Canonical form of Euler–Lagrange equations and gauge symmetries

2003 J. Phys. A: Math. Gen. 36 6587

(http://iopscience.iop.org/0305-4470/36/23/321)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) 6587–6609 PII: S0305-4470(03)60780-8

Canonical form of Euler–Lagrange equations and gauge symmetries

B Geyer[1], **D M Gitman**[2] **and I V Tyutin**[3]

[1] Naturwissenschaftlich-Theoretisches Zentrum und Institut für Theoretische Physik, Universität Leipzig, Leipzig, Germany
[2] Institute of Physics, University of Sao Paulo, Sao Paulo, Brazil
[3] Lebedev Physics Institute, Moscow, Russia

E-mail: geyer@itp.uni-leipzig.de, gitman@fma.if.usp.br and tyutin@lpi.ru

Received 12 March 2003
Published 29 May 2003
Online at stacks.iop.org/JPhysA/36/6587

Abstract

The structure of the Euler–Lagrange equations for a general Lagrangian theory (e.g. singular, with higher derivatives) is studied. For these equations we present a reduction procedure to the so-called canonical form. In the canonical form the equations are solved with respect to highest-order derivatives of nongauge coordinates, whereas gauge coordinates and their derivatives enter the right-hand sides of the equations as arbitrary functions of time. The reduction procedure reveals constraints in the Lagrangian formulation of singular systems and, in that respect, is similar to the Dirac procedure in the Hamiltonian formulation. Moreover, the reduction procedure allows one to reveal the gauge identities between the Euler–Lagrange equations. Thus, a constructive way of finding all the gauge generators within the Lagrangian formulation is presented. At the same time, it is proved that for local theories all the gauge generators are local in time operators.

PACS numbers: 03.70.+k, 11.10.Ef, 11.15.−q

1. Introduction

At present increasingly complicated gauge models are used in field and string theories. Generally a comprehensive analysis of their structure is not a simple task. In the Lagrangian formulation the problem includes, in particular, finding generators of gauge symmetries and their algebra, revealing the hidden structure of the equations of motion and so on. One ought to say that in the Hamiltonian formulation there exists a relatively well-developed scheme of constraint finding (Dirac procedure [1]) and reorganization [1–4]. The constraint structure can be, in principle, related to the symmetry properties of the initial gauge theory in the Lagrangian formulation [5]. However, in the general case, this relation cannot be considered as a constructive method of studying the Lagrangian symmetries (it is indirect and complicated). Moreover, the modern tendency is to avoid the non-covariant Hamiltonization step and to use the Lagrangian quantization [6] for constructing quantum theory. Such an approach incorporates all the Lagrangian structures (in particular, the total gauge algebra). That is why it seems important to develop a reduction procedure within the Lagrangian formulation—in a sense similar to the Dirac procedure in the Hamiltonian formulation—that may allow one to reveal the hidden structure of the Euler–Lagrange equations (ELE) of motion in a constructive manner and to find all the gauge identities and therefore the generators of all the gauge transformations. An idea of such a procedure was first mentioned in publications of the authors (DG and IT) [7, 8] (see also appendix C in [2]), but was not appropriately elaborated and some important points were not revealed.

In the present paper we return to this idea studying the structure of the ELE for a general Lagrangian theory (singular, with higher derivatives, and with external fields). In section 2 we introduce some notation and definitions. In section 3, we reduce the ELE of nonsingular theories to the so-called canonical form (in the canonical form the equations are solved with respect to highest-order derivatives of nongauge coordinates, whereas gauge coordinates and their derivatives enter the right-hand sides of the equations as arbitrary functions of time, see below). In section 4 we formulate the reduction procedure for the singular case. In a sense, the reduction procedure reveals constraints in the Lagrangian formulation of singular systems and, in that respect, is similar to the Dirac procedure in the Hamiltonian formulation. In section 5 we demonstrate how the reduction procedure reveals the gauge identities between the ELE. Thus, a constructive way of finding all the gauge generators within the Lagrangian formulation is presented. At the same time it is proved that for local theories all the gauge generators are local in time operators. In the appendix we collect some lemmas that are useful for our consideration.

2. General ELE

2.1. Notation, definitions and conventions

We consider a system with finite degrees of freedom (classical mechanics). These degrees of freedom are described by the generalized coordinates q^a, $a = 1, \ldots, n,$ which depend on the time t. The following notation is used:

$$\dot{q}^a = \frac{\mathrm{d}q^a}{\mathrm{d}t} \qquad \ddot{q}^a = \frac{\mathrm{d}^2 q^a}{\mathrm{d}t^2} \quad \ldots \qquad \text{or} \qquad q^{a[l]} = \frac{\mathrm{d}^l q^a}{\mathrm{d}t^l}$$
$$l = 0, 1, \ldots \qquad (q^{a[0]} = q^a). \tag{1}$$

The coordinates $q^a = q^{a[0]}$ are sometimes called velocities of zeroth order; the velocities $\dot{q}^a = q^{a[1]}$ are called velocities of first order; the accelerations $\ddot{q}^a = q^{a[2]}$ are called velocities of second order and so on. The space of all the velocities is often called the jet space, see [9].

As local functions (LF) we call those functions that are defined on the jet space and depend on the velocities $q^{a[l]}$ up to some finite orders $N_a \geqslant 0$ $(l \leqslant N_a)$. Further, we call N_a the order of the coordinate q^a in the LF. For the LF we use the following notation[4]:

$$F(q^a, \dot{q}^a, \ddot{q}^a, \ldots) = F(q^{a[0]}, q^{a[1]}, q^{a[2]}, \ldots) = F(q^{[l]}) \qquad q^{[l]} = (q^{a[l]}, 0 \leqslant l \leqslant N_a)$$
$$\text{or sometimes :} \quad F(q^{[l]}) = F(\cdots q^{a[N_a]}). \tag{2}$$

In the latter form, we indicate only the highest-order derivatives in the arguments of the LF.

[4] The functions F may depend on time explicitly, however, we do not include t in the arguments of the functions.

The following notation is often used: $[a]$ is the number of the indices a, namely, if $a = 1, \ldots, n$, then $[a] = n$. Similarly, suppose $F_a(\eta)$, $a = 1, \ldots, n$, are some functions, then $[F]$ is the number of these functions, $[F] = n$, etc. However, differently, writing $q^{a[l]}$ we denote by $[l]$ the order of the time derivatives, see (1).

On the jet space, we define local operators (LO) to be matrix operators $\hat{U}$ of the form

$$\hat{U}_{Aa} = \sum_{k=0}^{K} u^k_{Aa} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^k \tag{3}$$

where K is a finite number and u^k_{Aa} are some LF. The LO act on columns of LF f_a producing columns of LF $F_A = \hat{U}_{Aa} f_a$ as well. We define the transposed operator to $\hat{U}$ as

$$(\hat{U}^T)_{aA} = \sum_{k=0}^{K} \left(-\frac{\mathrm{d}}{\mathrm{d}t}\right)^k u^k_{Aa} = \sum_{k=0}^{K} (-1)^k \sum_{l=0}^{k} \binom{k}{l} u^{k[l]}_{Aa} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{k-l}. \tag{4}$$

Then the following relation holds true:

$$F_A \hat{U}_{Aa} f_a = [(\hat{U}^T)_{aA} F_A] f_a + \frac{\mathrm{d}}{\mathrm{d}t} Q \tag{5}$$

where F_A, f_a and Q are LF. The LO $\hat{U}_{ab}$ is symmetric (+) or skewsymmetric ($-$) whenever the relation $(\hat{U}^T)_{ab} = \pm\hat{U}_{ab}$ holds true.

Suppose a set of LF $F_A\left(\cdots q^{a[N_a^A]}\right)$, or a set of equations $F_A\left(\cdots q^{a[N_a^A]}\right) = 0$, are given. In the general case the orders N_a^A of the coordinates q^a in the functions F_A (in the equations $F_A = 0$) are different, i.e. these orders depend both on a and A. The number $\mathcal{N}_a = \max_A N_a^A$ is called the order of the coordinate q^a in the set of the functions F_A (in the set of the equations $F_A = 0$).

When a subset $F_{A'} = 0, A' \subset A$ has orders $\mathcal{N}'_a$ of the coordinates less than the corresponding orders of the complete set, namely, $\forall a : \mathcal{N}'_a < \mathcal{N}_a$, we call this subset the constraint equations.

Generally two sets of equations, $F_A(q^{[l]}) = 0$ and $f_\alpha(q^{[l]}) = 0$ are equivalent whenever they have the same set of solutions. In what follows we denote this fact as $F = 0 \Longleftrightarrow f = 0$.

Suppose that two sets of LF $F_A(q^{[l]})$ and $\chi_A(q^{[l]})$, $[F] = [\chi]$, are related by some LO,

$$F = \hat{U}\chi \qquad \chi = \hat{V}F \qquad \hat{U}\hat{V} = 1. \tag{6}$$

Then we call such functions equivalent and denote this fact as $F \sim \chi$. Obviously,

$$F \sim \chi \quad \Longrightarrow \quad F = 0 \quad \Longleftrightarrow \quad \chi = 0. \tag{7}$$

If (7) holds true, we will call the equations $F_A = 0$ and $f_\alpha = 0$ strong equivalent.

In what follows we often come across the case where

$$\chi_A = \begin{pmatrix} f_\alpha \\ 0_G \end{pmatrix} \qquad A = (\alpha, G) \qquad \forall G : 0_G \equiv 0. \tag{8}$$

Here the equivalence $F \sim \chi$ implies the equivalence of the equations $F = 0$ and $f = 0$ and the existence of the identities $\hat{V}_{GA} F_A \equiv 0$. Namely,

$$F \sim \chi \quad \Longrightarrow \quad \begin{cases} F = 0 \quad \Longleftrightarrow \quad f = 0 \\ \hat{V}_{GA} F_A \equiv 0. \end{cases} \tag{9}$$

2.2. ELE

In this section we restrict our consideration to the Lagrange functions L that are LF on the jet space, and depend on some external coordinates (fields) u^μ (we call the coordinates u^μ external ones in contrast to the coordinates q^a, which we call inner coordinates) which are some given functions of time. Thus,

$$L = L(\cdots q^{a[N_a]}; u^\mu) \qquad a = 1, \ldots, n \qquad N_a \geqslant 0. \tag{10}$$

The orders N_a of the inner coordinates q^a in the Lagrange function will be further called the proper orders of the coordinates. We call the coordinates q^a with the proper orders $N_a = 0$, the degenerate coordinates [10].

Equations of motion of a Lagrangian theory (the ELE) follow from the action principle $\delta S = 0$, $S = \int L\,\mathrm{d}t$, and have the form (merely the inner coordinates have to be varied):

$$\frac{\delta S}{\delta q^a} = \sum_{l=0}^{N_a} \left(-\frac{\mathrm{d}}{\mathrm{d}t}\right)^l \frac{\partial L}{\partial q^{a[l]}} = 0 \qquad a = 1, \ldots, n. \tag{11}$$

Following [10], we classify the Lagrangian theories as nonsingular ($M \neq 0$) and singular ($M = 0$) ones with the help of the generalized Hessian $M = \det \|M_{ab}\|$, where

$$M_{ab} = \frac{\partial^2 L}{\partial q^{a[N_a]} \partial q^{b[N_b]}} \tag{12}$$

is the generalized Hessian matrix.

In what follows the ELE of a nonsingular (singular) theory will be called the nonsingular (singular) ELE.

Sometimes, it is convenient to enumerate the inner coordinates and organize them into groups such that $q^a = (q^{a_0}, \ldots, q^{a_I})$, where a_i are groups of indices that enumerate coordinates having the same proper orders, $N_{a_k} = n_k$. Besides, we organize these groups such that $n_I > n_{I-1} > \cdots > n_0 = 0$ (max $N_a = N_{a_I} = n_I$, and q^{a_0} are the degenerate coordinates, $N_{a_0} = n_0 = 0$). Thus,

$$a = (a_k, k = 0, 1, \ldots, I) \qquad [a] = \sum_i [a_i] \qquad [a_i] \geqslant 0 \qquad n_I > n_{I-1} > \cdots > n_0 = 0. \tag{13}$$

Taking into account the notation (13), we may write the Lagrange function and the ELE as

$$L = L\,(\cdots q^{a_k[n_k]}; u^\mu) \qquad k = 0, 1, \ldots, I \tag{14}$$

$$F_{a_k}(\cdots q^{b[N_b+n_k]}; \cdots u^{\mu[n_k]}) = 0 \tag{15}$$

$$F_{a_k} = \begin{cases} M_{a_k b} q^{b[N_b+n_k]} + K_{a_k}(\cdots q^{b[N_b+n_k-1]}; \cdots u^{\mu[n_k]}) & k = 1, \ldots, I \\ M_{a_0}(\cdots q^{b[N_b]}; u^\mu) = \partial L / \partial q^{a_0}. \end{cases} \tag{16}$$

Here $M_{a_k b}$ is the generalized Hessian matrix and K_{a_k} and M_{a_0} are some LF of the indicated arguments.

Consider the orders of the inner coordinates in the complete set of the ELE. These orders are $\mathcal{N}_a = N_a + n_I$. One can see that these orders are, in fact, defined by a subset of (15) with $k = I$. In any subset of equations (15) with $k < I$ the orders of the coordinates are less than those in the complete set. Then according to the above definition, all the ELE with $k < I$ are constraints. The set (15) has the following specific structure: in each equation of the complete set the order of a coordinate q^a is the sum of the proper order N_a and of the order n_k. The latter is the same for all the coordinates and depends only on the number a_k of the equation.

2.3. Canonical form

Let a set of equations

$$F_A(\cdots q^{a[\mathcal{N}_a]}) = 0 \tag{17}$$

be given. Suppose that these equations can be transformed to the following equivalent form:

$$q^{\alpha[l_\alpha]} = \varphi^\alpha(\cdots q^{\alpha[l_\alpha-1]}; \cdots q^{g[l_g]}) \qquad q^a = (q^\alpha, q^g) \qquad a = (\alpha, g) \qquad l_a \leqslant \mathcal{N}_a. \tag{18}$$

Equation (18) presents the canonical form of the initial set (17). In the canonical form the equations are solved with respect to the highest-order time derivatives $q^{\alpha[l_\alpha]}$ of the coordinates q^α. The coordinates q^g (if they exist) and their derivatives $q^{g[l_g]}$ enter the set (18) as arbitrary functions of time. In fact, there are no equations for these coordinates. In what follows we call these coordinates the gauge coordinates whereas we call q^α the nongauge coordinates. The orders of the coordinates in the canonical forms may be less than those in the initial set. In the general case, the same set of equations can have different canonical forms. Generally there are many canonical forms of the same set of equations.

Below, we are going to formulate a general procedure of reducing the ELE to the canonical form (in what follows it is called the reduction procedure). Our consideration is always local in the vicinity of a given consideration point $q_0^{a[l]}$ (in the jet space), which is on shell w.r.t. the ELE. We consider theories and coordinates where the consideration point could be selected as zero point. Thus, we suppose that the zero point is on shell. Further we always suppose that the ranks of the encountered Jacobi matrices[5] are constant in the vicinity of the consideration point. We call such suppositions 'suppositions of the ranks'. Stating that some suppositions hold true in the consideration point, we always suppose that they hold true in the vicinity of the consideration point. In the course of the reduction procedure we perform several typical transformations with LF or with the corresponding equations. Each such transformation leads to equivalent sets of equations or to equivalent sets of LF (definitions of such equivalences are given above). The proof of these equivalences is based on two lemmas which are presented in the appendix. Any statement of the form 'the following equivalence holds true' can be easily justified by these lemmas.

3. Canonical form of nonsingular ELE

3.1. A particular case

Consider theories without external coordinates and with only two different proper orders of the inner coordinates. In such a case all the indices a can be divided into two groups: $a = (a_1, a_2)$, such that $N_{a_2} = n_2 > N_{a_1} = n_1$, $L = L(\cdots q^{a_2[n_2]}, \cdots q^{a_1[n_1]})$. Consider first the case $n_1 > 0$. Then equation (15) can be written as

$$F_{a_2} = M_{a_2 a} q^{a[N_a+n_2]} + K_{a_2}(\cdots q^{b[N_b+n_2-1]}) = 0 \tag{19}$$

$$F_{a_1} = M_{a_1 a} q^{a[N_a+n_1]} + K_{a_1}(\cdots q^{b[N_b+n_1-1]}) = 0. \tag{20}$$

Equations (20) are constraints. Consider the set

$$M_{a_1 a} q^{a[N_a+n_2]} + K_{a_1}^{(1)}(\cdots q^{b[N_b+n_2-1]}) = 0 \tag{21}$$

obtained from the constraints after being differentiated $n_2 - n_1$ times with respect to the time t. Since $M \neq 0$, the rectangular matrix $M_{a_1 a}$ has a maximal rank, therefore there exists another

[5] A rectangular matrix with elements $\partial A_\alpha/\partial x^i$ is often denoted as $\partial A/\partial x$ and called the Jacobi matrix.

division of the indices:

$$a = (a_{|i}) \qquad [a_{|i}] = [a_i] \qquad i = 1, 2 \qquad \det\, M_{a_1 b_{|1}} \neq 0. \tag{22}$$

Note that

$$a_i = (a_{i|1}, a_{i|2}) \qquad a_{|i} = (a_{1|i}, a_{2|i}) \qquad [a_{1|2}] = [a_{2|1}]. \tag{23}$$

The set (21) can be solved with respect to the derivatives $q^{a_{|1}[N_{a_{|1}}+n_2]}$ as follows:

$$q^{a_{|1}[N_{a_{|1}}+n_2]} = -\left(M_1^{-1}\right)^{a_{|1}a_1}\left[(M_3)_{a_1 a_{|2}} q^{a_{|2}[N_{a_{|2}}+n_2]} + K^{(1)}_{a_1}(\cdots q^{b[N_b+n_2-1]})\right]. \tag{24}$$

Here the matrices M_1 and M_3 are defined by the following block representation of the matrix M:

$$M_{ab} = \begin{pmatrix} (M_2)_{a_2 b_{|1}} & (M_4)_{a_2 b_{|2}} \\ (M_1)_{a_1 b_{|1}} & (M_3)_{a_1 b_{|2}} \end{pmatrix} \qquad \det M_{a_1 b_{|1}} \neq 0 \quad \Longrightarrow \quad \det M_1 \neq 0.$$

Excluding the derivatives $q^{a_{|1}[N_{a_{|1}}+n_2]}$ from equation (19) with the help of (24), we get the equations

$$(M_5)_{a_2 b_{|2}} q^{b_{|2}[N_{b_{|2}}+n_2]} + K^{(2)}_{a_2}(\cdots q^{b[N_b+n_2-1]}) = 0. \tag{25}$$

Taking into account a useful relation

$$\begin{aligned} \det M &= \det \begin{pmatrix} M_2 & M_4 \\ M_1 & M_3 \end{pmatrix} = \det \begin{pmatrix} 0 & M_4 - M_2 M_1^{-1} M_3 \\ M_1 & M_3 \end{pmatrix} \\ &= \det M_1 \det\left(M_4 - M_2 M_1^{-1} M_3\right) \end{aligned} \tag{26}$$

which is related to the Gaussian reduction of matrices [11], we get

$$\left.\begin{aligned} \det M \neq 0 \\ \det M_1 \neq 0 \end{aligned}\right\} \quad \Longrightarrow \quad \det M_5 \neq 0 \qquad M_5 = M_4 - M_2 M_1^{-1} M_3. \tag{27}$$

Therefore, (25) can be solved with respect to the highest-order derivatives $q^{a_{|2}[N_{a_{|2}}+n_2]}$ as

$$\begin{aligned} q^{a_{|2}[N_{a_{|2}}+n_2]} &= -\left(M_5^{-1}\right)^{a_{|2}a_2}\left[K_{a_2}(\cdots q^{b[N_b+n_2-1]}) - \left(M_2 M_1^{-1}\right)^{a_1}_{a_2} K^{(1)}_{a_1}(\cdots q^{b[N_b+n_2-1]})\right] \\ &\equiv \varphi^{a_{|2}}(\cdots q^{b_{|2}[N_{b_{|2}}+n_2-1]}, \cdots q^{b_{|1}[N_{b_{|1}}+n_2-1]}). \end{aligned} \tag{28}$$

Thus, we get a set

$$\psi^{a_{|2}} = q^{a_{|2}[N_{a_{|2}}+n_2]} - \varphi^{a_{|2}}(\cdots q^{b_{|2}[N_{b_{|2}}+n_2-1]}, \cdots q^{b_{|1}[N_{b_{|1}}+n_2-1]}) = 0 \tag{29}$$

$$F_{a_1} = (M_1)_{a_1 a_{|1}} q^{a_{|1}[N_{a_{|1}}+n_1]} - K^{(3)}_{a_1}(\cdots q^{b_{|2}[N_{b_{|2}}+n_1]}, \cdots q^{b_{|1}[N_{b_{|1}}+n_1-1]}) = 0 \tag{30}$$

which is strong equivalent to the initial ELE by virtue of lemma 1.

Due to the condition $\det M_1 \neq 0$, equations (30) can be solved with respect to $q^{a_{|1}[N_{a_{|1}}+n_1]}$ and we obtain

$$\begin{aligned} q^{a_{|1}[N_{a_{|1}}+n_1]} &= -\left(M_1^{-1}\right)^{a_{|1}a_1}\left[(M_3)_{a_1 a_{|2}} q^{a_{|2}[N_{a_{|2}}+n_1]} + K_{a_1}(\cdots q^{b[N_b+n_2-1]})\right] \\ &\equiv f^{a_{|1}}(\cdots q^{b_{|2}[N_{b_{|2}}+n_1]}, \cdots q^{b_{|1}[N_{b_{|1}}+n_1-1]}). \end{aligned} \tag{31}$$

Equations (29) and (31) are not of canonical form since the functions $\varphi^{a_{|2}}$ contain derivatives $q^{b_{|1}[N_{b_{|1}}+n_2-1]}$ exceeding the 'allowed' order $[N_{b_{|1}} + n_1 - 1]$. Now we exclude all the surplus derivatives $q^{a_{|1}[N_{a_{|1}}+n_1]}, \ldots, q^{a_{|1}[N_{a_{|1}}+n_2-1]}$ from the right-hand side of (29) with the help of (31) and its corresponding derivatives. To this end we need to differentiate (31) not more than $n_2 - n_1 - 1$ times. Finally, we obtain the following strong equivalent form (the equivalence is justified by lemma 1) of the ELE:

$$\begin{aligned} q^{a_{|2}[N_{a_{|2}}+n_2]} &= f^{a_{|2}}(\cdots q^{b_{|2}[N_{b_{|2}}+n_2-1]}, \cdots q^{b_{|1}[N_{b_{|1}}+n_1-1]}) \\ q^{a_{|1}[N_{a_{|1}}+n_1]} &= f^{a_{|1}}(\cdots q^{b_{|2}[N_{b_{|2}}+n_1]}, \cdots q^{b_{|1}[N_{b_{|1}}+n_1-1]}). \end{aligned} \tag{32}$$

It is just the canonical form. Taking into account the division of the indices w.r.t. proper orders of the coordinates, one gets

$$\begin{aligned}
q^{a_{2|2}[2n_2]} &= f^{a_{2|2}}(\cdots q^{b_{2|2}[2n_2-1]}, \cdots q^{b_{1|2}[n_1+n_2-1]}, \cdots q^{b_{2|2}[n_2+n_1-1]}, \cdots q^{b_{1|1}[2n_1-1]})\\
q^{a_{1|2}[n_1+n_2]} &= f^{a_{1|2}}(\cdots q^{b_{2|2}[2n_2-1]}, \cdots q^{b_{1|2}[n_1+n_2-1]}, \cdots q^{b_{2|1}[n_2+n_1-1]}, \cdots q^{b_{1|1}[2n_1-1]})\\
q^{a_{2|1}[n_2+n_1]} &= f^{a_{2|1}}(\cdots q^{b_{2|2}[n_1+n_2]}, \cdots q^{b_{1|2}[2n_1]}, \cdots q^{b_{2|1}[n_2+n_1-1]}, \cdots q^{b_{1|1}[2n_1-1]})\\
q^{a_{1|1}[2n_1]} &= f^{a_{1|1}}(\cdots q^{b_{2|2}[n_1+n_2]}, \cdots q^{b_{1|2}[2n_1]}, \cdots q^{b_{2|1}[n_2+n_1-1]}, \cdots q^{b_{1|1}[2n_1-1]}).
\end{aligned} \tag{33}$$

Note that the number of initial data is equal to $2\sum_a N_a$. Indeed,

$$\begin{aligned}
&[a_{2|2}](n_2+n_2) + [a_{1|2}](n_1+n_2) + [a_{2|1}](n_2+n_1) + [a_{1|1}](n_1+n_1)\\
&\qquad = 2[a_2]n_2 + 2[a_1]n_1 = 2\sum_a N_a.
\end{aligned}$$

One ought to mention that the canonical form (33) was obtained in [12]. However, the procedure that was used for that purpose did not provide the proof of the equivalence between the initial ELE and the form (33).

Suppose now that the Lagrange function contains degenerate coordinates q^{a_0}, $a = (a_0, a_1)$. Thus, $L = L(q^{a_0}, \cdots q^{a_1[n_1]})$ and the ELE read

$$F_{a_1} \equiv M_{a_1 a} q^{a[N_a+n_1]} + K_{a_1}(\cdots q^{b[N_b+n_1-1]}) = 0 \tag{34}$$

$$F_{a_0} \equiv \frac{\partial L}{\partial q^{a_0}} = M_{a_0}(\cdots q^{b[N_b]}) = 0. \tag{35}$$

Despite these equations being formally different from the above case, the whole procedure of reductions goes through without any essential change. In fact, differentiating equations (35) n_1 times, one obtains

$$M_{a_0 a} q^{a[N_a+n_1]} + K^{(1)}_{a_0}(\cdots q^{b[N_b+n_1-1]}) = 0 \tag{36}$$

and all the previous steps may be done as before. Namely, one obtains

$$q^{a_{|1}[N_{a_{|1}}+n_1]} = \varphi^{a_{|1}}(\cdots q^{b_{|1}[N_{b_{|1}}+n_1-1]}, \cdots q^{b_{|0}[N_{b_{|0}}+n_0-1]}) \tag{37}$$

and, since $\det \|(M_1)_{a_0 a_{|0}}\| \neq 0$, equations (35) can be solved with respect to the variable $q^{a_{|0}[N_{a_{|0}}]}$ as follows:

$$q^{a_{|0}[N_{a_{|0}}]} = f^{a_{|0}}(\cdots q^{b_{|1}[N_{b_{|1}}]}, \cdots q^{b_{|0}[N_{b_{|0}}-1]}).$$

Finally, after eliminating the 'bad' derivatives on the right-hand side of (37) for $q^{a_{|1}[N_{a_{|1}}+n_1]}$ one ends up again with equations (33) but now with $n_2 \to n_1, n_1 \to 0$ (by convention $q^{b_{1|1}[-1]} \equiv 0$).

3.2. General nonsingular ELE

Consider the general nonsingular ELE. Here the Lagrange function may contain some degenerate inner coordinates, higher derivatives of some inner coordinates, and, moreover, may depend on some external coordinates, $L = L(\cdots q^{a[N_a]}; u^\mu)$, $N_a \geqslant 0$. Thus, we are going to deal with the nonsingular ELE of the form (15). Our aim is to present these equations in an equivalent canonical form.

Theorem 1. *The nonsingular ELE (15) can be transformed to the following equivalent canonical form:*

$$f^{a_{i|k}} = q^{a_{i|k}[n_i+n_k]} - \varphi^{a_{i|k}}(\cdots q^{b_{j|k_-}[n_j+n_{k_-}-1]}, \cdots q^{b_{j|k_+}[n_j+n_k]}; \cdots u^{\mu[n_k]}) = 0$$
$$I \geqslant k_+ \geqslant k+1 \qquad k \geqslant k_- \geqslant 0 \qquad i, j, k = 0, 1, \ldots, I \tag{38}$$

where the indices of the coordinates are divided into groups as follows: $a = (a_i)$ is the division of the indices w.r.t. the proper orders of the coordinates, and besides

$$a_i = (a_{i|k}, i, k = 0, 1, \ldots, I) \qquad [a_{i|k}] \geqslant 0 \qquad \sum_k [a_{i|k}] = \sum_k [a_{k|i}] = [a_i] = [a_{|i}].$$

Moreover, the equivalence $F \sim f$ between the corresponding LF holds true. That implies

$$F_a = \hat{U}_{ab} f^b \qquad f^b = \hat{V}^{ba} F_a \qquad \hat{U}_{ab} \hat{V}^{bc} = \delta_a^c$$

where $\hat{U}$ and $\hat{V}$ are LO. Besides, that implies the strong equivalence between the ELE and their canonical form (38).

The proof of theorem 1 may be considered, in fact, as the general reduction procedure to the canonical form for the nonsingular ELE.

It is reasonable to divide the reduction procedure into two parts. These parts may be called conditionally 'the preliminary resolution' and 'the subordination procedure'.

3.2.1. Preliminary resolution. Let us introduce the notation $a = (\underline{a}, a_I), \underline{a} = (a_k, k = 0, 1, \ldots, I-1)$, $N_{\underline{a}} < n_I$, such that the ELE read

$$F_{a_I}(\cdots q^{b[N_b+n_I]}; \cdots u^{\mu[n_I]}) = M_{a_I b} q^{b[N_b+n_I]} + K_{a_I}(\cdots q^{b[N_b+n_I-1]}; \cdots u^{\mu[n_I]}) = 0 \tag{39}$$

$$F_{\underline{a}}(\cdots q^{b[N_b+N_{\underline{a}}]}; \cdots u^{\mu[N_{\underline{a}}]}) = 0. \tag{40}$$

Recall that equations (40) can be considered as constraints.

The first step of the procedure is the following: we consider the consistency conditions of the constraints. Namely, we consider the equations that are obtained from the constraints by differentiating them $n_I - n_{\underline{a}}$ times,

$$F_{\underline{a}}^{[n_I - N_{\underline{a}}]} = M_{\underline{a}b} q^{b[N_b+n_I]} + K_{\underline{a}}^{(1)}(\cdots q^{b[N_b+n_I-1]}; \cdots u^{\mu[n_I]}) = 0. \tag{41}$$

Here $K_{\underline{a}}^{(1)}$ are some LF of the indicated arguments. Note that the orders of all the coordinates in the set (41) coincide with those in the complete set. For $M \neq 0$, the matrix

$$\frac{\partial F_a^{[n_I - N_a]}}{\partial q^{b[N_b+n_I]}} = \frac{\partial^2 L}{\partial q^{a[N_a]} \partial q^{b[N_b]}} = M_{ab} \tag{42}$$

is invertible. At the same time, the rectangular matrix $M_{\underline{a}a}$ has the maximal rank $[\,\underline{a}\,]$. Therefore, there exists a division of the indices a such that

$$a = (\bar{a}, a_{|I}) \qquad [\bar{a}] = [\underline{a}] \qquad [a_{|I}] = [a_I] \qquad \det M_{\underline{a}\bar{a}} \neq 0. \tag{43}$$

Thus, the division (13) of the indices a w.r.t. the proper orders of the coordinates becomes more detailed,

$$a_i = (\bar{a}_i, a_{i|I}) \qquad \bar{a} = (\bar{a}_i) \qquad a_{|I} = (a_{i|I}) \qquad [a_{i|I}] \geqslant 0$$
$$\sum_i [a_{i|I}] = [a_{|I}] = [a_I] \qquad \sum_i [\bar{a}_i] = [\bar{a}] = [\underline{a}].$$

Due to (43), the set (41) can be solved with respect to the derivatives $q^{\bar{a}[N_{\bar{a}}+n_I]}$ as

$$q^{\bar{a}[N_{\bar{a}}+n_I]} = -\left(M_1^{-1}\right)^{\bar{a}\underline{a}}\left[(M_3)_{\underline{a}b_{|I}} q^{b_{|I}[N_{b_{|I}}+n_I]} + K^{(1)}_{\underline{a}}(\cdots q^{b[N_b+n_I-1]}; \cdots u^{\mu[n_I]})\right] \tag{44}$$

where

$$M_{ab} = \begin{pmatrix} (M_2)_{a_I\bar{b}} & (M_4)_{a_I b_{|I}} \\ (M_1)_{\underline{a}\bar{b}} & (M_3)_{\underline{a}b_{|I}} \end{pmatrix}.$$

Excluding the derivatives $q^{\bar{a}[N_{\bar{a}}+n_I]}$ from equations (39) with the help of (44), we get the set

$$\begin{aligned} &(M_5)_{a_I b_{|I}} q^{b_{|I}[N_{b_{|I}}+n_I]} + K^{(2)}_{a_I}(\cdots q^{b[N_b+n_I-1]}; \cdots u^{\mu[n_I]}) = 0 \\ &M_5 = M_4 - M_2 M_1^{-1} M_3 \qquad \det M_5 \neq 0 \end{aligned} \tag{45}$$

where $K^{(2)}_{a_I}$ are some LF of the indicated arguments. The set (45) can be solved with respect to its highest-order derivatives $q^{a_{|I}[N_{a_{|I}}+n_I]}$ as

$$q^{a_{|I}[N_{a_{|I}}+n_I]} = \phi^{a_{|I}}(\cdots q^{b_{|I}[N_{b_{|I}}+n_I-1]}, \cdots q^{\bar{b}[N_{\bar{b}}+n_I-1]}; \cdots u^{\mu[n_I]}) \tag{46}$$

where $\varphi^{a_{|I}}$ are some LF. Thus, after the first step we get a set of equations

$$\psi^{a_{|I}} = q^{a_{|I}[N_{a_{|I}}+n_I]} - \phi^{a_{|I}}(\cdots q^{b_{|I}[N_{b_{|I}}+n_I-1]}, \cdots q^{\bar{b}[N_{\bar{b}}+n_I-1]}; \cdots u^{\mu[n_I]}) = 0 \tag{47}$$

$$F_{\underline{a}}(\cdots q^{b[N_b+N_{\underline{a}}]}; \cdots u^{\mu[N_{\underline{a}}]}) = 0 \qquad \underline{a} = (a_k, k = 0, 1, \ldots, I-1) \qquad N_{\underline{a}} < N_I \tag{48}$$

which are strong equivalent to the initial ELE by virtue of lemma 1 in the appendix.

In the second step we turn to the subset (48). We remark that this subset has the same structure as the complete initial set of the ELE if one considers the coordinates $q^{\bar{a}}$ as inner ones and the variables $q^{a_{|I}}$ as external ones. Namely, let us denote

$$\begin{aligned} &\overset{1}{F}_{\underline{a}}(\cdots q^{\bar{b}[N_{\bar{b}}+N_{\underline{a}}]}; \cdots u^{\mu_1[N_{\underline{a}}]}) = F_{\underline{a}}(\cdots q^{b[N_b+N_{\underline{a}}]}; \cdots u^{\mu[N_{\underline{a}}]}) \\ &u^{\mu_1} = (u^{\mu}, \cdots q^{a_{|I}[N_{a_{|I}}]}) \qquad \mu_1 = (\mu, a_{|I}). \end{aligned}$$

Then the set (48) can be written as

$$\overset{1}{F}_{\underline{a}}(\cdots q^{\bar{b}[N_{\bar{b}}+N_{\underline{a}}]}; \cdots u^{\mu_1[N_{\underline{a}}]}) = 0 \qquad \underline{a} = (a_k, k = 0, 1, \ldots, I-1) \qquad N_{\underline{a}} < N_I \tag{49}$$

where

$$\overset{1}{F}_{a_k} = \begin{cases} M_{a_k\bar{b}} q^{\bar{b}[N_{\bar{b}}+n_k]} + \underline{K}_{a_k}(\cdots q^{\bar{b}[N_{\bar{b}}+n_k-1]}; \cdots u^{\mu_1[n_k]}) & k = 1, \ldots, I-1 \\ \underline{M}_{a_0}(\cdots q^{\bar{b}[N_{\bar{b}}]}; u^{\mu_1}) = \partial L/\partial q^{a_0}. \end{cases}$$

Here $q^{\bar{a}}$ are the inner coordinates, and u^{μ_1} are the external coordinates. The order of the set (49) is $2n_{I-1}$. Furthermore, by virtue of (43), the matrix

$$\frac{\partial \overset{1}{F}{}^{[n_{I-1}-N_{\underline{a}}]}_{\underline{a}}}{\partial q^{\bar{b}[N_{\bar{b}}+n_{I-1}]}} = M_{\underline{a}\bar{b}} \tag{50}$$

is invertible. Thus, the structure (15), (16) is repeated completely.

At the same time, the number of inner variables, the number of equations and the order of the set (49) are less than those of the initial set of the ELE (15), (16).

Now, we apply the same procedure as in the first step to the reduced set (49). That will be the second step of the reduction procedure. It will produce equations of similar structure with lesser inner variables and of lower order. After the last $(I+1)$th step the ELE (15) may be written in the following strong equivalent form:

$$\begin{aligned} &q^{a_{i|k}[n_i+n_k]} = \phi^{a_{i|k}}(\cdots q^{b_{j|k+}[n_j+n_k]}, \cdots q^{b_{j|k-}[n_j+n_k-1]}; \cdots u^{\mu[n_k]}) \\ &I \geqslant k_+ \geqslant k+1 \qquad k \geqslant k_- \geqslant 0 \qquad I \geqslant i, j \geqslant 0 \end{aligned} \tag{51}$$

where $\phi^{a_{i|k}}$ are some LF of the indicated arguments (the arguments $\cdots q^{b_{j|k_+}[n_j+n_k]}$ result from those coordinates that intermediately have been considered as external ones), and the indices a_i are divided into the following groups:

$$a_i = (a_{i|k}) \qquad [a_{i|k}] \geqslant 0 \qquad \sum_k [a_{i|k}] = \sum_k [a_{k|i}] = [a_i] \qquad i, k = 0, 1, \ldots, I.$$

The set (51) is still not the canonical form of the ELE. The reason is that the right-hand sides of the set contain derivatives of orders that may exceed the orders $n_i + n_k$ of the (highest) derivatives $q^{a_{i|k}[n_i+n_k]}$ appearing on the left-hand side of the set. We recall that by the definition in the canonical form there is a subordination of derivative orders, namely, the orders of all the derivatives on the right-hand sides have to be less than those on the left-hand side. Explicitly, this subordination would require that the following inequalities should hold:

$$n_j + n_{k_+} > n_j + n_k \qquad n_j + n_{k_-} > n_j + n_k - 1$$

which, because of the inequalities $n_I > n_{I-1} > \cdots > n_1 > n_0$, is true for the first line and the case $k_- = k$ of the second line, and it is definitely not true for the cases $k_- < k$. Arranging equations (51) (for fixed value of i) in descending order w.r.t. k, and the arguments in the functions φ (for fixed value of j) also in descending order w.r.t. the values of k_+ and k_-, we get, when disregarding the common value n_j, a quadratic matrix whose main diagonal (i.e. elements with $k = k_-$) contains the entries $n_k - 1$, whereas the entries to the left of that diagonal are equal to n_k, and to the right of that diagonal are equal to $n_k - 1$. Therefore, below the main diagonal 'good' derivatives occur, and above it occur 'bad' derivatives not obeying the subordination requirement.

3.2.2. Subordination procedure. One can see that these 'bad' derivatives can be excluded from the right-hand sides with the help of the corresponding 'lower' equations of the set and their differential consequences (compare equations (29) and (31) for the simple case $I = 2$). In what follows we call such an exclusion the subordination procedure.

In order to be more definite let us write down two arbitrary lines, $\ell > k$, of the right-hand sides of the set of equations (51) (for the highest derivatives only):

$$\begin{aligned}
&\phi^{a_{i|\ell}}\big(q^{b_{j|I}[n_j+n_\ell]}, \ldots, q^{b_{j|\ell+1}[n_j+n_\ell]}, q^{b_{j|\ell}[n_j+n_\ell-1]}, \ldots, q^{b_{j|k+1}[n_j+n_\ell-1]},\\
&\qquad\qquad q^{b_{j|k}[n_j+n_\ell-1]}, \ldots, q^{b_{j|0}[n_j+n_\ell-1]}\big)\\
&\qquad\qquad\qquad\vdots\\
&\phi^{a_{i|k}}\big(q^{b_{j|I}[n_j+n_k]}, \ldots, q^{b_{j|\ell+1}[n_j+n_k]}, q^{b_{j|\ell}[n_j+n_k]}, \ldots, q^{b_{j|k+1}[n_j+n_k]}, q^{b_{j|k}[n_j+n_k-1]}, \ldots, q^{b_{j|0}[n_j+n_k-1]}\big).
\end{aligned}$$

Obviously, because $n_\ell > n_k$ all the derivatives of the equation for $q^{a_{i|\ell}[n_i+n_\ell]}$ with $k \geqslant \ell_- \geqslant 0$ are 'bad' with respect to the derivatives $q^{a_{i|k}[n_i+n_k]}$ (recall that $\ell \geqslant \ell_- \geqslant 0$). However, these 'bad' derivatives can be eliminated by the equations for the latter ones, $q^{a_{i|k}[n_i+n_k]}$, and their differential consequences up to the order $n_\ell - n_k - 1$. Thereby, the function $\phi^{a_{i|\ell}}$ changes into some new function $\tilde{\phi}^{a_{i|\ell}}$. One can see that by doing this we do not change the highest orders of derivatives of the other coordinates, both proper and external ones, on the right-hand side of the equation for $q^{a_{i|\ell}[n_i+n_\ell]}$. (Recall that the derivatives of the external coordinates are $u^{\mu[n_\ell]}$ and $u^{\mu[n_k]}$, respectively.)

This subordination procedure, starting with $\ell = I$, may be done for any $k < I$, thereby 'cleaning' every entry on the right-hand side of equations for $q^{a_{i|I}[n_i+n_I]}$. Namely, the highest orders of derivatives on the rhs become $q^{b_{j|k_-}[n_i+n_{k_-}-1]}$ with $I \geqslant k_- \geqslant 0$ (for the case $\ell = I$ no k_+ appears). Then the procedure will be applied to the equations for $q^{a_{i|I-1}[n_i+n_{I-1}]}$, and so forth, up to $q^{a_{i|0}[n_i+n_0]}$, where nothing is to be changed.

After having eliminated all the 'bad' derivatives, we transformed the set (51), and therefore the initial ELE, to the following strong equivalent (the equivalence is justified by lemma 1) canonical form:

$$\begin{aligned} &q^{a_{i|k}[n_i+n_k]} = \varphi^{a_{i|k}}(\cdots q^{b_{j|k_+}[n_j+n_k]}, \cdots q^{b_{j|k_-}[n_j+n_{k_-}-1]}; \cdots u^{\mu[n_k]}) \\ &I \geqslant k_+ \geqslant k+1 \qquad k \geqslant k_- \geqslant 0 \qquad i, j, k = 0, 1, \ldots, I \end{aligned}$$

where $\varphi^{a_{i|k}}$ are some LF of the indicated arguments. This proves theorem 1.

We see that there are no gauge coordinates in the nonsingular ELE.

The number of initial data is equal to $2\sum_a N_a$. Indeed,

$$\sum_{i,k}[a_{i|k}](N_i + N_k) = \sum_i \left(N_i \sum_k [a_{i|k}]\right) + \sum_k \left(N_k \sum_i [a_{i|k}]\right) = 2\sum_a N_a.$$

One ought to remark that in the general case there exist many different canonical forms of the nonsingular ELE. This uncertainty is related to the possibility of different choices of nonzero minors of a matrix with a given rank (different divisions of the indices a_i in course of the reduction procedure). However, as was demonstrated above, the number of equations in the canonical form (which is equal to the number of ELE in the nonsingular case) and the number of initial data are the same for all possible canonical forms.

4. Canonical form of singular ELE

Studying the canonical form of nonsingular ELE, we have demonstrated that the equations in the canonical form are solved with respect to the highest-order derivatives $q^{a_{i|k}[n_i+n_k]}$, where n_i are the proper orders of the coordinates q^{a_i}. However, considering specific examples, one can see that this is not always true for singular ELE. Namely, in the canonical form of the latter case, the highest orders of the derivatives $q^{a_i[l]}$ may take on all the values from zero to $n_i + I$. The reduction procedure to the canonical form for the general singular ELE is considered below. In the singular case, already after the first step of the reduction procedure, the ELE cease to have their initial specific structure (15), (16). Namely, the simple structure of terms with highest-order derivatives in the equations may be lost. That is why in the singular case it is more convenient to formulate the reduction procedure for a more general set of ordinary differential equations, which contains the ELE as a particular case. Namely, further we are going to consider a set of the form[6]

$$F_{A_\mu}(\cdots q^{a_i[i+\mu]}) = 0 \qquad i = 0, 1, \ldots, I \qquad \mu = 0, \ldots, J. \tag{52}$$

Here $F_{A_\mu}(\cdots q^{a_i[i+\mu]})$ are some LF. a_i and A_μ are used to denoted sets of indices, $[a_i] \geqslant 0$, $[A_\mu] \geqslant 0$, and the complete set of the inner coordinates in equations (52) is $q^a = (q^{a_0}, \ldots, q^{a_I})$, $a = (a_i, i = 0, 1, \ldots, I)$. The indices $A = (A_\mu)$ enumerate the equations. In the general case the number of indices A (the number of all the equations) is not equal to the number of indices a (the number of coordinates). The division of the indices A into groups is not related to the division of the indices a into groups. The orders of the coordinates q^{a_i} in the complete set (52) are $\mathcal{N}_{a_i} = i + J$. In fact, these orders are defined by a subset of (52) with $\mu = J$. In all the other equations with $\mu < J$ the coordinates q^{a_i} have orders less than $i + J$. Thus, the latter equations are constraints.

Similar to the ELE (15), the set (52) has the following specific structure: in each equation of the set the order of a coordinate q^{a_i} is the sum of the proper order i and of the order μ. The latter is the same for all the coordinates and is related to the number of the equation in the set.

[6] We do not indicate here the possible external coordinates.

Below we consider the reduction procedure to the canonical form for equations (52). In fact, this reduction procedure is formulated in the proof of theorem 2. Theorem 2 holds true under certain suppositions of the structure of the functions F_{A_μ}. These suppositions are formulated as suppositions of the ranks of some Jacobi matrices involving the functions F_{A_μ}. First of all, the complete matrix

$$M_{A_\mu a_i} = \frac{\partial F_{A_\mu}}{\partial q^{a_i[i+\mu]}} = \frac{\partial F_{A_\mu}^{[J-\mu]}}{\partial q^{a_i[i+J]}} \tag{53}$$

has to have a constant rank in the vicinity of the consideration point (one can see that the matrix $M_{A_\mu a_i}$ coincides with the generalized Hessian matrix if the set (52) is the Lagrangian one).

Theorem 2. *Under certain suppositions of the ranks, equations (52) can be transformed to the following equivalent canonical form:*

$$\begin{gathered} f^{a_{i|\sigma}} = q^{a_{i|\sigma}[i+\sigma]} - \varphi^{a_{i|\sigma}}(\cdots q^{a_{j|\sigma_-}[j+\sigma_- -1]}, \cdots q^{a_{j|\sigma_+}[j+\sigma]}) = 0 \\ i, j = 0, 1, \ldots, I \qquad \sigma = -I, \ldots, J \qquad -I \leqslant \sigma_- \leqslant \sigma \qquad \sigma + 1 \leqslant \sigma_+ \leqslant J + 1 \end{gathered} \tag{54}$$

where all the indices a are divided into groups as follows:

$$a_i = (a_{i|\sigma}) \qquad [a_{i|\sigma}] \geqslant 0 \qquad \sigma = -I, \ldots, J+1 \qquad ([a_{i|\sigma}] = 0 \text{ if } i + \sigma < 0) \tag{55}$$

and it is assumed that negative powers of the time derivatives do not exist, that is $[q^{a[p]}] = 0$ *for* $p < 0$.

Moreover, the following equivalence between the corresponding LF holds true:

$$\begin{gathered} F_A \sim \bar{F}_A = \begin{pmatrix} f^{a_{i|\sigma}} \\ 0_G \end{pmatrix} \qquad A = (a_{i|\sigma}, G) \qquad i = 0, 1, \ldots, I \qquad \sigma = -I, \ldots, J \\ 0_G \equiv 0 \quad \forall G \qquad [G] = [A] - [a] + \sum_i [a_{i|J+1}]. \end{gathered} \tag{56}$$

That implies

$$F_A = \hat{U}_A^B \bar{F}_B \qquad \bar{F}_B = \hat{V}_B^A F_A \qquad \hat{U}_A^B \hat{V}_B^C = \delta_A^C \tag{57}$$

where $\hat{U}$ *and* $\hat{V}$ *are LO.*

Let us make some comments on theorem 2. The canonical form (54) of the singular ELE differs from that (38) of the nonsingular ELE. As was demonstrated in the previous section, in the latter case the spectrum of the orders of the variables q^{a_i} in the canonical form extends from $i + \mu_{\min}$ to $i + J$. In the singular case, we have to admit (and one can see this from specific examples) the spectrum extends from 0 to $i + J$. Under such a supposition we can justify by induction the structure (54) of the canonical form. One can see from (55) that each group of the indices a_i is divided into subgroups $a_i \to a_{i|\sigma}, \sigma = -I, \ldots, J+1$. In the canonical form the singular ELE are solved with respect to the highest-order derivatives $q^{a_{i|\sigma}[i+\sigma]}, \sigma = -I, \ldots, J$ ($[a_{i|\sigma}] = 0$ for $i + \sigma < 0$). There are no equations for the coordinates $q^{a_{i|J+1}}$. These coordinates enter the set (54) as arbitrary functions of time. They are gauge coordinates according to the general definition. As in the nonsingular case, it is supposed that no coordinate $q^{a_{k|\sigma}}$ in the function $\varphi^{a_{i|\sigma}}$ has an order greater than $k + \sigma$ (the proper order plus σ). Besides, the order of the coordinates $q^{b_{k|\sigma_-}}$ in the function $\varphi^{a_{i|\sigma}}$ has to be less than $k + \sigma_-$.

We are going to prove theorem 2 by induction w.r.t. $\mathcal{N} = I + J$. To this end, we consider first equations of lower orders, then we use induction to prove the general case.

4.1. Equations of lower orders

Note that the case $\mathcal{N} = 0$ implies $I = J = 0$ and the set (52) is reduced to the form

$$F_A(q) = 0 \qquad q = (q^a). \tag{58}$$

Here theorem 2 holds true by virtue of lemma 3 in the appendix.

Let $\mathcal{N} = 1$. That implies either $I = 1,\ J = 0$ or $I = 0,\ J = 1$. Consider, for example, the first case. Here $(i = 0, 1, \mu = 0)$ and the set (52) reads

$$F_A(q^{a_0}, q^{a_1}, \dot{q}^{a_1}) = 0 \qquad [a_1] > 0 \qquad [a_0] \geqslant 0. \tag{59}$$

In the case under consideration the supposition (53) reads

$$\operatorname{rank}\frac{\partial F_A}{\partial q^{a_i[i]}} = r. \tag{60}$$

Then there exists a division of the indices $A = (A_{/1}, A_{/2}),\ a_i = (a_{i/1}, a_{i/2}),\ [A_{/1}] = [a_{0/1}] + [a_{0/1}] = r$, such that

$$\det\left|\frac{\partial F_{A_{/1}}}{\partial q^{a_{i/1}[i]}}\right| \neq 0.$$

Thus, we may solve the equations $F_{A_{/1}} = 0$ with respect to $q^{a_{i/1}[i]}$,

$$F_{A_{/1}} = 0 \quad\Longleftrightarrow\quad q^{a_{i/1}[i]} = \phi^{a_{i/1}}(q^{b_{i/1}[i-1]}, q^{b_{i/2}[i-1]}, q^{b_{i/2}[i]}). \tag{61}$$

Then we exclude the arguments $q^{a_{i/1}[i]}$ from the functions $F_{A_{/2}}$ with the help of (61),

$$\bar{F}_{A_{/2}} = F_{A_{/2}}|_{q^{a_{i/1}[i]}=\phi^{a_{i/1}}} = \bar{F}_{A_{/2}}(q^{a_1}).$$

By virtue of lemma 2 in the appendix, the functions $\bar{F}_{A_{/2}}$ depend on q^{a_1} only. Thus, we have the equivalence[7]

$$F_A \sim \bar{F}_A = \begin{pmatrix} F_{A_{/1}}(q^{a_0}, q^{a_1}, \dot{q}^{a_1}) \\ \bar{F}_{A_{/2}}(q^{a_1}) \end{pmatrix}. \tag{62}$$

Now we suppose that the matrix $\partial\bar{F}_{A_{/2}}/\partial q^{a_1}$ has a constant rank. Therefore (see lemma 3)

$$\bar{F}_{A_{/2}} \sim \begin{pmatrix} q^{\underline{a}_1} - \varphi^{\underline{a}_1}(q^{\bar{a}_1}) \\ 0_{G_1} \end{pmatrix} \qquad a_1 = (\underline{a}_1, \bar{a}_1).$$

Let us exclude the arguments $q^{\underline{a}_1}, \dot{q}^{\underline{a}_1}$ from the functions $F_{A_{/1}}$, with the help of the equations $q^{\underline{a}_1} = \varphi^{\underline{a}_1}(q^{\bar{a}_1})$,

$$\overset{1}{F}_{A_{/1}}(q^{a_0}, q^{\bar{a}_1}, \dot{q}^{\bar{a}_1}) = F_{A_{/1}}|_{q^{\underline{a}_1}=\varphi^{\underline{a}_1}}.$$

Then the following equivalence holds true:

$$F \sim \overset{1}{F} = \begin{pmatrix} \overset{1}{F}_{A_{/1}}(q^{a_0}, q^{\bar{a}_1}, \dot{q}^{\bar{a}_1}) \\ q^{\underline{a}_1} - \varphi^{\underline{a}_1}(q^{\bar{a}_1}) \\ 0_{G_1} \end{pmatrix} \qquad a = (a_0, a_1) \qquad a_1 = (\underline{a}_1, \bar{a}_1). \tag{63}$$

The set of functions $\overset{1}{F}$ has the same structure as the initial set F. However, the number of the nonzero functions $\overset{1}{F}$ is less than the number of the functions F. Moreover, some of the functions $\overset{1}{F}$ depend linearly on a part of the variables. That is why the supposition of type (60) for the functions $F^{(1)}$ is reduced to the supposition about the rank of the matrix

[7] Here, and in what follows, we use lemma 1 to justify the equivalence.

$\partial \overset{1}{F}_{A_{/1}} \big/ \partial(q^{a_0}, \dot{q}^{\bar{a}_1})$. Accepting the latter supposition we apply the above reduction procedure to the functions $\overset{1}{F}$ and so on. After the ith stage we have the following equivalence:

$$F \sim \overset{i}{F} = \begin{cases} \overset{i}{F}_{A_{/i}}(q^{a_0}, q^{\bar{a}_i}, \dot{q}^{\bar{a}_i}) \\ q^{\underline{a}_i} - \varphi^{\underline{a}_i}(q^{\bar{a}_i}) \\ 0_{G_i} \end{cases} \qquad a = (a_0, a_1) \qquad a_1 = (\underline{a}_i, \bar{a}_i).$$

The procedure ends at the kth stage when

$$\operatorname{rank} \frac{\partial \overset{k}{F}_{A_{/k}}}{\partial(\dot{q}^{\bar{a}_k}, q^{a_0})} = [A_{/k}].$$

Then there exists a division of the indices $\bar{a}_k = (a_{1|0}, a_{g_1})$, $a_0 = (a_{0|0}, a_{g_0})$, $[a_{1|0}] + [a_{0|0}] = [A_{/k}]$, such that

$$\det \frac{\partial \overset{k}{F}_{A_{/k}}}{\partial(\dot{q}^{a_{1|0}}, q^{a_{0|0}})} \neq 0 \quad \Longrightarrow \quad \overset{k}{F}_{A_{/k}} \sim \begin{pmatrix} \dot{q}^{a_{1|0}} - \varphi^{a_{1|0}}(q^{a_{1|0}}, q^{a_{g_0}}, q^{a_{g_1}}, \dot{q}^{a_{g_1}}) \\ q^{a_{0|0}} - \varphi^{a_{0|0}}(q^{a_{1|0}}, q^{a_{g_0}}, q^{a_{g_1}}, \dot{q}^{a_{g_1}}) \end{pmatrix}.$$

Denoting $\underline{a}_k \equiv a_{1|-1}$, $G = G_k$, such that $a = (a_{1|-1}, a_{0|0}, a_{1|0}, a_g)$, and $a_g = (a_{g_0}, a_{g_1})$, $[G] = [A] - [a] + [a_g]$, we get finally the equivalence

$$F \sim \begin{pmatrix} \dot{q}^{a_{1|0}} - \varphi^{a_{1|0}}(q^{a_{1|0}}, q^{a_{g_0}}, q^{a_{g_1}}, \dot{q}^{a_{g_1}}) \\ q^{a_{0|0}} - \varphi^{a_{0|0}}(q^{a_{1|0}}, q^{a_{g_0}}, q^{a_{g_1}}, \dot{q}^{a_{g_1}}) \\ q^{a_{1|-1}} - \varphi^{a_{1|-1}}(q^{a_{1|0}}, q^{a_{g_1}}) \\ 0_G \end{pmatrix}. \tag{64}$$

Here $q^{a_g} = (q^{a_{g_0}}, q^{a_{g_1}})$ are gauge coordinates. Thus, theorem 2 holds true in this case.

The case $I = 0$, $J = 1$ ($i = 0$, $\mu = 0, 1$) corresponds to the equations of the form

$$F_{A_1}(q^{a_1}, \dot{q}^{a_1}) = 0 \qquad F_{A_0}(q^{a_1}) = 0. \tag{65}$$

Such equations present a particular case ($[a_0] = 0$) of the equations $\bar{F}_A = 0$ with the LF $\bar{F}_A$ defined in (62). The reduction procedure for the latter case was considered above. It leads to the following equivalence:

$$F \sim \begin{pmatrix} \dot{q}^{a_{|1}} - \varphi^{a_{|1}}(q^{a_{|1}}, q^{a_g}, \dot{q}^{a_g}) \\ q^{a_{|0}} - \varphi^{a_{|0}}(q^{a_{|1}}, q^{a_g}) \\ 0_G \end{pmatrix} \qquad a = (a_{|0}, a_{|1}, a_g) \qquad [G] = [A] - [a] + [a_g].$$

Here q^{a_g} are the gauge coordinates. Thus, the theorem holds true in this case as well.

4.2. Equations of arbitrary orders

We have verified that theorem 2 holds true for $\mathcal{N} = 0, 1$. Now we are going to prove the theorem for $\mathcal{N} = I + J = K$ (where K is some fixed number) supposing that the theorem holds true for any $\mathcal{N} < K$.

In the first step we consider the set

$$F_{A_\mu}^{[J-\mu]}(\cdots q^{a_i[i+J]}) = 0 \qquad i = 0, 1, \ldots, I \qquad \mu = 0, \ldots, J \tag{66}$$

which is obtained from the initial set (52) by substituting the constraints by the corresponding consistency conditions (conditions obtained from the constraints F_{A_μ} by $J - \mu$ time differentiations). According to the supposition (53), there exists a division of the indices A_μ and a_i as $A_\mu = (A_{\mu/1}, A_{\mu/2})$, $a_i = (a_{i/1}, a_{i/2})$, $\sum_\mu [A_{\mu/1}] = \sum_i [a_{i/1}] = r$, such that

$$\det \left| \frac{\partial F_{A_{\mu/1}}^{[J-\mu]}}{\partial q^{a_{i/1}[i+J]}} \right| \neq 0. \tag{67}$$

Thus, we may solve the equations $F^{[J-\mu]}_{A_{\mu/1}}=0$ with respect to the derivatives $q^{a_{i/1}[i+J]}$. Namely,

$$F^{[J-\mu]}_{A_{\mu/1}}=0 \quad\Longleftrightarrow\quad q^{a_{i/1}[i+J]}=\varphi^{a_{i/1}}(\cdots q^{b_{j/1}[j+J-1]},\cdots q^{b_{j/2}[j+J]}). \tag{68}$$

Now we pass from the functions $F_{A_{J/2}}$ to the $\bar F_{A_{J/2}}$ excluding the arguments $q^{b_{i/1}[i+J]}$ from the former,

$$\bar F_{A_{J/2}}=F_{A_{J/2}}|_{f=0}=\bar F_{A_{J/2}}(\cdots q^{b_i[i+J-1]}). \tag{69}$$

The fact that the functions $\bar F_{A_{J/2}}$ do not depend on both $q^{b_{i/1}[i+J]}$ and $q^{b_{i/2}[i+J]}$ is based on lemma 2 in the appendix. Thus, we have the equivalence (see lemma 1 in the appendix)

$$F_A\sim\begin{pmatrix}F_{A_{J/1}}\\ \bar F_{A_{J/2}}\\ F_{A_\nu},\ \nu=0,\ldots,J-1\end{pmatrix}\sim\begin{pmatrix}F_{A_{J/1}}\\ F'_{A'}\end{pmatrix} \tag{70}$$

where

$$F'_{A'}=\left(F'_{A'_\nu},\ \nu=0,\ldots,J-1\right)=\begin{cases}F_{A_\nu}(\cdots q^{b_i[i+\nu]}) & \nu=0,\ldots,J-2\\ F'_{A'_{J-1}}(\cdots q^{b_i[i+J-1]})=\begin{cases}F_{A_{J-1}}\\ \bar F_{A_{J/2}}.\end{cases}\end{cases} \tag{71}$$

Let us turn to the functions $F'_{A'_\nu}$. They have the same structure as in (52) and correspond to the case $\mathcal{N}=I+J<K$. In accordance with the induction hypothesis, supposing, in particular, that the matrix

$$M'_{A'_\nu a_i}=\frac{\partial F'_{A'_\nu}}{\partial q^{a_i[i+\nu]}}$$

has a constant rank in the consideration point, the following equivalence holds true:

$$\begin{gathered}F'_{A'}\sim\begin{pmatrix}q^{a_{i|\sigma}[i+\sigma]}-\varphi^{a_{i|\sigma}}(\cdots q^{b_{j|\sigma_-}[j+\sigma_--1]},\cdots q^{b_{j|\sigma_+}[j+\sigma]},\cdots q^{b_{j|J}[j+\sigma]})\\ 0_{G'}\end{pmatrix}\\ i,j=0,1,\ldots,I \qquad [G']=[A']-[a]+\sum_i[a_{i|J}]\\ \sigma=-I,\ldots,J-1 \qquad -I\leqslant\sigma_-\leqslant\sigma \qquad \sigma+1\leqslant\sigma_+\leqslant J-1.\end{gathered} \tag{72}$$

Taking into account (70), we obtain

$$\begin{gathered}F\sim\begin{pmatrix}F_{A_{J/1}}(\cdots q^{b_i[i+J]})\\ q^{a_{i|\sigma}[i+\sigma]}-\varphi^{a_{i|\sigma}}(\cdots q^{b_{j|\sigma_-}[j+\sigma_--1]},\cdots q^{b_{j|\sigma_+}[j+\sigma]},\cdots q^{b_{j|J}[j+\sigma]})\\ 0_{G'}\end{pmatrix}\\ i,j=0,1,\ldots,I \qquad \sigma=-I,\ldots,J-1 \qquad -I\leqslant\sigma_-\leqslant\sigma \qquad \sigma+1\leqslant\sigma_+\leqslant J-1.\end{gathered} \tag{73}$$

Now we pass from the functions $F_{A_{J/1}}$ to the $\bar F_{A_{J/1}}$ excluding the arguments $q^{a_{i|\sigma}[p_i]}$, $p_i\geqslant i+\sigma$, $\sigma=-I,\ldots,J-1$ from the former. As a result, we have the following equivalence:

$$F\sim\tilde F=\begin{pmatrix}\bar F_{A_{J/1}}(\cdots q^{b_{i|J}[i+J]},\cdots q^{b_{i|\sigma}[i+\sigma-1]})\\ q^{a_{i|\sigma}[i+\sigma]}-\varphi^{a_{i|\sigma}}(\cdots q^{b_{j|\sigma_-}[j+\sigma_--1]},\cdots q^{b_{j|\sigma_+}[j+\sigma]},\cdots q^{b_{j|J}[j+\sigma]})\\ 0_{G'}\end{pmatrix}. \tag{74}$$

The functions $\tilde F$ have the same structure as in (52), however, they depend linearly on a part of highest-order derivatives. Here the supposition of the rank for the matrix

$$\frac{\partial\tilde F_A}{\partial(q^{a_{i|J}[i+J]},q^{a_{i|\sigma}[i+\sigma]})} \qquad A=(A_{J/1},a_{i|\sigma},G') \tag{75}$$

is equivalent to the same supposition for the matrix

$$\frac{\partial \bar{F}_{A_{J/1}}}{\partial q^{b_{i|J}[i+J]}}. \tag{76}$$

Let this rank be equal to $[A_{J/1}]$. In this case there exists a final division of indices,

$$a_{i|J} \to (a_{i|J}, a_{i|J+1}) \qquad \text{with} \quad [a_{i|J}] = [A_{J/1}]$$

such that the equations $\bar{F}_{A_{J/1}} = 0$ can be solved with respect to the derivatives $q^{a_{i|J}[i+J]}$ and we obtain, instead of the first two lines of (74), the following expressions:

$$q^{a_{i|J}[i+J]} - \varphi^{a_{i|J}}(\cdots q^{b_{j|J}[j+J-1]}, \cdots q^{b_{j|\sigma}[j+\sigma-1]}, \cdots q^{b_{j|J+1}[j+J]})$$
$$q^{a_{i|\sigma}[i+\sigma]} - \varphi^{a_{i|\sigma}}(\cdots q^{b_{j|\sigma_-}[j+\sigma_- -1]}, \cdots q^{b_{j|\sigma_+}[j+\sigma]}, \cdots q^{b_{j|J}[j+\sigma]})$$
$$i, j = 0, 1, \ldots, I \qquad \sigma = -I, \ldots, J-1 \qquad -I \leqslant \sigma_- \leqslant \sigma \qquad \sigma + 1 \leqslant \sigma_+ \leqslant J-1.$$

Now, let us put together the first two entries of $\varphi^{a_{i|J}}$ as $\cdots q^{b_{j|\sigma}[j+\sigma-1]}$, $-I \leqslant \sigma \leqslant J$ and recall that for $\sigma = J$ no corresponding σ_+ occurs. Furthermore, let us replace the last entry of $\varphi^{a_{i|\sigma}}$ as follows: $\cdots q^{b_{j|J}[j+\sigma]} \to \cdots q^{b_{j|J}[j+\sigma]} \cdots q^{b_{j|J+1}[j+\sigma]}$, $-I \leqslant \sigma \leqslant J-1$, then we get the missing contribution to σ_+ for the case under consideration. So, we end up exactly with equation (54) and theorem 2 is proved.

If the rank is less than $[A_{J/1}]$ then the above procedure is applied to the functions $\bar{F}_{A_{J/1}}$. Doing that we lower the number of the equations that are not yet reduced to the canonical form (the equations of the type $\bar{F}_{A_{J/1}} = 0$). Note that such a diminution does not happen at the first stage if $[A_{J/2}] = 0$. At a certain stage the procedure does not lower the number of above-mentioned equations. This can happen when the rank of the matrix of type (76) is maximal, i.e. is equal to the number of the functions of the type $\bar{F}_{A_{J/1}}$. In such a case we may reduce them to the canonical form as was mentioned above. This can also happen when we do not obtain the functions of the type $\bar{F}_{A_{J/1}}$ in the reduction procedure. That means that already in the previous step the set is reduced to the case $\mathcal{N} = K - 1$, i.e. the possibility of the reduction to the canonical form is proved.

Finally, we stress that the reduction procedure is formulated for sets of equations of the type (52) (the ELE are a particular case of such sets). The procedure holds true under certain suppositions of ranks. These suppositions demand various Jacobi matrices of the type $\partial F_s/\partial q^{a[l]}$ to have constant ranks in the vicinity of the consideration point. Here $F_s = 0$ are equations obtained at a given stage of the procedure and $q^{a[l]}$ are highest-order derivatives in these equations. It is important to realize that proving the equivalence (56) we prove at the same time the locality of the operators $\hat{U}$ and $\hat{V}$ from (57). In fact, the latter proof is provided by the applicability of the lemmas in the appendix.

5. Gauge identities and action symmetries

It was demonstrated above that in the general case of singular ELE the number of equations in the canonical form is less than the number of equations in the initial set of the differential equations. This reduction is related to the fact that in the canonical form we retain the independent equations only, whereas the initial equations may be dependent. The dependence of the equations in the initial set may be treated as the existence of some identities between the initial equations. The identities between the ELE imply the existence of gauge transformations of the corresponding action. Below we discuss this interrelationship in detail.

First, we introduce some relevant definitions: the relation of the form

$$\hat{R}^a F_a \equiv 0 \tag{77}$$

where $\hat{R}^a$ are some LO, and $F_a(q^{[l]})$ are some LF, is called the identity between the equations $F_a(q^{[l]}) = 0$. The identity sign $\equiv$ means that the left-hand side of (77) is zero for any arguments $q^{[l]}$.

Any set $\hat{\mathbf{R}} = (\hat{R}^a)$ of LO that obeys relation (77) is called the generator of an identity. Whenever $\hat{\mathbf{R}}$ is a generator then $\hat{n}\hat{\mathbf{R}}$ with some LO $\hat{n}$ is a generator as well. Any linear combination $\hat{n}^i\hat{\mathbf{R}}_i$ of some generators $\hat{\mathbf{R}}_i$ with operator coefficients $\hat{n}^i$ is a generator.

A generator $\hat{\mathbf{R}}$ will be called nontrivial if the relation[8] $\hat{n}\hat{\mathbf{R}} = \hat{O}(F)$ can only be provided by a LO $\hat{n}$ of the form $\hat{n} = \hat{O}(F)$.

A set of generators $\hat{\mathbf{R}}_i$ will be called independent if the relation $\hat{n}^i\hat{\mathbf{R}}_i = \hat{O}(F)$ can only be provided by $\hat{n}^i$ of the form $\hat{n}^i = \hat{O}(F)$. Identities generated by independent generators will be called independent.

Note that for any set of LF F_a, there always exist trivial generators. Namely, the generators $\hat{\mathbf{R}}_{\text{triv}} = \left(\hat{R}^a_{\text{triv}}\right) = \hat{O}(F)$ of the form

$$\hat{R}^a_{\text{triv}} = \sum_{k,l} F_b^{[k]} u^{bk|al} \frac{\mathrm{d}^l}{\mathrm{d}t^l} \qquad u^{bk|al} = -u^{al|bk} \tag{78}$$

with arbitrary antisymmetric LF $u^{bk|al}$ obviously lead to the identities (77). These identities are not, however, connected to the mutual dependence of the functions F_a.

An independent set of generators $\hat{\mathbf{R}}_g$ is complete whenever any generator $\hat{\mathbf{R}}$ can be represented in the form $\hat{\mathbf{R}} = \hat{\lambda}^g\hat{\mathbf{R}}_g + \hat{\mathbf{R}}_{\text{triv}}$ with some LO $\hat{\lambda}^g$. Any two complete sets of independent generators $\hat{\mathbf{R}}_g$ and $\hat{\mathbf{R}}'_g$ are related as $\hat{\mathbf{R}}'_g = \hat{U}_g^{g'}\hat{\mathbf{R}}_{g'} + \hat{\mathbf{R}}_{\text{triv}}$, where $\hat{U}$ is an invertible LO.

Supposing now that F_a in equation (77) are functional derivatives of an action, $F_a = \delta S/\delta q^a$, such that $F_a = 0$ are ELE. Let the functions F_a obey all the necessary suppositions of ranks such that ELE can be reduced to the canonical form (54). Let us write here this canonical form as follows[9]:

$$f^\alpha = q^{\alpha[l_\alpha]} - \varphi^\alpha(\cdots q^{\alpha[l_\alpha-1]}; \cdots q^{g[l_g]}) = 0 \qquad a = (\alpha, g) \tag{79}$$

where q^g are gauge coordinates. Moreover, according to theorem 2, there exists the equivalence

$$F_a \sim \bar{F}_a = \begin{pmatrix} f^\alpha \\ 0_g \end{pmatrix} \implies F_a = \hat{U}_a^b \bar{F}_b \qquad \bar{F}_a = \hat{V}_a^b F_b \qquad \hat{U}_a^b \hat{V}_b^c = \delta_a^c \tag{80}$$

where $\hat{U}$ and $\hat{V}$ are LO. Now we may consider the identity (77) as an equation for finding the general form for the generator $\hat{\mathbf{R}}$. Using (80) we transform this problem to the one for finding the operators $\hat{\xi}^a$,

$$\hat{\xi}^a \bar{F}_a \equiv 0 \qquad \hat{R}^a = \hat{\xi}^b \hat{V}_b^a. \tag{81}$$

Using the explicit form (80) of the functions $\bar{F}_a$, we get $\hat{\xi}^a = (\hat{\xi}^\alpha \hat{\xi}^g)$, $a = (\alpha, g)$, where $\hat{\xi}^\alpha$ obey the equation

$$\hat{\xi}^\alpha f^\alpha \equiv 0 \tag{82}$$

and $\hat{\xi}^g$ is a set of arbitrary LO. Since the functions f have the canonical form (80), any solution of equation (82) is presented by trivial generators of the form

$$\hat{\xi}^\alpha = \hat{\xi}^\alpha_{\text{triv}} = \sum_{k,l} \left(\frac{\mathrm{d}^l}{\mathrm{d}t^l} f_{\alpha'}\right) u^{l\alpha'|k\alpha} \frac{\mathrm{d}^k}{\mathrm{d}t^k} \qquad u^{l\alpha'|k\alpha} = -u^{k\alpha|l\alpha'} \tag{83}$$

[8] We denote by $\hat{O}(F)$ LO of the form (3) with all the LF $u^k_{Aa} = O(F)$, where $O(F)|_{F=0} = 0$.

[9] Here, we do not distinguish possible different proper orders of the coordinates.

where $u^{l\alpha'|k\alpha}$ are arbitrary antisymmetric LF. To demonstrate that we present the generators $\hat{\xi}^\alpha$ as $\hat{\xi}^\alpha = \sum_{k=0}^K \xi^{\alpha k}\mathrm{d}^k/\mathrm{d}t^k$, where $\xi^{\alpha k}$ are some LF. Then, in equation (82), we pass from the variables $q^{\alpha[k]}, q^{g[l]}, k, l = 0, 1, \ldots$, to the $q^{\alpha[k_\alpha]}, f_\alpha^{[l]}, q^{g[l]}, k_\alpha = 0, 1, \ldots, l_\alpha - 1, l = 0, 1, \ldots$. Such a variable change is not singular. In terms of the new variables, equation (82) reads

$$\sum_{k=0}^K \xi^{\alpha k} f_\alpha^{[k]} = 0 \qquad K < \infty.$$

Its general solution is well known

$$\xi^{\alpha k} = \sum_l f_{\alpha'}^{[l]} u^{l\alpha'|k\alpha} \qquad u^{l\alpha'|k\alpha} = -u^{k\alpha|l\alpha'}.$$

Now we can write the general solution of equation (81) as

$$\hat{\xi}^a = \hat{\xi}^g \delta_g^a + \hat{\xi}_{\mathrm{triv}}^a \qquad \hat{\xi}_{\mathrm{triv}}^a = \sum_{k,l}\left(\frac{\mathrm{d}^l}{\mathrm{d}t^l}\bar{F}_b\right) u^{lb|ka}\frac{\mathrm{d}^k}{\mathrm{d}t^k} \qquad u^{lb|ka} = -u^{ka|lb}. \tag{84}$$

Let $b = (\alpha', g'), a = (\alpha, g)$ in (84). Then $u^{lg'|k\alpha}, u^{l\alpha'|kg} = -u^{kg|l\alpha'}$ and $u^{lg'|kg}$ are arbitrary LF (e.g., they can be selected to be zero). Indeed, the functions $u^{lg'|k\alpha}$ and $u^{lg'|kg}$ do not enter the expressions for the generators $\hat{\xi}^a$. Besides, terms with $u^{l\alpha'|kg}$ affect only the generators $\hat{\xi}^g$, which are arbitrary by construction. Accordingly, the general solution of equation (77) reads

$$\hat{\mathbf{R}} = \hat{\xi}^g \hat{\mathbf{R}}_g + \hat{\mathbf{R}}_{\mathrm{triv}} \qquad \hat{\mathbf{R}}_g = \left(\hat{R}_g^a = \delta_g^b \hat{V}_b^a = \hat{V}_g^a\right) \tag{85}$$

and

$$\hat{R}_{\mathrm{triv}}^a = \hat{\xi}_{\mathrm{triv}}^b \hat{V}_b^a = \sum_{k,l}\left[\frac{\mathrm{d}^l}{\mathrm{d}t^l}\left(\hat{V}_b^c F_c\right)\right] u^{lb|kd}\frac{\mathrm{d}^k}{\mathrm{d}t^k}\hat{V}_d^a = \sum_{k,l}\left(\frac{\mathrm{d}^l}{\mathrm{d}t^l}F_b\right) T^{lb|ka}\frac{\mathrm{d}^k}{\mathrm{d}t^k}$$

where $T^{lb|ka} = -T^{ka|lb}$ are some LF. The set of generators $\hat{\mathbf{R}}_g = \left(\hat{R}_g^a = \hat{V}_g^a\right)$ is complete and is presented by LO. Moreover, these generators are independent. Indeed, multiplying the equation $\hat{n}^g \hat{R}_g^a = \hat{O}(F)$ from the right by $\hat{U}_a^b$, we get $\hat{n}^g \delta_g^b = \hat{O}(F) \Longrightarrow \hat{n}^g = \hat{O}(F)$.

Thus, there exist the following nontrivial identities between the ELE:

$$\hat{R}_g^a \frac{\delta S}{\delta q^a} \equiv 0 \qquad g = 1, \ldots, r \tag{86}$$

with generators $\hat{\mathbf{R}}_g$ that are LO. These identities are called the gauge identities. As is well known (see, for example, [2, 3]), the existence of the gauge identities (86) implies the existence of infinitesimal gauge transformations of the form

$$q^a \to q^a + \delta q^a \qquad \delta q^a = (\hat{R}^T)_g^a \epsilon^g \tag{87}$$

where r parameters $\epsilon^g = \epsilon^g(t)$ are arbitrary functions of time t. Note that $\hat{\mathbf{R}}^T$ are LO as well.

Thus, it was demonstrated that for theories that obey appropriate suppositions of the ranks there exists a constructive procedure for revealing the gauge generators. For such theories all the generators are LO. The number of independent generators and, therefore, the number of independent gauge transformations is equal to the number of gauge coordinates in the ELE.

As a simple mechanical example, consider the action of the form[10]

$$S = \int L\,\mathrm{d}t \qquad L = \frac{1}{2}(\dot{x} - y)^2 + \frac{a}{2}(y^2 - x^2). \tag{88}$$

[10] At $a \neq 0$ we have a finite-dimensional analogue of the Proca action, and at $a = 0$ we have the analogue of the Maxwell action.

The corresponding ELE are

$$F_1 = \ddot{x} - \dot{y} + ax = 0 \qquad F_2 = \dot{x} - (1+a)y = 0 \tag{89}$$

where $F_2 = 0$ is a constraint. The generalized Hessian reads

$$M = \begin{vmatrix} \frac{\partial^2 L}{\partial \dot{x}^2} = 1 & \frac{\partial^2 L}{\partial \dot{x}\partial y} = -1 \\ \frac{\partial^2 L}{\partial y \partial \dot{x}} = -1 & \frac{\partial^2 L}{\partial y^2} = a+1 \end{vmatrix} = a. \tag{90}$$

Let $a \neq 0$, $M \neq 0$. In such a nonsingular case the reduction procedure looks as follows: with the help of the consistency condition $\dot{F}_2 = 0 \Longrightarrow \ddot{x} = (1+a)\dot{y}$, we eliminate $\ddot{x}$ from the first ELE. Thus, we get an equivalent set, which has the canonical form

$$\dot{y} = -x \qquad \dot{x} = (1+a)y. \tag{91}$$

Another canonical form

$$\ddot{x} = -(1+a)x \qquad y = (1+a)^{-1}\dot{x} \tag{92}$$

we obtain by eliminating $\dot{y}$ from the equation $F_1 = 0$ with the help of the consistency condition $\dot{F}_2 = 0 \Longrightarrow \dot{y} = \ddot{x}/(1+a)$.

Let $a = 0$. The case is singular, $M = 0$, and the rank of the Hessian matrix is equal to 1. One can easily see that the equivalence

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = \hat{U}\begin{pmatrix} \dot{x} - y \\ 0 \end{pmatrix} \qquad \hat{U} = \begin{pmatrix} \mathrm{d}/\mathrm{d}t & 1 \\ 1 & 0 \end{pmatrix} \qquad \hat{U}^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -\mathrm{d}/\mathrm{d}t \end{pmatrix}$$

holds true. Then the canonical form of the ELE reads $\dot{x} = y$ and there is a gauge identity

$$\hat{R}^a F_a \equiv 0 \qquad \hat{R}^1 = 1 \qquad \hat{R}^2 = -\mathrm{d}/\mathrm{d}t.$$

The operators transposed to $\hat{R}^a$ are $(\hat{R}^T)^a = \left((\hat{R}^T)^1 = 1, (\hat{R}^T)^2 = \frac{\mathrm{d}}{\mathrm{d}t}\right)$. Thus, at $a = 0$, the action (88) is invariant under the gauge transformation $x \to x + \epsilon$, $y \to y + \dot{\epsilon}$. In the case under consideration, the ELE have two canonical forms: $\dot{x} = y$ and $y = \dot{x}$.

6. Concluding remarks

We have formulated the reduction procedure which allows one to transform the ELE to the canonical form as well as to establish possible gauge identities between the equations. The latter part of the procedure can be considered as a constructive way of finding all the gauge generators within the Lagrangian formulation. At the same time, it is proved that, for local theories, all the gauge generators are local in time operators. The canonical form of the ELE reveals their hidden structure, in particular, it presents the spectrum of possible initial data, and it allows one to separate coordinates into nongauge and gauge ones. One also ought to remark that the reduction procedure can be, in particular, treated as a procedure for finding constraints in the Lagrangian formulation.

In that respect one can compare the reduction procedure with the well-known Dirac procedure in the Hamiltonian formulation of constrained systems [1–3]. Recall that the Dirac procedure is applicable to the Hamilton equations with primary constraints, namely to equations of the form

$$F(\eta, \dot{\eta}) = \dot{\eta} - \{\eta, H^{(1)}\} = 0 \qquad \Phi^{(1)}(\eta) = 0 \qquad H^{(1)} = H(\eta) + \lambda\Phi^{(1)}(\eta). \tag{93}$$

Here $\eta = (q^a, p_a)$ are phase-space variables; $\Phi^{(1)}(\eta) = 0$ are primary constraints, λ are Lagrange multipliers to the primary constraints, and $H^{(1)}$ is the total Hamiltonian . By $\{\cdot, \cdot\}$ the Poisson bracket is denoted. The aim of the procedure is to eliminate as many λ as possible from the equations, to find all the constraints in the theory. The procedure is based on the consistency conditions $\dot{\Phi}^{(1)} = 0$. Using the equations $F(\eta, \dot{\eta}) = 0$, we may transform any consistency condition to the following form:

$$\dot{\Phi}^{(1)} = \{\Phi^{(1)}, H^{(1)}\} = 0.$$

From these equations one can define some λ as functions of η and reveal some new constraints. Then the procedure has to be applied to the latter constraints and so on.

Equations (93) present a particular case of differential equations considered in the present paper (indeed, these equations are ELE for a Hamiltonian action). Thus, our reduction procedure may be applied to these equations. Namely, first one has to consider the equations $F_A = 0$, $\dot{\Phi}^{(1)} = 0$ and select independent w.r.t. $\dot{\eta}$ equations. Since equations of the primary constraints are independent by construction, we pass to the next step and solve the constraint equations $\Phi^{(1)} = 0$ with respect to a part of the variables η, as $\Phi^{(1)} = 0 \to \eta_1 - \varphi_1(\eta_2) = 0$. Then we exclude η_1 and $\dot{\eta}_1$ from the equations $F = 0$. Thus, we get $F = 0 \to \bar{F}_A(\eta_2, \dot{\eta}_2) = 0$. Then one has to select independent w.r.t. $\dot{\eta}_2$ functions $\bar{F}_{A_{/1}}$. At the same time one finds new constraints $\bar{F}_{A_{/2}}(\eta_2) = 0$ and so on (see section 4.1).

We see that the Dirac procedure differs from our reduction procedure. Indeed, as was mentioned above, in the Dirac procedure one excludes all the derivatives $\dot{\eta}$ with the help of the equations $F = 0$ from the consistency conditions $\dot{\Phi}^{(1)} = 0$. Thus, one gets equations for the Lagrange multipliers and new constraints. Besides, one of the aims of the Dirac procedure is to maintain the canonical Hamiltonian structure of the equations $F = 0$. The possibility of the Dirac reduction is due to the specific structure of equations (93). Namely, here the consistency conditions never involve $\dot{\lambda}$ and rank $\partial F/\partial\dot{\eta} = [F] = [\eta]$.

Besides, one ought to mention the work [14] where an alternative (to the Dirac procedure) way of reducing the equations of motion was proposed for theories with actions of the form $S = \int[\varphi_A(\eta)\dot{\eta}^A - V(\eta)]\,\mathrm{d}t$. One can verify that, in fact, the procedure of that work, in part (the procedure does not reveal the gauge identities), is similar to our reduction procedure in the case of the first-order equations (see section 4).

However, the reduction procedure proposed in the present paper is formulated for a wider class of Lagrangian systems (differential equations). It does not need the introduction of new variables such as momenta and Lagrange multipliers, and is defined in the framework of the initial Lagrangian formulation. Moreover, its aim is twofold: to reduce ELE to their canonical form and to reveal the gauge identities between the ELE equations.

The consideration in the present paper is restricted by finite-dimensional systems. Its application to field theories (theories with infinite number of degrees of freedom) demands additional study. We hope to present the corresponding formulation in future publications. However, in simple cases, one can apply the present reduction procedure with some natural modifications in the infinite-dimensional case. Consider the Maxwell action $S = -\frac{1}{4}\int \mathcal{F}_{\mu\nu}\mathcal{F}^{\mu\nu}\,\mathrm{d}x$, $\mathcal{F}_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, as a common example of a singular field theory. The ELE read

$$F^i = \frac{\partial S}{\partial A_i} = \partial_\nu\mathcal{F}^{i\nu} = \ddot{A}^i + \partial_i\dot{A}^0 - \triangle A^i + \partial_i\varphi = 0 \tag{94}$$

$$F^0 = -\frac{\partial S}{\partial A^0} = \partial_\nu\mathcal{F}^{\nu 0} = \dot{\varphi} + \triangle A^0 = 0, \varphi = \partial_k A^k. \tag{95}$$

The equation $F_0 = 0$ is a constraint. Following the reduction procedure, we have to consider the set $F_i = 0$, $\dot{F}_0 = 0$. The Jacobi matrix $\partial F^\mu/\partial\ddot{A}_\nu$ has the constant rank 3. We can, for example, select equations (94) as independent with respect to the derivatives $\ddot{A}^i$. The equation

$\dot{F}_0 = 0$ is their result. No more constraints appear. Now we exclude A^0 and $\dot{A}^0$ from (94) with the help of (95). That creates the equivalence

$$\begin{pmatrix} F^i \\ F^0 \end{pmatrix} = \begin{pmatrix} \delta^i_k & -\frac{\partial_i \partial_0}{\triangle} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{F}^k \\ F^0 \end{pmatrix} \qquad \bar{F}^k = F^k|_{F^0=0} = \Box(A^k + \partial_k \varphi). \tag{96}$$

Now we discover that the functions $\bar{F}^k$ are dependent, $\partial_k \bar{F}^k \equiv 0$. In our terms that reads, for example, as the following equivalence:

$$\begin{pmatrix} \bar{F}^1 \\ \bar{F}^2 \\ \bar{F}^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\partial_3^{-1}\partial_1 & -\partial_3^{-1}\partial_2 & 1 \end{pmatrix} \begin{pmatrix} \bar{F}^1 \\ \bar{F}^2 \\ 0 \end{pmatrix}. \tag{97}$$

The equations $\bar{F}^1 = 0$, $\bar{F}^2 = 0$, $F^0 = 0$ present one of the canonical forms of the Maxwell equations. The identity that follows from the presence of the zero in the right column of (97) reads $\partial_\mu F^\mu = 0$ in terms of the initial functions F^μ and implies the invariance of the Maxwell action under gradient gauge transformations.

Acknowledgments

BG thanks the foundations FAPESP and DAAD for support and Institute of Physics USP for hospitality. DG is grateful to the foundations FAPESP, CNPq, DAAD for support as well as to the Lebedev Physics Institute (Moscow) and to the Institute of Theoretical Physics (University of Leipzig) for hospitality. IT thanks INTAS 00-00262, RFBR 02-02-16944, 00-15-96566 for partial support.

Appendix

Here we present three lemmas which are used in the reduction procedure to justify the equivalence of equations and LF. In this respect it is useful to recall here the relevant definitions from section 2.

Two sets of equations, $F_A(q^{[l]}) = 0$ and $f_\alpha(q^{[l]}) = 0$ are equivalent $F = 0 \Longleftrightarrow f = 0$ whenever they have the same set of solutions. If two sets of LF $F_A(q^{[l]})$ and $\chi_A(q^{[l]})$, $[F] = [\chi]$, are related by some LO $\hat{U}$ and $\hat{V}$ as $F = \hat{U}\chi$, $\chi = \hat{V}F$, $\hat{U}\hat{V} = 1$, then we call such LF equivalent and denote this fact as $F \sim \chi$. In this case the corresponding equations are strongly equivalent.

Lemma 1. *Let a set of equations*

$$\Phi_\mu(x, y^{[l]}) = 0 \qquad F_a(x, y) = 0 \qquad x = (x^\mu) \qquad y = (y^a) \tag{98}$$

be given, where Φ *are some LF. And let* $\det \partial F_a/\partial y^b|_{x_0,y_0} \neq 0$, *where the consideration point* (x_0, y_0) *is on shell. Then:*

(a) The equations $F_a(x, y) = 0$ *can be solved w.r.t.* y *as* $y^a = \varphi^a(x)$, *where* $\varphi^a(x)$ *are some single-valued functions of* x *in the vicinity of the point* x_0. *In other words, there is the equivalence*

$$F_a(x, y) \sim y^a - \varphi^a(x) \tag{99}$$

which implies strong equivalence between the equations $F_a(x, y) = 0$ *and* $y^a = \varphi^a(x)$.

(b) The following equivalence between the LF holds true:

$$\begin{pmatrix} \Phi_\mu(x, y^{[l]}) \\ F_a(x, y) \end{pmatrix} \sim \begin{pmatrix} \bar{\Phi}_\mu(x^{[l]}) \\ y^a - \varphi^a(x) \end{pmatrix} \qquad \bar{\Phi}_\mu = \Phi_\mu|_{y^{[l]}=\varphi^{[l]}}. \tag{100}$$

The first statement is, in fact, the well-known implicit function theorem [13]. Taking into account (99), we have $F_a(x, y) = u_{ab}(y^a - \varphi^a(x))$, $\det u|_{x_0,y_0} \neq 0$. On the other hand one can write $\Phi_\mu = \bar{\Phi}_\mu + \hat{V}_{\mu a}[y^a - \varphi^a(x)]$, where $\hat{V}_{Aa}$ is a LO. Thus,

$$\begin{pmatrix}\Phi\\ F\end{pmatrix} = \hat{U}\begin{pmatrix}\bar{\Phi}\\ y-\varphi\end{pmatrix} \qquad \hat{U} = \begin{pmatrix}1 & \hat{V}\\ 0 & u\end{pmatrix} \qquad \hat{U}^{-1} = \begin{pmatrix}1 & -\hat{V}\\ 0 & u^{-1}\end{pmatrix} \tag{101}$$

and the equivalence (100) is justified.

Lemma 2. *Let a set of equations*

$$F_A(q, z) = 0 \qquad q = (q^a) \qquad z = (z^i) \qquad A = 1, \ldots, m$$
$$a = 1, \ldots, n \qquad i = 1, \ldots, l$$

be given. And let the Jacobi matrix $\partial F_A/\partial q^a$ have a constant rank in the vicinity D_0 of the consideration point (q_0, z_0), which is on shell $(F_A(q_0, z_0) = 0)$,

$$\operatorname{rank}\frac{\partial F_A}{\partial q^a}\bigg|_{q,z\in D_0} = r. \tag{102}$$

Then there exists an equivalence

$$F_A \sim \bar{F}_A = \begin{pmatrix}y^\mu - \varphi^\mu(x, z)\\ \Omega_G(z)\end{pmatrix} \qquad q^a = (x^g, y^\mu) \qquad A = (\mu, G) \qquad [\mu] = r. \tag{103}$$

We begin the proof with the remark that, due to (102), there exists a division of the indices $A = (\mu, G)$, $a = (\mu, g)$, $[\mu] = r$, $q^a = (x^g, y^\mu)$, such that

$$\det\frac{\partial F_\mu}{\partial y^\nu}\bigg|_{q_0,z_0} \neq 0. \tag{104}$$

Then by virtue of lemma 1 we can write

$$F_\mu = u_{\mu\nu} f^\nu \qquad f^\nu = y^\nu - \varphi^\nu(x, z) \qquad \det u|_{q_0,z_0} \neq 0. \tag{105}$$

Let us present the functions F_G in the form $F_G(x, y, z) = \Omega_G(x, z) + \Pi_{G\mu} f^\mu(x, y, z)$, where $\Omega_G(x, z) = F_G|_{y=\varphi(z,x)}$, such that $\Omega_G(x_0, z_0) = 0$. Then

$$F_A = \begin{pmatrix}F_\mu\\ F_G\end{pmatrix} = U_{AB}\chi_B \qquad \chi_B = \begin{pmatrix}f^\mu\\ \Omega_G\end{pmatrix} U = \begin{pmatrix}u & 0\\ \Pi & 1\end{pmatrix} \qquad \det U|_{q_0,z_0} \neq 0. \tag{106}$$

By virtue of (102) and (106)

$$\operatorname{rank}\frac{\partial \chi_A}{\partial q^a}\bigg|_{q,z\in D_0} = r. \tag{107}$$

The Jacobi matrix $\partial\chi_A/\partial q^a$ has the following structure:

$$\frac{\partial\chi_A}{\partial q^a} = \frac{\partial(f^\mu, \Omega_G)}{\partial(y^\nu, x^g)} = \begin{pmatrix}\delta^\mu_\nu & -\partial\varphi^\mu/\partial x^g\\ 0 & \partial\Omega_G/\partial x^g\end{pmatrix}.$$

Therefore,

$$\operatorname{rank}\frac{\partial\Omega_G}{\partial x^g}\bigg|_{x\in D_0} = 0 \implies \frac{\partial\Omega_G}{\partial x^g}\bigg|_{x,z\in D_0} = 0. \tag{108}$$

Equation (108), together with the relation $\Omega_G(x_0, z_0) = 0$, implies

$$\Omega_G|_{x,z\in D_0} = \Omega_G(z) \qquad \Omega_G(z_0) = 0.$$

Finally, we may write

$$F_A = U_{AB}\chi_B \qquad \chi_B = \begin{pmatrix} f^\mu(x, y, z) \\ \Omega_G(z) \end{pmatrix} \qquad \det U|_{q_0,z_0} \neq 0 \tag{109}$$

Thus, the equivalence (103) is justified.

As a result of lemma 2 the following lemma holds true.

Lemma 3. *Let a set of equations*

$$F_A(q^a) = 0 \qquad A = 1, \ldots, m \qquad a = 1, \ldots, n$$

be given. And let the Jacobi matrix $\partial F_A/\partial q^a$ *have a constant rank in the vicinity* D_0 *of the consideration point* q_0 *which is on shell* ($F(q_0) = 0$),

$$\operatorname{rank} \left. \frac{\partial F_A}{\partial q^a} \right|_{q\in D_0} = r.$$

Then there exists an equivalence

$$F_A \sim \bar{F}_A = \begin{pmatrix} y^\mu - \varphi^\mu(x) \\ 0_G \end{pmatrix} \qquad A = (\mu, G) \qquad 0_G \equiv 0 \quad \forall G \qquad [\mu] = r. \tag{110}$$

The proof of this lemma follows that of lemma 2 if one selects $z = z_0$ there.

References

[1] Dirac P A M 1964 *Lectures on Quantum Mechanics* (New York: Belfer Graduate School of Science, Yeshiva University)

[2] Gitman D M and Tyutin I V 1990 *Quantization of Fields with Constraints* (Berlin: Springer)

[3] Henneaux M and Teitelboim C 1992 *Quantization of Gauge Systems* (Princeton, NJ: Princeton University Press)

[4] Gitman D M and Tyutin I V 2002 Fine structure of constraints in Hamiltonian formulation *Grav. Cosmol.* **8** 138–40

Olshanetsky M and Vainstein A 2002 *Michael Marinov Memorial Volume: Multiple Facets of Quantization and Supersymmetry* (Singapore: World Publishing) pp 184–204

[5] Borochov V A and Tyutin I V 1998 *Yad. Fiz.* **61** 1715 (Engl. transl. *Phys. At. Nucl.* **61** 1603)

Borochov V A and Tyutin I V 1999 *Yad. Fiz.* **62** 1137 (Engl. transl. *Phys. At. Nucl.* **62** 1070)

[6] Batalin I A and Vilkovisky G A 1981 *Phys. Lett.* B **102** 27

Batalin I A and Vilkovisky G A 1983 *Phys. Rev.* D **28** 2567

[7] Gitman D M and Tyutin I V 1983 Canonical quantization of singular theories *Izv. Vuzov Fiz.* **5** 3 (Engl. transl. *1983 Sov. Phys. J.* **5** 423)

[8] Gitman D M and Tyutin I V 1987 The structure of gauge theories in the Lagrangian and Hamiltonian formalisms *Quantum Field Theory and Quantum Statistics* vol I ed I Batalin, C Isham and G Vilkovisky (Bristol: Hilger) pp 143–64

[9] Olver P 1986 *Applications of Lie Groups to Differential Equations (Graduate Texts in Mathematics* vol 107) (Berlin: Springer)

Kupershmidt B A 1980 *Geometry of Jet Bundles and Structure of Lagrangian and Hamiltonian Formalism (Lecture Notes in Mathematics vol 775)* (Berlin: Springer) pp 162–218

Krasilshchik I S, Lychagin V V and Vinogradov A M 1986 Geometry of jet spaces and nonlinear partial differential equations *Adv. Studies Contemp. Math.* (New York: Gordon and Breach)

[10] Gitman D M and Tyutin I V 2002 *Nucl. Phys.* B **630** 509

[11] Gantmacher F R 1959 *The Theory of Matrices* vol 1 (New York: Chelsea)

Gantmacher F R 1986 *Matrizentheorie* (Berlin: Dt. Verlag Wiss)

[12] Gitman D M, Lyakhovich S L and Tyutin I V 1983 *Izv. Vuzov Fiz.* **8** 61 (Engl. transl. *Sov. Phys. J.* **8** 730)

[13] Dieudonné J 1968 *Foundations of Modern Analysis* (New York: Academic)

Fichtenholz G M 1989 *Differential- und Integralrechnung* vol I (Berlin: Dt. Verlag Wiss)

[14] Faddeev L and Jackiw R 1988 *Phys. Rev. Lett.* **60** 1692