Differential Geometry and its Applications 28 (2010) 436–453

Contents lists available at ScienceDirect

Differential Geometry and its Applications

www.elsevier.com/locate/difgeo

# Chaplygin systems associated to Cartan decompositions of semi-simple Lie groups

Simon Hochgerner

*Section de Mathematiques, Station 8, EPFL, CH-1015 Lausanne, Switzerland*

ARTICLE INFO

*Article history:*
Received 3 June 2009
Received in revised form 22 September 2009
Communicated by T.S. Ratiu

Dedicated to Peter Michor on the occasion of his 60th birthday

*MSC:*
37J60
37J15

*Keywords:*
Non-holonomic systems
Symmetries

ABSTRACT

We relate a Chaplygin type system to a Cartan decomposition of a real semi-simple Lie group. The resulting system is described in terms of the structure theory associated to the Cartan decomposition. It is shown to possess a preserved measure and when internal symmetries are present these are factored out via a process called truncation. Furthermore, a criterion for Hamiltonizability of the system on the so-called ultimate reduced level is given. As important special cases we find the Chaplygin ball rolling on a table and the rubber ball rolling over another ball.



## 1. Introduction

We generalize the $n$-dimensional Chaplygin ball problem [8,13,10,9,16,14] to non-holonomic systems associated to semi-simple Lie groups, and show how the Chaplygin ball system arises as a special case. That is, we consider a real semi-simple Lie group $G$ and a Cartan decomposition $G \cong K \times \mathfrak{p}$ in the common notation of [17]. On the Lie algebra level we have $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ together with the usual bracket relations. In $\mathfrak{p}$ we fix a maximal abelian subspace $\mathfrak{a}$ and an element $w_0 \in \mathfrak{a}$. In Section 3 we define a non-holonomic system that is naturally associated to these data: the configuration space is

$$Q := K \times V$$

where $V$ is orthogonal to $Z_{\mathfrak{p}}(w_0) = \{x \in \mathfrak{p}\colon\ [w_0, x] = 0\}$ within $\mathfrak{p}$, the constraint distribution is

$$\mathcal{D} := \bigl\{\bigl(s, u, x, \bigl[w_0, \mathrm{Ad}(s)u\bigr]\bigr) \in K \times \mathfrak{k} \times V \times V\bigr\} \subset TQ,$$

and the Lagrangian is the obvious left invariant kinetic energy function on $TQ$. Then we use the restricted roots of the pair $(\mathfrak{g}, \mathfrak{a})$ to give a detailed description of this model. We will see that the $n$-dimensional Chaplygin ball corresponds to taking $G = \mathrm{SO}(n, 1)$.

We extend some of the results of [13,16,14] to this setting. In particular this yields a geometrization of these results since we follow the philosophy of [10] in working with a global trivialization of the compressed phase space and using (almost) symplectic techniques.

*E-mail address:* simon.hochgerner@epfl.ch.

0926-2245/$ – see front matter 
doi:10.1016/j.difgeo.2010.04.003

More precisely, by making use of the restricted root space decomposition associated to $(\mathfrak{g}, \mathfrak{a})$ we directly show the existence of a preserved measure for these types of systems at the compressed level – Proposition 3.4.

Then we pass to the ultimate reduced phase space by means of truncation and reduction of internal symmetries. This involves changing the non-holonomic two-form in a certain way that is better adapted to the symmetries – Section 3.6. The passage from the original non-holonomic system to this reduced phase space via compression followed by reduction of internal symmetries is reminiscent of the Hamiltonian reduction in stages theory which also lends the terminology 'ultimate reduced space'.

Moreover, in Theorem 3.6 we derive a necessary and sufficient condition for Hamiltonization of the ultimate reduced system when the angular momentum with respect to the internal symmetries is fixed to 0. This condition is of algebraic nature and in some simple cases it allows to decide (non-) Hamiltonizability by looking at the root system of $(\mathfrak{g}, \mathfrak{a})$. This result is a statement which only holds at the ultimate reduced level and thus depends crucially on the reduction by truncation described in Section 3.6.

Section 4 contains some examples. We return to the $n$-dimensional Chaplygin ball system corresponding to $G = \mathrm{SO}(n, 1)$ and apply Theorem 3.6 to verify the recent result of Jovanovic [16] on Hamiltonizability of this system at the ultimate reduced level when the angular momentum is fixed to 0 and the inertia tensor is of special form. Then we give two examples related to $\mathrm{SL}(n, \mathbb{R})$ and $\mathrm{Sp}(n, \mathbb{R})$.

Finally, we show how the rubber rolling sphere-on-sphere system arises in this setting. This is not so straightforward as for the ball on a table: We start with the split real form of the complex semi-simple Lie group $G_2$ and consider, according to the recipe of Section 3, its Cartan decomposition. The resulting system is shown to be never Hamiltonizable, not even for homogeneous inertia tensor $\mathbb{I} = 1$. However, from Koiller and Ehlers [19] we know that the rubber rolling system is Hamiltonizable. Thus we are motivated to find a subsystem which is an obvious candidate for allowing Hamiltonizability. This subsystem is then recognized as the rubber ball arrangement for the case in which the ratio of the radii of the balls is 1 : 3. However, we are not claiming that we provide any new insights into the dynamics of this system; we only find a new way to see this as being part of a non-holonomic system that is naturally defined on some bigger phase space.

In Section 2 we recall the notion of Hamiltonization of a non-holonomic system. Then we reformulate the Chaplygin multiplier theorem in terms of a characterization of conformally closed almost symplectic forms which is due to Libermann [20,21]. This characterization extends to higher dimensions whence we also formulate a higher-dimensional analogon of the multiplier theorem. In Section 3.7 this is used as a preparation for Theorem 3.6.

## 2. Remarks on Hamiltonization

Non-holonomic systems can be seen as a generalization of Hamiltonian mechanics. A natural question that arises is: when is a non-holonomic system Hamiltonian or *Hamiltonizable*?

As a toy example to illustrate some key ideas and also to set up notation we consider the vertical rolling disk. For more information on this, and also on more complicated examples, see Bloch [4]. The configuration space is

$$Q = S^1 \times S^1 \times \mathbb{R}^2$$

with coordinates $q = (\theta, \varphi, x, y)$. Here $(x, y)$ denotes the contact point of the disk on the table, $\theta$ its internal orientation, and $\phi$ its orientation with respect to a fixed axis on the table. The Lagrangian is the kinetic energy

$$L = \frac{1}{2}\mathbb{I}\dot{\theta}^2 + \frac{1}{2}\mathbb{J}\dot{\varphi}^2 + \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right)$$

where $m$ is the mass of the disk and $\mathbb{I}$ and $\mathbb{J}$ are the different moments of inertia of the disk. The motion is to satisfy a no slip constraint which means that

$$\dot{x} = R\dot{\theta}\cos\varphi \quad \text{and} \quad \dot{y} = R\dot{\theta}\sin\varphi$$

where $R$ is the radius of the disk. To rewrite these constraints in a more geometric manner consider the $\mathbb{R}^2$-valued one-form $\mathcal{A} \in \Omega^1(S, \mathbb{R}^2)$ on $S := S^1 \times S^1$ given by

$$\mathcal{A}_{(\theta,\varphi)} = \begin{pmatrix} -R\cos\varphi\, d\theta \\ -R\sin\varphi\, d\theta \end{pmatrix}.$$

Let $\pi : Q = S \times \mathbb{R}^2 \to S$ denote the Cartesian projection. The constraint space is thus defined by the smooth distribution

$$\mathcal{D} = \left\{\left(q, \dot{\theta}, \dot{\varphi}, -\mathcal{A}_{\pi(q)}(\dot{\theta}, \dot{\varphi})\right)\right\} \subset TQ.$$

Now it is important to notice that $L$ and $\mathcal{D}$ are invariant under the free and proper action of the abelian Lie group $\mathbb{R}^2$ on $TQ$. This action defines a (trivial) principal fiber bundle $\mathbb{R}^2 \hookrightarrow Q \twoheadrightarrow S$. Moreover, $\mathcal{D}$ is complementary to the vertical space $\ker T\pi$ of this bundle. In other words $\mathcal{D}$ defines a principal connection with connection form $\mathcal{A}$ and the non-holonomic system $(Q, L, \mathcal{D})$ is a *G-Chaplygin system* with $G = \mathbb{R}^2$. This system is truly non-holonomic since $\mathcal{D}$ is non-integrable since the curvature $\mathrm{Curv}_0^{\mathcal{A}} = d\mathcal{A}$ is non-zero.

$G$-Chaplygin systems are very well behaved in the sense that they allow for a natural reduction of symmetries. For this our main reference is [10] where this reduction is termed *compression*. See also [3] for a more general reduction and [14] for an account of these facts in the present notation. The compressed system turns out to be an almost Hamiltonian system on $T^*S$ with compressed Hamiltonian $\mathcal{H}_c$. Of course, $\mathcal{H}_c$ is obtained by taking the Legendre transform of $L$, restricting to the appropriate constraint subspace and factoring out the symmetries. The dynamics $X_{nh} = (\Omega_{nh})^{-1} d\mathcal{H}_c$ of the compressed system are encoded in the almost symplectic form

$$\Omega_{nh} := \Omega^S - \big\langle J \circ \mathrm{horLift}^{\mathcal{A}}, \mathrm{Curv}_0^{\mathcal{A}}\big\rangle = \Omega^S + \langle \mathcal{A}, d\mathcal{A}\rangle$$

where $\Omega^S$ is the canonical symplectic form on $T^*S = TS$ (identified via induced Legendre transform), $\mathrm{horLift}^{\mathcal{A}} : TS \to TQ$ is the horizontal lift, $J : TQ = T^*Q \to \mathbb{R}^{2*} = \mathbb{R}^2$ (Legendre transform) is the standard momentum map associated to the $\mathbb{R}^2$-action, and $\mathrm{Curv}_0^{\mathcal{A}}$ is the induced curvature form on $S$ pulled-back to $TS$. Note that $\langle \mathcal{A}, d\mathcal{A}\rangle$ is a semi-basic two-form on $TS$ which depends linearly on the fibers; the $\mathcal{A}$ in the left hand side of the pairing is viewed as a function on $TS$. In general, the term $\langle J \circ \mathrm{horLift}^{\mathcal{A}}, \mathrm{Curv}_0^{\mathcal{A}}\rangle$ is non-closed thus preventing the system form being Hamiltonian. However, in this special example we have

$$\langle \mathcal{A}, d\mathcal{A}\rangle_{(\theta,\varphi,\dot\theta,\dot\varphi)} = R^2 \left\langle \begin{pmatrix} \dot\theta \cos\varphi \\ \dot\theta \sin\varphi \end{pmatrix}, \begin{pmatrix} -\sin\varphi\, d\varphi \wedge d\theta \\ \cos\varphi\, d\varphi \wedge d\theta \end{pmatrix} \right\rangle = 0.$$

Thus the compressed system $(TS, \Omega^S, \mathcal{H}_c)$ is Hamiltonian even though we started from a truly non-holonomic system $(Q, L, \mathcal{D})$. Of course, this fact is neither new nor surprising: the constraint forces for this system are trivial.

More generally it may turn out that $\Omega_{nh}$ is *conformally symplectic* with respect to a positive function $F : S \to \mathbb{R}$, that is, $d(F\Omega_{nh}) = 0$. If this is the case we consider the rescaled vectorfield $F^{-1}X_{nh}$ which is now Hamiltonian with respect to $F\Omega_{nh}$, and we say that the system $(T^*S, \Omega_{nh}, \mathcal{H}_c)$ is *Hamiltonizable* or that $(Q, L, \mathcal{D})$ is *Hamiltonizable at the compressed level*. The idea is that one reparametrizes the time $dt = F^{-1} d\tau$ in an $F$-dependent manner so that the system is Hamiltonian in the new time $\tau$.

### 2.1. Chaplygin's multiplier theorem via Libermann's criterion

Let $(M, \sigma)$ be an almost symplectic manifold of dimension $2m$, that is, $\sigma$ is non-degenerate. Then we will make use of the codifferential operator

$$\delta : \Omega^k(M) \to \Omega^{k-1}(M)$$

which is built out of $\sigma$ in the same way that the Hodge codifferential is built out of a metric. This operator is explained in the first chapter of the book of Libermann and Marle [21] and we use the same conventions.

**Theorem 2.1** *(Chaplygin). Let $B$ be a 2-dimensional Riemannian manifold. Consider the natural kinetic energy Hamiltonian $\mathcal{H} : T^*B \to \mathbb{R}$ associated to the metric. Let $(T^*B, \sigma, \mathcal{H})$ be an almost Hamiltonian system such that:*

(1) *$\sigma = \Omega + \Lambda$ where $\Lambda$ is semi-basic with respect to $T^*B \to B$ and linear in the fiber. That is, locally, $\Lambda = l(q,p)\, dq^1 \wedge dq^2$ with $l$ linear in $p$. Further, $\Omega = \Omega^B + \Xi$ with $\Xi$ magnetic, that is, closed and basic.*
(2) *There is a function $F : B \to \mathbb{R}_{>0}$ such that $L_X(F\sigma^2) = 0$ where $X$ is the vector field associated to $\mathcal{H}$ via $\sigma$.*

*Then*

$$\delta\sigma = -d(\log F) \quad \textit{and} \quad d(F\sigma) = 0.$$

**Proof.** The following formula can be found in [21]:

$$d\sigma = \delta\sigma \wedge \sigma$$

which holds since $\dim B = 2$, and thus

$$d(f\sigma) = \big(\delta\sigma + d(\log f)\big) \wedge f\sigma \tag{2.1}$$

for an arbitrary smooth function $f : T^*B \to \mathbb{R}$. Therefore,

$$0 = L_X\big(F\sigma^2\big) = 2d(F\, d\mathcal{H} \wedge \sigma) = 2(dF + F\delta\sigma) \wedge d\mathcal{H} \wedge \sigma.$$

Using the special structure of $\Lambda$ we can show that $\delta\sigma$ is basic. (See Lemma 2.4.) Therefore, since $\mathcal{H}$ is natural it follows that $dF + F\delta\sigma = 0$. Thus $d(F\sigma) = 0$ by (2.1). □

In particular, this proves Hamiltonization of the 3-dimensional Chaplygin ball at the ultimate reduced level – the $T^*S^2$-level which can be attained after truncation. It is remarkable that this theorem as well as its crucial assumption – the preserved measure – had already been found by Chaplygin. Nevertheless, he could not apply these facts to conclude Hamiltonizability of the problem. This is probably due to the fact that it is not entirely straightforward to reduce all the relevant structure in a coherent manner to the $T^*S^2$-level. See [14]. Indeed, it was Borisov and Mamaev [6,7] who invented a proof of Hamiltonizability of this system.

### 2.2. A multiplier theorem for higher dimensions

Let $(M,\sigma)$ be a $2m$-dimensional almost symplectic manifold with codifferential $\delta$. According to [20], [21, Proposition I.16.5] there is a certain (*effective*) three-form $\psi$ such that

$$d\sigma = \psi + \frac{1}{m-1}\delta\sigma\wedge\sigma. \tag{2.2}$$

Moreover, $\sigma$ is *locally* conformal symplectic if and only if $\psi = 0$.

Thus for an almost Hamiltonian system $(T^*B = M, \sigma, \mathcal{H})$ with dynamics given by $X = \sigma^{-1}d\mathcal{H}$ there are two obvious necessary conditions for a function to be a conformal factor.

**Lemma 2.2.** *If a function $F : B \to \mathbb{R}_{>0}$ is a conformal factor, that is $d(F\sigma) = 0$, then $\psi = 0$ and there is a preserved measure with density $F^{m-1}$, that is $L_X(F^{m-1}\sigma^m) = 0$.*

The following statement attempts to reverse the situation: When $\psi$ vanishes we know that the structure is locally conformally symplectic; when there is additionally a preserved measure then we can turn this local statement to a global one.

In fact, we will consider a slightly more general situation by allowing the almost Hamiltonian system to have additional internal degrees of freedom: Let $H \hookrightarrow S \twoheadrightarrow B$ be a principal fiber bundle which is at the same time a Riemannian submersion. That is, $(S,\mu_S)$ and $(B,\mu_B)$ are Riemannian manifolds, $\mu_S$ is $H$-invariant and the bundle projection map induces an isometry $\mathrm{Hor}(\mu_S) = \mathrm{Ver}^\perp \to TB$. Let us denote the connection form corresponding to $\mathrm{Hor}(\mu_S)$ by $A : TS \to \mathfrak{h}$. This is the mechanical connection on $(S,\mu_S)$ (and should not be confused with the $\mathcal{A}$ appearing in Section 3). We suppose that $T^*S$ is equipped with an almost symplectic form $\widetilde{\Omega} := \Omega^S + \Lambda$ where $\Lambda$ is $H$-basic with respect to $T^*S \twoheadrightarrow (T^*S)/H$, semi-basic with respect to $T^*S \to S$ and linear in the fibers of $T^*S$. Thus $\widetilde{\Omega}$ admits a momentum map $J_H : T^*S \to \mathfrak{h}^*$ which is the standard one, since $\Lambda$ vanishes upon insertion of infinitesimal generators of the $H$-action.

Further, assume that there is a right Hamiltonian $H$-space $(F,\Omega^F)$ with equivariant momentum map $J_F : F \to \mathfrak{h}^*$.

Then we consider the diagonal action of $H$ on $T^*S \times F$ where the $H$-action on the second factor is inverted to give a left action. This action admits a momentum map which is given by $J := J_H - J_F$. Notice that $(s,u,f) \in J^{-1}(0)$ if and only if $u = u_0 + A_s^*(J_F(f))$ with $u_0 \in \mathrm{Hor}_s^*$. Thus we may pass to the reduced space

$$J^{-1}(0)/H \cong T^*B \times_B (S \times_H F) =: \mathcal{W}$$

where the isomorphism is defined in terms of the connection $A$. In particular, the reduced space $\mathcal{W}$ is a (symplectic) fiber bundle over $T^*B$ with fiber $F$. By construction the form $\widetilde{\Omega} + \Omega^F$ is basic when restricted to $J^{-1}(0)$ and passes to an almost symplectic form on $T^*B \times_B (S \times_H F)$ which we shall denote by $\sigma_A$ to emphasize the $A$-dependence. This is the Weinstein construction rewritten for a semi-basic perturbation of the standard symplectic form on $T^*S$. By the usual computation one sees that

$$\sigma_A = \Omega^B - \langle J_F, \mathrm{Curv}^A\rangle + \Lambda_0 + \Omega^F \tag{2.3}$$

where $\Omega^B$ is the canonical symplectic form on $T^*B$, the second term is magnetic and $\Lambda_0$ is the non-closed semi-basic term induced from $\Lambda$.

The situation which we have in mind is that of [14, Corollary 4.2].

**Theorem 2.3.** *Consider the natural kinetic energy Hamiltonian $\mathcal{H} : T^*S \to \mathbb{R}$ associated to the metric $\mu_S$ and let $\mathcal{H} : \mathcal{W} \to \mathbb{R}$ also denote the induced function. Let $m = \frac{1}{2}\dim\mathcal{W}$, $n = \dim B$ and $k = \frac{1}{2}\dim F$, whence $m = n + k$. Assume that*:

(1) *There is a function $F : B \to \mathbb{R}_{>0}$ such that $L_X(F^{m-1}\sigma_A^m) = 0$ where $X$ is the vector field associated to $\mathcal{H}$ via $\sigma_A$.*
(2) *$\psi = 0$, or, equivalently $d\sigma_A = \frac{1}{m-1}\delta\sigma_A\wedge\sigma_A$.*

*Then*

$$(m-1)\,d\log F = -\delta\sigma_A \quad \text{and} \quad d(F\sigma_A) = 0,$$

*that is, the almost Hamiltonian system $(\mathcal{W},\sigma_A,\mathcal{H})$ with dynamics given by $X = \sigma_A^{-1}d\mathcal{H}$ can be transformed to a Hamiltonian system $(\mathcal{W}, F\sigma_A, \mathcal{H})$ with rescaled dynamics $F^{-1}X$.*

**Proof.** According to (2.2) we have

$$d(f\sigma_A) = \frac{1}{m-1}\big(\delta\sigma_A + (m-1)d\log f\big) \wedge f\sigma_A \tag{2.4}$$

for all smooth functions $f : \mathcal{W} \to \mathbb{R}_{>0}$.

We use local Darboux coordinates $q^a, p_a$ on $T^*B$. Because of Lemma 2.4 the one-form $\delta\sigma_A$ is basic. Thus we have

$$(m-1)\, d\log F + \delta\sigma_A = \sum \phi_a(q)\, dq^a$$

in the local coordinates. Since $\sigma_A^m = (\Omega^B)^n \wedge (\Omega^F)^k$,

$$\begin{aligned}
0 &= di_X\big(F^{m-1}\sigma_A^m\big) = md\big(F^{m-1}\, d\mathcal{H} \wedge \sigma_A^{m-1}\big)\\
&= m\big((m-1)F^{m-2}\, dF \wedge d\mathcal{H} \wedge \sigma_A^{m-1} - F^{m-1}\, d\mathcal{H} \wedge \delta\sigma_A \wedge \sigma_A \wedge \sigma_A^{m-2}\big)\\
&= mF^{m-1}\big((m-1)\, d\log F + \delta\sigma_A\big) \wedge d\mathcal{H} \wedge \sigma_A^{m-1}\\
&= mF^{m-1} \sum \phi_a\, dq^a \wedge \sum \frac{\partial \mathcal{H}}{\partial p_b}\, dp_b \wedge \Big(\sum dq^c \wedge dp_c\Big)^{n-1} \wedge \big(\Omega^F\big)^k\\
&= \frac{mF^{m-1}}{(m-1)!} \sum \phi_a \frac{\partial \mathcal{H}}{\partial p_a} dq^1 \wedge dp_1 \wedge \cdots \wedge dq^n \wedge dp_n \wedge \big(\Omega^F\big)^k.
\end{aligned}$$

Since $\phi_a$ depends only on $q$ and $\mathcal{H}$ is regular it follows that $\phi_a = 0$. Because $\psi = 0$ in (2.4) this finishes the proof. □

**Lemma 2.4.** *Under the assumptions of Theorem 2.3, $\delta\sigma_A$ is basic with respect to the projection $\mathcal{W} \to T^*B \to B$.*

**Proof.** We use local Darboux coordinates $q^a, p_a$ on $T^*B$ and coordinates $f^i$ on $F$. According to (2.3) we may write $\sigma_A$ terms of

$$\begin{aligned}
&\Omega^B = \sum dq^a \wedge dp_a, && \langle J_F, \mathrm{Curv}^A\rangle = \sum \Xi_{ab}\, dq^a \wedge dq^b,\\
&\Lambda_0 = \sum \Lambda_{ab}\, dq^a \wedge dq^b, && \Omega^F = \sum \Omega^F_{ij}\, df^i \wedge df^j.
\end{aligned} \tag{2.5}$$

Let us write $\delta\sigma_A$ as

$$\delta\sigma_A = \sum \big(C_a(q,p,f)\, dq^a + C^a(q,p,f)\, dp_a + D_i(q,p,f)\, df^i\big).$$

We need to show that $C^a = 0$, $D_i = 0$ and $C_a = C_a(q)$. Using the relation

$$d\sigma_A = d\Lambda_0 = \frac{1}{m-1}\delta\sigma_A \wedge \sigma_A,$$

expanding it in terms of (2.5), and inserting a pair $\frac{\partial}{\partial p_a}, \frac{\partial}{\partial p_b}$ of vertical vectors on both sides we see that $C^a = 0$ for all $a$. Similarly one sees that $D_i = 0$. Now we insert vectors $\frac{\partial}{\partial q^b}, \frac{\partial}{\partial q^a}, \frac{\partial}{\partial p_a}$ on both sides, and see that $C_a(q,p) = d_v\Lambda_{ba}(\frac{\partial}{\partial p_a}) = C_a(q)$. (It is here that we use that $\Lambda$ is linear in the fiber.) □

## 3. Chaplygin systems associated to semi-simple Lie groups

We associate a Chaplygin type system to a Cartan decomposition (and choice of a restricted root system) of an arbitrary (real) semi-simple Lie group. In Section 3.3 it is shown that this construction generalizes the classical $n$-dimensional Chaplygin ball system. For background on semi-simple Lie groups we refer to Knapp [17].

The systems considered in this paper belong to the class of non-holonomic systems defined on semi-direct products [23] and coupled $LR$-systems [15]. The compression of a coupled $LR$-system is an $L+R$-system [15], so the compressed vectorfield $X_{\mathrm{nh}}$ defines an $L+R$-flow on $TK$. See [11,12].

### 3.1. Configuration space and constraints

Let $G$ be a semi-simple Lie group with Lie algebra $\mathfrak{g}$ and Killing form $B$. Consider a Cartan decomposition $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ associated to the Cartan involution $\theta$, and let $G \cong K \times \mathfrak{p}$, $g = k\exp x \leftarrow (k, x)$ be the corresponding decomposition of the group. Thus:

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \qquad [\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}, \qquad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}.$$

Fix a maximal abelian subspace $\mathfrak{a} \subset \mathfrak{p}$, and put $\mathfrak{m} = Z_{\mathfrak{k}}(\mathfrak{a})$ and $M = Z_K(\mathfrak{a})$. Fix also an element $w_0 \in \mathfrak{a}$.[1] Define $Z_K(w_0) = H$ to be the stabilizer of this vector, and note that

$$\mathrm{ad}(w_0)|\mathfrak{h}^{\perp} : \mathfrak{h}^{\perp} := \mathfrak{h}^{B\perp} \cap \mathfrak{k} \to \mathrm{ad}(w_0)\big(\mathfrak{h}^{\perp}\big) =: V \subset \mathfrak{p} \tag{3.6}$$

is an isomorphism onto its image $V$. If $w_0$ is regular then $H = M$ and $V = \mathfrak{a}^{\perp} \cap \mathfrak{p}$.

The configuration space is now defined to be

$$Q := K \times V.$$

The Lagrangian is the natural kinetic energy Lagrangian $L$ which is associated to the positive definite inner product $B_\theta = -B(.,\theta.) = -B|\mathfrak{k} + B|V$ taking into account the inertia tensor which is a symmetric positive definite endomorphism $\mathbb{I}$ of $(\mathfrak{k}, -B|\mathfrak{k})$. Thus

$$L = \frac{1}{2}\langle \mathbb{I}u, u\rangle + \frac{1}{2}\langle x', x'\rangle$$

where $\langle .,. \rangle = B_\theta$. This Lagrangian is *left*-invariant (i.e., invariant with respect to left multiplication of $K$ on the first factor of $Q$) since we identify $TK = K \times \mathfrak{k}$ via the *left* multiplication, $u = s^{-1}s'$.

The distribution is

$$\mathcal{D} = \big\{\big(s, u, x, -\mathcal{A}_s(u)\big)\big\} \subset TK \times TV$$

where

$$\mathcal{A} : (s,u) \longmapsto -\big[\mathrm{Ad}(s)u, w_0\big] = -\mathrm{pr}_V\big(\big[\mathrm{Ad}(s)u, w_0\big]\big), \qquad TK \to V \tag{3.7}$$

and $w_0$ has been fixed to define the isomorphism (3.6). It is customary to define also $\mathcal{A}^{\mathrm{up}} \in \Omega^1(Q, V)$ by $\mathcal{A}^{\mathrm{up}}(s,u,x,x') = x' + \mathcal{A}_s(u)$. Then $\mathcal{D} = (\mathcal{A}^{\mathrm{up}})^{-1}(0)$ and one says that $\mathcal{A}$ is the local connection form associated to the global (or upstairs) connection form $\mathcal{A}^{\mathrm{up}}$ on $Q \twoheadrightarrow Q/V$.

$(Q, \mathcal{D}, L)$ is a $V$-Chaplygin system with abelian Lie group $V$. This precisely means that $(Q, \mathcal{D}, L)$ is a non-holonomic system which is invariant under the free and proper action of the abelian Lie group $V$ and that the distribution $\mathcal{D}$ determines a principal bundle connection on $Q \twoheadrightarrow Q/V$. The following are essential observations.

(1) $\mathcal{A} : TK \to V$ is the connection form associated to $\mathcal{D}$ on the principal fiber bundle $V \hookrightarrow Q \twoheadrightarrow K$.
(2) $\mathcal{A}$ is right invariant.

The group $H = \{h \in K:\ \mathrm{Ad}(h)w_0 = w_0\}$ acts through *two* different actions on $Q$:

(3) The $l$-action: $l_h(s,x) = (hs, x)$. This action generates *internal* symmetries: $\mathcal{A}\zeta_Y^l = 0$ for all $Y \in \mathfrak{h}$ ($\zeta_Y^l(s) = \mathrm{Ad}(s^{-1}).Y$).
(4) The $d$-action: $d_h(s,x) = (hs, hx)$. This action generates *external* symmetries. $\mathcal{A}(hs, u) = h.\mathcal{A}(s,u)$ for all $h \in H$. Thus $\mathcal{D}$ is invariant under the $d$-action.

This should be compared to the set-up in [14].

### 3.2. Non-holonomic reduction: The compressed system

Compression refers to the passage from the non-holonomic system $(Q, \mathcal{D}, L)$ with (external) symmetry group $V$ to an almost Hamiltonian system $(T^*(Q/V), \Omega_{\mathrm{nh}}, \mathcal{H}_{\mathrm{c}})$. The metric $\mu$ on $Q$ is invariant under the $V$-action. By requiring $Q \twoheadrightarrow Q/V$ to be a Riemannian submersion there is a well-defined induced metric $\mu_0$ on $Q/V = K$. We will henceforth use $\mu_0$ to identify $T^*K = TK$. According to general results on compression in the presence of internal symmetries (e.g., [10,14,3,18]):

The compressed Hamiltonian is

$$\mathcal{H}_{\mathrm{c}}(s,u) = \frac{1}{2}\langle \mathbb{I}u, u\rangle + \frac{1}{2}\langle \mathcal{A}_s(u), \mathcal{A}_s(u)\rangle$$

which is $H$-invariant. The compressed almost symplectic form is

$$\Omega_{\mathrm{nh}} = \Omega^K - \big\langle J_V \circ \mathrm{hl}^{\mathcal{A}}, \mathrm{Curv}_0^{\mathcal{A}}\big\rangle_V = \Omega^K + \langle \mathcal{A}, d\mathcal{A}\rangle_V$$

which is also $H$-invariant. The dynamics are given by $X_{\mathrm{nh}}$:

[1] This corresponds to the vertical vector orthogonal to the table in the case of the $n$-dimensional Chaplygin ball.

$$i(X_{\mathrm{nh}})\Omega_{\mathrm{nh}} = d\mathcal{H}_{\mathrm{c}}.$$

Finally, according to the non-holonomic Noether Theorem there is a conserved quantity:

$$J_H : TK \to \mathfrak{h}^*$$

which is the standard momentum map.

What about reduction? Can this data be reproduced on a quotient of the form $J_H^{-1}(\lambda)/H_\lambda$ for some value $\lambda \in \mathfrak{h}^*$. Just like in, e.g., [14] the problem that arises is that $J_H$ is (for $w_0 \neq 0$) not a momentum map with respect to $\Omega_{\mathrm{nh}}$. Thus the restriction of $\Omega_{\mathrm{nh}}$ to a level set $J_H^{-1}(\lambda)$ is not horizontal with respect to the induced action of the stabilizer subgroup $H_\lambda$. We will return to this problem in Section 3.6.

### 3.3. *Example:* SO$(p,q)$ *and Chaplygin's ball*

Let $G = \mathrm{SO}(p,q)_0$ with $p \geqslant q$. Then the spaces under consideration are the following.

$$K = \bigl\{\mathrm{diag}(A, D):\ A \in \mathrm{SO}(p),\ D \in \mathrm{SO}(q)\bigr\},$$
$$\mathfrak{p} = \left\{\begin{pmatrix} 0_{p\times p} & b \\ b^t & 0_{q\times q}\end{pmatrix}:\ b \in \mathfrak{gl}(p\times q, \mathbb{R})\right\}$$

and

$$\mathfrak{a} = \left\{\begin{pmatrix} 0_{p\times p} & b \\ b^t & 0_{q\times q}\end{pmatrix}:\ b \text{ has only lower antidiagonal non-zero}\right\} = \mathbb{R}^q,$$
$$M = \bigl\{\mathrm{diag}\bigl(\mathrm{SO}(p-q), \theta_q, \dots, \theta_1, \theta_1, \dots, \theta_q\bigr):\ \theta_i = \pm 1,\ \Pi\theta_i = 1\bigr\} = \mathrm{SO}(p-q) \times \{\pm 1\}^{q-1}.$$

Therefore,

$$K/M = \bigl(\mathrm{SO}(p)/\mathrm{SO}(p-q)\times \mathrm{SO}(q)\bigr)/\{\pm 1\}^{q-1} \cong V(q,p) \times \mathrm{SO}(q)/\{\pm 1\}^{q-1}$$

which is the ultimate reduced configuration space.

#### 3.3.1. *Special case* $q = 1$, $p \geqslant 3$

In this case there is only one positive root and assuming that $w_0 \neq 0$ yields the following.

$$K = \mathrm{SO}(p) \times \{1\},$$
$$\mathfrak{p} = \left\{\begin{pmatrix} 0_{p\times p} & b \\ b^t & 0\end{pmatrix}:\ b \in \mathfrak{gl}(p\times 1, \mathbb{R}) = \mathbb{R}^p\right\},$$
$$\mathfrak{a} \cong \mathbb{R}^1 \quad \text{and} \quad V = \mathfrak{a}^\perp \cong \mathbb{R}^{p-1},$$
$$H = M \cong \mathrm{SO}(p-1).$$

Thus,

$$\mathfrak{g} = \begin{pmatrix} \mathfrak{so}(p) & \mathbb{R}^p \\ (\mathbb{R}^p)^* & 0\end{pmatrix} \quad \text{and} \quad w_0 := \begin{pmatrix} 0 & e_p \\ e_p^t & 0\end{pmatrix} \in \mathfrak{a} \subset \mathfrak{g}$$

yield

$$\mathcal{A}_s(u) = -\mathrm{pr}_V\bigl[\mathrm{Ad}(s)u, w_0\bigr] = \begin{pmatrix} 0 & -(\mathrm{Ad}(s)u).e_p \\ -((\mathrm{Ad}(s)u).e_p)^t & 0\end{pmatrix} \in V$$

which can be identified with the connection form

$$T\mathrm{SO}(p) \to \mathbb{R}^{p-1}, \qquad (s,u) \longmapsto -\mathrm{pr}_{\mathbb{R}^{p-1}}\bigl(\bigl(\mathrm{Ad}(s)u\bigr).e_p\bigr)$$

describing the $p$-dimensional Chaplygin system when mass and radius of the ball are both set to 1. See [13,10,14]. Moreover,

$$K/M = V(1,p) = S^{p-1}$$

whence we recover the $p$-dimensional Chaplygin ball. (The Lagrangian $L$ also identifies in the expected way.)

*3.4. Describing the system*

In this section we introduce notation and formulae that will be used very much in the subsequent. Let $\Sigma$ be the set of restricted roots associated to the pair $(\mathfrak{g},\mathfrak{a})$ and $\Sigma_+ \subset \Sigma$ a choice of positive roots. Then the associated root space decomposition is

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \bigoplus_{\lambda\in\Sigma} \mathfrak{g}_\lambda \quad \text{where } \mathfrak{g}_0 = \mathfrak{m}\oplus\mathfrak{a}.$$

Moreover, we choose an orthonormal system

$$Y_\alpha, \quad \alpha = 1,\dots,\dim\mathfrak{m} \quad \text{and} \quad Z_{(\lambda,a)}, \quad \lambda\in\Sigma_+,\ a = 1,\dots,\dim\mathfrak{g}_\lambda$$

that is adapted to the decomposition $\mathfrak{k} = \mathfrak{m}\oplus\mathfrak{m}^\perp$, and an orthonormal basis

$$e_{(\lambda,a)}, \quad \lambda\in\Sigma_+,\ a = 1,\dots,\dim\mathfrak{g}_\lambda$$

of $\mathfrak{a}^\perp\cap\mathfrak{p}$. We assume further the relations

$$\operatorname{ad}(w)Z_{(\lambda,a)} = \lambda(w)e_{(\lambda,a)} \quad \text{and} \quad \operatorname{ad}(w)e_{(\lambda,a)} = \lambda(w)Z_{(\lambda,a)} \tag{3.8}$$

for all $w\in\mathfrak{a}$. Such a basis always exists. In the following we will use the convention that $\alpha,\beta,\gamma,\dots$ take values $1,\dots,\dim\mathfrak{m}$, and pairs $(\lambda,a),(\mu,b),(\nu,c)$ have their first component in $\Sigma_+$ while the second component runs from 1 to the dimension of the corresponding root space. The basis vectors $Y_\alpha$, $Z_{(\lambda,a)}$ as well as their dual basis are right extended to give a right invariant frame and coframe

$$\xi_\alpha,\ \zeta_{(\lambda,a)} \quad \text{and} \quad \rho^\alpha,\ \eta^{(\lambda,a)}$$

of $K$. With respect to the left trivialization this frame and coframe become

$$\xi_\alpha(s) = \operatorname{Ad}\big(s^{-1}\big)Y_\alpha = s^{-1}Y_\alpha \quad \text{and} \quad \rho^\alpha(s)(u) = \big\langle\operatorname{Ad}\big(s^{-1}\big)Y_\alpha, u\big\rangle = \big\langle s^{-1}Y_\alpha, u\big\rangle,$$

etc. (We will often suppress the Ad-notation and simply write $s^{-1}Y$ for $\operatorname{Ad}(s^{-1})Y$.) It will be convenient to use the notation

$$l_\alpha = \rho^\alpha : TK \to \mathbb{R} \quad \text{and} \quad g_{(\lambda,a)} = \eta^{(\lambda,a)} : TK \to \mathbb{R}$$

when we view the one-forms as functions on the tangent bundle. These functions are the components of the angular velocity of the ball with respect to the space frame. Thus the component of $X_{\mathrm{nh}}$ which is tangent to the group can be written as

$$T\tau.X_{\mathrm{nh}} = \sum l_\alpha\xi_\alpha + \sum g_{(\lambda,a)}\zeta_{(\lambda,a)} \tag{3.9}$$

where $\tau : TK = K\times\mathfrak{k}\to K$. This is just the reformulation of $u = s^{-1}s'$. Moreover, it will be convenient to define

$$\widetilde{l}_\alpha := l_\alpha\circ\mu_0 \quad \text{and} \quad G_{(\lambda,a)} := g_{(\lambda,a)}\circ\mu_0 : TK\to\mathbb{R}$$

where we view $\mu_0$ as a bundle endomorphism $TK = K\times\mathfrak{k}\to K\times\mathfrak{k}^* =_{\langle.,.\rangle} K\times\mathfrak{k}$. The Liouville one-form can now be written as

$$\theta^K = \sum\widetilde{l}_\alpha\rho^\alpha + \sum G_{(\lambda,a)}\eta^{(\lambda,a)}.$$

With this notation we derive the following simple formula for the connection form $\mathcal{A}$ which will be central to the subsequent. Namely,

$$\mathcal{A} = \sum_{\lambda\in\Phi}\lambda(w_0)\eta^{(\lambda,a)}e_{(\lambda,a)} \tag{3.10}$$

where

$$\Phi := \big\{\lambda\in\Sigma_+ :\ \lambda(w_0)\neq 0\big\} \tag{3.11}$$

is the set of relevant roots. For reference we also note that

$$\mathfrak{h} = \mathfrak{m}\oplus\bigoplus_{\lambda(w_0)=0}\operatorname{span}\{Z_{(\lambda,a)}\}.$$

This subalgebra is reminiscent of the $\mathfrak{k}$-part of the Langlands decomposition of a parabolic subalgebra of $\mathfrak{g}$. Indeed the possible choices of $\Phi$ correspond in a one-to-one fashion to the possible parabolics in $\mathfrak{g}$. In fact, according to Knapp [17, Section VII.7] every parabolic is specified by a set $\Gamma \subset \Sigma$ which contains $\Sigma_+$. The correspondence is now given by setting $\Gamma = \Sigma \setminus (-\Phi)$. Equivalently $\Gamma$ can be defined by requiring the identity $-(\Gamma \cap \Sigma_-) = \Sigma_+ \setminus \Phi$. We will make use of this observation in Section 4.4.

The induced metric

$$\mu_0(u_1, u_2)|_s = \langle \mathbb{I}u_1, u_2\rangle + \big\langle \big[w_0, \mathrm{Ad}_s(u_1)\big], \big[w_0, \mathrm{Ad}_s(u_2)\big]\big\rangle$$

becomes in this notation

$$\mu_0 = \langle \mathbb{I}.,.\rangle + \sum_{\lambda\in\Phi} \lambda(w_0)^2 \eta^{(\lambda,a)} \otimes \eta^{(\lambda,a)}.$$

Alternatively $\mu_0$ can be considered as an endomorphism

$$\mu_0 = \mathbb{I} + \mathcal{A}^*\mathcal{A} = \mathbb{I} + \sum \lambda(w_0)^2 g_{(\lambda,a)} \zeta_{(\lambda,a)}$$

of $TK = K \times \mathfrak{k}$, and we will use the same symbol $\mu_0$ to denote both instances. The compressed Hamiltonian is

$$\mathcal{H}_c(s,u) = \frac{1}{2}\langle \mathbb{I}u, u\rangle + \frac{1}{2}\sum_{\lambda\in\Phi} \lambda(w_0)^2 g_{(\lambda,a)}(s,u)^2.$$

Furthermore,

$$\Omega_{\mathrm{nh}} = \Omega^K + \langle \mathcal{A}, d\mathcal{A}\rangle = \Omega^K + \sum_{\lambda\in\Phi} \lambda(w_0)^2 g_{(\lambda,a)}\, d\eta^{(\lambda,a)}$$

and a formula for $d\eta^{(\lambda,a)}$ is given in (3.12).

**Lemma 3.1.** *$\langle \mathcal{A}, d\mathcal{A}\rangle(X_{\mathrm{nh}}, \zeta_Y) = 0$ for all $Y \in \mathfrak{h}$.*

**Proof.** This follows either form direct calculation using the above formula. Alternatively one can use that $\mathcal{H}_c$ is $H$-invariant and that $J_H$ is a preserved quantity. Thus $\Omega_{\mathrm{nh}}(X_{\mathrm{nh}}, \zeta_Y) = 0 = -\Omega^K(X_{\mathrm{nh}}, \zeta_Y)$. □

The structure constants are defined by $c^\alpha_{(\lambda,a)(\mu,b)} = \langle Y_\alpha, [Z_{(\lambda,a)}, Z_{(\mu,b)}]\rangle$ etc.

**Lemma 3.2.** *Let $\lambda, \mu, \nu \in \Sigma_+$ and $1 \leqslant \alpha \leqslant \dim \mathfrak{m}$.*

(1) *If $c^\alpha_{(\lambda,a)(\mu,b)} \neq 0$ then $\lambda = \mu$.*
(2) *If $c^{(\lambda,a)}_{(\mu,b)(\nu,c)} \neq 0$ then $\lambda = \pm\mu \pm \nu$.*

**Proof.** To see this one notices that the $Z_{(\lambda,a)}$ can be written as $Z_{(\lambda,a)} = -X^a_{-\lambda} - \theta X^a_{-\lambda} \in \mathfrak{k}$ for a suitably normalized orthogonal basis $X^a_\lambda$ of $\mathfrak{g}$ consisting of root vectors. (Recall that $\theta$ denotes the Cartan involution.) The assertions now follow directly from the properties of the root system with respect to the action of the Lie bracket together with the fact that $Y_\alpha \in \mathfrak{m} = \mathfrak{g}_0 \cap \mathfrak{k}$. □

Taking into account the change of sign in the map $\zeta : \mathfrak{k} \to \mathfrak{X}(K)$, $[X,Y] \mapsto \zeta_{[X,Y]} = -[\zeta_X, \zeta_Y]$ we obtain the formulas

$$\begin{aligned} d\rho^\alpha &= \frac{1}{2}\sum c^\alpha_{\beta\gamma}\rho^\beta \wedge \rho^\gamma + \frac{1}{2}\sum c^\alpha_{(\lambda,a)(\mu,b)}\eta^{(\lambda,a)} \wedge \eta^{(\mu,b)}, \\ d\eta^{(\lambda,a)} &= \sum c^{(\lambda,a)}_{\beta(\lambda,b)}\rho^\beta \wedge \eta^{(\lambda,b)} + \frac{1}{2}\sum c^{(\lambda,a)}_{(\mu,b)(\nu,c)}\eta^{(\mu,b)} \wedge \eta^{(\nu,c)}. \end{aligned} \tag{3.12}$$

*3.5. The preserved measure*

The $n$-dimensional Chaplygin ball problem has a preserved measure which was found by Fedorov and Kozlov [13]. We consider the Chaplygin system $(TK, \Omega_{\mathrm{nh}}, \mathcal{H}_c)$ introduced above and show that the existence of a preserved measure continues to hold.

Let $d = \dim K$ and $g := \det \mu_0$ where we view $\mu_0$ as a function $K \to \mathrm{End}(\mathfrak{k})$. Consider the volume form

$$\mathrm{vol} = \mathrm{vol}\big(\mu_0 \times \langle .,.\rangle\big) = \frac{1}{d!}\sqrt{g}\big(\Omega^K\big)^d$$

on $TK = K \times \mathfrak{k}$.

**Lemma 3.3.** *Let $f: K \to \mathbb{R}_{>0}$. Then*

$$L_{X_{\mathrm{nh}}}\big(f\big(\Omega^K\big)^d\big) = d! L_{X_{\mathrm{nh}}}\big(f g^{-\frac{1}{2}}\,\mathrm{vol}\big) = 0 \quad \Longleftrightarrow \quad d(\log f) X_{\mathrm{nh}} = -\sum \frac{\partial}{\partial p_i}\langle J, K\rangle\left(X_{\mathrm{nh}}, \frac{\partial}{\partial q^i}\right)$$

*where $(q^i, p_i)$ are canonical coordinates on $TK$.*

**Proof.** $L_{X_{\mathrm{nh}}}(f g^{-\frac{1}{2}}\,\mathrm{vol}) = d(f g^{-\frac{1}{2}}).X_{\mathrm{nh}}\,\mathrm{vol} + f g^{-\frac{1}{2}}\,\mathrm{div}_{\mathrm{vol}} X_{\mathrm{nh}}\,\mathrm{vol}$. Thus $f$ is a preserved density corresponding to the volume $(\Omega^K)^d = \Omega^d_{\mathrm{nh}}$ iff

$$d(\log f).X_{\mathrm{nh}} = -\,\mathrm{div}_{\mathrm{vol}} X_{\mathrm{nh}} + \frac{1}{2} d(\log g).X_{\mathrm{nh}}.$$

Now,

$$\mathrm{div}_{\mathrm{vol}} X_{\mathrm{nh}} = \sum\left(\frac{\partial}{\partial q^i}\frac{\partial \mathcal{H}_c}{\partial p^i} + \frac{\partial}{\partial p_i}\left(-\frac{\partial \mathcal{H}_c}{\partial q^i} + \langle J, K\rangle\left(X_{\mathrm{nh}}, \frac{\partial}{\partial q^i}\right)\right)\right) + \frac{1}{2} d(\log g).X_{\mathrm{nh}}$$

where we use the general formula for the divergence and the equations of motion of the almost Hamiltonian system. □

By (3.9) we can identify $d(\log f) X_{\mathrm{nh}}$ with the function $TK \to \mathbb{R}$ that corresponds to the one-form $d(\log f)$ on $K$. In particular, $f$ is unique up to multiplication by positive constants. We will use the notation

$$f := \frac{1}{\sqrt{g}}$$

and refer to this (after Proposition 3.4) as the preserved density of the system. When $G = \mathrm{SO}(n,1)$ and we are dealing with the $n$-dimensional Chaplygin ball then $f$ coincides with the density found by [13]. Using the rule for the differential of the determinant, $\zeta_{(\lambda,a)} \det \mu_0 = \det(\mu_0)\,\mathrm{Tr}(\mu_0^{-1}\zeta_{(\lambda,a)}\mu_0)$, one obtains

$$d(\log f).\zeta_{(\lambda,a)} = -\sum_{(\mu,b)} \mu(w_0)^2 \big\langle \mu_0^{-1}[\zeta_{(\lambda,a)}, \zeta_{(\mu,b)}], \zeta_{(\mu,b)}\big\rangle \tag{3.13}$$

where the notation is as in Section 3.4.

**Proposition 3.4** *(The preserved measure). $L_{X_{\mathrm{nh}}}(f(\Omega^K)^d) = 0$.*[2]

**Proof.** Of course, we will use Lemma 3.3. Choose coordinates $q^i$ with $i \in J \cup I$ around a point in $K$ such that $\frac{\partial}{\partial q^i}(s) = \xi_\alpha$ for all $i \in J$ where $i$ corresponds to $\alpha$, and $\frac{\partial}{\partial q^i}(s) = \zeta_{(\lambda,a)}(s)$ for all $i \in I$ where $i$ corresponds to $(\lambda, a)$. The conjugate momenta corresponding to $i = (\lambda, a)$ are then given by $\frac{\partial}{\partial p_i} = (0, \mu_0^{-1}\zeta_{(\lambda,a)})$. The first equality in the following calculation uses Lemma 3.1.

$$\begin{aligned}
&\sum_{i\in I\cup J} \frac{\partial}{\partial p_i}\langle J, K\rangle\left(X_{\mathrm{nh}}, \frac{\partial}{\partial q^i}\right)\\
&\quad= \sum \frac{\partial}{\partial p_{(\lambda,a)}}\langle J, K\rangle\left(X_{\mathrm{nh}}, \frac{\partial}{\partial q^{(\lambda,a)}}\right)\\
&\quad= \sum \frac{\partial}{\partial p_{(\lambda,a)}} \mu(w_0)^2 g_{(\mu,b)} d\eta^{(\mu,b)}\left(\sum (l_\alpha \xi_\alpha + g_{(\nu,c)}\zeta_{(\nu,c)}), \frac{\partial}{\partial q^{(\lambda,a)}}\right)\\
&\quad= -\sum \frac{\partial}{\partial p_{(\lambda,a)}} \mu(w_0)^2 g_{(\mu,b)} c^{(\mu,b)}_{\alpha(\lambda,a)} l_\alpha - \sum \frac{\partial}{\partial p_{(\lambda,a)}} \mu(w_0)^2 g_{(\mu,b)} c^{(\mu,b)}_{(\nu,c)(\lambda,a)} g_{(\nu,c)}\\
&\quad= -\sum \mu(w_0)^2 \big\langle \zeta_{(\mu,b)}, \mu_0^{-1}\zeta_{(\lambda,a)}\big\rangle l_\alpha c^{(\mu,b)}_{\alpha(\lambda,a)} - \sum \mu(w_0)^2 g_{(\mu,b)} \big\langle \xi_\alpha, \mu_0^{-1}\zeta_{(\lambda,a)}\big\rangle c^{(\mu,b)}_{\alpha(\lambda,a)}\\
&\qquad - \sum \mu(w_0)^2 \big\langle \zeta_{(\mu,b)}, \mu_0^{-1}\zeta_{(\lambda,a)}\big\rangle g_{(\nu,c)} c^{(\mu,b)}_{(\nu,c)(\lambda,a)} - \sum \mu(w_0)^2 g_{(\mu,b)} \big\langle \zeta_{(\nu,c)}, \mu_0^{-1}\zeta_{(\lambda,a)}\big\rangle c^{(\mu,b)}_{(\nu,c)(\lambda,a)}\\
&\quad= \sum \mu(w_0)^2 g_{(\mu,b)} \big\langle [\zeta_{(\mu,b)}, \zeta_{(\lambda,a)}]^{\xi}, \mu_0^{-1}\zeta_{(\lambda,a)}\big\rangle + \sum \mu(w_0)^2 g_{(\nu,c)} \big\langle \zeta_{(\mu,b)}, \mu_0^{-1}[\zeta_{(\nu,c)}, \zeta_{(\mu,b)}]^{\zeta}\big\rangle\\
&\quad= \sum \mu(w_0)^2 g_{(\lambda,a)} \big\langle [\zeta_{(\lambda,a)}, \zeta_{(\mu,b)}], \mu_0^{-1}\zeta_{(\mu,b)}\big\rangle = -d(\log f) X_{\mathrm{nh}},
\end{aligned}$$

[2] I am grateful to the referees for pointing out that this also follows from the expression of the density of an invariant measure given by Fedorov [11,12].

where we have used that $c^{(\mu,b)}_{\alpha(\lambda,a)} = c^{(\mu,b)}_{\alpha(\lambda,a)}\delta_{\lambda\mu}$. Further, $(\_)^{\xi}$, $(\_)^{\zeta}$ denote the projections onto the subspaces spanned by $\xi_\alpha$, $\zeta_{(\lambda,a)}$ respectively. Finally note that $f$ is a pull-back of a function on the base $K$ and we have made use of some formulas from Section 3.4. $\square$

**Remark.** When $\mathcal{D}$ is mechanical, that is orthogonal to the vertical bundle via $\mu$, then we know that compression equals symplectic reduction at 0. (This case can be realized by setting $w_0 = 0$.) Thus $X_{\mathrm{nh}}$ is the reduced Hamiltonian vector field and as such it preserves $(\Omega^K)^d$. This is consistent with the above since, now, $J = 0$ whence $\frac{\partial}{\partial p_i}\langle J, K\rangle(X_{\mathrm{nh}}, \frac{\partial}{\partial q^i}) = 0$ and thus $\operatorname{div}_{\mathrm{vol}} X_{\mathrm{nh}} = \frac{1}{2}d(\log g)X_{\mathrm{nh}}$. This can be used as a roundabout way to reach the obvious conclusion $f = 1$.

### 3.6. Truncation

The system $(TK, \Omega_{\mathrm{nh}}, \mathcal{H}_{\mathrm{c}})$ is $H$-invariant and has a preserved quantity which is just the standard momentum map $J_H : TK \to \mathfrak{h}^*$. Thus it is natural to ask whether this set of data can be reduced to $J_H^{-1}(\mathcal{O})/H \cong J_H^{-1}(\alpha)/H_\alpha$ where $\mathcal{O}$ is the $\mathrm{Ad}^*(H)$-orbit through $\alpha \in \mathfrak{h}^*$ and $H_\alpha$ is the stabilizer of $\alpha$ in the group. The answer to this question is negative: the momentum map equation

$$i(\zeta_Y)\Omega_{\mathrm{nh}} = d\langle J_H, Y\rangle$$

with $Y \in \mathfrak{h}$ is *not* satisfied in general. Thus the restriction of $\Omega_{\mathrm{nh}}$ to $J_H^{-1}(\alpha)$ is not horizontal in general whence it cannot induce a form on the reduced space. The situation here is identical with that of [14]. Thus by [14, Theorem 3.3] we also know that there is a solution: the form $\langle J, K\rangle$ is not optimal for describing the system; it sees vertical directions that are inessential (Lemma 3.1) whence it needs to be replaced by an entity which is horizontal. As an aside, we remark that [14, Theorem 3.3] is only of philosophical value here: It does tell us that a horizontal perturbation $\Lambda$ of $\langle J, K\rangle$ such that $i(X_{\mathrm{nh}})(\Lambda - \langle J, K\rangle) = 0$ exists but it does not provide a very practical way for finding one. The cited theorem yields a two-form which is well defined only on a dense open subset of $TK$. To get a form which is globally well defined we propose definition (3.14) below which has been found by trial and error, and it seems like this form cannot be constructed according to the recipe of [14, Theorem 3.3]. Nevertheless we retain the name truncation since the idea of replacing $\langle J.K\rangle$ by $\Lambda$ is to chop off the vertical directions – and since vertical vectors and horizontal forms are canonically defined we do not need the notion of a connection on $TK$ to make sense of this.

Let

$$\Lambda := -\frac{1}{2}\sum \lambda(w_0)^2 c^{\alpha}_{(\lambda,a)(\lambda,b)} l_\alpha \eta^{(\lambda,a)} \wedge \eta^{(\lambda,b)} - \frac{1}{2}\sum_{\lambda\notin\Phi,\mu,\nu\in\Phi} \mu(w_0)^2 c^{(\lambda,a)}_{(\mu,b)(\nu,c)} g_{(\lambda,a)} \eta^{(\mu,b)} \wedge \eta^{(\nu,c)}$$
$$+ \frac{1}{2}\sum_{\mu,\nu\in\Phi} \lambda(w_0)^2 c^{(\lambda,a)}_{(\mu,b)(\nu,c)} g_{(\lambda,a)} \eta^{(\mu,b)} \wedge \eta^{(\nu,c)}. \tag{3.14}$$

Notice that the coefficients of the second summand of $\Lambda$ are skew-symmetric: when $c^{(\lambda,a)}_{(\mu,b)(\nu,c)} \neq 0$ with $\lambda \notin \Phi$ and $\mu, \nu \in \Phi$ then $\mu(w_0)^2 = \nu(w_0)^2$ by Lemma 3.2. One makes a choice here: in principle one could add to $\Lambda$ any $\tau$-semi-basic $H$-basic two-from which vanishes upon contraction with $X_{\mathrm{nh}}$. However, in the proof of Theorem 3.6 we will see that this choice for $\Lambda$ seems to be preferred by the problem at hand.

The following theorem generalizes [14, Theorem 4.1] but the situation here is more tricky: In [14, Theorem 4.1] we defined the perturbed form to be $\langle L, \mathrm{Curv}^\omega\rangle$ where $L = \sum l_\alpha Y_\alpha$ and $\mathrm{Curv}^\omega$ is the curvature of $\omega = \sum \rho^\alpha Y_\alpha \in \Omega^1(K, \mathfrak{h})$. However, while $\Lambda$ and $-\langle L, \mathrm{Curv}^\omega\rangle$ coincide in the case of the Chaplygin ball, $i(X_{\mathrm{nh}})\langle J, K\rangle = -i(X_{\mathrm{nh}})\Lambda \neq i(X_{\mathrm{nh}})\langle L, \mathrm{Curv}^\omega\rangle$ in general.

**Theorem 3.5** *(Truncation). The system $(TK, \widetilde{\Omega}, \mathcal{H}_{\mathrm{c}})$ where*

$$\widetilde{\Omega} := \Omega^K + \Lambda$$

*has the following properties.*

(1) *$\widetilde{\Omega}$ is non-degenerate and $H$-basic.*
(2) *$i(X_{\mathrm{nh}})\widetilde{\Omega} = d\mathcal{H}_{\mathrm{c}}$.*
(3) *$i(\zeta_Y)\widetilde{\Omega} = d\langle J_H, Y\rangle$ for all $Y \in \mathfrak{h}$.*

**Proof.** Non-degeneracy is clear. Observe that

$$\left(\frac{1}{2}\sum \lambda(w_0)^2 c^{\alpha}_{(\lambda,a)(\lambda,b)} l_\alpha \eta^{(\lambda,a)} \wedge \eta^{(\lambda,b)} + \frac{1}{2}\sum_{\lambda\notin\Phi,\mu,\nu\in\Phi} \mu(w_0)^2 c^{(\lambda,a)}_{(\mu,b)(\nu,c)} g_{(\lambda,a)} \eta^{(\mu,b)} \wedge \eta^{(\nu,c)}\right)_{(s,u)} (u_1', u_2')$$
$$= \langle [\mathrm{ad}(w_0)^2 s.u_1', s.u_2']^{\mathfrak{h}}, s.u\rangle$$

where $(\_)^{\mathfrak{h}}$ denotes projection onto $\mathfrak{h}$. Clearly this is $H$-invariant since, by definition, $H$ commutes with $\mathrm{ad}(w_0)$. On the other hand,

$$\left(\frac{1}{2}\sum_{\mu,\nu\in\Phi}\lambda(w_0)^2 c^{(\lambda,a)}_{(\mu,b)(\nu,c)} g_{(\lambda,a)}\eta^{(\mu,b)}\wedge\eta^{(\nu,c)}\right)_{(s,u)}(u_1',u_2') = \langle [su_1', su_2']^{\mathfrak{h}^\perp}, \mathrm{ad}(w_0)^2 su\rangle$$

which is also $H$-independent. Thus $\Lambda$ is $H$-invariant. Obviously $\Lambda$ is also $H$-horizontal since the $\eta^{(\mu,b)}$ for $\mu\in\Phi$ are horizontal by construction. To see that $\widetilde{\Omega}$ produces the right dynamics note simply that

$$\begin{aligned}\langle\mathcal{A}, d\mathcal{A}\rangle(X_{\mathrm{nh}},\zeta_{(\nu,c)}) &= \sum\mu(w_0)^2 g_{(\mu,b)} l_\alpha c^{(\mu,b)}_{\alpha(\nu,c)}\delta_{\mu,\nu} + \sum\lambda(w_0)^2 g_{(\lambda,a)} g_{(\mu,b)} c^{(\lambda,a)}_{(\mu,b)(\nu,c)}\\ &= -\sum\mu(w_0)^2 g_{(\mu,b)} l_\alpha c^{\alpha}_{(\mu,b)(\nu,c)}\delta_{\mu,\nu} - \sum_{\mu\notin\Phi}\lambda(w_0)^2 g_{(\lambda,a)} g_{(\mu,b)} c^{(\mu,b)}_{(\lambda,a)(\nu,c)}\\ &\quad + \sum_{\mu\in\Phi}\lambda(w_0)^2 g_{(\lambda,a)} g_{(\mu,b)} c^{(\lambda,a)}_{(\mu,b)(\nu,c)}\\ &= \Lambda(X_{\mathrm{nh}},\zeta_{(\nu,c)})\end{aligned}$$

for all $\nu\in\Phi$. Finally, we can use the momentum map equation with respect to $\Omega^K$ and horizontality of $\Lambda$ to obtain the momentum map equation for $\widetilde{\Omega}$. $\square$

Thus one can pass to the description $(TK,\widetilde{\Omega},\mathcal{H}_c)$ of the system and do (almost) Hamiltonian reduction with respect to the symmetry group $H$ and the momentum map $J_H$. Using the mechanical connection associated to $\mu_0$ the reduced space can be realized as a symplectic fiber bundle over $T^*(K/H)$ with fiber a coadjoint orbit $\mathcal{O}\subset\mathfrak{h}^*$ whence Theorem 2.3 is applicable.

### 3.7. Cases of Hamiltonization for multidimensional systems

In this setting multidimensional means that the dimension of the ultimate reduced configuration space $K/H$ is greater than 2.

By Theorem 3.5 we regard the compressed system as being described by the almost Hamiltonian system $(TK,\widetilde{\Omega},\mathcal{H}_c)$ and we recall that we identify $TK = T^*K$ via the induced metric $\mu_0$. According to Proposition 3.4 this system admits a preserved measure: $L_{X_{\mathrm{nh}}}(f\Omega_K^d)=0$ where $d=\dim K$ and

$$f = (\det\mu_0)^{-\frac{1}{2}}.$$

(From Lemma 3.3 it is not hard to see that $f$ factors also to a density on $T^*(K/H) = J_H^{-1}(0)/H$.) Let $\iota: J_H^{-1}(\alpha)\hookrightarrow TK$, $\alpha\in\mathfrak{h}^*$, $\pi: J_H^{-1}(\alpha)\twoheadrightarrow H_H^{-1}(\alpha)/H_\alpha$ where $H_\alpha$ is the isotropy subgroup of $\alpha$ in $H$, and

$$F := f^{\frac{1}{m-1}}$$

with $m=\frac{1}{2}\dim J_H^{-1}(\alpha)/H_\alpha$. Then the reduced almost symplectic form $\sigma$ is characterized by the equation $\pi^*\sigma = \iota^*\widetilde{\Omega}$. Note that we may use the metric $\mu_0$ to identify

$$J_H^{-1}(\alpha)/H_\alpha \cong J_H^{-1}(\mathcal{O})/H \cong T^*(K/H)\times_{K/H}(K\times_H\mathcal{O}) \tag{3.15}$$

where $\mathcal{O}$ is the $\mathrm{Ad}^*(H)$-orbit through $\alpha$ and $\sigma$ is of the form 'canonical plus magnetic plus semi-basic' with the semi-basic part linear in the fibers whence we are in the situation of Theorem 2.3. Up to multiplication by positive constants, the only possible candidate for a conformal factor of $\sigma$ will be $F$ which we can view as a function $K/H_\alpha\to\mathbb{R}_{>0}$. (Because $\delta\sigma = -(m-1)d\log F$ in this case.) However, we do not know beforehand whether $\psi = d\sigma - \frac{1}{m-1}\delta\sigma\wedge\sigma$ vanishes. This only follows a posteriori since we prove in Theorem 3.6 that $\sigma$ is exact and we know from Lemma 2.2 that $\psi=0$ is a necessary condition for $\sigma$ to be (locally conformally) closed. Thus the only practical value of Theorem 2.3 here is that it says where to look for a conformal factor, and even this value is limited by the fact that the same conclusion can be reached by considering Lemma 2.2.

It is a trivial observation to note that $F$ indeed is a conformal factor if and only if

$$\iota^* d\Lambda = -\iota^*\big(d(\log F)\wedge\widetilde{\Omega}\big). \tag{3.16}$$

Analyzing this equation for $\alpha=0$ leads to the following result.

**Theorem 3.6** *(Hamiltonization at 0 momentum). Let $m=\dim K/H$. The induced almost symplectic structure $\sigma$ on $J_H^{-1}(0)/H\cong T^*(K/H)$ is Hamiltonizable if and only if the metric tensor $\mu_0 = \mathbb{I}+\sum\lambda(w_0)^2 g_{(\lambda,a)}\zeta_{(\lambda,a)}:\mathfrak{k}\to\mathfrak{k}$ satisfies*

$$\begin{aligned}&\big\langle s\mu_0(s)^{-1}s^{-1}Z_{(\kappa,d)},\big[\mathrm{ad}(w_0)^2Z_{(\mu,b)},Z_{(\nu,c)}\big]^{\mathfrak h}-\mathrm{ad}(w_0)^2[Z_{(\mu,b)},Z_{(\nu,c)}]\big\rangle\\&\quad=\frac{1}{m-1}\sum\big\langle s\mu_0(s)^{-1}s^{-1}Z_{(\lambda,a)},\big[Z_{(\mu,b)},\mathrm{ad}(w_0)^2Z_{(\lambda,a)}\big]\delta_{(\nu,c),(\kappa,d)}-\big[Z_{(\nu,c)},\mathrm{ad}(w_0)^2Z_{(\lambda,a)}\big]\delta_{(\mu,b),(\kappa,d)}\big\rangle\end{aligned}\tag{3.17}$$

*for all $\kappa,\mu,\nu\in\Phi$. Here $(\_)^{\mathfrak h}$ denotes the projection onto $\mathfrak h$ with respect to the* Ad*-invariant inner product. As usual $\delta_{(\nu,c),(\kappa,d)}$ is* 1 *if $(\nu,c)=(\kappa,d)$ and* 0 *else. Moreover, if this condition is satisfied then*

$$\pi^*(F\sigma)=\iota^*(F\widetilde{\Omega})=-\iota^*d\Big(F\sum G_{(\lambda,a)}\eta^{(\lambda,a)}\Big)=-\pi^*d\big(F\theta^{K/H}\big)\tag{3.18}$$

*where $\theta^{K/H}$ is the Liouville one-form on $T^*(K/H)$. That is, $F\sigma$ is even exact.*

We remark that $\mathbb{I}=1$ implies that $s\mu_0(s)s^{-1}=\mu_0(e)$. Notice that the condition simplifies when $|\Phi|=1$ as is the case for the $n$-dimensional Chaplygin ball. When $\dim K/H=2$ then the condition is empty in agreement with the Chaplygin multiplier theorem.

**Proof.** Let us first prove that (3.17) implies (3.18). Since $\iota^*(F\widetilde{\Omega})=\iota^*(-Fd(\sum G_{(\lambda,a)}\eta^{(\lambda,a)})+F\Lambda)$ it suffices to show that $\Lambda=-d(\log F)\wedge\sum G_{(\lambda,a)}\eta^{(\lambda,a)}$ along $J_H^{-1}(0)$.[3] Consider an element $(s,u)\in J_H^{-1}(0)$ with $u=\mu_0^{-1}\zeta_{(\kappa,d)}$ where $\kappa\in\Phi$. (Notice that we sometimes drop the base point $s$ in order not to make the notation too cumbersome.) Then with $\mu,\nu\in\Phi$ we have

$$\begin{aligned}\Lambda_{(s,u)}(\zeta_{(\mu,b)},\zeta_{(\nu,c)})&=-\sum_\alpha\mu(w_0)^2\delta_{\mu,\nu}c^{\alpha}_{(\mu,b)(\nu,c)}\big\langle Y_\alpha s\mu_0^{-1}s^{-1}Z_{(\kappa,d)}\big\rangle\\&\quad-\sum_{\lambda\notin\Phi}\mu(w_0)^2c^{(\lambda,a)}_{(\mu,b)(\nu,c)}\big\langle Z_{(\lambda,a)},s\mu_0^{-1}s^{-1}Z_{(\kappa,d)}\big\rangle\\&\quad+\sum_{\lambda\in\Phi}\lambda(w_0)^2c^{(\lambda,a)}_{(\mu,b)(\nu,c)}\big\langle Z_{(\lambda,a)},s\mu_0^{-1}s^{-1}Z_{(\kappa,d)}\big\rangle\\&=-\big\langle\big[\mathrm{ad}(w_0)^2Z_{(\mu,b)},Z_{(\nu,c)}\big]^{\mathfrak m},s\mu_0^{-1}s^{-1}Z_{(\kappa,d)}\big\rangle\\&\quad-\big\langle\big[\mathrm{ad}(w_0)^2Z_{(\mu,b)},Z_{(\nu,c)}\big]^{\mathfrak h\cap\mathfrak m^\perp},s\mu_0^{-1}s^{-1}Z_{(\kappa,d)}\big\rangle\\&\quad+\big\langle\mathrm{ad}(w_0)^2[Z_{(\mu,b)},Z_{(\nu,c)}],s\mu_0^{-1}s^{-1}Z_{(\kappa,d)}\big\rangle\\&=-\big\langle s\mu_0^{-1}s^{-1}Z_{(\kappa,d)},\big[\mathrm{ad}(w_0)^2Z_{(\mu,b)},Z_{(\nu,c)}\big]^{\mathfrak h}-\mathrm{ad}(w_0)^2[Z_{(\mu,b)},Z_{(\nu,c)}]\big\rangle.\end{aligned}$$

As before, the superscript $(\_)^{\mathfrak m}$ denotes projection onto $\mathfrak m$ with respect to the Ad-invariant inner product $\langle.,.\rangle$. On the other hand,

$$\begin{aligned}-\Big(d(\log F)\wedge\sum G_{(\lambda,c)}\eta^{(\lambda,a)}\Big)_{(s,u)}(\zeta_{(\mu,b)},\zeta_{(\nu,c)})&=\frac{1}{m-1}\sum\lambda(w_0)^2\big\langle\mu_0^{-1}[\zeta_{(\mu,b)},\zeta_{(\lambda,a)}],\zeta_{(\lambda,a)}\big\rangle\delta_{(\kappa,d),(\nu,c)}\\&\quad-\frac{1}{m-1}\sum\lambda(w_0)^2\big\langle\mu_0^{-1}[\zeta_{(\nu,c)},\zeta_{(\lambda,a)}],\zeta_{(\lambda,a)}\big\rangle\delta_{(\kappa,d),(\mu,b)}\\&=-\frac{1}{m-1}\sum\big\langle s\mu_0^{-1}s^{-1}Z_{(\lambda,a)},\big[Z_{(\mu,b)},\mathrm{ad}(w_0)^2Z_{(\lambda,a)}\big]\big\rangle\delta_{(\nu,c),(\kappa,d)}\\&\quad+\frac{1}{m-1}\sum\big\langle s\mu_0^{-1}s^{-1}Z_{(\lambda,a)},\big[Z_{(\nu,c)},\mathrm{ad}(w_0)^2Z_{(\lambda,a)}\big]\big\rangle\delta_{(\mu,b),(\kappa,d)}.\end{aligned}$$

Since the two-forms in question are semi-basic and linear in the fibers this proves that they are equal along the 0 level set of $J_H$. Note also that the pull-back of the Liouville one-form on $T^*(K/H)$ equals $\iota^*\sum G_{(\lambda,a)}\eta^{(\lambda,a)}=\iota^*\sum_{\lambda\in\Phi}G_{(\lambda,a)}\eta^{(\lambda,a)}$. To see that the condition is also necessary one evaluates Eq. (3.16) on a triple of the form $(\zeta_{(\mu,b)},\zeta_{(\nu,c)},\frac{\partial}{\partial G_{(\kappa,d)}}=(0,\mu_0^{-1}\zeta_{(\kappa,d)}))$. The resulting calculation is very similar to the one above. □

## 4. Examples

This section contains examples of the class of non-holonomic systems introduced in the previous section. We continue all the notation from above, most of which has been introduced in Section 3.4. In particular, $\Sigma$ will be the set of restricted roots associated to a pair $(\mathfrak g,\mathfrak a)$ and $\Sigma_+\subset\Sigma$ a choice of positive roots. Then the associated root space decomposition is $\mathfrak g=\mathfrak g_0\oplus\bigoplus_{\lambda\in\Sigma}\mathfrak g_\lambda$ where $\mathfrak g_0=\mathfrak m\oplus\mathfrak a$. Moreover, we choose an orthonormal system $Y_\alpha$ and $Z_{(\lambda,a)}$, that is adapted to the decomposition $\mathfrak k=\mathfrak m\oplus\mathfrak m^\perp$, and an orthonormal basis $e_{(\lambda,a)}$ of $\mathfrak a^\perp\cap\mathfrak p$. We will in each example fix an element $w_0\in\mathfrak a$ and consider the set $\Phi:=\{\lambda\in\Sigma_+:\ \lambda(w_0)\neq0\}$.

[3] We view this as 'compelling evidence' that the choice for $\Lambda$ in (3.14) is in a sense optimal.

*4.1.* $\mathrm{SO}(n,1)$, *Hamiltonization of Chaplygin's ball*

According to Section 3.3 the above Theorem 3.6 should have some bearing on the $n$-dimensional Chaplygin ball system with angular momentum $\alpha = 0$. Moreover, for this system there is only 1 positive root (and we assume that $\lambda(w_0) = 1$ for this root) whence condition (3.17) simplifies to

$$\big\langle s\mu_0^{-1}s^{-1}Z_d, [Z_b, Z_c]\big\rangle = \frac{1}{m-1}\sum_a \big\langle s\mu_0^{-1}s^{-1}Z_a, [Z_b, Z_a]\delta_{cd} - [Z_c, Z_a]\delta_{bd}\big\rangle. \tag{4.19}$$

Let us decompose $s\mu_0^{-1}s^{-1}Z_d = Y(d) + Z(d) \in \mathfrak{h} \oplus \mathfrak{h}^\perp$. Then (4.19) does not induce a restriction on $Z(d) \in \mathfrak{h}^\perp$. On the other hand, $\langle Y(d), [Z_b, Z_c]\rangle = 0$ if $b \neq d \neq c$, and $\langle Y(d), [Z_b, Z_d]\rangle = \frac{1}{m-1}\sum_a \langle s\mu_0^{-1}s^{-1}Z_a, [Z_b, Z_a]\rangle =: \mathcal{M}_b(s)$. Thus $Y(d) = \sum_b \mathcal{M}_b(s)[Z_b, Z_d]$ where $\mathcal{M}_b(s)$ depends only on $s$ and $b$. Using that $\mu = \mathbb{I} + \mathcal{A}^*\mathcal{A}$, $\mathcal{A}^*\mathcal{A}|s^{-1}\mathfrak{h} = 0$ and $\mathcal{A}^*\mathcal{A}|s^{-1}\mathfrak{h}^\perp = \mathrm{id}|s^{-1}\mathfrak{h}^\perp$ implies that the system is Hamiltonizable at the $T^*(K/H) = T^*(\mathrm{SO}(n)/\mathrm{SO}(n-1))$-level if and only if $\mathbb{I}$ satisfies

$$s^{-1}Z_d = \mu_0 s^{-1}\big(Z(d) + Y(d)\big) = (\mathbb{I}+1)s^{-1}Z(d) + \mathbb{I}\sum_b \mathcal{M}_b(s)s^{-1}[Z_b, Z_d] \tag{4.20}$$

for certain $\mathcal{M}_b(s)$. We will identify $\mathfrak{so}(n)$ with $\mathbb{R}^n \wedge \mathbb{R}^n$ and hence $Z_d = e_d \wedge e_n$ and $[Z_b, Z_d] = e_b \wedge e_d$ where $e_1, \dots, e_n$ is the standard basis of $\mathbb{R}^n$. Simultaneously we revert to writing $\mathrm{Ad}(s)$ for the adjoint action of $s$ on $\mathfrak{so}(n)$.

Making the simplifying assumption that $\mathbb{I}$ is diagonal with respect to the basis $Y_\alpha, Z_a$ of $\mathfrak{k} = \mathfrak{so}(n)$ and evaluating (4.20) at $s = e$ then implies that $Z(d) = (\mathbb{I}+1)^{-1}Z_d = \varphi_d Z_d$ for some $\varphi_d > 0$. Therefore,

$$\mathbb{I}e_d \wedge e_n = \frac{1-\varphi_d}{\varphi_d}e_d \wedge e_n.$$

A choice of a number $a_n > 0$ then induces a prescription

$$\varphi_d \mapsto \frac{1-\varphi_d}{a_n} = a_d, \quad a_d \mapsto \varphi_d = 1 - a_d a_n$$

which can be taken as a motivation to define

$$\mathbb{I}e_i \wedge e_j = \frac{a_i a_j}{1 - a_i a_j}e_i \wedge e_j \quad \text{with } 0 < a_i a_j < 1 \text{ for } 1 \leqslant i,\ j \leqslant n. \tag{4.21}$$

This is the inertia tensor of Jovanovic [16, Section 4]. Another equivalent way to write (4.20) is

$$\mu_0^{-1}\mathrm{Ad}\big(s^{-1}\big)(e_d \wedge e_n) = \mathrm{Ad}\big(s^{-1}\big)Z(d) + \sum \mathcal{M}_b(s)\mathrm{Ad}\big(s^{-1}\big)(e_b \wedge e_d) \tag{4.22}$$

with the same notation as above. Going through the proof of Theorem 3 of [16] one sees that

$$\begin{aligned}&\mu_0^{-1}\mathrm{Ad}\big(s^{-1}\big)(e_d \wedge e_n)\\ &\quad = \big\langle s^{-1}e_n, A^{-1}s^{-1}e_n\big\rangle\Big(\big(-As^{-1}e_d + \big\langle A^{-1}s^{-1}e_n, s^{-1}e_n\big\rangle s^{-1}e_d\big) \wedge s^{-1}e_n + \sum \big\langle A^{-1}s^{-1}e_n, s^{-1}e_b\big\rangle s^{-1}e_b \wedge s^{-1}e_d\Big)\end{aligned}$$

where $A := \mathrm{diag}(a_1, \dots, a_n)$. With

$$Z(d) = \big\langle s^{-1}e_n, A^{-1}s^{-1}e_n\big\rangle\big(-sAs^{-1}e_d + \big\langle A^{-1}s^{-1}e_n, s^{-1}e_n\big\rangle e_d\big) \wedge e_n$$

and

$$\mathcal{M}_b(s) = \big\langle s^{-1}e_n, A^{-1}s^{-1}e_n\big\rangle\big\langle A^{-1}s^{-1}e_n, s^{-1}e_b\big\rangle$$

this clearly satisfies (4.22). Thus the system defined by the inertia tensor (4.21) is Hamiltonizable at the $T^*(K/H)$-level which reproduces the result of [16, Theorem 5]. In fact, the rescaled form is given by (3.18) whence it is not only symplectic but even exact.

*4.2.* $\mathrm{SL}(n,\mathbb{R})$

Let $\mathfrak{g} = \mathfrak{sl}(n,\mathbb{R})$. Then $\mathfrak{k} = \mathfrak{so}(n)$, $\mathfrak{p} = \{x \in \mathfrak{sl}(n,\mathbb{R}):\ x^t = x\}$, $\mathfrak{a} = \{\mathrm{diag}(w^1, \dots, w^n) \in \mathfrak{sl}(n,\mathbb{R})\}$, and $\mathfrak{m} = \{0\}$. Thus there are no internal symmetries when $w_0$ is regular. Let $f_i : \mathfrak{m} \to \mathbb{R}$, $w = \mathrm{diag}(w^1, \dots, w^n) \mapsto w^i$ for $1 \leqslant i \leqslant n$. Similarly to the Cartan case the restricted root system $\Sigma = \{\lambda_{ij} := f_i - f_j:\ i \neq j\}$ associated to $(\mathfrak{g}, \mathfrak{m})$ is of type $A_{n-1}$. A choice of a positive system is $\Sigma_+ = \{\lambda_{ij}:\ i < j\}$.

Let $n = 3$. According to (3.10) the constraints are determined by the connection form $\mathcal{A} : TK \to V = \{x \in \mathfrak{sl}(3,\mathbb{R}):\ x^t = x \text{ and } x^{ii} = 0\}$,

$$\mathcal{A}:(s,u)\mapsto \mathrm{Ad}(s)u=\widetilde{u}=\begin{pmatrix}\widetilde{u}^1\\ \widetilde{u}^2\\ \widetilde{u}^3\end{pmatrix}\mapsto -\mathrm{ad}(w_0)\widetilde{u}=-\begin{pmatrix}\lambda_3(w_0)\widetilde{u}^1\\ \lambda_1(w_0)\widetilde{u}^2\\ \lambda_2(w_0)\widetilde{u}^3\end{pmatrix}\tag{4.23}$$

where $\lambda_1=\lambda_{13}>\lambda_2=\lambda_{12}>\lambda_3=\lambda_{23}$ are the ordered positive roots. Note that $\lambda_2+\lambda_3=\lambda_1$. The basis vectors $Z_{(\lambda,a)}$, $e_{(\lambda,a)}$ introduced in Section 3.4 can now be identified with $Z_{\lambda_1}=(0,1,0)^t$, etc., considered as an element of $\mathfrak{k}\cong\mathbb{R}^3$ and $e_{\lambda_1}=Z_{\lambda_1}=(0,1,0)^t$, etc., considered as an element of $V\cong\mathbb{R}^3$.

For generic $w_0$, $Q\cong\mathrm{SO}(3)\times\mathbb{R}^3$, and the system (4.23) could be viewed as a three-axial ellipsoid with constraints moving through space. There are no internal symmetries, $\mathfrak{h}=\mathfrak{m}=0$, in this case. Using the relation $[Z_{\lambda_1},Z_{\lambda_2}]=Z_{\lambda_3}$ condition (3.17) with $\kappa=\lambda_3$, $\mu=\lambda_1$ and $\nu=\lambda_2$ thus becomes $\lambda_3(w_0)^2\langle\mu_0^{-1}Z_{\lambda_3},Z_{\lambda_3}\rangle=0$. Since $\mu_0$ is positive definite this implies $\lambda(w_0)=0$ contradicting genericity of $w_0$. Thus this case is *never* Hamiltonizable, not even for the homogeneous case $\mathbb{I}=1$. This is in contrast with the $n$-dimensional Chaplygin ball system [14, Corollary 4.3].

However, when $\lambda_2(w_0)=0$ and $\lambda_1(w_0)=\lambda_3(w_0)\neq 0$ then $H=S^1$ and we recover the 3-dimensional Chaplygin ball system.

### *4.3.* $\mathrm{Sp}(n,\mathbb{R})$

Let $G=\mathrm{Sp}(n,\mathbb{R})=\{g\in\mathrm{SL}(2n,\mathbb{R})\colon\ g^tJg=J\}$ where $J$ is the standard complex structure on $\mathbb{R}^{2n}$. Thus $\mathfrak{g}=\mathfrak{sp}(n,\mathbb{R})$ consists of matrices of the form

$$\begin{pmatrix}X_1 & X_2\\ X_3 & -X_1^t\end{pmatrix}$$

with $X_i\in\mathfrak{gl}(n,\mathbb{R})$ such that $X_2$ and $X_3$ are symmetric. The constituents of the Cartan decomposition are $\mathfrak{k}=\mathfrak{so}(2n)\cap\mathfrak{sp}(n,\mathbb{R})\cong\mathfrak{u}(n)$, $K=\mathrm{U}(n)$, and $\mathfrak{p}=\{x\in\mathfrak{g}\colon\ x^t=x\}$, and $\mathfrak{a}$ is the subspace of diagonal matrices in $\mathfrak{p}$ and $\mathfrak{m}=\{0\}$.

For convenience we will restrict now to the case $n=2$. For $i=1,2$ define $f_i\in\mathfrak{a}^*$ to be the mapping $f_i$: $\mathrm{diag}(w^1,w^2,-w^1,-w^2)\mapsto w^i$. Then the positive restricted roots associated to $(\mathfrak{g},\mathfrak{a})$ are

$$\Sigma_+=\{f_1-f_2,f_1+f_2,2f_1,2f_2\}.$$

Note that $\{f_1-f_2,2f_2\}$ forms a simple system. Since we are interested in having internal symmetries we fix an element $w_0=\mathrm{diag}(a,a,-a,-a)\in\mathfrak{a}$ with $a>0$. Thus $(f_1-f_2)(w_0)=0$, $\Phi=\{f_1+f_2,2f_1,2f_2\}$ and $\lambda(w_0)=2a$ for all $\lambda\in\Phi$. Therefore,

$$\mathcal{A}:(s,u)\mapsto \mathrm{Ad}(s)u=\widetilde{u}=\begin{pmatrix}\widetilde{u}^1\\ \widetilde{u}^2\\ \widetilde{u}^3\\ \widetilde{u}^4\end{pmatrix}\mapsto -\mathrm{ad}(w_0)\widetilde{u}=-2a\begin{pmatrix}0\\ \widetilde{u}^2\\ \widetilde{u}^3\\ \widetilde{u}^4\end{pmatrix}.$$

Further, the configuration space is $Q=K\times V\cong U(2)\times\mathbb{R}^3$ and $\mathfrak{k}=\mathfrak{h}\oplus\mathfrak{h}^\perp=\{yZ_{f_1-f_2}\colon\ y\in\mathbb{R}\}\oplus\{z^{11}Z_{2f_1}+z^{12}Z_{f_1+f_2}+z^{22}Z_{2f_2}\colon\ z^{ij}\in\mathbb{R}\}$ where

$$Z_{f_1-f_2}=\begin{pmatrix}0&-1&&\\1&0&&\\&&0&-1\\&&1&0\end{pmatrix}\quad\text{and}\quad z^{11}Z_{2f_1}+z^{12}Z_{f_1+f_2}+z^{22}Z_{2f_2}=\begin{pmatrix}&&z^{11}&z^{12}\\&&z^{12}&z^{22}\\-z^{11}&-z^{12}&&\\-z^{12}&-z^{22}&&\end{pmatrix}.$$

Notice also that one can read off from the properties of the root system that $[\mathfrak{h}^\perp,\mathfrak{h}^\perp]\subset\mathfrak{h}$ whence the left and right hand side of (3.17) are both identically 0 for the homogeneous case $\mathbb{I}=1$. Thus the homogeneous case is Hamiltonian ($F$ is constant) at the ultimate reduced level $T^*(U(2)/S^1)$.

For general $n$ one can use that the root system $\Sigma(\mathfrak{g},\mathfrak{a})$ is of type $C_n$ whence the positive system will be of the form $\Sigma_+=\{f_i\pm f_j\colon\ 1\leqslant i<j\leqslant n\}\cup\{2f_i\colon\ 1\leqslant i\leqslant n\}$ and the simple roots are $f_i-f_j$ with $1\leqslant i<j\leqslant n$ and $2f_n$. A choice of $w_0$ can now be determined by letting appropriately many simple roots vanish on $w_0$. E.g., one can conclude just as above that choosing a non-zero $w_0$ in the joint kernel of $f_i-f_j$ with $1\leqslant i<j\leqslant n$ yields a system which is Hamiltonian at the ultimate reduced level $T^*(K/H)=T^*(U(n)/(U(1)^{n-1}))$.

### *4.4. Split $G_2$, 2–3–5, 1/3 and rubber rolling*

Let $G$ be the split real form of the exceptional complex semi-simple Lie group $G_2$. This group is 14-dimensional and can be realized as the automorphism group of the split octonions. We refer to [24,22,17] for background. The Cartan decomposition data are the following,

$$K=\mathrm{SU}(2)\times_{(\pm1)}\mathrm{SU}(2)\cong\mathrm{SO}(4),\qquad \mathfrak{p}\cong\mathbb{R}^8,\qquad \mathfrak{a}\cong\mathbb{R}^2,\quad\text{and}\quad \mathfrak{m}=\{0\}.$$

The restricted roots are of type $G_2$ whence a positive system can be written as

$$\Sigma_+ = \{\lambda_1, \lambda_2, \lambda_1 + \lambda_2, \lambda_1 + 2\lambda_2, 2\lambda_1 + 3\lambda_2, \lambda_1 + 3\lambda_2\}$$

with $\lambda_1$ and $\lambda_2$ simple. We choose $w_0 \in \mathfrak{a}$ such that $\lambda_1(w_0) = 0$ and $\lambda_2(w_0) \neq 0$. Thus the set of relevant roots is $\Phi = \{\lambda_2, \lambda_1 + \lambda_2, \lambda_1 + 2\lambda_2, 2\lambda_1 + 3\lambda_2, \lambda_1 + 3\lambda_2\}$ and the infinitesimal internal symmetries are

$$\mathfrak{h} = \operatorname{span}\{Z_{\lambda_1}\} = \mathbb{R}$$

which we view as the Lie algebra of the connected component $H$ of $Z_K(w_0)$,

$$H \cong S^1.$$

According to Section 3 we have $V = \operatorname{ad}(w_0)\mathfrak{k} = \operatorname{span}\{e_\lambda\colon \lambda \in \Phi\} \cong \mathbb{R}^5$ and therefore

$$Q \cong K \times \mathbb{R}^5 \quad \text{and} \quad Q/\big(\mathbb{R}^5 \times H\big) = K/H \cong \mathrm{SU}(2) \times \mathrm{SO}(3)/S^1 \cong \mathrm{SU}(2) \times S^2.$$

We remark that $K/H \cong G/P_{w_0}$ where $P_{w_0}$ is the parabolic subgroup of $G$ associated to the subset of simple roots $\Pi$ consisting of $\{\lambda \in \Pi\colon \lambda(w_0) = 0\} = \{\lambda_1\}$.

What about Hamiltonization? Suppose $\mathbb{I} = 1$ which implies that $s\mu_0(s)s^{-1} = \mu_0(e)$ and $\mu_0(e)^{-1}Z_\kappa = (1 + \kappa(w_0)^2)^{-1}Z_\kappa$ for all $\kappa \in \Sigma_+$. Thus the left hand side of (3.17) is non-zero for, e.g., $\kappa = \lambda_1 + \lambda_2$, $\mu = \lambda_1 + 2\lambda_2$ and $\nu = 2\lambda_1 + 3\lambda_2$. Thus the system is not Hamiltonizable at the $T(K/H)$-level corresponding to reduction of $(TK, \widetilde{\Omega}, \mathcal{H}_c)$ at the 0-level set of the $J_H$-momentum map.

On the other hand we recognize $K/H$ as the double cover configuration space $\mathrm{SO}(3) \times S^2$ of the sphere-on-sphere-rolling system. This system is a natural generalization of the Chaplygin ball on a table when one forbids slipping. One can also introduce a no-twist constraint and the resulting non-holonomic system has been shown to be Hamiltonizable by Koiller and Ehlers [19]. Moreover, it seems to be known since Cartan that $G_2$ is related to this no-twist no-slip sphere-on-sphere system. Therefore, one might expect some relation between this system and the one defined by $(TK, \widetilde{\Omega}, \mathcal{H}_c)$ even though the non-Hamiltonizability of the latter is apparently an obstruction to any such relation.

Recall from Theorem 3.5 that $\widetilde{\Omega} = \Omega^K + \Lambda$. In order to stand a chance at obtaining a Hamiltonizable system we consider the set $\{(s,u) \in TK\colon i(X_{\mathrm{nh}})\Lambda_{(s,u)} = 0\}$. By (3.14) we have

$$i(X_{\mathrm{nh}})\Lambda(\zeta_\nu) = -\sum_\mu \mu(w_0)^2 c^{\lambda_1}_{\mu\nu} g_\mu g_{\lambda_1} + \sum_{\lambda,\mu \in \Phi} \mu(w_0)^2 c^\lambda_{\mu\nu} g_\lambda g_\mu.$$

Setting $\nu = \lambda_1 + 2\lambda_2$ the possibilities for $\{\lambda, \mu\}$ are $\{\lambda_2, \lambda_1 + \lambda_2\}$ and $\{\lambda_2, \lambda_1 + 3\lambda_2\}$. The resulting condition for $i(X_{\mathrm{nh}})\Lambda(\zeta_\nu) = 0$ is then

$$c^{\lambda_2}_{\lambda_1+\lambda_2,\nu}\big((\lambda_1+\lambda_2)(w_0)^2 - (\lambda_2)(w_0)^2\big) g_{\lambda_2} g_{\lambda_1+\lambda_2} + c^{\lambda_2}_{\lambda_1+3\lambda_2,\nu}\big((\lambda_1+3\lambda_2)(w_0)^2 - (\lambda_2)(w_0)^2\big) g_{\lambda_2} g_{\lambda_1+3\lambda_2} = 0.$$

Since $\lambda_1(w_0) = 0$ this is satisfied if $g_{\lambda_1+3\lambda_2} = 0$. We find that $i(X_{\mathrm{nh}})\Lambda_{(s,u)}$ vanishes when $(s,u)$ belongs to the right invariant distribution

$$\mathcal{D}_{\mathrm{new}} := \ker\big(\eta^{\lambda_1}, \eta^{\lambda_1+2\lambda_2}, \eta^{\lambda_1+3\lambda_2}, \eta^{2\lambda_1+3\lambda_2}\big) = \operatorname{span}\{\zeta_{\lambda_2}, \zeta_{\lambda_1+\lambda_2}\}. \tag{4.24}$$

This is a rank two distribution with growth 2–3–5–6 on a 6-dimensional configuration space. Notice that $[\zeta_{\lambda_1}, \mathcal{D}_{\mathrm{new}}] \subset \mathcal{D}_{\mathrm{new}}$, i.e., $\mathcal{D}_{\mathrm{new}}$ is invariant under the action of the connected Lie group $H$ on $K$. Via the Langlands decomposition $H$ coincides with $P_{w_0} \cap K \cong H \cong S^1$. Along $\mathcal{D}_{\mathrm{new}}$ the equations of motion are thus given by the canonical equation

$$i(X_{\mathrm{nh}})\Omega^K = d\mathcal{H}_c.$$

Moreover, $X_{\mathrm{nh}}$ lies in the kernel of $\tau_K^*\mathcal{A}_{\mathrm{new}}$ where $\tau_K: TK \to K$ and

$$\mathcal{A}_{\mathrm{new}} = \big(\eta^{\lambda_1}, \eta^{\lambda_1+2\lambda_2}, \eta^{\lambda_1+3\lambda_2}, \eta^{2\lambda_1+3\lambda_2}\big): TK \to \mathbb{R}^4.$$

However, it is not true that $X_{\mathrm{nh}}$ is tangent to $\mathcal{D}_{\mathrm{new}}$. (One could say that the constraint forces introduced by $\mathcal{D}_{\mathrm{new}}$ on the Hamiltonian system $(TK, \Omega^K, \mathcal{H}_c)$ are non-trivial.) By invariance $\mathcal{D}_{\mathrm{new}}$ factors to a rank two distribution $\mathcal{D}_{\mathrm{new}}/H$ of growth 2–3–5 on $K/H \cong \mathrm{SU}(2) \times \mathrm{SO}(3)/S^1 \cong S^3 \times S^2$. Indeed, passing to the right trivialization of $TK$ for a moment, $\mathcal{D}_{\mathrm{new}}/H$ can be realized as

$$K \times_H \operatorname{span}\{Z_{\lambda_2}, Z_{\lambda_1+\lambda_2}\}.$$

Further, the restriction of the compressed Hamiltonian

$$\mathcal{H}_c|\mathcal{D}_{\mathrm{new}} = \frac{1}{2}\langle \mathbb{I}u, u\rangle + \frac{1}{2}\lambda_2(w_0)^2\big(g^2_{\lambda_2} + g^2_{\lambda_1+\lambda_2}\big)$$

is $K$-independent. E.g., $\zeta_{\lambda_1}(g^2_{\lambda_2} + g^2_{\lambda_1+\lambda_2}) = -2c^{\lambda_1+\lambda_2}_{\lambda_1,\lambda_2}(g_{\lambda_2}g_{\lambda_1+\lambda_2} - g_{\lambda_1+\lambda_2}g_{\lambda_2}) = 0$. That is, $\mathcal{H}_c|\mathcal{D}_{\mathrm{new}}$ is actually left invariant and therefore

$$\mathcal{H}_c|\mathcal{D}_{\mathrm{new}} = \frac{1}{2}\langle(\mathbb{I}+\lambda(w_0)^2)u, u\rangle.$$

Let us now follow [22] and define $\mathfrak{g}_i \subset \mathfrak{g}$ for $i \neq 0$ to be the sum of all restricted root spaces $\mathfrak{g}_\lambda$ such that $\lambda_2$ occurs with coefficient $i$ in the decomposition of $\lambda$ into simple roots $\lambda_1, \lambda_2$; $\mathfrak{g}_0$ is defined to be the sum of $\mathfrak{a}$ and all restricted root spaces $\mathfrak{g}_\lambda$ such that $\lambda_2$ occurs with coefficient 0 in the decomposition of $\lambda$ into simple roots $\lambda_1, \lambda_2$. Thus

$$\mathfrak{g} = \mathfrak{g}_{-3} \oplus \mathfrak{g}_{-2} \oplus \mathfrak{g}_{-1} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \mathfrak{g}_3$$

which is the grading of $\mathfrak{g}$ with respect to the parabolic subalgebra $\mathfrak{p}_{w_0} = \mathrm{Lie}(P_{w_0}) = \bigoplus_{i=0,\dots,3} \mathfrak{g}_i$. Choose an orthonormal basis $X_\lambda$ of $\bigoplus_{\lambda\in\Sigma} \mathfrak{g}_\lambda$ consisting of root vectors. Then the prescription $Z_\lambda \mapsto X_{-\lambda}$ and $e_\lambda \mapsto X_\lambda$ for $\lambda \in \Sigma_+$ induces isomorphisms

$$\mathfrak{h}^\perp \cong \mathfrak{g}_- := \mathfrak{g}_{-3} \oplus \mathfrak{g}_{-2} \oplus \mathfrak{g}_{-1} \quad \text{and} \quad V \cong \mathfrak{g}_+ := \mathfrak{g}_1 \oplus \mathfrak{g}_2 \oplus \mathfrak{g}_3 = \mathfrak{p}_{w_0}/\mathfrak{g}_0.$$

This corresponds effectively to the passage from the Cartan to the Iwasawa decomposition. Moreover, the isomorphism $\mathfrak{h}^\perp \cong \mathfrak{g}_-$ is equivariant with respect to the $H$-action on $\mathfrak{h}^\perp$ and the $P_{w_0}$-action on $\mathfrak{g}_-$. This follows from the Langlands decomposition of the parabolic $P_{w_0}$. Associated to the grading there is a $P_{w_0}$-invariant filtration

$$\mathfrak{g}/\mathfrak{p}_{w_0} \supset \mathfrak{g}^{-2}/\mathfrak{p}_{w_0} \supset \mathfrak{g}^{-1}/\mathfrak{p}_{w_0}$$

of $\mathfrak{g}/\mathfrak{p}_{w_0}$ where the filter components are $\mathfrak{g}^i = \bigoplus_{j=i,\dots,3} \mathfrak{g}_j$. With this notation and the isomorphism $\mathfrak{h}^\perp \cong \mathfrak{g}_-$ we obtain

$$\mathcal{D}_{\mathrm{new}}/H \cong K \times_H \mathrm{span}\{Z_{\lambda_2}, Z_{\lambda_1+\lambda_2}\} \cong G \times_{P_{w_0}} \mathfrak{g}^{-1}/\mathfrak{p}_{w_0} \subset G \times_{P_{w_0}} \mathfrak{g}/\mathfrak{p}_{w_0} \cong T(S^3 \times S^2).$$

The growth of the distribution is of course reflected in the way in which the filtration reacts to the Lie bracket: $[\mathfrak{g}^{-1}/\mathfrak{p}_{w_0}, \mathfrak{g}^{-1}/\mathfrak{p}_{w_0}] = \mathfrak{g}^{-2}/\mathfrak{p}_{w_0}$ and $[\mathfrak{g}^{-1}/\mathfrak{p}_{w_0}, \mathfrak{g}^{-2}/\mathfrak{p}_{w_0}] = \mathfrak{g}/\mathfrak{p}_{w_0}$. This distribution corresponds to the homogeneous model of Cartan geometries of type $(G, P_{w_0})$.

Bor and Montgomery [5] have explained that $G \times_{P_{w_0}} \mathfrak{g}^{-1}/\mathfrak{p}_{w_0} \subset G \times_{P_{w_0}} \mathfrak{g}/\mathfrak{p}_{w_0}$ can be identified with the no-twist no-slip distribution when one passes over the two fold covering $S^3 \times S^2 = K/H \to \mathrm{SO}(3) \times S^2$ and when the ratio of the radii of the two balls is 1/3. Along similar lines Sagerschnig [22] has explained some of the Cartan geometric background and proved that it is isomorphic to a certain 'divisors of 0 distribution', and Agrachev [1] has shown that this 'divisors of 0 distribution' can be realized as the 'rubber rolling distribution' for ratio 1/3.

## 5. Questions

Hamiltonization at non-zero momentum $\alpha \in \mathfrak{h}^*$ remains open. Generalizing Theorem 3.6 to this setting is a problem for future work. The difficulty here is that one has to take into account the extra structure coming from the non-zero orbit $\mathcal{O} = \mathrm{Ad}^*(H).\alpha$ in (3.15).

Integrability? Very little is known about integrability of $n$-dimensional Chaplygin systems, and we have not touched at all the question of integrating the systems introduced in Section 3. Jovanovic [16] has just shown very recently that the $n$-dimensional Chaplygin ball is integrable when the inertia tensor is of special type as in (4.21). Chaplygin [8] has explicitly integrated the 3-dimensional problem.

## Acknowledgements

I would like to thank the organizers of the Mikulov meeting, Special edition in honor of Peter Michor's 60th birthday, for two pleasant days and the opportunity to present parts of the present paper. The idea of relating Chaplygin systems to semi-simple Lie groups is, of course, taken from the paper [2] of Peter et al. where a similar programme is carried out for spin Calogero–Moser systems. I am also grateful to Tudor Ratiu for helpful discussions and to Katja Sagerschnig for important remarks concerning Section 4.4.

## References

[1] A.A. Agrachev, Rolling balls and octonions, Proc. Steklov Inst. Math. 258 (1) (2007) 17–27.
[2] D. Alekseevsky, A. Kriegl, M. Losik, P.W. Michor, The Riemann geometry of orbit spaces – the metric, geodesics and integrable systems, Publ. Math. 62 (3–4) (2003) 247–276.
[3] L. Bates, J. Sniatycki, Nonholonomic reduction, Rep. Math. Phys. 32 (1) (1993) 99–115.
[4] A.M. Bloch, Nonholonomic Mechanics and Control, Springer, 2003.
[5] G. Bor, R. Montgomery, $G_2$ and the rolling distribution, arXiv:math/0612469v1.
[6] A.V. Borisov, I.S. Mamaev, Chaplygin's ball rolling problem is Hamiltonian, Mathematical Notes 70 (2001) 793–795.
[7] A.V. Borisov, I.S. Mamaev, Hamiltonization of nonholonomic systems, arXiv:nlin/0509036v1, 2005.
[8] S.A. Chaplygin, On a ball's rolling on a horizontal plane, Regul. Chaotic Dyn. 7 (2002) 131–148; translation of original in: Mathematical Collection of the Moscow Mathematical Society 24 (1903) 139–168 (in Russian).
[9] J.J. Duistermaat, Chaplygin's sphere, arXiv:math/0409019v1.

[10] K. Ehlers, J. Koiller, R. Montgomery, P.M. Rios, Nonholonomic systems via moving frames: Cartan equivalence and Chaplygin Hamiltonization, in: The Breath of Symplectic and Poisson Geometry, in: Progress in Mathematics, vol. 232, 2004, pp. 75–120.
[11] Yu. Fedorov, Dynamical systems with an invariant measure on the Riemannian symmetric pairs $(GL(N), SO(N))$, Reg. Ch. Dyn. 1 (1) (1996) 38–44 (in Russian).
[12] Yu. Fedorov, Systems with an invariant measure on Lie groups, in: Hamiltonian Systems with Three or More Degrees of Freedom, in: NATO ASI Series C, vol. 533, Kluwer Acad. Publ., 1999, pp. 350–357.
[13] Y.N. Fedorov, V.V. Kozlov, Various aspects of $n$-dimensional rigid body dynamics, in: Amer. Math. Soc. Transl., Ser. (2), vol. 168, 1995, pp. 141–171.
[14] S. Hochgerner, L. Garcia-Naranjo, $G$-Chaplygin systems with internal symmetries, truncation, and an (almost) symplectic view of Chaplygin's ball, J. Geom. Mech. 1 (1) (2009) 35–53.
[15] B. Jovanovic, LR and $\mathrm{L}+\mathrm{R}$ systems, J. Phys. A: Math. Theor. 42 (22) (2009).
[16] B. Jovanovic, Hamiltonization and integrability of the Chaplygin sphere in $\mathbb{R}^n$, arXiv:math-ph/0902.4397v1.
[17] A.W. Knapp, Lie Groups Beyond an Introduction, 2nd edition, Birkhäuser, 2002.
[18] J. Koiller, Reduction of some classical nonholonomic systems with symmetry, Arch. Rat. Mech. An. 118 (1992) 113–148.
[19] J. Koiller, K. Ehlers, Rubber rolling over a sphere, Regul. Chaotic Dyn. 7 (2007) 127–152.
[20] P. Libermann, Sur les structures presque complexes et autres structures infinitesimales regulieres, Bulletin de la S.M.F. 83 (1955) 195–224.
[21] P. Libermann, C.-M. Marle, Symplectic Geometry and Analytical Mechanics, D. Reidel, Kluwer, 1987.
[22] K. Sagerschnig, Split octonions and generic rank two distributions in dimension five, Arch. Math. (Brno) 42 (2006) 329–339.
[23] D. Schneider, Nonholonomic Euler–Poincaré equations and stability in Chaplygin's sphere, Dynamical Systems: An International J. 17 (2) (2002) 87–130.
[24] T.A. Springer, F.D. Veldkamp, Octonions, Jordan Algebras and Exceptional Groups, Springer, 2000.