recall that when the area constant is non-zero, the case of the classical Chaplygin ball can be integrated by means of a non-trivial transformation which reduces the problem to the case of a vanishing area constant. This transformation was considered by Chaplygin and its geometric meaning was considered in [30].

Besides the problem considered by Chaplygin, other integrable systems related to a rolling ball have been found. We consider them in more detail. In 1988 Fedorov [69] showed that the problem of a body moving in a spherical suspension is integrable by quadratures (see Fig. 10, (a)). This system was explicitly integrated in [30] using Chaplygin’s non-trivial transformation. Another problem, concerning a straight-line motion of a Chaplygin ball, was considered in [128] (1988); such a motion can be realized with the help of absolutely smooth walls (see Fig. 10, (b)). This system was explicitly integrated only in the case when the area constant

\(^4\)The Brun field is connected with the case when each part of the ball is attracted by the plane with a force proportional to the distance to the plane. This field can be regarded as a quadrupole approximation in the expansion of the Newton potential.
is equal to zero [30]. Although both systems are very close to the problem of a Chaplygin ball, no explicit representation of their equations of motion has been found in a conformally Hamiltonian form.

In [24] Borisov and Fedorov (1995) considered the problem of the motion of a rigid body in a spherical suspension (see Fig. 11, (a)). A representation of this system in a conformally Hamiltonian form was recently found [121], but an explicit integration was carried out only when the area constant is zero [25], [56]. A topological analysis was carried out in [47]. It turns out that in this problem the foliation by invariant tori is equivalent to the foliation in Clebsh’s integrable system in rigid body dynamics.

In 1986 Veselova [129] considered the motion of a rigid body with a fixed point subject to a non-integrable constraint: the projection of the angular velocity on a space-fixed axis is equal to zero. Veselova integrated this system explicitly, and an isomorphism with the problem of a Chaplygin ball was found in [22]. Two problems leading to this system are known: in the first case a body is attached to a spherical shell and small wheels (disks) with space-fixed axes are in contact with the shell (see Fig. 11, (b)); in the second case a Chaplygin ball rolls on a plane under the condition that there is no sliding nor spinning at the point of contact (see [45] and [51] for details). Various generalizations of Veselova’s system were considered in [44].

2. Let us now turn to other results on non-holonomic systems which have a general theoretical nature. Tatarinov [120] showed in 1988 that two-dimensional integral surfaces other than tori can occur in non-holonomic mechanics. At the same time one of the dynamical effects obstructing integrability and connected with the absence of a continuous invariant measure was pointed out by Kozlov (1985) in [92]. In addition, Yaroshchuk found in 1992 and 1995 new cases when there is a smooth invariant measure depending on the position variables [132], [133]. In [102] (1993)

---

5Interestingly enough, the modern explanations of the theory of bicycles in [86] and [107] are not based on the gyroscopic effect alone, but are also connected with such non-trivial effects as the asymptotic stability of non-holonomic systems.
Kozlov and Yaroshchuk posed the question of multidimensional non-holonomic systems (the Euler–Poincaré–Suslov equations) and of the existence of an invariant measure for such systems.

**Remark 21.** We remark that not many integrable multidimensional problems are known in non-holonomic dynamics. The reason is that the \((L-A)\)-pair method is inefficient in this case. We mention only the papers [70]–[72], where Veselova’s system (for physical parameters) was represented in Hamiltonian form. The question of a multidimensional (and in particular, four-dimensional) Chaplygin ball has not yet been resolved.

In 1994 van der Schaft and Maschke [125] represented the equations of non-holonomic mechanics in skew-symmetric form. However, this is of no use when it comes to representing the equations of motion in Hamiltonian form, because the corresponding skew-symmetric tensor satisfies the Jacobi identity only for holonomic constraints (see §4.3 for details).

At the same time a more abstract line of research, connected with geometric formalism in non-holonomic mechanics and combining geometric concepts from Lie algebras and differential geometry, was also developing, mostly outside Russia. We note that only a few concrete problems were considered in this direction, and general results concerning mostly reduction theory were discussed in [49].

**3.** The next stage in the development of non-holonomic mechanics was connected with the papers [39] and [54] by Borisov, Mamaev, and Kilin, who showed that non-holonomic systems exhibit much more diverse dynamical behaviour than Hamiltonian systems. The groundwork here is due to Kozlov [92].

The variety of behaviours (which was called the hierarchy of dynamics) is governed by the existence or absence of various tensor invariants (conservation laws): invariant measures, first integrals, Poisson structure, and so on.

**Remark 22.** The method of a reducing multiplier, considered originally by Chaplygin in [63] and later developed in [22] and [45], proved to be very efficient in determining the Poisson structure of equations in non-holonomic mechanics. Nevertheless, in [29] and [46] explicit algebraic investigations were used to find Poisson brackets (non-linear in the momenta) whose existence could not be explained using Chaplygin’s method. Note that there exists no smooth invariant measure in these examples.

It turned out that these Poisson brackets could be found based on a certain algorithm, called the Hojman construction (see [14] for details). Poisson brackets obtained using the Hojman construction are usually degenerate, do not have a continuous (or even singular) invariant measure, and do not necessarily have global Casimir functions. Furthermore, the behaviour of the system can be strongly different from Hamiltonian behaviour (in particular, limit cycles are possible).

The most important recent results in non-holonomic mechanics have been proved using methods based on topological and qualitative investigations. We can mention several non-holonomic systems for which this approach was most successful.

Of special note among new non-holonomic systems is the problem of the free motion of two bodies joined by a *non-holonomic hinge* [12], [10] (see Fig. 12). This is a generalization of Suslov’s problem on the motion of a body with a fixed point in
the presence of a non-holonomic constraint. Other generalizations were considered in [16]. It is an interesting feature of this problem that it reduces to analysis of an integrable geodesic flow on a solvable Lie group. Three types of integral surfaces occur in this problem: a sphere $S^2$, a torus $T^2$, and an oriented 2-dimensional surface of genus three $M^2_3$.

There are several examples of non-holonomic systems [21], [8] in which the phase space is foliated by 2-tori with limit cycles on some of them.

Besides tensor invariants, discrete symmetries (invertibilities) giving rise to non-trivial involutions (see Fig. 13) are essential for non-holonomic systems. Depending on the number of involutions [51], non-holonomic systems can possess both ordinary attractors (fixed points, limit cycles) [15] and strange attractors [40].

Various dynamical effects and strange attractors have been discovered in problems in non-holonomic mechanics by using approaches based on numerical investigation of the Poincaré map (bifurcation analysis, calculation of Lyapunov exponents). For instance, in addition to the non-holonomic rattleback model [40], other systems exhibiting the phenomenon of reversal have been found: the Chaplygin top [28] and the Suslov problem in a gravity field [11], [98]. A detailed numerical simulation of a rattleback was recently carried out in the paper [115]. Moreover, strange attractors such as the Lorenz attractor [74], the figure-eight attractor [28], and Feigenbaum-type attractors [11], [27], [26] have been found in the above-mentioned systems.

Nevertheless, the dynamics of a wheeled vehicle and its diverse versions remains poorly understood. Some preliminary results in this area were obtained in [9], [55], and [59]. We point out [34], where it was shown that the problem of an arbitrary wheeled vehicle is equivalent to a problem of an analogous vehicle with a wheel pair replaced by a weightless knife edge (skate).

In conclusion we comment on certain directions of research in non-holonomic mechanics which are being discussed in the literature.

1) The authors of several papers (for instance, see [122] and [123]) advocate an alternative method for finding and investigating equations in non-holonomic mechanics with the help of Moore–Penrose matrices. However, this approach is
equivalent to the classical method based on deriving equations with undetermined multipliers and thus cannot lead to new results.

2) Zung (for instance, see [134]) investigates normal forms and bifurcations of tori in general integrable systems, and he claims, in particular, that many of his methods can be applied to non-holomorphic systems. However, from a general standpoint this is equivalent to Lie’s approach, and moreover, no new integrable systems have been found in this way. A more general approach, developed specifically for the integration of multidimensional non-holonomic systems, was proposed by Kozlov in [95]. Unfortunately, no general theory of integration of multidimensional non-holonomic systems (analogous to the theory in Hamiltonian mechanics) has yet been developed.

3) In recent years, there has been a large amount of research on discrete versions of non-holonomic systems, but the variety of discretizations is so great that it is difficult to see any rational justification for such analyses. Unless discretization is treated as a goal in itself, one possible motivation for it may be to develop numerical methods for the integration of non-holonomic systems [73]. On the other hand, it is interesting that no author of such ‘non-holonomic integrators’ has demonstrated their advantages in the solution of complicated applied problems. Moreover, the very idea of this approach is in contradiction to the idea of a hierarchy of dynamics: according to the latter we must first find the whole set of tensor invariants, which depends on the dynamical and geometric parameters of the problem, and only then use specific methods of investigation, including numerical methods.

The authors acknowledge fruitful discussions with A.V. Bolsinov, A.A. Agrachev, Yu.L. Sachkov, and A.A. Kilin and their useful remarks. The authors are particularly grateful to V.V. Kozlov, who has read the manuscript and made quite a few useful comments.
Bibliography


A. V. Borisov, I. S. Mamaev, and I. A. Bizyaev


A. V. Borisov, I. S. Mamaev, and I. A. Bizyaev


Dynamical systems with non-integrable constraints


[129] Л. Е. Веселова, “Новые случаи интегрируемости уравнений движения твердого тела при наличии неголономной связи”, Геометрия, дифференциальные уравнения и механика (Москва, 1985) (В. В. Козлов,
Alexey V. Borisov
Udmurtian State University
E-mail: borisov@rcd.ru

Ivan S. Mamaev
Izhevsk State Technical University
E-mail: mamaev@rcd.ru

Ivan A. Bizyaev
National Research University
“Higher School of Economics”
E-mail: bizaev_90@mail.ru

Received 09/JUN/17
Translated by N. KRUZHILIN