Nonholonomic systems via moving frames: Cartan equivalence and Chaplygin Hamiltonization*

Kurt Ehlers[1], Jair Koiller[2], Richard Montgomery[3], and Pedro M. Rios[4]

[1] Department of Mathematics
Truckee Meadows Community College
7000 Dandini Boulevard
Reno, NV 89512-3999
USA
kehlers@tmcc.edu
[2] Fundação Getulio Vargas
Praia de Botafogo 190,
Rio de Janeiro 22253-900
Brazil
jkoiller@fgv.br
[3] Mathematics Department
University of California at Santa Cruz
Santa Cruz, CA 95064
USA
rmont@math.ucsc.edu
[4] Department of Mathematics
University of California at Berkeley
Berkeley, CA 94720
USA
prios@math.berkeley.edu

Dedicated to Alan Weinstein on his 60*th birthday.*

Abstract. A *nonholonomic system*, for short "NH," consists of a configuration space Q^n, a Lagrangian $L(q, \dot{q}, t)$, a nonintegrable constraint distribution $\mathcal{H} \subset TQ$, with dynamics governed by Lagrange–d'Alembert's principle. We present here two studies, both using *adapted moving frames*. In the first we explore the *affine connection viewpoint*. For natural Lagrangians $L = T - V$, where we take $V = 0$ for simplicity, NH-trajectories are geodesics of a (nonmetric) connection ∇_{NH} which mimics Levi-Civita's. *Local geometric invariants* are obtained by Cartan's method of equivalence. As an example, we analyze Engel's (2–4) distribution. This is the first such study for a distribution that is not strongly nonholonomic. In the second part we study

* The authors thank the Brazilian funding agencies CNPq and FAPERJ: a CNPq research fellowship (JK), a CNPq post-doctoral fellowship at Berkeley (PMR), a FAPERJ visiting fellowship to Rio de Janeiro (KE). (JK) thanks the E. Schrödinger Institute, Vienna, for financial support during Alanfest and the Poisson Geometry Program, August 2003.

G-Chaplygin systems; for those, the constraints are given by a connection $\phi : TQ \to \mathrm{Lie}(G)$ on a principal bundle $G \hookrightarrow Q \to S = Q/G$ and the Lagrangian L is G-equivariant. These systems compress to an almost Hamiltonian system $(T^*S, H^\phi, \Omega_{NH})$, $\Omega_{NH} = \Omega_{\mathrm{can}} + (J.K)$, with $d(J.K) \neq 0$ in general; the momentum map $J : T^*Q \to \mathrm{Lie}(G)$ and the curvature form $K : TQ \to \mathrm{Lie}(G)^*$ are matched via the Legendre transform. Under an $s \in S$ dependent time reparametrization, a number of compressed systems become Hamiltonian, i.e., Ω_{NH} is sometimes conformally symplectic. A necessary condition is the existence of an invariant volume for the original system. Its density produces a candidate for conformal factor. Assuming an invariant volume, we describe the obstruction to Hamiltonization. An example of a Hamiltonizable system is the "rubber" Chaplygin's sphere, which extends Veselova's system in $T^*SO(3)$. This is a ball with unequal inertia coefficients rolling without slipping on the plane, with vertical rotations forbidden. Finally, we discuss reduction of internal symmetries. Chaplygin's "marble," where vertical rotations are allowed, is not Hamiltonizable at the compressed $T^*SO(3)$ level. We conjecture that it is also *not* Hamiltonizable when *reduced* to T^*S^2.

> "Nonholonomic mechanical systems (such as systems with rolling contraints) provide a very interesting class of systems where the reduction procedure has to be modified. In fact this provides a class of systems that give rise to an almost Poisson structure, i.e., a bracket which does not necessarily satisfy the Jacobi identity" (Marsden and Weinstein [2001]).

1 Introduction and outline

Cartan's moving frames method is a standard tool in Riemannian geometry.[1] In analytical mechanics, the method goes back to Poincaré [1901], perhaps earlier, to Euler's rigid body equations, perhaps much earlier, to the cave person who invented the wheel. Let $q \in \mathbb{R}^n$ be local coordinates on a configuration space Q^n, and consider a local frame, defined by an $n \times n$ invertible matrix $B(q)$,

$$X_j = \frac{\partial}{\partial \pi_j} = \sum_{i=1}^{n} b_{ij} \frac{\partial}{\partial q_i}, \qquad \sum \dot{\pi}_j X_j = \sum \dot{q}_i \frac{\partial}{\partial q_i}, \quad \dot{\pi} = A(q)\dot{q}, \quad A = B^{-1}. \tag{1.1}$$

In mechanical engineering (Hamel [1949], Papastavridis [2002]), moving frames are disguised under the keyword *quasi-coordinates*, nonexisting entities π such that

$$\frac{\partial f}{\partial \pi_j} = \sum_i \frac{\partial f}{\partial q_i} \frac{\partial q_i}{\partial \pi_j} = \sum_i \frac{\partial f}{\partial q_i} b_{ij} = X_j(f).$$

Let $\{\epsilon_i\}_{i=1,\ldots,n}$ be the dual coframe to $\{X_j\}$, $\epsilon_i =$ "$d\pi_i$" $= \sum_j a_{ij} dq_j$.

[1] Cartan [1926]; there is a recent English translation from the Russian translation (Cartan [2001]). One of the most important applications was the construction of characteristic classes by Alan's advisor, S. S. Chern. Our taste for moving frames in mechanics is a small tribute to his influence.

1.1 Moving frames: Lagrangian and Hamiltonian mechanics

The Euler–Lagrange 1-form may be rewritten as:[2]

$$\sum_{r=1}^{n}\left(\frac{d}{dt}\frac{\partial L}{\partial \dot q_r}-\frac{\partial L}{\partial q_r}-F_r\right)dq_r$$
$$=\sum_{k=1}^{n}\left(\frac{d}{dt}\frac{\partial L^*}{\partial \dot\pi_k}-\frac{\partial L^*}{\partial \pi_k}+\sum_{i=1}^{n}\frac{\partial L^*}{\partial \dot\pi_i}\sum_{j=1}^{n}\gamma^i_{kj}\dot\pi_j-R_k\right)\epsilon_k=0, \qquad (1.2)$$

where $L^*(q,\dot\pi,t) = L(q,B(q)\dot\pi,t)$ is the Lagrangian written in "quasi-coordinates" and $R_k=\sum_s F_s b_{sk}$ are the covariant components of the total force (external, F_{ext}, and constraint force λ). The so-called *Hamel transpositional symbols* $\gamma^i_{kj} = \gamma^i_{jk} = \sum_{s,\ell=1}^{n} b_{sk}b_{\ell j}(\partial a_{is}/\partial q_\ell - \partial a_{i\ell}/\partial q_s)$ are precisely the moving frame structure coefficients (Koiller [1992]).

If the velocities are restricted to a subbundle $\mathcal{H}\subset TQ$, a *constraint force* λ appears. The *d'Alembert–Lagrange principle*[3] implies that λ belongs to the anihilator $\mathcal{H}^o\subset T^*Q$ of $\mathcal{H}$, hence exerting zero work on admissible motions $\dot q\in\mathcal{H}$:

$$[L] := \frac{d}{dt}\frac{\partial L}{\partial \dot q}-\frac{\partial L}{\partial q}-F_{\text{ext}}=\lambda\in\mathcal{H}^o,\quad \dot q\in\mathcal{H}. \qquad (1.3)$$

Using moving frames, constraints can be eliminated directly. If $\mathcal{H}^o$ is spanned by the last r forms ϵ_J, $s+1\le J\le n$ $(s=n-r)$, then equations of motion result from setting the first s Euler–Lagrange differentials equal to zero:

$$\frac{d}{dt}\frac{\partial L^*}{\partial \dot\pi_k}-\frac{\partial L^*}{\partial \pi_k}+\sum_{i=1}^{n}\frac{\partial L^*}{\partial \dot\pi_i}\sum_{j=1}^{n}\gamma^i_{kj}\dot\pi_j-F^{\text{ext}}_k=0\quad (1\le k\le s). \qquad (1.4)$$

Strikingly, *the Hamiltonian counterparts of* (1.2) *and* (1.4) *are simpler, although less known*.[4] The philosophy is to fight against Darboux's dictatorship. In terms of the local coframe $\{\epsilon_i\}_{1\le i\le n}$, any element $p_q\in T^*Q$ can be written as $p_q=\sum m_i\epsilon_i(q)$. The natural 1-form α on T^*Q keeps the familiar confusing expression $\alpha := pdq = m\epsilon$. Consequently, the canonical symplectic form $\Omega := d\alpha$ may be written as

$$\Omega := dp\wedge dq = dm\wedge\epsilon + md\epsilon. \qquad (1.5)$$

The second term $md\epsilon$, which deviates from Darboux's format, is *not* a nuisance; it carries most valuable information. For instance, Kostant–Arnold–Kirillov–Souriau's bracket in T^*G, G a Lie group, can be immediately visualized: take a (left- or right-)

[2] Atributed to Hamel, but certainly known by Poincaré. Quasi-coordinates can be found in Whittaker [1937] and were first used in mechanics by Gibbs; see Pars [1965].

[3] According to Sommerfeld [1952], this gives the most natural foundation for mechanics.

[4] A "moving frames operational system" for Hamiltonian mechanics in T^*Q was given in Koiller, Rios, and Ehlers [2002].

invariant coframe and apply H. Cartan's "magic formula" on $d\epsilon$. So moving frames are ideally suited when a Lie symmetry group G is present.[5]

Example: Mechanics in $SO(3)$

To fix notation, we now review the standard example. The Lie algebra basis $X_i \in sO(3) = T_I SO(3)$, $i = 1, 2, 3$ (infinitesimal rotations around the x, y, z-axis at the identity) can either be right or left transported, producing moving frames on $SO(3)$ denoted $\{X_i^r\}$ and $\{X_i^\ell\}$, respectively. Let $\{\rho_i\}_{1\le i\le 3}$ and $\{\lambda_i\}_{1\le i\le 3}$ denote their dual coframes (right- and left-invariant forms in $SO(3)$). To represent angular momenta, we use Arnold's notations (Arnold [1989]): capital letters mean objects in the body frame, lowercase objects in the space frame. Thus for instance, $\ell = RL$, where L is the angular momentum in the body frame and ℓ is the angular momentum in space; likewise $\omega = R\Omega$ relate the angular velocities. The canonical 1-form in $T^*SO(3)$ is given by

$$\alpha = \ell_1\rho_1 + \ell_2\rho_2 + \ell_3\rho_3 = L_1\lambda_1 + L_2\lambda_2 + L_3\lambda_3,$$

so

$$\begin{aligned}\Omega_{\text{can}} &= \sum d\ell_i\rho_i + \ell_1 d\rho_1 + \ell_2 d\rho_2 + \ell_3 d\rho_3\\ &= \sum dL_i\lambda_i + L_1 d\lambda_1 + L_2 d\lambda_2 + L_3 d\lambda_3,\end{aligned}$$

where by Cartan's structure equations, $d\lambda_1 = -\lambda_2 \wedge \lambda_3, \dots$ and $d\rho_1 = \rho_2 \wedge \rho_3, \dots$ (cyclic). A left-invariant metric is given by an inertia operator $L = A\Omega$. Euler's rigid body equations follow immediately.

Poisson action of S^1 on $SO(3)$

Consider the *left* S^1 action on $SO(3)$ given by $\exp(i\phi) \cdot R := S(\phi)R$, where $S(\phi)$ is the rotation matrix about the z-axis:

$$S(\phi) := \begin{pmatrix} \cos(\phi) & -\sin(\phi) & 0\\ \sin(\phi) & \cos(\phi) & 0\\ 0 & 0 & 1\end{pmatrix}, \qquad S(-\phi)S'(\phi) = \begin{pmatrix} 0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{pmatrix} = X_3.$$

Two matrices are in the same equivalence class iff their third rows, which we denote by γ, called the *Poisson vector*, are the same: $R_1 \sim R_2 \iff R_1^{-1}\hat{k} = R_2^{-1}\hat{k} = \gamma \in S^2$. So we have a principal bundle $\pi : SO(3) \to S^2$, $\gamma = \pi(R) = R^{-1}\hat{k} = R^\dagger\hat{k}$. The derivative of π is

$$\dot\gamma = \pi_*(\dot R) = -(R^{-1}\dot R R^{-1})k = -(R^{-1}\dot R)(R^{-1})k = -[\Omega]\gamma = -\mathbf{\Omega}\times\gamma = \gamma\times\mathbf{\Omega}, \tag{1.6}$$

[5] As we learned from Alan at the banquet, the etymology for *symplectic* is "capable to join," themes and people. The latter is one of the most important aspects of the symplectic "creed." *Provocation*: taking moving frames, *adapted* to some other mathematical structure for Q, would the non-Darboux term provide a *local* symplectic invariant?

where we used the customary identification[6] $[\Omega] \in sO(3) \leftrightarrow \mathbf{\Omega} \in \mathbb{R}^3$, Arnold [1989]. The lifted action to $T^*SO(3)$ has momentum map $J = \ell_3$.

Connection on $S^1 \hookrightarrow SO(3) \to S^2$

Take the usual bi-invariant metric $\langle\langle\,,\rangle\rangle$ on $SO(3)$ so that both $\{X_i^\ell\}$ and $\{X_i^r\}$ are orthonormal moving frames. The tangent vectors to the fibers are $(d/d\phi)S(\phi)\cdot R = X_3^{\text{right}}$. Consider the *mechanical connection* associated to $\langle\langle\,,\rangle\rangle$, namely, that horizontal and vertical spaces are orthogonal. The horizontal spaces are generated by X_1^{right} and X_2^{right}. The connection form is $\phi = \rho_3$. The horizontal lift of $\dot\gamma$ to R is the tangent vector $\dot R$ such that

$$\Omega_{\text{hor}} = R^{-1}\dot R = [\dot\gamma \times \gamma]. \tag{1.7}$$

Note that Ω_{hor} *is the* -90 *degrees rotation of* $\dot\gamma$ *inside* $T_\gamma S^2$. The curvature of this connection $\kappa = d\rho_3$ is the area form of the sphere.

Reduction of S^1 symmetry

It is convenient for reduction to use (a, ℓ_3), $a \in \mathbb{R}^3$, $a \perp \gamma$,

$$L := a \times \gamma + \ell_3\gamma, \tag{1.8}$$

where a is a vector perpendicular to γ. The vector a has an intrinsic meaning: Consider a moving frame e_1, e_2 in S^2, with dual coframe θ_1, θ_2. Then $v_\gamma = v_1e_1 + v_2e_2$ parametrizes TS^2, and $p_\gamma = a\cdot d\gamma = p_1\theta_1 + p_2\theta_2$ parametrizes T^*S^2, $a = p_1e_1 + p_2e_2$. Here $a\cdot d\gamma$, $\sum\gamma_i d\gamma_i = 0$ denotes both an element of T^*S^2 and the canonical 1-form. Our parametrization for $SO(3)$ is $R(\phi,\gamma) = S(\phi)\cdot R(\gamma)$, $R(\gamma) = \text{rows}(e_1, e_2, \gamma)$. Then $L = p_2e_1 - p_1e_2 + \ell_3\gamma$ corresponds to $\ell = (p_2, -p_1, \ell_3)$ *along the section* $\phi = 0$. The right-invariant forms are compactly represented as

$$\rho_3 = d\phi - (de_1, e_2), \qquad \rho_1 + i\rho_2 = -i\exp(i\phi)(\theta_1 + i\theta_2). \tag{1.9}$$

Lifting $v \in TS^2$ to an horizontal vector in $TSO(3)$ is simple:

$$\Omega_{\text{hor}} = [(v_1e_1 + v_2e_2)\times\gamma] = [v_2e_1 - v_1e_2] \quad \text{or} \quad \text{hor}(v) = v_2X_1^r - v_1X_2^r. \tag{1.10}$$

Hence any vector $\dot R \in TSO(3)$ can be written as $\dot R = \omega_1X_1^\ell + \omega_2X_2^\ell + \omega_3X_3^\ell$ with $\omega_1 = v_2, \omega_2 = -v_1$. Any covector $p_R \in T^*SO(3)$ can be written as $p_R = p_1\pi^*(\theta_1) + p_2\pi^*(\theta_2) + \ell_3\rho_3$.

The reduced symplectic manifold $J^{-1}(\ell_3)/S^1 \equiv T^*S^2$ can be *explicitly* constructed, taking the section $\phi = 0$. Let $i : T^*S^2 \to T^*SO(3)$,

$$i(\gamma, p_1, p_2) = (R(\gamma), \ell), \quad \ell = (p_2, -p_1, \ell_3). \tag{1.11}$$

[6] We will drop the $[\bullet]$ and $\bullet$ in what follows and mix all notation, hoping no confusion will arise. Equation (1.6) is one half of every system of ODEs for S^1-equivariant mechanics in $SO(3)$. Of course, we also obtain $\dot\gamma = -\Omega\times\gamma$ by differentiating $R\gamma = k$ (we could use the notation $\gamma = K$, but we won't).

Then from (1.9) we get $i^*\rho_2 = -\theta_1$, $i^*\rho_1 = \theta_2$, and $i^*d = di^*$ yields

$$i^*d\rho_1 = d\theta_2, \qquad i^*d\rho_2 = -d\theta_1, \qquad i^*d\rho_3 = i^*\rho_1 \wedge i^*\rho_2 = -\theta_2\theta_1 = \theta_1\theta_2.$$

We get immediately

$$\Omega^{\mathrm{red}}_{T^*S^2} = i^*(\Omega_{T^*SO(3)}) = d(p_1\theta_1 + p_2\theta_2) + \ell_3 \mathrm{area} = \Omega^{\mathrm{can}}_{T^*S^2} + \ell_3 \mathrm{area}_{S_2}. \quad (1.12)$$

All references to the moving frame disappear, but the expression $\Omega^{\mathrm{can}}_{T^*S^2} = d(p_1\theta_1 + p_2\theta_2)$ suggests that whenever a natural mechanical system in $T^*SO(3)$ reduces to $T^*S^2 \equiv TS^2$, there is a preferred choice for the moving frame $\{e_1, e_2\}_\gamma$: namely, that which diagonalizes the Legendre transform $T_\gamma S^2 \to T^*_\gamma S^2 \equiv T_\gamma S^2$ of the reduced (Routh) Lagrangian.

1.2 Nonholonomic systems

A NH system $(Q, L, \mathcal{H})$ consists of a configuration space Q^n, a Lagrangian $L : TQ \times \mathbb{R} \to \mathbb{R}$, and a totally nonholonomic constraint distribution $\mathcal{H} \subset TQ$. The dynamics are governed by Lagrange–d'Alembert's principle.[7] Usually L is *natural*, $L = T - V$, where T is the kinetic energy associated to a Riemannian metric $\langle\,,\rangle$, and $V = V(q)$ is a potential. By *totally nonholonomic*, we mean that the filtration $\mathcal{H} \subset \mathcal{H}_1 \subset \mathcal{H}_2 \subset \cdots$ ends in TQ. Each subbundle $\mathcal{H}_{i+1}$ is obtained from the previous one by adding to $\mathcal{H}_i$ combinations of all possible Lie brackets of vector fields in $\mathcal{H}_i$. To avoid interesting complications we assume that all have constant rank. Equivalently, let $\mathcal{H}^o \subset T^*Q$ the codistribution of "admissible constraints" annihilating $\mathcal{H}$; dually, one has a decreasing filtration of *derived ideals* ending in zero.

Internal symmetries of NH systems: Noether's theorem

An *internal symmetry* occurs whenever a vector field $\xi_Q \in \mathcal{H}$ preserves the Lagrangian. For natural systems ξ_Q is a Killing vector field for the metric. Noether's theorem from unconstrained mechanics remains true. The argument (see Arnold, Kozlov, and Neishtadt [1988]) goes as follows: denote by $\phi_\xi(s)$ the one-parameter group generated by ξ and let $\phi(s,t) = \phi_\xi(s) \cdot q(t)$, so $\phi' = \frac{d}{ds}\phi = \xi_Q(\phi)$, where $q(t)$ is chosen as a trajectory of the nonholonomic system. Differentiating with respect to s the identitly $L(\phi(s,t), \frac{d}{dt}\phi(s,t)) = \text{const}$ after a standard integration by parts we get $\frac{d}{dt}(\frac{\partial L}{\partial \dot q}\phi') = [L]\phi'$. This vanishes precisely when $\phi' = \xi_Q \in \mathcal{H}$, so $I_\xi := \frac{\partial L}{\partial \dot q} \cdot \xi = \text{const}$.

[7] "Vakonomic" mechanics uses the same ingredients, but the dynamics are governed by the variational principle with constraints, and produce different equations; see, e.g., Cortés, de Léon, de Diego, and Martínez [2003]. The equations coincide if and only if the distribution is integrable. In spite of many similarities, there are striking differences between NH and holonomic systems. For instance, NH systems do not have (in general) a smooth invariant measure. Necessary and sufficient conditions for the existence of the invariant measure were first given (explicitly in coordinates) by Blackall [1941].

External symmetries: G-Chaplygin systems

External (*or transversal*) *symmetries* occur when group G acts on Q, preserving the Lagrangian and the distribution $\mathcal{H}$, meaning that $g_*\mathcal{H}_q = \mathcal{H}_{gq}$. In the most favorable case one has a principal bundle action $G^r \hookrightarrow Q^n \to S^m$, $m + r = n$, where $\mathcal{H}$ forms the horizontal spaces of a connection with 1-form $\phi : TQ \to \mathrm{Lie}(G)$. These systems are called *G-Chaplygin*.[8]

Terminology

Since Bates and Śniatycki [1993], and Bloch, Krishnaprasad, Marsden, and Murray [1996], several authors have called attention to these two types of symmetries. Reduction of internal symmetries was already described in Śniatycki [1998]. To stress the difference, reduction of external symmetries is called *compression* here. The word *reduction* will be used for internal symmetries.

LR systems

Veselov and Veselova [1986], Veselov and Veselova [1988] considered Lie groups $Q = G$ with *left*-invariant metrics, with constraint distributions given by *right* translation of $\mathcal{D} \subset \mathrm{Lie}(G)$, i.e., the constraints are given by right-invariant forms. For a *LR-Chaplygin system*, in addition there is a decomposition $\mathrm{Lie}(G) = \mathrm{Lie}(H) \oplus \mathcal{D}$, where H is a Lie subgroup such that $\mathrm{Ad}_{h^{-1}} D = h^{-1} D h = D$. Therefore, $H \hookrightarrow G \to S = G/H$ is a H-Chaplygin system; the base S is the homogeneous space of cosets Hg. Fedorov and Jovanovic [2003] considered the case where G is compact and $\mathrm{Lie}(H)$ is *orthogonal* to D with respect to the bi-invariant metric.[9]

Compression of G-Chaplygin systems

From symmetry, it is clear that the Lagrange–d'Alembert equations *compress* to the base TS.[10] In covariant form, the dynamics take the form $[L^\phi] = F(s, \dot{s})$, where $L^\phi(s, \dot{s}) = L(s, h(\dot{s}))$ is the compressed Lagrangian in TS; $h(\dot{s})$ is the horizontal

8 A "historical" remark (by JK). Chaplygin considered the abelian case. During a post-doctoral year in Berkeley, way back in 1982, I became interested in NH systems with symmetries. Alan directed me to two wonderful books: Hertz [1899] *Foundation of Mechanics* and Neimark and Fufaev [1972]. In the latter I learned about (abelian) Chaplygin systems, presented in coordinates. I said to Alan that I would like to examine nonabelian group symmetries, and Alan immediately made a diagram on his blackboard, and told me: "well, then, the constraints are given by a connection on a principal bundle." This was the starting point of Koiller [1992].

9 These conditions are not met in the marble and rubber Chaplygin spheres (see Section 3.2); however, Veselov's result (Theorem 3.3 below) on invariant volume forms still holds.

10 The full dynamics can be reconstructed from the compressed solutions, horizontal lifting the trajectories via ϕ, since the admissible paths are horizontal relative to the connection. This last step is not "just" a quadrature; in the nonabelian case, a *path*-ordered integral is in order. For $G = SO(3)$, Levi [1996] found an interesting geometric construction.

lift to any local section and F is a *pseudogyroscopic force*.[11] In order to write F explicitly, take group quasi-coordinates $(s, \dot{s}, g, \dot{\pi})$. Write $q = g\sigma(s)$, with $g \in G$ and a local section $\sigma(s)$ of $Q \to S$. Fix a basis X_k for the Lie algebra, $[X_K, X_L] = \sum c^J_{KL} X_J$, $X(\dot{\pi}) = \sum \dot{\pi}_I X_I$. Any tangent vector $\dot{q} \in T_{\sigma(s)}Q$ can be written as $\dot{q} = d\sigma(s)\cdot\dot{s} + X(\dot{(\pi)})\cdot\sigma(s)$. Horizontal vectors are represented by $\dot{\pi} = b(s)\cdot\dot{s}$, where $b(s)$ is an $r \times m$ matrix. The connection 1-form may be written as $\phi(\dot{q}) = \dot{\pi} - b(s)\cdot\dot{s}$. Then

$$[L^\phi] = F(s, \dot{s}), \quad F = \sum_{K=1}^{r} \left(\frac{\partial L}{\partial \dot{\pi}_k}\right)^* \sum_{j=1}^{m} \left(\frac{b_{Ki}}{\partial q_j} - \frac{b_{Kj}}{\partial q_i} + \sum_{U,V=1}^{r} b_{Ui} b_{Vj} c^K_{UV}\right) \dot{s}_j. \tag{1.13}$$

1.3 Main results

Using the moving frames method we present results on two aspects of nonholonomic systems.

- Cartan's equivalence, using Cartan's geometric description of NH systems via affine connections (Cartan [1928]). The objective is to find all local invariants.
- Chaplygin systems: compression of external symmetries, reduction of internal symmetries. The objective is to generalize Chaplygin's "reducing factor" method (Chaplygin [1911]), namely, verify if Hamiltonization is possible (via *conformally symplectic structures*).

Results on Cartan's equivalence

In Section 2 we analyze NH systems under the *affine connection* perspective. We pursue the (local) classification program proposed by Cartan [1928] using his equivalence method. See Koiller, Rodrigues, and Pitanga [2001] and Tavares [2002] for a rewrite of Cartan's paper in modern language. Cartan's method of equivalence is a powerful method for uncovering and interpreting all differential invariants and symmetries in a given geometric structure. In Ehlers [2002] NH systems in a 3-manifold with a contact distribution were classified. Here we go one step further, looking at Engel's distribution in 4-manifolds (see definition below). Our results are summarized in Theorem 2.3. The "role model" here is the rolling penny example (no pun intended). *This is the first such study for a distribution that is not strongly nonholonomic*. Next in line is studying the famous Cartan 2–3–5 distribution.

Results on G-Chaplygin systems

Instead of using (1.13) in TS, we may describe the compressed system in T^*S as an

[11] This nonholonomic force represents, philosophically, a *concealed force* in the sense of Hertz [1899], having a geometric origin. This force vanishes in some special cases, not necessarily requiring the constraints to be holonomic. Equivalently, the dynamics in TS is the geodesic spray of a *modified* affine connection. One adds to the induced Levi-Civita connection in TS a certain tensor $B(X, Y)$. This NH connection in general is nonmetric (Koiller [1992]).

almost Hamiltonian system[12]

$$i_X \Omega_{NH} = dH, \quad H = H^{\phi} : T^*S \to \mathbb{R}, \quad \Omega_{NH} = \Omega_{\text{can}}^{T^*S} + (\text{J.K}), \tag{1.14}$$

where H^{ϕ} is the Legendre transform of the compressed Lagrangian. (J.K) is a semibasic 2-form on T^*S which in general is not closed. As one may guess, J is the momentum map, and K is the curvature of the connection. Ambiguities cancel, since J is Ad*-equivariant while K is Ad-equivariant. The construction is independent of the point q on the fiber over s.

Under an $s \in S$ dependent time reparametrization, $d\tau = f(s)dt$, several interesting compressed G-Chaplygin systems become Hamiltonian. A necessary condition is the existence of an invariant volume (Theorem 3.3) whose density F produces a candidate $f = F^{1/(m-1)}$, $m = \dim(S)$ for a conformal factor. Chaplygin's "rubber" ball (vertical rotations forbidden) is, as far we know, a new example, and generalizes the well-known Veselova system in $SO(3)$ (Proposition 3.6). We describe the obstruction to Hamiltonization as the 2-form $i_X d(f\Omega_{NH})$ (Theorem 3.4) and we discuss further reduction by internal symmetries. An example of the latter situation is Chaplygin's "marble" (a hard ball with unequal inertia coefficients rolling without slipping on the plane). It is non-Hamiltonizable in $T^*SO(3)$, and our calculations suggest that it is also non-Hamiltonizable when *reduced* to T^*S^2 (Theorem 3.8). Compare with Borisov and Mamaev [2001].

What does Hamiltonization accomplish?

Why do we focus so much on the question of Hamiltonizability? The example of the reduced equations for Chaplygin's skate (after a two-dimensional Euclidean symmetry is removed) shows that changing time scale in a nonholonomic systems can completely change its character. In this example (see, e.g., Koiller [1992]) the fully reduced equations of motion are not Hamiltonian because every solution is asymptotic in forward and backward time to a point which depends on which solution you choose. However, after rescaling time the fully reduced equations become Hamiltonian, namely, the harmonic oscillator. However, this Hamiltonian vector field is incomplete because along one of the coordinate axes, the time rescaling is not defined.[13] In light of this example, why is time rescaling interesting? The answer is that it is interesting mostly in the context of integrability, where no singularities are removed in the phase space. See Section 3.

2 Nonholonomic geometry: Cartan equivalence

A *Cartan nonholonomic structure* is a triple $(Q, \mathcal{G} = \langle\cdot,\cdot\rangle, \mathcal{H})$, where Q is an n-dimensional manifold endowed with a Riemannian metric $\mathcal{G}$ and a rank r totally

[12] For details, see Koiller, Rios, and Ehlers [2002], Koiller and Rios [2001]. The Hamiltonian compression for Chaplygin systems was first explored, in the abelian case, by Stanchenko [1985]. The nonclosed term was described as a semibasic 2-form, depending linearly on the fiber coordinate in T^*S, but its geometric content was not indicated there.

[13] We thank one of the referees for this observation.

nonholonomic distribution $\mathcal{H}$. Our motivation for studying such a structure is a free particle moving in Q, nonholonomically constrained to $\mathcal{H}$, with kinetic energy $T = \frac{1}{2}\langle\cdot,\cdot\rangle$. The nonholonomic geodesic equations are obtained by computing accelerations using the Levi-Civita connection associated with $\mathcal{G}$ and orthogonally projecting the result onto $\mathcal{H}$. The projected connection is called a *nonholonomic connection* (Lewis [1998]), and was introduced by Cartan [1928]. A distribution $\mathcal{H}$ is *strongly nonholonomic* if any basis of vector fields spanning $\mathcal{H}$ on $U \subset Q$, together with their Lie brackets, span the entire tangent space over U. The equivalence problem for nonholonomic geometry was revisited in Koiller, Rodrigues, and Pitanga [2001] and the generalization to arbitrary nonholonomic distributions was discussed. *Engel manifolds provide the simplest example involving distributions that are not strongly nonholonomic*.[14]

The main question we address is the following. Given two nonholonomic structures $(Q, \mathcal{G}, \mathcal{H})$ and $(\bar{Q}, \bar{\mathcal{G}}, \bar{\mathcal{H}})$, is there a (local) diffeomorphism $f : U \subset Q \to \bar{U} \subset \bar{Q}$ carrying nonholonomic geodesics in Q to nonholonomic geodesics in $\bar{Q}$? In Cartan's approach, this question is recast as an equivalence problem. The nonholonomic structure is encoded into a subbundle of the frame bundle over Q called a G-structure. *The diffeomorphism f exists if the two corresponding G-structures are locally equivalent*. Necessary and sufficient conditions for the G-structures to be equivalent are given in terms of differential invariants found using the *method of equivalence*.

Outline

Our main example is the equivalence problem for nonholonomic geometry on an Engel manifold. Let Q be a four-dimensional manifold and $\mathcal{H}$ a rank two distribution. $\mathcal{H}$ is an Engel distribution if and only if, for any vector fields X and Y locally spanning $\mathcal{H}$, and some functions $a, b : Q \to \mathbb{R}$, the vector fields $X, Y, Z = [X, Y]$, and $W = a[X, Z] + b[Y, Z]$ form a local basis for TQ. By an Engel manifold, we mean a four-dimensional manifold endowed with an Engel distribution. We begin by describing the nonholonomic geodesic equations. In the spirit of Cartan's program, we express them in terms of connection 1-forms and (co)frames adapted to the distribution. This formulation is particularly well suited to the problem at hand; the nonholonomic geodesic equations are obtained by writing the ordinary geodesic equations in terms of the Levi-Civita connection 1-form and crossing out terms corresponding to directions complementary to $\mathcal{H}$. We then set up the equivalence problem for nonholonomic

[14] Historical remarks. Cartan [1928] introduced the equivalence problem for nonholonomic geometry and studied the case of manifolds endowed with *strongly nonholonomic distributions*. In his address, Cartan warned against attempts to study other cases because of the "plus compliqués" computations involved. In the meantime strides have been made in the equivalence method by Robert Gardner and his students that allow computations to be made at the Lie algebra level rather than at the group level (Gardner [1989]). This together with symbolic computation packages such as Mathematica™ make equivalence problems tractable in many important cases. See Gardner [1989], Bryant [1994], Montgomery [2002], Grossman [2000], Ehlers [2002], Hughen [1995], and Moseley [2001] for some applications.

geometry and give a brief description of the equivalence method as it is applied to our main example. We conclude this section by applying the method of equivalence to the case of nonholonomic geometry on an Engel manifold. We derive all differential invariants associated with the nonholonomic structure and show that the symmetry group of such a structure has dimension at most four.

2.1 Nonholonomic geodesics: Straightest paths

Totally nonholonomic distributions

A distribution $\mathcal{H}$ is a rank r vector subbundle of the tangent bundle $T(Q)$ over Q. Let $\mathcal{H}^1 = \mathcal{H} + [\mathcal{H}, \mathcal{H}]$ and $\mathcal{H}^i = [\mathcal{H}, \mathcal{H}^i]$, and consider the filtration

$$\mathcal{H} \subset \mathcal{H}^1 \subset \cdots \mathcal{H}^i \subset \cdots \subset TQ.$$

$\mathcal{H}$ is totally nonholonomic if and only if, for some k, $\mathcal{H}^k = TQ$ at all points in Q. For the present discussion we will assume that each $\mathcal{H}^i$ has constant rank over Q. As a specific example, consider the Engel distribution $\mathcal{H}$ on $\mathbb{R}^4$ with coordinates (x, y, z, w), spanned by $\{X_1 = \frac{\partial}{\partial w}, X_2 = \frac{\partial}{\partial x} + w\frac{\partial}{\partial y} + y\frac{\partial}{\partial z}\}$. There are, in fact, local coordinates on any Engel manifold so that the distribution is given by this normal form, see Montgomery [2002]. Then $\{X_1, X_2, X_3 = [X_1, X_2]\}$ spans the three-dimensional distribution $\mathcal{H}^1$, and $\{X_1, X_2, X_3, X_4 = [X_2, X_3]\}$ spans the entire $T\mathbb{R}^4$.

A path $c : \mathbb{R} \to Q$ is *horizontal* if $\dot{c}(t) \in \mathcal{H}_{c(t)}$ for all t. Chow's theorem implies that if $\mathcal{H}$ is totally nonholonomic, then any two points in Q can be joined by a horizontal path (see Montgomery [2002]). At the other extreme, the classical theorem of Frobenius implies that $\mathcal{H}$ is *integrable*, which is to say that Q is foliated by submanifolds whose tangent spaces coincide with $\mathcal{H}$ at each point, if and only if $[X_i, X_j] \in \mathcal{H}$ for all i and j (Warner [1971]).

In what follows we will need a description of distributions in terms of differential ideals. Details can be found in Warner [1971] or Montgomery [2002]. Let $\mathcal{I} = \mathcal{H}^{\perp}$ be the ideal in $\Lambda^*(Q)$ consisting of the differential forms annihilating $\mathcal{H}$. If $\mathcal{H}$ is rank r, then $\mathcal{I}$ is generated by $n - r$ independent 1-forms. The *first derived ideal* of $\mathcal{I}$ is the ideal

$$(\mathcal{I})' := \{\theta \in \mathcal{I} | d\theta \equiv 0 \,\mathrm{mod}(\mathcal{I})\}. \tag{2.1}$$

If we set $\mathcal{I}^{(0)} = \mathcal{I}$ and $\mathcal{I}^{(n+1)} = (\mathcal{I}^{(n)})'$ we obtain a decreasing filtration

$$\mathcal{I} = \mathcal{I}^{(0)} \supset \mathcal{I}^{(1)} \supset \cdots \supset 0.$$

The filtration terminating with the 0 ideal is equivalent to the assumption that the distribution is completely nonholonomic. We note that $I^{(j)} = (\mathcal{H}^j)^{\perp}$ for $j = 1$, but this is not true in general for $j > 1$ (see Montgomery [2002]). At the other extreme, the differential ideal version of the Frobenius theorem implies that $\mathcal{H}$ is integrable if and only if $(\mathcal{I})' \subset \mathcal{I}$ (Warner [1971]).

For the Engel example, the 1-forms $\eta^1 = dy - w dx$ and $\eta^2 = dz - y dx$ generate the ideal $\mathcal{I}$. Notice that $d\eta^2 = \eta^1 \wedge dx$ so $\eta^2 \in \mathcal{I}^{(1)}$ but $d\eta^1$ cannot be written in terms of η^1 or η^2; therefore, $\eta^1 \notin \mathcal{I}^{(1)}$.

The nonholonomic geodesic equations

There are two different geometries commonly defined on a nonholonomic structure $(Q, \mathcal{G} = \langle\cdot,\cdot\rangle, \mathcal{H})$: *sub-Riemannian geometry and nonholonomic geometry*. In sub-Riemannian geometry one is interested in *shortest paths.* The length of a path $c : [a, b] \to Q$ joining points x and y is $\ell(c) = \int \sqrt{\langle\dot{c}, \dot{c}\rangle}dt$. The distance from x to y is $d(x, y) = \inf(\ell(c))$ taken over all horizontal paths joining x to y. In nonholonomic geometry one is interested in *straightest paths*, which are solutions to the nonholonomic geodesic equations. Hertz [1899] was the first to notice that shortest $\neq$ straightest unless the constraints are holonomic.[15]

The nonholonomic geodesic equations are obtained by computing the acceleration of a horizontal path $c : \mathbb{R} \to Q$ using the Levi-Civita connection associated with $\mathcal{G}$ and orthogonally projecting the result onto $\mathcal{H}$. It is convenient to adopt the following indicial conventions:

$$\begin{aligned} 1 \leq I, J, K &\leq n, \\ 1 \leq i, j, k &\leq r \quad (= \operatorname{rank}(\mathcal{H})), \\ r + 1 \leq \nu &\leq n. \end{aligned} \tag{2.2}$$

Let $e = \{e_I\}$ be a local orthonormal frame for which the e_i span $\mathcal{H}$, and let $\eta = \{\eta_I\}$ be the dual coframe defined by $\eta_I(e_J) = \delta_{IJ}$, the Kronecker delta function. We note that the η^ν annihilate $\mathcal{H}$ and the metric, restricted to $\mathcal{H}$ is $g_{|\mathcal{H}} = \eta^1 \otimes \eta^1 + \cdots + \eta^r \otimes \eta^r$. The Levi-Civita connection can be expressed in terms of local 1-forms $\omega_{IJ} = -\omega_{JI}$ satisfying Cartan's structure equation $d\eta = -\omega \wedge \eta$ (Hicks [1965]).

A horizontal path $c : \mathbb{R} \to M$ is a *nonholonomic geodesic* if it satisfies the nonholonomic geodesic equations

$$\left[\frac{d}{dt}(v_i) + \sum_j v_j \omega_{ij}(\dot{c})\right] e_i = 0, \tag{2.3}$$

where $1 \leq i, j, \leq r$ and $v_i = \eta^i(\dot{c})$ are the quasi-velocities.

Example: The vertical rolling penny

A standard example of a mechanical system modeled by a nonholonomic Engel system is that of a coin rolling without sliping on the Euclidean plane. Consider a coin of radius a rolling vertically on the xy-plane. The location of the coin is represented by the coordinates (x, y, θ, ϕ). The point of contact of the coin with the plane is (x, y), the angle made by the coin with respect to the positive x-axis is θ, and the angle made by the point of contact, the center of the coin, and a point marked on the outer edge of the coin is ϕ. The state space can be identified with the Lie group $SE(2) \times SO(2)$ where the first factor is the group of Euclidean motions locally parametrized by x,

[15] The terminology *straightest path* for a nonholonomic geodesic was, in fact, coined by Hertz himself.

y and θ. The mass of the coin is m, the moment of inertia in the θ direction is J and the moment of inertia in the ϕ direction is I. The kinetic energy, which defines a Riemannian metric on the state space, is

$$T = \frac{m}{2}(dx \otimes dx + dy \otimes dy) + \frac{J}{2} d\theta \otimes d\theta + \frac{I}{2} d\phi \otimes d\phi. \tag{2.4}$$

The penny rolls without slipping giving rise to the constraints

$$\dot{x} = (a \cos\theta)\dot{\phi}, \qquad \dot{y} = (a \sin\theta)\dot{\phi}. \tag{2.5}$$

Consider the orthonormal frame (X_1, X_2, X_3, X_4), where

$$\begin{aligned} X_1 &:= \sqrt{\frac{2}{ma^2+I}}\left(a\cos\theta\frac{\partial}{\partial x} + a\sin\theta\frac{\partial}{\partial y} + \frac{\partial}{d\phi}\right), \\ X_2 &:= \sqrt{\frac{2}{J}}\frac{\partial}{\partial\theta}, \\ X_3 &:= \sqrt{\frac{2}{m}}\left(-\sin\theta\frac{\partial}{\partial x} + \cos\theta\frac{\partial}{\partial y}\right), \\ X_4 &:= \sqrt{\frac{2}{m}}\left(\cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}\right). \end{aligned} \tag{2.6}$$

Note that the constraint subspace is $\mathcal{H} = \operatorname{span}\{X_1, X_2\}$, and $\mathcal{H}^{(1)} = \operatorname{span}\{X_1, X_2, X_3\}$. The dual coframe is $(\eta^1, \eta^2, \eta^3, \eta^4)$, where

$$\begin{aligned} \eta^1 &:= \sqrt{\frac{ma^2+I}{2}}d\phi, & \eta^2 &:= \sqrt{\frac{J}{2}}d\theta, \\ \eta^3 &:= \sqrt{\frac{m}{2}}(-\sin\theta dx + \cos\theta dy), & \eta^4 &:= \sqrt{\frac{m}{2}}(\cos\theta dx + \sin\theta dy - d\phi). \end{aligned} \tag{2.7}$$

To compute the Levi-Civita connection form, we determine $\omega = [\omega_{IJ}]$ such that $\omega_{IJ} = -\omega_{JI}$ and $d\eta = -\omega \wedge \eta$. Using simple linear algebra, we find

$$\omega = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}}\sqrt{\frac{m}{J(ma^2+I)}}\eta^3 & \frac{1}{\sqrt{2}}\sqrt{\frac{m}{J(ma^2+I)}}\eta^2 & 0 \\ -\frac{1}{\sqrt{2}}\sqrt{\frac{m}{J(ma^2+I)}}\eta^3 & 0 & -\frac{1}{\sqrt{2}}\sqrt{\frac{m}{J(ma^2+I)}}\eta^1 & 0 \\ -\frac{1}{\sqrt{2}}\sqrt{\frac{m}{J(ma^2+I)}}\eta^2 & \frac{1}{\sqrt{2}}\sqrt{\frac{m}{J(ma^2+I)}}\eta^1 & 0 & -\frac{\sqrt{2}}{\sqrt{J}}\eta^2 \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{J}}\eta^2 & 0 \end{pmatrix} \tag{2.8}$$

so, in particular,

$$\omega_{12} = -\omega_{21} = \frac{1}{2}\sqrt{\frac{m}{J(ma^2+I)}}\eta^3.$$

Let $c : \mathbb{R} \to Q$ be a nonholonomic geodesic given by $\dot{c}(t) = v_1(t)X_1 + v_2(t)X_2$. From the structure equations we see immediately that $\omega_{12}(\dot{c}(t)) = -\omega_{21}(\dot{c}(t)) = 0$ and the nonholonomic geodesic equations reduce to $\frac{d}{dt}(v_1) = \frac{d}{dt}(v_2) = 0$. The nonholonomic geodesics are solutions to $(\dot{x}, \dot{y}, \dot{\phi}, \dot{\theta}) = AX_1 + BX_2$. In particular,

$$\begin{aligned} \dot{x} &= \frac{\sqrt{2}Aa\cos\theta(t)}{\sqrt{ma^2+I}}, & \dot{y} &= \frac{\sqrt{2}Aa\sin\theta(t)}{\sqrt{ma^2+I}}, \\ \dot{\phi} &= \frac{\sqrt{2}A}{\sqrt{ma^2+I}}, & \dot{\theta} &= \frac{\sqrt{2}B}{\sqrt{J}}. \end{aligned} \tag{2.9}$$

The trajectories are spinning in place ($A = 0$), rolling along a line ($B = 0$), or circles ($A, B \neq 0$).

2.2 Equivalence problem of nonholonomic geometry

Cartan's method of equivalence starts by encoding a geometric structure in terms of a subbundle of the coframe bundle called a G-structure. We begin this section by describing the G-structure for nonholonomic geometry[16]. We then give a brief outline of some of the main ideas behind the method of equivalence as it is applied in our example of nonholonomic geometry on an Engel manifold. Details about the method of equivalence can be found in Gardner [1989], Montgomery [2002], or Bryant [1994]. We then derive the local invariants associated with a nonholonomic structure on a four-dimensional manifold endowed with an Engel distribution.

Initial G-structure for nonholonomic geometry

A *coframe* $\eta(x)$ at $x \in Q^n$ is a basis for the cotangent space $T^*_x(Q)$. Alternatively, we can regard a coframe as a linear isomorphism $\eta(x) : T_x(Q) \to \mathbb{R}^n$ where $\mathbb{R}^n$ is represented by column vectors. A coframe can then be multiplied by a matrix on the left in the usual way. The set of all coframes at x is denoted $F^*_x(Q)$ and has the projection mapping $\pi : F^*_x(Q) \mapsto x$. The coframe bundle $F^*(Q)$ is the union of the $F^*_x(Q)$ as x varies over Q. A coframe on Q is a smooth (local) section $\eta : Q \to F^*(Q)$ and is represented by a column vector of 1-forms $(\eta^1, \ldots, \eta^n)^{\mathrm{tr}}$, where "tr" indicates transpose. $F^*(Q)$ is a right $Gl(n)$-bundle with action $R_g\eta = g^{-1}\eta$ where g is a matrix in $Gl(n)$.

Let G be a matrix subgroup of $Gl(n)$. A *G-structure* is a G-subbundle of $F^*(Q)$. We now describe the G-structure encoding the nonholonomic geometry associated with a nonholonomic structure $(Q, \mathcal{G}, \mathcal{H})$. Given a nonholonomic structure $(Q, \mathcal{G}, \mathcal{H})$ we can choose an orthonormal coframe $\eta = (\eta^i, \eta^\nu)^{\mathrm{tr}}$ on $U \subset Q$ so that the η^ν annihilate $\mathcal{H}$ and use this coframe to write down the nonholonomic geodesic equations as described above. On the other hand, given a coframe $\bar{\eta} = (\bar{\eta}^i, \bar{\eta}^\nu)^{\mathrm{tr}}$ on Q we can construct a nonholonomic structure $(Q, \bar{\mathcal{G}} = \sum \bar{\eta}^i \otimes \bar{\eta}^i + \bar{\eta}^\nu \otimes \bar{\eta}^\nu, \bar{\mathcal{H}})$ where $\bar{\mathcal{H}}$ is

16 This G-structure was first presented by Cartan in his 1928 address to the International Congress of Mathematicians (Cartan [1928])

annihilated by the $\bar\eta^\nu$. How is $\bar\eta$ related to η if it is to lead to the same nonholonomic geodesic equations as η? In order to preserve $\mathcal{H}$ we must have $\eta^\nu - \bar\eta^\nu = 0 \pmod{I}$. In matrix notation, any modified coframe $\bar\eta$ must be related to η by

$$\begin{pmatrix} \bar\eta^i \\ \bar\eta^\nu \end{pmatrix} = \begin{pmatrix} A & b \\ 0 & a \end{pmatrix} \begin{pmatrix} \eta^i \\ \eta^\nu \end{pmatrix}, \tag{2.10}$$

where $A \in Gl(r)$, $a \in Gl(n-r)$, and $b \in M(k, n-r)$. If we were studying the geometry of distributions, there would be no further restrictions. In order to preserve the metric restricted to $\mathcal{H}$, we must further insist that $A \in O(r)$. We would then have the starting point for the study of sub-Riemannian geometry (see Montgomery [2002], Hughen [1995], or Moseley [2001]).

It is important to observe that in nonholonomic geometry we need the full metric and not just its restriction to $\mathcal{H}$ (as in sub-Riemannian geometry) to obtain the equations of motion. Cartan [1928] showed that *in order to preserve the nonholonomic geodesic equations, we can only add covectors that are in the first derived ideal to the η^i.*

Since this fact is central to our analysis, we sketch the argument here (see Koiller, Rodrigues, and Pitanga [2001] for details). Suppose $\bar\eta = g\eta$ with connection 1-form defined by $d\bar\eta = -\bar\omega \wedge \bar\eta$. For simplicity, assume that $A = id$; then $\eta^j \equiv \bar\eta^j \pmod{\mathcal{I}}$. The geodesic equations are preserved if and only if $\omega_{ij}(T) = \bar\omega_{ij}(T)$ for all $T \in \mathcal{H}$, in other words $\omega_{ij} \equiv \bar\omega_{ij} \pmod{\mathcal{I}}$. Note also that $\bar\eta^\nu \equiv 0 \pmod{\mathcal{I}}$. Subtracting the structure equations for $d\eta^i$ and $d\bar\eta^i$, we get

$$d\eta^i - d\bar\eta^i = -\omega_{ij} \wedge \eta^j - \omega_{i\nu} \wedge \eta^\nu + \bar\omega_{ij} \wedge \bar\eta^j + \bar\omega_{i\nu} \wedge \bar\eta^\nu \equiv 0 \pmod{\mathcal{I}}.$$

Now $\bar\eta^i = \eta^i + b_{i\nu}\eta^\nu$ so we also have

$$d\eta^i - d\bar\eta^i = d\eta^i - (d\eta^i + db_{i\nu}\eta^\nu + b_{i\nu} d\eta^\nu) \equiv -b_{i\nu} d\eta^\nu \pmod{\mathcal{I}}.$$

Therefore, $b_{i\nu} d\eta^\nu \equiv 0 \pmod{\mathcal{I}}$ or, equivalently, $b_{i\nu}\eta^\nu \in I^{(1)}$. This completes the argument.

We further subdivide our indicial notation: let

$$r + 1 \le \phi \le s \quad (= \operatorname{rank} \mathcal{H}^1), \qquad s + 1 \le \Phi \le n.$$

Adapted coframes

A covector $\eta = (\eta^i, \eta^\phi, \eta^\Phi)^{\mathrm{tr}})$ arranged so that

1. The η^ϕ and η^Φ generate I,
2. $ds^2|_{\mathcal{H}} = \sum \eta^i \otimes \eta^i$,
3. The η^Φ generate the first derived ideal $I^{(1)}$,

is said to be *adapted to the nonholonomic structure*. In matrix notation, the most general change of coframes that preserves the nonholonomic geodesic equations is of the form $\bar\eta = g\eta$ where

$$g = \begin{pmatrix} A & 0 & b \\ 0 & a_1 & a_2 \\ 0 & 0 & a_3 \end{pmatrix} \tag{2.11}$$

with $A \in O(k)$, $b \in M(n-s, k)$, $a_1 \in Gl(s-k)$, $a_2 \in M(n-s, s-k)$, and $a_3 \in Gl(n-s)$. The set of all such block matrices form a matrix subgroup of $Gl(n)$ which we shall denote G_0.

The *initial G-structure* for nonholonomic geometry on $(Q, ds^2, \mathcal{H})$ is a subbundle $B_0(Q) \subset F^*(Q)$ (or simply B_0 if there is no risk of confusion) with structure group G_0 defined above. All local sections of $B_0(Q)$ lead to the same nonholonomic geodesic equations. In this way, the initial G-structure $B_0(Q)$ completely characterizes the nonholonomic geometry.

Two G-structures, $B(Q) \xrightarrow{\pi_Q} Q$ and $B(N) \xrightarrow{\pi_N} N$, are said to be *equivalent* if there is a diffeomorphism $f : Q \to N$ for which $f_1(B(Q)) = B(N)$ where f_1 is the induced bundle map. (If we think of $b \in B(Q)$ as a linear isomorphism $b : T_{\pi_Q(b)}Q \to \mathbb{R}^n$ then $f_1(b) = b \circ (f_*)^{-1}$ where f_* is the differential of f.) Our original question as to whether there is a local diffeomorphism that carries nonholonomic geodesics to nonholonomic geodesics can be answered by determining whether the associated G-structures are locally equivalent.

2.3 A tutorial on the method of equivalence

Necessary and sufficient conditions for the equivalence between G-structures are given in terms of *differential invariants* which are derived using the method of equivalence. In this section, we briefly describe some of the main ideas behind the method of equivalence as it is applied in our example. Details and other facets of the method together with many examples can be found in the excellent text by Robert Gardner (Gardner [1989]). One of the principal objects used in the method of equivalence is the *tautological* 1*-form*. Let $B(Q) \xrightarrow{\pi} Q$ be a G-structure with structure group G whose Lie algebra is Lie(G). The tautological 1-form Ω on $B(Q)$ is an $\mathbb{R}^n$-valued 1-form defined as follows. Let $\eta : U \subset Q \to B(Q)$ be a local section of $B(Q)$ and consider the inverse trivialization $U \times G_0 \to B(Q)$ defined by $(x, g) \to g^{-1}\eta(x)$. Relative to this section, the tautological 1-form is defined by

$$\Omega(b) = g^{-1}(\pi^*\eta), \tag{2.12}$$

where $b = g^{-1}\eta$. From (2.12) one can verify that the tautological 1-form is semibasic (i.e., $\Omega(v) = 0$ for all $v \in \ker(\pi_*)$), has the reproducing property $\bar{\eta}^*\Omega = \bar{\eta}$, where $\bar{\eta}$ is any local section of $B(Q)$, and is equivariant: $R_g^*\Omega = g^{-1}\Omega$. The components of the tautological 1-form provide a partial coframing for $B(Q)$ and form a basis for the semibasic forms on $B(Q)$.

The following proposition reduces the problem of finding an equivalence between G-structures to finding a smooth map that preserves the tautological 1-form. (See Gardner [1989] or Bryant [1994] for a proof.)

Proposition 2.1. *Let $B(Q)$ and $B(N)$ be two G-structures with corresponding tautological 1-forms Ω_Q and Ω_N, and let $F : B(Q) \to B(N)$ be a smooth map. If G is connected and $F^*(\Omega_N) = \Omega_Q$, then there exists a local diffeomorphism $f : Q \to N$ for which $F = f_1$, i.e., the two G-structures are equivalent.*

To find the map F in this proposition we would like to apply *Cartan's technique of the graph* (see Warner [1971], p. 75): if we could find an integral manifold $\Sigma \subset B(Q) \times B(N)$ of the 1-form $\theta = \Omega_Q - \Omega_N$ that projects diffeomorphically onto each factor, then Σ would be the graph of a function $h : Q \to N$ for which $h_1^*\Omega_N = \Omega_Q$. By the above proposition the G-structures would then be equivalent. We generally cannot apply this idea directly because Ω_Q and Ω_N do not provide full coframes on $B(Q)$ and $B(N)$ as is required in the technique of the graph. In the example of nonholonomic geometry on Engel manifolds, and indeed in many important examples (see Gardner [1989], Hughen [1995], Moseley [2001], Montgomery [2002], Ehlers [2002]), applying the method of equivalence leads to a new G-structure called an e-structure. An e-structure is a G-structure endowed with a canonical coframe.

Differentiating both sides of (2.12) one can verify that $d\Omega$ satisfies the structure equation

$$d\Omega = -\alpha \wedge \Omega + T, \tag{2.13}$$

where T is a semibasic 2-form on $B(Q)$ and α is called a *pseudoconnection*: a $\mathrm{Lie}(G)$-valued 1-form on $B(Q)$ that agrees with the Mauer–Cartan form on vertical vector fields. Here, $\mathrm{Lie}(G)$ is the Lie algebra of G. Summarizing,

$$\text{Pseudoconnection:}\quad \alpha = g^{-1}dg + \text{semibasic } \mathrm{Lie}(G)\text{-valued 1-form.} \tag{2.14}$$

The components of the pseudoconnection together with the tautological 1-form do provide a full coframe on the G-structure, but unlike the tautological 1-form, the pseudoconnection is not canonically defined. *Understanding how changes in the pseudoconnection affect the torsion is at the heart of the method of equivalence*.

For any G-structure, that part of the torsion that is left unchanged under all possible changes of pseudoconnection is known as the *intrinsic torsion*. The intrinsic torsion is the only first-order differential invariant of the G-structure (Gardner [1989]). As an example, the intrinsic torsion for the G-structure B of a general distribution (equation 2.10) is the dual curvature of the distribution (Cartan [1910], see also Montgomery [2002]). In the case of a rank two distribution on a four-dimensional manifold, the structure equations for the tautological 1-form Ω are

$$d\begin{pmatrix}\Omega_1\\ \Omega_2\\ \Omega_3\\ \Omega_4\end{pmatrix} = -\begin{pmatrix}\mathcal{A}_{11} & \mathcal{A}_{12} & \beta_{13} & \beta_{14}\\ \mathcal{A}_{21} & \mathcal{A}_{22} & \beta_{23} & \beta_{24}\\ 0 & 0 & \alpha_{33} & \alpha_{34}\\ 0 & 0 & \alpha_{34} & \alpha_{44}\end{pmatrix} \wedge \begin{pmatrix}\Omega^1\\ \Omega^2\\ \Omega^3\\ \Omega^4\end{pmatrix} + \begin{pmatrix}T^1\\ T^2\\ T^3\\ T^4\end{pmatrix}, \tag{2.15}$$

where $T^I = \sum_{J<K} T^I_{JK}\Omega^J \wedge \Omega^K$ with $T^I_{JK} : B \to \mathbb{R}$. The intrinsic torsion consists of the terms $T^3_{12}\Omega^1 \wedge \Omega^2$ and $T^4_{12}\Omega^1 \wedge \Omega^2$. Note that the distribution is integrable if and only if $T^3_{12} = T^4_{12} = 0$.

Reduction and prolongation

There are two major steps in the equivalence method: *prolongation* and *reduction* (see Gardner [1989] or Montgomery [2002]). In the case of nonholonomic geometry on an Engel manifold a sequence of reductions lead to an e-structure. A brief outline of the reduction procedure is as follows. The first step involves writing out the structure equations for the tautological 1-form Ω. A semibasic Lie(G)-valued 1-form is added to the pseudoconnection to make the torsion as simple as possible. Gardner [1989] calls this step *absorption of torsion.* The action of G on the torsion is deduced by differentiating both sides of the identity $R_g^*(\Omega) = g^{-1}\Omega$. The action of G is used to simplify part of the torsion. The isotropy subgroup of that choice of simplified torsion is then the structure group of the reduced G-structure. In the case of nonholonomic geometry on an Engel manifold this procedure is repeated until an e-structure is obtained.

Suppose that Ω is the canonical coframing on the resulting manifold B. The Ω^i form a basis for the 1-forms on B so we can write

$$d\Omega^I = \sum_{J<K} c^I_{JK}\Omega^J \wedge \Omega^K. \tag{2.16}$$

Relationships between the c^I_{JK} are found by differentiating this equation. The resulting torsion functions provide the "complete invariants" for the geometric structure (see Gardner [1989], p. 59, Bryant [1994], pp. 9–10, or Cartan [2001]).

Many important examples have *integrable e-structures*. An e-structure is integrable if the c^I_{JK} are constant (Gardner [1989]). In this case we can apply the following result from Montgomery [2002].

Lemma 2.2. *Let B be an n-dimensional manifold endowed with a coframing Ω. Then the* (*local*) *group G of diffeomorphisms of B that preserves this coframing is a finite-dimensional* (*local*) *Lie group of dimension at most n. The bound n is achieved if and only if the e-structure is integrable. In this case the c^I_{JK} are the structure constants of G, G acts freely and transitively on B, and the coframe can be identified with the left-invariant* 1*-forms on G.*

The Jacobi identities are found by differentiating $d\Omega = \sum_{J<K} c^I_{JK}\Omega^J \wedge \Omega^K$. Lie's third fundamental theorem then implies that we can, at least in principle, reconstruct the group G using the structure constants. In some circumstances one can also conclude that B itself is a Lie group (see Gardner [1989], p. 72).

2.4 The nonholonomic geometry of an Engel manifold

The initial G structure for nonholonomic geometry on $\{Q, \mathcal{G}, \mathcal{H}\}$, where $\mathcal{H}$ is an Engel distribution on a four-dimensional manifold M is the subbundle $B_0 \subset F^*(Q)$ with structure group G_0 consisting of matrices of the form

$$\begin{pmatrix} A_{11} & A_{12} & 0 & B_{14} \\ A_{21} & A_{22} & 0 & B_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix}, \tag{2.17}$$

where $A = [A_{IJ}] \in O(2)$, $a_{33}a_{44} \neq 0$, and B_{14} and B_{24} are arbitrary.

Let $\Omega = (\Omega^1, \Omega^2, \Omega^3, \Omega^4)^{tr}$ be the tautological 1-form on B_0. The structure equations are

$$d\begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \\ \Omega_4 \end{pmatrix} = -\begin{pmatrix} 0 & \gamma & 0 & \beta_{14} \\ -\gamma & 0 & 0 & \beta_{24} \\ 0 & 0 & \alpha_{33} & \alpha_{34} \\ 0 & 0 & 0 & \alpha_{44} \end{pmatrix} \wedge \begin{pmatrix} \Omega^1 \\ \Omega^2 \\ \Omega^3 \\ \Omega^4 \end{pmatrix} + \begin{pmatrix} T^1_{13}\Omega^1 \wedge \Omega^3 + T^1_{23}\Omega^2 \wedge \Omega^3 \\ T^2_{13}\Omega^1 \wedge \Omega^3 + T^2_{23}\Omega^2 \wedge \Omega^3 \\ T^3_{12}\Omega^1 \wedge \Omega^2 \\ T^4_{13}\Omega^1 \wedge \Omega^3 + T^4_{23}\Omega^2 \wedge \Omega^3 \end{pmatrix}, \tag{2.18}$$

where we have chosen the pseudoconnection so that the remaining T^i_{jk} are zero. $\Omega^4 \in I^{(1)}$ so $d\Omega^4 = 0 \bmod (\Omega^3, \Omega^4)$ and we must therefore have $T^4_{12} = 0$. Also, $\Omega^3 \notin I^{(1)}$ so $d\Omega^3 \neq 0 \bmod (\Omega^3, \Omega^4)$; therefore, the torsion function T^3_{12} cannot equal zero. The pseudoconnection for this choice of torsion is not unique. We can, for instance, add arbitrary multiples of Ω^4 to the β_{i4} and α_{i4}.

Following Cartan's prescription, we investigate the induced action of G_0 on the torsion. Let $g \in G_0$. To simplify notation, functions and forms pulled back by R_g will be indicated by a hat so, for instance, $R_g^*\Omega = \hat{\Omega} = (\hat{\Omega}^1, \hat{\Omega}^2, \hat{\Omega}^3, \hat{\Omega}^4)^{\mathrm{tr}}$ and $R_g^*(T^k_{ij}) = \hat{T}^k_{ij}$. We have

$$\begin{pmatrix} \hat{\Omega}^1 \\ \hat{\Omega}^2 \\ \hat{\Omega}^3 \\ \hat{\Omega}^4 \end{pmatrix} = \begin{pmatrix} \# \\ \# \\ \det(a^{-1})(a_{44}\Omega^3 - a_{34}\Omega^4) \\ \det(a^{-1})(a_{33}\Omega^4) \end{pmatrix}. \tag{2.19}$$

To determine the induced action of G_0 on the torsion we differentiate both sides of the identity $R_g^*\Omega^3 = \hat{\Omega}^3$. For Ω^3, we compute

$$\begin{aligned} R_g^*(d\Omega^3) &= \hat{\alpha}_{33} \wedge \hat{\Omega}^3 + \hat{\alpha}_{34} \wedge \hat{\Omega}^4 + \hat{T}^3_{12}\hat{\Omega}^1 \wedge \hat{\Omega}^2, \\ &= \det(A^{-1})\hat{T}^3_{12}\Omega^1 \wedge \Omega^2 \qquad (\mathrm{mod}\ \Omega^3, \Omega^4), \end{aligned}$$

and

$$\begin{aligned} d\hat{\Omega}^3 &= \det(a^{-1})(a_{44}d\Omega^3 - a_{34}d\Omega^4) \qquad (\mathrm{mod}\ \Omega^3, \Omega^4) \\ &= \det(a^{-1})(a_{44}T^3_{12}\Omega^1 \wedge \Omega^2) \qquad (\mathrm{mod}\ \Omega^3, \Omega^4). \end{aligned}$$

The induced action of G_0 on T^3_{12} is therefore

$$R_g^*(T^3_{12}) = \frac{\det(A)}{a_{33}} T^3_{12}. \tag{2.20}$$

Since $T^3_{12} \neq 0$ we can force it to equal 1 using the action of G_0. The stabilizer subgroup G_1 for this choice of torsion consists of matrices of the form (2.17) with $a_{33} = \epsilon$ where $\epsilon = \det(A)$. Note that $T^3_{12}\Omega^1 \wedge \Omega^2$ is the (normalized) dual curvature of the distribution.

The structure equations for the G_1-structure B_1 are

$$\begin{aligned} d\begin{pmatrix}\Omega_1\\ \Omega_2\\ \Omega_3\\ \Omega_4\end{pmatrix} &= -\begin{pmatrix} 0 & \gamma & 0 & \beta_{14}\\ -\gamma & 0 & 0 & \beta_{24}\\ 0 & 0 & 0 & \alpha_{34}\\ 0 & 0 & 0 & \alpha_{44}\end{pmatrix} \wedge \begin{pmatrix}\Omega^1\\ \Omega^2\\ \Omega^3\\ \Omega^4\end{pmatrix} \\ &\quad + \begin{pmatrix} T^1_{13}\Omega^1\wedge\Omega^3 + T^1_{23}\Omega^2\wedge\Omega^3\\ T^2_{13}\Omega^1\wedge\Omega^3 + T^2_{23}\Omega^2\wedge\Omega^3\\ T^3_{13}\Omega^1\wedge\Omega^3 + T^3_{23}\Omega^2\wedge\Omega^3 + \Omega^1\wedge\Omega^2\\ T^4_{13}\Omega^1\wedge\Omega^3 + T^4_{23}\Omega^2\wedge\Omega^3\end{pmatrix}. \end{aligned} \tag{2.21}$$

Let $g \in G_1$. We write the inverse of g as

$$g^{-1} = \begin{pmatrix} A_{11} & A_{21} & 0 & \bar{B}_{14}\\ A_{12} & A_{22} & 0 & \bar{B}_{24}\\ 0 & 0 & \bar{a}_{33} & \bar{a}_{34}\\ 0 & 0 & 0 & \bar{a}_{44}\end{pmatrix}, \tag{2.22}$$

so, in particular, $\bar{a}_{33} = \epsilon$, $\bar{a}_{34} = -\epsilon a_{34}(a_{44})^{-1}$, and $\bar{a}_{44} = (a_{44})^{-1}$. We have

$$R_g^*\Omega = \begin{pmatrix}\hat{\Omega}^1\\ \hat{\Omega}^2\\ \hat{\Omega}^3\\ \hat{\Omega}^4\end{pmatrix} = \begin{pmatrix} A_{11}\Omega^1 + A_{21}\Omega^2 + \bar{B}_{14}\Omega^4\\ A_{12}\Omega^1 + A_{22}\Omega^2 + \bar{B}_{24}\Omega^4\\ \bar{a}_{33}\Omega^3 + \bar{a}_{34}\Omega^4\\ \bar{a}_{44}\Omega^4\end{pmatrix}. \tag{2.23}$$

For the next reduction we differentiate both sides of the identity $R_g^*\Omega^4 = \hat{\Omega}^4$. We have

$$\begin{aligned} R_g^* d\Omega^4 &= \hat{\alpha}_{44}\wedge\hat{\Omega}^4 + \hat{T}^4_{13}\hat{\Omega}^1\wedge\hat{\Omega}^3 + \hat{T}^4_{23}\hat{\Omega}^2\wedge\hat{\Omega}^3 && \pmod{\Omega^4},\\ &= \bar{a}_{33}((A_{11}\hat{T}^4_{13} + A_{12}\hat{T}^4_{23})\Omega^1\wedge\Omega^3 \\ &\quad + (A_{21}\hat{T}^4_{13} + A_{22}\hat{T}^4_{23})\Omega^2\wedge\Omega^3) && \pmod{\Omega^4}. \end{aligned}$$

On the other hand,

$$d\hat{\Omega}^4 = \bar{a}_{44} d\Omega^4 = \bar{a}_{44}(T^4_{13}\Omega^1\wedge\Omega^3 + T^4_{23}\Omega^2\wedge\Omega^3) \pmod{\Omega^4}.$$

The induced action of G_1 on the torsion plane (T^4_{13}, T^4_{23}) is therefore

$$\begin{pmatrix} \hat{T}^4_{13} \\ \hat{T}^4_{23} \end{pmatrix} = \frac{\epsilon}{a_{44}} A^{-1} \begin{pmatrix} T^4_{13} \\ T^4_{23} \end{pmatrix}. \tag{2.24}$$

The torsion plane $(T^4_{13}, T^4_{23}) \neq (0, 0)$ since $I^{(2)} = 0$ implies that $d\Omega^4 \wedge \Omega^4 \neq 0$ and we have already established that $T^4_{12} = 0$. We can therefore use the action to force $(T^4_{13}, T^4_{23}) = (0, 1)$. The torsion $T^4_{23}\Omega^2 \wedge \Omega^3$ can be interpreted as the (normalized) dual curvature of the rank three distribution $\mathcal{H}^1$. The statement that $(T^4_{13}, T^4_{23}) \neq (0, 0)$ is equivalent to $\mathcal{H}$ not being integrable. To determine the subgroup that stabilizes this choice of torsion, we investigate

$$R_g^* \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{\epsilon}{a_{44}} A^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.25}$$

As $A \in O(2)$, it must be of the form

$$\begin{pmatrix} \epsilon_1\epsilon_2 & 0 \\ 0 & \epsilon_2 \end{pmatrix}, \tag{2.26}$$

where $\epsilon_1, \epsilon_2 \in \{-1, 1\}$. We must also have $a_{44} = \epsilon_1\epsilon_2$ so that the stabilizer subgroup G_2 consists of matrices of the form

$$\begin{pmatrix} \epsilon_1\epsilon_2 & 0 & 0 & B_{14} \\ 0 & \epsilon_2 & 0 & B_{24} \\ 0 & 0 & \epsilon_1 & a_{34} \\ 0 & 0 & 0 & \epsilon_1\epsilon_2 \end{pmatrix}, \tag{2.27}$$

where $\epsilon_1, \epsilon_2 \in \{-1, 1\}$ and B_{14}, B_{24} and $a_{34} \in \mathbb{R}$. We compute

$$R_g^*\Omega = \begin{pmatrix} \hat{\Omega}^1 \\ \hat{\Omega}^2 \\ \hat{\Omega}^3 \\ \hat{\Omega}^4 \end{pmatrix} = \begin{pmatrix} \epsilon_1\epsilon_2\Omega^1 - B_{14}\Omega^4 \\ \epsilon_2\Omega^2 - B_{24}\Omega^4 \\ \epsilon_1\Omega^3 - \epsilon_2 a_{34}\Omega^4 \\ \epsilon_1\epsilon_2\Omega^4 \end{pmatrix}. \tag{2.28}$$

The structure equations are now

$$\begin{aligned} d\begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \\ \Omega_4 \end{pmatrix} &= -\begin{pmatrix} \beta_{14} \wedge \Omega^4 \\ \beta_{24} \wedge \Omega^4 \\ \alpha_{34} \wedge \Omega^4 \\ 0 \end{pmatrix} \\ &\quad + \begin{pmatrix} T^1_{12}\Omega^1 \wedge \Omega^2 + T^1_{13}\Omega^1 \wedge \Omega^3 + T^1_{23}\Omega^2 \wedge \Omega^3 \\ T^2_{12}\Omega^1 \wedge \Omega^2 + T^2_{13}\Omega^1 \wedge \Omega^3 + T^2_{23}\Omega^2 \wedge \Omega^3 \\ \Omega^1 \wedge \Omega^2 + T^3_{13}\Omega^1 \wedge \Omega^3 + T^3_{23}\Omega^2 \wedge \Omega^3 \\ T^4_{14}\Omega^1 \wedge \Omega^4 + \Omega^2 \wedge \Omega^3 + T^4_{24}\Omega^2 \wedge \Omega^4 + T^4_{34}\Omega^3 \wedge \Omega^4 \end{pmatrix}. \end{aligned} \tag{2.29}$$

B_2 is not an e-structure, so again we differentiate both sides of the identity $R_g^*\Omega = \hat{\Omega}$ to determine the action of G_2 on the torsion. After some computation, we find that

$$\begin{aligned}
d\hat{\Omega}^1 &= \epsilon_1\epsilon_2(T^1_{13}\Omega^1\wedge\Omega^3 + T^1_{23}\Omega^2\wedge\Omega^3 + T^1_{12}\Omega^1\wedge\Omega^2)\\
&\qquad - B_{14}\Omega^2\wedge\Omega^3 \pmod{\Omega^4},\\
d\hat{\Omega}^2 &= \epsilon_2(T^2_{13}\Omega^1\wedge\Omega^3 + T^2_{23}\Omega^2\wedge\Omega^3 + T^2_{12}\Omega^1\wedge\Omega^2)\\
&\qquad - \epsilon_1 B_{24}\Omega^2\wedge\Omega^3 \pmod{\Omega^4},\\
d\hat{\Omega}^3 &= \epsilon_1(T^3_{13}\Omega^1\wedge\Omega^3 + T^3_{23}\Omega^2\wedge\Omega^3 + \Omega^1\wedge\Omega^2)\\
&\qquad - \epsilon_2 a_{34}\Omega^2\wedge\Omega^3 \pmod{\Omega^4}.
\end{aligned}$$

Also,

$$\begin{aligned}
R^*_g(d\Omega^1) &= \epsilon_2\hat{T}^1_{13}\Omega^1\Omega^3 + \epsilon_1\epsilon_2\hat{T}^1_{23} + \epsilon_1\hat{T}^1_{12}\Omega^1\wedge\Omega^2 \pmod{\Omega^4},\\
R^*_g(d\Omega^2) &= \epsilon_2\hat{T}^2_{13}\Omega^1\Omega^3 + \epsilon_1\epsilon_2\hat{T}^2_{23} + \epsilon_1\hat{T}^2_{12}\Omega^1\wedge\Omega^2 \pmod{\Omega^4},\\
R^*_g(d\Omega^3) &= \epsilon_1\hat{T}^3_{13}\Omega^1\Omega^3 + \epsilon_1\epsilon_2\hat{T}^3_{23} + \epsilon_1\Omega^1\wedge\Omega^2 \pmod{\Omega^4}.
\end{aligned}$$

Matching the $\Omega^2\wedge\Omega^3$ terms, we find that

$$\begin{aligned}
\hat{T}^1_{23} &= T^1_{23} - \epsilon_1\epsilon_2 B_{14},\\
\hat{T}^2_{23} &= \epsilon_1 T^2_{23} - \epsilon_2 B_{24},\\
\hat{T}^3_{23} &= \epsilon_2 T^3_{23} - \epsilon_1 a_{34}.
\end{aligned} \tag{2.30}$$

We can therefore use the action of G_1 to force $T^1_{23} = T^2_{23} = T^3_{23} = 0$. The stabilizer subgroup G_{final} for this choice of torsion consists of matrices of the form

$$\begin{pmatrix} \epsilon_1\epsilon_2 & 0 & 0 & 0\\ 0 & \epsilon_2 & 0 & 0\\ 0 & 0 & \epsilon_1 & 0\\ 0 & 0 & 0 & \epsilon_1\epsilon_2 \end{pmatrix}. \tag{2.31}$$

The reduced structure group is discrete so we now have an e-structure B_{final}. The tautological 1-form $(\Omega_1, \Omega_2, \Omega_3, \Omega_4)^{\text{tr}}$ provides a full coframing for B_{final}. The B_{final} structure equations are

$$d\begin{pmatrix}\Omega_1\\ \Omega_2\\ \Omega_3\\ \Omega_4\end{pmatrix} = \begin{pmatrix} T^1_{12} & T^1_{13} & T^1_{14} & 0 & T^1_{24} & T^1_{34}\\ T^2_{12} & T^2_{13} & T^2_{14} & 0 & T^2_{24} & T^2_{34}\\ 1 & T^3_{13} & T^3_{14} & 0 & T^3_{24} & T^3_{34}\\ 0 & 0 & T^4_{14} & 1 & T^4_{24} & T^4_{34} \end{pmatrix}\begin{pmatrix}\Omega^1\wedge\Omega^2\\ \Omega^1\wedge\Omega^3\\ \Omega^1\wedge\Omega^4\\ \Omega^2\wedge\Omega^3\\ \Omega^2\wedge\Omega^4\\ \Omega^3\wedge\Omega^4\end{pmatrix}, \tag{2.32}$$

where the T^k_{ij} are functions on B_{final}. What remains is to determine any second order relations between the torsion functions. To determine them we use the fact that $d^2 = 0$. After some computation, we find that $T^4_{14} = T^2_{12} + T^3_{13}$. We summarize these results in the following theorem:

Theorem 2.3. *Associated to any nonholonomic Engel structure* $\{Q, \mathcal{G} = \langle\cdot,\cdot\rangle, \mathcal{H}\}$ *there is a canonical* $G \cong \mathbf{Z}_2 \times \mathbf{Z}_2$*-structure* B_{final}. *The tautological* 1*-form* $(\Omega_1, \Omega_2, \Omega_3, \Omega_4)^{tr}$ *provides a canonical coframing for* B_{final}. *The* B_{final} *structure equations are*

$$d\begin{pmatrix}\Omega_1\\ \Omega_2\\ \Omega_3\\ \Omega_4\end{pmatrix} = \begin{pmatrix} T^1_{12} & T^1_{13} & T^1_{14} & 0 & T^1_{24} & T^1_{34}\\ T^2_{12} & T^2_{13} & T^2_{14} & 0 & T^2_{24} & T^2_{34}\\ 1 & T^3_{13} & T^3_{14} & 0 & T^3_{24} & T^3_{34}\\ 0 & 0 & T^2_{12}+T^3_{13} & 1 & T^4_{24} & T^4_{34}\end{pmatrix}\begin{pmatrix}\Omega^1\wedge\Omega^2\\ \Omega^1\wedge\Omega^3\\ \Omega^1\wedge\Omega^4\\ \Omega^2\wedge\Omega^3\\ \Omega^2\wedge\Omega^4\\ \Omega^3\wedge\Omega^4\end{pmatrix}. \tag{2.33}$$

According to the *framing lemma* (Lemma 2.2) the largest Lie group of symmetries of a nonholonomic structure on an Engel manifold is the dimension of B_{final} which is four. In this case the T^I_{JK} are constants and can be identified with the structure constants of the four-dimensional Lie algebra of the symmetry group. The Jacobi identities are obtained using the identity $d^2 = 0$. We have computed them, and it appears that the set of possible symmetry algebras form a rather complicated subvariety of the variety of all four-dimensional Lie algebras. We leave as an open problem the classification of all possible four-dimensional symmetry algebras for nonholonomic structures on an Engel manifold.

The rolling penny (*continued*)

An example of a structure with maximal symmetry is given by the rolling penny. A B_{final}-adapted coframe for the penny-table system is

$$\begin{aligned}
\eta^1 &= \sqrt{\frac{ma^2+I}{2}}\,d\phi,\\
\eta^2 &= \sqrt{\frac{J}{2}}\,d\theta,\\
\eta^3 &= \frac{\sqrt{J(ma^2+I)}}{2}(-\sin\theta dx + \cos\theta dy),\\
\eta^4 &= \sqrt{\frac{m}{2}}(\cos\theta dx + \sin\theta dy - d\phi).
\end{aligned} \tag{2.34}$$

The structure equations are

$$\begin{aligned}
d\eta^1 &= 0,\\
d\eta^2 &= 0,\\
d\eta^3 &= \eta^1\wedge\eta^2 - \sqrt{\frac{ma^2+I}{m}}\,\eta^2\wedge\eta^4,\\
d\eta^4 &= \frac{2}{J}\sqrt{\frac{m}{ma^2+I}}\,\eta^2\wedge\eta^3.
\end{aligned} \tag{2.35}$$

The torsion functions are constant, so by the framing lemma (Lemma 2.2) we can identify these constants with the structure constants Lie group of symmetries of this system. We recognize them as the structure constants for the Lie algebra of the group $SE(2) \times SO(2)$ which is isomorphic to the configuration space of the penny-table system.

B_{final}-adapted frames and coframes

The e-structure B_{final} has a canonical coframing which descends to a coframing and hence a framing, up to signs, on Q. There should be a relationship between this framing and a canonical line field possessed by any Engel manifold. In this section we briefly describe this relationship. If Q and $\mathcal{H}$ are both oriented, then Q is parallelizable and the following constructions can be made globally (Montgomery [2002]).

Let η be a B_{final} adapted coframe on $U \subset Q$ with dual frame $X = \{X_I\}$ defined by $\eta^I(X_J) = \delta_{IJ}$. If $\bar{\eta}$ is any other B_{final}-adapted coframe with dual frame $\bar{X}$on U, then by Theorem 2.3, $\bar{\eta}$ is related to η by $\bar{\eta}^1 = \epsilon_1\epsilon_2\eta^1$, $\bar{\eta}^2 = \epsilon_2\eta^1$, $\bar{\eta}^3 = \epsilon_1\eta^3$, $\bar{\eta}^4 = \epsilon_1\epsilon_2\eta^4$. The dual frames are related in precisely the same way: $\bar{X}_1 = \epsilon_1\epsilon_2 X_1$, $\bar{X}_2 = \epsilon_2 X_2$, $\bar{X}_3 = [\bar{X}_1, \bar{X}_2] = \epsilon_1 X_3$, and $\bar{X}_4 = [\bar{X}_2, \bar{X}_3] = \epsilon_1\epsilon_2 X_4$.

An important feature of an Engel distribution is the presence of a canonical line field $L \subset \mathcal{H}$ (Montgomery [2002], Kazarian, Montgomery, and Shapiro [1997]). L is defined by the condition that $[L, \mathcal{H}^1] \subset \mathcal{H}^1$. Here we are abusing notation, using L for the line field or a vector field spanning L. We have

Corollary 2.4. *Let $\eta = \eta^I$ be a B_{final}-adapted coframe. Let $X = \{X_I\}$ be the dual frame defined by $\eta^I(X_J) = \delta_{IJ}$; then $L = \text{span}(X_1)$.*

Proof. Suppose L is spanned by the vector field $Y = aX_1 + bX_2$. Since η^4 annihilates $\mathcal{H}^1$ we have $\eta^4([X_3, Y]) = 0$. Then

$$0 = \eta^4([X_3, Y]) = X_3\eta^4(Y) - Y\eta^4(X_3) - d\eta^4(X_3, Y) = -d\eta^4(X_3, Y).$$

But $d\eta^4 \equiv \eta^2 \wedge \eta^3 \bmod(\eta^4)$, so we must have

$$\begin{aligned} 0 = \eta^2 \wedge \eta^3(X_3, Y) &= \eta^2(X_3)\eta^3(Y) - \eta^3(X_3)\eta_2(Y), \\ &= -\eta^3(X_3)\eta_2(Y), \\ &= -b. \end{aligned}$$

L is therefore spanned by X_1. This concludes the arguement.

There is a natural metric, associated with B_{final}, on Q given by $g_{\text{nat}} = \tilde{\eta}^1 \otimes \tilde{\eta}^1 + \cdots \tilde{\eta}^4 \otimes \tilde{\eta}^4$ where $\tilde{\eta}$ is any B_{final}-adapted coframe. Clearly all B_{final}-adapted coframes induce this same metric; using the sub-Riemannian metric $g_{\text{nat}}|_{\mathcal{H}}$ we form $L^{\perp}$ within $\mathcal{H}$ so that $\mathcal{H} = L \oplus L^{\perp}$. By construction, X_2 spans $L^{\perp}$.

3 Nonholonomic dynamics: Chaplygin Hamiltonization

Historically, Hamiltonization of nonholonomic systems started with Chaplygin's *last multiplier method*. In the new time, the dynamics obeys Euler–Lagrange equations *without extra terms*; the gyroscopic force (1.13) "magically" disappears! When, after a time reparametrization the compressed system can be described as a Hamiltonian system, symplectic techniques can be employed. A number of NH systems have been Hamiltonized, and some interesting ones are Liouville-integrable; see Veselov and Veselova [1988], Kozlov [2002], Fedorov and Jovanovic [2003], Fedorov [1989], Dragovic, Gajić, and Jovanovic [1998], Borisov and Mamaev [2002a], Borisov and Mamaev [2002b], Borisov, Mamaev, and Kilin [2002], Borisov and Mamaev [2001], Jovanovic [2003].

3.1 Compression to T^*S, $S = Q/G$; existence of invariant measures

We recall from the introduction that the compressed system has a concise almost Hamiltonian

$$dH^\phi = i_{X_{NH}}\Omega_{NH}, \quad \Omega_{NH} := \Omega_{\text{can}}^{T^*S} + (\text{J.K}), \quad d\Omega_{NH} \neq 0 \quad \text{(in general)},$$

where $\Omega_{\text{can}}^{T^*S}$ is the canonical 2-form of T^*S and the (J.K) term is a semibasic 2-form, which in general is nonclosed. It combines the momentum J of the G-action on T^*Q, and the curvature K of the connection. As this is important for the remaining, we outline the derivation (see Koiller, Rios, and Ehlers [2002] for details). Given the coframe coordinates $m, \epsilon(q)$ in T^*Q (see 1.5) the Poisson bracket matrix relative to ϵ_I, dm_I is

$$[\Lambda] = [\Omega]^{-1} = \begin{pmatrix} 0_n & I_n \\ -I_n & E \end{pmatrix} \tag{3.1}$$

with

$$E_{JK} = m_I d\epsilon_I(e_J, e_K) = -m_I\epsilon_I[e_J, e_K]. \tag{3.2}$$

Let us consider the case of a principal bundle $\pi : Q^n \to S^s$ with Lie group G^r acting on the left, $r = n - s$. Recall our convention: capital roman letters I, J, K, etc., run from 1 to n. Lower case roman characters i, j, k run from 1 to s. Greek characters α, β, γ, etc., run from $s + 1$ to n.

Fix a connection $\lambda = \lambda(q) : T_qQ \to \text{Lie}(G)$ defining a G-invariant distribution $\mathcal{H}$ of horizontal subspaces. Denote by $K(q) = d\lambda \circ \text{Hor} : T_qQ \times T_qQ \to \text{Lie}(G)$ the curvature 2-form (which is, as is well known, Ad-equivariant). Choose a local frame $\overline{e}_i$ on S. For simplicity, we may assume that

$$\overline{e}_i = \partial/\partial s_i \tag{3.3}$$

are the coordinate vector fields of a chart $s : S \to \mathbb{R}^s$.

Let $e_i = h(\overline{e}_i)$ be their horizontal lift to Q. We complete to a moving frame of Q with vertical vectors e_α which we will specify in a moment. The dual basis will be denoted $\epsilon_i, \epsilon_\alpha$ and we write $p_q = m_i\epsilon_i + m_\alpha\epsilon_\alpha$. These are in a sense the "least moving" among all the moving frames adapted to this structure. We now describe how the $n \times n$ matrix $E = (E_{IJ})$ looks like in this setting.

(i) *The $s \times s$ block* (E_{ij}). Decompose $[e_i, e_j] = h[\overline{e}_i, \overline{e}_j] + V[e_i, e_j] = V[e_i, e_j]$ into vertical and horizontal parts. The choice (3.3) is convenient, since $\overline{e}_i$ and $\overline{e}_j$ commute: $[e_i, e_j]$ is vertical. Hence

$$E_{ij} = -p_q[e_i, e_j] = -m_\alpha \epsilon_\alpha [e_i, e_j]. \tag{3.4}$$

Now by Cartan's rule,

$$K(e_i, e_j) = e_i \lambda(e_j) - e_j \lambda(e_i) - \lambda[e_i, e_j] = -\lambda[e_i, e_j] \in \mathrm{Lie}(G).$$

Thus we showed that

$$[e_i, e_j]_q = -K(e_i, e_j) \cdot q \tag{3.5}$$

Moreover, let $J : T^*Q \to \mathrm{Lie}(G)^*$ the momentum mapping. We have

$$(J(p_q), K_q(e_i, e_j)) = p_q(K(e_i, e_j).q) = -p_q[e_i, e_j] \quad (= E_{ij}).$$

Theorem 3.1 (the J.K formula).

$$E_{ij} = (J(p_q), K_q(e_i, e_j)). \tag{3.6}$$

This gives a nice description for this block, under the choice $[\overline{e}_i, \overline{e}_j] = 0$. Notice that the functions E_{ij} depend on s and the components m_α, but do not depend on g. This is because the Ad^*-ambiguity of the momentum mapping J is cancelled by the Ad-ambiguity of the curvature K. The other blocks are not needed here, but we include for completeness.

(ii) *The $r \times r$ block* $(E_{\alpha\beta})$. Choose a basis X_α for $\mathrm{Lie}(G)$. We take $e_\alpha(q) = X_\alpha \cdot q$ as the vertical distribution. Choosing a point q_o allows identifying the Lie group G with the fiber containing Gq_o, so that $\mathrm{id} \mapsto q_o$. Through the mapping $g \in G \mapsto gq_o \in Gq_o$ the vector field e_α is identified with a *right*-invariant (not left-invariant!) vector field in G. The commutation relations for the e_α $[e_\alpha, e_\beta] = -c^\gamma_{\alpha\beta} e_\gamma$ appear with a minus sign. Therefore,

$$E_{\alpha\beta} = m_\gamma c^\gamma_{\alpha\beta}. \tag{3.7}$$

(iii) *The $s \times n$ block* $(E_{i\alpha})$. The vectors $[e_i, e_\alpha]$ are vertical, but their values depend on the specific principal bundle one is working with, and there are some noncanonical choices. Given a section $\sigma : U_S \to Q$ over the coordinate chart $s : U_S \to \mathbb{R}^m$ on S, we need to know the coefficients $b^\gamma_{i\alpha}$ in the expansion

$$[e_i, e_\alpha](\sigma(s)) = b^\gamma_{i\alpha}(s) e_\gamma.$$

Then

$$E_{i\alpha}(\sigma(s)) = -m_\gamma b^\gamma_{i\alpha}(s). \tag{3.8}$$

At another point on the fiber, we need the adjoint representation $\mathrm{Ad}_g : \mathrm{Lie}(G) \to \mathrm{Lie}(G)$, $X \mapsto g_*^{-1} X g$, described by a matrix $(A_{\mu\alpha}(g))$ such that

$$\mathrm{Ad}_g(X_\alpha) = A_{\mu\alpha}(g) X_\mu. \tag{3.9}$$

Then

$$[e_i, e_\alpha](g \cdot \sigma(s)) = -m_\gamma b^\gamma_{i\mu}(s) A_{\mu\alpha}(g). \tag{3.10}$$

The clockwise diagram

Starting on $p_s \in T^*S$ we go clockwise to $P_q \in \mathrm{Leg}(\mathcal{H}) \subset T^*Q$, for some q on the fiber $\pi^{-1}(s)$ of Q over s.

$$\begin{tikzcd}
\mathcal{H} \subset TQ \arrow[r, "\mathrm{Leg}"'] & \mathrm{Leg}(\mathcal{H}) \subset T^*Q \\
TS \arrow[u, "h"] & T^*S. \arrow[l, "(\mathrm{Leg}^{\phi})^{-1}"]
\end{tikzcd} \tag{3.11}$$

Taking differentials of all maps in (3.11) we obtain an induced principal connection $\hat{\phi}$ in the bundle $G \hookrightarrow \mathrm{Leg}(\mathcal{H}) \to T^*S$. Let $v, w, z \in T_{p_s}(T^*S)$, V, W, Z horizontal lifts at $P_q \in \mathrm{Leg}(\mathcal{H})$, and denote by $\hat{K}$ the curvature of this induced connection. The following proposition is basically a rephrasing of a result in Bates and Śniatycki [1993].

Proposition 3.2.

$$d(\mathrm{J.K})(v, w, z) = \mathrm{cyclic}(dJ(V), K(W, Z)). \tag{3.12}$$

Densities of invariant measures and a dimension dependent exponent

A necessary and sufficient condition for the existence of an invariant measure for compressed Chaplygin systems was obtained by Cantrijn, Cortés, de Léon, and de Diego [2002] (Theorem 7.5). Since in T^*S there is a natural Liouville measure $d\mathrm{vol} = ds_1 \cdots ds_m dp_1 \cdots dp_m$, where (s, p) are coordinates in T^*S, the *density function F produces an educated guess for a time reparametrization* which may Hamiltonize the compressed system. If $\dim(S) = m$ and $f\Omega_{NH}$ is closed, the time-reparametrized vector field X_{NH}/f has the invariant measure $f^m d\mathrm{vol}$. X_{NH} will have the invariant measure $f^{m-1} ds_1 \cdots ds_m dp_1 \cdots dp_m$. Working backwards, if a measure density F is known so that $F(s)d\mathrm{vol}$ is an invariant measure for X_{NH}, then the obvious *candidate* for conformal factor is

$$f = F(s)^{\frac{1}{m-1}}. \tag{3.13}$$

This dimension dependent exponent will be relevant in the Chaplygin marble; see Section 3.2.

Invariant measures for LR systems

Let $Q = G$ be a unimodular Lie group and identify $TG \equiv T^*G$ via the bi-invariant metric. Assume that $H \subset G$ is a subgroup acting on the left and preserving the distribution: $\mathcal{D}_{hg} = h\mathcal{D}_g = h\mathcal{D}g$ (which boils down to $\mathrm{Ad}_{h^{-1}}\mathcal{D} = h^{-1}\mathcal{D}h = \mathcal{D}$). The

Legendre transform Leg : $\mathrm{Lie}(G) \to \mathrm{Lie}(G) \equiv \mathrm{Lie}^*(G)$ of a natural, left-invariant Lagrangian, is represented by a positive symmetric transformation $A : \mathrm{Lie}(G) \to \mathrm{Lie}(G)$, the *inertia operator*.

For each $g \in G$, let P_g^1 and P_g^2 be, respectively, the projections of $\mathrm{Lie}(G)$ relative to the decomposition $\mathrm{Lie}(G) = \mathrm{Ad}_{g^{-1}} \mathrm{Lie}(H) \oplus \mathrm{Ad}_{g^{-1}} \mathcal{D}$. We can also think of P_g^2 as a map $P_g^2 : T_g G \to \mathcal{D}g$, projection parallel to the vertical spaces $\mathrm{Lie}(H)g$. Let $P_g^2 o \,\mathrm{Leg}_g : \mathcal{D}g \to \mathcal{D}g$. This map descends to the compressed Legendre transform $\mathrm{Leg}_s^\phi : T_s S \to T_s S \equiv T_s^* S$, where $S = G/H$ is the homogeneous space whose metric is induced by the bi-invariant metric on G. Consider the function

$$F(s) = \det \mathrm{Leg}_s^\phi\,. \tag{3.14}$$

The following result is a rephrasing of a theorem by Veselov and Veselova [1988]; see also Fedorov and Jovanovic [2003], Theorem 3.3.[17]

Theorem 3.3. *The reduced LR-Chaplygin system in the homogeneous space $T^*(G/H)$ always has the invariant measure*

$$\nu = F(s)^{-1/2} ds_1 \cdots ds_m dp_1 \cdots dp_m, \quad F(s) = \det \mathrm{Leg}_s^\phi\,. \tag{3.15}$$

The density can be also calculated by the "dual" formula

$$F(s) = \det(A) \det(P_g^2 o A^{-1}|_{g^{-1}\,\mathrm{Lie}(H) g}) \tag{3.16}$$

(P_g^1 *is the projection over* $g^{-1}\,\mathrm{Lie}(H) g$ *parallel to* $g^{-1}\mathcal{D}g$).

The second formula may be easier to use if there are few constraints.

Almost Hamiltonian systems

Let Ω be a nondegenerate (but in general, nonclosed) 2-form on M^{2n}, and H a function on M. Denote (as usual) by $X = X_H$ the skew-gradient vector field defined by $i_X \Omega = dH$. We say X_H is *almost Hamiltonian*. If α is a closed 1-form, the vector field $X = X_\alpha$ defined by $i_X \Omega = \alpha$ is called *locally almost Hamiltonian*. Distilling a construction in Stanchenko [1985], we formalize an extension of the notion of a conformally symplectic structure.

The 2-form Ω is called H *(or α)-affine symplectic* if there is a function $f > 0$ on M and a 2-form Ω_o such that (i) $i_X \Omega_o \equiv 0$; (ii) $\Omega - \Omega_o$ is nondegenerate; and (iii) $\tilde{\Omega} = f(\Omega - \Omega_o)$ is closed.[18]

[17] We do not need to assume $\mathcal{D}$ and $\mathrm{Lie}(H)$ to be orthogonal with respect to the bi-invariant metric.

[18] We must admit, however, that we found no example yet where the affine term is really needed. This notwithstanding, at any point where $X \neq 0$, the contraction condition yields $d = 2n$ equations on $d(d-1)/2$ unknowns (local coordinate coefficients of Ω_o). This allows additional freedom to Hamiltonize X rather than just requiring conformality of Ω.

The first condition implies that X does not "see" Ω_o. Together with the third, we get $\tilde{\Omega}(X/f, \bullet) = dH$ so the vector field X/f is (truly) Hamiltonian with respect to the symplectic form $\tilde{\Omega}$.

The closedness condition can be restated as

$$d(\Omega - \Omega_o) = (\Omega - \Omega_o) \wedge \theta, \quad \text{where } \theta = df/f. \tag{3.17}$$

When (3.17) holds with α a closed (but not necessarily exact) 1-form, we say that Ω is *locally* affine symplectic. The following proposition describes the obstruction to Hamiltonization once f is given.

Theorem 3.4. *Given a locally almost hamiltonian system* (Ω, α) *and an educated guess* $f > 0$*, an affine term* Ω_o *exists with* $d(f\Omega - \Omega_o) = 0$ *if and only if* $i_X d(f\Omega) = 0$*.*

The proof is quite easy. The vector field X satisfies $i_X\Omega = \alpha$. Since the same equation holds by replacing X by X/f and Ω by $f\Omega$, to expedite notation we may assume $f \equiv 1$. Let us prove that Ω_o exists if $i_X d\Omega = 0$. Since $d(i_X\Omega) = d\alpha = 0$, we see that the Lie derivative $L_X\Omega = 0$. Consider a regular point of X. By the flow box theorem there are coordinates so that $X = \partial/\partial x_1$. Since $L_X\Omega = 0$, the coefficients of this 2-form do not depend on the coordinate x_1 (but there may exist terms with a dx_1 factor). However, our hypothesis $i_{\partial/\partial x_1} d\Omega = 0$ ensures that there are no terms containing a dx_1 factor in $d\Omega$. Thus $d\Omega$ can be thought of as a 3-form in the space of the remaining coordinates. By Poincaré's theorem $d\Omega = d\Omega_o$, where Ω_o is a 2-form in the space of the remaining coordinates. Hence $i_X\Omega_o = 0$ and $d(\Omega - \Omega_o) = 0$, as desired. The converse is even easier.

3.2 Examples: Veselova's system and Chaplygin spheres (marble or rubber)

Veselov and Veselova [1986, 1988] considered one of the simplest nonholonomic LR-Chaplygin systems, $Q = SO(3)$ with a left-invariant metric $L = T = \frac{1}{2}(A\Omega, \Omega)$, and subject to a right-invariant constraint which, without loss of generality, can be assumed to be $\rho_3 = 0$. Hence the admissible motions satisfy $\omega_3 = 0$, where ω is the angular velocity viewed in the space frame. This is a LR Chaplygin system on $S^1 \hookrightarrow SO(3) \to S^2$.

Chaplygin's ball is a sphere of radius r and mass μ, whose center of mass is assumed to be at the geometric center, but the inertia matrix $A = \mathrm{diag}(I_1, I_2, I_3)$ may have unequal entries. Thus its Lagrangian is given by $2L = (A\Omega, \Omega) + \mu(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$. The configuration space is the Euclidean group $Q = SE(3)$.

In the case of the *marble*, the ball rolls without slipping on a horizontal plane, with rotations about the z-axis allowed.[19] Thus the distribution of admissible velocities is

[19] Chaplygin [2002] showed that the 3D problem is integrable using elliptic coordinates in the sphere; for $n > 3$ the problem is open. For basic informations, see Fedorov and Kozlov [1995], pp. 147–149, on the 3D case and pp. 153–156 for the general n-dimensional case. For a detailed account of the algebraic integrability of "Chaplygin's Chaplygin sphere"; see Duistermaat [2000]. Schneider [2002] analyzed control theoretical aspects.

defined by $\mathcal{D} : \dot{z} = 0$, $\dot{x} = r\omega_1$, $\dot{y} = -r\omega_2$. Both Lagrangian and constraints are preserved under the action of the Euclidean motions in the plane, together with the vertical translations. $G = SE(2) \times \mathbb{R}$ acts on Q via

$$(\phi, u, v, w).(R, x, y, z) = (S(\phi)R, e^{i\phi}(u + iv), z + w).$$

The dynamics could be directly reduced to $\mathcal{D}/G$ (see, e.g., Zenkov and Bloch [2003]), but we will proceed in two stages. First, we Chaplygin-compress the dynamics from TQ to $TSO(3)$ using the translation subgroup of $SE(3)$, regarding the constraint distribution as an abelian *connection* on Q with base space $S = SO(3)$ and fiber $\mathbb{R}^3$; the connection form is given by

$$\alpha_{\text{marble}} := (dx - r\rho_2, dy + r\rho_1, dz). \tag{3.18}$$

There is *another* S^1 action on Q, this time acting on the first factor only: $e^{i\phi}(R, z) = (S(\phi)R, z)$. This action preserves the Lagrangian but does *not* preserve the distribution: $D_{(S(\phi)R,z)} \neq e_*^{i\phi} D_{(R,z)}$. However, its infinitesimal action is given by the right vector field $X_3^r \in \mathcal{D}$. Noether's theorem applies, so $p_\phi = \ell_3$ is a constant of motion. Therefore, Chaplygin's marble equations can be reduced, on each level set ℓ_3, to $T(SO(3)/S^1) = TS^2$.

In the case of *Chaplygin's rubber ball*,[20] rotations about the vertical axis are forbidden (since such rotations would cause energy dissipation). Here the constraints are defined by a subdistribution $\mathcal{H} \subset \mathcal{D}$ with Cartan's 2-3-5 growth numbers and, in fact, defining a connection on $SE(2) \times \mathbb{R} \hookrightarrow Q \to S^2$ with 1-form

$$\alpha_{\text{rubber}} := (\rho_3\hat{k}, dx - r\rho_2, dy + r\rho_1, dz). \tag{3.19}$$

The extrinsic viewpoint

For clarity we present the classical, direct derivation of the equations of motion, following the "extrinsic viewpoint" advocated by the Russian Geometric Mechanics school (Borisov and Mamaev [2002a]).

• For the rubber Chaplygin ball (and Veselova's): in the space frame one has $\dot{\ell} = \tau$, where $\tau = \lambda\hat{k}$ is the torque exerted by the constraint force. The torque is vertical because $(\tau, \omega) = 0$ for all ω with third component equal to zero. Viewed in the body frame,

$$\dot{L} + \Omega \times L = \lambda\gamma, \tag{3.20}$$

Together (1.6), one gets a closed system of ODEs in the space $(L, \gamma) \in \mathbb{R}^3 \times \mathbb{R}^3$, *provided* the relation between Ω and L is obtained. In Veselova's example, $\Omega = A^{-1}L$. The multiplier can be eliminated by differentiating the constraint equation $(\Omega, \gamma) = 0$. After a simple computation, one gets

[20] This problem was not studied by Chaplygin. For the physical justification, see Neimark and Fufaev [1972] and Cendra, Ibort, de Léon, de Diego [2004]. As far as we know its integrability has not yet been established. Formally, Veselova's system is the limit of Chaplygin's rubber ball as $r \to 0$.

$$\lambda = \frac{(L, A^{-1}\gamma \times A^{-1}L)}{(\gamma, A^{-1}\gamma)}. \tag{3.21}$$

Besides the standard integrals of motion $2H = (A^{-1}L, L), (\gamma, \gamma) = 1, (A^{-1}L, \gamma) = 0$, Veselov and Veselova [1988] showed that there is a quartic polynomial integral

$$G = (L, L) - (L, \gamma)^2 \tag{3.22}$$

and an invariant measure[21]

$$\mu = f(\gamma) dL_1 \wedge dL_2 \wedge dL_3 \wedge d\gamma_1 \wedge d\gamma_2 \wedge d\gamma_3, \quad f(\gamma) = (A^{-1}\gamma, \gamma)^{-1/2}. \tag{3.23}$$

• For Chaplygin's marble: the angular momentum *at the contact point* in the space frame $\boldsymbol{\ell}$ is constant. An engineer would argue that both gravity and friction produce no torque at that point; a mathematician would use the fact that the admissible vector fields $V_i \in \mathcal{H}$ given by

$$V_1 := -r\partial/\partial y + X_1^{\text{right}}, \qquad V_2 := r\partial/\partial x + X_2^{\text{right}}, \qquad V_3 := X_3^{\text{right}} \tag{3.24}$$

preserve the Lagrangian, and would invoke NH-Noether's theorem. Whichever explanation chosen, differentiating $RL = \ell = RL$ and $R\gamma = k$, one gets Chaplygin's equations

$$\dot{L} = -\Omega \times L, \qquad \dot{\gamma} = -\Omega \times \gamma. \tag{3.25}$$

These two form a coupled system, since again Ω *is a linear function of* L *depending only on* γ:

$$L = L_\gamma(\Omega) = A\Omega + \mu r^2 \gamma \times (\Omega \times \gamma) = \tilde{A}\Omega - \mu r^2 (\gamma, \Omega)\gamma, \quad \tilde{A} := A + \mu r^2 \text{id}. \tag{3.26}$$

A simple way to get this map is to look at the total energy

$$2T = (\omega, \ell) = (\Omega, L) = (A\Omega, \Omega) + \mu(\dot{x}^2 + \dot{y}^2) = (A\Omega, \Omega) + \mu r^2(\omega_1^2 + \omega_2^2), \tag{3.27}$$

which can be also written as

$$2T = (\Omega, L) = (A\Omega, \Omega) + \mu r^2 (\Omega, \gamma \times (\Omega \times \gamma)) = (\Omega, A\Omega + \mu r^2 \gamma \times (\Omega \times \gamma))). \tag{3.28}$$

The expression $\gamma \times (\bullet \times \gamma)$ represents the projection in the plane perpendicular to γ, and we get (3.26). An *ansatz* for the inverse of the map (3.26) is (Duistermaat [2000]),

$$\Omega = \Omega(L, \gamma) = (L_\gamma)^{-1}(L) = \tilde{A}^{-1}L + \alpha(L)\tilde{A}^{-1}(\gamma), \tag{3.29}$$

and one gets the interesting expression for $\alpha(L)$ (which will be used in equation (3.48) and Proposition 3.8):

---

[21] The level sets of the four integrals are 2-tori, since there are no fixed points in the dynamics. The existence of an invariant measure in the tori allows the explicit integration via Jacobi's theorem. Veselov and Veselova [1988] found a "rather unexpected connection with Neumann's problem."

$$\alpha(L) = \mu r^2 \frac{(\gamma, \tilde{A}^{-1}L)}{1 - \mu r^2(\gamma, \tilde{A}^{-1}\gamma)}. \tag{3.30}$$

The function

$$f(\gamma) := [1 - \mu r^2(\gamma, \tilde{A}^{-1}\gamma)]^{-1/2} \tag{3.31}$$

was found by Chaplygin to be the density of an invariant measure in $\mathbb{R}^6$:

$$\nu_{\mathbb{R}^6} = f(\gamma) d\gamma_1 d\gamma_2 d\gamma_3 dL_1 dL_2 dL_3. \tag{3.32}$$

This follows from Veselova's theorem, as $F(\gamma) = 1 - \mu r^2(\gamma, \tilde{A}^{-1}\gamma)$ is (up to a constant factor) the determinant of the linear map $\Omega \mapsto L = L(\Omega; \gamma)$. For direct proofs of invariance of the measure, see Duistermaat [2000] or Fedorov and Kozlov [1995].

A system of ODE's for the rubber ball can be derived in a similar fashion. For the angular momentum ℓ at the contact point, we get the same equation (3.20) from Veselova's system, but the relation between Ω and L is (3.26), the same as in Chaplygin's marble. Differentiating $(\Omega, \gamma) = 0$ the multiplier can be eliminated.

Hamiltonization of Veselova's system

The compressed Lagrangian is

$$L_{\text{comp}} = \frac{1}{2}(A(\dot{\gamma} \times \gamma), \dot{\gamma} \times \gamma), \tag{3.33}$$

since $\Omega = \dot{\gamma} \times \gamma$; the momentum map corresponding to the S^1-action is $J = \ell_3 = (L, \gamma)$. Thus (J.K) $= \ell_3 d\rho_3 = (A\Omega, \gamma) d\rho_3$, where $d\rho_3$ is the area form of S^2. The compressed Legendre transform is

$$\dot{\gamma} \mapsto a = \frac{\partial L^*}{\partial \dot{\gamma}} = \gamma \times A(\dot{\gamma} \times \gamma).$$

The nonholonomic 2-form in T^*S^2 is

$$\Omega_{NH} = da \wedge d\gamma + (A(\dot{\gamma} \times \gamma), \gamma) d\rho_3. \tag{3.34}$$

Being a two-degrees of freedom system, a general result from Fedorov and Jovanovic [2003] (Theorem 3.5) guarantees that this system is Hamiltonizable. In order to verify that Ω_{NH} is conformally symplectic, it is simpler to use $\dot{\gamma}$ as coordinates, that is, we pull back Ω_{NH} to TS^2 via Leg*. We get

$$\Omega_{NH} = d(\gamma \times A(\dot{\gamma} \times \gamma))) \wedge d\gamma + (\gamma, A(\dot{\gamma} \times \gamma)) d\rho_3.$$

Proposition 3.5. *Veselova's system is conformally symplectic, $d(f\Omega_{NH}) = 0$, with conformal factor*

$$f = f(\gamma) = (A^{-1}\gamma, \gamma)^{-1/2}. \tag{3.35}$$

As expected, it is the density of the Veselova invariant measure $\mu = f(\gamma) dL d\gamma$ obtained via Proposition 3.3. The orthonormal frame in S^2 diagonalizing (3.33) provides explicit coordinates for integration via the Hamilton–Jacobi method.

Chaplygin's rubber ball

The dynamics compress to T^*S^2, and by the same general result in Fedorov and Jovanovic [2003], we know in advance that the system is Hamiltonizable. Choose a moving frame e_1, e_2 in S^2. The horizontal lift from $\dot{\gamma} = v_1 e_1 + v_2 e_2$ to $\mathrm{Hor}(\dot{\gamma}) \in T(SE(3))$ is easily done via (1.10):

$$\mathrm{Hor}(\dot{\gamma}) = v_2(X_1^r - r\partial/\partial y) - v_1(X_2^r + r\partial/\partial x).$$

Composing $d\alpha_{\text{rubber}} = (\rho_1 \wedge \rho 2, -r\rho_3 \wedge \rho_1, r\rho_2 \wedge \rho_3, 0)$ with Hor, we get $K_{\text{rubber}} = (dS\hat{k}, 0, 0, 0)$, where dS is the S^2 area form. Thus for the term $(J.K)$ we need only the third component of the angular momentum, $m_3 = (M, \gamma) = (A\Omega, \gamma)$, where we insert (1.10) $\Omega = \dot{\gamma} \times \gamma = v_2 e_1 - v_1 e_2$. Therefore,

$$\Omega_{NH} = \Omega_{T^*S^2} + (A(\dot{\gamma} \times \gamma), \gamma) \cdot dS. \tag{3.36}$$

Here $\dot{\gamma} = v_1 e_1 + v_2 e_2 \in TS^2$ corresponds to $p_\gamma = p_1\theta_1 + p_2\theta_2$ via the Legendre map $\mathrm{Leg}^{\mathrm{comp}}$ of the compressed Lagrangian

$$L_{\text{comp}} = \frac{1}{2} A(v_2 e_1 - v_1 e_2, v_2 e_1 - v_1 e_2) + \frac{1}{2}\mu r^2(v_1^2 + v_2^2). \tag{3.37}$$

Clearly, this system becomes Veselova for $r = 0$. Using Proposition 3.3 and Fedorov's result for two degrees of system, we get Proposition 3.6.

Proposition 3.6. *The compressed rubber ball system is Hamiltonizable. The conformal factor is*

$$\begin{aligned} f &= [\det \mathrm{Leg}^{\mathrm{comp}}]^{-1/2} \\ &= (I_1 I_2 I_3)^{-1/2}\left((A^{-1}\gamma, \gamma) + \mu r^2\left[\frac{\gamma_2^2 + \gamma_3^3}{I_2 I_3} + \frac{\gamma_1^2 + \gamma_3^3}{I_1 I_3} + \frac{\gamma_1^2 + \gamma_2^3}{I_1 I_2}\right] + \frac{\mu^2 r^4}{I_1 I_2 I_3}\right)^{-1/2}. \end{aligned} \tag{3.38}$$

Proof. We checked using spherical coordinates and Mathematica™.[22]

3.3 Chaplygin's marble is not Hamiltonizable at the $T^*SO(3)$ level

The homogeneous sphere

In a nutshell, the dynamics in the homogeneous case are embarrassingly simple. The angular velocity in space is constant, so the attitude matrix R evolves as a one-parameter group $R = \exp([\omega]t)$, so Ω and ω are constant. The vector $\gamma(t)$ describes a circle in the sphere perpendicular to ω, and $L(t)$ the curve given by $L(t) = (I + \mu r^2)\omega - \omega_3\gamma(t)$. Provided ℓ is not vertical, L and γ are never parallel. The invariant tori are always foliated by closed curves and the two frequencies coincide. From

[22] We can provide the (short) notebook under request. It should be investigated if the rubber ball problem is integrable. Does a (quartic) integral still exist?

the constraint equations we see that the motion of the contact point in the plane is a straight line. Shooting pool with a perfect Chaplygin ball is very dull.[23] *Let us use these simple results as template for our operational system.* In terms of the right coframe, we have

$$\begin{aligned}\Omega_{NH} = d\ell_1\rho_1 + d\ell_2\rho_2 + d\ell_3\rho_3 + \ell_1\rho_2\rho_3 + \ell_2\rho_3\rho_1 + \ell_3\rho_1\rho_2 \\ - \mu r^2(\omega_2\rho_3\rho_1 + \omega_1\rho_2\rho_3).\end{aligned} \tag{3.39}$$

This formula holds in general. In the nonhomogeneous case one must write ω_1 and ω_2 in terms of ℓ and $R \in SO(3)$: $\omega = R\Omega = R\Omega_\gamma(R^{-1}\ell)$ which seems to be a quite involved expression, a haunting monster we will avoid, until a final confrontation in Proposition 3.8. In the homogeneous case, life is much easier: $\omega = R\frac{1}{\kappa}IR^{-1}m = \frac{1}{I}m$, so the dependence of ω on R disappears. The Hamiltonian is given by

$$H = \frac{1}{2}\left(\frac{\ell_1^2 + \ell_2^2}{I + \mu r^2}\right) + \frac{\ell_3^2}{I},$$

where

$$\ell_1 = \left(1 + \frac{\mu r^2}{I}\right)m_1, \qquad \ell_2 = \left(1 + \frac{\mu r^2}{I}\right)m_2, \qquad \ell_3 = m_3,$$
$$\omega_1 = \frac{m_1}{I}, \qquad \omega_2 = \frac{m_2}{I}, \qquad \omega_3 = \frac{m_3}{I}.$$

To obtain the equations of motion we solve

$$\begin{pmatrix}\omega_1\\ \omega_2\\ \omega_3\\ \dot\ell_1\\ \dot\ell_2\\ \dot\ell_3\end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\\ -1 & 0 & 0 & 0 & \ell_3 & -I\omega_2\\ 0 & -1 & 0 & -\ell_3 & 0 & -I\omega_1\\ 0 & 0 & -1 & I\omega_2 & -I\omega_1 & 0\end{pmatrix} \cdot \begin{pmatrix}0\\0\\0\\ \omega_1\\ \omega_2\\ \omega_3\end{pmatrix}, \tag{3.40}$$

where we have used $H_{\ell_1} = \ell_1/(I + \mu r^2) = m_1/I = \omega_1$, and similarly, $H_{\ell_2} = \omega_2$, $H_{\ell_2} = \omega_2$. This gives, as expected

$$\dot\ell_1 = (I\omega_3)H_{\ell_2} - I\omega_2 H_{\ell_3} = 0, \qquad \dot\ell_2 = \cdots = 0, \qquad \dot\ell_3 = \cdots = 0.$$

Thus $\omega_i = m_i/I = \text{const}$, $i = 1, 2, 3$, and the vector field is simply $X = \omega_1 X_1^{\text{right}} +$

[23] We found the following relevant information in www.ot.com/skew/five/myths. html (Top Ten Myths in Pool or the Laws of Physics Do Apply): "4. If the cue is kept level, contacting the cueball purely left or right of its center will make it curve as it rolls. (*No!* The rolling cue ball can have two completely independent components to its angular momentum. Basically, this means that it can rotate in the manner of a top while rolling slowly forward along a straight line. In general, spin on a cue ball is of two types; follow/draw is the spin like tires on a car, while English is the spin like a child's toy 'top.' Separately, neither one will make a ball curve! If they are combined—e.g., strike low-left giving left English and draw—then the spin is called masse ("mass-ay''), and the ball will curve as it travels.)''

$\omega_2 X_2^{\text{right}} + \omega_3 X_3^{\text{right}}$ (no components in the fiber directions $\partial/\partial m_i$). We now use Theorem 3.4. Using m as coordinates, the nonholonomic 2-form is given by

$$\Omega_{NH} = \left(1 + \frac{\mu r^2}{I}\right)(dm_1\rho_1 + dm_2\rho_2) \\ + dm_3\rho_3 + (m_1\rho_2\rho_3 + m_2\rho_3\rho_1 + m_3\rho_1\rho_2)$$

so that $d\Omega_{NH} = -\frac{\mu r^2}{I}(dm_1\rho_2\rho_3 + dm_2\rho_3\rho_1)$. It is easy to see that the equation $d\Omega_{NH} = \Omega_{NH} \wedge \alpha$ has no solution. Indeed, suppose $\alpha = A_1 dm_1 + A_2 dm_2 + A_3 dm_3 + B_1\rho_1 + B_2\rho_2 + B_3\rho_3$. Taking the exterior product, and looking at terms like $dm_1 dm_2 \rho_2$ we see that all the A's must be zero. Examining the coefficient of $\rho_1\rho_2\rho_3$ we get $B_1 m_1 + B_2 m_2 + B_3 m_3 \equiv 0$ so all the B's are also zero.

Hence the homogeneous Chaplygin sphere, as simple at it can be, has no conformal symplectic structure! In fact, it does not have an affine symplectic structure either. A short calculation shows that

$$i_X d\Omega_{NH} = \frac{\mu r^2}{I^2}(-dm_1 m_2\rho_3 + dm_1\rho_2 m_3 - dm_2 m_3\rho_1 + dm_2\rho_3 m_1) \neq 0.$$

By continuity, for sufficiently close but different inertia coefficients the inequalities persist. We have also done the calculation for the nonhomogeneous case and things only get worse. But, it still remains a possibility: is the *reduced* system to T^*S^2 Hamiltonizable? The impatient reader can go directly to Theorem 3.8.

3.4 Chaplygin's marble: Reduction to T^*S^2

Using (1.8), $L = a \times \gamma + \ell_3\gamma$, Chaplygin's marble equations in (L, γ)-space directly reduce to T^*S^2:

$$\dot{\gamma} = \gamma \times \Omega, \qquad \dot{a} = -2H\gamma + (\gamma, \Omega)\cdot(a \times \gamma + \ell_3\gamma) \tag{3.41}$$

with

$$\Omega = \Omega(a, \gamma; \ell_3) = \tilde{A}^{-1}L + \mu r^2 \frac{(\gamma, \tilde{A}^{-1}L)}{1 - \mu r^2(\gamma, \tilde{A}^{-1}\gamma)}\tilde{A}^{-1}(\gamma).$$

S^1 reduction of the homogeneous sphere to T^*S^2

The homogeneous Chaplygin sphere when reduced to T^*S^2 produces a more interesting system. Equations (3.41) become

$$\dot{\gamma} = \frac{1}{I + \mu r^2}a, \qquad \dot{a} = \omega_3 a \times \gamma - \frac{1}{I + \mu r^2}|a|^2\gamma. \tag{3.42}$$

One observes that $(a, \gamma) = 0$ and that $|a|^2$ is conserved. So at each level set, we get an isotropic 3D oscillator with a Lorentz force.[24]

[24] Alan Weinstein commented on more than one occasion that "unreduction" sometimes is even nicer than reduction: unreducing a nontrivial system may lead to a trivial one. Alan credits this to Guillemin and Sternberg; one reference could be Guillemin and Sternberg [1980].

Dimension count argument

In hindsight, we can give two simple arguments why the Chaplygin marble could not be Hamiltonizable at the $T^*SO(3)$ level. First, if $(T^*SO(3), \Omega_{NH}, H)$ were Hamiltonizable, the system would be Liouville integrable *by "mere" symmetries*, due to the existence of three independent first integrals H, ℓ_3, $\ell_1^2 + \ell_2^2$, ℓ_3^2. But it is known that integrability of Chaplygin's marble stems not from symmetries, but from a special choice of separating coordinates (Duistermaat [2000]), namely, elliptic coordinates on the sphere. Second, Stanchenko [1985] verified that Chaplygin's density function F (3.31) of the system in $\mathbb{R}^6$ also gives an invariant measure on $T^*SO(3)$ (see also Duistermaat [2000], Section 7),

$$\nu_{T^*SO(3)} = F(\gamma) d\lambda_1 d\lambda_2 d\lambda_3 dL_1 dL_2 dL_3, \quad F = [1 - \mu r^2(\gamma, \tilde{A}^{-1}\gamma)]^{-1/2}. \tag{3.43}$$

Were the compressed system Hamiltonizable in $T^*SO(3)$, the conformal factor (time reparametrization) would be $F(\gamma)^{\frac{1}{m-1}}$, with $m = 3$; see (3.13). But the correct time reparametrization holds with $m = 2$ instead of $m = 3$. This strongly suggests that Hamiltonization should be attempted *after* reduction of the internal S^1 symmetry.

Phase locking

The fact that Chaplygin's sphere is integrable implies an interesting *phase locking* property. For simplicity, consider a resonant torus and a periodic solution, $\gamma(T) = \gamma(0)$, $L(T) = L(0)$. We may assume that $R(0) =$ identity, so $R(T)$ preserves both k and ℓ. If we assume $\ell \neq \pm k$, then $R(T)$ must also be the identity (there is only one orthogonal matrix with two different eigenvectors with equal eigenvalues 1). Since the rotational conditions are reproduced after time T, there is a "planar geometric phase" (meaning a translation), $\Delta z = (\Delta x, \Delta y)$. From Duistermaat [2000], Section 11, one knows this direction.

Proposition 3.7. *On average, Δz moves in the direction of $\ell \times k$.*

In the normal direction $k \times (\ell \times k)$ there is a "swaying motion," with zero average, see Duistermaat [2000], (11.71), and Remark 11.11. This result depends on the explicit solution in terms of elliptic coordinates, but the zero average can be proved in a more elementary way, see Duistermaat, Section 8.2. In the direction $\ell \times k$ one has

$$\frac{d}{dt}(z(t), \ell \times k) = r(\omega \times k, \ell \times k) = r(\omega, \ell - \ell_3 k) = r(2T - \ell_3\omega_3) > 0.$$

Duistermaat [2000] shows (in Section 9.2) that by a suitable change of coordinates, one may assume that $\ell_3 = 0$, so in this equivalent problem, the velocity in this direction is simply $2rT$.

Chaplygin's marble via the almost Hamiltonian structure

After this detour, we hope the reader will appreciate a concise way of describing this system. The clockwise map is

$$
\begin{array}{ccccc}
T(SO(3)\times\mathbb{R}^2) & \underset{\text{Leg}}{\longrightarrow} & T^*(SO(3)\times\mathbb{R}^2) & (\Omega,\dot x,\dot y) & \mapsto (M=A\Omega,\, P_x=\mu\dot x,\, P_y=\mu\dot y)\\
\uparrow & & & \uparrow & \\
h & & & h & \\
| & & & | & \\
TSO(3) & \underset{(\text{Leg}^{\phi})^{-1}}{\longleftarrow} & T^*SO(3) & \Omega \quad \leftarrow & L,
\end{array}
\tag{3.44}
$$

where $(\dot x, \dot y) = r\omega \times k$, and $(P_x, P_y) = \mu r\omega \times k$. We now compute the "gyroscopic" 2-form

$$
(\text{J.K}) = r(-P_x d\rho_2 + P_y d\rho_1) = \mu r(-\dot x d\rho_2 + \dot y d\rho_1) = -\mu r^2(\omega_2 d\rho_2 + \omega_1 d\rho_1).
\tag{3.45}
$$

To obtain ω_1 and ω_2 as functions in $T^*SO(3)$, we use the Legendre transformation: $\omega = R\Omega = RA^{-1}M$, so (J.K) is a combination of the basic forms $\rho_3 \wedge \rho_1, \rho_2 \wedge \rho_3$ (coefficients linear in M and functions of R).

S^1 invariance

We claim that Ω_{NH} is S^1-invariant (S^1 acting only in the first factor of $Q = SO(3) \times \mathbb{R}^3$). For the canonical term this is a standard symplectic fact. The $(J.K)$ term is invariant as well: (3.45), written in terms of the left-invariant forms, depends only on the Poisson vector γ:

$$
(\text{J.K}) = \mu r^2\,(\gamma \times (\Omega(L,\gamma)\times\gamma), d\lambda)\,.
\tag{3.46}
$$

In fact, the S^1 action generated by the right-invariant vector field X_3^{right} maintains the projection γ fixed. We know (general nonsense) that the right-invariant vector fields preserve the left-invariant forms: $R_\phi^*\lambda_i = \lambda_i$. (Proof: $(R_\phi^*\lambda_i)(\dot R) = \lambda_i(R_\phi R[\Omega]) = \Omega_i$.) Since under the left S^1 action (actually under the left action of $SO(3)$ on $SO(3)$) the *value* of Ω remains unchanged, the (J.K) term is preserved.

The twisted action generator and S^1 reduction

MW reduction method works fine, although X_3^r is the Hamiltonian vector field of $J = \ell_3$, relative to the canonical symplectic form, but not relative to Ω_{NH}. We just change to the *twisted* S^1-action generator $\tilde X_3$, defined by $i_{\tilde X_3}\Omega_{NH} = -d\ell_3$. A simple computation gives $\tilde X_3 = X_3^r - m_2\frac{\partial}{\partial \ell_1} + m_1\frac{\partial}{\partial \ell_2}$, where $m_1 = \ell_1 - \mu r^2\omega_1$, $m_2 = \ell_2 - \mu r^2\omega_2$. The reduced manifold is the quotient of a level ℓ_3 in $T^*SO(3)$, identifying the flow lines $\tilde\phi$. A concrete realization is achieved using (1.11). Taking the pullback via i^*, the reduced form is then

$$
\Omega_{\text{red}} = \Omega^{\text{can}}_{T^*S^2} + \ell_3\text{area}_{S^2} - \mu r^2(\omega_1 d\theta_2 - \omega_2 d\theta_1),
\tag{3.47}
$$

where we recall the parametrizations $p_1\theta_1 + p_2\theta_2 \in T^*S^2$, $R(\gamma) = \text{rows}(e_1, e_2, \gamma)$.

In (3.47) we must write ω_1, ω_2 explicitly in terms of p_1, p_2, ℓ_3. To write this explicitly, there is no other option than to confront the monster (which actually is not that terrible): from $\Omega = R^{-1}\omega = \omega_1 e_1 + \omega_2 e_2 + \omega_3 \gamma$ and (3.26), we get

$$\omega = R(\gamma)\Omega_\gamma [R(\gamma)]^\dagger \ell, \quad \ell = (p_2, -p_1, \ell_3), \tag{3.48}$$

where Ω_γ is explicitly given by (3.30).

Theorem 3.8. $i_X d(f\Omega_{\rm red}) \neq 0,\ f(\gamma) = [1 - \mu r^2(\gamma, \tilde{A}^{-1}\gamma)]^{-1/2}$.

Proof. We used spherical coordinates (*fâute de mieux*) and a Mathematica™ notebook. It misses being conformally symplectic by very little (even in the homogeneous case).[25]

Our calculation shows that Chaplygin's sphere is *not* affine symplectic even at the T^*S^2 level, so Chaplygin's sphere integrability is due to a specific nonholonomic phenomenon. This observation is in accordance with the opening statement in Duistermaat [2000]:

> "Although the system is integrable in every sense of the word, it neither arises as a Hamiltonian system, nor is the integrability an immediate consequence of the symmetries."

4 Recent developments and final comments

NH systems have a reputation of having peculiar (even rebellious) dynamic behavior (Arnold, Kozlov, and Neishtadt [1988]). In spite of good progress, the general theory for NH systems is way behind the theory of Hamiltonian systems. For instance, although the groundwork for a Hamilton–Jacobi theory for NH systems has been set up in Weber [1986], not much has been achieved since.

We have no intention (or competence) to make a survey of recent developments in NH systems, especially regarding reduction of symmetries; nevertheless it may be worth registering the intense activity going on. Recent books of interest are Cushman and Bates [1997], Cortés [2002], Oliva [2002], Bloch [2003] and a treatise in the mechanical engineering tradition is Papastavridis [2002].[26] *Reports on Mathematical Physics* has been publishing NH papers regularly, and *Regular and Chaotic Dynamics* devoted large parts of vols. 1/2 (2002) to NH systems. For older eastern European literature, see *P. M. M. USSR, J. Appl. Math. Mechanics*, which has strongly influenced Chinese mechanics as well. For a historical account of NH systems, from a somewhat "antireductionist" perspective, see Borisov and Mamaev [2002a].

[25] Borisov and Mamaev [2002] showed by a subtle numerical evidence that, in the original time, Chaplygin's marble is not Hamiltonizable at any level of reduction. The question whether Chaplygin's marble is Hamiltonizable in the new time $dt/d\tau = f(\gamma)$ was addressed in Borisov and Mamaev [2001]. They provide a bracket structure in terms of the coordinates (L, γ) or the coordinates $(\tilde{L}, \gamma)$, with $\tilde{L} = L/f(\gamma)$. Using a computer algebra program we checked that the second brackets satisfy the Jacobi identity. However, we could not recover Chaplygin's equations for the L coordinates, even in the homogenous case.

[26] Reviewed in Koiller [2003].

4.1 Invariant measures and integrability

Kupka and Oliva [2001] and Kobayashi and Oliva [2003] find conditions ensuring a special, but very interesting situation, where the Riemann measure in TQ induced by the metric in Q is an invariant measure for the NH system. Invariant measures for systems with distributional symmetries were characterized in Zenkov and Bloch [2003].

Curiously, although a number of interesting NH systems have been solved using Abelian functions, a precise definition for integrability of a NH system is still lacking (Bates and Cushman [1999]). These examples suggest that the presence of an invariant measure must be imposed as a necessary (although not sufficient) condition for integrability (whatever it may be), see Kozlov [2002]. Most of them have enough integrals of motion that the dynamics occur on invariant two-dimensional tori. Due to the invariant measure, *the flow becomes linear in these tori after a time rescaling*. This follows from Jacobi's multiplier method and Kolmogorov's theorem (Arnold [1989]). Time reparametrization indicates the possibility of an affine symplectic structure. We believe that characterizing NH systems possessing an affine symplectic structure (if needed, after some reduction stage) could be an interesting project. As a first step, one may examine the existing literature to see which examples fit. We list a few papers for that purpose: Veselov and Veselova [1988], Veselov and Veselova [1986], Fedorov [1989], Cushman, Hermans, and Kemppainen [1995], Zenkov [1995], Zenkov and Bloch [2000], Dragovic, Gajić, and Jovanovic [1998], Jovanovic [2003], Fedorov and Jovanovic [2003]. One can hope that the manifestly geometric character of (1.14) can be instrumental to understand when, where and why *Hamiltonization* is possible. Moreover, a prior geometric understanding of the invariant volume form conditions is a more general question. It would be also interesting to tie the "Hamiltonizable" question with the invariants from the Cartan equivalence viewpoint, see below.

4.2 Nonholonomic reduction

The difficulties in reduction for general NH systems are explained in Śniatycki [2002]. There are four current theories of reduction of symmety for nonholonomic systems[27]:

(i) *projection methods*; see Marle [1995], Dazord [1994];
(ii) *the distributional Hamiltonian approach*, initiated by Bocharov and Vinogradov [1977] and developed in Bates and Śniatycki [1993], Cushman, Kemppainen, Śniatycki, and Bates [1995], Śniatycki [1998], and Cushman and Śniatycki [2002].
(iii) *bracket methods*, initiated by Mashke and van der Schaft [1994], and developed by Koon and Marsden [1998] and Sńiatycki [2001];
(iv) *Lagrangian reduction*; see Cendra, Marsden, and Ratiu [2001].

A few other references in this rapidly developing theme, besides those already mentioned are Bates [2002], Koon and Marsden [1997], Cantrijn, de Léon, Marrero, and de Diego [1998], Cortés and de Léon [1999], Marle [1998], Marle [2003].

[27] We thank one of the referees for this information.

Almost Poisson, almost Dirac approaches

Mashke and van der Schaft [1994] were the first to describe a NH system using an almost-Poisson structure,[28] $\dot{x}_i = \{x_i, H\}_{MS}$. This bracket, defined on the manifold $P = \mathrm{Leg}(\mathcal{H}) \subset T^*Q$, where $\mathrm{Leg} : TQ \to T^*Q$ is the Legendre transformation, in general does not satisfy the Jacobi identity. They proved that the Jacobi identity holds if and only if the constraints are integrable. In Koiller, Rios, and Ehlers [2002] we gave a moving frames based derivation of the bracket structure. For some recent work on the MS-bracket and also Dirac estructures (the latter introduced in Courant [1990]), see Cantrijn, de Léon, de Diego [1999], Koon and Marsden [1998], Ibort, de Léon, Marrero, and de Diego [1999], Clemente-Gallardo, Maschke, and van der Schaft [2001]. In spite of these advances, a complete understanding of the NH bracket geometry is still in order.[29]

4.3 *G*-Chaplygin systems via affine connections

Trajectories of the compressed system can be described as geodesics of an affine connection ∇^{NH} in S (Vershik and Fadeev [1981], Koiller [1992]). For background in this approach, see Lewis [1998] and references therein. Consider the parallel transport operator along closed curves; if the holonomy group is always conjugate to a subgroup of $S0(m)$, then the connection is metrizable. This means that there is a metric such that ∇^{NH} is precisely the Levi-Civita connection of this metric. More generally, one may want to know when the geodesics of ∇^{NH} are, up to time reparametrization, the geodesics of a Riemannian metric. This is a traditional area in differential geometry, whose roots go back to the 19th century, and goes under the name of *projectively equivalent connections* (Cartan [1937], Eisenhart [1925], Kobayashi and Nomizu [1963], Sharpe [1997]). Grossman [2000] studies integrability of geodesics equations via the equivalence method. Our problem, then, is to find conditions for the NH connection to be projectively equivalent to a Riemannian connection. It would be also interesting to tie the Hamiltonization question with the canonical system and invariants of the Cartan equivalence method. When an internal symmetry group is present, it would be desirable to construct a projected connection in S for each set of conserved momenta, and address these issues in the reduced level.

Acknowledgments Our thanks to Hans Duistermaat for information on Chaplygin's sphere and the referees for very good criticism and suggestions. This is also a special occasion to thank our mathematical family: Alan, of course; mathematical brothers and sisters from all continents (especially Yilmaz Akyildyz and Henrique Bursztyn) and cousins (especially Tudor Ratiu and

[28] Physicists are never shy to use the word "super'' in their endeavors; on the other hand we, mathematicians, prefer to use low-key terminology, like "almost-quasi-twisted-(freakaz-)-'oid's''; this certainly does not help our image problem with applied people, see Papastavridis [2002] and Koiller [2003].

[29] Observations by J. Marsden (joint work with H. Yoshimura), and by C. Marle in their Alanfest talks are important steps in this direction.

Debra Lewis), "uncle" Jerry Marsden, and, last but not least, "grandfather" Chern (with his gentle voice, commanding us to keep interested in Math, up to his age).

We shall toast in many more Alanfests: *Madadayo!*

References

Abraham, R., Marsden, J. [1994], *Foundations of Mechanics*, 2nd revised ed., Perseus Publishing, Boulder, CO.

Arnold, V. I. [1989], *Mathematical Methods of Classical Mechanics*, 2nd ed., Graduate Texts in Mathematics, Vol. 60, Springer-Verlag, New York; 1st ed., 1978.

Arnold, V. I., Kozlov, V. V., Neishtadt, A. I. [1988], Mathematical aspects of classical and celestial mechanics, in Arnold, V. I., ed., *Dynamical Systems* III, Springer-Verlag, New York.

Bates, L. [2002], Problems and progress in nonholonomic reduction, *Rep. Math. Phys.*, **49**:2–3, 143–149.

Bates, L., Cushman, R. [1999], What is a completely integrable nonholonomic dynamical system?, *Rep. Math. Phys.*, **44**:1–2, 29–35.

Bates, L., Śniatycki, J. [1993], Nonholonomic reduction, *Rep. Math. Phys.*, **32**:1, 99–115.

Blackall, C. J. [1941], On volume integral invariants of non-holonomic dynamical systems, *Amer. J. Math.*, **63**:1, 155–168.

Bloch, A. M. [2003], *Nonholonomic Mechanics and Control*, Springer-Verlag, New York.

Bloch, A. M., Krishnaprasad, P. S., Marsden, J. E., Murray, R. M. [1996], Nonholonomic mechanical systems with symmetry, *Arch. Rational Mech. Anal.*, **136**, 21–99.

Bocharov, A. V., Vinogradov, A. M. [1977], The Hamiltonian form of mechanics with friction, nonholonomic mechanics, invariant mechanics, the theory of refraction and impact, Appendix II in Kuperschmidt, B. A. and Vinogradov, A. M., "The structure of Hamiltonian mechanics," *Russian Math. Surveys*, **42**:2, 177–243.

Borisov, A. V., Mamaev, I. S. [2001], Chaplygin's ball rolling problem is Hamiltonian, *Math. Notes* (*Mat. Z.*), **70**:5, 793–795.

Borisov, A. V., Mamaev, I. S. [2002], Obstacle to the reduction of nonholonomic systems to the Hamiltonian form, *Dokl. Phys. USSR*, **47**:12, 892–894.

Borisov, A. V., Mamaev, I. S. [2002a], On the history of the development of the nonholonomic dynamics, *Regular Chaotic Dynam.*, **7**:1, 43–47.

Borisov, A. V., Mamaev, I. S. [2002a], The rolling body motion of a rigid body on a plane and a sphere: Hierarchy of dynamics, *Regular Chaotic Dynam.*, **7**:2, 177–200.

Borisov A. V., Mamaev I. S., Kilin A. A. [2002], The rolling motion of a ball on a surface: New integrals and hierarchy of dynamics, *Regular Chaotic Dynam.*, **7**:2, 201–218.

Bryant, R. [1994], *Lectures on the Geometry of Differential Equations*, Fall 1994 André Aisenstadt Lectures, Centre de Recherche Mathématique, Montreal, in preparation; unpublished lecture notes available at http://www.cimat.mx/gil.

Cantrijn, F., de Léon, M., Marrero, J. C., de Diego, D. [1998], Reduction of nonholonomic mechanical systems with symmetries, *Rep. Math. Phys.*, **42**:1–2, 25–45.

Cantrijn, F., de Léon, M., de Diego, D. [1999], On almost-Poisson structures in nonholonomic mechanics, *Nonlinearity*, **12**, 721–737.

Cantrijn, F., Cortés, J., de Léon, M., de Diego, D. [2002], On the geometry of generalized Chaplygin systems, *Math. Proc. Cambridge Philos. Soc.*, **132**, 323–351.

Cartan, É. [1910], Le système de Pfaff à cinq variables et les équations aux dérivées partielles du second ordre, *Ann. Sci. École Norm.*, **27**:3, 109–192.

Cartan, É. [1926], *La théorie des groupes finis et continus et la géométrie différentielle, traitées par la méthode du repère mobile: Leçons professées à la Sorbonne* (rédigées par Jean Leray), Gauthier-Villars, Paris, 1951.

Cartan, É. [1928], Sur la represéntation géométrique des systèmes matériels non holonomes, *Proc. Internat. Congr. Math. Bologna*, **4**, 253–261.

Cartan, É [1937], *Leçons sur la théorie des espaces a connexion projective*, Gauthier-Villars, Paris.

Cartan, É. [2001], *Riemannian Geometry in an Orthogonal Frame: From Lectures Delivered by Élie Cartan at the Sorbonne in* 1926–27, World Scientific, River Edge, NJ, 2001 (translated from the 1960 Russian edition by Vladislav V. Goldberg; foreword by S. S. Chern; preface to the Russian edition by S. P. Finikov).

Cendra, H., Lacomba, E. A., Reartes, W. [2001], The Lagrange-d'Alembert-Poincaré equations for the symmetric rolling sphere, in *Proceedings of the* 6*th Dr. Antonio A. R. Monteiro Congress of Mathematics* (*Bahía Blanca,* 2001), 19–32, Departamento de Matemática, Instituto de Matemática, Universidad Nacional del Sur, Bahía Blanca, Argentina; MR1919465 (2003g:70021).

Cendra, H., Marsden, J. E., Ratiu, T. S. [2001], Geometric mechanics, Lagrangian reduction and nonholonomic systems, in Enguist, B., and Schmid, W., eds., *Mathematics Unlimited:* 2001 *and Beyond*, Springer-Verlag, New York, 221–273.

Cendra, H., Ibort, A., de Léon, M., de Diego, D. [2004] A generalization of Chetaev's principle for a class of higher order non-holonomic constraints, *J. Math. Phys.*, submitted.

Chaplygin, S. A. [1911], On the theory of the motion of nonholonomic systems: Theorem on the reducing factor, *Mat. Sb.*, **28**, 303–314.

Chaplygin, S. A. [1981], *Selected Works on Mechanics and Mathematics*, Nauka, Moscow.

Chaplygin, S. A. [2002], On a ball's rolling on a horizontal plane, *Regular Chaotic Dynam.*, **7**:2, 131–148; original paper in *Math. Sb.*, **24** (1903), 139–168.

Bryant, R. L., Chern, S. S., Gardner, R. B., Goldschmidt, H. L., and Griffiths, P. A. [1991], *Exterior Differential Systems*, Mathematical Sciences Research Institute Publications, Vol. 18, Springer-Verlag, New York.

Clemente-Gallardo, J., Mashke, B., van der Schaft, A. J. [2001], Kinematical constraints and algebroids, *Rep. Math. Phys.*, **47**:3, 413–429.

Cortés, J. [2002], *Geometric, Control and Numerical Aspects of Nonholonomic Systems*, Springer-Verlag, New York.

Cortés, J., de Léon, M. [1999], Reduction and reconstruction of the dynamics of nonholonomic systems, *J. Phys.* A *Math. Gen.*, **32**, 8615–8645.

Cortés, J., de Léon, M., de Diego, D., Martínez, S. [2003], Geometric description of vakonomic and nonholonomic dynamics: Comparison of solutions, *SIAM J. Control Optim.*, **41**:5, 1389–1412.

Courant, T. J. [1990], Dirac manifolds, *Trans. Amer. Math. Soc.*, **319**:2, 631–661.

Cushman, R., Hermans, J., Kemppainen, D. [1995], The rolling disk, in *Nonlinear Dynamical Systems and Chaos*, Progress in Nonlinear Differential Equations and Their Applications, Vol. 19, 21–60, Birkhäuser, Basel.

Cushman, R., Kemppainen, D., Śniatycki, J., Bates, L. [1995], Geometry of nonholonomic constraints, *Rep. Math. Phys.*, **36**:2–3, 275–286.

Cushman, R., Bates, L. [1997], *Global Aspects of Classical Integrable Systems*, Birkhäuser, Basel, 1997.

Cushman, R., Śniatycki, J. [2002], Nonholonomic reduction for free and proper actions, *Regular Chaotic Dynam.*, **7**, 61–72.

Dazord, P. [1994], Mécanique hamiltonienne en présence de contraintes, *Illinois J. Math.*, **38**, 148–175.

Dragović, V., Gajić, B., Jovanović, B. [1998], Generalizations of classical integrable nonholonomic rigid body systems, *J. Phys.* A, **31**:49, 9861–9869.

Duistermaat, J. J. [2000], Chaplygin's sphere, in Cushman, R., Duistermaat, J. J., and Śniatycki, J., *Chaplygin and the Geometry of Nonholonomically Constrained Systems*, in preparation; arXiv.org/abs/math.DS/0409019.

Eisenhart, L. P. [1925], *Riemannian Geometry*, Princeton University Press, Princeton, NJ; 5th printing, 1964.

Ehlers, K. [2002], The geometry of nonholonomic three-manifolds, in *Proceedings of the* 4*th International Conference on Dynamical Systems and Differential Equations*, University of North Carolina at Wilmington, Wilmington, NC.

Fedorov, Yu. N. [1989], Two integrable nonholonomic systems in classical dynamics, *Vestnik Moskov. Univ. Ser.* I *Mat. Mekh.*, **4**, 38–41 (in Russian).

Fedorov, Yu. N., Kozlov, V. V. [1995], Various aspects of n-dimensional rigid body dynamics, in Kozlov, V. V., ed., *Dynamical Systems in Classical Mechanics*, AMS Translations (Series 2), Vol. 168, AMS, Providence, RI.

Fedorov, Yu. N., Jovanovic, B. [2003], Nonholonomic LR systems as generalized Chaplygin systems with an invariant measure and geodesic flows on homogeneous spaces, *J. Nonlinear Sci.*, to appear; arXiv: math-ph/0307016.

Flanders, H. [1963], *Differential Forms with Applications to the Physical Sciences*, Academic Press, New York.

Gardner, R. [1989], *The Method of Equivalence and its Applications*, SIAM, Philadelphia.

Grossman, D. A. [2000], Torsion-free path geometries and integrable second order ODE systems, *Selecta Math.* (*N.S.*), **6**, 399–442.

Guillemin, V., Sternberg, S. [1980], The moment map and collective motion, *Ann. Phys.*, **1278**, 220–253.

Hamel, G. [1949], *Theoretische Mechanik: Eine einheitliche Einführung in die gesamte Mechanik*, revised ed., Grundlehren der Mathematischen Wissenschaften, Vol. 57, Springer-Verlag, Berlin, New York, 1978.

Haller, S., Rybicki, T. [1999], On the group of diffeomorphisms preserving a locally conformal symplectic structure, *Ann. Global Anal. Geom.*, **17**, 475–502.

Haller, S., Rybicki, T. [2001], Symplectic reduction for locally conformal symplectic manifolds, *J. Geom. Phys.*, **37**, 262–271.

Hertz, H. [1899], *The Principles of Mechanics Presented in a New Form by Heinrich Hertz, with an Introduction by H. von Helmholtz*, Macmillan, London, New York.

Hicks, N. J. [1965], *Notes on Differential Geometry*, Van Nostrand, New York.

Hughen, K. [1995], *The Geometry of Sub-Riemannian Three-Manifolds*, Ph.D. thesis, Duke University, Durham, NC.

Ibort, A., de Léon, M., Marrero, J. C., de Diego, D. [1999], Dirac brackets in constraind dynamics, *Fortschr. Phys.*, **47**:5, 459–492.

Iliyev, Il. [1985], On the conditions for the existence of the reducing Chaplygin factor, *P. M. M. USSR*, **49**:2, 295–301.

Jovanovic, B. [2003], Some multidimensional integrable cases of nonholonomic rigid body dynamics, *Regular Chaotic Dynam.*, **8**:1, 125–132.

Kazarian, M. R., Montgomery, R., Shapiro, B. [1997], Characteristic classes for the degenerations of two-plane fields in four dimensions, *Pacific J. Math.*, **179**:2, 355–370.

Kobayashi, S., Nomizu, K. [1963], *Foundations of Differential Geometry*, Vol. 1, Wiley, New York, London.

Kobayashi, M. H., Oliva, W. M. [2003], A note on the conservation of energy and volume in the setting of nonholonomic mechanical systems, *Qual. Theory Dynam. Systems*, **5** (2004), 255–283.

Koiller, J. [1992], Reduction of some classical non-holonomic systems with symmetry, *Arch. Rational Mech. Anal.*, **118**, 113–148.

Koiller, J., Rodrigues, P. R., Pitanga, P. [2001], Nonholonomic connections following Élie Cartan, *An. Acad. Brasileira Ciencias*, **73**:2, 165–190.

Koiller, J., Rios, P. M. [2001], Nonholonomic systems with symmetry allowing a conformally symplectic reduction, in Delgado, J., Lacomba, E. A., and Pérez-Chavela, E., eds., *Proceedings of the* 4*th International Symposium on Hamiltonian Systems and Celestial Mechanics* (*Mexico* 2001), Kluwer, Dordrecht, the Netherlands, to appear; arXiv: math-ph/0203013.

Koiller, J., Rios, P. M., Ehlers, K. [2002], Moving frames for cotangent bundles, *Rep. Math. Phys.*, **49**:2–3, 225–238.

Koiller, J. [2003], Analytical Mechanics, *Bull. Amer. Math. Soc.* (*N.S.*), **40**:3, 405–419.

Koon, W. S., Marsden, J. E. [1997], The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems, *Rep. Math. Phys.*, **40**, 21–62.

Koon, W. S., Marsden, J. E. [1998], Poisson reduction for nonholonomic mechanical systems with symmetry, *Rep. Math. Phys.*, **42**, 101–134.

Kozlov, V. V. [2002], On the integration theory of equations of nonholonomic mechanics, *Regular Chaotic Dynam.*, **7**:2, 161–176.

Kummer, M. [1981], On the construction of the reduced phase space of a Hamiltonian system with symmetry, *Indiana Univ. Math. J.*, **30**, 281–291.

Kupka, I., Oliva, W. M. [2001], The Non-Holonomic Mechanics, *J. Differential Equations*, **169**, 169–189.

Levi, M. [1996], Composition of rotations and parallel transport, *Nonlinearity*, **9**, 413–419.

Lewis, A. D. [1998], Affine connections and distributions with applications to nonholonomic mechanics, *Rep. Math. Phys.*, **42**:1–2, 135–164.

Marle, C. M. [1995], Reduction of constrained mechanical systems and stability of relative equilibria, *Comm. Math. Phys.*, **174**, 295–318.

Marle, C. M. [1998], Various approaches to conservative and nonconservative nonholonomic systems, *Rep. Math. Phys.*, **42**:1–2, 211–229.

Marle, C. M. [2003], On symmetries and constants of motion in Hamiltonian systems with nonholonomic constraints, *Banach Center Publ.*, **59**, 223–242.

Marsden, J. E., Weinstein, A. [2001], Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.*, **5**, 121–130.

Marsden, J. E., Weinstein, A. [2001], Some comments on the history, theory, and applications of symplectic reduction, in Landsman, N. P., and Pflaum, M., eds., *Quantization of Singular Symplectic Quotients*, Birkhäuser, Boston, 2001.

Montgomery, R. [1991], How much does the rigid body rotate? A Berry's phase from the 18th century, *Amer. J. Phys.*, **59**:5, 394–398.

Montgomery, R. [2002], *A Tour of Sub-Riemannian Geometries, Their Geodesics, and Applications*, Mathematical Surveys and Monographs, Vol. 91, AMS, Providence, RI.

Moseley, K. [2001], *The Geometry of Sub-Riemannian Engel Manifolds*, Ph.D. thesis, Duke University, Durham, NC, 2001.

Neimark, J. I., Fufaev, N. A. [1972], *Dynamics of Nonholonomic Systems*, Translations of Mathematical Monographs, Vol. 33, AMS, Providence, RI.

Oliva, W. M. [2002], *Geometric Mechanics*, Lecture Notes in Mathematics 1798, Springer-Verlag, New York.

Papastavridis, J. G. [2002], *Analytical Mechanics: A Comprehensive Treatise on the Dynamics of Constrained Systems: For Engineers, Physicists and Mathematicians*, Oxford University Press, Oxford, UK (reviewed in Koiller [2003]).

Pars, L. A. [1965], *A Treatise on Analytical Dynamics*, Wiley, New York.

Poincaré, H. [1901], Sur une forme nouvelle des équations de la mécanique, *C. R. Acad. Sci. Paris*, **132**, 369–371.

Ramos, A. [2004], *Poisson Structures for Reduced Non-Holonomic Systems*, *J. Phys.* A *Math. Gen.*, **37** (2004), 4821–4842.

Salamon, S. [1989], *Riemannian Geometry and Holonomy Groups*, Pitman Research Notes in Mathematics 201, Longman Scientific and Technical, Essex, UK.

Schneider, D., Nonholonomic Euler-Poincare equations and stability in Chaplygin's sphere, *Dynam. Systems Intern. J.*, **17**:2, 87–130.

Sharpe, R. W. [1997], *Differential Geometry: Cartan's Generalization of Klein's Erlangen Program*, Graduate Texts in Mathematics, Vol. 166, Springer-Verlag, New York, 1997.

Śniatycli, J. [1998], Nonholonomic Noether theorem and reduction of symmetries, *Rep. Math. Phys.*, **42**:1–2, 5–23.

Sńiatycki, J. [2001], Almost Poisson spaces and nonholonomic singular reduction, *Rep. Math. Phys.*, **48**:1–2, 235–248.

Śniatycki, J. [2002], The momentum equation and the second order differential condition, *Rep. Math. Phys.*, **213**, 371–394.

Sommerfeld, A. [1952], *Lectures on Theoretical Physics: Mechanics*, Academic Press, New York.

Stanchenko, S. V. [1985], Non-holonomic Chaplygin systems, *P. M. M. USSR*, **53**:1, 11–17.

Tavares, J. N. [2002], *About Cartan Geometrization of Non-Holonomic Mechanics*, preprint, Centro de Matemática, Universidade do Porto, Porto, Portugal; available at www.fc.up.pt/cmup.

van der Schaft, A., Mashke, B. M. [1994], On the hamiltonian formulation of nonholonomic mechanical systems, *Rep. Math. Phys.*, **34**, 225–233.

Vaisman, I., Locally conformal symplectic manifolds, *Internat. J. Math. Math. Sci.*, **8**:3, 521–536.

Vershik, A. M., Gerhskovich, V. [1994], Nonholonomic dynamical systems, geometry of distributions and variational problems, in Arnol'd, V. I., and Novikov, S. P., eds., *Dynamical Systems* VII, Encyclopedia of Mathematical Sciences, Vol. 16, Springer-Verlag, New York.

Vershik, A. M., Fadeev, L. D. [1981], Lagrangian mechanics in invariant form, *Selecta Math. Sov.*, **1**:4, 339–350.

Veselov, A. P., Veselova, L. E. [1986], Flows on Lie groups with a nonholonomic constraint and integrable non-Hamiltonian systems, *Funkt. Anal. Prilozhen.*, **20**:4, 65–66 (in Russian); *Functional Anal. Appl.*, **20**:4, 308–309 (in English).

Veselov, A. P., Veselova, L. E. [1988], Integrable nonholonomic systems on Lie groups, *Mat. Z.*, **44**:5, 604–619, 701 (in Russian); *Math. Notes*, **44**:5–6 (1989), 810–819 (in English).

Wade, A. [2000], Conformal Dirac structures, *Lett. Math. Phys.*, **53**, 331–348.

Warner, F. [1971], *Foundations of Differentiable Manifolds and Lie Groups*, Scott Foresman, Glenview, IL.

Weber, R. W. [1986], Hamiltonian systems with constraints and their meaning in mechanics, *Arch. Rational Mech. Anal.*, **91**, 309–335.

Whittaker, E. T. [1988], *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies, with an Introduction to the Problem of Three Bodies*, reprint of the 1937 ed., Cambridge University Press, Cambridge, UK.

Zenkov, D. V. [1995], The geometry of the Routh problem, *J. Nonlinear Sci.*, **5**:6, 503–519.

Zenkov, D. V., Bloch, A. M. [2000], Dynamics of the n-dimensional Suslov problem, *J. Geom. Phys.*, **34**:2, 121–136.

Zenkov, D. V., Bloch, A. M. [2003], Invariant measures of nonholonomic flows with internal degrees of freedom, *Nonlinearity*, **16**, 1793–1807.