TECHNISCHE MECHANIK, Band 20, Heft 2, (2000), 137-146
Manuskripteingang: 24. November 1999

Involution and Constrained Dynamics III: Intrinsic Degrees of Freedom Count

W. M. Seiler

The formal theory of differential equations is applied to constrained dynamics in order to give an intrinsic definition of the number of degrees of freedom of a mechanical system or a field theory. We show how the action of a gauge (pseudo) group can be easily taken into account by some purely formal and combinatorial calculations. As a by-product we exhibit the connection between the Hilbert functions of two differential equations related by a differential relation.

1 Introduction

This is the third part in a series of articles (Seiler and Tucker, 1995; Seiler, 1995a) devoted to the application of the formal theory of differential equations in the context of constrained dynamics. It continues the work described in the first part, correcting and clarifying some points concerning the count of degrees of freedom. Finite- and infinite-dimensional systems are now treated in the same manner. In addition, we give a completely intrinsic introduction into formal theory in contrast to the somewhat more coordinate oriented presentation in the previous parts.

We will study in some detail the problem of counting the degrees of freedom of arbitrary mechanical systems and field theories, respectively. We will not make any assumptions like the system possesses a Lagrangian or Hamiltonian structure. Special emphasis will be put on the correct treatment of gauge symmetries. Of course, this requires first some discussion of what we should understand by degrees of freedom and by gauge symmetries.

We will take the fairly simple (but very effective) point of view of taking the size of the formal solution space, i.e. the space of all formal power series solutions without consideration of their convergence, as a measure for the number of degrees of freedom. With this approach we can avoid a number of interpretation problems when we go beyond standard Lagrangian or Hamiltonian mechanics and it can also handle constrained systems without any modification.

For finite-dimensional systems the size of the formal solution space can be readily measured by the number of free parameters in the power series solution (we neglect here the possible existence of singular integrals and count only the dimension of the general integral). For field theories this question becomes more delicate, as their solution spaces are generally of infinite dimension. We apply ideas from Commutative Algebra by introducing a Hilbert polynomial and define its leading coefficient as the number of degrees of freedom. This is closely related to an approach proposed by Einstein (1955).

Gauge theories require some slight modifications. In such theories a group is acting on the solution space and solutions lying in the same orbit are identified. In general, one is neither able to derive explicitly reduced field equations for some gauge invariant fields nor to perform a full gauge fixing (i.e. to distinguish exactly one solution in each orbit). We will show how one can nevertheless predict the size of the reduced solution space by a simple formal manipulation of Hilbert functions provided the gauge symmetries are represented as a Lie pseudogroup.

The article is organised as follows. The next section briefly reviews the basics of the formal theory of differential equations with special emphasis on the notion of an involutive system. Sect. 3 introduces our basic tools for measuring the size of the formal solution space: the Cartan characters and the Hilbert function. Sect. 4 discusses Lie pseudogroups and the effect of gauge symmetries. In the following two sections the theory is first applied to the finite-dimensional case, i.e. to particle mechanics, and then to field theories. Finally, some conclusions are given in Sect. 7.

2 The Formal Geometry of Differential Equations

The formal theory represents a powerful geometric framework for analysing differential equations based on the jet bundle formalism (Saunders, 1989). In this article we can only briefly review some basic notions. For more details we must refer to the literature, e.g. (Dubois-Violette, 1984; Pommaret, 1978; Seiler, 1994a). We use here an intrinsic coordinate-free approach; in (Seiler and Tucker, 1995) and references therein it is described how the theory can be applied in concrete computations.

Let $\pi : \mathcal{E} \to \mathcal{B}$ be a fibred manifold. Geometrically, the first order jet bundle $\pi_0^1 : J_1\mathcal{E} \to \mathcal{E}$ can most easily be described by saying that it is an *affine* bundle and that its fibre over a point $\xi = (x,u) \in \mathcal{E}$ is the following affine space

$$(J_1\mathcal{E})_\xi = \{\ \gamma \in T_x^*\mathcal{B} \otimes T_\xi\mathcal{E} \mid T\pi \circ \gamma = \mathrm{id}_{T_x\mathcal{B}}\ \} \tag{1}$$

modelled on the vector space $T_x^*\mathcal{B} \otimes V_\xi\mathcal{E}$. As $J_1\mathcal{E}$ may again be considered as a fibred manifold over $\mathcal{B}$, we can iterate this construction. Higher order jet bundles are then obtained by identifying in an obvious manner $J_{q+r}\mathcal{E}$ with a subbundle of $J_r(J_q\mathcal{E})$. For any $q, r \geq 0$ they possess bundle structures $\pi_q^{q+r} : J_{q+r}\mathcal{E} \to J_q\mathcal{E}$ where we set $J_0\mathcal{E} = \mathcal{E}$ and a structure as fibred manifold $\pi^q : J_q\mathcal{E} \to \mathcal{B}$.

A section $\sigma : \mathcal{B} \to \mathcal{E}$ can be *prolonged* to a section $j_1(\sigma) : \mathcal{B} \to J_1\mathcal{E}$ by $j_1(\sigma)(x) = (\sigma(x), T_x\sigma)$. As above, this construction can naturally be extended to higher order jet bundles. We define now a *differential equation* (of order q) as a fibred submanifold $\mathcal{R}_q$ of $\pi^q : J_q\mathcal{E} \to \mathcal{B}$. A section $\sigma : \mathcal{U} \subset \mathcal{B} \to \mathcal{E}$ is a *(local) solution* of the equation $\mathcal{R}_q$, if $j_q(\sigma)(\mathcal{U}) \subset \mathcal{R}_q$.

There exist two natural geometric operations with a differential equation $\mathcal{R}_q$: *projection* and *prolongation*. The first one is inherited from the canonical projections between jet bundles of different order. If $\mathcal{R}_{q+r}$ is an equation of order $q+r$, we define the projected equation of order q by $\mathcal{R}_q^{(r)} = \pi_q^{q+r}(\mathcal{R}_{q+r})$. Conversely, a q^{th} order differential equation $\mathcal{R}_q$ can be prolonged to one of order $q+r$: $\mathcal{R}_{q+r} = J_r(\mathcal{R}_q) \cap J_{q+r}\mathcal{E}$ (the intersection is understood to take place in $J_r(J_q\mathcal{E})$). In general, we cannot expect that either projection or prolongation leads again to a fibred manifold. However, for simplicity, we will make this assumption in the sequel, i.e. we restrict to so-called *regular* equations.

For our purposes, equations of the form $\mathcal{R}_{q+r}^{(s)} = \pi_{q+r}^{q+r+s}(\mathcal{R}_{q+r+s})$, i.e. equations which were first prolonged $r+s$ times and then projected back s times, are especially important. Note that for $r = 0$ in general $\mathcal{R}_q^{(s)} \subset \mathcal{R}_q$. This indicates the presence of *integrability conditions.* We call $\mathcal{R}_q$ *formally integrable*, if for all $r \geq 0$ the equality $\mathcal{R}_{q+r}^{(1)} = \mathcal{R}_{q+r}$ holds, i.e. at no order of prolongation integrability conditions occur. The name stems from the fact that for such equations it is straight-forward to construct formal power series solutions. Unfortunately, no finite criterion for formal integrability is known so far.

An important property of jet bundles is that $\pi_{q-1}^q : J_q\mathcal{E} \to J_{q-1}\mathcal{E}$ defines an *affine* bundle modelled on the vector bundle $S_qT^*\mathcal{B} \otimes V\mathcal{E}$ with S_q denoting the symmetric product. This leads naturally to the concept of the *symbol* of a differential equation. Let $\xi \in \mathcal{R}_q \subset J_q\mathcal{E}$; the symbol $\mathcal{M}_q$ of $\mathcal{R}_q$ is a family of vector spaces over $\mathcal{R}_q$ with $(\mathcal{M}_q)_\xi = V_\xi^{(q)}\mathcal{R}_q \subset V_\xi^{(q)}J_q\mathcal{E}$ where $V_\xi^{(q)}J_q\mathcal{E} = \ker T_\xi\pi_{q-1}^q$ is the vertical space with respect to the projection π_{q-1}^q. We can thus identify the symbol with a subspace of $S_qT^*\mathcal{B} \otimes V\mathcal{E}$. For simplicity, we will assume in the sequel that the symbol is not just a family of vector spaces but actually a vector bundle over $\mathcal{R}_q$. Note that while our definition of a symbol is closely related to the standard one in text books on differential equations, it is not the same! Written out in local coordinates, our symbol corresponds to a much larger matrix; the classical (principal) symbol is obtained by a kind of contraction with a one-form $\chi \in T^*\mathcal{B}$.

Consider the map $\delta : S_{r+1}T^*\mathcal{B} \to T^*\mathcal{B} \otimes S_rT^*\mathcal{B}$ defined by composition of the natural inclusion $S_{r+1}T^*\mathcal{B} \hookrightarrow T^*\mathcal{B} \otimes \bigotimes^r T^*\mathcal{B}$ with the canonical projection $T^*\mathcal{B} \otimes \bigotimes^r T^*\mathcal{B} \to T^*\mathcal{B} \otimes S_rT^*\mathcal{B}$. By wedging with $\Lambda_sT^*\mathcal{B}$ (the s-fold exterior product of $T^*\mathcal{B}$) and tensoring with $V\mathcal{E}$ we can extend δ to a map $\Lambda_sT^*\mathcal{B} \otimes S_{r+1}T^*\mathcal{B} \otimes V\mathcal{E} \to \Lambda_{s+1}T^*\mathcal{B} \otimes S_rT^*\mathcal{B} \otimes V\mathcal{E}$. This leads to the δ-sequences

$$\begin{aligned} 0 \longrightarrow S_rT^*\mathcal{B} \otimes V\mathcal{E} \longrightarrow T^*\mathcal{B} \otimes S_{r-1}T^*\mathcal{B} \otimes V\mathcal{E} \longrightarrow \cdots \\ \longrightarrow \Lambda_sT^*\mathcal{B} \otimes S_{r-s}T^*\mathcal{B} \otimes V\mathcal{E} \longrightarrow \cdots \longrightarrow \Lambda_nT^*\mathcal{B} \otimes S_{r-n}T^*\mathcal{B} \otimes V\mathcal{E} \longrightarrow 0 \end{aligned} \tag{2}$$

where we set $S_iT^*\mathcal{B} = 0$ for $i < 0$ and where $n = \dim \mathcal{B}$. The so-called *formal Poincaré lemma* states that these sequences are exact for all $r > 0$.

The prolongation of a symbol $\mathcal{M}_q \subset S_qT^*\mathcal{B} \otimes V\mathcal{E}$ can be directly computed as the intersection $\mathcal{M}_{q+r} = (S_rT^*\mathcal{B} \otimes \mathcal{M}_q) \cap (S_{q+r}T^*\mathcal{B} \otimes V\mathcal{E})$ (which is understood to take place in $\bigotimes^{q+r} T^*\mathcal{B} \otimes V\mathcal{E}$). Setting $\mathcal{M}_i = 0$ for $i < 0$ and $\mathcal{M}_i = S_iT^*\mathcal{B} \otimes V\mathcal{E}$ for $0 \le i < q$ the δ-sequence (2) can be restricted to a sequence

$$0 \longrightarrow \mathcal{M}_{q+r} \longrightarrow T^*\mathcal{B} \otimes \mathcal{M}_{q+r-1} \longrightarrow \cdots \longrightarrow \Lambda_n T^*\mathcal{B} \otimes \mathcal{M}_{q+r-n} \longrightarrow 0 \tag{3}$$

which is still a complex but in general no longer exact. Its (bigraded) cohomology is called the *Spencer cohomology* of the symbol $\mathcal{M}_q$. We denote by $H^{s,r}(\mathcal{M}_q)$ the cohomology group at $\Lambda_s T^*\mathcal{B} \otimes \mathcal{M}_r$ and define $\mathcal{M}_q$ to be *involutive*, if $H^{s,q+r}(\mathcal{M}_q) = 0$ holds for all $0 \le s \le n$ and all $r \ge 0$. The differential equation $\mathcal{R}_q$ is *involutive*, if it is formally integrable and if its symbol $\mathcal{M}_q$ is involutive.

It is important to note that all these constructions based on the symbol are to be understood pointwise on $\mathcal{R}_q$. In general, one must expect that the dimensions of the cohomology classes vary from point to point and so it may happen that $\mathcal{R}_q$ is involutive at some points and not at other ones. Again we will assume for simplicity that this does not happen.

Involutive equations share many special properties; e. g. in the analytic category a very general existence and uniqueness theorem for (non-characteristic) initial value problems holds: the *Cartan-Kähler theorem.* According to the *Cartan-Kuranishi theorem* any regular equation $\mathcal{R}_q$ either can be completed in a finite number of prolongations and projections to an equivalent involutive equation $\mathcal{R}^{(s)}_{q+r}$ or it is inconsistent. Computational aspects of this completion and its implementation in a computer algebra system are studied in (Schü et al., 1994); its relation to the classical constraint algorithm of Dirac is shown in (Seiler and Tucker, 1995).

3 Cartan Characters and the Hilbert Function

Let $\{e_1, \dots, e_n\}$ be an ordered basis of $T\mathcal{B}$. For a given symbol $\mathcal{M}_q$ we define for $1 \le k \le n$ the following subspaces (again everything should be understood pointwise on $\mathcal{R}_q$)

$$\mathcal{M}_{q,k} = \{\, \rho \in \mathcal{M}_q \mid \rho(e_i, v_1, \dots, v_{q-1}) = 0\,, \quad \forall\, 1 \le i \le k\,,\ \forall v_1, \dots, v_{q-1} \in T\mathcal{B} \,\} \tag{4}$$

and $\mathcal{M}_{q,0} = \mathcal{M}_q$. We call the basis *quasi-regular* (for $\mathcal{M}_q$), if

$$\dim \mathcal{M}_{q+1} = \sum_{k=0}^{n-1} \dim \mathcal{M}_{q,k} \tag{5}$$

One can show that a symbol is involutive, if and only if quasi-regular bases exist for it. Given such a basis for the involutive symbol $\mathcal{M}_q$, we can compute its *Cartan characters* $\alpha_q^{(k)}$ (for $1 \le k \le n$) as the differences $\alpha_q^{(k)} = \dim \mathcal{M}_{q,k-1} - \dim \mathcal{M}_{q,k}$. They form a descending sequence

$$\alpha_q^{(1)} \ge \alpha_q^{(2)} \ge \cdots \ge \alpha_q^{(n)} \ge 0 \tag{6}$$

We define the *Hilbert function* $H(s)$ of the differential equation $\mathcal{R}_q$ as the number of arbitrary coefficients of order s of the general formal power series solution of $\mathcal{R}_q$. It is probably the most natural and useful measure for the size of the formal solution space. If $\mathcal{R}_q$ is involutive, the Hilbert function is polynomial for $s \ge q$ and given by $H(s) = \dim \mathcal{M}_s$. One can derive a closed form expression for the Hilbert polynomial in terms of the Cartan characters of $\mathcal{M}_q$:

$$H(q+s) = \sum_{k=1}^{n} \alpha_q^{(k)} \binom{q+s}{k}, \quad \forall s \ge 0 \tag{7}$$

For $0 < s < q$ the values of the Hilbert function can be computed as $H(s) = \dim \mathcal{R}_s^{(q-s)} - \dim \mathcal{R}_{s-1}^{(q-s+1)}$ and $H(0) = \dim \mathcal{R}_0^{(q)}$. With a bit of combinatorics one can express H as given by (7) explicitly as a polynomial in s. Conversely, if the Hilbert polynomial is given in the explicit form $H(q+s) = \sum_{i=0}^{n-1} h_i s^i$, the corresponding Cartan characters $\alpha_q^{(k)}$ are determined by the recurrence relation

$$\alpha_q^{(k)} = (k-1)!\, h_{k-1} - \sum_{i=k+1}^{n} \frac{(k-1)!}{(i-1)!}\, s_{i-k}^{(i-1)}\, \alpha_q^{(i)} \tag{8}$$

with the modified Stirling numbers $s_k^{(j)} = \sigma_k^{(j)}(1,2,\dots,j)$ where $\sigma_k^{(j)}$ denotes the elementary symmetric polynomial of degree k in j variables. Note that (8) depends indirectly on the order q of the differential equation, since we start with $H(q+s)$ written as polynomial in s.

As intrinsically defined objects the Cartan characters $\alpha_q^{(k)}$ and the Hilbert function $H(s)$ are of course invariant under arbitrary changes of coordinates in $\mathcal{E}$. For more general transformations — like rewriting a higher order system as a first order one — this is not necessarily the case. We exhibit now for two differential equations connected by a differential relation a simple equation for their Hilbert polynomials.

Let $\pi : \mathcal{E} \to \mathcal{B}$ and $\bar{\pi} : \bar{\mathcal{E}} \to \mathcal{B}$ be two fibred manifolds over the same base manifold $\mathcal{B}$. We define a *differential relation* between $\mathcal{E}$ and $\bar{\mathcal{E}}$ as a submanifold $\mathcal{T}_{r,\bar{r}} \subset J_r\mathcal{E} \times J_{\bar{r}}\bar{\mathcal{E}}$ such that $\mathcal{T}_{r,\bar{r}}$ is a fibred submanifold with respect to the two fibrations induced by π and $\bar{\pi}$, namely $\pi^r \times \mathrm{id} : J_r\mathcal{E} \times J_{\bar{r}}\bar{\mathcal{E}} \to \mathcal{B} \times J_{\bar{r}}\bar{\mathcal{E}}$ and $\mathrm{id} \times \bar{\pi}^{\bar{r}} : J_r\mathcal{E} \times J_{\bar{r}}\bar{\mathcal{E}} \to J_r\mathcal{E} \times \mathcal{B}$. This allows us for a given section $\sigma : \mathcal{B} \to \mathcal{E}$ to consider $\mathcal{T}_{\bar{r}}[\sigma] = \mathcal{T}_{r,\bar{r}} \cap \left(j_r\sigma(\mathcal{B}) \times J_{\bar{r}}\bar{\mathcal{E}}\right)$ as a differential equation in $J_{\bar{r}}\bar{\mathcal{E}}$ and similarly for a section $\bar{\sigma} : \mathcal{B} \to \bar{\mathcal{E}}$ to consider $\mathcal{T}_r[\bar{\sigma}] = \mathcal{T}_{r,\bar{r}} \cap \left(J_r\mathcal{E} \times j_{\bar{r}}\bar{\sigma}(\mathcal{B})\right)$ as a differential equation in $J_r\mathcal{E}$.

Now let $\mathcal{R}_q \subset J_q\mathcal{E}$ and $\bar{\mathcal{R}}_{\bar{q}} \subset J_{\bar{q}}\bar{\mathcal{E}}$ be two differential equations. We say that they are *differentially related* via $\mathcal{T}_{r,\bar{r}}$, if the following condition is satisfied. If the section σ is a solution of $\mathcal{R}_q$, then every solution of $\mathcal{T}_{\bar{r}}[\sigma]$ is also a solution of $\bar{\mathcal{R}}_{\bar{q}}$, and conversely if $\bar{\sigma}$ is a solution of $\bar{\mathcal{R}}_{\bar{q}}$, then every solution of $\mathcal{T}_r[\bar{\sigma}]$ is also a solution of $\mathcal{R}_q$. One may consider this as a very general form of Bäcklund transformations. For simplicity, we assume that all differential equations involved are involutive.

Denote by $H(s)$ the Hilbert function of $\mathcal{R}_q$ and by $\bar{H}(s)$ the one of $\bar{\mathcal{R}}_{\bar{q}}$. For $\mathcal{T}_{r,\bar{r}}$ we can introduce two Hilbert functions $G(s)$ and $\bar{G}(s)$; in the first case it is considered as a differential equation $\mathcal{T}_r[\bar{\sigma}] \subset J_r\mathcal{E}$, in the latter one as an equation $\mathcal{T}_{\bar{r}}[\sigma] \subset J_{\bar{r}}\bar{\mathcal{E}}$ (assuming that the Hilbert functions are independent of σ and $\bar{\sigma}$). If $\mathcal{T}_{r,\bar{r}}$ is a differential relation between $\mathcal{R}_q$ and $\bar{\mathcal{R}}_{\bar{q}}$, the four Hilbert polynomials satisfy

$$H(s) - G(s) = \bar{H}(s + \bar{r} - r) - \bar{G}(s + \bar{r} - r) \tag{9}$$

Note that in general this equation holds only for the Hilbert polynomials and not for the full Hilbert functions, i.e. it holds only, if all arguments are greater than or equal to $\max\{r, \bar{r}\}$.

This can be seen as follows. The general formal power series solution of $\mathcal{R}_q$ is parametrised at order s by $H(s)$ Taylor coefficients. To each solution σ of $\mathcal{R}_q$ there corresponds via the differential relation $\mathcal{T}_{r,\bar{r}}$ a family of solutions of $\bar{\mathcal{R}}_{\bar{q}}$ parametrised at order s by $\bar{G}(s)$ coefficients. However, each of these solutions of $\bar{\mathcal{R}}_{\bar{q}}$ can be obtained not only by starting with σ but also by starting with any solution in a whole family parameterised at order s by $G(s)$ coefficients (obtained by applying $\mathcal{T}_{r,\bar{r}}$ "backwards"). Since for $s \geq \max\{r, \bar{r}\}$ a coefficient of σ of order s corresponds via $\mathcal{T}_{r,\bar{r}}$ to coefficients of order $s + \bar{r} - r$, the general formal power series solution of $\bar{\mathcal{R}}_{\bar{q}}$ is parameterised by $\bar{H}(s+\bar{r}-r) = H(s) - G(s) + \bar{G}(s+\bar{r}-r)$ Taylor coefficients of order $s + \bar{r} - r$.

For lower values of s (9) does not need to hold. This is due to the fact that then in general we cannot say that the coefficients of σ of order s are related to coefficients of $\bar{\sigma}$ of order $s + \bar{r} - r$. We only know that the coefficients of σ *up to order* r correspond to the coefficients of $\bar{\sigma}$ *up to order* $\bar{r}$. Only when $\mathcal{T}_{r,\bar{r}}$ has a special form, we may be able to make stronger statements. Especially, if both $\mathcal{T}_r[\bar{\sigma}]$ and $\mathcal{T}_{\bar{r}}[\sigma]$ are of Cauchy-Kowalevsky form, (9) holds for the full Hilbert functions.

An important special case is $G(s) = \bar{G}(s + \bar{r} - r)$. In this case there exists a one-to-one correspondence between solutions of $\mathcal{R}_q$ and $\bar{\mathcal{R}}_{\bar{q}}$ and their Hilbert polynomials are related by a simple shift of the argument. This happens for example for the classical Bäcklund transformations. If $H(s) = \bar{H}(s+\bar{r}-r)$, there also exists a simple relation between the Cartan characters $\alpha_q^{(k)}$ and $\bar{\alpha}_{\bar{q}}^{(k)}$ of $\mathcal{R}_q$ and $\bar{\mathcal{R}}_{\bar{q}}$, respectively. If $\bar{r} \geq r$, the $\bar{\alpha}_{\bar{q}}^{(k)}$ coincide with the characters of the $(\bar{r} - r)$-fold prolongation of $\mathcal{R}_q$. Otherwise, the $\alpha_q^{(k)}$

coincide with the characters of the $(r - \bar{r})$-fold prolongation of $\bar{\mathcal{R}}_{\bar{q}}$.

Under prolongations the highest non-vanishing Cartan character $\alpha_q^{(k_0)}$ (i. e. $\alpha_q^{(k)} = 0$ for $k > k_0$ and $\alpha_q^{(k_0)} \neq 0$) remains unchanged: if $\alpha_{\bar{q}}^{(k)}$ are the Cartan characters of the prolongation $\mathcal{R}_{\bar{q}}$ for $\bar{q} > q$, then $\alpha_{\bar{q}}^{(k)} = 0$ for all $k > k_0$ and $\alpha_{\bar{q}}^{(k_0)} = \alpha_q^{(k_0)}$. Thus for all differential equations with identical Hilbert polynomials, k_0 and $\alpha_q^{(k_0)}$ are equal. We call $\alpha_q^{(k_0)}$ the *index of generality* of $\mathcal{R}_q$ and k_0 its *Cartan genus.*

4 Gauge Symmetries

An important aspect of counting degrees of freedom is to take into account the effect of gauge symmetries. They lead to a reduction of the number of degrees of freedom, as one identifies solutions related by a gauge transformation or, more precisely, such solutions are considered as representing the same physical state. Thus gauge symmetries are a matter of physical interpretation; for arbitrary equations of motion there exists no mathematical criterion which tells us when a symmetry should be "promoted" to a gauge symmetry (the situation is somewhat different, if we restrict to Lagrangian or Hamiltonian systems). We will therefore assume in the sequel that we know already the gauge symmetries and treat only the problem of computing the corresponding correction of the number of degrees of freedom. In fact, this is the common situation in most physical applications.

(Seiler, 1994b) and (Seiler, 1995b) present already solutions to this problem. There it is assumed that the gauge symmetries are given in the form of explicit transformations. While this is the usual form in physics, it is rather cumbersome from a mathematical point of view. One problem is that the action of the gauge group may not be effective. In addition, a number of artificial assumptions has to be made about the dependence of the transformation on the gauge functions or parameter.

In this article we will instead assume that the gauge symmetries are given in the form of a Lie pseudogroup (Pommaret, 1978), i. e. in form of a differential equation. This allows us to apply the same techniques to determine the size of the group (independent of whether or not its action is effective) and of the formal solution space of the field equations. In this approach the gauge correction becomes a simple subtraction of Hilbert functions very similar to (9).

Let $\mathcal{E}$ be an arbitrary manifold. Symmetry transformations are diffeomorphisms $\mathcal{E} \to \mathcal{E}$ or alternatively sections of the trivial bundle $\mathrm{pr}_1 : \mathcal{E} \times \mathcal{E} \to \mathcal{E}$ where pr_1 denotes the source projection on the first factor. Taking the latter point of view, we can build jet bundles $J_r(\mathcal{E} \times \mathcal{E})$ over this bundle; the jets of invertible maps define an open subbundle $I_r(\mathcal{E} \times \mathcal{E})$. This allows us to represent transformations as solutions of differential equations.

We define now a *Lie pseudogroup* as a differential equation $\mathcal{G}_r \subset I_r(\mathcal{E} \times \mathcal{E})$ (often called a *finite Lie equation*) satisfying the following two conditions which correspond to the classical axioms of a (local) group: (i) if the sections $\sigma : \mathcal{U} \subset \mathcal{E} \to \mathcal{E}$ and $\tau : \mathcal{V} \subset \mathcal{E} \to \mathcal{E}$ are two (local) solutions of $\mathcal{G}_r$ with $\sigma(U) \cap \mathcal{V} \neq \emptyset$, then the section $\tau \circ \sigma : \sigma(U) \cap \mathcal{V} \to \mathcal{E}$ is also a (local) solution of $\mathcal{G}_r$; (ii) the same holds for the section $\sigma^{-1} : \sigma(\mathcal{U}) \to \mathcal{U}$.

In our case $\mathcal{E}$ is not an arbitrary manifold but the total space of a fibred manifold $\pi : \mathcal{E} \to \mathcal{B}$. A *gauge transformation* preserves the fibration, i. e. it factors through π. We will further assume that the induced mapping on $\mathcal{B}$ is the identity. Given a differential equation $\mathcal{R}_q \subset J_q\mathcal{E}$, such a transformation is called a *symmetry,* if it maps solutions of $\mathcal{R}_q$ again into solutions.

Let $H(s)$ be the Hilbert function of $\mathcal{R}_q$ and let $\mathcal{G}_r$ be a Lie pseudogroup of gauge symmetries with Hilbert function $G(s)$. Assume that we were able to perform a complete gauge fixing, i. e. to derive a differential equation $\bar{\mathcal{R}}_{\bar{q}}$ such that any solution of $\bar{\mathcal{R}}_{\bar{q}}$ is also a solution of $\mathcal{R}_q$, but that of each orbit of the gauge pseudogroup in the solution space of $\mathcal{R}_q$ precisely one solution is also a solution of $\bar{\mathcal{R}}_{\bar{q}}$. Locally, this is (in principle) always possible. We show now by an argument very similar to the proof of (9) that the Hilbert function $\bar{H}(s)$ of the gauge fixed equation $\bar{\mathcal{R}}_{\bar{q}}$ is given by

$$\bar{H}(s) = H(s) - G(s) \tag{10}$$

If we expand the general solution of $\mathcal{R}_q$ in a formal power series, the value $H(s)$ of the Hilbert function tells the number of arbitrary Taylor coefficients of order s. On the other hand, $G(s)$ equals the number of coefficients of order s that can be given arbitrary values via gauge transformations. Thus modulo this gauge freedom only $\bar{H}(s)$ arbitrary coefficients remain. In contrast to the last section, (10) holds for the full Hilbert functions and not just for the Hilbert polynomials.

5 Particle Mechanics

We use here the term "particle mechanics" somewhat loosely as an abbreviation for finite-dimensional systems. In this case the base manifold $\mathcal{B}$ is one-dimensional and represents the time (or more generally some evolution parameter). In many applications the total space $\mathcal{E}$ will be of the simple form $\mathcal{E} = \mathcal{B} \times \mathcal{Q}$ where $\mathcal{Q}$ denotes either the classical configuration space or the phase space of the system, i. e. some m-dimensional manifold. The former case corresponds to Lagrangian mechanics where one is dealing with some differential equation $\mathcal{R}_{2k} \subset J_{2k}\mathcal{E}$ for a Lagrangian of order k; the latter case appears in Hamiltonian mechanics where one studies a first order equation $\mathcal{R}_1 \subset J_1\mathcal{E}$.

Completion to involution takes here a particular simple form, as the symbol of an ordinary differential equation is always involutive. It amounts simply to checking whether prolonging lower order equations yields some new equations (this is equivalent to the tangency condition in the classical approaches based on vector fields). Note that there is absolutely no difference whether we complete Lagrangian or Hamiltonian, holonomic or anholonomic, first or higher order systems etc.

For a mechanical system without gauge symmetry one expects that the Cartan character vanishes. Otherwise the system is underdetermined which in general makes sense only if the indeterminacy stems from a gauge symmetry. If $\mathcal{R}_q$ is an involutive equation of motion with vanishing Cartan character, we define the number N of degrees of freedom by the simple formula

$$N = \dim \mathcal{R}_q \tag{11}$$

N counts the number of independent conditions that can be imposed on the equation of motion, as under the made assumptions we have in the general formal power series solution $\dim \mathcal{R}_q$ Taylor coefficients that are undetermined by the differential equation and must be prescribed by initial or boundary conditions.

If one applies (11) to a simple system like a pendulum, one obtains double the number of degrees of freedom than in the classical approach, as we treat position and velocity (or momentum, respectively) as separate degrees of freedom. If a natural "pairing" of the coordinates exists (in the case of higher order Lagrangians such a "pair" might consist of more than two coordinates), one can divide in (11) by the corresponding factor. Note, however, that the abstract theory cannot guarantee that the result of the division will be integral. This remains to be shown for the specific formalism one uses. A classical example for such a proof appears in the Dirac theory where one can easily show that there is always an even number of second class constraints (Dirac, 1950).

Besides this problem of obtaining an integral number of degrees of freedom, our main reason for counting positions and velocities separately is that this avoids interpretation problems in the treatment of more general systems like those subject to anholonomic constraints or systems described by Lagrangians linear in the velocities (or more generally the highest order derivatives). In the first case, one typically still has a natural pairing of position and velocity coordinates, but the constraints affect only the velocities. Thus it is not clear how the constraints should be subtracted. In the second case, the Euler-Lagrange equations are of lower order than in the generic case. Insisting on a pairing excludes the use of simpler approaches like the Faddeev-Jackiw formalism (Seiler, 1995a).

If the equations of motion form an underdetermined system, the gauge group should have precisely the same Hilbert polynomial as the equations of motion, i. e. the reduced Hilbert polynomial $\bar{H}(s)$ introduced in the last section must vanish. Let now $\mathcal{R}_q$ be the equation of motion, completed to involution, and $\mathcal{G}_r$ the gauge pseudogroup, also completed to involution. If $s = \max\{q, r\}$, then we obtain for the number $\bar{N}$ of degrees of freedom

$$\bar{N} = \dim \mathcal{R}_s - \dim \mathcal{G}_s \tag{12}$$

Often a mechanical system can be modelled in many different ways leading to very different equations of motion. However, if the various models are consistent, their solution spaces must be in one-to-one correspondence (after subtracting gauge symmetries). The results at the end of Sect. 3 ensure that our definition yields always the same number of degrees of freedom.

As an example we study a simple three-dimensional system due to Vladimir Gerdt (private communication) where the formula in (Seiler and Tucker, 1995) gives a wrong result. The error can be best explained in the language of the Dirac theory (Dirac, 1950): the effect of secondary first class constraints was ignored, as the gauge correction was solely based on the Cartan character counting only the primary first class constraints (only these correspond to arbitrary functions in the gauge transformations). The approach via pseudogroups avoids this problem. It also does not matter whether we use the Lagrangian or the Hamiltonian framework, so for simplicity we perform our computations on the Lagrangian side.

The system is defined by the Lagrangian $L = q_1(\dot{q}_2 - q_3) - \dot{q}_1 q_2$ which is linear in the velocities. Its Euler-Lagrange equations form an involutive first order system

$$\mathcal{R}_1 : \begin{cases} 2\dot{q}_2 - q_3 = 0\,, \quad \dot{q}_1 = 0\,, \\ q_1 = 0 \end{cases} \tag{13}$$

For $s \geq 1$ the Hilbert function is given by $H(s) = 1$ and $H(0) = \dim \mathcal{R}_0^{(1)} = 2$.

The Lagrangian L is invariant (up to a total derivative) under gauge transformations of the form $\bar{q}_1 = q_1$, $\bar{q}_2 = q_2 + \frac{1}{2}\eta$ and $\bar{q}_3 = q_3 + \dot{\eta}$ where $\eta = \eta(t)$ is an arbitrary function. This pseudogroup can be described as the solution space of the following involutive Lie equation

$$\mathcal{G}_1 : \begin{cases} \dfrac{\partial \bar{q}_i}{\partial q_j} = \delta_i^j\,, \quad \dfrac{\partial \bar{q}_1}{\partial t} = 0\,, \quad \dfrac{\partial \bar{q}_2}{\partial t} = \dfrac{1}{2}(\bar{q}_3 - q_3)\,, \\ \bar{q}_1 = q_1 \end{cases} \tag{14}$$

For $s \geq 1$ the Hilbert function is $G(s) = 1$ (thus as expected identical to the one of the equations of motion) and $G(0) = \dim \mathcal{G}_0^{(1)} = 2$. Since $\dim \mathcal{G}_1 = \dim \mathcal{R}_1 = 3$, we see that this dynamical system possesses no degrees of freedom. The same result can be obtained within the classical Dirac theory where one finds three first class constraints only one of which is primary. This reflects the fact that the gauge transformations depend on only one arbitrary function η.

6 Field Theories

Now we proceed to the case that the base manifold $\mathcal{B}$ is of higher dimension, say n. In physical applications, $\mathcal{B}$ will typically represent space-time. Very often, $\pi : \mathcal{E} \to \mathcal{B}$ will be a vector or a tensor bundle; however, we do not need this additional structure for our purposes. As explained in detail in (Seiler and Tucker, 1995), the naive generalisation of the Dirac algorithm often used in the physics literature does in general not suffice for a consistency check of the field equations. Only a full completion to involution provides such a check and thus a valid constraint algorithm.

Without loss of generality we may assume that (after completion to involution) the field equations form a first order system. For such systems the Cartan-Kähler theorem yields a direct interpretation of the Cartan characters $\alpha_1^{(k)}$ as the number of arbitrary functions which can be prescribed as initial conditions. Namely, we can prescribe $\alpha_1^{(n)}$ functions depending on n arguments and $\alpha_1^{(k)} - \alpha_1^{(k+1)}$ functions of k arguments for $1 \leq k < n$. This implies that the field equations are underdetermined, if and only if $\alpha_1^{(n)} > 0$, as then some fields are completely unconstrained by the field equations and can be chosen arbitrarily. Again we expect this to happen only for systems where a gauge symmetry is the reason for the indeterminacy.

In field theories one typically considers as a degree of freedom a field for which one can prescribe initial values on an $(n-1)$-dimensional Cauchy surface Σ. If we can choose for $T\mathcal{B}$ a quasi-regular basis $\{e_1, \ldots, e_n\}$ where e_n is transversal to Σ, the surface Σ is non-characteristic and there exist exactly $\alpha_1^{(n-1)}$ such fields for a not underdetermined system. This leads naturally to the following definition: for an involutive field equation $\mathcal{R}_q$ without gauge symmetries the number N of degrees of freedom is

$$N = \alpha_q^{(n-1)} \tag{15}$$

For a system with gauge symmetries we cannot directly use (15); we must first introduce *reduced* or *gauge corrected* Cartan characters $\bar{\alpha}_{\bar{q}}^{(k)}$. They are associated with the reduced Hilbert polynomial $\bar{H}(s)$ defined by (10) and measure the size of the reduced solution space obtained by quotienting by the gauge pseudogroup. So it is natural to define for the number $\bar{N}$ of true degrees of freedom

$$\bar{N} = \bar{\alpha}_{\bar{q}}^{(n-1)} \tag{16}$$

Again it follows from the results at the end of Sect. 3 that any alternative formulation of the field theory will lead to the same value $\bar{N}$ and that it does not matter what value we take for $\bar{q}$. Especially, if it is possible to find a complete set of gauge invariant fields and to rewrite the field equations in terms of these, one will obtain $\bar{N}$ as the index of generality of the new equations and up to a shift of the argument $\bar{H}(s)$ as their Hilbert function.

This approach to counting degrees of freedom is closely related to some concepts developed in an appendix of (Einstein, 1955). Einstein discussed this problem already much earlier with Cartan in a number of letters (Cartan and Einstein, 1979) where Cartan outlined his theory of involutive systems. For some reasons Einstein apparently never used this theory but instead published much later an approach of his own. He introduced a *compatibility coefficient* $Z^{(0)}$ and for compatible systems, i. e. systems with $Z^{(0)} = 0$, the *strength* $Z^{(1)}$ which he considered as a measure for the size of the solution space. It was shown in (Seiler, 1994b) that these quantities are related to the gauge corrected Cartan characters by the relations

$$Z^{(0)} = \bar{\alpha}_{\bar{q}}^{(n)}\,, \quad Z^{(1)} = (n-1)\left(\frac{1}{2}n\bar{\alpha}_{\bar{q}}^{(n)} + \bar{\alpha}_{\bar{q}}^{(n-1)}\right) \tag{17}$$

Thus a compatible system is just a system which, after subtraction of the gauge symmetry, is not underdetermined ($\bar{\alpha}_{\bar{q}}^{(n)} = 0$) and its strength corresponds up to the numerical factor $n-1$ determined by the dimension n of space-time to our number $\bar{N}$ of degrees of freedom.

In practice, it is usually easier and less error prone to determine the reduced Cartan characters than to compute directly the strength in the formalism of Einstein. His approach requires a careful and non-trivial counting of all "identities" between the field equations and "identities of the identities" etc. (for some concrete calculations see e. g. (Siklos, 1996) and references therein). In the literature one can find several cases where mistakes were made. In the formal theory the number of all these identities are encoded in the Cartan characters which are trivial to determine for an involutive symbol.

As a simple example we treat the $U(1)$-Yang-Mills theory where Maxwell's equations represent a gauge invariant form. In (Seiler, 1994b) Hilbert function and Cartan characters for both theories are determined (our proof that the Yang-Mills equations are involutive and our computation of the Cartan characters were performed in local coordinates; for an example of a more intrinsic approach see (Giachetti and Mangiarotti, 1996)). In a four-dimensional space-time the Hilbert function of the $U(1)$-Yang-Mills equations (a second order involutive system) is given by

$$H_{\mathrm{YM}}(0) = 4\,, \quad H_{\mathrm{YM}}(1) = 16\,, \quad H_{\mathrm{YM}}(2+r) = 36 + \frac{73}{3}r + \frac{9}{2}r^2 + \frac{1}{6}r^3 \tag{18}$$

The gauge transformation $\bar{A} = A + d\Lambda$ are the solutions of the following finite Lie equation

$$\mathcal{G}_1 : \left\{ \frac{\partial \bar{A}_\mu}{\partial x^\nu} - \frac{\partial \bar{A}_\nu}{\partial x^\mu} = 0\,, \quad \frac{\partial \bar{A}_\mu}{\partial A_\nu} = \delta_\mu^\nu \right. \tag{19}$$

$\mathcal{G}_1$ is already involutive and has the Hilbert function

$$G(0) = 4\,, \quad G(1) = 10\,, \quad G(2+r) = 20 + \frac{37}{3}r + \frac{5}{2}r^2 + \frac{1}{6}r^3 \tag{20}$$

In order to correct for the gauge freedom we must subtract the two Hilbert functions (18) and (20) and obtain the reduced Hilbert function $\bar{H}_{\rm YM}(0) = 0$, $\bar{H}_{\rm YM}(1) = 6$ and $\bar{H}_{\rm YM}(2+s) = 30+16s+2s^2$. Applying (8) to it yields

$$\bar{\alpha}^{(4)}_{\rm YM} = 0\,, \quad \bar{\alpha}^{(3)}_{\rm YM} = 4\,, \quad \bar{\alpha}^{(2)}_{\rm YM} = \bar{\alpha}^{(1)}_{\rm YM} = 6 \tag{21}$$

so that we can conclude that the theory has four degrees of freedom.

Up to a shift $\bar{H}_{\rm YM}(s)$ is precisely the Hilbert function of Maxwell's equations. This is not surprising in the light of our previous results: the field strength F and the vector potential A are connected by a differential relation in the sense of Sect. 3. Taking the Yang-Mills equations as $\mathcal{R}_2$ and Maxwell's equations as $\bar{\mathcal{R}}_1$, the equation $F = dA$ defines a relation $\mathcal{T}_{1,0}$. Associated with this relation are the two Hilbert functions $\bar{G}(s) = 0$ and $G(s)$ as given by (20). Thus (9) yields $H_{\rm Max}(s-1) = H_{\rm YM}(s) - G(s) = \bar{H}_{\rm YM}(s)$. In this example this holds even for the full Hilbert functions, as $\mathcal{T}_{1,0}$ is in both directions normal, i. e. of Cauchy-Kowalevsky type.

7 Conclusion

We have shown that the formal theory of differential equations provides a natural framework for counting degrees of freedom. Using a pseudogroup approach the correction for gauge symmetries becomes a trivial subtraction of Hilbert functions. Our definition is completely independent of any specific formalism for setting up the equations of motion or the field equations, respectively, and can thus be applied to any kind of mechanical system or field theory.

For field theories one may wonder why only the index of generality is used. After all, the lower Cartan characters also measure some freedom in the general solution, as one can see in the Cartan-Kähler theorem. So for a full comparison of the size of the solution spaces of two different field theories (the application Einstein had in mind when introducing the strength) one should look at the Hilbert function. One obvious reason for the use of the index of generality lies in the results of Sect. 3 where we have shown that it remains invariant even under very general transformations.

Siklos (1996) discussed the degrees of freedom count for various formulations of Einstein's equations based on different fields (metric, connection or curvature). For each case he performed a new count of Taylor coefficients. In our approach, it would suffice to analyse the relation between, say, connection and metric as described in Sect. 3 (which is usually easier than to analyse each time the full Einstein equations); the relation between the Hilbert polynomials is then given by (9).

In this article we always assumed that the gauge pseudogroup was known. This is the typical situation in concrete applications where usually the Lagrangians are constructed so that they possess certain gauge symmetries. As already mentioned above the identification of gauge symmetries is in general to a large extent a question of physical interpretation and not of mathematics. The situation is different for Lagrangian and Hamiltonian systems where one can explicitly derive the (infinitesimal) gauge transformations (see e. g. (Henneaux et al., 1990)). It would be useful to be able to directly obtain them in form of Lie equations, as otherwise the Lie equations must first be determined via a tedious elimination of the gauge parameters.

Acknowledgment. The author would like to thank Vladimir Gerdt for some interesting discussions and for pointing out problems with the degrees of freedom count in (Seiler and Tucker, 1995). This work has been supported by Deutsche Forschungsgemeinschaft.

Literature

1. Cartan, E.; Einstein, A.: Lettres sur le Parallélisme Absolu 1929–1932, (edited by R. Debever), Académie Royale de Belgique and Princeton University Press, Bruxelles, (1979).
2. Dirac, P.A.M.: Generalized Hamiltonian Dynamics, Can. J. Math., 2,(1950),129–148.
3. Dubois-Violette, M.: The Theory of Overdetermined Linear Systems and its Applications to Non-Linear Field Equations, J. Geom. Phys., 1,(1984),139–172.
4. Einstein, A.: The Meaning of Relativity, 5th ed., Princeton University Press, Princeton, (1955).

5. Giachetta, G.; Mangiarotti, L.: Gauge Invariance and Formal Integrability of the Yang-Mills-Higgs Equations, Int. J. Theor. Phys., 35, (1996), 1405–1422.

6. Henneaux, M.; Teitelboim, C.; Zanelli, J.: Gauge Invariance and Degree of Freedom Count, Nucl. Phys., B332, (1990), 169–188.

7. Pommaret, J.-F.: Systems of Partial Differential Equations and Lie Pseudogroups, Gordon & Breach, London, (1978).

8. Saunders, D.J.: The Geometry of Jet Bundles, Cambridge University Press, Cambridge, (1989).

9. Schü, J; Seiler, W.M.; Calmet, J: Algorithmic Methods for Lie Pseudogroups, Proc. Modern Group Analysis: Advanced Analytical and Computational Methods in Mathematical Physics, N. Ibragimov, M. Torrisi, A. Valenti (eds.), Kluwer, Dordrecht, (1993), pp. 337–344.

10. Seiler, W.M.: Analysis and Application of the Formal Theory of Partial Differential Equations, Ph.D. Thesis, School of Physics and Materials, Lancaster University, (1994).

11. Seiler, W.M.: On the Arbitrariness of the General Solution of an Involutive Partial Differential Equation, J. Math. Phys., 35, (1994) ,486–498.

12. Seiler, W.M.: Involution and Constrained Dynamics II: The Faddeev-Jackiw Approach, J. Phys., A28, (1995), 7315–7331.

13. Seiler, W.M.; Arbitrariness of the General Solution and Symmetries, Acta Appl. Math., 41, (1995), 311–322.

14. Seiler, W.M.; Tucker, R.W.: Involution and Constrained Dynamics I: The Dirac Approach, J. Phys., A28, (1995), 4431–4451.

15. Siklos, S.T.C.: Counting Solutions of Einstein's Equations, Class. Quant. Grav., 13,(1996), 1931–1948.

Address: Dr. Werner M. Seiler, Lehrstuhl für Mathematik I, Universität Mannheim, D-68131 Mannheim, Germany