Some Geometric Evolution Equations Arising as Geodesic Equations on Groups of Diffeomorphisms Including the Hamiltonian Approach

Peter W. Michor ⋆

Key words: Diffeomorphism group, connection, Jacobi field, symplectic structure, Burgers' equation, KdV equation.
2000 Mathematics Subject Classification: Primary 58B20, 58D05, 58D15, 58F07, 58E12, 35Q53

In: *Phase space analysis of Partial Differential Equations.*
Series: Progress in Non Linear Differential Equations and Their Applications, Vol.69. Bove, Antonio; Colombini, Ferruccio; Santo, Daniele Del (Eds.).
Birkhauser Verlag 2006. Pages 133-215

Fakultät für Mathematik, Universität Wien, Nordbergstrasse 15, A-1090 Wien, Austria; and Erwin Schrödinger International Institute of Mathematical Physics, Boltzmanngasse 9, A-1090 Wien, Austria.
peter.michor@esi.ac.at

⋆ Supported by 'Fonds zur Förderung der wissenschaftlichen Forschung', Projekt P 17108. Work partly done at the Program for Evolutionary Dynamics, Harvard University

Table of contents

Introduction

This is the extended version of a lecture course given at the University of Vienna in the spring term 2005. Many thanks to the audience of this course for many keen questions. The main aim of this course was to understand the papers [11] and [13].

The purpose of this review article is to give a complete account of existence and uniqueness of the solutions of the members of higher order of the hierarchies of Burgers' equation and the Korteweg-de Vries equation, including their derivation and all the necessary background. We do this both on the circle, and on the real line in the setting of rapidly decreasing functions. These are all geodesic equations of infinite dimensional regular Lie groups, namely the diffeomorphism group of the line or the circle and the corresponding Virasoro group.

Let us describe the content: Appendix A is a short description of convenient calculus in infinite dimensions (beyond Banach spaces) where everything is based on smooth curves: A mapping is C^∞ if it maps smooth curves to smooth curves. It is a theorem that smooth curves in a space of smooth functions are just smooth functions of one variable more; this is the basic assumption of variational calculus. Appendix B gives a short account of infinite dimensional regular Lie groups. Here regularity means that a smooth curve in the Lie algebra can be integrated to a smooth curve in the group whose right (or left) logarithmic derivative equals the given curve. No infinite dimensional Lie group is known which is not regular. Section 1, as a motivating example, computes the geodesics and the curvature of the most naive Riemannian metric on the space of embeddings of the real line to itself and shows that this can be converted into Burgers' equation. Section 2 treats Hamiltonian mechanics on infinite dimensional weak symplectic manifolds. Here 'weak' means that the symplectic 2-form is injective as a mapping from the tangent bundle to the cotangent bundle. Section 3 computes geodesics and curvatures of right invariant Riemannian metrics on regular Lie groups as done by Arnold [4]. Section 4 redoes this in the symplectic approach and computes the associated momentum mappings and conserved quantities. Section 5 shows that the geodesic distance vanishes on any full diffeomorphis group for the right invariant metric coming from the L^2-metric on the Lie algebra of vector fields for a given Riemannian metric on a manifold. In particular, Burgers' equation is the geodesic equation of such a metric. Section 6 treats the group of diffeomorphisms of the real line which decrease rapidly to the identity as a regular Lie group. This will be important for Burgers' equation as geodesic equation on this group, and also for the KdV equation. Here we also give a short presentation of Sobolev spaces on the real line and of the scale of HC^n-spaces for which we were able to give simple proofs of the results which we need later. Section 7 treats geodesic equations on the diffeomorphism groups of the real line or S^1 which leads to Burgers' hierarchy. We solve these equations starting at certain higher order, following [13]. Section

8 does this for the Virasoro groups on the real line or S^1. For the solution of the higher order equations we follow [11].

Note that in this paper we concentrate on in the smooth ($= C^\infty$) aspect. We also do not treat complete integrability for Burgers' and KdV equation, although we prepared almost all of the necessary background.

1. A general setting and a motivating example

1.1. The principal bundle of embeddings

Let M and N be smooth finite dimensional manifolds, connected and second countable without boundary, such that $\dim M \le \dim N$. Then the space $\mathrm{Emb}(M,N)$ of all embeddings (immersions which are homeomorphisms on their images) from M into N is an open submanifold of $C^\infty(M,N)$ which is stable under the right action of the diffeomorphism group of M. Here $C^\infty(M,N)$ is a smooth manifold modeled on spaces of sections with compact support $\Gamma_c(f^*TN)$. In particular the tangent space at f is canonically isomorphic to the space of vector fields along f with compact support in M. If f and g differ on a non-compact set then they belong to different connected components of $C^\infty(M,N)$. See [31] and [37]. Then $\mathrm{Emb}(M,N)$ is the total space of a smooth principal fiber bundle with structure group the diffeomorphism group of M; the base is called $B(M,N)$, it is a Hausdorff smooth manifold modeled on nuclear (LF)-spaces. It can be thought of as the "nonlinear Grassmannian" or "differentiable Chow variety" of all submanifolds of N which are of type M. This result is based on an idea implicitly contained in [51], it was fully proved in [7] for compact M and for general M in [36]. See also [37], section 13 and [31]. If we take a Hilbert space H instead of N, then $B(M,H)$ is the classifying space for $\mathrm{Diff}(M)$ if M is compact, and the classifying bundle $\mathrm{Emb}(M,H)$ carries also a universal connection. This is shown in [38].

1.2

If (N,g) is a Riemannian manifold then on the manifold $\mathrm{Emb}(M,N)$ there is a naturally induced weak Riemannian metric given, for $s_1, s_2 \in \Gamma_c(f^*TN)$ and $\varphi \in \mathrm{Emb}(M,N)$, by

$$G_\phi(s_1,s_2) = \int_M g(s_1,s_2)\,\mathrm{vol}(\phi^* g), \quad \phi \in \mathrm{Emb}(M,N),$$

where $\mathrm{vol}(g)$ denotes the volume form on N induced by the Riemannian metric g and $\mathrm{vol}(\phi^*g)$ the volume form on M induced by the pull back metric ϕ^*g. The covariant derivative and curvature of the Levi-Civita connection induced by G were investigated in [6] if $N = \mathbb{R}^{\dim M+1}$ (endowed with the standard inner product) and in [25] for the general case. In [40] it was shown that the geodesic distance (topological metric) on the base manifold $B(M,N) = \mathrm{Emb}(M,N)/\operatorname{Diff}(M)$ induced by this Riemannian metric vanishes.

This weak Riemannian metric is invariant under the action of the diffeomorphism group $\operatorname{Diff}(M)$ by composition from the right and hence it induces a Riemannian metric on the base manifold $B(M,N)$.

1.3. Example

Let us consider the special case $M = N = \mathbb{R}$, that is, the space $\mathrm{Emb}(\mathbb{R},\mathbb{R})$ of all embeddings of the real line into itself, which contains the diffeomorphism group $\operatorname{Diff}(\mathbb{R})$ as an open subset. The case $M = N = S^1$ is treated in a similar fashion and the results of this paper are also valid in this situation, where $\mathrm{Emb}(S^1,S^1) = \operatorname{Diff}(S^1)$. For our purposes, we may restrict attention to the space of orientation-preserving embeddings, denoted by $\mathrm{Emb}^+(\mathbb{R},\mathbb{R})$. The weak Riemannian metric has thus the expression

$$G_f(h,k) = \int_{\mathbb{R}} h(x)k(x)|f'(x)|\,dx, \quad f \in \mathrm{Emb}(\mathbb{R},\mathbb{R}), \quad h,k \in C_c^\infty(\mathbb{R},\mathbb{R}).$$

We shall compute the geodesic equation for this metric by variational calculus. The energy of a curve f of embeddings is

$$E(f) = \tfrac{1}{2}\int_a^b G_f(f_t,f_t)dt = \tfrac{1}{2}\int_a^b\int_{\mathbb{R}} f_t^2 f_x\,dxdt.$$

If we assume that $f(x,t,s)$ is a smooth function and that the variations are with fixed endpoints, then the derivative with respect to s of the energy is

$$\begin{aligned}
\partial s|_0 E(f(\quad,s)) &= \partial s|_0 \tfrac{1}{2}\int_a^b\int_{\mathbb{R}} f_t^2 f_x\,dxdt \\
&= \tfrac{1}{2}\int_a^b\int_{\mathbb{R}} (2f_t f_{ts} f_x + f_t^2 f_{xs})dxdt \\
&= -\tfrac{1}{2}\int_a^b\int_{\mathbb{R}} (2f_{tt} f_s f_x + 2f_t f_s f_{tx} + 2f_t f_{tx} f_s)dxdt \\
&= -\int_a^b\int_{\mathbb{R}} \left(f_{tt} + 2\frac{f_t f_{tx}}{f_x}\right) f_s f_x dxdt,
\end{aligned}$$

so that the geodesic equation with its initial data is:

$$\begin{aligned} f_{tt} &= -2\frac{f_t f_{tx}}{f_x}, \quad f(\quad,0) \in \operatorname{Emb}^+(\mathbb{R},\mathbb{R}), \quad f_t(\quad,0) \in C_c^\infty(\mathbb{R},\mathbb{R}) \\ &=: \Gamma_f(f_t,f_t), \end{aligned} \tag{1}$$

where the Christoffel symbol $\Gamma : \operatorname{Emb}(\mathbb{R},\mathbb{R}) \times C_c^\infty(\mathbb{R},\mathbb{R}) \times C_c^\infty(\mathbb{R},\mathbb{R}) \to C_c^\infty(\mathbb{R},\mathbb{R})$ is given by symmetrisation:

$$\Gamma_f(h,k) := -\frac{hk_x + h_x k}{f_x} = -\frac{(hk)_x}{f_x}\,. \tag{2}$$

For vector fields X, Y on $\operatorname{Emb}(\mathbb{R},\mathbb{R})$ the covariant derivative is given by the expression $\nabla^{\mathrm{Emb}}_X Y = dY(X) - \Gamma(X,Y)$. The Riemannian curvature $R(X,Y)Z = (\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]})Z$ is then determined in terms of the Christoffel form by

$$\begin{aligned} R(X,Y)Z &= (\nabla_X\nabla_Y - \nabla_Y\nabla_X - \nabla_{[X,Y]})Z \\ &= \nabla_X(dZ(Y) - \Gamma(Y,Z)) - \nabla_Y(dZ(X) - \Gamma(X,Z)) \\ &\qquad - dZ([X,Y]) + \Gamma([X,Y],Z) \\ &= d^2Z(X,Y) + dZ(dY(X)) - \Gamma(X, dZ(Y)) \\ &\quad - d\Gamma(X)(Y,Z) - \Gamma(dY(X),Z) - \Gamma(Y, dZ(X)) + \Gamma(X,\Gamma(Y,Z)) \\ &\quad - d^2Z(Y,X) - dZ(dX(Y)) + \Gamma(Y, dZ(X)) \\ &\quad + d\Gamma(Y)(X,Z) + \Gamma(dX(Y),Z) + \Gamma(X, dZ(Y)) - \Gamma(Y,\Gamma(X,Z)) \\ &\quad - dZ(dY(X) - dX(Y)) + \Gamma(dY(X) - dX(Y), Z) \\ &= -d\Gamma(X)(Y,Z) + \Gamma(X,\Gamma(Y,Z)) + d\Gamma(Y)(X,Z) - \Gamma(Y,\Gamma(X,Z) \end{aligned}$$

so that

$$\begin{aligned} &R_f(h,k)\ell = \\ &= -d\Gamma(f)(h)(k,\ell) + d\Gamma(f)(k)(h,\ell) + \Gamma_f(h,\Gamma_f(k,\ell)) - \Gamma_f(k,\Gamma_f(h,\ell)) \\ &= -\frac{h_x(k\ell)_x}{f_x^2} + \frac{k_x(h\ell)_x}{f_x^2} + \frac{\Big(h\frac{(k\ell)_x}{f_x}\Big)_x}{f_x} - \frac{\Big(k\frac{(h\ell)_x}{f_x}\Big)_x}{f_x} \\ &= \frac{1}{f_x^3}\Big(f_{xx}h_xk\ell - f_{xx}hk_x\ell + f_xhk_{xx}\ell - f_xh_{xx}k\ell + 2f_xhk_x\ell_x - 2f_xh_xk\ell_x\Big). \end{aligned} \tag{3}$$

Now let us consider the trivialisation of $T\operatorname{Emb}(\mathbb{R},\mathbb{R})$ by right translation (this is most useful for $\operatorname{Diff}(\mathbb{R})$). The derivative of the inversion $\operatorname{Inv} : g \mapsto g^{-1}$ is given by

$$T_g(\mathrm{Inv})h = -T(g^{-1}) \circ h \circ g^{-1} = -\frac{h \circ g^{-1}}{g_x \circ g^{-1}}$$

for $g \in \mathrm{Emb}(\mathbb{R},\mathbb{R})$, $h \in C_c^\infty(\mathbb{R},\mathbb{R})$. Defining

$$u := f_t \circ f^{-1}, \quad \text{or, in more detail,} \quad u(t,x) = f_t(t, f(t, \quad)^{-1}(x)),$$

we have

$$\begin{aligned}
u_x &= (f_t \circ f^{-1})_x = (f_{tx} \circ f^{-1}) \frac{1}{f_x \circ f^{-1}} = \frac{f_{tx}}{f_x} \circ f^{-1}, \\
u_t &= (f_t \circ f^{-1})_t = f_{tt} \circ f^{-1} + (f_{tx} \circ f^{-1})(f^{-1})_t \\
&= f_{tt} \circ f^{-1} - (f_{tx} \circ f^{-1}) \frac{1}{f_x \ f^{-1}} (f_t \ f^{-1})
\end{aligned}$$

which, by (1) and the first equation becomes

$$u_t = f_{tt} \circ f^{-1} - \left(\frac{f_{tx} f_t}{f_x}\right) \circ f^{-1} = -3 \left(\frac{f_{tx} f_t}{f_x}\right) \circ f^{-1} = -3 u_x u.$$

The geodesic equation on $\mathrm{Emb}(\mathbb{R},\mathbb{R})$ in right trivialization, that is, in Eulerian formulation, is hence

$$u_t = -3 u_x u\,, \tag{4}$$

which is just Burgers' equation.

Finally let us solve Burgers' equation and also describe its universal completion, see see [10], [2], or [26].

In $\mathbb{R}^2$ with coordinates (x,y) consider the vector field $Y(x,y) = (3y, 0) = 3y\partial_x$ with differential equation $\dot{x} = 3y, \dot{y} = 0$. It has the complete flow $\mathrm{Fl}_t^Y(x,y) = (x + 3ty, y)$.

Let now $t \mapsto u(t,x)$ be a curve of functions on $\mathbb{R}$. We ask when the graph of u can be reparametrized in such a way that it becomes a solution curve of the push forward vector field $Y_* : f \mapsto Y \circ f$ on the space of embeddings $\mathrm{Emb}(\mathbb{R},\mathbb{R}^2)$. Thus consider a time dependent reparametrization $z \mapsto x(t,z)$, i.e., $x \in C^\infty(\mathbb{R}^2,\mathbb{R})$. The curve $t \mapsto (x(t,z), u(t, x(z,t)))$ in $\mathbb{R}^2$ is an integral curve of Y if and only if

$$\begin{aligned}
&\begin{pmatrix} 3u \circ x \\ 0 \end{pmatrix} = \partial_t \begin{pmatrix} x \\ u \circ x \end{pmatrix} = \begin{pmatrix} x_t \\ u_t \circ x + (u_x \circ x) \cdot x_t \end{pmatrix} \\
\iff &\begin{cases} x_t = 3u \circ x \\ 0 = (u_t + 3uu_x) \circ x \end{cases}
\end{aligned}$$

This implies that the graph of $u(t,\cdot)$, namely the curve $t \mapsto (x \mapsto (x, u(t,x)))$, may be parameterized as a solution curve of the vector field Y_* on the space of embeddings $\mathrm{Emb}(\mathbb{R},\mathbb{R}^2)$ starting at $x \mapsto (x, u(0,x))$ if and only if u is a solution of the partial differential equation $u_t + 3uu_x = 0$. The parameterization $z \mapsto x(z,t)$ is then given by $x_t(z,t) = 3u(x(t,z))$ with $x(0,z) = z \in \mathbb{R}$.

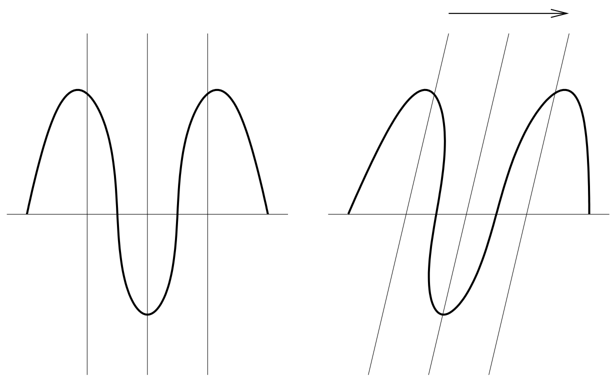


Fig. 1 The characteristic flow of the inviscid Burgers' equation tilts the plane.

This has a simple physical meaning. Consider freely flying particles in $\mathbb{R}$, and trace a trajectory $x(t)$ of one of the particles. Denote the velocity of a particle at the position x at the moment t by $u(t,x)$, or rather, by $3u(t,x) := \dot{x}(t)$. Due to the absence of interaction, the Newton equation of any particle is $\ddot{x}(t) = 0$.

Let us illustrate this: The flow of the vector field $Y = 3u\partial_x$ is tilting the plane to the right with constant speed. The illustration shows how a graph of an honest function is moved through a shock (when the derivatives become infinite) towards the graph of a multivalued function; each piece of it is still a local solution.

2. Weak symplectic manifolds

2.1. Review

For a finite dimensional symplectic manifold (M,ω) we have the following exact sequence of Lie algebras:

$$0 \to H^0(M) \to C^\infty(M,\mathbb{R}) \xrightarrow{\text{grad}^\omega} \mathfrak{X}(M,\omega) \longrightarrow H^1(M) \to 0.$$

Here $H^*(M)$ is the real De Rham cohomology of M, the space $C^\infty(M,\mathbb{R})$ is equipped with the Poisson bracket $\{\quad,\quad\}$, $\mathfrak{X}(M,\omega)$ consists of all vector fields ξ with $\mathcal{L}_\xi\omega = 0$ (the locally Hamiltonian vector fields), which is a Lie algebra for the Lie bracket. Furthermore, $\text{grad}^\omega f$ is the Hamiltonian vector field for $f \in C^\infty(M,\mathbb{R})$ given by $i(\operatorname{grad}^\omega f)\omega = df$ and $\gamma(\xi) = [i_\xi\omega]$. The spaces $H^0(M)$ and $H^1(M)$ are equipped with the zero bracket.

Consider a symplectic right action $r : M \times G \to M$ of a connected Lie group G on M; we use the notation $r(x,g) = r^g(x) = r_x(g) = x.g$. By $\zeta_X(x) = T_e(r_x)X$ we get a mapping $\zeta : \mathfrak{g} \to \mathfrak{X}(M,\omega)$ which sends each element X of the Lie algebra $\mathfrak{g}$ of G to the fundamental vector field X. This is a Lie algebra homomorphism (for right actions!).

$$\begin{tikzcd}
H^0(M) \arrow[r, "i"] & C^\infty(M,\mathbb{R}) \arrow[rr, "\operatorname{grad}^\omega"] & & \mathfrak{X}(M,\omega) \arrow[r, "\gamma"] & H^1(M) \\
& & \mathfrak{g} \arrow[ul, "j"] \arrow[ur, "\zeta"'] & &
\end{tikzcd}$$

A linear lift $j : \mathfrak{g} \to C^\infty(M,\mathbb{R})$ of ζ with $\operatorname{grad}^\omega \circ j = \zeta$ exists if and only if $\gamma \circ \zeta = 0$ in $H^1(M)$. This lift j may be changed to a Lie algebra homomorphism if and only if the 2-cocycle $\bar{j} : \mathfrak{g} \times \mathfrak{g} \to H^0(M)$, given by $(i \circ \bar{j})(X,Y) = \{j(X), j(Y)\} - j([X,Y])$, vanishes in the Lie algebra cohomology $H^2(\mathfrak{g}, H^0(M))$, for if $\bar{j} = \delta\alpha$ then $j - i \circ \alpha$ is a Lie algebra homomorphism.

If $j : \mathfrak{g} \to C^\infty(M,\mathbb{R})$ is a Lie algebra homomorphism, we may associate the *moment mapping* $\mu : M \to \mathfrak{g}' = L(\mathfrak{g},\mathbb{R})$ to it, which is given by $\mu(x)(X) = \chi(X)(x)$ for $x \in M$ and $X \in \mathfrak{g}$. It is G-equivariant for a suitably chosen (in general affine) action of G on $\mathfrak{g}'$.

2.2

We now want to carry over to infinite dimensional manifolds the procedure of subsection (2.1). First we need the appropriate notions in infinite dimensions. So let M be a manifold, which in general is infinite dimensional.

A 2-form $\omega \in \Omega^2(M)$ is called a *weak symplectic structure* on M if it is closed ($d\omega = 0$) and if its associated vector bundle homomorphism $\check{\omega} : TM \to T^*M$ is injective.

A 2-form $\omega \in \Omega^2(M)$ is called a *strong symplectic structure* on M if it is closed ($d\omega = 0$) and if its associated vector bundle homomorphism $\check{\omega} : TM \to T^*M$ is invertible with smooth inverse. In this case, the vector bundle TM has reflexive fibers T_xM: Let $i : T_xM \to (T_xM)''$ be the canonical mapping onto the bidual. Skew symmetry of ω is equivalent to the fact that the transposed $(\check{\omega})^t = (\check{\omega})^* \circ i : T_xM \to (T_xM)'$ satisfies $(\check{\omega})^t = -\check{\omega}$. Thus, $i = -((\check{\omega})^{-1})^* \circ \check{\omega}$ is an isomorphism.

2.3

Every cotangent bundle T^*M, viewed as a manifold, carries a canonical weak symplectic structure $\omega_M \in \Omega^2(T^*M)$, which is defined as follows. Let $\pi_M^* : T^*M \to M$ be the projection. Then the *Liouville form* $\theta_M \in \Omega^1(T^*M)$ is given by $\theta_M(X) = \langle \pi_{T^*M}(X), T(\pi_M^*)(X)\rangle$ for $X \in T(T^*M)$, where $\langle\ \ ,\ \ \rangle$ denotes the duality pairing $T^*M \times_M TM \to \mathbb{R}$. Then the symplectic structure on T^*M is given by $\omega_M = -d\theta_M$, which of course in a local chart looks like $\omega_E((v,v'),(w,w')) = \langle w', v\rangle_E - \langle v', w\rangle_E$. The associated mapping $\check{\omega} : T_{(0,0)}(E \times E') = E \times E' \to E' \times E''$ is given by $(v,v') \mapsto (-v', i_E(v))$, where $i_E : E \to E''$ is the embedding into the bidual. So the canonical symplectic structure on T^*M is strong if and only if all model spaces of the manifold M are reflexive.

2.4

Let M be a weak symplectic manifold. The first thing to note is that the Hamiltonian mapping $\operatorname{grad}^\omega : C^\infty(M,\mathbb{R}) \to \mathfrak{X}(M,\omega)$ does not make sense in general, since $\check{\omega} : TM \to T^*M$ is not invertible. Namely, $\operatorname{grad}^\omega f = (\check{\omega})^{-1} \circ df$ is defined only for those $f \in C^\infty(M,\mathbb{R})$ with $df(x)$ in the image of $\check{\omega}$ for all $x \in M$. A similar difficulty arises for the definition of the Poisson bracket on $C^\infty(M,\mathbb{R})$.

Definition

For a weak symplectic manifold (M,ω) let $T_x^\omega M$ denote the real linear subspace $T_x^\omega M = \check{\omega}_x(T_xM) \subset T_x^*M = L(T_xM,\mathbb{R})$, and let us call it the *smooth cotangent space* with respect to the symplectic structure ω of M at x in view of the embedding of test functions into distributions. These vector spaces fit together to form a subbundle of T^*M which is isomorphic to the tangent bundle TM via $\check{\omega} : TM \to T^\omega M \subseteq T^*M$. It is in general not a splitting subbundle.

2.5. Definition

For a weak symplectic vector space (E,ω) let

$$C^\infty_\omega(E,\mathbb{R}) \subset C^\infty(E,\mathbb{R})$$

denote the linear subspace consisting of all smooth functions $f : E \to \mathbb{R}$ such that each iterated derivative $d^k f(x) \in L^k_{\text{sym}}(E;\mathbb{R})$ has the property that

$$d^k f(x)(\quad, y_2, \dots, y_k) \in E^\omega$$

is actually in the smooth dual $E^\omega \subset E'$ for all $x, y_2, \dots, y_k \in E$, and that the mapping

$$\begin{aligned} \prod^k E &\to E \\ (x, y_2, \dots, y_k) &\mapsto (\check{\omega})^{-1}(df(x)(\quad, y_2, \dots, y_k)) \end{aligned}$$

is smooth. By the symmetry of higher derivatives, this is then true for all entries of $d^k f(x)$, for all x.

2.6. Lemma.
For $f \in C^\infty(E, \mathbb{R})$ the following assertions are equivalent:

(1) $df : E \to E'$ factors to a smooth mapping $E \to E^\omega$.
(2) f has a smooth ω-gradient $\operatorname{grad}^\omega f \in \mathfrak{X}(E) = C^\infty(E, E)$ which satisfies $df(x)y = \omega(\operatorname{grad}^\omega f(x), y)$.
(3) $f \in C^\infty_\omega(E, \mathbb{R})$.

Proof. Clearly, (3) $\Rightarrow$ (2) $\Leftrightarrow$ (1). We have to show that (2) $\Rightarrow$ (3).
Suppose that $f : E \to \mathbb{R}$ is smooth and $df(x)y = \omega(\operatorname{grad}^\omega f(x), y)$. Then

$$\begin{aligned} d^k f(x)(y_1, \dots, y_k) &= d^k f(x)(y_2, \dots, y_k, y_1) \\ &= (d^{k-1}(df))(x)(y_2, \dots, y_k)(y_1) \\ &= \omega\big(d^{k-1}(\operatorname{grad}^\omega f)(x)(y_2, \dots, y_k), y_1\big). \square \end{aligned}$$

2.7

For a weak symplectic manifold (M, ω) let

$$C^\infty_\omega(M, \mathbb{R}) \subset C^\infty(M, \mathbb{R})$$

denote the linear subspace consisting of all smooth functions $f : M \to \mathbb{R}$ such that the differential $df : M \to T^*M$ factors to a smooth mapping $M \to T^\omega M$. In view of lemma (2.6) these are exactly those smooth functions on M which admit a smooth ω-gradient $\operatorname{grad}^\omega f \in \mathfrak{X}(M)$. Also the condition (2.6.1) translates to a local differential condition describing the functions in $C^\infty_\omega(M, \mathbb{R})$.

2.8. Theorem.
The Hamiltonian mapping $\operatorname{grad}^\omega : C^\infty_\omega(M,\mathbb{R}) \to \mathfrak{X}(M,\omega)$*, which is given by*

$$i_{\operatorname{grad}^\omega f}\omega = df \quad \text{or} \quad \operatorname{grad}^\omega f := (\check{\omega})^{-1} \circ df$$

is well defined. Also the Poisson bracket

$$\begin{aligned}\{\quad,\quad\} &: C^\infty_\omega(M,\mathbb{R}) \times C^\infty_\omega(M,\mathbb{R}) \to C^\infty_\omega(M,\mathbb{R})\\ \{f,g\} &:= i_{\operatorname{grad}^\omega f} i_{\operatorname{grad}^\omega g}\omega = \omega(\operatorname{grad}^\omega g, \operatorname{grad}^\omega f) =\\ &= dg(\operatorname{grad}^\omega f) = (\operatorname{grad}^\omega f)(g)\end{aligned}$$

is well defined and gives a Lie algebra structure to the space $C^\infty_\omega(M,\mathbb{R})$*, which also fulfills*

$$\{f,gh\} = \{f,g\}h + g\{f,h\}.$$

We have the following long exact sequence of Lie algebras and Lie algebra homomorphisms:

$$0 \to H^0(M) \to C^\infty_\omega(M,\mathbb{R}) \xrightarrow{\operatorname{grad}^\omega} \mathfrak{X}(M,\omega) \xrightarrow{\gamma} H^1_\omega(M) \to 0,$$

where $H^0(M)$ *is the space of locally constant functions, and*

$$H^1_\omega(M) = \frac{\{\varphi \in C^\infty(M \leftarrow T^\omega M) : d\varphi = 0\}}{\{df : f \in C^\infty_\omega(M,\mathbb{R})\}}$$

is the first symplectic cohomology space of (M,ω)*, a linear subspace of the De Rham cohomology space* $H^1(M)$*.*

Proof. It is clear from lemma (2.6), that the Hamiltonian mapping $\operatorname{grad}^\omega$ is well defined and has values in $\mathfrak{X}(M,\omega)$, since by [31], 34.18.6 we have

$$\mathcal{L}_{\operatorname{grad}^\omega f}\omega = i_{\operatorname{grad}^\omega f} d\omega + d i_{\operatorname{grad}^\omega f}\omega = ddf = 0.$$

By [31], 34.18.7, the space $\mathfrak{X}(M,\omega)$ is a Lie subalgebra of $\mathfrak{X}(M)$. The Poisson bracket is well defined as a mapping $\{\quad,\quad\} : C^\infty_\omega(M,\mathbb{R}) \times C^\infty_\omega(M,\mathbb{R}) \to C^\infty(M,\mathbb{R})$; it only remains to check that it has values in the subspace $C^\infty_\omega(M,\mathbb{R})$.

This is a local question, so we may assume that M is an open subset of a convenient vector space equipped with a (non-constant) weak symplectic structure. So let f, $g \in C^\infty_\omega(M,\mathbb{R})$, then $\{f,g\}(x) = dg(x)(\operatorname{grad}^\omega f(x))$, and we have

$$\begin{aligned} d(\{f,g\})(x)y &= d(dg(\quad)y)(x).\operatorname{grad}^\omega f(x) + dg(x)(d(\operatorname{grad}^\omega f)(x)y)\\ &= d(\omega(\operatorname{grad}^\omega g(\quad),y)(x).\operatorname{grad}^\omega f(x) + \omega\Big(\operatorname{grad}^\omega g(x), d(\operatorname{grad}^\omega f)(x)y\Big)\\ &= \omega\Big(d(\operatorname{grad}^\omega g)(x)(\operatorname{grad}^\omega f(x)) - d(\operatorname{grad}^\omega f)(x)(\operatorname{grad}^\omega g(x)), y\Big),\end{aligned}$$

since $\operatorname{grad}^\omega f \in \mathfrak{X}(M,\omega)$ and for any $X \in \mathfrak{X}(M,\omega)$ the condition $\mathcal{L}_X\omega = 0$ implies $\omega(dX(x)y_1, y_2) = -\omega(y_1, dX(x)y_2)$. So (2.6.2) is satisfied, and thus $\{f,g\} \in C^\infty_\omega(M,\mathbb{R})$.

If $X \in \mathfrak{X}(M,\omega)$ then $di_X\omega = \mathcal{L}_X\omega = 0$, so $[i_X\omega] \in H^1(M)$ is well defined, and by $i_X\omega = \check{\omega}\, oX$ we even have $\gamma(X) := [i_X\omega] \in H^1_\omega(M)$, so γ is well defined.

Now we show that the sequence is exact. Obviously, it is exact at $H^0(M)$ and at $C^\infty_\omega(M,\mathbb{R})$, since the kernel of $\operatorname{grad}^\omega$ consists of the locally constant functions. If $\gamma(X) = 0$ then $\check{\omega} \circ X = i_X\omega = df$ for $f \in C^\infty_\omega(M,\mathbb{R})$, and clearly $X = \operatorname{grad}^\omega f$. Now let us suppose that $\varphi \in \Gamma(T^\omega M) \subset \Omega^1(M)$ with $d\varphi = 0$. Then $X := (\check{\omega})^{-1} \circ \varphi \in \mathfrak{X}(M)$ is well defined and $\mathcal{L}_X\omega = di_X\omega = d\varphi = 0$, so $X \in \mathfrak{X}(M,\omega)$ and $\gamma(X) = [\varphi]$.

Moreover, $H^1_\omega(M)$ is a linear subspace of $H^1(M)$ since for $\varphi \in \Gamma(T^\omega M) \subset \Omega^1(M)$ with $\varphi = df$ for $f \in C^\infty(M,\mathbb{R})$ the vector field $X := (\check{\omega})^{-1} \circ \varphi \in \mathfrak{X}(M)$ is well defined, and since $\check{\omega}\, oX = \varphi = df$ by (2.6.1) we have $f \in C^\infty_\omega(M,\mathbb{R})$ with $X = \operatorname{grad}^\omega f$.

The mapping $\operatorname{grad}^\omega$ maps the Poisson bracket into the Lie bracket, since by [31], 34.18 we have

$$\begin{aligned} i_{\operatorname{grad}^\omega\{f,g\}}\omega &= d\{f,g\} = d\mathcal{L}_{\operatorname{grad}^\omega f}g = \mathcal{L}_{\operatorname{grad}^\omega f}dg = \\ &= \mathcal{L}_{\operatorname{grad}^\omega f} i_{\operatorname{grad}^\omega g}\omega - i_{\operatorname{grad}^\omega g}\mathcal{L}_{\operatorname{grad}^\omega f}\omega \\ &= [\mathcal{L}_{\operatorname{grad}^\omega f}, i_{\operatorname{grad}^\omega g}]\omega = i_{[\operatorname{grad}^\omega f, \operatorname{grad}^\omega g]}\omega. \end{aligned}$$

Let us now check the properties of the Poisson bracket. By definition, it is skew symmetric, and we have

$$\begin{aligned} \{\{f,g\},h\} &= \mathcal{L}_{\operatorname{grad}^\omega\{f,g\}}h = \mathcal{L}_{[\operatorname{grad}^\omega f, \operatorname{grad}^\omega g]}h = [\mathcal{L}_{\operatorname{grad}^\omega f}, \mathcal{L}_{\operatorname{grad}^\omega g}]h = \\ &= \mathcal{L}_{\operatorname{grad}^\omega f}\mathcal{L}_{\operatorname{grad}^\omega g}h - \mathcal{L}_{\operatorname{grad}^\omega g}\mathcal{L}_{\operatorname{grad}^\omega f}h = \{f,\{g,h\}\} - \{g,\{f,h\}\} \\ \{f,gh\} &= \mathcal{L}_{\operatorname{grad}^\omega f}(gh) = (\mathcal{L}_{\operatorname{grad}^\omega f}g)h + g\mathcal{L}_{\operatorname{grad}^\omega f}h = \\ &= \{f,g\}h + g\{f,h\}. \end{aligned}$$

Finally, it remains to show that all mappings in the sequence are Lie algebra homomorphisms, where we put the zero bracket on both cohomology spaces. For locally constant functions we have $\{c_1,c_2\} = \mathcal{L}_{\operatorname{grad}^\omega c_1}c_2 = 0$. We have already checked that $\operatorname{grad}^\omega$ is a Lie algebra homomorphism. For $X, Y \in \mathfrak{X}(M,\omega)$

$$i_{[X,Y]}\omega = [\mathcal{L}_X, i_Y]\omega = \mathcal{L}_X i_Y\omega + 0 = di_Xi_Y\omega + i_X\mathcal{L}_Y\omega = di_Xi_Y\omega$$

is exact. $\square$

2.9. Weakly symplectic group actions

Let us suppose that an infinite dimensional regular Lie group G with Lie algebra $\mathfrak{g}$ acts from the right on a weak symplectic manifold (M,ω) by $r : M \times G \to M$ in a way which respects ω, so that each transformation r^g is a symplectomorphism. This is called a *symplectic group action.* We shall use the notation $r(x,g) = r^g(x) = r_x(g)$. Let us list some immediate consequences:

(1) The space $C^\infty_\omega(M)^G$ of G-invariant smooth functions with ω-gradients is a Lie subalgebra for the Poisson bracket, since for each $g \in G$ and $f,h \in C^\infty(M)^G$ we have $(r^g)^*\{f,h\} = \{(r^g)^*f,(r^g)^*h\} = \{f,h\}$.

(2) For $x \in M$ the pullback of ω to the orbit $x.G$ is a 2-form, invariant under the action of G on the orbit. In the finite dimensional case the orbit is an initial submanifold. In our case this has to be checked directly in each example. In any case we have something like a tangent bundle $T_x(x.G) = T(r_x)\mathfrak{g}$. If $i : x.G \to M$ is the embedding of the orbit then $r^g \circ i = i \circ r^g$, so that $i^*\omega = i^*(r^g)^*\omega = (r^g)^*i^*\omega$ holds for each $g \in G$ and thus $i^*\omega$ is invariant.

(3) The fundamental vector field mapping $\zeta : \mathfrak{g} \to \mathfrak{X}(M,\omega)$, given by $\zeta_X(x) = T_e(r_x)X$ for $X \in \mathfrak{g}$ and $x \in M$, is a homomorphism of Lie algebras, where $\mathfrak{g}$ is the Lie algebra of G (for a left action we get an anti homomorphism of Lie algebras). Moreover, ζ takes values in $\mathfrak{X}(M,\omega)$. Let us consider again the exact sequence of Lie algebra homomorphisms from (2.8):

$$\begin{tikzcd}
0 \arrow[r] & H^0(M) \arrow[r, "\alpha"] & C^\infty_\omega(M) \arrow[r, "\mathrm{grad}^\omega"] & \mathfrak{X}(M,\omega) \arrow[r, "\gamma"] & H^1_\omega(M) \arrow[r] & 0 \\
 & & & \mathfrak{g} \arrow[u, "\zeta"] \arrow[ul, dotted, "j"] & &
\end{tikzcd}$$

One can lift ζ to a linear mapping $j : \mathfrak{g} \to C^\infty(M)$ if and only if $\gamma \circ \zeta = 0$. In this case the action of G is called a *Hamiltonian group action*, and the linear mapping $j : \mathfrak{g} \to C^\infty(M)$ is called a *generalized Hamiltonian function* for the group action. It is unique up to addition of a mapping $\alpha \circ \tau$ for $\tau : \mathfrak{g} \to H^0(M)$.

(4) If $H^1_\omega(M) = 0$ then any symplectic action on (M,ω) is a Hamiltonian action. But if $\gamma\circ\zeta \neq 0$ we can replace $\mathfrak{g}$ by its Lie subalgebra $\ker(\gamma\circ\zeta) \subset \mathfrak{g}$ *and consider the corresponding Lie subgroup G which then admits a Hamiltonian action.*

(5) If the Lie algebra $\mathfrak{g}$ is equal to its commutator subalgebra $[\mathfrak{g},\mathfrak{g}]$, the linear span of all $[X,Y]$ for $X,Y \in \mathfrak{g}$ (true for all full diffeomorphism groups), then any infinitesimal symplectic action $\zeta : \mathfrak{g} \to \mathfrak{X}(M,\omega)$ is a Hamiltonian action, since then any $Z \in \mathfrak{g}$ can be written as $Z = \sum_i [X_i,Y_i]$ so that $\zeta_Z = \sum[\zeta_{X_i},\zeta_{Y_i}] \in \mathrm{im}(\mathrm{grad}^\omega)$ since $\gamma : \mathfrak{X}(M,\omega) \to H^1(M)$ is a homomorphism into the zero Lie bracket.

(6) If $j : \mathfrak{g} \to (C^\infty_\omega(M), \{\ \ ,\ \ \})$ happens to be not a homomorphism of Lie algebras then $c(X,Y) = \{j(X), j(Y)\} - j([X,Y])$ lies in $H^0(M)$, and indeed $c : \mathfrak{g} \times \mathfrak{g} \to H^0(M)$ is a cocycle for the Lie algebra cohomology: $c([X,Y],Z) + c([Y,Z],X) + c([Z,X],Y) = 0$. If c is a coboundary, i.e., $c(X,Y) = -b([X,Y])$, then $j + \alpha \circ b$ is a Lie algebra homomorphism. If the cocycle c is non-trivial we can use the central extension $H^0(M) \times_c \mathfrak{g}$ with bracket $[(a,X),(b,Y)] = (c(X,Y),[X,Y])$ in the diagram

$$\begin{tikzcd}
0 \arrow[r] & H^0(M) \arrow[r, "\alpha"] & C^\infty_\omega(M) \arrow[r, "\mathrm{grad}^\omega"] & \mathfrak{X}(M,\omega) \arrow[r, "\gamma"] & H^1_\omega(M) \arrow[r] & 0 \\
 & & H^1(M) \times_c \mathfrak{g} \arrow[u, "\bar{j}"] \arrow[r, "\mathrm{pr}_2"] & \mathfrak{g} \arrow[u, "\zeta"'] & &
\end{tikzcd}$$

where $\bar{j}(a,X) = j(X) + \alpha(a)$. Then $\bar{j}$ is a homomorphism of Lie algebras.

2.10. Momentum mapping.
For an infinitesimal symplectic action, i.e. a homomorphism $\zeta : \mathfrak{g} \to \mathfrak{X}(M,\omega)$ of Lie algebras, we can find a linear lift $j : \mathfrak{g} \to C^\infty_\omega(M)$ if and only if there exists a mapping

$$J \in C^\infty_\omega(M, \mathfrak{g}^*) := \{f \in C^\infty(M, \mathfrak{g}^*) : \langle f(\ \ ), X\rangle \in C^\infty_\omega(M) \text{ for all } X \in \mathfrak{g}\}$$

such that

$$\mathrm{grad}^\omega(\langle J, X\rangle) = \zeta_X \qquad \text{for all } X \in \mathfrak{g}.$$

The mapping $J \in C^\infty_\omega(M, \mathfrak{g}^*)$ is called the *momentum mapping* for the infinitesimal action $\zeta : \mathfrak{g} \to \mathfrak{X}(M,\omega)$. Let us note again the relations between the generalized Hamiltonian j and the momentum mapping J:

$$\begin{gathered}
J : M \to \mathfrak{g}^*, \quad j : \mathfrak{g} \to C^\infty_\omega(M), \quad \zeta : \mathfrak{g} \to \mathfrak{X}(M,\omega) \\
\langle J, X\rangle = j(X) \in C^\infty_\omega(M), \quad \mathrm{grad}^\omega(j(X)) = \zeta(X), \quad X \in \mathfrak{g}, \\
i_{\zeta(X)}\omega = dj(X) = d\langle J, X\rangle,
\end{gathered} \tag{1}$$

where $\langle\ ,\ \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$ is the duality pairing.

2.11. Basic properties of the momentum mapping

Let $r : M \times G \to M$ be a Hamiltonian right action of an infinite dimensional regular Lie group G on a weak symplectic manifold M, let $j : \mathfrak{g} \to C^\infty_\omega(M)$ be a generalized Hamiltonian and let $J \in C^\infty_\omega(M, \mathfrak{g}^*)$ be the associated momentum mapping.

(1) For $x \in M$, the transposed mapping of the linear mapping $dJ(x) : T_xM \to \mathfrak{g}^*$ is

$$dJ(x)^\top : \mathfrak{g} \to T_x^*M, \qquad dJ(x)^\top = \check{\omega}_x \circ \zeta,$$

since for $\xi \in T_xM$ and $X \in \mathfrak{g}$ we have

$$\langle dJ(\xi), X\rangle = \langle i_\xi dJ, X\rangle = i_\xi d\langle J, X\rangle = i_\xi i_{\zeta_X}\omega = \langle \check{\omega}_x(\zeta_X(x)), \xi\rangle.$$

(2) The closure of the image $dJ(T_xM)$ of $dJ(x) : T_xM \to \mathfrak{g}*$ is the annihilator $\mathfrak{g}_x^\circ$ of the isotropy Lie algeba $\mathfrak{g}_x := \{X \in \mathfrak{g} : \zeta_X(x) = 0\}$ in $\mathfrak{g}^*$, since the annihilator of the image is the kernel of the transposed mapping,

$$\operatorname{im}(dJ(x))^\circ = \ker(dJ(x)^\top) = \ker(\check{\omega}_x \circ \zeta) = \ker(\operatorname{ev}_x \circ \zeta) = \mathfrak{g}_x.$$

(3) The kernel of $dJ(x)$ is the symplectic orthogonal

$$(T(r_x)\mathfrak{g})^{\perp,\omega} = (T_x(x.G))^{\perp,\omega} \subseteq T_xM,$$

since for the annihilator of the kernel we have

$$\begin{aligned}\ker(dJ(x))^\circ = \overline{\operatorname{im}(dJ(x)^\top)} = \overline{\operatorname{im}(\check{\omega}_x \circ \zeta)} = \\ = \overline{\{\check{\omega}_x(\zeta_X(x)) : X \in \mathfrak{g}\}} = \overline{\check{\omega}_x(T_x(x.G))}.\end{aligned}$$

(4) If G is connected, $x \in M$ is a fixed point for the G-action if and only if x is a critical point of J, i.e. $dJ(x) = 0$.

(5) (Emmy Noether's theorem) Let $h \in C^\infty_\omega(M)$ be a Hamiltonian function which is invariant under the Hamiltonian G action. Then $dJ(\operatorname{grad}^\omega(h)) = 0$. Thus the momentum mapping $J : M \to \mathfrak{g}^*$ is constant on each trajectory (if it exists) of the Hamiltonian vector field $\operatorname{grad}^\omega(h)$. Namely,

$$\begin{aligned}\langle dJ(\operatorname{grad}^\omega(h)), X\rangle = d\langle J, X\rangle(\operatorname{grad}^\omega(h)) = dj(X)(\operatorname{grad}^\omega(h)) = \\ = \{h, j(X)\} = -dh(\operatorname{grad}^\omega j(X)) = dh(\zeta_X) = 0.\end{aligned}$$

E. Noether's theorem admits the following generalization.

2.12. Theorem.
Let G_1 and G_2 be two regular Lie groups which act by Hamiltonian actions r_1 and r_2 on the weakly symplectic manifold (M, ω), with momentum mappings J_1 and J_2, respectively. We assume that J_2 is G_1-invariant, i.e. J_2 is constant along all G_1-orbits, and that G_2 is connected.

Then J_1 is constant on the G_2-orbits and the two actions commute.

Proof. Let $\zeta^i : \mathfrak{g}_i \to \mathfrak{X}(M, \omega)$ be the two infinitesimal actions. Then for $X_1 \in \mathfrak{g}_1$ and $X_2 \in \mathfrak{g}_2$ we have

$$\begin{aligned}\mathcal{L}_{\zeta^2_{X_2}}\langle J_1, X_1\rangle = i_{\zeta^2_{X_2}} d\langle J_1, X_1\rangle = i_{\zeta^2_{X_2}} i_{\zeta^1_{X_1}}\omega = \{\langle J_2, X_2\rangle, \langle J_1, X_1\rangle\} \\ = -\{\langle J_1, X_1\rangle, \langle J_2, X_2\rangle\} = -i_{\zeta^1_{X_1}} d\langle J_2, X_2\rangle = -\mathcal{L}_{\zeta^1_{X_1}}\langle J_2, X_2\rangle = 0\end{aligned}$$

since J_2 is constant along each G_1-orbit. Since G_2 is assumed to be connected, J_1 is also constant along each G_2-orbit. We also saw that each Poisson bracket $\{\langle J_2, X_2\rangle, \langle J_1, X_1\rangle\}$ vanishes; by $\operatorname{grad}^\omega \langle J_i, X_i\rangle = \zeta^i_{X_i}$ we conclude that $[\zeta^1_{X_1}, \zeta^2_{X_2}] = 0$ for all $X_i \in \mathfrak{g}_i$ which implies the result if also G_1 is connected. In the general case we can argue as follows:

$$\begin{aligned}(r_1^{g_1})^*\zeta^2_{X_2} &= (r_1^{g_1})^* \operatorname{grad}^\omega \langle J_2, X_2\rangle = (r_1^{g_1})^*(\check{\omega}^{-1} d\langle J_2, X_2\rangle)\\ &= (((r_1^{g_1})^*\omega)\check{})^{-1} d\langle (r_1^{g_1})^* J_2, X_2\rangle = (\check{\omega}^{-1} d\langle J_2, X_2\rangle = \operatorname{grad}^\omega \langle J_2, X_2\rangle = \zeta^2_{X_2}.\end{aligned}$$

Thus $r_1^{g_1}$ commutes with each $r_2^{\exp(tX_2)}$ and thus with each $r_2^{g_2}$, since G_2 is connected. $\square$

3. Right invariant weak Riemannian metrics on Lie groups

3.1. Notation on Lie groups

Let G be a Lie group which may be infinite dimensional, but then is supposed to be regular, with Lie algebra $\mathfrak{g}$. See appendix (B) for more information. Let $\mu : G \times G \to G$ be the multiplication, let μ_x be left translation and μ^y be right translation, given by $\mu_x(y) = \mu^y(x) = xy = \mu(x, y)$.

Let $L, R : \mathfrak{g} \to \mathfrak{X}(G)$ be the left and right invariant vector field mappings, given by $L_X(g) = T_e(\mu_g).X$ and $R_X = T_e(\mu^g).X$, respectively. They are related by $L_X(g) = R_{\operatorname{Ad}(g)X}(g)$. Their flows are given by

$$\operatorname{Fl}^{L_X}_t(g) = g.\exp(tX) = \mu^{\exp(tX)}(g), \quad \operatorname{Fl}^{R_X}_t(g) = \exp(tX).g = \mu_{\exp(tX)}(g).$$

We also need the right Maurer-Cartan form $\kappa = \kappa^r \in \Omega^1(G, \mathfrak{g})$, given by $\kappa_x(\xi) := T_x(\mu^{x^{-1}}) \cdot \xi$. It satisfies the right Maurer-Cartan equation $d\kappa - \frac{1}{2}[\kappa, \kappa]_\wedge = 0$, where $[\ ,\ ]_\wedge$ denotes the wedge product of $\mathfrak{g}$-valued forms on G induced by the Lie bracket. Note that $\frac{1}{2}[\kappa, \kappa]_\wedge(\xi, \eta) = [\kappa(\xi), \kappa(\eta)]$. The (exterior) derivative of the function $\operatorname{Ad} : G \to GL(\mathfrak{g})$ can be expressed by

$$d \operatorname{Ad} = \operatorname{Ad}.(\operatorname{ad} \circ \kappa^l) = (\operatorname{ad} \circ \kappa^r). \operatorname{Ad},$$

since we have $d \operatorname{Ad}(T\mu_g.X) = \frac{d}{dt}|_0 \operatorname{Ad}(g.\exp(tX)) = \operatorname{Ad}(g). \operatorname{ad}(\kappa^l(T\mu_g.X))$.

3.2. Geodesics of a right invariant metric on a Lie group

Let $\gamma = \langle\ ,\ \rangle : \mathfrak{g}\times\mathfrak{g} \to \mathbb{R}$ be a positive definite bounded (weak) inner product. Then

$$\gamma_x(\xi,\eta) = \langle T(\mu^{x^{-1}})\cdot\xi, T(\mu^{x^{-1}})\cdot\eta)\rangle = \langle\kappa(\xi),\kappa(\eta)\rangle \tag{1}$$

is a right invariant (weak) Riemannian metric on G, and any (weak) right invariant bounded Riemannian metric is of this form, for suitable $\langle\ ,\ \rangle$.

Let $g : [a,b] \to G$ be a smooth curve. The velocity field of g, viewed in the right trivializations, coincides with the right logarithmic derivative

$$\delta^r(g) = T(\mu^{g^{-1}})\cdot\partial_t g = \kappa(\partial_t g) = (g^*\kappa)(\partial_t),\ \text{ where } \partial_t = \frac{\partial}{\partial t}.$$

The energy of the curve $g(t)$ is given by

$$E(g) = \tfrac{1}{2}\int_a^b G_g(g',g')dt = \tfrac{1}{2}\int_a^b \langle (g^*\kappa)(\partial_t),(g^*\kappa)(\partial_t)\rangle\, dt.$$

For a variation $g(s,t)$ with fixed endpoints we have then, using the right Maurer-Cartan equation and integration by parts,

$$\begin{aligned}
\partial_s E(g) &= \tfrac{1}{2}\int_a^b 2\langle \partial_s(g^*\kappa)(\partial_t),\ (g^*\kappa)(\partial_t)\rangle\, dt\\
&= \int_a^b \langle \partial_t(g^*\kappa)(\partial_s) - d(g^*\kappa)(\partial_t,\partial_s),\ (g^*\kappa)(\partial_t)\rangle\, dt\\
&= \int_a^b \left(-\langle (g^*\kappa)(\partial_s),\ \partial_t(g^*\kappa)(\partial_t)\rangle - \langle [(g^*\kappa)(\partial_t),(g^*\kappa)(\partial_s)],\ (g^*\kappa)(\partial_t)\rangle\right)\, dt\\
&= -\int_a^b \langle (g^*\kappa)(\partial_s),\ \partial_t(g^*\kappa)(\partial_t) + \operatorname{ad}((g^*\kappa)(\partial_t))^\top((g^*\kappa)(\partial_t))\rangle\, dt
\end{aligned}$$

where $\operatorname{ad}((g^*\kappa)(\partial_t))^\top : \mathfrak{g}\to\mathfrak{g}$ is the adjoint of $\operatorname{ad}((g^*\kappa)(\partial_t))$ with respect to the inner product $\langle\ ,\ \rangle$. In infinite dimensions one also has to check the existence of this adjoint. In terms of the right logarithmic derivative $u : [a,b]\to\mathfrak{g}$ of $g : [a,b]\to G$, given by $u(t) := g^*\kappa(\partial_t) = T_{g(t)}(\mu^{g(t)^{-1}})\cdot g'(t)$, the geodesic equation has the expression:

$$\boxed{u_t = -\operatorname{ad}(u)^\top u} \tag{2}$$

This is, of course, just the Euler-Poincaré equation for right invariant systems using the Lagrangian given by the kinetic energy (see [34], section 13).

3.3. The covariant derivative

Our next aim is to derive the Riemannian curvature and for that we develop the basis-free version of Cartan's method of moving frames in this setting, which also works in infinite dimensions. The right trivialization, or framing, $(\pi_G, \kappa) : TG \to G \times \mathfrak{g}$ induces the isomorphism $R : C^\infty(G, \mathfrak{g}) \to \mathfrak{X}(G)$, given by $R(X)(x) := R_X(x) := T_e(\mu^x) \cdot X(x)$, for $X \in C^\infty(G, \mathfrak{g})$ and $x \in G$. Here $\mathfrak{X}(G) := \Gamma(TG)$ denote the Lie algebra of all vector fields. For the Lie bracket and the Riemannian metric we have

$$\begin{aligned}
[R_X, R_Y] &= R(-[X,Y]_{\mathfrak{g}} + dY \cdot R_X - dX \cdot R_Y), \qquad (1)\\
R^{-1}[R_X, R_Y] &= -[X,Y]_{\mathfrak{g}} + R_X(Y) - R_Y(X),\\
\gamma_x(R_X(x), R_Y(x)) &= \gamma(X(x), Y(x))\,,\ x \in G.
\end{aligned}$$

In the sequel we shall compute in $C^\infty(G, \mathfrak{g})$ instead of $\mathfrak{X}(G)$. In particular, we shall use the convention

$$\nabla_X Y := R^{-1}(\nabla_{R_X} R_Y) \quad \text{ for } X, Y \in C^\infty(G, \mathfrak{g}).$$

to express the Levi-Civita covariant derivative.

Lemma.
Assume that for all $\xi \in \mathfrak{g}$ the adjoint $\operatorname{ad}(\xi)^\top$ with respect to the inner product $\langle\ ,\ \rangle$ exists and that $\xi \mapsto \operatorname{ad}(\xi)^\top$ is bounded. Then the Levi-Civita covariant derivative of the metric (3.2.1) exists and is given for any $X, Y \in C^\infty(G, \mathfrak{g})$ in terms of the isomorphism R by

$$\nabla_X Y = dY.R_X + \tfrac{1}{2}\operatorname{ad}(X)^\top Y + \tfrac{1}{2}\operatorname{ad}(Y)^\top X - \tfrac{1}{2}\operatorname{ad}(X)Y. \qquad (2)$$

Proof. Easy computations show that this formula satisfies the axioms of a covariant derivative, that relative to it the Riemannian metric is covariantly constant, since

$$R_X\gamma(Y,Z) = \gamma(dY.R_X, Z) + \gamma(Y, dZ.R_X) = \gamma(\nabla_X Y, Z) + \gamma(Y, \nabla_X Z),$$

and that it is torsion free, since

$$\nabla_X Y - \nabla_Y X + [X,Y]_{\mathfrak{g}} - dY.R_X + dX.R_Y = 0.\square$$

For $\xi \in \mathfrak{g}$ define $\alpha(\xi) : \mathfrak{g} \to \mathfrak{g}$ by $\alpha(\xi)\eta := \operatorname{ad}(\eta)^\top\xi$. With this notation, the previous lemma states that for all $X \in C^\infty(G, \mathfrak{g})$ the covariant derivative of the Levi-Civita connection has the expression

$$\nabla_X = R_X + \tfrac{1}{2}\operatorname{ad}(X)^\top + \tfrac{1}{2}\alpha(X) - \tfrac{1}{2}\operatorname{ad}(X). \qquad (3)$$

3.4. The curvature

First note that we have the following relations:

$$\begin{aligned}
[R_X, \operatorname{ad}(Y)] &= \operatorname{ad}(R_X(Y)), & [R_X, \alpha(Y)] &= \alpha(R_X(Y)), \\
[R_X, \operatorname{ad}(Y)^\top] &= \operatorname{ad}(R_X(Y))^\top, & [\operatorname{ad}(X)^\top, \operatorname{ad}(Y)^\top] &= -\operatorname{ad}([X,Y]_{\mathfrak g})^\top.
\end{aligned} \tag{1}$$

The Riemannian curvature is then computed by

$$\begin{aligned}
\mathcal{R}(X,Y) &= [\nabla_X, \nabla_Y] - \nabla_{-[X,Y]_{\mathfrak g} + R_X(Y) - R_Y(X)} \\
&= [R_X + \tfrac12 \operatorname{ad}(X)^\top + \tfrac12 \alpha(X) - \tfrac12 \operatorname{ad}(X), R_Y + \tfrac12 \operatorname{ad}(Y)^\top + \tfrac12\alpha(Y) - \tfrac12 \operatorname{ad}(Y)] \\
&\quad - R_{-[X,Y]_{\mathfrak g} + R_X(Y) - R_Y(X)} - \tfrac12 \operatorname{ad}(-[X,Y]_{\mathfrak g} + R_X(Y) - R_Y(X))^\top \\
&\quad - \tfrac12 \alpha(-[X,Y]_{\mathfrak g} + R_X(Y) - R_Y(X)) + \tfrac12 \operatorname{ad}(-[X,Y]_{\mathfrak g} + R_X(Y) - R_Y(X)) \\
&= -\tfrac14 [\operatorname{ad}(X)^\top + \operatorname{ad}(X), \operatorname{ad}(Y)^\top + \operatorname{ad}(Y)] \qquad (2) \\
&\quad + \tfrac14 [\operatorname{ad}(X)^\top - \operatorname{ad}(X), \alpha(Y)] + \tfrac14 [\alpha(X), \operatorname{ad}(Y)^\top - \operatorname{ad}(Y)] \\
&\quad + \tfrac14 [\alpha(X), \alpha(Y)] + \tfrac12 \alpha([X,Y]_{\mathfrak g}).
\end{aligned}$$

If we plug in all definitions and use 4 times the Jacobi identity we get the following expression

$$\begin{aligned}
\gamma(4\mathcal{R}(X,Y)Z,U) &= +2\gamma([X,Y],[Z,U]) - \gamma([Y,Z],[X,U]) + \gamma([X,Z],[Y,U]) \\
&\quad - \gamma(Z,[U,[X,Y]]) + \gamma(U,[Z,[X,Y]]) - \gamma(Y,[X,[U,Z]]) - \gamma(X,[Y,[Z,U]]) \\
&\quad + \gamma(\operatorname{ad}(X)^\top Z, \operatorname{ad}(Y)^\top U) + \gamma(\operatorname{ad}(X)^\top Z, \operatorname{ad}(U)^\top Y) \\
&\quad + \gamma(\operatorname{ad}(Z)^\top X, \operatorname{ad}(Y)^\top U) - \gamma(\operatorname{ad}(U)^\top X, \operatorname{ad}(Y)^\top Z) \qquad (3) \\
&\quad - \gamma(\operatorname{ad}(Y)^\top Z, \operatorname{ad}(X)^\top U) - \gamma(\operatorname{ad}(Z)^\top Y, \operatorname{ad}(X)^\top U) \\
&\quad - \gamma(\operatorname{ad}(U)^\top X, \operatorname{ad}(Z)^\top Y) + \gamma(\operatorname{ad}(U)^\top Y, \operatorname{ad}(Z)^\top X).
\end{aligned}$$

This yields the following expression which is useful for computing the sectional curvature:

$$\begin{aligned}
4\gamma(\mathcal{R}(X,Y)X,Y) &= 3\gamma(\operatorname{ad}(X)Y, \operatorname{ad}(X)Y) - 2\gamma(\operatorname{ad}(Y)^\top X, \operatorname{ad}(X)Y) \\
&\quad - 2\gamma(\operatorname{ad}(X)^\top Y, \operatorname{ad}(Y)X) + 4\gamma(\operatorname{ad}(X)^\top X, \operatorname{ad}(Y)^\top Y) \qquad (4) \\
&\quad - \gamma(\operatorname{ad}(X)^\top Y + \operatorname{ad}(Y)^\top X, \operatorname{ad}(X)^\top Y + \operatorname{ad}(Y)^\top X).
\end{aligned}$$

3.5. Jacobi fields, I

We compute first the Jacobi equation directly via variations of geodesics. So let $g : \mathbb{R}^2 \to G$ be smooth, $t \mapsto g(t,s)$ a geodesic for each s. Let again $u = \kappa(\partial_t g) = (g^*\kappa)(\partial_t)$ be the velocity field along the geodesic in right trivialization which satisfies the geodesic equation $u_t = -\operatorname{ad}(u)^\top u$. Then $y := \kappa(\partial_s g) = (g^*\kappa)(\partial_s)$ is the Jacobi field corresponding to this variation, written in the right trivialization. From the right Maurer-Cartan equation we then have:

$$\begin{aligned}
y_t &= \partial_t(g^*\kappa)(\partial_s) = d(g^*\kappa)(\partial_t, \partial_s) + \partial_s(g^*\kappa)(\partial_t) + 0 \\
&= [(g^*\kappa)(\partial_t), (g^*\kappa)(\partial_s)]_{\mathfrak{g}} + u_s \\
&= [u, y] + u_s.
\end{aligned}$$

Using the geodesic equation, the definition of α, and the fourth relation in (3.4.1), this identity implies

$$\begin{aligned}
u_{st} = u_{ts} = \partial_s u_t &= -\partial_s(\operatorname{ad}(u)^\top u) = -\operatorname{ad}(u_s)^\top u - \operatorname{ad}(u)^\top u_s \\
&= -\operatorname{ad}(y_t + [y,u])^\top u - \operatorname{ad}(u)^\top (y_t + [y,u]) \\
&= -\alpha(u) y_t - \operatorname{ad}([y,u])^\top u - \operatorname{ad}(u)^\top y_t - \operatorname{ad}(u)^\top([y,u]) \\
&= -\operatorname{ad}(u)^\top y_t - \alpha(u) y_t + [\operatorname{ad}(y)^\top, \operatorname{ad}(u)^\top] u - \operatorname{ad}(u)^\top \operatorname{ad}(y) u\,.
\end{aligned}$$

Finally we get the Jacobi equation as

$$\begin{aligned}
y_{tt} &= [u_t, y] + [u, y_t] + u_{st} \\
&= \operatorname{ad}(y)\operatorname{ad}(u)^\top u + \operatorname{ad}(u) y_t - \operatorname{ad}(u)^\top y_t \\
&\qquad - \alpha(u) y_t + [\operatorname{ad}(y)^\top, \operatorname{ad}(u)^\top] u - \operatorname{ad}(u)^\top \operatorname{ad}(y) u\,, \\
y_{tt} &= [\operatorname{ad}(y)^\top + \operatorname{ad}(y), \operatorname{ad}(u)^\top] u - \operatorname{ad}(u)^\top y_t - \alpha(u) y_t + \operatorname{ad}(u) y_t\,. \qquad (1)
\end{aligned}$$

3.6. Jacobi fields, II

Let y be a Jacobi field along a geodesic g with right trivialized velocity field u. Then y should satisfy the analogue of the finite dimensional Jacobi equation

$$\nabla_{\partial_t}\nabla_{\partial_t} y + \mathcal{R}(y,u)u = 0$$

We want to show that this leads to same equation as (3.5.1). First note that from (3.3.2) we have

$$\nabla_{\partial_t} y = y_t + \tfrac{1}{2}\operatorname{ad}(u)^\top y + \tfrac{1}{2}\alpha(u) y - \tfrac{1}{2}\operatorname{ad}(u) y$$

so that, using $u_t = -\operatorname{ad}(u)^\top u$, we get:

$$\begin{aligned}
\nabla_{\partial_t}\nabla_{\partial_t} y &= \nabla_{\partial_t}\Big(y_t + \tfrac{1}{2}\operatorname{ad}(u)^\top y + \tfrac{1}{2}\alpha(u) y - \tfrac{1}{2}\operatorname{ad}(u) y\Big) \\
&= y_{tt} + \tfrac{1}{2}\operatorname{ad}(u_t)^\top y + \tfrac{1}{2}\operatorname{ad}(u)^\top y_t + \tfrac{1}{2}\alpha(u_t) y
\end{aligned}$$

$$
\begin{aligned}
&+\tfrac{1}{2}\alpha(u)y_t-\tfrac{1}{2}\operatorname{ad}(u_t)y-\tfrac{1}{2}\operatorname{ad}(u)y_t\\
&+\tfrac{1}{2}\operatorname{ad}(u)^\top\Big(y_t+\tfrac{1}{2}\operatorname{ad}(u)^\top y+\tfrac{1}{2}\alpha(u)y-\tfrac{1}{2}\operatorname{ad}(u)y\Big)\\
&+\tfrac{1}{2}\alpha(u)\Big(y_t+\tfrac{1}{2}\operatorname{ad}(u)^\top y+\tfrac{1}{2}\alpha(u)y-\tfrac{1}{2}\operatorname{ad}(u)y\Big)\\
&-\tfrac{1}{2}\operatorname{ad}(u)\Big(y_t+\tfrac{1}{2}\operatorname{ad}(u)^\top y+\tfrac{1}{2}\alpha(u)y-\tfrac{1}{2}\operatorname{ad}(u)y\Big)\\
=\;&y_{tt}+\operatorname{ad}(u)^\top y_t+\alpha(u)y_t-\operatorname{ad}(u)y_t\\
&-\tfrac{1}{2}\alpha(y)\operatorname{ad}(u)^\top u-\tfrac{1}{2}\operatorname{ad}(y)^\top\operatorname{ad}(u)^\top u-\tfrac{1}{2}\operatorname{ad}(y)\operatorname{ad}(u)^\top u\\
&+\tfrac{1}{2}\operatorname{ad}(u)^\top\Big(\tfrac{1}{2}\alpha(y)u+\tfrac{1}{2}\operatorname{ad}(y)^\top u+\tfrac{1}{2}\operatorname{ad}(y)u\Big)\\
&+\tfrac{1}{2}\alpha(u)\Big(\tfrac{1}{2}\alpha(y)u+\tfrac{1}{2}\operatorname{ad}(y)^\top u+\tfrac{1}{2}\operatorname{ad}(y)u\Big)\\
&-\tfrac{1}{2}\operatorname{ad}(u)\Big(\tfrac{1}{2}\alpha(y)u+\tfrac{1}{2}\operatorname{ad}(y)^\top u+\tfrac{1}{2}\operatorname{ad}(y)u\Big)\,.
\end{aligned}
$$

In the second line of the last expression we use

$$
-\tfrac{1}{2}\alpha(y)\operatorname{ad}(u)^\top u=-\tfrac{1}{4}\alpha(y)\operatorname{ad}(u)^\top u-\tfrac{1}{4}\alpha(y)\alpha(u)u
$$

and similar forms for the other two terms to get:

$$
\begin{aligned}
\nabla_{\partial_t}\nabla_{\partial_t}y=\;&y_{tt}+\operatorname{ad}(u)^\top y_t+\alpha(u)y_t-\operatorname{ad}(u)y_t\\
&+\tfrac{1}{4}[\operatorname{ad}(u)^\top,\alpha(y)]u+\tfrac{1}{4}[\operatorname{ad}(u)^\top,\operatorname{ad}(y)^\top]u+\tfrac{1}{4}[\operatorname{ad}(u)^\top,\operatorname{ad}(y)]u\\
&+\tfrac{1}{4}[\alpha(u),\alpha(y)]u+\tfrac{1}{4}[\alpha(u),\operatorname{ad}(y)^\top]u+\tfrac{1}{4}[\alpha(u),\operatorname{ad}(y)]u\\
&-\tfrac{1}{4}[\operatorname{ad}(u),\alpha(y)]u-\tfrac{1}{4}[\operatorname{ad}(u),\operatorname{ad}(y)^\top+\operatorname{ad}(y)]u,
\end{aligned}
$$

where in the last line we also used $\operatorname{ad}(u)u=0$. We now compute the curvature term using (3.4.2):

$$
\begin{aligned}
\mathcal{R}(y,u)u=\;&-\tfrac{1}{4}[\operatorname{ad}(y)^\top+\operatorname{ad}(y),\operatorname{ad}(u)^\top+\operatorname{ad}(u)]u\\
&+\tfrac{1}{4}[\operatorname{ad}(y)^\top-\operatorname{ad}(y),\alpha(u)]u+\tfrac{1}{4}[\alpha(y),\operatorname{ad}(u)^\top-\operatorname{ad}(u)]u\\
&+\tfrac{1}{4}[\alpha(y),\alpha(u)]+\tfrac{1}{2}\alpha([y,u])u\\
=\;&-\tfrac{1}{4}[\operatorname{ad}(y)^\top+\operatorname{ad}(y),\operatorname{ad}(u)^\top]u-\tfrac{1}{4}[\operatorname{ad}(y)^\top+\operatorname{ad}(y),\operatorname{ad}(u)]u\\
&+\tfrac{1}{4}[\operatorname{ad}(y)^\top,\alpha(u)]u-\tfrac{1}{4}[\operatorname{ad}(y),\alpha(u)]u+\tfrac{1}{4}[\alpha(y),\operatorname{ad}(u)^\top-\operatorname{ad}(u)]u\\
&+\tfrac{1}{4}[\alpha(y),\alpha(u)]u+\tfrac{1}{2}\operatorname{ad}(u)^\top\operatorname{ad}(y)u\,.
\end{aligned}
$$

Summing up we get

$$
\nabla_{\partial_t}\nabla_{\partial_t}y+\mathcal{R}(y,u)u=y_{tt}+\operatorname{ad}(u)^\top y_t+\alpha(u)y_t-\operatorname{ad}(u)y_t
$$

$$
\begin{aligned}
&- \tfrac{1}{2}[\operatorname{ad}(y)^\top + \operatorname{ad}(y), \operatorname{ad}(u)^\top]u \\
&+ \tfrac{1}{2}[\alpha(u), \operatorname{ad}(y)]u + \tfrac{1}{2}\operatorname{ad}(u)^\top \operatorname{ad}(y)u\,.
\end{aligned}
$$

Finally we need the following computation using (3.4.1):

$$
\begin{aligned}
\tfrac{1}{2}[\alpha(u), \operatorname{ad}(y)]u &= \tfrac{1}{2}\alpha(u)[y,u] - \tfrac{1}{2}\operatorname{ad}(y)\alpha(u)u \\
&= \tfrac{1}{2}\operatorname{ad}([y,u])^\top u - \tfrac{1}{2}\operatorname{ad}(y)\operatorname{ad}(u)^\top u \\
&= -\tfrac{1}{2}[\operatorname{ad}(y)^\top, \operatorname{ad}(u)^\top]u - \tfrac{1}{2}\operatorname{ad}(y)\operatorname{ad}(u)^\top u\,.
\end{aligned}
$$

Inserting we get the desired result:

$$
\begin{aligned}
\nabla_{\partial_t}\nabla_{\partial_t} y + \mathcal{R}(y,u)u = y_{tt} &+ \operatorname{ad}(u)^\top y_t + \alpha(u)y_t - \operatorname{ad}(u)y_t \\
&- [\operatorname{ad}(y)^\top + \operatorname{ad}(y), \operatorname{ad}(u)^\top]u.
\end{aligned}
$$

3.7. The weak symplectic structure on the space of Jacobi fields

Let us assume now that the geodesic equation in $\mathfrak{g}$

$$
u_t = -\operatorname{ad}(u)^\top u
$$

admits a unique solution for some time interval, depending smoothly on the choice of the initial value $u(0)$. Furthermore we assume that G is a regular Lie group (B.9) so that each smooth curve u in $\mathfrak{g}$ is the right logarithmic derivative of a smooth curve g in G which depends smoothly on u, so that $u = (g^*\kappa)(\partial_t)$. Furthermore we have to assume that the Jacobi equation along u admits a unique solution for some time, depending smoothly on the initial values $y(0)$ and $y_t(0)$. These are non-trivial assumptions: in (A.4) there are examples of ordinary linear differential equations 'with constant coefficients' which violate existence or uniqueness. These assumptions have to be checked in the special situations. Then the space $\mathcal{J}_u$ of all Jacobi fields along the geodesic g described by u is isomorphic to the space $\mathfrak{g} \times \mathfrak{g}$ of all initial data.

There is the well known symplectic structure on the space $\mathcal{J}_u$ of all Jacobi fields along a fixed geodesic with velocity field u, see e.g. [28], II, p.70. It is given by the following expression which is constant in time t:

$$
\begin{aligned}
\omega(y,z) :&= \langle y, \nabla_{\partial_t} z\rangle - \langle \nabla_{\partial_t} y, z\rangle \\
&= \langle y, z_t + \tfrac{1}{2}\operatorname{ad}(u)^\top z + \tfrac{1}{2}\alpha(u)z - \tfrac{1}{2}\operatorname{ad}(u)z\rangle \\
&\qquad - \langle y_t + \tfrac{1}{2}\operatorname{ad}(u)^\top y + \tfrac{1}{2}\alpha(u)y - \tfrac{1}{2}\operatorname{ad}(u)y, z\rangle \\
&= \langle y, z_t\rangle - \langle y_t, z\rangle + \langle [u,y], z\rangle - \langle y, [u,z]\rangle - \langle [y,z], u\rangle \\
&= \langle y, z_t - \operatorname{ad}(u)z + \tfrac{1}{2}\alpha(u)z\rangle - \langle y_t - \operatorname{ad}(u)y + \tfrac{1}{2}\alpha(u)y, z\rangle.
\end{aligned}
$$

It is worth while to check directly from the Jacobi field equation (3.5.1) that $\omega(y,z)$ is indeed constant in t. Clearly ω is a weak symplectic structure on the relevant vector space $\mathcal{J}_u \cong \mathfrak{g}\times\mathfrak{g}$, i.e., ω gives an injective (but in general not surjective) linear mapping $\mathcal{J}_u \to \mathcal{J}_u^*$. This is seen most easily by writing

$$\omega(y,z) = \langle y, z_t - \Gamma_g(u,z)\rangle|_{t=0} - \langle y_t - \Gamma_g(u,y), z\rangle|_{t=0}$$

which is induced from the standard symplectic structure on $\mathfrak{g}\times\mathfrak{g}^*$ by applying first the automorphism $(a,b)\mapsto(a,b-\Gamma_g(u,a))$ to $\mathfrak{g}\times\mathfrak{g}$ and then by injecting the second factor $\mathfrak{g}$ into its dual $\mathfrak{g}^*$.

For regular (infinite dimensional) Lie groups variations of geodesics exist, but there is no general theorem stating that they are uniquely determined by $y(0)$ and $y_t(0)$. For concrete regular Lie groups, this needs to be shown directly.

4. The Hamiltonian approach

4.1. The symplectic form on T^*G and $G\times\mathfrak{g}^*$

For an (infinite dimensional regular) Lie group G with Lie algebra $\mathfrak{g}$, elements in the cotangent bundle $\pi : (T^*G, \omega_G) \to G$ are said to be in *material* or *Lagrangian representation.* The cotangent bundle T^*G has two trivializations, the left one

$$\begin{aligned}&(\pi_G, \kappa^l) : T^*G \to G\times\mathfrak{g}^*,\\ &T_g^*G \ni \alpha_g \mapsto (g, T_e(\mu_g)^*\alpha_g = T_g^*(\mu_{g^{-1}})\alpha_g),\end{aligned}$$

also called the *body coordinate chart*, and the right one,

$$\begin{aligned}&(\pi_G, \kappa^r) : T^*G \to G\times\mathfrak{g}^*,\\ &T^*G \ni \alpha_g \mapsto (g, T_e(\mu^g)^*\alpha_g = T_g^*(\mu^{g^{-1}})\alpha_g),\\ &T_g(\mu^{g^{-1}})^*\alpha \leftarrow (g,\alpha)\in G\times\mathfrak{g}^*\end{aligned} \tag{1}$$

also called the *space* or *Eulerian coordinate chart.* We will use only this from now on. The canonical 1-form in the Eulerian chart is given by (where $\langle\ \ ,\ \ \rangle : \mathfrak{g}^*\times\mathfrak{g}\to\mathbb{R}$ is the duality pairing):

$$\begin{aligned}\theta_{G\times\mathfrak{g}^*}(\xi_g,\alpha,\beta) &:= (((\pi,\kappa^r)^{-1})^*\theta_G)_{(g,\alpha)}(\xi_g,\alpha,\beta)\\ &= \theta_G(T_{(g,\alpha)}(\pi,\kappa^r)^{-1}(\xi_g,\alpha,\beta))\\ &= \Big\langle \pi_{T^*G}(T_{(g,\alpha)}(\pi,\kappa^r)^{-1}(\xi_g,\alpha,\beta)), T(\pi)(T_{(g,\alpha)}(\pi,\kappa^r)^{-1}(\xi_g,\alpha,\beta))\Big\rangle\end{aligned}$$

$$= \Big\langle (\pi,\kappa^r)^{-1}(\pi_G,\pi_{\mathfrak g^*})(\xi_g,\alpha,\beta), T(\pi\circ(\pi,\kappa^r)^{-1})(\xi_g,\alpha,\beta))\Big\rangle$$
$$= \Big\langle (\pi,\kappa^r)^{-1}(g,\alpha), T(\mathrm{pr}_1)(\xi_g,\alpha,\beta))\Big\rangle = \Big\langle T_g(\mu^{g^{-1}})^*\alpha,\xi_g\Big\rangle$$
$$= \langle \alpha, T_g(\mu^{g^{-1}})\xi_g\rangle = \langle\alpha,\kappa^r(\xi_g)\rangle \tag{2}$$

Now it is easy to to take the exterior derivative: For $X_i\in G$, thus $R_{X_i}\in\mathfrak{X}(G)$ right invariant vector fields, and $\mathfrak g^*\ni\beta_i\in\mathfrak X(\mathfrak g^*)$ constant vector fields, we have

$$\begin{aligned}
&\theta_{G\times\mathfrak g^*}(R_{X_i}(g),(\alpha,\beta_i)) = \langle\alpha,X_i\rangle\\
&\theta_{G\times\mathfrak g^*}(R_{X_i},\beta_i) = \langle \mathrm{Id}_{\mathfrak g^*},X_i\rangle = \langle\quad,X_i\rangle\\
&\omega_{G\times\mathfrak g^*}((R_{X_1},\beta_1),(R_{X_2},\beta_2)) = -d\theta_{G\times\mathfrak g^*}((R_{X_1},\beta_1),(R_{X_2},\beta_2))\\
&\qquad = -(R_{X_1},\beta_1)(\theta_{G\times\mathfrak g^*}(R_{X_2},\beta_2)) + (R_{X_2},\beta_2)(\theta_{G\times\mathfrak g^*}(R_{X_1},\beta_1))\\
&\qquad\quad + (\theta_{G\times\mathfrak g^*}([(R_{X_1},\beta_1),R_{X_2},\beta_2)])\\
&\qquad = -(R_{X_1},\beta_1)(\langle\quad,X_2\rangle) + (R_{X_2},\beta_2)(\langle\quad,X_1\rangle)\\
&\qquad\quad + (\theta_{G\times\mathfrak g^*}(-R_{[X_1,X_2]},0_{\mathfrak g^*})\\
&\qquad = -\langle\beta_1,X_2\rangle + \langle\beta_2,X_1\rangle - \langle\quad,[X_1,X_2]\rangle\\
&(\omega_{G\times\mathfrak g^*})_{(g,\alpha)}((T(\mu^g).X_1,\beta_1),(T(\mu^g)X_2,\beta_2))\\
&\qquad = \langle\beta_2,X_1\rangle - \langle\beta_1,X_2\rangle - \langle\alpha,[X_1,X_2]\rangle
\end{aligned}\tag{3}$$

4.2. The symplectic form on TG and $G\times\mathfrak g$ and the momentum mapping

We consider an (infinite dimensional regular) Lie group G with Lie algebra $\mathfrak g$ and a bounded weak inner product $\gamma:\mathfrak g\times\mathfrak g\to\mathbb R$ with the property the transpose of the adjoint action of G on $\mathfrak g$,

$$\gamma(\mathrm{Ad}(g)^\top X,Y) = \gamma(X,\mathrm{Ad}(g)X),$$

exists. It is then unique and a right action of G on $\mathfrak g$. By differentiating it follows that then also the transpose of the adjoint operation of $\mathfrak g$ exists:

$$\gamma(\mathrm{ad}(X)^\top Y,Z) = \partial_t|_0\gamma(\mathrm{Ad}(\exp(tX))^\top Y,Z) = \gamma(Y,\mathrm{ad}(X)Z)$$

exists.

We exted γ to a right invariant Riemannian metric, again called γ on G and consider $\gamma: TG\to T^*G$. Then we pull back the canonical symplectic structure ω_G to $G\times\mathfrak g$ in the right or Eulerian trivialization:

$$\begin{aligned}
&\gamma: G\times\mathfrak g\to G\times\mathfrak g^*, (g,X)\mapsto(g,\gamma(X))\\
&(\gamma^*\omega)_{(g,X)}((T(\mu^g).X_1,X,Y_1),(T(\mu^g)X_2,X,Y_2))
\end{aligned}$$

$$
\begin{aligned}
&= \omega_{(g,\gamma(X))}((T(\mu^g).X_1, \gamma(X), \gamma(Y_1)), (T(\mu^g)X_2, \gamma(X), \gamma(Y_2)))\\
&= \langle \gamma(Y_2), X_1\rangle - \langle \gamma(Y_1), X_2\rangle - \langle \gamma(X), [X_1, X_2]\rangle\\
&= \gamma(Y_2, X_1) - \gamma(Y_1, X_2) - \gamma(X, [X_1, X_2]) \qquad (1)
\end{aligned}
$$

Since γ is a weak inner product, $\gamma^*\omega$ is again a weak symplectic structure on $TG \cong G \times \mathfrak{g}$. We compute the Hamiltonian vector field mapping (symplectic gradient) for functions $f \in C^\infty_{\gamma^*\omega}(G \times \mathfrak{g})$ admitting such gradients:

$$
\begin{aligned}
&(\gamma^*\omega)_{(g,X)}\Big(\operatorname{grad}^{\gamma^*\omega}(f)(g,X), (T(\mu^g)X_2, X, Y_2)\Big) = df(T(\mu^g)X_2; X, Y_2)\\
&\quad = d_1 f(g,X)(T(\mu^g)X_2) + d_2 f(g,X)(Y_2)\\
&\quad = \gamma(\kappa^r(\operatorname{grad}_1^\gamma(f)(g,X)), X_2) + \gamma(\operatorname{grad}_2^\gamma(f)(g,X), Y_2)\\
&\quad = \gamma(X_1, Y_2) + \gamma(-Y_1 - \operatorname{ad}(X_1)^\top X, X_2) \qquad \text{by } ((1)).
\end{aligned}
$$

Thus the Hamiltonian vector field of $f \in C^\infty_{\gamma^*\omega}(G \times \mathfrak{g}) = C^\infty_\gamma(G \times \mathfrak{g})$ is

$$
\begin{aligned}
&\operatorname{grad}^{\gamma^*\omega}(f)(g,X) = \qquad (2)\\
&\big(T(\mu^g)\operatorname{grad}_2^\gamma(f)(g,X), X, -\operatorname{ad}(\operatorname{grad}_2^\gamma(f)(g,X))^\top X - \kappa^r(\operatorname{grad}_1^\gamma(f)(g,X))\big)
\end{aligned}
$$

In particular, the Hamiltonian vector field of the function $(g,X) \mapsto \gamma(X,X) = \|X\|^2_\gamma$ on TG is given by:

$$
\operatorname{grad}^{\gamma^*\omega}(\tfrac{1}{2}\|\quad\|^2_\gamma)(g,X) = (T(\mu^g)X; X, -\operatorname{ad}(X)^\top X) \qquad (3)
$$

We can now compute again the flow equation of the Hamiltonian vector field $\operatorname{grad}^{\gamma^*\omega}(\frac{1}{2}\|\quad\|^2_\gamma)$: For $g_t(t) \in TG$ we have

$$
(\pi_G, \kappa^r)(g_t(t)) = (g(t), u(t)) = (g(t), T(\mu^{g(t)^{-1}})g_t(t))
$$

and

$$
\partial_t(g,u) = \operatorname{grad}^{\gamma^*\omega}(\tfrac{1}{2}\|\quad\|^2_\gamma)(g,u) = (T(\mu^g)u, u, -\operatorname{ad}(u)^\top u). \qquad (4)
$$

which reproduces the geodesic equation from (3.2).

4.3. The momentum mapping

Under the assumptions of (4.2), consider the right action of G on G and its prolongation to a right action of G on TG in the Eulerian chart. The corresponding fundamental vector fields are then given by:

$$
\begin{aligned}
&T(\mu^g) : TG \to TG,\\
&(\pi, \kappa^r)T(\mu^g)T(\mu^h)X = (\pi, \kappa^r)T(\mu^{hg})X = (h.g, X), \quad (h, X) \mapsto (hg, X)
\end{aligned}
$$

$$\zeta_X^{G\times\mathfrak{g}}(h,Y) = \partial_t|_0(h.\exp(tX),Y) = (T(\mu_h)X, 0_Y) \in TG\times T\mathfrak{g} \quad (1)$$

Consider now the diagram from (2.1) in the case of the weak symplectic manifold $(M = G\times\mathfrak{g}, \gamma^*\omega)$:

$$\begin{tikzcd} H^0 \arrow[r] & C^\infty_{\gamma^*\omega}(G\times\mathfrak{g},\mathbb{R}) \arrow[rr, "\mathrm{grad}^{\gamma^*\omega}"] & & \mathfrak{X}(G\times\mathfrak{g},\gamma^*\omega) \arrow[r] & H^1_{\gamma^*\omega} \\ & & \mathfrak{g} \arrow[ul, "j"] \arrow[ur, "\zeta"'] & & \end{tikzcd}$$

From the formulas derived above we see that for $j(X)(h,Y) := \gamma(\mathrm{Ad}(h)X, Y)$ we have:

$$\begin{aligned}
&\gamma(\mathrm{grad}_2^\gamma(j(X))(h,Y),Z) = d_2(j(X))(h,Y)(Z) = \gamma(\mathrm{Ad}(h)X,Z)\\
&\ \mathrm{grad}_2^\gamma(j(X))(h,Y) = \mathrm{Ad}(h)X\\
&\gamma(\mathrm{grad}_1^\gamma(j(X))(h,Y),T(\mu^h)Z) = d(j(X))(T(\mu^h)Z,Y,0)\\
&\quad = \gamma(d\,\mathrm{Ad}(T(\mu^h)Z)(X),Y) = \gamma(((\mathrm{ad}\circ\kappa^r)\,\mathrm{Ad})(T(\mu^h)Z)(X),Y)\\
&\quad = \gamma(\mathrm{ad}(Z)\,\mathrm{Ad}(h)X,Y) = -\gamma([\mathrm{Ad}(h)X,Z],Y) = -\gamma(Z,\mathrm{ad}(\mathrm{Ad}(h)X)^\top Y)\\
&\kappa^r(\mathrm{grad}_1^\gamma(j(X))(h,Y)) = -\,\mathrm{ad}(\mathrm{Ad}(h)X)^\top Y
\end{aligned}$$

Thus the momentum mapping is

$$\begin{aligned}
J : G\times\mathfrak{g}\to\mathfrak{g}^*,&\quad J\in C^\infty_{\gamma^*\omega}(G\times\mathfrak{g},\mathfrak{g}^*) =\\
&= \{f\in C^\infty(G\times\mathfrak{g},\mathfrak{g}^*) : \langle f(\quad),X\rangle \in C^\infty_{\gamma^*\omega}(G\times\mathfrak{g})\ \forall X\in\mathfrak{g}\}\\
\langle J(h,Y),X\rangle &= j(X)(h,Y) = \gamma(\mathrm{Ad}(h)X,Y) = \gamma(\mathrm{Ad}(h)^\top Y,X)\\
&= \langle\gamma(\mathrm{Ad}(h)^\top Y),X\rangle,\\
J(h,Y) &= \gamma(\mathrm{Ad}(h)^\top Y)\in\mathfrak{g}*\\
\bar J := \gamma^{-1}\circ J &: G\times\mathfrak{g}\to\mathfrak{g},\\
\bar J(h,Y) &= \mathrm{Ad}(h)^\top Y\in\mathfrak{g}. \qquad (2)
\end{aligned}$$

(3) Note that the momentum mapping $J : G\times\mathfrak{g}\to\mathfrak{g}^*$ is equivariant for the right G-action and the coadjoint action, and that $\bar J : G\times\mathfrak{g}\to\mathfrak{g}$ is equivariant for the right action $\mathrm{Ad}(\quad)^\top$ on $\mathfrak{g}$:

$$\begin{aligned}
\langle J(hg,Y),X\rangle &= \langle\gamma(\mathrm{Ad}(hg)^\top Y),X\rangle = \gamma(\mathrm{Ad}(g)^\top\,\mathrm{Ad}(h)^\top Y,X)\\
&= \gamma(\mathrm{Ad}(h)^\top Y,\mathrm{Ad}(g)X) = \langle\gamma(\mathrm{Ad}(h)^\top Y),\mathrm{Ad}(g)X\rangle\\
&= \langle\mathrm{Ad}(g)^*\gamma(\mathrm{Ad}(h)^\top Y),X\rangle = \langle\mathrm{Ad}(g)^*J(h,Y),X\rangle\\
\bar J(hg,Y) &= \mathrm{Ad}(hg)^\top Y = \mathrm{Ad}(g)^\top\bar J(h,Y).
\end{aligned}$$

(4) For $x\in G\times\mathfrak{g}$, the transposed mapping of $d\bar J(x) : T_x(G\times\mathfrak{g})\to\mathfrak{g}$ is

$$d\bar J(x)^\top : \mathfrak g \to T^*_x(G\times\mathfrak g), \qquad d\bar J(x)^\top = (\gamma^*\omega)_x\circ\zeta,$$

since for $\xi\in T_x(G\times\mathfrak g)$ and $X\in\mathfrak g$ we have

$$\gamma(d\bar J(\xi),X) = d\gamma(\bar J,X)(\xi) = dj(X)(\xi) = \langle(\gamma^*\omega)(\zeta_X),\xi\rangle.$$

(5) For $x\in G\times\mathfrak g$, the closure $\overline{d\bar J(T_x(G\times\mathfrak g))}$ of the image of $d\bar J(x): T_x(G\times\mathfrak g)\to\mathfrak g$ is the γ-orthogonal space $\mathfrak g_x^{\perp,\gamma}$ of the isotropy Lie algeba $\mathfrak g_x := \{X\in\mathfrak g : \zeta_X(x)=0\}$ in $\mathfrak g$, since the annihilator of the image is the kernel of the transposed mapping,

$$\operatorname{im}(dJ(x))^\circ = \ker(dJ(x)^\top) = \ker((\gamma^*\omega)_x\circ\zeta) = \ker(\operatorname{ev}_x\circ\zeta) = \mathfrak g_x.$$

Attention: the orthogonal space with respect to a weak inner product need not be a complement.

(6) For $(h,Y)\in G\times\mathfrak g$, the G-orbit $(h,Y).G = G\times\{Y\}$ is a submanifold of $G\times\mathfrak g$. *The kernel of $d\bar J(h,Y)$ is the symplectic orthogonal space*

$$(T_{(h,Y)}(G\times\{Y\}))^{\perp,\gamma^*\omega}\subset T(\mu^h)\mathfrak g\times\mathfrak g$$

since for the annihilator of the kernel we have

$$\begin{aligned}\ker(d\bar J(h,Y))^\circ &= \overline{\operatorname{im}(d\bar J(h,Y)^\top)} = \overline{\operatorname{im}((\gamma^*\omega_{(h,Y)}\circ\zeta)}, \qquad \text{by ((4))},\\ &= \overline{\{(\gamma^*\omega)_{(h,Y)}(\zeta_X(x)) : X\in\mathfrak g\}} = \overline{(\gamma^*\omega)_{(h,Y)}(T_{(h,Y)}(G\times\{Y\}))},\\ &= \big((T_{(h,Y)}(G\times\{Y\}))^{\perp,\gamma^*\omega}\big)^\circ.\end{aligned}$$

The last equality holds by the bipolar theorem for the usual duality pairing.

(7) Thus, for $(h,Y)\in G\times\mathfrak g$,

$$\begin{aligned}&T(\mu^h)X_1,Y_1)\in\ker(d\bar J(h,Y))\\ &\iff (\gamma^*\omega)_{(h,Y)}((T(\mu^h)X_1,Y_1),(T(\mu^h)Z,0)) = 0 \text{ for all } Z\in\mathfrak g\\ &\iff 0 = 0-\gamma(Y_1,Z)-\gamma(Y,[X_1,Z]) = -\gamma(Y_1+\operatorname{ad}(X_1)^\top Y,Z)\ \forall\, Z\in\mathfrak g\\ &\iff Y_1 = -\operatorname{ad}(X_1)^\top Y.\end{aligned}$$

(8) (Emmy Noether's theorem) Let $h\in C^\infty_\omega(G\times\mathfrak g)$ be a Hamiltonian function which is invariant under the right G-action. Then $d\bar J(\operatorname{grad}^{\gamma^*\omega}(h)) = 0\in\mathfrak g$ and also $dJ(\operatorname{grad}^{\gamma^*\omega}(h)) = 0\in\gamma(\mathfrak g)\subseteq\mathfrak g^*$. Thus the momentum mappings $\bar J: G\times\mathfrak g\to\mathfrak g$ and $J: G\times\mathfrak g\to\gamma(\mathfrak g)\subset\mathfrak g^*$ are constant on each trajectory (if it exists) of the Hamiltonian vector field $\operatorname{grad}^{\gamma^*\omega}(h)$. Namely, consider the function $\gamma(\bar J,X) = \langle J,X\rangle = j(X)$.

$$\begin{aligned}\gamma(d\bar J(\operatorname{grad}^{\gamma^*\omega}(h)),X) &= \operatorname{grad}^{\gamma^*\omega}(h)(\gamma(\bar J,X)) =\\ &= \{h,\gamma(\bar J,X)\} = -\{j(X),h\} = -\zeta_X(h) = 0.\\ \langle dJ(\operatorname{grad}^{\gamma^*\omega}(h)),X\rangle &= \operatorname{grad}^{\gamma^*\omega}(h)(\langle J,X\rangle) =\end{aligned}$$

$$= \{h, j(X)\} = -\{j(X), h\} = -\zeta_X(h) = 0.$$

4.4. The geodesic equation via conserved momentum

We consider a smooth curve $t \mapsto g(t)$ in G and $(\pi_G, \kappa^r)g_t(t) = (g(t), u(t)) = (g(t), T(\mu^{g(t)^{-1}})g_t(t))$ as in (4.2.4). Applying $\bar{J} : G \times \mathfrak{g} \to \mathfrak{g}$ to it we get $\bar{J}(g, u) = \operatorname{Ad}(g)^\top u$. We claim that *the curves $t \mapsto g(t)$ in G for which $\bar{J}(g(t), u(t))$ is constant in t are exactly the geodesics in (G, γ).* Namely, by (3.1) we have

$$\begin{aligned}
0 &= \partial_t \operatorname{Ad}(g(t))^\top u(t) = \big((\operatorname{ad} \circ \kappa^r)(\partial_t g(t)).\operatorname{Ad}(g(t))\big)^\top u(t) + \operatorname{Ad}(g(t))^\top \partial_t u(t) \\
&= \operatorname{Ad}(g(t))^\top \big(\operatorname{ad}(u(t))^\top u(t) + u_t(t)\big) \\
&\iff \qquad u_t = -\operatorname{ad}(u)^\top u.
\end{aligned}$$

4.5. Symplectic reduction to transposed adjoint orbits

Under the assumptions of (4.2) we have the following:

(1) *For $X \in \bar{J}(G \times \mathfrak{g})$ the inverse image $\bar{J}^{-1}(X) \subset G \times \mathfrak{g}$ is a manifold.* Namely, it is the graph of a smooth mapping:

$$\begin{aligned}
\bar{J}^{-1}(X) &= \{(h, Y) \in G \times \mathfrak{g} : \operatorname{Ad}(h)^\top Y = X\} \\
&= \{(h, \operatorname{Ad}(h^{-1})^\top X) : h \in G\} \xleftarrow{\cong} G. \square
\end{aligned}$$

(2) *At any point of $\bar{J}^{-1}(X)$, the kernel of the pullback of the symplectic form $\gamma^*\omega$ on $G \times \mathfrak{g}$ from (4.2.1) equals the tangent space to the orbit of the isotropy group $G_X := \{g \in G : \operatorname{Ad}(g)^\top X = X\}$ through that point.*

For $(h, Y = \operatorname{Ad}(h^{-1})^\top X) \in \bar{J}^{-1}(X)$ the G_X-orbit is $h.G_X \times \{Y\}$ and its tangent space at (h, Y) is $T(\mu_h)\mathfrak{g}_X \times 0$ where $\mathfrak{g}_X = \{Z \in \mathfrak{g} : \operatorname{ad}(Z)^\top X = 0\}$. The tangent space at (h, Y) of $\bar{J}^{-1}(X)$ is

$$\begin{aligned}
T_{(h, \operatorname{Ad}(h^{-1})^\top X)}\bar{J}^{-1}(X) &= \{\partial_t|_0(\exp(tZ).h, \operatorname{Ad}((\exp(tZ).h)^{-1})^\top X) : Z \in \mathfrak{g}\} \\
&= \{(T(\mu^h)Z, -\operatorname{ad}(Z)^\top \operatorname{Ad}(h^{-1})^\top X) : Z \in \mathfrak{g}\} \subset T_h G \times \mathfrak{g}.
\end{aligned}$$

For $Z_1, Z_2 \in \mathfrak{g}$ consider the tangent vectors $(T(\mu^h)\operatorname{Ad}(h)Z_1, Y, -\operatorname{ad}(Z_1)X)$ and $(T(\mu^h)Z, Y, -\operatorname{ad}(Z)^\top \operatorname{Ad}(h^{-1})^\top X)$ in $T_{(h,Y)}\bar{J}^{-1}(X)$. From (4.2.1), we get

$$\begin{aligned}
&(\gamma^*\omega)_{(h,Y)}\big((T(\mu^h)\operatorname{Ad}(h)Z_1, -\operatorname{ad}(Z_1)^\top X), (T(\mu^h)Z_2, -\operatorname{ad}(Z_2)^\top \operatorname{Ad}(h^{-1})^\top X)\big) \\
&= \gamma(-\operatorname{ad}(Z_2)^\top \operatorname{Ad}(h^{-1})^\top X, \operatorname{Ad}(h)Z_1) - \gamma(-\operatorname{ad}(Z_1)^\top X, Z_2)
\end{aligned}$$

$$
\begin{aligned}
&\quad - \gamma(Y, [\mathrm{Ad}(h)Z_1, Z_2]) \\
&= -\gamma(\mathrm{Ad}(h^{-1})^\top X, \mathrm{ad}(Z_2)\,\mathrm{Ad}(h)Z_1) + \gamma(\mathrm{ad}(Z_1)^\top X, Z_2) - \\
&\quad - \gamma(\mathrm{Ad}(h^{-1})^\top X, [\mathrm{Ad}(h)Z_1, Z_2]) \\
&= \gamma(\mathrm{ad}(Z_1)^\top X, Z_2) = 0 \quad \forall Z_2 \in \mathfrak{g} \iff Z_1 \in \mathfrak{g}_X. \qquad \square
\end{aligned}
$$

(3) *The reduced symplectic manifold $\bar{J}^{-1}(X)/G_X$ with symplectic form induced by $\gamma^*\omega|_{\bar{J}^{-1}(X)}$ is symplectomorphic to the adjoint orbit $\mathrm{Ad}(G)^\top X \subset \mathfrak{g}$ with symplectic form the pullback via $\gamma : \mathfrak{g} \to \mathfrak{g}^*$ of the Kostant Kirillov Souriou form*

$$\omega_\alpha(\mathrm{ad}(Y_1)^*\alpha, \mathrm{ad}(Y_2)^*\alpha) = \langle \alpha, [Y_1, Y_2] \rangle$$

which is given by

$$
\begin{aligned}
\omega_Z(\mathrm{ad}(Y_1)^\top Z, \mathrm{ad}(Y_2)^\top Z) &= \omega_{\gamma(Z)}(\gamma\,\mathrm{ad}(Y_1)^\top Z, \gamma\,\mathrm{ad}(Y_2)^\top Z) \\
&= \omega_{\gamma(Z)}(\mathrm{ad}(Y_1)^*\gamma Z, \mathrm{ad}(Y_2)^*\gamma Z) = \langle \gamma(Z), [Y_1, Y_2] \rangle = \gamma(Z, [Y_1, Y_2]),
\end{aligned}
$$

since for $Y, Z, U \in \mathfrak{g}$ we get

$$
\begin{aligned}
\langle \gamma\,\mathrm{ad}(Y)^\top Z, U \rangle = \gamma(\mathrm{ad}(Y)^\top Z, U) = \gamma(Z, \mathrm{ad}(Y)U) = \\
= \langle \gamma(Z), \mathrm{ad}(Y)U \rangle = \langle \mathrm{ad}(Y)^*\gamma(Z), U \rangle.
\end{aligned}
$$

The quotient space is $\bar{J}^{-1}(X)/G_X = \{(h.G_X, \mathrm{Ad}(h^{-1})^\top X) : h \in G\} \cong \mathrm{Ad}(G)^\top X \cong G/G_X$. The 2-form $\gamma^*\omega|_{\bar{J}^{-1}(X)}$ induces a symplectic form on the quotient by (2) and it remains to check that it agrees with the pullback of the Kirillov Kostant Souriou symplectic form. But this is obvious from the last computation in (2) (for the special case $h = e$ if the reader insists). $\square$

(4) Reconsider the geodesic equation on the reduced space $\bar{J}^{-1}(X)/G_X \cong \mathrm{Ad}(G)^\top X$. The energy function is $E(\mathrm{Ad}(g)^\top X) = \frac{1}{2}\|\mathrm{Ad}(g)^\top X\|^2_\gamma$. For $Z = \mathrm{Ad}(g)^\top X \in \mathrm{Ad}(G)^\top X$ the tangent space is given by $T_Z(\mathrm{Ad}(G)^\top X) = \{\mathrm{ad}(Y)^\top Z : Y \in \mathfrak{g}\}$. We look for the Hamiltonian vector field of E in the form $\mathrm{grad}^\omega E(Z) = \mathrm{ad}(H_E(Z))^\top Z$, for a vector field H_E. The differential of the energy function is $dE(Z)(\mathrm{ad}(Y)^\top Z) = \gamma(Z, \mathrm{ad}(Y)^\top Z) = \gamma([Y, Z], Z)$ which equals $\omega_Z(\mathrm{grad}^\omega E(Z), \mathrm{ad}(Y)^\top Z) = \omega_Z(\mathrm{ad}(H_E(Z))^\top Z, \mathrm{ad}(Y)^\top Z) = \gamma(Z, [H_E(Z), Y])$ from which we conclude that $H_E(Z) = -Z$ will do (which is defined up to annihilator of Z). Thus $\mathrm{grad}^\omega E(Z) = -\mathrm{ad}(Z)^\top Z$ which leads us back to the geodesic equation $u_t = -\mathrm{ad}(u)^\top u$ again.

5. Vanishing H^0-geodesic distance on groups of diffeomorphisms

This section is based on [40].

5.1 The H^0-metric on groups of diffeomorphisms

Let (N, g) be a smooth connected Riemannian manifold, and let $\mathrm{Diff}_c(N)$ be the group of all diffeomorphisms with compact support on N, and let $\mathrm{Diff}_0(N)$ be the subgroup of those which are diffeotopic in $\mathrm{Diff}_c(N)$ to the identity; this is the connected component of the identity in $\mathrm{Diff}_c(N)$, which is a regular Lie group in the sense of [42], section 38. This is proved in [31], section 42. The Lie algebra is $\mathfrak{X}_c(N)$, the space of all smooth vector fields with compact support on N, with the negative of the usual bracket of vector fields as Lie bracket. Moreover, $\mathrm{Diff}_0(N)$ is a simple group (has no nontrivial normal subgroups), see [18], [50], [35]. The *right invariant* H^0-metric on $\mathrm{Diff}_0(N)$ is then given as follows, where $h, k : N \to TN$ are vector fields with compact support along φ and where $X = h \circ \varphi^{-1}, Y = k \circ \psi^{-1} \in \mathfrak{X}_c(N)$:

$$\begin{aligned} \gamma^0_\varphi(h,k) &= \int_N g(h,k)\,\mathrm{vol}(\varphi^* g) = \int_N g(X\circ\varphi, Y\circ\varphi)\varphi^*\,\mathrm{vol}(g) \\ &= \int_N g(X,Y)\,\mathrm{vol}(g). \end{aligned} \tag{1}$$

5.2. Theorem.
Geodesic distance on $\mathrm{Diff}_0(N)$ *with respect to the* H^0*-metric vanishes.*

Proof. Let $[0,1] \ni t \mapsto \varphi(t, \quad)$ be a smooth curve in $\mathrm{Diff}_0(N)$ between φ_0 and φ_1. Consider the curve $u = \varphi_t \circ \varphi^{-1}$ in $\mathfrak{X}_c(N)$, the right logarithmic derivative. Then for the length and the energy we have:

$$L_{\gamma^0}(\varphi) = \int_0^1 \sqrt{\int_N \|u\|_g^2\,\mathrm{vol(g)}}\; dt \tag{1}$$

$$E_{\gamma^0}(\varphi) = \int_0^1 \int_N \|u\|_g^2\,\mathrm{vol(g)}\; dt \tag{2}$$

$$L_{\gamma^0}(\varphi)^2 \le E_{\gamma^0}(\varphi) \tag{3}$$

(4) Let us denote by $\mathrm{Diff}_0(N)^{E=0}$ the set of all diffeomorphisms $\varphi \in \mathrm{Diff}_0(N)$ with the following property: For each $\varepsilon > 0$ there exists a smooth curve from the identity to φ in $\mathrm{Diff}_0(N)$ with energy $\le \varepsilon$.
(5) *We claim that* $\mathrm{Diff}_0(N)^{E=0}$ *coincides with the set of all diffeomorphisms which can be reached from the identity by a smooth curve of arbitraily short* γ^0*-length.* This follows by (3).
(6) *We claim that* $\mathrm{Diff}_0(N)^{E=0}$ *is a normal subgroup of* $\mathrm{Diff}_0(N)$. Let $\varphi_1 \in \mathrm{Diff}_0(N)^{E=0}$ and $\psi \in \mathrm{Diff}_0(N)$. For any smooth curve $t \mapsto \varphi(t, \quad)$ from the identity to φ_1 with energy $E_{\gamma^0}(\varphi) < \varepsilon$ we have

$$E_{\gamma^0}(\psi^{-1}\circ\varphi\circ\psi) = \int_0^1 \int_N \|T\psi^{-1}\circ\varphi_t\circ\psi\|_g^2\,\mathrm{vol}((\psi^{-1}\circ\varphi\circ\psi)^* g)$$

$$\begin{aligned}
&\le \sup_{x\in N}\|T_x\psi^{-1}\|^2\cdot\int_0^1\int_N \|\varphi_t\circ\psi\|_g^2(\varphi\circ\psi)^*\operatorname{vol}((\psi^{-1})^*g)\\
&\le \sup_{x\in N}\|T_x\psi^{-1}\|^2\cdot\sup_{x\in N}\frac{\operatorname{vol}((\psi^{-1})^*g)}{\operatorname{vol}(g)}\cdot\int_0^1\int_N \|\varphi_t\circ\psi\|_g^2\,(\varphi\circ\psi)^*\operatorname{vol}(g)\\
&\le \sup_{x\in N}\|T_x\psi^{-1}\|^2\cdot\sup_{x\in N}\frac{\operatorname{vol}((\psi^{-1})^*g)}{\operatorname{vol}(g)}\cdot E_{\gamma^0}(\varphi).
\end{aligned}$$

Since ψ is a diffeomorphism with compact support, the two suprema are bounded. Thus $\psi^{-1}\circ\varphi_1\circ\psi\in\operatorname{Diff}_0(N)^{E=0}$.
(7) *We claim that* $\operatorname{Diff}_0(N)^{E=0}$ *is a non-trivial subgroup.* In view of the simplicity of $\operatorname{Diff}_0(N)$ mentioned in (5.1) this concludes the proof.

It remains to find a non-trivial diffeomorphism in $\operatorname{Diff}_0(N)^{E=0}$. The idea is to use compression waves. The basic case is this: take any non-decreasing smooth function $f:\mathbb{R}\to\mathbb{R}$ such that $f(x)\equiv 0$ if $x\ll 0$ and $f(x)\equiv 1$ if $x\gg 0$. Define

$$\varphi(t,x)=x+f(t-\lambda x)$$

where $\lambda<1/\max(f')$. Note that

$$\varphi_x(t,x)=1-\lambda f'(t-\lambda x)>0,$$

hence each map $\varphi(t,\quad)$ is a diffeomorphism of $\mathbb{R}$ and we have a path in the group of diffeomorphisms of $\mathbb{R}$. These maps are not the identity outside a compact set however. In fact, $\varphi(x)=x+1$ if $x\ll 0$ and $\varphi(x)=x$ if $x\gg 0$. As $t\to-\infty$, the map $\varphi(t,\quad)$ approaches the identity uniformly on compact subsets, while as $t\to+\infty$, the map approaches translation by 1. This path is a moving compression wave which pushes all points forward by a distance 1 as it passes. We calculate its energy between two times t_0 and t_1:

$$\begin{aligned}
E_{t_0}^{t_1}(\varphi)&=\int_{t_0}^{t_1}\int_{\mathbb{R}}\varphi_t(t,\varphi(t,\quad)^{-1}(x))^2dx\,dt=\int_{t_0}^{t_1}\int_{\mathbb{R}}\varphi_t(t,y)^2\varphi_y(t,y)dy\,dt\\
&=\int_{t_0}^{t_1}\int_{\mathbb{R}}f'(z)^2\cdot(1-\lambda f'(z))\frac{dz}{\lambda}\,dt\\
&\le\frac{\max f'^2}{\lambda}\cdot(t_1-t_0)\cdot\int_{\operatorname{supp}(f')}(1-\lambda f'(z))dz
\end{aligned}$$

If we let $\lambda=1-\varepsilon$ and consider the specific f given by the convolution

$$f(z)=\max(0,\min(1,z))\star G_\varepsilon(z),$$

where G_ε is a smoothing kernel supported on $[-\varepsilon,+\varepsilon]$, then the integral is bounded by 3ε, hence

$$E_{t_0}^{t_1}(\varphi)\le(t_1-t_0)\tfrac{3\varepsilon}{1-\varepsilon}.$$

We next need to adapt this path so that it has compact support. To do this we have to start and stop the compression wave, which we do by giving it variable length. Let:

$$f_\varepsilon(z,a) = \max(0, \min(a,z)) \star (G_\varepsilon(z)G_\varepsilon(a)).$$

The starting wave can be defined by:

$$\varphi_\varepsilon(t,x) = x + f_\varepsilon(t - \lambda x, g(x)), \quad \lambda < 1, \quad g \text{ increasing}.$$

Note that the path of an individual particle x hits the wave at $t = \lambda x - \varepsilon$ and leaves it at $t = \lambda x + g(x) + \varepsilon$, having moved forward to $x + g(x)$. Calculate the derivatives:

$$\begin{aligned}
(f_\varepsilon)_z &= I_{0\le z\le a} \star (G_\varepsilon(z)G_\varepsilon(a)) \in [0,1] \\
(f_\varepsilon)_a &= I_{0\le a\le z} \star (G_\varepsilon(z)G_\varepsilon(a)) \in [0,1] \\
(\varphi_\varepsilon)_t &= (f_\varepsilon)_z(t - \lambda x, g(x)) \\
(\varphi_\varepsilon)_x &= 1 - \lambda (f_\varepsilon)_z(t - \lambda x, g(x)) + (f_\varepsilon)_a(t - \lambda x, g(x)) \cdot g'(x) > 0.
\end{aligned}$$

This gives us:

$$\begin{aligned}
E_{t_0}^{t_1}(\varphi) &= \int_{t_0}^{t_1} \int_{\mathbb{R}} (\varphi_\varepsilon)_t^2 (\varphi_\varepsilon)_x dx\, dt \\
&\le \int_{t_0}^{t_1} \int_{\mathbb{R}} (f_\varepsilon)_z^2(t - \lambda x, g(x)) \cdot (1 - \lambda (f_\varepsilon)_z(t - \lambda x, g(x)))dx\, dt \\
&\quad + \int_{t_0}^{t_1} \int_{\mathbb{R}} (f_\varepsilon)_z^2(t - \lambda x, g(x)) \cdot (f_\varepsilon)_a(t - \lambda x, g(x)) g'(x) dx\, dt
\end{aligned}$$

The first integral can be bounded as in the original discussion. The second integral is also small because the support of the z-derivative is $-\varepsilon \le t - \lambda x \le g(x) + \varepsilon$, while the support of the a-derivative is $-\varepsilon \le g(x) \le t - \lambda x + \varepsilon$, so together $|g(x) - (t - \lambda x)| \le \varepsilon$. Now define x_1 and x_2 by $g(x_1) + \lambda x_1 = t + \varepsilon$ and $g(x_0) + \lambda x_0 = t - \varepsilon$. Then the inner integral is bounded by

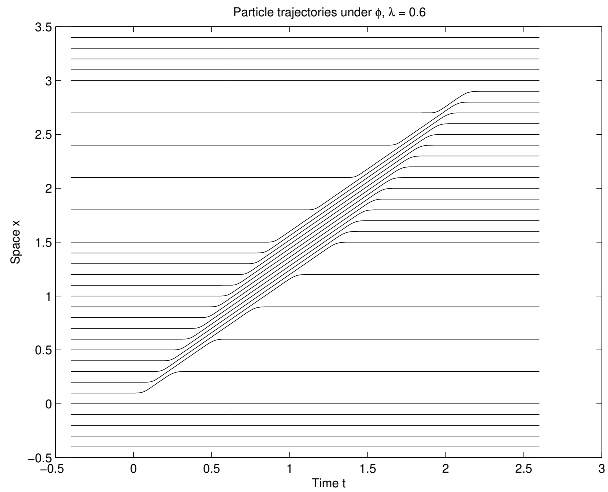
$$\int_{|g(x)+\lambda x - t| \le \varepsilon} g'(x) dx = g(x_1) - g(x_0) \le 2\varepsilon,$$

and the whole second term is bounded by $2\varepsilon(t_1 - t_0)$. Thus the length is $O(\varepsilon)$.

The end of the wave can be handled by playing the beginning backwards. If the distance that a point x moves when the wave passes it is to be $g(x)$, so that the final diffeomorphism is $x \mapsto x + g(x)$, then let $b = \max(g)$ and use the above definition of φ while $g' > 0$. The modification when $g' < 0$ (but $g' > -1$ in order for $x \mapsto x + g(x)$ to have positive derivative) is given by:

$$\varphi_\varepsilon(t,x) = x + f_\varepsilon(t - \lambda x - (1-\lambda)(b - g(x)), g(x)).$$

Consider the figure showing the trajectories $\varphi_\varepsilon(t,x)$ for sample values of x.



It remains to show that $\mathrm{Diff}_0(N)^{E=0}$ is a nontrivial subgroup for an arbitrary Riemannian manifold. We choose a piece of a unit speed geodesic containing no conjugate points in N and Fermi coordinates along this geodesic; so we can assume that we are in an open set in $\mathbb{R}^m$ which is a tube around a piece of the u^1-axis. Now we use a small bump function in the slice orthogonal to the u^1-axis and multiply it with the construction from above for the coordinate u^1. Then it follows that we get a nontrivial diffeomorphism in $\mathrm{Diff}_0(N)^{E=0}$ again. $\square$

Remark

Theorem (5.2) can be proved directly without the help of the simplicity of $\mathrm{Diff}_0(N)$. For $N = \mathbb{R}$ one can use the method of (5.2.7) in the parameter space of a curve, and for general N one can use a Morse function on N to produce a special coordinate for applying the same method.

5.3. Geodesics and sectional curvature for γ^0 on $\mathbf{Diff}(N)$

According to (3.2), (3.4), or (4.4), for a right invariant weak Riemannian metric G on an (possibly infinite dimensional) Lie group the geodesic equation and the curvature are given in terms of the transposed operator (with respect to G, if it exists) of the Lie bracket by the following formulas:

$$\begin{aligned}
u_t &= -\operatorname{ad}(u)^* u, \quad u = \varphi_t \circ \varphi^{-1} \\
G(\operatorname{ad}(X)^* Y, Z) &:= G(Y, \operatorname{ad}(X) Z) \\
4G(R(X,Y)X,Y) &= 3G(\operatorname{ad}(X)Y, \operatorname{ad}(X)Y) - 2G(\operatorname{ad}(Y)^* X, \operatorname{ad}(X)Y) \\
&\quad - 2G(\operatorname{ad}(X)^* Y, \operatorname{ad}(Y)X) + 4G(\operatorname{ad}(X)^* X, \operatorname{ad}(Y)^* Y) \\
&\quad - G(\operatorname{ad}(X)^* Y + \operatorname{ad}(Y)^* X, \operatorname{ad}(X)^* Y + \operatorname{ad}(Y)^* X)
\end{aligned}$$

In our case, for $\operatorname{Diff}_0(N)$, we have $\operatorname{ad}(X)Y = -[X,Y]$ (the bracket on the Lie algebra $\mathfrak{X}_c(N)$ of vector fields with compact support is the negative of the usual one), and:

$$\begin{aligned}
\gamma^0(X,Y) &= \int_N g(X,Y)\operatorname{vol}(g) \\
\gamma^0(\operatorname{ad}(Y)^* X, Z) &= \gamma^0(X, -[Y,Z]) = \int_N g(X, -\mathcal{L}_Y Z)\operatorname{vol}(g) \\
&= \int_N g\Big(\mathcal{L}_Y X + (g^{-1}\mathcal{L}_Y g)X + \operatorname{div}^g(Y)X, Z\Big)\operatorname{vol}(g) \\
\operatorname{ad}(Y)^* &= \mathcal{L}_Y + g^{-1}\mathcal{L}_Y(g) + \operatorname{div}^g(Y)\operatorname{Id}_T N = \mathcal{L}_Y + \beta(Y),
\end{aligned}$$

where the tensor field $\beta(Y) = g^{-1}\mathcal{L}_Y(g) + \operatorname{div}^g(Y)\operatorname{Id} : TN \to TN$ is self adjoint with respect to g. Thus the geodesic equation is

$$u_t = -(g^{-1}\mathcal{L}_u(g))(u) - \operatorname{div}^g(u)u = -\beta(u)u, \qquad u = \varphi_t \circ \varphi^{-1}.$$

The main part of the sectional curvature is given by:

$$\begin{aligned}
&4G(R(X,Y)X,Y) = \\
&= \int_N \Big(3\|[X,Y]\|_g^2 + 2g((\mathcal{L}_Y + \beta(Y))X, [X,Y]) + 2g((\mathcal{L}_X + \beta(X))Y, [Y,X]) \\
&\qquad + 4g(\beta(X)X, \beta(Y)Y) - \|\beta(X)Y + \beta(Y)X\|_g^2\Big)\operatorname{vol}(g) \\
&= \int_N \Big(-\|\beta(X)Y - \beta(Y)X + [X,Y]\|_g^2 - 4g([\beta(X), \beta(Y)]X, Y)\Big)\operatorname{vol}(g)
\end{aligned}$$

So sectional curvature consists of a part which is visibly non-negative, and another part which is difficult to decompose further.

5.4 Example: n-dimensional analog of Burgers' equation

For $(N,g)=(\mathbb{R}^n,\mathrm{can})$ or $((S^1)^n,\mathrm{can})$ we have:

$$(\mathrm{ad}(X)Y)^k = \sum_i((\partial_i X^k)Y^i - X^i(\partial_i Y^k))$$
$$(\mathrm{ad}(X)^*Z)^k = \sum_i\Big((\partial_k X^i)Z^i + (\partial_i X^i)Z^k + X^i(\partial_i Z^k)\Big),$$

so that the geodesic equation is given by

$$\partial_t u^k = -(\mathrm{ad}(u)^\top u)^k = -\sum_i\Big((\partial_k u^i)u^i + (\partial_i u^i)u^k + u^i(\partial_i u^k)\Big),$$

the n-dimensional analog of Burgers' equation.

5.5. Stronger metrics on $\mathrm{Diff}_0(N)$

A very small strengthening of the weak Riemannian H^0-metric on $\mathrm{Diff}_0(N)$ makes it into a true metric. We define the stronger right invariant semi-Riemannian metric by the formula:

$$G^A_\varphi(X\circ\varphi, Y\circ\varphi) = \int_N (g(X,Y) + A\,\mathrm{div}_g(X).\,\mathrm{div}_g(Y))\,\mathrm{vol}(g).$$

Then the following holds:

Theorem.
For any distinct diffeomorphisms φ_0,φ_1, the infimum of the lengths of all paths from φ_0 to φ_1 with respect to G^A is positive.

Proof. We may suppose that $\varphi_0=\mathrm{Id}_N$. If $\varphi_1\neq\mathrm{Id}_N$, there are two functions ρ and f on N with compact support such that:

$$\int_N \rho(y)f(\varphi_1(y))\,\mathrm{vol}(g)(y) \neq \int_N \rho(y)f(y)\,\mathrm{vol}(g)(y).$$

Now consider any path $\varphi(t,y)$ between $\varphi_0=\mathrm{Id}_N$ to φ_1 with left logarithmic derivative $u=T(\varphi)^{-1}\circ\varphi_t$ and a path in $\mathfrak{X}_c(N)$. Then we have:

$$\int_N \rho(f\circ\varphi_1)\,\mathrm{vol}(g) - \int_N \rho f\,\mathrm{vol}(g) = \int_0^1\int_N \rho\partial t f(\varphi(t,\quad)\,\mathrm{vol}(g)dt$$
$$= \int_0^1\int_N \rho(df.\varphi_t)\,\mathrm{vol}(g)\,dt = \int_0^1\int_N \rho(df.T\varphi.u)\,\mathrm{vol}(g)dt$$

$$= \int_0^1 \int_N (df.T\varphi.(\varphi u)) \operatorname{vol}(g) dt$$

Locally, on orientable pieces of N, we have:

$$\begin{aligned}
\operatorname{div}((f\circ\varphi)\rho u)\operatorname{vol}(g) &= \mathcal{L}_{(f\circ\varphi)\rho u}\operatorname{vol}(g) = (i_{(f\circ\varphi)\rho u}d + di_{(f\circ\varphi)\rho u})\operatorname{vol}(g)\\
&= d((f\circ\varphi)i_{\rho u}\operatorname{vol}(g)) = d(f\circ\varphi)\wedge i_{\rho u}\operatorname{vol}(g) + \rho\operatorname{div}(u)\operatorname{vol}(g),\\
&= d(f\circ\varphi)(\rho u)\operatorname{vol}(g) + (f\circ\varphi)\operatorname{div}(\rho u)\operatorname{vol}(g), \qquad \text{since}
\end{aligned}$$

$$d(f\circ\varphi)\wedge i_{\rho u}\operatorname{vol}(g) = -i_{\rho u}(d(f\circ\varphi)\wedge\operatorname{vol}(g)) + (i_{\rho u}d(f\circ\varphi))\operatorname{vol}(g)).$$

Thus on N we have:

$$\begin{aligned}
0 &= \int_N \operatorname{div}((f\circ\varphi)\rho u)\operatorname{vol}(g)\\
&= \int_N d(f\circ\varphi)(\rho u)\operatorname{vol}(g) + \int_N (f\circ\varphi)\operatorname{div}(\rho u)\operatorname{vol}(g)
\end{aligned}$$

and hence

$$\begin{aligned}
0 &\le \Big|\int_N \rho(f\circ\varphi_1)\operatorname{vol}(g) - \int_N \rho f\operatorname{vol}(g)\Big| = \Big|\int_0^1\int_N d(f\circ\varphi)(\varphi u))\operatorname{vol}(g)dt\Big|\\
&= \Big|\int_0^1\int_N -(f\circ\varphi)\operatorname{div}(\rho u)\operatorname{vol}(g)dt\Big|\\
&\le \sup|f|\cdot\int_0^1\sqrt{\int_N C_\rho\|u\|^2 + C'_\rho|\operatorname{div}(u)|^2\operatorname{vol}(g)}\;dt
\end{aligned}$$

for constants C_ρ, C'_ρ depending only on ρ. Clearly the right hand side gives a lower bound for the length of any path from φ_0 to φ_1. $\square$

5.6. Geodesics and sectional curvature for G^A on $\mathrm{Diff}(\mathbb{R})$

We consider the groups $\mathrm{Diff}_c(\mathbb{R})$ or $\mathrm{Diff}(S^1)$ with Lie algebras $\mathfrak{X}_c(\mathbb{R})$ or $\mathfrak{X}(S^1)$ whose Lie brackets are $\operatorname{ad}(X)Y = -[X,Y] = X'Y - XY'$. The G^A-metric equals the H^1-metric on $\mathfrak{X}_c(\mathbb{R})$, and we have:

$$\begin{aligned}
G^A(X,Y) &= \int_{\mathbb{R}}(XY + AX'Y')dx = \int_{\mathbb{R}} X(1-A\partial_x^2)Y\,dx,\\
G^A(\operatorname{ad}(X)^*Y,Z) &= \int_{\mathbb{R}}(YX'Z - YXZ' + AY'(X'Z - XZ')')dx\\
&= \int_{\mathbb{R}} Z(1-\partial_x^2)(1-\partial_x^2)^{-1}(2YX' + Y'X - 2AY''X' - AY'''X)dx,
\end{aligned}$$

$$\begin{aligned}\operatorname{ad}(X)^*Y &= (1-\partial_x^2)^{-1}(2YX' + Y'X - 2AY''X' - AY'''X)\\ \operatorname{ad}(X)^* &= (1-\partial_x^2)^{-1}(2X' + X\partial_x)(1-A\partial_x^2)\end{aligned}$$

so that the geodesic equation in Eulerian representation $u = (\partial_t f)\circ f^{-1} \in \mathfrak{X}_c(\mathbb{R})$ or $\mathfrak{X}(S^1)$ is

$$\begin{aligned}\partial_t u &= -\operatorname{ad}(u)^*u = -(1-\partial_x^2)^{-1}(3uu' - 2Au''u' - Au'''u), \text{ or}\\ u_t - u_{txx} &= Au_{xxx}.u + 2Au_{xx}.u_x - 3u_x.u,\end{aligned}$$

which for $A = 1$ is the dispersionless version of the *Camassa-Holm equation*, see (7.3.4). Note that here geodesic distance is a well defined metric describing the topology.

6. The regular Lie group of rapidly decreasing diffeomorphisms

6.1. Lemma.
For smooth functions of one variable we have:

$$\begin{aligned}(f\circ g)^{(p)}(x) &= p!\sum_{m\ge 0}\frac{f^{(m)}(g(x))}{m!}\sum_{\substack{\alpha\in\mathbb{N}_{>0}^m\\ \alpha_1+\dots+\alpha_m=p}}\prod_{i=1}^m\frac{g^{(\alpha_i)}(x)}{\alpha_i!}\\ &= \sum_{m\ge 0} f^{(m)}(g(x))\sum_{\substack{\lambda=(\lambda_n)\in\mathbb{N}_{\ge 0}^{\mathbb{N}_{>0}}\\ \sum_n\lambda_n=m\\ \sum_n\lambda_n n=p}}\frac{p!}{\lambda!}\prod_{n>0}\left(\frac{g^{(n)}(x)}{n!}\right)^{\lambda_n}\end{aligned}$$

Let $f\in C^\infty(\mathbb{R}^k)$ and let $g = (g_1,\dots,g_k)\in C^\infty(\mathbb{R}^n,\mathbb{R}^k)$. Then for a multiindex $\gamma\in\mathbb{N}^n$ the partial derivative $\partial^\gamma(f\circ g)(x)$ of the composition is given by the following formula, where we use multiindex-notation heavily.

$$\begin{aligned}&\partial^\gamma(f\circ g)(x) =\\ &= \sum_{\beta\in\mathbb{N}^k}(\partial^\beta f)(g(x))\sum_{\substack{\lambda=(\lambda_{i\alpha})\in\mathbb{N}^{k\times(\mathbb{N}^n\setminus 0)}\\ \sum_\alpha\lambda_{i\alpha}=\beta_i\\ \sum_{i\alpha}\lambda_{i\alpha}\alpha=\gamma}}\frac{\gamma!}{\lambda!}\prod_{\substack{\alpha\in\mathbb{N}^n\\ \alpha>0}}\left(\frac{1}{\alpha!}\right)^{\sum_i\lambda_{i\alpha}}\prod_{i,\alpha>0}(\partial^\alpha g_i(x))^{\lambda_{i\alpha}}\\ &= \sum_{\substack{\lambda=(\lambda_{i\alpha})\in\mathbb{N}^{k\times(\mathbb{N}^n\setminus 0)}\\ \sum_{i\alpha}\lambda_{i\alpha}\alpha=\gamma}}\frac{\gamma!}{\lambda!}\prod_{\substack{\alpha\in\mathbb{N}^n\\ \alpha>0}}\left(\frac{1}{\alpha!}\right)^{\sum_i\lambda_{i\alpha}}\left(\partial^{\sum_\alpha\lambda_\alpha}f\right)(g(x))\prod_{i,\alpha>0}(\partial^\alpha g_i(x))^{\lambda_{i\alpha}}\end{aligned}$$

The one dimensional version is due to Faà di Bruno [19], the only beatified mathematician.

Proof. We compose the Taylor expansions of

$$
\begin{aligned}
f(g(x)+h): &\quad j^\infty_{g(x)} f(h) = \sum_{m\ge 0} \frac{f^{(m)}(g(x))}{m!} h^m,\\
g(x+t): &\quad j^\infty_x g(t) = g(x) + \sum_{n\ge 1} \frac{g^{(n)}(x)}{n!} t^n,\\
f(g(x+t)): &\quad j^\infty_x (f\circ g)(t) = \sum_{m\ge 0} \frac{f^{(m)}(g(x))}{m!} \left(\sum_{n\ge 1} \frac{g^{(n)}(x)}{n!} t^n \right)^m\\
&\qquad = \sum_{m\ge 0} \frac{f^{(m)}(g(x))}{m!} \sum_{\alpha_1,\dots,\alpha_m>0} \left(\prod_{i=1}^m \frac{g^{(\alpha_i)}(x)}{\alpha_i!} \right) t^{\alpha_1+\dots+\alpha_m}.
\end{aligned}
$$

Or we use the multinomial expansion

$$
\Big(\sum_{j=1}^q a_j\Big)^m = \sum_{\substack{\lambda_1,\dots,\lambda_q\in\mathbb{N}_{\ge 0}\\ \lambda_1+\dots+\lambda_q=m}} \frac{m!}{\lambda_1!\dots\lambda_q!} a_1^{\lambda_1}\dots a_q^{\lambda_q}
$$

to get

$$
j^\infty_x (f\circ g)(t) = \sum_{m\ge 0} \frac{f^{(m)}(g(x))}{m!} \sum_{\substack{\lambda=(\lambda_n)\in\mathbb{N}_{\ge 0}^{\mathbb{N}_{>0}}\\ \sum_n \lambda_n = m}} \frac{m!}{\lambda!} \left(\prod_{n>0} \left(\frac{g^{(n)}(x)}{n!} \right)^{\lambda_n} \right) t^{\sum_n \lambda_n n}
$$

where $\lambda! = \lambda_1!\,\lambda_2!\dots$; most of the λ_i are 0. The multidimensional formula just uses more indices. $\square$

6.2

The space $\mathcal{S}(\mathbb{R})$ of all rapidly decreasing smooth functions f for which $x \mapsto (1+|x|^2)^k \partial_x^n f(x)$ is bounded for all $k\in\mathbb{N}$ and all $n\in\mathbb{N}_{\ge 0}$, with the locally convex topology described by these conditions, is a nuclear Fréchet space. The dual space $\mathcal{S}'(\mathbb{R})$ is the space of tempered distributions.

$\mathcal{S}(\mathbb{R})$ is a commutative algebra under pointwise multiplication and convolution $(u*v)(x) = \int u(x-y)v(y)dy$. The Fourier transform

$$
\mathcal{F}(u)(\xi) = \hat{u}(\xi) = \int e^{-ix\xi} u(x)dx, \qquad \mathcal{F}^{-1}(a)(x) = \frac{1}{2\pi}\int e^{ix\xi} a(\xi)d\xi,
$$

is an isomorphism of $\mathcal{S}(\mathbb{R})$ and also of $L^2(\mathbb{R})$ and has the following further properties:

$$\begin{aligned}
\widehat{\partial_x u}(\xi) &= -i\xi \cdot \hat{u}(\xi), \quad \widehat{x \cdot u}(\xi) = -i\partial_\xi \hat{u}(\xi),\\
u\widehat{(x-a)}(\xi) &= e^{ia\xi}\hat{u}(\xi), \quad e^{iax}\widehat{u(x)}(\xi) = e^{ia\xi}\hat{u}(\xi),\\
\widehat{u(ax)}(\xi) &= \tfrac{1}{|a|}\hat{u}(\tfrac{\xi}{a}), \quad u\widehat{(-x)}(\xi) = \hat{u}(-\xi),\\
\widehat{u \cdot v} &= \hat{u} * \hat{v}, \qquad \widehat{u * v} = \hat{u} \cdot \hat{v}.
\end{aligned}$$

In particular, for any polynomial P with constant coefficients we have

$$\mathcal{F}(P(-i\partial_x)u)(\xi) = P(\xi)\hat{u}(\xi).$$

$\mathcal{S}(\mathbb{R})$ satisfies the uniform $\mathcal{V}$-boundedness principle for every point separating set $\mathcal{V}$ of bounded linear functionals by [31], 5.24, since it is a Fréchet space; in particular for the set of all point evaluations $\{\mathrm{ev}_x : \mathcal{S}(\mathbb{R}) \to \mathbb{R}, \quad x \in \mathbb{R}\}$. Thus a linear mapping $\ell : E \to \mathcal{S}(\mathbb{R})$ is bounded (smooth) if and only if $\mathrm{ev}_x \circ f$ is bounded for each $x \in \mathbb{R}$.

6.3. Lemma.
The space $C^\infty(\mathbb{R}, \mathcal{S}(\mathbb{R}))$ of smooth curves in $\mathcal{S}(\mathbb{R})$ consists of all functions $f \in C^\infty(\mathbb{R}^2, \mathbb{R})$ satisfying the following property:

- *For all $n, m \in \mathbb{N}_{\geq 0}$ and each $t \in \mathbb{R}$ the expression $(1 + |x|^2)^k \partial_t^n \partial_x^m f(t, x)$ is uniformly bounded in x, locally in t.*

Proof. We use (A.3) for the set $\{\mathrm{ev}_x : x \in \mathbb{R}\}$ of point evaluations in $\mathcal{S}'(\mathbb{R})$. Note that $\mathcal{S}(\mathbb{R})$ is reflexive. Here $c^k(t) = \partial_t^k f(t, \quad)$. $\square$

6.4. Diffeomorphisms which decrease rapidly to the identity

Any orientation preserving diffeomorphism $\mathbb{R} \to \mathbb{R}$ can be written as $\mathrm{Id} + f$ for f a smooth function with $f'(x) > -1$ for all $x \in \mathbb{R}$. Let us denote by $\mathrm{Diff}_{\mathcal{S}}(\mathbb{R})_0$ the space of all diffeomorphisms $\mathrm{Id} + f : \mathbb{R} \to \mathbb{R}$ (so $f'(x) > -1$ for all $x \in \mathbb{R}$) for $f \in \mathcal{S}(\mathbb{R})$.

Theorem.
$\mathrm{Diff}_{\mathcal{S}}(\mathbb{R})_0$ *is a regular Lie group.*

Proof. Let us first check that $\mathrm{Diff}_{\mathcal{S}}(\mathbb{R})_0$ is closed under multiplication. We have

$$((\mathrm{Id} + f) \circ (\mathrm{Id} + g))(x) = x + g(x) + f(x + g(x)), \tag{1}$$

and $x \mapsto f(x + g(x))$ is in $\mathcal{S}(\mathbb{R})$ by the Faà di Bruno formula (6.1) and the following estimate:

$$f^{(m)}(x+g(x)) = O\Big(\frac{1}{(1+|x+g(x)|^2)^k}\Big) = O\Big(\frac{1}{(1+|x|^2)^k}\Big) \tag{2}$$

which holds since $g(x) \to 0$ for $|x| \to \infty$ and thus

$$\frac{1+|x|^2}{1+|x+g(x)|^2} \quad \text{is globally bounded.}$$

Let us check next that multiplication is smooth. Suppose that the curves $t \mapsto \operatorname{Id} + f(t, \quad), \operatorname{Id} + g(t, \quad)$ are in $C^\infty(\mathbb{R}, \operatorname{Diff}_{\mathcal{S}}(\mathbb{R})_0)$ which means that the functions $f, g \in C^\infty(\mathbb{R}^2, \mathbb{R})$ satisfy the conditions of lemma (6.2). Then

$$(1+|x|^2)^k \partial_t^n \partial_x^m f(t, x+g(t,x))$$

is bounded in $x \in \mathbb{R}$, locally in t, by the 2-dimensional Faá di Bruno formula (6.1) and the more elaborate version of estimate (2)

$$(\partial^{(n,m)} f)(t, x+g(t,x)) = O\Big(\frac{1}{(1+|x+g(t,x)|^2)^k}\Big) = O\Big(\frac{1}{(1+|x|^2)^k}\Big) \tag{3}$$

which follows from (6.3) for f and g. Thus the multiplication respects smooth curves and is smooth.

To check that the inverse $(\operatorname{Id} + g)^{-1}$ is again an element in $\operatorname{Diff}_{\mathcal{S}}(\mathbb{R})_0$ for $g \in \mathcal{S}(\mathbb{R})$, we write $(\operatorname{Id} + g)^{-1} = \operatorname{Id} + f$ and we have to check that $f \in \mathcal{S}(\mathbb{R})$.

$$\begin{aligned}(\operatorname{Id} + f) \circ (\operatorname{Id} + g) = \operatorname{Id} &\implies x + g(x) + f(x+g(x)) = x \\ &\implies x \mapsto f(x+g(x)) = -g(x) \text{ is in } \mathcal{S}(\mathbb{R}).\end{aligned} \tag{4}$$

Now consider

$$\begin{aligned}
\partial_x(f(x+g(x))) &= f'(x+g(x))(1+g'(x)) \\
\partial_x^2(f(x+g(x))) &= f''(x+g(x))(1+g'(x))^2 + f'(x+g(x))g''(x) \\
\partial_x^3(f(x+g(x))) &= f^{(3)}(x+g(x))(1+g'(x))^3 + \\
&\quad + 3f''(x+g(x))(1+g'(x))g''(x) + f'(x+g(x))g^{(3)}(x) \\
\partial_x^m(f(x+g(x))) &= f^{(m)}(x+g(x))(1+g'(x))^m + \\
&\quad + \sum_{k=1}^{m-1} f^{(m-k)}(x+g(x)) a_{mk}(x),
\end{aligned} \tag{5}$$

where $a_{nk} \in \mathcal{S}(\mathbb{R})$ for $n \geq k \geq 1$. We have $1 + g'(x) \geq \varepsilon > 0$ thus $\frac{1}{1+g'(x)}$ is bounded and its derivative is in $\mathcal{S}(\mathbb{R})$. Hence we can conclude that $(1+|x|^2)^k f^{(n)}(x+g(x))$ is bounded for each k. Since $(1+|x+g(x)|^2)^k = O(1+|x|^2)$ we conclude that $(1+|x+g(x)|^2)^k f^{(n)}(x+g(x))$ is bounded for all k and n. Inserting $y = x+g(x)$ it follows that $f \in \mathcal{S}(\mathbb{R})$. Thus inversion maps $\operatorname{Diff}_{\mathcal{S}}(\mathbb{R})$ into itself.

Let us check that inversion is also smooth. So we assume that $g(t,x)$ is a smooth curve in $\mathcal{S}(\mathbb{R})$, satisfies (6.3), and we have to check that then f does the same. Retracing our considerations we see from (4) that $f(t,x+g(t,x)) = -g(t,x)$ satisfies (6.3) as a function of t,x, and we claim that f then does the same. Applying ∂_t^n to the equations in (5) we get

$$\partial_t^n\partial_x^m(f(t,x+g(t,x))) = (\partial^{(n,m)}f)(t,x+g(t,x))(1+\partial_x g(t,x))^m + \\ + \sum_{\substack{k_1\le n\\ k_2\le m+n}} (\partial^{(k_1,k_2)}f)(t,x+g(t,x))a_{k_1,k_2}(t,x),$$

where $a_{k_1,k_2}(t,x) = O(\frac{1}{(1+|x|^2)^k})$ uniformly in x and locally in t. Again $1 + \partial_x g(t,x) \ge \varepsilon > 0$, locally in t and uniformly in x, thus the function $\frac{1}{1+\partial_x g(t,x)}$ is bounded with any derivative in $S(\mathbb{R})$ with respect to x. Thus we can conclude f satisfies (6.3). So the inversion is smooth and $\operatorname{Diff}_S(\mathbb{R})$ is a Lie group.

We claim that $\operatorname{Diff}_S(\mathbb{R})$ is also a regular Lie group. So let $t\mapsto X(t,\quad)$ be a smooth curve in the Lie algebra $\mathcal{S}(\mathbb{R})\partial_x$,i.e., X satisfies (6.3). The evolution of this time dependent vector field is the function given by the ODE

$$\begin{aligned}&\operatorname{Evol}(X)(t,x) = x + f(t,x),\\ &\begin{cases}\partial_t(x+f(t,x)) = f_t(t,x) = X(t,x+f(t,x)),\\ f(0,x)=0.\end{cases}\end{aligned} \tag{6}$$

We have to show that f satisfies (6.3). For $0\le t\le C$ we consider

$$|f(t,x)| \le \int_0^t |f_t(s,x)|ds = \int_0^t |X(s,x+f(s,x))|\,ds. \tag{7}$$

Since $X(t,x)$ is uniformly bounded in x, locally in t, the same is true for $f(t,x)$ by (7). But then we may insert $X(s,x+f(s,x)) = O(\frac{1}{(1+|x+f(s,x)|^2)^k}) = O(\frac{1}{(1+|x|^2)^k})$ into (7) and can conclude that $f(t,x) = O(\frac{1}{(1+|x|^2)^k})$ globally in x, locally in t, for each k. For $\partial_t^n\partial_x^m f(t,x)$ we differentiate equation (6) and arrive at a system of ODE's with functions in $\mathcal{S}(\mathbb{R})$ which we can estimate in the same way. $\square$

6.5. Sobolev spaces and HC^n-spaces

The differential operator

$$A_k = P_k(-i\partial_x) = \sum_{i=0}^k (-1)^i\partial_x^{2i}, \qquad P(\xi) = \sum_{i=0}^k \xi^{2i},$$

will play an important role later on. We consider the *Sobolev spaces*, namely the Hilbert spaces

$$H^n(\mathbb{R}) = \{f \in \mathcal{S}'(\mathbb{R}) : f, f', f^{(2)}, \dots f^{(n)} \in L^2(\mathbb{R})\}.$$

In terms of the Fourier transform $\hat{f}$ we have, by the properties listed in (6.2):

$$\begin{aligned} f \in H^n &\iff (1+|\xi|)^n \hat{f}(\xi) \in L^2 \iff (1+|\xi|^2)^{n/2} \hat{f}(\xi)) \in L^2 \\ &\iff (1+|\xi|)^{n-2k} P_k(\xi)\hat{f}(\xi) \in L^2 \iff A_k(f) \in H^{n-2k}. \end{aligned}$$

We shall use the norm

$$\|f\|_{H^n} := \|\hat{f}(\xi)(1+|\xi|)^n\|_{L^2}$$

on $H^n(\mathbb{R})$. Moreover, for $0 < \alpha \leq 1$ we consider the Banach space

$$C_b^{0,\alpha}(\mathbb{R}) = \{f \in C^0(\mathbb{R}) : \sup_{x\in\mathbb{R}} |f(x)| + \sup_{x\neq y\in\mathbb{R}} \frac{|f(x)-f(y)|}{|x-y|^\alpha} < \infty\}$$

of bounded Hölder continuous functions on $\mathbb{R}$, and the Banach spaces

$$C_b^{n,\alpha}(\mathbb{R}) = \{f \in C^n(\mathbb{R}) : f, f', \dots, f^{(n-1)} \text{ bounded, and } f^{(n)} \in C_b^{0,\alpha}(\mathbb{R})\}.$$

Finally we shall consider the space

$$HC^n(\mathbb{R}) = H^n(\mathbb{R}) \cap C_b^n(\mathbb{R}), \quad \|f\|_{HC^c} = \|f\|_{H^n} + \|f\|_{C_b^n}.$$

6.6. Lemma.
Consider the differential operator $A_k = \sum_{i=0}^k (-1)^i \partial_x^{2i}$.

(1) $A_k : \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ is a linear isomorphism of the Fréchet space of rapidly decreasing smooth functions.
(2) $A_k : H^{n+2k}(S^1) \to H^n(S^1)$ is a linear isomorphism of Hilbert spaces for each $n \in \mathbb{Z}$, where $H^n(S^1) = \{f \in L^2(S^1) : A_n(f) \in L^2(S^1)\}$. Note that $H^n(S^1) \subseteq C^k(S^1)$ if $n > k + 1/2$ (Sobolev inequality).
(3) $A_k : C^\infty(S^1) \to C^\infty(S^1)$ is a linear isomorphism.
(4) $A_k : HC^{n+2k}(\mathbb{R}) \to HC^n(\mathbb{R})$ is a linear isomorphism of Banach spaces for each $n \geq 0$.

Proof. Without loss we may consider complex-valued functions.

(1) Let $\mathcal{F} : C^\infty(S^1) \to s(\mathbb{Z})$ be the Fourier transform which is an isomorphism on the space of rapidly decreasing sequences. Since $\mathcal{F}(f_{xx})(n) = -(2\pi n)^2 \mathcal{F}(f)(n)$ we have $\mathcal{F} \circ A_k \circ \mathcal{F}^{-1} : (c_n) \mapsto ((1+(2\pi n)^2+\dots+(2\pi n)^{2k})\, c_n)$ which is a linear bibounded isomorphism.

(2) This is obvious from the definition.

(3) can be proved similarly to (1), using that the Fourier series expansion is an isomorphism between $C^\infty(S^1)$ and the space $\int$ of rapidly decreasing sequences.

(4) follows from (2). $\square$

6.7. Sobolev inequality.
We have bounded linear embeddings ($0 < \alpha \le 1$):

$$H^n(\mathbb{R}) \subset C_b^k(\mathbb{R}) \text{ if } n > k + \tfrac{1}{2},$$
$$H^n(\mathbb{R}) \subset C_b^{k,\alpha}(\mathbb{R}) \text{ if } n > k + \tfrac{1}{2} + \alpha.$$

Proof. Since $\partial_x^k : H^n(\mathbb{R}) \to H^{n-k}(\mathbb{R})$ is bounded we may assume that $k = 0$. So let $n > \frac{1}{2}$. Then we use the Cauchy-Schwartz inequality:

$$\begin{aligned} 2\pi|u(x)| &= \left|\int e^{ix\xi}\hat{u}(\xi)\,d\xi\right| \le \int |\hat{u}(\xi)|\,d\xi = \int |\hat{u}(\xi)|(1+|\xi|)^n \frac{1}{(1+|\xi|)^n}\,d\xi \\ &\le \left(\int |\hat{u}(\xi)|^2(1+|\xi|)^{2n}\,d\xi\right)^{\frac{1}{2}} \left(\int \frac{1}{(1+|\xi|)^{2n}}\,d\xi\right)^{\frac{1}{2}} = C\|u\|_{H^n} \end{aligned}$$

where

$$C = \left(\int \frac{1}{(1+|\xi|)^{2n}}\,d\xi\right)^{\frac{1}{2}} < \infty$$

depends only on $n > \frac{1}{2}$. For the second assertion we use $x > y$ and

$$\begin{aligned} e^{ix\xi} - e^{iy\xi} &= (x-y)\int_0^1 i\xi e^{i(y+t(x-y))\xi}dt, \\ \left|e^{ix\xi} - e^{iy\xi}\right| &\le |x-y|.|\xi| \end{aligned}$$

to obtain

$$\begin{aligned} &2\pi\left|\frac{u(x)-u(y)}{(x-y)^\alpha}\right| \le \int \left|\frac{e^{ix\xi}-e^{iy\xi}}{x-y}\right|^\alpha .\left|e^{ix\xi}-e^{iy\xi}\right|^{1-\alpha}|\hat{u}(\xi)|\,d\xi \\ &\le 2\int |\hat{u}(\xi)|(1+|\xi|)^n \frac{|\xi|^\alpha}{(1+|\xi|)^n}\,d\xi \\ &\le 2\left(\int |\hat{u}(\xi)|^2(1+|\xi|)^{2n}d\xi\right)^{\frac{1}{2}} \left(\int \frac{|\xi|^{2\alpha}}{(1+|\xi|)^{2n}}\,d\xi\right)^{\frac{1}{2}} = C_1\|u\|_{H^n} \end{aligned}$$

where C_1 depends only on $n - \alpha > \frac{1}{2}$. $\square$

6.8. Banach algebra property.
If $n > \frac{1}{2}$ then pointwise multiplication $\mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ extends to a bounded bilinear mapping $H^n(\mathbb{R}) \times H^n(\mathbb{R}) \to H^n(\mathbb{R})$.

For $n \ge 0$ multiplication $HC^n(\mathbb{R}) \times HC^n(\mathbb{R}) \to HC^n(\mathbb{R})$ is bounded bilinear.

See [17] for the most general version of this on open Riemannian manifolds with bounded geometry.

Proof. For $f, g \in H^n(\mathbb{R})$ we have to show that for $0 \le k \le n$ we have

$$(f.g)^{(k)} = \sum_{l=0}^{k} \binom{k}{l} f^{(l)}.g^{(k-l)} \qquad \in L^2(\mathbb{R})$$

with norm bounded by a constant times $\|f\|_{H^n}.\|g\|_{H^n}$. If $l < n$ then $f^{(l)} \in C_b^0(\mathbb{R})$ by the Sobolev inequality and $g^{(k-l)} \in H^l \subset L^2$ so the product is in L^2 with the required bound on the norm. If $l = 0$ we exchange f and g.

In the case of HC^n, the L^2-norm of each product in the sum is bounded by the sup-norm of the first factor times the L^2-norm of the second one. And the sup-norm is clearly submultiplicative. $\square$

6.9. Differentiability of composition.
If $n \ge 0$ then composition $\mathcal{S}(\mathbb{R}) \times \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R})$ extends to a weakly C^k-mapping $HC^{n+k}(\mathbb{R}) \times (\mathrm{Id}_{\mathbb{R}} + HC^n(\mathbb{R})) \to HC^n(\mathbb{R})$.

A mapping $f : E \to F$ is weakly C^1 for Banach spaces E, F if $df : E \times E \to F$ exists and is continuous. We call it strongly C^1 if $df : E \to L(E, F)$ is continuous for the operator norm on the image space. Similarly for C^k. Since I could not find a convincing proof of this result for the spaces H^n under the assumtion $n > \frac{1}{2}$, I decided to use the spaces $HC^n(\mathbb{R})$. This also inproves on the degree n which we need.

Proof. We consider the Taylor expansion

$$\begin{aligned} f(x + g(x)) = &\sum_{p=0}^{k} \frac{1}{p!} f^{(p)}(x).g(x)^p + \\ &+ \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} \big(f^{(k)}(x + tg(x)) - f^{(k)}(x) \big)\, dt\, .g(x)^k \end{aligned}$$

For fixed f this is weakly C^k in g by invoking the Banach algebra property and by estimating the integral in the remainder term. We have to show that the integrand is continuous at $(f^{(k)}, g = 0)$ as a mapping $H^n \times H^n \to H^n$. The integral from 0 to 1 does not disturb this so we disregard it. By (6.1) we have

$$\begin{aligned} &\partial_x^p (f^{(k)}(x + g(x)) - f^{(k)}(x)) = \\ &= p! \sum_{m=0}^{p} \frac{f^{(k+m)}(x + g(x))}{m!} \sum_{\substack{\alpha_1, \dots, \alpha_m > 0 \\ \alpha_1 + \dots + \alpha_m = p}} \frac{\partial_x^{\alpha_1}(x + g(x))}{\alpha_1!} \cdots \frac{\partial_x^{\alpha_m}(x + g(x))}{\alpha_m!} \end{aligned}$$

The most dangerous term is the one for $p = n$. As soon as a derivative of g of order ≥ 2 is present, this is easily estimated. The most difficult term is

$$f^{(k+n)}(x+g(x)) - f^{(k+n)}(x)$$

which should go to 0 in $L^2 \cap C^0_b$ for fixed f and for $g \to 0$ in HC^n. $f^{(k)}$ is continuous and in L^2. Off some big compact intervall it has small H^n-norm and small sup-norm (the latter by the lemma of Riemann-Lebesque). On this compact intervall $f^{(k)}$ is uniformly continuous and if we choose $\|g\|_{C^n}$ small enough, $f^{(k)}(x+tg(x)) - f^{(k)}(x)$ is uniformly small there, thus small in the sup-norm, and also small in L^2 (which involves the length of the compact intervall – but we can still choose g smaller). $\square$

The last result cannot be improved to strongly C^k since we have:

6.10. Attention.
Composition $HC^n(\mathbb{R}) \times (\mathrm{Id}_{\mathbb{R}} + HC^n(\mathbb{R})) \to HC^n(\mathbb{R})$ is only continuous and not Lipschitz in the first variable.

Proof. To see this, consider $(f,t) \mapsto f(\quad - t.g)$ for a given bump function g which equals 1 on a large intervall. For each $t > 0$ we consider a bump function f with support in $(-\frac{t}{2}, \frac{t}{2})$ with $\|f\|_{L^2} = 1$. Then we have $\|f - f(\quad - t)\|_{L^2} = \sqrt{2}$ by Pythagoras, and consequently $\|f - f(\quad - t.g)\|_{HC^n} \geq \|f - f(\quad - t)\|_{L^2} = \sqrt{2}$. $\square$

6.11. The topological group Diff$(\mathbb{R})$

For $n \geq 1$ we consider $f : \mathbb{R} \to \mathbb{R}$ of the form $f(x) = x + g(x)$ for $g \in HC^n$. Then f is a C^n-diffeomorphism iff $g'(x) > -1$ for all x. The inverse is also of the form $f^{-1}(y) = y + h(y)$ for $h \in HC^n(\mathbb{R})$ iff $g'(x) \geq -1 + \varepsilon$ for a constant ε. Indeed, $h(y) = -g(f^{-1}(y))$. Let us call $\mathrm{DiffHC}^n(\mathbb{R})$ the group of all these diffeomorphsms.

Lemma.
Inversion $\mathrm{DiffHC}^{n+k}(\mathbb{R}) \to \mathrm{DiffHC}^n(\mathbb{R})$ is weakly C^k.

Proof. As we saw above, $\mathrm{DiffHC}^{n+k}(\mathbb{R})$ is stable under inversion. $(f,g) \mapsto f \circ g$ is a weak C^k submersion by (6.9). So we can use the implicit function theorem for the equation $f \circ f^{-1} = \mathrm{Id}$. $\square$

6.12 Proposition.
For $n \geq 1$ and $a \in HC^n(\mathbb{R})$, the mapping $HC^n(\mathbb{R}) \times \mathrm{DiffHC}^n(\mathbb{R}) \to HC^{n-1}(\mathbb{R})$ given by $(f,g) \mapsto (a\partial_x(f \circ g^{-1})) \circ g$ is continuous and Lipschitz in f.

For $n > k + \frac{1}{2}$ and for each linear differential operator D of order k, the mapping $HC^n(\mathbb{R}) \times \mathrm{DiffHC}^n(\mathbb{R}) \to HC^{n-k}(\mathbb{R})$ given by $(f,g) \mapsto (D(f \circ g^{-1})) \circ g$ is continuous and Lipschitz in f.

Here $\mathrm{Diff}(\mathbb{R}) = \{\mathrm{Id}_{\mathbb{R}} + h : \|h'\|_{C^0_b} > -1\}$.

Proof. We have

$$(a\partial_x(f\circ g^{-1}))\circ g = \Big(a.(f_x\circ g^{-1})\frac{1}{g_x\circ g^{-1}}\Big)\circ g = (a\circ g).f_x.\frac{1}{g_x}$$

which is Lipschitz by the results above. □

6.13 Proposition.
For the operator $A_k = \sum_{i=0}^k (-1)^i \partial_x^{2i}$ and for $n \ge 2k$, the mapping $(f,g)\mapsto (A_k^{-1}(f\circ g^{-1}))\circ g$ is Lipschitz $HC^n(\mathbb{R})\times \mathrm{DiffHC}^n(\mathbb{R}) \to HC^{n+2k}(\mathbb{R})$.

Proof. The inverse of A_k is given by the pseudo differential operator

$$(A_k^{-1}f)(x) = \int_{\mathbb{R}^2} e^{i(x-y)\xi} f(y)\frac{1}{1+\xi^2+\quad+\xi^{2n}}d\xi\,dy$$

Thus the mapping is given by

$$\begin{aligned}(A_k^{-1}(f\circ g^{-1}))(g(x)) &= \int_{\mathbb{R}^2} e^{i(g(x)-y)\xi} f(g^{-1}(y))\frac{1}{1+\xi^2+\quad+\xi^{2n}}d\xi\,dy\\ &= \int_{\mathbb{R}^2} e^{i(g(x)-g(z))\xi} f(z)\frac{g'(z)}{1+\xi^2+\quad+\xi^{2n}}d\xi\,dz\end{aligned}$$

which is a genuine Fourier integral operator. By the foregoing results this is visibly locally Lipschitz. □

7. The diffeomorphism group of S^1 or $\mathbb{R}$, and Burgers' hierarchy

7.1. Burgers' equation and its curvature

We consider the Lie groups $\mathrm{Diff}_{\mathcal{S}}(\mathbb{R})$ and $\mathrm{Diff}(S^1)$ with Lie algebras $\mathfrak{X}_{\mathcal{S}}(\mathbb{R})$ and $\mathfrak{X}(S^1)$ where the Lie bracket $[X,Y]=X'Y-XY'$ is the negative of the usual one. For the L^2-inner product $\gamma(X,Y)=\langle X,Y\rangle_0 = \int X(x)Y(x)\,dx$ integration by parts gives

$$\begin{aligned}\langle [X,Y],Z\rangle_0 &= \int_{\mathbb{R}} (X'YZ-XY'Z)dx\\ &= \int_{\mathbb{R}} (2X'YZ+XYZ')dx = \langle Y, \mathrm{ad}(X)^\top Z\rangle,\end{aligned}$$

which in turn gives rise to

$$\mathrm{ad}(X)^\top Z = 2X'Z + XZ', \tag{1}$$

$$\alpha(X)Z = \mathrm{ad}(Z)^\top X = 2Z'X + ZX', \tag{2}$$

$$(\mathrm{ad}(X)^\top + \mathrm{ad}(X))Z = 3X'Z, \tag{3}$$

$$(\mathrm{ad}(X)^\top - \mathrm{ad}(X))Z = X'Z + 2XZ' = \alpha(X)Z. \tag{4}$$

Equation (4) states that $-\frac{1}{2}\alpha(X)$ is the skew-symmetrization of $\mathrm{ad}(X)$ with respect to the inner product $\langle \quad , \quad \rangle_0$. From the theory of symmetric spaces one then expects that $-\frac{1}{2}\alpha$ is a Lie algebra homomorphism and indeed one can check that

$$-\tfrac{1}{2}\alpha([X,Y]) = \left[-\tfrac{1}{2}\alpha(X), -\tfrac{1}{2}\alpha(Y)\right]$$

holds for any vector fields X, Y. From (1) we get the geodesic equation, whose second part is Burgers' equation [10]:

$$\begin{cases} g_t(t,x) = u(t, g(t,x)) \\ u_t = -\,\mathrm{ad}(u)^\top u = -3u_x u \end{cases} \tag{5}$$

Using the above relations and the general curvature formula (3.4.2), we get

$$\begin{aligned} \mathcal{R}(X,Y)Z &= -X''YZ + XY''Z - 2X'YZ' + 2XY'Z' \\ &= -2[X,Y]Z' - [X,Y]'Z = -\alpha([X,Y])Z. \end{aligned} \tag{6}$$

Sectional curvature is non-negative and unbounded:

$$\begin{aligned} -G^0_a(R(X,Y)X,Y) &= \langle \alpha([X,Y])(X), Y\rangle = \langle \mathrm{ad}(X)^\top([X,Y]), Y\rangle \\ &= \langle [X,Y],[X,Y]\rangle = \|[X,Y]\|^2, \\ k(X\wedge Y) &= -\frac{G^0_a(R(X,Y)X,Y)}{\|X\|^2\|Y\|^2 - G^0_a(X,Y)^2} \\ &= \frac{\|[X,Y]\|^2}{\|X\|^2\|Y\|^2 - \langle X,Y\rangle^2} \ge 0. \end{aligned} \tag{7}$$

Let us check invariance of the momentum mapping $\bar J$ from (4.3):

$$\begin{aligned} \gamma(\bar J(g,X),Y) &= \gamma(\mathrm{Ad}(g)^\top X, Y) = \gamma(X, \mathrm{Ad}(g)Y) = \int X((g'Y)\circ g^{-1})dx \\ &= \int X(g'\circ g^{-1})(Y\circ g^{-1})dx = \mathrm{sign}(g')\int (X\circ g)(g')^2 Y dx \\ &= \mathrm{sign}(g')\gamma((g')^2(X\circ g), Y) \\ \bar J(g,X) &= \mathrm{sign}(g_x).(g_x)^2(X\circ g). \end{aligned} \tag{8}$$

Along a geodesic $t \mapsto g(t, \quad)$, according to (5) and (4.3), the momentum

$$\bar J(g, u = g_t\circ g^{-1}) = g_x^2 g_t \quad \text{ is constant.} \tag{9}$$

This is what we found in (1.3) by chance.

7.2. Jacobi fields for Burgers' equation

A Jacobi field y along a geodesic g with velocity field u is a solution of the partial differential equation (3.5.1), which in our case becomes:

$$\begin{aligned} y_{tt} &= [\mathrm{ad}(y)^\top + \mathrm{ad}(y), \mathrm{ad}(u)^\top]u - \mathrm{ad}(u)^\top y_t - \alpha(u)y_t + \mathrm{ad}(u)y_t \\ &= -3u^2 y_{xx} - 4uy_{tx} - 2u_x y_t \\ u_t &= -3u_x u. \end{aligned} \tag{1}$$

If the geodesic equation has smooth solutions locally in time it is to be expected that the space of all Jacobi fields exists and is isomorphic to the space of all initial data $(y(0), y_t(0)) \in C^\infty(S^1, \mathbb{R})^2$ or $C_c^\infty(\mathbb{R}, \mathbb{R})^2$, respectively. The weak symplectic structure on it is given by (3.7):

$$\begin{aligned} \omega(y,z) &= \langle y, z_t - \tfrac{1}{2}u_x z + 2uz_x\rangle - \langle y_t - \tfrac{1}{2}u_x y + 2uy_x, z\rangle \\ &= \int_{S^1 \text{or } \mathbb{R}} (yz_t - y_t z + 2u(yz_x - y_x z))\, dx. \end{aligned} \tag{2}$$

7.3. The Sobolev H^k-metric on $\mathrm{Diff}(S^1)$ and $\mathrm{Diff}(\mathbb{R})$

On the Lie algebras $\mathfrak{X}_c(\mathbb{R})$ and $\mathfrak{X}(S^1)$ with Lie bracket $[X,Y] = X'Y - XY'$ we consider the H^k-inner product

$$\begin{aligned} \gamma(X,Y) = \langle X, Y\rangle_k &= \sum_{i=0}^k \int (\partial_x^i X)(\partial_x^i Y)\, dx = \int A_k(X)(Y)\, dx \\ &= \int X A_k(Y)\, dx, \qquad \text{where} \qquad A_k = \sum_{i=0}^k (-1)^i \partial_x^{2i} \end{aligned} \tag{1}$$

is a linear isomorphism $\mathfrak{X}_c(\mathbb{R}) \to \mathfrak{X}_c(\mathbb{R})$ or $\mathfrak{X}(S^1) \to \mathfrak{X}(S^1)$ whose inverse is a pseudo differential operator. A_k is also a bounded linear isomorphism between the Sobolev spaces $H^{l+2k}(S^1) \to H^l(S^1)$, see lemma (6.5). On the real line we have to consider functions with fixed support in some compact set $[-K, K] \subset \mathbb{R}$.

Integration by parts gives

$$\begin{aligned} \langle [X,Y], Z\rangle_k &= \int_{\mathbb{R}} (X'Y - XY')A_k(Z)dx = \int_{\mathbb{R}} (2X'YA_k(Z) + XYA_k(Z'))dx \\ &= \int_{\mathbb{R}} YA_k A_k^{-1}\big(2X'A_k(Z) + XA_k(Z')\big)dx = \langle Y, \mathrm{ad}(X)^{\top,k}, Z\rangle_k, \end{aligned}$$

which in turn gives rise to

$$
\begin{aligned}
\operatorname{ad}(X)^{\top,k} Z &= A_k^{-1}\big(2X' A_k(Z) + X A_k(Z')\big),\\
\alpha_k(X) Z = \operatorname{ad}(Z)^{\top,k}(X) &= A_k^{-1}\big(2Z' A_k(X) + Z A_k(X')\big)
\end{aligned} \tag{2}
$$

Thus the geodesic equation is

$$
\begin{cases}
g_t(t,x) = u(t, g(t,x)) \\
u_t = -\operatorname{ad}(u)^{\top,k} u = -A_k^{-1}\big(2u_x A_k(u) + u A_k(u_x)\big) \\
\quad = -A_k^{-1}\Big(2u_x \sum_{i=0}^{k} (-1)^i \partial_x^{2i} u + u \sum_{i=0}^{k} (-1)^i \partial_x^{2i+1} u\Big).
\end{cases} \tag{3}
$$

For $k = 0$ the second part is Burgers' equation, and for $k = 1$ it becomes

$$
\begin{aligned}
u_t - u_{txx} &= -3uu_x + 2u_x u_{xx} + u u_{xxx} \\
\iff u_t + uu_x &+ (1 - \partial_x^2)^{-1}(u^2 + \tfrac{1}{2}u_x^2)_x = 0
\end{aligned} \tag{4}
$$

which is the dispersionfree version of the *Camassa-Holm equation*, see [11], [44], [29]. We met it already in (5.6), and will meet the full equation in (8.7). Let us check the invariant momentum mapping from (4.3.2):

$$
\begin{aligned}
\gamma(\bar J(g,X), Y) &= \langle \operatorname{Ad}(g)^\top X, Y\rangle_k = \langle X, \operatorname{Ad}(g) Y\rangle_k \\
&= \int A_k(X)(g' \circ g^{-1})(Y \circ g^{-1}) dx = \operatorname{sign}(g') \int (A_k(X) \circ g)(g')^2 Y\, dx \\
&= \operatorname{sign}(g') \Big\langle A_k^{-1}\Big((g')^2 (A_k(X) \circ g)\Big), Y \Big\rangle_k \\
\bar J(g,X) &= \operatorname{sign}(g_x).A_k^{-1}\Big((g_x)^2 (A_k(X) \circ g)\Big).
\end{aligned} \tag{5}
$$

Along a geodesic $t \mapsto g(t, \quad)$, by (3) and (4.3), the expressions

$$
\operatorname{sign}(g_x) \bar J(g, u = g_t \circ g^{-1}) = A_k^{-1}\Big((g_x)^2 (A_k(u) \circ g)\Big) \tag{6}
$$

and thus also $(g_x)^2 (A_k(u) \circ g)$ are constant in t.

7.4. Theorem.
Let $k \ge 1$. There exists a HC^{2k+1}-open neighborhood V of $(\mathrm{Id}, 0)$ in $\mathrm{Diff}(S^1) \times \mathfrak{X}(S^1)$ such that for each $(g_0, u_0) \in V$ there exists a unique C^3 geodesic $g \in C^3((-2,2), \mathrm{Diff}(S^1))$ for the right invariant H^k Riemann metric, starting at $g(0) = g_0$ in the direction $g_t(0) = u_0 \circ g_0 \in T_{g_0} \mathrm{Diff}(S^1)$. Moreover, the solution depends C^1 on the initial data $(g_0, u_0) \in V$.

The same result holds if we replace $\mathrm{Diff}(S^1)$ by $\mathrm{Diff}_{\mathcal{S}(\mathbb{R})}$ and $\mathfrak{X}(S^1)$ by $\mathfrak{X}_{\mathcal{S}}(\mathbb{R}) = \mathcal{S}(\mathbb{R})\partial_x$.

This result is stated in [13], and also this proof follows essentially [13]. But there is a mistake in [13], p 795, where the authors assume that composition and inversion on $H^n(S^1)$ are smooth. This is wrong. One needs to use (6.12) and (6.13). The mistake was corrected in [12], for the more general case of the Virasoro group.

In the following proof, Diff, $\mathfrak{X}$, DiffHC^n, HC^n should stand for either Diff(S^1), $\mathfrak{X}(S^1)$, DiffHC$^n(S^1)$, $HC^n(S^1)$ or for Diff$_{\mathcal{S}}(\mathbb{R})$, $\mathfrak{X}_{\mathcal{S}}(\mathbb{R})$, DiffHC$^n(\mathbb{R})$, $HC^n(\mathbb{R})$, respectively.

Proof. For $u \in HC^n$, $n \geq 2k+1$, we have

$$\begin{aligned} A_k(uu_x) &= \sum_{i=0}^k (-1)^i \partial_x^{2i}(uu_x) = \sum_{i=0}^k (-1)^i \sum_{j=0}^{2i} \tbinom{2i}{j} (\partial_x^j u)(\partial_x^{2i-j+1} u) \\ &= uA_k(u_x) + \sum_{i=0}^k (-1)^i \sum_{j=1}^{2i} \tbinom{2i}{j} (\partial_x^j u)(\partial_x^{2i-j+1} u) \\ &=: u\,A_k(u_x) + B_k(u), \end{aligned}$$

where $B_k : HC^n \to HC^{n-2k}$ is a bounded quadratic operator. Recall that we have to solve

$$\begin{aligned} u_t &= -\operatorname{ad}(u)^{\top,k} u = -A_k^{-1}\big(2u_x A_k(u) + uA_k(u_x)\big) \\ &= -A_k^{-1}\big(2u_x A_k(u) + A_k(uu_x) - B_k(u)\big) \\ &= -uu_x - A_k^{-1}\big(2u_x A_k(u) - B_k(u)\big) \\ &=: -uu_x + A_k^{-1} C_k(u), \end{aligned}$$

where $C_k : HC^n \to HC^{n-2k}$ is a bounded quadratic operator, and where $u = g_t \circ g^{-1} \in \mathfrak{X}$. Note that

$$\begin{aligned} C_k(u) &= -2u_x A_k(u) + B_k(u) \\ &= -2u_x A_k(u) + \sum_{i=0}^k (-1)^i \sum_{j=1}^{2i} \tbinom{2i}{j} (\partial_x^j u)(\partial_x^{2i-j+1} u). \end{aligned}$$

We put

$$\left\{ \begin{aligned} & g_t =: v = u \circ g \\ & v_t = u_t \circ g + (u_x \circ g) g_t = u_t \circ g + (uu_x) \circ g = A_k^{-1} C_k(u) \circ g \\ & \quad = A_k^{-1} C_k(v \circ g^{-1}) \circ g =: \mathrm{pr}_2(D_k \circ E_k)(g,v), \qquad \text{where} \end{aligned} \right. \tag{1}$$

$$E_k(g,v) = (g, C_k(v \circ g^{-1}) \circ g), \qquad D_k(g,v) = (g, A_k^{-1}(v \circ g^{-1}) \circ g).$$

Now consider the topological group and Banach manifold DiffHC^n described in (6.11).

(2) *Claim.* The mapping D_k : DiffHC$^n \times HC^{n-2k} \to$ DiffHC$^n \times HC^n$ is strongly C^1.

First we check that all directional derivatives exist and are in the right spaces.

For $w \in HC^n$ we have

$$\begin{aligned}
&\partial_s|_0(u\circ(g+sw)) = (u_x\circ g)w\\
&\partial_s|_0(g+sw)^{-1} = -\frac{w\circ g^{-1}}{g_x\circ g^{-1}}\\
&\partial_s|_0 \operatorname{pr}_2 D_k(g+sw,v) =\\
&= \partial_s|_0 A_k^{-1}(v\circ g^{-1})\circ(g+sw) + \partial_s|_0(A_k^{-1}(v\circ(g+sw)^{-1}))\circ g\\
&= ((\partial_x A_k^{-1}(v\circ g^{-1}))\circ g)\,w - (A_k^{-1}((v_x\circ g^{-1})\tfrac{w\circ g^{-1}}{g_x\circ g^{-1}}))\circ g\\
&= (A_k^{-1}(v\circ g^{-1})_x.(w\circ g^{-1}))\circ g - (A_k^{-1}((v\circ g^{-1})_x(w\circ g^{-1})))\circ g.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&A_k((\partial_s|_0 \operatorname{pr}_2 D_k(g+sw,v))\circ g^{-1}) =\\
&= A_k(A_k^{-1}(v\circ g^{-1})_x.(w\circ g^{-1})) - (v\circ g^{-1})_x(w\circ g^{-1})\\
&= (v\circ g^{-1})_x.(w\circ g^{-1}) + \sum_{i=0}^{k}\sum_{j=0}^{2i-1}\binom{2i}{j}\partial_x^{j+1}A_k^{-1}(v\circ g^{-1}).\partial_x^{2k-j}(w\circ g^{-1})\\
&\quad - (v\circ g^{-1})_x(w\circ g^{-1}) \in HC^{n-2k}.
\end{aligned}$$

By (6.12) and (6.13) this is locally Lipschitz jointly in v, g, w. Moreover we have $\partial_s|_0 \operatorname{pr}_2 D_k(g+sw,v) \in HC^n$, and D_k is linear in v. Thus D_k is strongly C^1.

(3) *Claim.* The mapping $E_k : \mathrm{DiffHC}^n \times HC^n \to \mathrm{DiffHC}^n \times HC^{n-2k}$ is strongly C^1. This can be proved similarly, again using (6.12) and (6.13).

By the two claims equation (1) can be viewed as the flow equation of a C^1-vector field on the Hilbert manifold $\mathrm{DiffHC}^n \times HC^n$. Here an existence and uniqueness theorem holds. Since $v = 0$ is a stationary point, there exist an open neighborhood W_n of $(\mathrm{Id}, 0)$ in $\mathrm{DiffHC}^n \times HC^n$ such that for each initial point $(g_0, v_0) \in W_n$ equation (1) has a unique solution $\mathrm{Fl}^n_t(g_0, v_0) = (g(t), v(t))$ defined and C^2 in $t \in (-2, 2)$. Note that $v(t) = g_t(t)$, thus $g(t)$ is even C^3 in t. Moreover, the solution depends C^1 on the initial data.

We start with the neighborhood

$$W_{2k+1} \subset \mathrm{DiffHC}^{2k+1} \times HC^{2k+1} \supset \mathrm{DiffHC}^n \times HC^n \quad \text{for } n \ge 2k+1$$

and consider the neighborhood $V_n := W_{2k+1} \cap \mathrm{DiffHC}^n \times HC^n$ of $(\mathrm{Id}, 0)$

(4) *Claim.* For any initial point $(g_0, v_0) \in V_n$ the unique solution $\mathrm{Fl}^n_t(g_0, v_0) = (g(t), v(t))$ exists, is C^2 in $t \in (-2, 2)$, and depends C^1 on the initial point in V_n.

We use induction on $n \ge 2k+1$. For $n = 2k+1$ the claim holds since $V_{2k+1} = W_{2k+1}$. Let $(g_0, v_0) \in V_{2k+2}$ and let $\mathrm{Fl}^{2k+2}_t(g_0, v_0) = (\tilde g(t), \tilde v(t))$ be maximally defined for $t \in (t_1, t_2) \ni 0$. Suppose for contradiction that $t_2 < 2$. Since $(g_0, v_0) \in V_{2k+2} \subset V_{2k+1}$ the curve $\mathrm{Fl}^{2k+2}_t(g_0, v_0) = (\tilde g(t), \tilde v(t))$

solves (1) also in $\mathrm{DiffHC}^{2k+1} \times HC^{2k+1}$, thus $\mathrm{Fl}^{2k+2}_t(g_0, v_0) = (\tilde{g}(t), \tilde{v}(t)) = (g(t), v(t)) := \mathrm{Fl}^{2k+1}_t(g_0, v_0)$ for $t \in (t_1, t_2) \cap (-2, 2)$. By (7.3.6), the expression

$$\tilde{J}(t) = \tilde{J}(g, v, t) = g_x(t)^2 A_k(u(t)) \circ g(t) = g_x(t)^2 A_k(v(t) \circ g(t)^{-1}) \circ g(t) \quad (5)$$

is constant in $t \in (-2, 2)$. Actually, since we used C^∞-theory for deriving this, one should check it again by differentiating. Since $u = g_t \circ g^{-1}$ we get the following (the exact formulas can be computed with the help of Faà di Bruno's formula (6.1).

$$\begin{aligned}
u_x &= (g_{tx} \circ g^{-1})(g^{-1})_x = \frac{g_{tx}}{g_x} \circ g^{-1} \\
\partial_x^2 u &= (\frac{\partial_x^2 g_t}{g_x^2} - g_{tx} \frac{\partial_x^2 g}{g_x^3}) \circ g^{-1} \\
\partial_x(g^{-1}) &= \frac{1}{g_x} \circ g^{-1} \\
\partial_x^2(g^{-1}) \circ g &= -\frac{\partial_x^2 g}{g_x^3} \\
\partial_x^{2k}(g^{-1}) \circ g &= -\frac{\partial_x^{2k} g}{g_x^{2k+1}} + \text{ lower order terms in } g \\
(\partial_x^{2k} u) \circ g &= \frac{\partial_x^{2k} g_t}{g_x^{2k}} - g_{tx} \frac{\partial_x^{2k} g}{g_x^{2k+1}} + \text{ lower order terms in } g, g_t = v.
\end{aligned}$$

Thus

$$(-1)^k g_x^{2k-1} \tilde{J}(t) = g_x \partial_x^{2k} g_t - g_{tx} \partial_x^{2k} g + \text{ lower order terms in } g, g_t = v.$$

Hence for each $t \in (-2, 2)$:

$$\begin{aligned}
g_x \partial_x^{2k} g_t - g_{tx} \partial_x^{2k} g &= (-1)^k g_x^2 \left(g_x^{2k-3} \tilde{J}(t) + P_k(g, v) \right), \text{ where} \\
P_k(g, v) &= \frac{Q_k(g, \partial_x g, \dots, \partial_x^{2k-1} g, v, \partial_x v, \dots, \partial_x^{2k-1} v)}{g_x^2}
\end{aligned}$$

for a polynomial Q_k. Since $\tilde{J}(t) = \tilde{J}(0)$ we obtain that

$$\left(\frac{\partial_x^{2k} g(t)}{g_x(t)} \right)_t = (-1)^k \left(g_x^{2k-3}(t) \tilde{J}(0) + P_k(g(t), v(t)) \right) \text{ for all } t \in (-2, 2).$$

This implies

$$\frac{\partial_x^{2k} g(t)}{g_x(t)} = \frac{\partial_x^{2k} g(0)}{g_x(0)} + (-1)^k \int_0^t \left(g_x^{2k-3}(s) \tilde{J}(0) + P_k(g(s), v(s)) \right) ds.$$

For $t \in (t_1, t_2)$ we have

$$\partial_x^{2k}\tilde g(t) = \frac{\partial_x^{2k} g_0}{\partial_x g_0} g_x(t) + \tag{6}$$
$$+ (-1)^k g_x(t) \int_0^t \Big(g_x^{2k-3}(s)\tilde J(0) + P_k(g(s), v(s))\Big)\, ds.$$

Since $(g_0, v_0) \in V_{2k+2}$ we have $\tilde J(0) = \tilde J(g_0, v_0, 0) \in HC^2$ by (5). Since $k \geq 1$, by (6) we see that $\partial_x^{2k}\tilde g(t) \in HC^2$. Moreover, since $t_2 < 2$, the limit $\lim_{t\to t_2-} \partial_x^{2k}\tilde g(t)$ exists in HC^2, so $\lim_{t\to t_2-}\tilde g(t)$ exists in HC^{2k+2}. As this limit equals $g(t_2)$, we conclude that $g(t_2) \in \mathrm{DiffHC}^{2k+2}$. Now $\tilde v = \tilde g_t$; so we may differentiate both sides of (6) in t and obtain similarly that $\lim_{t\to t_2-}\tilde v(t)$ exists in HC^{2k+2} and equals $v(t_2)$. But then we can prolong the flow line $(\tilde g, \tilde v)$ in $\mathrm{DiffHC}^{2k+2} \times HC^{2k+2}$ beyond t_2, so (t_1, t_2) was not maximal.

By the same method we can iterate the induction. □

8. The Virasoro-Bott group and the Korteweg-de Vries hierarchy

8.1. The Virasoro-Bott group

Let Diff denote any of the groups $\mathrm{DiffHC}^+(S^1)$, $\mathrm{Diff}(\mathbb{R})_0$ (diffeomorphisms with compact support), or $\mathrm{Diff}_{\mathcal S}(\mathbb{R})$ of section (6). For $\varphi \in \mathrm{Diff}$ let $\varphi' : S^1$ or $\mathbb{R} \to \mathbb{R}^+$ be the mapping given by $T_x\varphi\cdot\partial_x = \varphi'(x)\partial_x$. Then

$$c : \mathrm{Diff}\times\mathrm{Diff} \to \mathbb{R}$$
$$c(\varphi,\psi) := \frac12\int_{S^1} \log(\varphi\circ\psi)' d\log\psi' = \frac12\int_{S^1}\log(\varphi'\circ\psi)d\log\psi'$$

satisfies $c(\varphi,\varphi^{-1}) = 0$, $c(\mathrm{Id},\psi) = 0$, $c(\varphi,\mathrm{Id}) = 0$, and is a smooth group cocycle, i.e.,

$$c(\varphi_2,\varphi_3) - c(\varphi_1\circ\varphi_2,\varphi_3) + c(\varphi_1,\varphi_2\circ\varphi_3) - c(\varphi_1,\varphi_2) = 0,$$

called the Bott cocycle.

Proof. Let us check first:

$$\int \log(\varphi\circ\psi)'d\log\psi' = \int \log((\varphi'\circ\psi)\psi')d\log\psi' =$$
$$= \int\log(\varphi'\circ\psi)d\log\psi' + \int\log(\psi')d\log\psi',$$
$$\int\log(\psi')d\log\psi' = \tfrac12\int d\,\log(\psi')^2 = 0.$$

$$
\begin{aligned}
2c(\mathrm{Id},\psi) &= \int \log(1) d\log\psi' = 0.\\
2c(\varphi,\mathrm{Id}) &= \int \log(\varphi') d\log(1) = 0.\\
2c(\varphi^{-1},\varphi) &= \int \log((\varphi^{-1}\circ\varphi)') d\log\varphi' = \int \log(1) d\log\varphi' = 0.\\
c(\varphi,\varphi^{-1}) &= 0.
\end{aligned}
$$

For the cocycle condition we add the following terms:

$$
\begin{aligned}
2c(\varphi_2,\varphi_3) &= \int \log(\varphi_2'\circ\varphi_3) d\log\varphi_3'\\
-2c(\varphi_1\circ\varphi_2,\varphi_3) &= -\int \log((\varphi_1\circ\varphi_2)'\circ\varphi_3) d\log\varphi_3'\\
&= -\int \log((\varphi_1'\circ\varphi_2\circ\varphi_3)(\varphi_2'\circ\varphi_3)) d\log\varphi_3'\\
&= -\int \log(\varphi_1'\circ\varphi_2\circ\varphi_3) d\log\varphi_3' - \int \log(\varphi_2'\circ\varphi_3) d\log\varphi_3'\\
2c(\varphi_1,\varphi_2\circ\varphi_3) &= \int \log(\varphi_1'\circ\varphi_2\circ\varphi_3) d\log(\varphi_2\circ\varphi_3)'\\
&= \int \log(\varphi_1'\circ\varphi_2\circ\varphi_3) d\log((\varphi_2'\circ\varphi_3)\varphi_3')\\
&= \int \log(\varphi_1'\circ\varphi_2\circ\varphi_3) d\log(\varphi_2'\circ\varphi_3) + \int \log(\varphi_1'\circ\varphi_2\circ\varphi_3) d\log\varphi_3'\\
&= \int \log(\varphi_1'\circ\varphi_2) d\log\varphi_2' + \int \log(\varphi_1'\circ\varphi_2\circ\varphi_3) d\log\varphi_3'\\
-2c(\varphi_1,\varphi_2) &= -\int \log(\varphi_1'\circ\varphi_2) d\log\varphi_2' \qquad \square
\end{aligned}
$$

The corresponding central extension group $S^1 \times_c \mathrm{DiffHC}^+(S^1)$, called the periodic Virasoro-Bott group, is a trivial S^1-bundle $S^1 \times \mathrm{DiffHC}^+(S^1)$ that becomes a regular Lie group relative to the operations

$$
\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}\begin{pmatrix}\psi\\ \beta\end{pmatrix} = \begin{pmatrix}\varphi\circ\psi\\ \alpha\beta\, e^{2\pi i c(\varphi,\psi)}\end{pmatrix}, \quad \begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^{-1} = \begin{pmatrix}\varphi^{-1}\\ \alpha^{-1}\end{pmatrix}
$$

for $\varphi,\psi \in \mathrm{DiffHC}^+(S^1)$ and $\alpha,\beta\in S^1$. Likewise we have the central extension group with compact supports $\mathbb{R}\times_c \mathrm{Diff}(\mathbb{R})_0$ with group operations

$$
\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}\begin{pmatrix}\psi\\ \beta\end{pmatrix} = \begin{pmatrix}\varphi\circ\psi\\ \alpha+\beta+c(\varphi,\psi)\end{pmatrix}, \quad \begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^{-1} = \begin{pmatrix}\varphi^{-1}\\ -\alpha\end{pmatrix}
$$

for $\varphi,\psi\in\operatorname{DiffHC}^+(\mathbb{R})$ and $\alpha,\beta\in\mathbb{R}$. Finally there is the central extension of the rapidly decreasing Virasoro-Bott group $\mathbb{R}\times_c \operatorname{Diff}^+_{\mathcal{S}}(\mathbb{R})$ which is given by the same formulas.

8.2. The Virasoro Lie algebra

Let us compute the Lie algebra of the two versions of the the Virasoro-Bott group. Consider $\mathbb{R}\times_c \operatorname{Diff}$, where again Diff denotes any one of the groups $\operatorname{DiffHC}^+(S^1)$, $\operatorname{Diff}(\mathbb{R})_0$, or $\operatorname{Diff}_{\mathcal{S}}(\mathbb{R})$. So let $\varphi,\psi:\mathbb{R}\to\operatorname{Diff}$ with $\varphi(0)=\psi(0)=\operatorname{Id}$ and $\varphi_t(0)=X$, $\psi_t(0)=Y\in X_c(\mathbb{R})$, $\mathfrak{X}(S^1)$, or $\mathcal{S}(\mathbb{R})\partial_x$. For completeness' sake we also consider $\alpha,\beta:\mathbb{R}\to\mathbb{R}$ with $\alpha(0)=0$, $\beta(0)=0$. Then we compute:

$$\begin{aligned}
&\operatorname{Ad}\begin{pmatrix}\varphi(t)\\ \alpha(t)\end{pmatrix}\begin{pmatrix}Y\\ \beta'(0)\end{pmatrix}=\partial_s|_0\begin{pmatrix}\varphi(t)\\ \alpha(t)\end{pmatrix}\begin{pmatrix}\psi(s)\\ \beta(s)\end{pmatrix}\begin{pmatrix}\varphi(t)^{-1}\\ -\alpha(t)\end{pmatrix}\\
&=\partial_s|_0\begin{pmatrix}\varphi(t)\circ\psi(s)\circ\varphi(t)^{-1}\\ \alpha(t)+\beta(s)+c(\varphi(t),\psi(s))-\alpha(t)+c(\varphi(t)\circ\psi(s),\varphi(t)^{-1})\end{pmatrix}\\
&=\begin{pmatrix}\varphi(t)_*Y=\operatorname{Ad}(\varphi(t))Y\\ \beta_t(0)+\partial_s|_0c(\varphi(t),\psi(s))+\partial_s|_0c(\varphi(t)\circ\psi(s),\varphi(t)^{-1})\end{pmatrix} \qquad (1)\\
&\Big[\begin{pmatrix}X\\ \alpha_t(0)\end{pmatrix},\begin{pmatrix}Y\\ \beta_t(0)\end{pmatrix}\Big]=\\
&=\partial_t|0\begin{pmatrix}(\operatorname{Fl}^X_t)_*Y=\operatorname{Ad}(\varphi(t))Y\\ \beta_t(0)+\partial_s|_0c(\varphi(t),\psi(s))+\partial_s|_0c(\varphi(t)\circ\psi(s),\varphi(t)^{-1})\end{pmatrix}\\
&=\begin{pmatrix}-[X,Y]\\ \partial_t|_0\partial_s|_0c(\varphi(t),\psi(s))+\partial_t|_0\partial_s|_0c(\varphi(t)\circ\psi(s),\varphi(t)^{-1})\end{pmatrix} \qquad (2)
\end{aligned}$$

Now we differentiate the Bott cocycle, where sometimes $f'=\partial_x f$:

$$\begin{aligned}
2\partial_s|_0c(\varphi(t),\psi(s))&=\partial_s|_0\int\log(\varphi(t)'\circ\psi(s))\,d\log(\psi(s)')\\
&=\int\frac{(\varphi(t)''\circ\psi(0))Y}{\varphi(t)'\circ\psi(0)}\,d\log(\underbrace{\psi(0)'}_{=1})+\int\log(\varphi(t)')\,dY'\\
&=\int\log(\varphi(t)')Y''\,dx\\
2\partial_t|_0\partial_s|_0c(\varphi(t),\psi(s))&=\partial_t|_0\int\log(\varphi(t)')Y''\,dx=\int\frac{X'Y''}{\varphi(0)'}dx=\int X'Y''dx.
\end{aligned}$$

For the second term we first check:

$$(\varphi^{-1})_x=\frac{1}{\varphi_x\circ\varphi^{-1}},\qquad(\varphi^{-1})_{xx}=-\frac{\varphi_{xx}\circ\varphi^{-1}}{(\varphi_x\circ\varphi^{-1})^3},$$

$$\varphi^{-1}(x) = y, \qquad \frac{1}{\varphi_x \circ \varphi^{-1}} dx = dy$$
$$d\log((\varphi^{-1})_x) = -\frac{\varphi'' \circ \varphi^{-1}}{(\varphi' \circ \varphi^{-1})^2} dx = -\frac{\varphi''}{\varphi'} dy$$

and continue to compute

$$2\partial_s|_0 c(\varphi(t) \circ \psi(s), \varphi(t)^{-1}) = \partial_s|_0 \int \log((\varphi(t) \circ \psi(s))_x \circ \varphi(t)^{-1})\, d\log(\varphi(t)_x^{-1})$$
$$= \int \frac{(\varphi(t)'' \circ \varphi(t)^{-1})(Y \circ \varphi(t)^{-1}) + (\varphi(t)' \circ \varphi(t)^{-1})(Y' \circ \varphi(t)^{-1})}{(\varphi(t)' \circ \varphi(t)^{-1})(\psi(0)' \circ \varphi(t)^{-1})}\, d\log(\varphi(t)_x^{-1})$$
$$= -\int \frac{(\varphi(t)'')^2\, Y + \varphi(t)'\varphi(t)''\, Y'}{(\varphi(t)')^2}\, dy$$
$$2\partial_t|_0 \partial_s|_0 c(\varphi(t) \circ \psi(s), \varphi(t)^{-1}) = -\partial_t|_0 \int \frac{(\varphi(t)'')^2\, Y + \varphi(t)'\varphi(t)''\, Y'}{(\varphi(t)')^2}\, dy$$
$$= -\int \frac{0 + 0 + \varphi(0)' X''\, Y' - 0}{(\varphi(0)' = 1)^4}\, dy$$
$$= -\int X'' Y'\, dy = \int X' Y''\, dx.$$

Finally we get from (2):

$$\left[\begin{pmatrix} X \\ a \end{pmatrix}, \begin{pmatrix} Y \\ b \end{pmatrix}\right] = \begin{pmatrix} -[X,Y] \\ \omega(X,Y) \end{pmatrix} = \begin{pmatrix} X'Y - XY' \\ \omega(X,Y) \end{pmatrix} \tag{3}$$

where

$$\omega(X,Y) = \omega(X)Y = \int X' dY' = \int X'Y'' dx = \tfrac{1}{2} \int \det \begin{pmatrix} X' & Y' \\ X'' & Y'' \end{pmatrix} dx,$$

is the *Gelfand-Fuchs Lie algebra cocycle* $\omega : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, which is a bounded skew-symmetric bilinear mapping satisfying the cocycle condition

$$\omega([X,Y],Z) + \omega([Y,Z],X) + \omega([Z,X],Y) = 0.$$

It is a generator of the 1-dimensional bounded Chevalley cohomology $H^2(\mathfrak{g},\mathbb{R})$ for any of the Lie algebras $\mathfrak{g} = \mathfrak{X}(S^1)$, $\mathfrak{X}_c(\mathbb{R})$, or $\mathcal{S}(\mathbb{R})\partial_x$. The Lie algebra of the Virasoro-Bott Lie group is thus the central extension $\mathbb{R} \times_\omega \mathfrak{g}$ of $\mathfrak{g}$ induced by this cocycle. We have $H^2(\mathfrak{X}_c(M),\mathbb{R}) = 0$ for each finite dimensional manifold of dimension ≥ 2 (see [21]), which blocks the way to find a higher dimensional analog of the Korteweg – de Vries equation in a way similar to that sketched below.

For further use we also note the expression for the adjoint action on the Virasoro-Bott groups which we computed along the way. For the integral in the central term in (1) we have:

$$\frac{1}{2}\int\Big(\log(\varphi')Y'' - \frac{(\varphi'')^2Y+\varphi'\varphi''Y'}{(\varphi')^2}\Big)\,dx = \frac{1}{2}\int\Big(-2\frac{\varphi''}{\varphi'}Y' - \Big(\frac{\varphi''}{\varphi'}\Big)^2Y\Big)\,dx =$$
$$= \int\Big(\Big(\frac{\varphi''}{\varphi'}\Big)' - \frac{1}{2}\Big(\frac{\varphi''}{\varphi'}\Big)^2\Big)Y\,dx = \int S(\varphi)Y\,dx,$$

where a new character appears on stage, the *Schwartzian derivative*:

$$S(\varphi) = \Big(\frac{\varphi''}{\varphi'}\Big)' - \frac{1}{2}\Big(\frac{\varphi''}{\varphi'}\Big)^2 = \frac{\varphi'''}{\varphi'} - \frac{3}{2}\Big(\frac{\varphi''}{\varphi'}\Big)^2 = \log(\varphi')'' - \frac{1}{2}(\log(\varphi')')^2 \tag{4}$$

which measures the deviation of φ from being a Moebius transformation:

$$S(\varphi)=0 \iff \varphi(x) = \frac{ax+b}{cx+d} \text{ for } \begin{pmatrix} a & b\\ c & d\end{pmatrix} \in SL(2,\mathbb{R}).$$

Indeed, $S(\varphi)=0$ if and only if $g = \log(\varphi')' = \frac{\varphi''}{\varphi'}$ satisfies the differential equation $g' = g^2/2$, so that $\frac{2\,dg}{g^2} = dx$ or $\frac{-2}{g} = x + \frac{d}{c}$ which means $\log(\varphi')'(x) = g(x) = \frac{-2}{x+d/c}$ or again $\log(\varphi'(x)) = \int \frac{-2dx}{x+d/c} = -2\log(x + d/c) - 2\log(c) = \log(\frac{1}{(cx+d)^2})$. Therefore, $\varphi'(x) = \frac{1}{(cx+d)^2} = \partial_x \frac{ax+b}{cx+d}$.

For completeness' sake, let us note here the Schwartzian derivative of a composition and an inverse (which follow since the adjoint action (5) below is an action):

$$S(\varphi\circ\psi) = (S(\varphi)\circ\psi)(\psi')^2 + S(\psi), \quad S(\varphi^{-1}) = -\frac{S(\varphi)}{(\varphi')^2}\circ\varphi^{-1}$$

So finally, the adjoint action is given by:

$$\mathrm{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}\begin{pmatrix}Y\\ b\end{pmatrix} = \begin{pmatrix}\mathrm{Ad}(\varphi)Y = \varphi_*Y = T\varphi\circ Y\circ\varphi^{-1}\\ b + \int S(\varphi)Y\,dx\end{pmatrix} \tag{5}$$

8.3. H^0-Geodesics on the Virasoro-Bott groups

We shall use the L^2-inner product on $\mathbb{R}\times_\omega\mathfrak{g}$, where $\mathfrak{g} = \mathfrak{X}(S^1), \mathfrak{X}_c(\mathbb{R}), \mathcal{S}(\mathbb{R})\partial_x$:

$$\Big\langle\begin{pmatrix}X\\ a\end{pmatrix},\begin{pmatrix}Y\\ b\end{pmatrix}\Big\rangle_0 := \int XY\,dx + ab. \tag{1}$$

Integrating by parts we get

$$\Big\langle \mathrm{ad}\begin{pmatrix}X\\ a\end{pmatrix}\begin{pmatrix}Y\\ b\end{pmatrix},\begin{pmatrix}Z\\ c\end{pmatrix}\Big\rangle_0 = \Big\langle\begin{pmatrix}X'Y - XY'\\ \omega(X,Y)\end{pmatrix},\begin{pmatrix}Z\\ c\end{pmatrix}\Big\rangle_0$$
$$= \int (X'YZ - XY'Z + cX'Y'')\,dx$$

$$= \int (2X'Z + XZ' + cX''')Y\,dx$$

$$= \left\langle \begin{pmatrix} Y \\ b \end{pmatrix}, \operatorname{ad}\begin{pmatrix} X \\ a \end{pmatrix}^{\top} \begin{pmatrix} Z \\ c \end{pmatrix} \right\rangle_0, \quad \text{where}$$

$$\operatorname{ad}\begin{pmatrix} X \\ a \end{pmatrix}^{\top} \begin{pmatrix} Z \\ c \end{pmatrix} = \begin{pmatrix} 2X'Z + XZ' + cX''' \\ 0 \end{pmatrix}.$$

Using matrix notation we get therefore (where $\partial := \partial_x$)

$$\operatorname{ad}\begin{pmatrix} X \\ a \end{pmatrix} = \begin{pmatrix} X' - X\partial & 0 \\ \omega(X) & 0 \end{pmatrix}$$

$$\operatorname{ad}\begin{pmatrix} X \\ a \end{pmatrix}^{\top} = \begin{pmatrix} 2X' + X\partial & X''' \\ 0 & 0 \end{pmatrix}$$

$$\alpha\begin{pmatrix} X \\ a \end{pmatrix} = \operatorname{ad}\begin{pmatrix} \ \end{pmatrix}^{\top} \begin{pmatrix} X \\ a \end{pmatrix} = \begin{pmatrix} X' + 2X\partial + a\partial^3 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\operatorname{ad}\begin{pmatrix} X \\ a \end{pmatrix}^{\top} + \operatorname{ad}\begin{pmatrix} X \\ a \end{pmatrix} = \begin{pmatrix} 3X' & X''' \\ \omega(X) & 0 \end{pmatrix}$$

$$\operatorname{ad}\begin{pmatrix} X \\ a \end{pmatrix}^{\top} - \operatorname{ad}\begin{pmatrix} X \\ a \end{pmatrix} = \begin{pmatrix} X' + 2X\partial & X''' \\ -\omega(X) & 0 \end{pmatrix}.$$

Formula (3.2.2) gives the H^0 geodesic equation on the Virasoro-Bott group:

$$\begin{pmatrix} u_t \\ a_t \end{pmatrix} = -\operatorname{ad}\begin{pmatrix} u \\ a \end{pmatrix}^{\top} \begin{pmatrix} u \\ a \end{pmatrix} = \begin{pmatrix} -3u_x u - a u_{xxx} \\ 0 \end{pmatrix} \quad \text{where} \tag{2}$$

$$\begin{pmatrix} u(t) \\ a(t) \end{pmatrix} = \partial_s \begin{pmatrix} \varphi(s) \\ \alpha(s) \end{pmatrix} . \begin{pmatrix} \varphi(t)^{-1} \\ -\alpha(t) \end{pmatrix}\Big|_{s=t}$$

$$= \partial_s \begin{pmatrix} \varphi(s) \circ \varphi(t)^{-1} \\ \alpha(s) - \alpha(t) + c(\varphi(s), \varphi(t)^{-1}) \end{pmatrix}\Big|_{s=t}$$

$$= \begin{pmatrix} \varphi_t \circ \varphi^{-1} \\ \alpha_t - \int \frac{\varphi_{tx}\varphi_{xx}}{2\varphi_x^2}\,dx \end{pmatrix}$$

since we have

$$2\partial_s c(\varphi(s), \varphi(t)^{-1})|_{s=t} = \partial_s \int \log(\varphi(s)' \circ \varphi(t)^{-1})\, d\log((\varphi(t)^{-1})')|_{s=t}$$

$$= \int \frac{\varphi_t(t)' \circ \varphi(t)^{-1}}{\varphi(t)' \circ \varphi(t)^{-1}} \left(-\frac{\varphi(t)'' \circ \varphi(t)^{-1}}{(\varphi(t)' \circ \varphi(t)^{-1})^2} \right) dx \quad \text{by (8.2)}$$

$$= -\int \frac{\varphi_t' \varphi''}{(\varphi')^2}\,dy = -\int \frac{\varphi_{tx}\varphi_{xx}}{\varphi_x^2}\,dx.$$

Thus a is a constant in time and the geodesic equation is hence the *Korteweg-de Vries equation*

$$u_t + 3u_x u + a u_{xxx} = 0. \tag{3}$$

with its natural companions

$$\varphi_t = u \circ \varphi, \qquad \alpha_t = a + \int \frac{\varphi_{tx}\varphi_{xx}}{2\varphi_x^2} dx.$$

It is the periodic equation, if we work on S^1.

The derivation above is direct and does not use the Euler-Poincaré equations; for a derivation of the Korteweg-de Vries equation from this point of view see [34], section 13.8.

Let us compute the invariant momentum mapping from (4.3.2). First we need the transpose of the adjoint action (8.2.5):

$$\begin{aligned}
\left\langle \operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^\top \begin{pmatrix}Y\\ b\end{pmatrix}, \begin{pmatrix}Z\\ c\end{pmatrix}\right\rangle_0 &= \left\langle \begin{pmatrix}Y\\ b\end{pmatrix}, \operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}\begin{pmatrix}Z\\ c\end{pmatrix}\right\rangle_0 \\
&= \left\langle \begin{pmatrix}Y\\ b\end{pmatrix}, \begin{pmatrix}\varphi_* Z\\ c + \int S(\varphi) Z\, dx\end{pmatrix}\right\rangle_0 \\
&= \int Y((\varphi'\circ\varphi^{-1})(Z\circ\varphi^{-1})\, dx + bc + \int bS(\varphi) Z\, dx \\
&= \int ((Y\circ\varphi)(\varphi')^2 + bS(\varphi))Z\, dx + bc \\
\operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^\top \begin{pmatrix}Y\\ b\end{pmatrix} &= \begin{pmatrix}(Y\circ\varphi)(\varphi')^2 + bS(\varphi)\\ b\end{pmatrix}.
\end{aligned}$$

Thus the invariant momentum mapping (4.3.2) turns out as

$$\bar J\left(\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}, \begin{pmatrix}Y\\ b\end{pmatrix}\right) = \operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^\top \begin{pmatrix}Y\\ b\end{pmatrix}\begin{pmatrix}(Y\circ\varphi)(\varphi')^2 + bS(\varphi)\\ b\end{pmatrix}. \tag{4}$$

Along a geodesic $t \mapsto g(t,\ \ ) = \binom{\varphi(t,\ \ )}{\alpha(t)}$, according to (3) and (4.3), the momentum

$$\bar J\left(\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}, \begin{pmatrix}u = \varphi_t\circ\varphi^{-1}\\ a\end{pmatrix}\right) = \begin{pmatrix}(u\circ\varphi)\varphi_x^2 + aS(\varphi)\\ a\end{pmatrix} = \begin{pmatrix}\varphi_t\varphi_x^2 + aS(\varphi)\\ a\end{pmatrix} \tag{5}$$

is constant in t.

8.4. The curvature

The computation of the curvature at the identity element has been done independently by [41] and Misiolek [42]. Here we proceed with a completely general computation that takes advantage of the formalism introduced so far. Inserting the matrices of differential- and integral operators $\mathrm{ad}\binom{X}{a}^{\top}$, $\alpha\binom{X}{a}$, and $\mathrm{ad}\binom{X}{a}$ etc. given above into formula (3.4.2) and recalling that the matrix is applied to vectors of the form $\binom{Z}{c}$, where c is a constant, we see that $4\mathcal{R}\left(\binom{X_1}{a_1},\binom{X_2}{a_2}\right)$ is the following 2×2-matrix whose entries are differential- and integral operators:

$$\begin{pmatrix} \begin{array}{l} 4(X_1X_2''-X_1''X_2)+2(a_1X_2^{(4)}-a_2X_1^{(4)}) \\ +(8(X_1X_2'-X_1'X_2)+10(a_1X_2'''-a_2X_1'''))\partial \\ +18(a_1X_2''-a_2X_1'')\partial^2 \\ +(12(a_1X_2'-a_2X_1')+2\omega(X_1,X_2))\partial^3 \\ -X_1'''\omega(X_2)+X_2'''\omega(X_1) \end{array} & \begin{array}{l} 2(X_1'''X_2'-X_1'X_2''') \\ +2(X_1X_2^{(4)}-X_1^{(4)}X_2) \\ +(a_1X_2^{(6)}-a_2X_1^{(6)}) \end{array} \\ \begin{array}{l} \omega(X_2)(4X_1'+2X_1\partial+a_1\partial^3) \\ -\omega(X_1)(4X_2'+2X_2\partial+a_2\partial^3) \end{array} & 0 \end{pmatrix}$$

Therefore, $4\mathcal{R}\left(\binom{X_1}{a_1},\binom{X_2}{a_2}\right)\binom{X_3}{a_3}$ has the following expression

$$\begin{pmatrix} \begin{array}{l} 4(X_1X_2''-X_1''X_2)X_3+2(a_1X_2^{(4)}-a_2X_1^{(4)})X_3 \\ +\left(8(X_1X_2'-X_1'X_2)+10(a_1X_2'''-a_2X_1''')\right)X_3' \\ +18(a_1X_2''-a_2X_1'')X_3''+12(a_1X_2'-a_2X_1')X_3''' \\ +2X_3'''\int X_1'X_2''dx-X_1'''\int X_2'X_3''dx+X_2'''\int X_1'X_3''dx \\ +2a_3(X_1'''X_2'-X_1'X_2''')+2a_3(X_1X_2^{(4)}-X_1^{(4)}X_2)+a_3(a_1X_2^{(6)}-a_2X_1^{(6)}) \end{array} \\ \begin{array}{l} \int X_3'''(a_1X_2'''-a_2X_1''')dx \\ +\int 2X_3'(X_1X_2'''-X_1'''X_2-2X_1'X_2''+2X_1''X_2')dx \end{array} \end{pmatrix}$$

which coincides with formula (2.3) in Misiolek [42]. This in turn leads to the following expression for the sectional curvature

$$\left\langle 4\mathcal{R}\left(\binom{X_1}{a_1},\binom{X_2}{a_2}\right)\binom{X_1}{a_1},\binom{X_2}{a_2}\right\rangle_0 =$$
$$=\int\Big(4(X_1X_2''-X_1''X_2)X_1X_2+8(X_1X_2'-X_1'X_2)X_1'X_2$$

$$
\begin{aligned}
&+ 2(a_1X_2^{(4)} - a_2X_1^{(4)})X_1X_2 + 10(a_1X_2''' - a_2X_1''')X_1'X_2\\
&+ 18(a_1X_2'' - a_2X_1'')X_1''X_2\\
&+ 12(a_1X_2' - a_2X_1')X_1'''X_2 + 2\omega(X_1,X_2)X_1'''X_2\\
&- X_1'''\omega(X_2,X_1)X_2 + X_2'''\omega(X_1,X_1)X_2\\
&+ 2(X_1'''X_2' - X_1'X_2''')a_1X_2\\
&+ 2(X_1X_2^{(4)} - X_1^{(4)}X_2)a_1X_2\\
&+ (a_1X_2^{(6)} - a_2X_1^{(6)})a_1X_2\\
&+ (4X_1'X_1X_2''' + 2X_1X_1'X_2''' + a_1X_1'''X_2'''\\
&\qquad - 4X_2'X_1X_1''' - 2X_2X_1'X_1''' - a_2X_1'''X_1''')a_2\Big)\,dx\\
= &\int\Big(- 4[X_1,X_2]^2 + 4(a_1X_2 - a_2X_1)(X_1X_2^{(4)} - X_1'X_2''' + X_1'''X_2' - X_1^{(4)}X_2)\\
&- (X_2''')^2a_1^2 + 2X_1'''X_2'''a_1a_2 - (X_1''')^2a_2^2\Big)\,dx\\
&+3\omega(X_1,X_2)^2.
\end{aligned}
$$

This formula shows that the sign of the sectional curvature is not constant. Indeed, choosing $h_1(x) = \sin x, h_2(x) = \cos x$ we get $-\pi(8+a_1^2+a_2^2-3\pi)$ which can be positive and negative by choosing the constants a_1, a_2 judiciously.

8.5. Jacobi fields

A Jacobi field $y = \binom{y}{b}$ along a geodesic with velocity field $\binom{u}{a}$ is a solution of the partial differential equation (3.5.1) which in our case looks as follows.

$$
\begin{aligned}
\begin{pmatrix} y_{tt}\\ b_{tt}\end{pmatrix} &= \left[\operatorname{ad}\begin{pmatrix} y\\ b\end{pmatrix}^\top + \operatorname{ad}\begin{pmatrix} y\\ b\end{pmatrix}, \operatorname{ad}\begin{pmatrix} u\\ a\end{pmatrix}^\top\right]\begin{pmatrix} u\\ a\end{pmatrix}\\
&\quad - \operatorname{ad}\begin{pmatrix} u\\ a\end{pmatrix}^\top\begin{pmatrix} y_t\\ b_t\end{pmatrix} - \alpha\begin{pmatrix} u\\ a\end{pmatrix}\begin{pmatrix} y_t\\ b_t\end{pmatrix} + \operatorname{ad}\begin{pmatrix} u\\ a\end{pmatrix}\begin{pmatrix} y_t\\ b_t\end{pmatrix}\\
&= \left[\begin{pmatrix} 3y_x & y_{xxx}\\ \omega(y) & 0\end{pmatrix}, \begin{pmatrix} 2u_x + u\partial_x & u_{xxx}\\ 0 & 0\end{pmatrix}\right]\begin{pmatrix} u\\ a\end{pmatrix}\\
&\quad + \begin{pmatrix} -2u_x - 4u\partial_x - a\partial_x^3 & -u_{xxx}\\ \omega(u) & 0\end{pmatrix}\begin{pmatrix} y_t\\ b_t\end{pmatrix},
\end{aligned}
$$

which leads to

$$
y_{tt} = -u(4y_{tx} + 3uy_{xx} + ay_{xxxx}) - u_x(2y_t + 2ay_{xxx}) \tag{1}
$$

$$
\begin{aligned}
&- u_{xxx}(b_t + \omega(y,u) - 3ay_x) - ay_{txxx}, \\
b_{tt} &= \omega(u, y_t) + \omega(y, 3u_x u) + \omega(y, au_{xxx}).
\end{aligned} \tag{2}
$$

Equation (2) is equivalent to:

$$
b_{tt} = \int (-y_{txxx}u + y_{xxx}(3u_x u + au_{xxx}))dx. \tag{2'}
$$

Next, let us show that the integral term in equation (1) is constant:

$$
b_t + \omega(y,u) = b_t + \int y_{xxx}u\,dx =: B_1. \tag{3}
$$

Indeed its t-derivative along the geodesic for u (that is, u satisfies the Korteweg-de Vries equation) coincides with (2′):

$$
b_{tt} + \int (y_{txxx}u + y_{xxx}u_t)\,dx = b_{tt} + \int (y_{txxx}u + y_{xxx}(-3u_x u - au_{xxx}))\,dx = 0.
$$

Thus $b(t)$ can be explicitly solved from (3) as

$$
b(t) = B_0 + B_1 t - \int_a^t \int y_{xxx}u\,dx\,dt. \tag{4}
$$

The first component of the Jacobi equation on the Virasoro-Bott group is a genuine partial differential equation. Thus the Jacobi equations are given by the following system:

$$
\begin{aligned}
y_{tt} &= -u(4y_{tx} + 3uy_{xx} + ay_{xxxx}) - u_x(2y_t + 2ay_{xxx}) \\
&\quad - u_{xxx}(B_1 - 3ay_x) - ay_{txxx}, \\
u_t &= -3u_x u - au_{xxx}, \\
a &= \text{ constant},
\end{aligned} \tag{5}
$$

where $u(t,x), y(t,x)$ are either smooth functions in $(t,x) \in I \times S^1$ or in $(t,x) \in I \times \mathbb{R}$, where I is an interval or $\mathbb{R}$, and where in the latter case u, y, y_t have compact support with respect to x.

Choosing $u = c \in \mathbb{R}$, a constant, these equations coincide with (3.1) in Misiolek [42] where it is shown by direct inspection that there are solutions of this equation which vanish at non-zero values of t, thereby concluding that there are conjugate points along geodesics emanating from the identity element of the Virasoro-Bott group on S^1.

8.6. The weak symplectic structure on the space of Jacobi fields on the Virasoro Lie algebra

Since the Korteweg - de Vries equation has local solutions depending smoothly on the initial conditions (and global solutions if $a \neq 0$), we expect that the space of all Jacobi fields exists and is isomorphic to the space of all initial data $(\mathbb{R} \times_\omega \mathfrak{X}(S^1)) \times (\mathbb{R} \times_\omega \mathfrak{X}(S^1))$. The weak symplectic structure is given in section (3.7):

$$\begin{aligned}
\omega\left(\begin{pmatrix} y \\ b\end{pmatrix}, \begin{pmatrix} z \\ c\end{pmatrix}\right) &= \left\langle \begin{pmatrix} y \\ b\end{pmatrix}, \begin{pmatrix} z_t \\ c_t\end{pmatrix}\right\rangle_0 - \left\langle \begin{pmatrix} y_t \\ b_t\end{pmatrix}, \begin{pmatrix} z \\ c\end{pmatrix}\right\rangle_0 + \left\langle \left[\begin{pmatrix} u \\ a\end{pmatrix}, \begin{pmatrix} y \\ b\end{pmatrix}\right], \begin{pmatrix} z \\ c\end{pmatrix}\right\rangle_0 \\
&\quad - \left\langle \begin{pmatrix} y \\ b\end{pmatrix}, \left[\begin{pmatrix} u \\ a\end{pmatrix}, \begin{pmatrix} z \\ b\end{pmatrix}\right]\right\rangle_0 - \left\langle \left[\begin{pmatrix} y \\ b\end{pmatrix}, \begin{pmatrix} z \\ c\end{pmatrix}\right], \begin{pmatrix} u \\ a\end{pmatrix}\right\rangle_0 \\
&= \int (yz_t - y_t z + 2u(yz_x - y_x z))\, dx \\
&\quad + b(c_t + \omega(z,u)) - c(b_t + \omega(y,u)) - a\omega(y,z) \\
&= \int (yz_t - y_t z + 2u(yz_x - y_x z))\, dx \qquad (1) \\
&\quad + bC_1 - cB_1 - a\int y' z''\, dx, \qquad (1)
\end{aligned}$$

where the constant C_1 relates to c as B_1 does to b, see (8.5.3) and (8.5.4).

8.7. The geodesics of the H^k-metric on the Virasoro group

We shall use the H^k-inner product on $\mathbb{R} \times_\omega \mathfrak{g}$, where $\mathfrak{g}$ is any of the Lie algebras $\mathfrak{X}(S^1)$ or $\mathfrak{X}_{\mathcal{S}}(\mathbb{R}) = \mathcal{S}(\mathbb{R})\partial_x$. The Lie algebra $\mathfrak{X}_c(\mathbb{R})$ does not work here any more since $A_k = \sum_{j=0}^k (-1)^j \partial_x^{2j}$ is no longer a linear isomorphism here.

$$\begin{aligned}
\left\langle \begin{pmatrix} X \\ a\end{pmatrix}, \begin{pmatrix} Y \\ b\end{pmatrix}\right\rangle_k &:= \int (XY + X'Y' + \dots + X^{(k)}Y^{(k)})\, dx + ab \qquad (1) \\
&= \int A_k(X) Y\, dx + ab = \int X A_k(Y)\, dx + ab, \\
&\text{where } A_k = \sum_{i=0}^k (-1)^i \partial_x^{2i} \text{ as in (7.3.1).}
\end{aligned}$$

Integrating by parts we get

$$
\begin{aligned}
\left\langle \operatorname{ad}\begin{pmatrix}X\\a\end{pmatrix}\begin{pmatrix}Y\\b\end{pmatrix}, \begin{pmatrix}Z\\c\end{pmatrix}\right\rangle_k &= \left\langle \begin{pmatrix}X'Y-XY'\\ \omega(X,Y)\end{pmatrix}, \begin{pmatrix}Z\\c\end{pmatrix}\right\rangle_k \\
&= \int (X'YA_k(Z) - XY'A_k(Z) + cX'Y'')\,dx \\
&= \int (2X'YA_k(Z) + XYA_k(Z') + cX''')\,dx \\
&= \int YA_kA_k^{-1}(2X'A_k(Z) + XA_k(Z') + cX''')\,dx \\
&= \left\langle \begin{pmatrix}Y\\b\end{pmatrix}, \operatorname{ad}\begin{pmatrix}X\\a\end{pmatrix}^\top \begin{pmatrix}Z\\c\end{pmatrix}\right\rangle_0, \quad \text{where}
\end{aligned}
$$

$$
\operatorname{ad}\begin{pmatrix}X\\a\end{pmatrix}^\top \begin{pmatrix}Z\\c\end{pmatrix} = \begin{pmatrix}A_k^{-1}(2X'A_k(Z) + XA_k(Z') + cX''')\\ 0\end{pmatrix}. \tag{2}
$$

Using matrix notation we get therefore (where $\partial := \partial_x$)

$$
\begin{aligned}
\operatorname{ad}\begin{pmatrix}X\\a\end{pmatrix} &= \begin{pmatrix}X'-X\partial & 0\\ \omega(X) & 0\end{pmatrix}\\
\operatorname{ad}\begin{pmatrix}X\\a\end{pmatrix}^\top &= \begin{pmatrix}A_k^{-1}.(2X'.A_k + XA_k.\partial_x) & A_k^{-1}(X''')\\ 0 & 0\end{pmatrix}\\
\alpha\begin{pmatrix}X\\a\end{pmatrix} = \operatorname{ad}\begin{pmatrix} \quad \end{pmatrix}^\top\begin{pmatrix}X\\a\end{pmatrix} &= \begin{pmatrix}A_k^{-1}.(A_k(X') + 2A_k(X)\partial_x + a\partial^3) & 0\\ 0 & 0\end{pmatrix}.
\end{aligned}
$$

Formula (3.2.2) gives the geodesic equation on the Virasoro-Bott group:

$$
\begin{pmatrix}u_t\\a_t\end{pmatrix} = -\operatorname{ad}\begin{pmatrix}u\\a\end{pmatrix}^\top\begin{pmatrix}u\\a\end{pmatrix} = \begin{pmatrix}-A_k^{-1}(2u_xA_k(u) + uA_k(u_x) + au_{xxx})\\ 0\end{pmatrix}, \tag{3}
$$

$$
\text{where } \begin{pmatrix}u(t)\\a(t)\end{pmatrix} = \begin{pmatrix}\varphi_t\circ\varphi^{-1}\\ \alpha_t - \int \frac{\varphi_{tx}\varphi_{xx}}{2\varphi_x^2}\,dx\end{pmatrix}
$$

as in (8.3.2) Thus a is a constant in time and the geodesic equation contains the equation from the Korteweg-de Vries hierarchy:

$$
A_k(u_t) = -2u_xA_k(u) - uA_k(u_x) - au_{xxx} \tag{4}
$$

For $k=0$ this gives the Korteweg–de Vries equation.

For $k=1$ we get the equation

$$
u_t - u_{txx} = -3uu_x + 2u_xu_{xx} + uu_{xxx} - au_{xxx},
$$

the *Camassa-Holm equation*, [13], [36]. See (7.3.4) for the dispersionfree version.

Let us compute the invariant momentum mapping from (4.3.2). First we need the transpose of the adjoint action (8.2.5):

$$
\begin{aligned}
\Big\langle \operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^{\top}\begin{pmatrix}Y\\ b\end{pmatrix},\begin{pmatrix}Z\\ c\end{pmatrix}\Big\rangle_k &= \Big\langle \begin{pmatrix}Y\\ b\end{pmatrix},\operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}\begin{pmatrix}Z\\ c\end{pmatrix}\Big\rangle_k\\
&= \Big\langle \begin{pmatrix}Y\\ b\end{pmatrix},\begin{pmatrix}\varphi_*Z\\ c+\int S(\varphi)Z\,dx\end{pmatrix}\Big\rangle_k\\
&= \int A_k(Y)(\varphi_*Z)\,dx + bc + \int bS(\varphi)Z\,dx\\
&= \int A_k(Y)((\varphi' Z)\circ\varphi^{-1})\,dx + bc + \int bS(\varphi)Z\,dx\\
&= \int (A_k(Y)\circ\varphi)(\varphi')^2 Z\,dy + bc + \int bS(\varphi)Z\,dx\\
&= \int ((A_k(Y)\circ\varphi)(\varphi')^2 + bS(\varphi))Z\,dx + bc\\
&= \int A_kA_k^{-1}((A_k(Y)\circ\varphi)(\varphi')^2 + bS(\varphi))Z\,dx + bc\\
&= \Big\langle \begin{pmatrix}A_k^{-1}((A_k(Y)\circ\varphi)(\varphi')^2 + bS(\varphi))\\ b\end{pmatrix},\begin{pmatrix}Z\\ c\end{pmatrix}\Big\rangle.\\
\operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^{\top}\begin{pmatrix}Y\\ b\end{pmatrix} &= \begin{pmatrix}A_k^{-1}((A_k(Y)\circ\varphi)(\varphi')^2 + bS(\varphi))\\ b\end{pmatrix}
\end{aligned}
$$

Thus the invariant momentum mapping (4.3.2) turns out as

$$
\bar J\left(\begin{pmatrix}\varphi\\ \alpha\end{pmatrix},\begin{pmatrix}Y\\ b\end{pmatrix}\right) = \operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^{\top}\begin{pmatrix}Y\\ b\end{pmatrix} = \begin{pmatrix}A_k^{-1}((A_k(Y)\circ\varphi)(\varphi')^2 + bS(\varphi))\\ b\end{pmatrix}. \tag{5}
$$

Along a geodesic $t\mapsto g(t,\quad) = \binom{\varphi(t,\quad)}{\alpha(t)}$, according to (4) and (4.3), the momentum

$$
\begin{aligned}
\bar J\left(\begin{pmatrix}\varphi\\ \alpha\end{pmatrix},\begin{pmatrix}u=\varphi_t\circ\varphi^{-1}\\ a\end{pmatrix}\right) &= \begin{pmatrix}A_k^{-1}\big(A_k(u)\circ\varphi)\varphi_x^2 + aS(\varphi)\big)\\ a\end{pmatrix}\\
&= \begin{pmatrix}A_k^{-1}\big((A_k(\varphi_t\circ\varphi^{-1})\circ\varphi)\varphi_x^2 + aS(\varphi)\big)\\ a\end{pmatrix}
\end{aligned} \tag{6}
$$

is constant in t, and thus also

$$
\tilde J(a,\varphi) := \big(A_k(\varphi_t\circ\varphi^{-1})\circ\varphi\big)\varphi_x^2 + aS(\varphi) \tag{7}
$$

is constant in t.

8.8. Theorem.
[12] *Let $k\ge 2$. There exists a HC^{2k+1}-open neighborhood V of $(\mathrm{Id},0)$ in the space $(S^1\times_c \operatorname{Diff}(S^1))\times(\mathbb{R}\times_\omega \mathfrak{X}(S^1))$ such that for each $(g_0,\alpha,u_0,a)\in V$ there exists a unique C^3 geodesic $g\in C^3((-2,2),S^1\times_c\operatorname{Diff}(S^1))$ for the right invariant H^k Riemann metric, starting at $g(0)=g_0$ in the direction*

$g_t(0) = u_0 \circ g_0 \in T_{g_0}\operatorname{Diff}(S^1)$. Moreover, the solution depends C^1 on the initial data $(g_0, u_0) \in V$.

The same result holds if we replace $S^1 \times_c \operatorname{Diff}(S^1)$ by $\mathbb{R} \times_c \operatorname{Diff}_{\mathcal{S}}(\mathbb{R})$ and $\mathfrak{X}(S^1)$ by $\mathcal{S}(\mathbb{R})\partial_x = \mathfrak{X}_{\mathcal{S}}(\mathbb{R})$.

In the following proof Diff, $\mathfrak{X}$, DiffHC^n, HC^n will mean either $\operatorname{Diff}(S^1)$, $\mathfrak{X}(S^1)$, $\operatorname{DiffHC}^n(S^1)$, $HC^n(S^1)$, or $\operatorname{Diff}_{\mathcal{S}}(\mathbb{R})$, $\mathfrak{X}_{\mathcal{S}}(\mathbb{R})$, $\operatorname{DiffHC}^n(\mathbb{R})$, $HC^n(\mathbb{R})$, respectively.

Proof. For $u \in HC^n$, $n \geq 2k+1$, we have as in the proof of (7.4)

$$A_k(uu_x) = uA_k(u_x) + \sum_{i=0}^{k}(-1)^i \sum_{j=1}^{2i} \tbinom{2i}{j}(\partial_x^j u)(\partial_x^{2i-j+1}u) =: u\,A_k(u_x) + B_k(u),$$

where $B_k : HC^n \to HC^{n-2k}$ is a bounded quadratic operator. Recall from (8.7.4) that we have to solve (where a is a real constant)

$$\begin{aligned}
u_t &= -A_k^{-1}\big(2u_x A_k(u) + uA_k(u_x) + au_{xxx}\big)\\
&= -A_k^{-1}\big(2u_x A_k(u) + A_k(uu_x) - B_k(u) + au_{xxx}\big)\\
&= -uu_x - A_k^{-1}\big(2u_x A_k(u) - B_k(u) + au_{xxx}\big)\\
&=: -uu_x + A_k^{-1}C_k(u,a),
\end{aligned}$$

where $u = g_t \circ g^{-1} \in \mathfrak{X}$, and where $C_k : HC^n \to HC^{n-2k}$ is a bounded polynomial operator, given by

$$\begin{aligned}
C_k(a,u) &= -2u_x A_k(u) + B_k(u) - au_{xxx}\\
&= -2u_x A_k(u) + \sum_{i=0}^{k}(-1)^i \sum_{j=1}^{2i} \tbinom{2i}{j}(\partial_x^j u)(\partial_x^{2i-j+1}u) - au_{xxx}.
\end{aligned}$$

Note that here we need $2k \geq 3$. In [43] this result was obtained for $k \geq 3/2$. We put

$$\left\{\begin{aligned}
g_t &=: v = u \circ g\\
v_t &= u_t \circ g + (u_x \circ g)g_t = u_t \circ g + (uu_x)\circ g = A_k^{-1}C_k(a,u)\circ g\\
&= A_k^{-1}C_k(a, v\circ g^{-1})\circ g =: \operatorname{pr}_2(D_k \circ E_k)(g,v), \qquad \text{where}
\end{aligned}\right. \tag{1}$$

$$E_k(a,g,v) = (g, C_k(a, v\circ g^{-1})\circ g), \qquad D_k(g,v) = (g, A_k^{-1}(v\circ g^{-1})\circ g).$$

Now consider the topological group and Banach manifold DiffHC^n.

Claim. The mapping $D_k : \operatorname{DiffHC}^n \times HC^{n-2k} \to \operatorname{DiffHC}^n \times HC^n$ is strongly C^1.

Let us assume that we have C^1-curves $s \mapsto g(s) \in \operatorname{DiffHC}^n$ and $s \mapsto v(s) \in HC^{n-2k}$. Then we have:

$$
\begin{aligned}
&\partial_s \operatorname{pr}_2 D_k(a, g(s), v(s)) = \partial_s A_k^{-1}(v\circ g^{-1})\circ g \\
&= A_k^{-1}(v_s\circ g^{-1})\circ g + A_k^{-1}\left((v_x\circ g^{-1})(-\frac{g_s\circ g^{-1}}{g_x\circ g^{-1}})\right)\circ g \\
&\qquad + (A_k^{-1}(v\circ g^{-1})_x\circ g)g_s \\
&A_k\Big(\Big(\partial_s \operatorname{pr}_2 D_k(a,g(s),v(s))\Big)\circ g^{-1}\Big) = \\
&= v_s\circ g^{-1} - (v\circ g^{-1})_x(g_s\circ g^{-1}) + A_k(A_k^{-1}(v\circ g^{-1})_x(g_s\circ g^{-1})) \\
&= v_s\circ g^{-1} - (v\circ g^{-1})_x(g_s\circ g^{-1}) + (v\circ g^{-1})_x(g_s\circ g^{-1}) + \\
&\quad + \sum_{i=0}^{k}\sum_{j=0}^{2i-1}\binom{2i}{j}(\partial_x^{j+1}(A_k^{-1}(v\circ g^{-1}))\partial_x^{2i-j}(g_s\circ g^{-1}) \in HC^{n-2k} \\
&\partial_s \operatorname{pr}_2 D_k(a,g(s),v(s)) = A_k^{-1}(v_s\circ g^{-1})\circ g \\
&\quad + \sum_{i=0}^{k}\sum_{j=0}^{2i-1}\binom{2i}{j}A_k^{-1}\Big((\partial_x^{j+1}(A_k^{-1}(v\circ g^{-1}))\partial_x^{2i-j}(g_s\circ g^{-1})\Big)\circ g
\end{aligned}
$$

and by (6.12) and (6.13) we can conclude that this is continuous in a, g, g_s, v, v_s jointly and Lipschitz in g_s and v_s. Thus D_k is strongly C^1.

Claim. The mapping $E_k : \operatorname{DiffHC}^n \times HC^n \to \operatorname{DiffHC}^n \times HC^{n-2k}$ is strongly C^1.

This can be proved in a similar way as the last claim.

By the two claims equation (1) can be viewed as the flow equation of a C^1-vector field on the Hilbert manifold $\operatorname{DiffHC}^n \times HC^n$. Here an existence and uniqueness theorem holds. Since $v = 0$ is a stationary point, there exists an open neighborhood W_n of $(\mathrm{Id}, 0)$ in $\operatorname{DiffHC}^n \times HC^n$ such that for each initial point $(g_0, v_0) \in W_n$ equation (1) has a unique solution $\mathrm{Fl}_t^n(g_0, v_0) = (g(t), v(t))$ defined and C^2 in $t \in (-2, 2)$. Note that $v(t) = g_t(t)$, thus $g(t)$ is even C^3 in t. Moreover, the solution depends C^1 on the initial data.

We start with the neighborhood

$$W_{2k+1} \subset \operatorname{DiffHC}^{2k+1} \times HC^{2k+1} \supset \operatorname{DiffHC}^n \times HC^n \qquad \text{for } n \ge 2k+1$$

and consider the neighborhood $V_n := W_{2k+1} \cap \operatorname{DiffHC}^n \times HC^n$ of $(\mathrm{Id}, 0)$.

Claim. For any initial point $(g_0, v_0) \in V_n$ the solution $\mathrm{Fl}_t^n(g_0, v_0) = (g(t), v(t))$ exists, is unique, is C^2 in $t \in (-2, 2)$, and depends C^1 on the initial point in V_n.

We use induction on $n \ge 2k+1$. For $n = 2k+1$ the claim holds since $V_{2k+1} = W_{2k+1}$. Let $(g_0, v_0) \in V_{2k+2}$ and let $\mathrm{Fl}_t^{2k+2}(g_0, v_0) = (\tilde g(t), \tilde v(t))$ be maximally defined for $t \in (t_1, t_2) \ni 0$. Suppose for contradiction that $t_2 < 2$. Since $(g_0, v_0) \in V_{2k+2} \subset V_{2k+1}$ the curve $\mathrm{Fl}_t^{2k+2}(g_0, v_0) = (\tilde g(t), \tilde v(t))$ solves (1) also in $\operatorname{DiffHC}^{2k+1} \times HC^{2k+1}$, thus $\mathrm{Fl}_t^{2k+2}(g_0, v_0) = (\tilde g(t), \tilde v(t)) = (g(t), v(t)) := \mathrm{Fl}_t^{2k+1}(g_0, v_0)$ for $t \in (t_1, t_2) \cap (-2, 2)$. By (7.3.6), the expression

$$\tilde J(t) = \tilde J(g, v, t) = g_x(t)^2 A_k(u(t))\circ g(t) = g_x(t)^t A_k(v(t)\circ g(t))\circ g(t) \qquad (2)$$

is constant in $t \in (-2, 2)$. Actually, since we used C^∞-theory for deriving this, one should check it again by differentiating. Since $u = g_t \circ g^{-1}$ we get the following (the exact formulas can be computed with the help of Faà di Bruno's formula (6.1):

$$\begin{aligned}
u_x &= (g_{tx} \circ g^{-1})(g^{-1})_x = \frac{g_{tx}}{g_x} \circ g^{-1} \\
\partial_x^2 u &= (\frac{\partial_x^2 g_t}{g_x^2} - g_{tx}\frac{\partial_x^2 g}{g_x^3}) \circ g^{-1} \\
\partial_x(g^{-1}) &= \frac{1}{g_x} \circ g^{-1} \\
\partial_x^2(g^{-1}) \circ g &= -\frac{\partial_x^2 g}{g_x^3} \\
\partial_x^{2k}(g^{-1}) \circ g &= -\frac{\partial_x^{2k} g}{g_x^{2k+1}} + \text{ lower order terms in } g \\
(\partial_x^{2k} u) \circ g &= \frac{\partial_x^{2k} g_t}{g_x^{2k}} - g_{tx}\frac{\partial_x^{2k} g}{g_x^{2k+1}} + \text{ lower order terms in } g, g_t = v.
\end{aligned}$$

Thus

$$(-1)^k g_x^{2k-1} \tilde{J}(t) = g_x \partial_x^{2k} g_t - g_{tx}\partial_x^{2k} g + \text{ lower order terms in } g, g_t = v.$$

Hence for each $t \in (-2, 2)$:

$$\begin{aligned}
g_x \partial_x^{2k} g_t - g_{tx}\partial_x^{2k} g &= (-1)^k g_x^2 \left(g_x^{2k-3} \tilde{J}(t) + P_k(g, v) \right), \text{ where} \\
P_k(g, v) &= \frac{Q_k(g, \partial_x g, \ldots, \partial_x^{2k-1} g, v, \partial_x v, \ldots, \partial_x^{2k-1} v)}{g_x^2}
\end{aligned}$$

for a polynomial Q_k. Since $\tilde{J}(t) = \tilde{J}(0)$ we obtain that

$$\left(\frac{\partial_x^{2k} g(t)}{g_x(t)} \right)_t = (-1)^k \left(g_x^{2k-3}(t)\tilde{J}(0) + P_k(g(t), v(t)) \right) \text{ for all } t \in (-2, 2).$$

This implies

$$\frac{\partial_x^{2k} g(t)}{g_x(t)} = \frac{\partial_x^{2k} g(0)}{g_x(0)} + (-1)^k \int_0^t \left(g_x^{2k-3}(s)\tilde{J}(0) + P_k(g(s), v(s)) \right) ds.$$

For $t \in (t_1, t_2)$ we have

$$\begin{aligned}
\partial_x^{2k} \tilde{g}(t) =& \frac{\partial_x^{2k} g_0}{\partial_x g_0} g_x(t) + \\
&+ (-1)^k g_x(t) \int_0^t \left(g_x^{2k-3}(s)\tilde{J}(0) + P_k(g(s), v(s)) \right) ds.
\end{aligned} \tag{3}$$

Since $(g_0, v_0) \in V_{2k+2}$ we have $\tilde{J}(0) = \tilde{J}(g_0, v_0, 0) \in HC^2$ by (2). Since $k \geq 1$, by (3) we see that $\partial_x^{2k}\tilde{g}(t) \in HC^2$. Moreover, since $t_2 < 2$, the limit $\lim_{t\to t_2-} \partial_x^{2k}\tilde{g}(t)$ exists in HC^2, so $\lim_{t\to t_2-} \tilde{g}(t)$ exists in HC^{2k+2}. As this limit equals $g(t_2)$, we conclude that $g(t_2) \in \mathrm{DiffHC}^{2k+2}$. Now $\tilde{v} = \tilde{g}_t$; so we may differentiate both sides of (3) in t and obtain similarly that $\lim_{t\to t_2-} \tilde{v}(t)$ exists in HC^{2k+2} and equals $v(t_2)$. But then we can prolong the flow line $(\tilde{g}, \tilde{v})$ in $\mathrm{DiffHC}^{2k+2} \times HC^{2k+2}$ beyond t_2, so (t_1, t_2) was not maximal.

By the same method we can iterate the induction. □

Appendix A. Smooth calculus beyond Banach spaces

The traditional differential calculus works well for finite dimensional vector spaces and for Banach spaces. For more general locally convex spaces we sketch here the convenient approach as explained in [20] and [30]. The main difficulty is that the composition of linear mappings stops to be jointly continuous at the level of Banach spaces, for any compatible topology. We use the notation of [30] and this is the main reference for the whole appendix. We list results in the order in which one can prove them, without proofs for which we refer to [30]. This should explain how to use these results. Later we also explain the fundamentals about regular infinite dimensional Lie groups.

A.1. Convenient vector spaces

Let E be a locally convex vector space. A curve $c : \mathbb{R} \to E$ is called *smooth* or C^∞ if all derivatives exist and are continuous - this is a concept without problems. Let $C^\infty(\mathbb{R}, E)$ be the space of smooth functions. It can be shown that $C^\infty(\mathbb{R}, E)$ does not depend on the locally convex topology of E, but only on its associated bornology (system of bounded sets).

E is said to be a *convenient vector space* if one of the following equivalent conditions is satisfied (called c^∞-completeness):

1. For any $c \in C^\infty(\mathbb{R}, E)$ the (Riemann-) integral $\int_0^1 c(t)dt$ exists in E.
2. A curve $c : \mathbb{R} \to E$ is smooth if and only if $\lambda \circ c$ is smooth for all $\lambda \in E'$, where E' is the dual consisting of all continuous linear functionals on E.
3. Any Mackey-Cauchy-sequence (i. e. $t_{nm}(x_n - x_m) \to 0$ for some $t_{nm} \to \infty$ in $\mathbb{R}$) converges in E. This is visibly a weak completeness requirement.

The final topology with respect to all smooth curves is called the c^∞-topology on E, which then is denoted by $c^\infty E$. For Fréchet spaces it coincides with the given locally convex topology, but on the space $\mathcal{D}$ of test functions with compact support on $\mathbb{R}$ it is strictly finer.

A.2. Smooth mappings

Let E and F be locally convex vector spaces, and let $U \subset E$ be c^∞-open. A mapping $f : U \to F$ is called *smooth* or C^∞, if $f \circ c \in C^\infty(\mathbb{R}, F)$ for all $c \in C^\infty(\mathbb{R}, U)$. *The main properties of smooth calculus are the following.*

1. *For mappings on Fréchet spaces this notion of smoothness coincides with all other reasonable definitions. Even on $\mathbb{R}^2$ this is non-trivial.*
2. *Multilinear mappings are smooth if and only if they are bounded.*
3. *If $f : E \supseteq U \to F$ is smooth then the derivative $df : U \times E \to F$ is smooth, and also $df : U \to L(E, F)$ is smooth where $L(E, F)$ denotes the space of all bounded linear mappings with the topology of uniform convergence on bounded subsets.*
4. *The chain rule holds.*
5. *The space $C^\infty(U, F)$ is again a convenient vector space where the structure is given by the obvious injection*

$$C^\infty(U, F) \to \prod_{c \in C^\infty(\mathbb{R}, U)} C^\infty(\mathbb{R}, F) \to \prod_{c \in C^\infty(\mathbb{R}, U), \lambda \in F'} C^\infty(\mathbb{R}, \mathbb{R}).$$

6. *The exponential law holds:*

$$C^\infty(U, C^\infty(V, G)) \cong C^\infty(U \times V, G)$$

   is a linear diffeomeorphism of convenient vector spaces. Note that this is the main assumption of variational calculus.
7. *A linear mapping $f : E \to C^\infty(V, G)$ is smooth (bounded) if and only if $E \xrightarrow{f} C^\infty(V, G) \xrightarrow{\mathrm{ev}_v} G$ is smooth for each $v \in V$. This is called the smooth uniform boundedness theorem and it is quite applicable.*

A.3. Theorem. [20], 4.1.19**.**
Let $c : \mathbb{R} \to E$ be a curve in a convenient vector space E. Let $\mathcal{V} \subset E'$ be a subset of bounded linear functionals such that the bornology of E has a basis of $\sigma(E, \mathcal{V})$-closed sets. Then the following are equivalent:

1. *c is smooth*
2. *There exist locally bounded curves $c^k : \mathbb{R} \to E$ such that $\ell \circ c$ is smooth $\mathbb{R} \to \mathbb{R}$ with $(\ell \circ c)^{(k)} = \ell \circ c^k$.*

If E is reflexive, then for any point separating subset $\mathcal{V} \subset E'$ the bornology of E has a basis of $\sigma(E, \mathcal{V})$-closed subsets, by [20], 4.1.23.

A.4. Counterexamples in infinite dimensions against common beliefs on ordinary differential equations

Let $E := s$ be the Fréchet space of rapidly decreasing sequences; note that by the theory of Fourier series we have $s = C^\infty(S^1,\mathbb{R})$. Consider the continuous linear operator $T : E \to E$ given by $T(x_0, x_1, x_2, \dots) := (0, 1^2x_1, 2^2x_2, 3^2x_3, \dots)$. The ordinary linear differential equation $x'(t) = T(x(t))$ with constant coefficients has no solution in s for certain initial values. By recursion one sees that the general solution should be given by

$$x_n(t) = \sum_{i=0}^{n} \left(\tfrac{n!}{i!}\right)^2 x_i(0)\frac{t^{n-i}}{(n-i)!}.$$

If the initial value is a finite sequence, say $x_n(0) = 0$ for $n > N$ and $x_N(0) \neq 0$, then

$$\begin{aligned}
x_n(t) &= \sum_{i=0}^{N} \left(\tfrac{n!}{i!}\right)^2 x_i(0)\frac{t^{n-i}}{(n-i)!}\\
&= \frac{(n!)^2}{(n-N)!}\; t^{n-N} \sum_{i=0}^{N} \left(\tfrac{1}{i!}\right)^2 x_i(0)\tfrac{(n-N)!}{(n-i)!}\; t^{N-i}\\
|x_n(t)| &\geq \frac{(n!)^2}{(n-N)!}\; |t|^{n-N} \left(|x_N(0)| \left(\tfrac{1}{N!}\right)^2 - \sum_{i=0}^{N-1} \left(\tfrac{1}{i!}\right)^2 |x_i(0)|\tfrac{(n-N)!}{(n-i)!}|t|^{N-i}\right)\\
&\geq \frac{(n!)^2}{(n-N)!}\; |t|^{n-N} \left(|x_N(0)| \left(\tfrac{1}{N!}\right)^2 - \sum_{i=0}^{N-1} \left(\tfrac{1}{i!}\right)^2 |x_i(0)||t|^{N-i}\right)
\end{aligned}$$

where the first factor does not lie in the space s of rapidly decreasing sequences and where the second factor is larger than $\varepsilon > 0$ for t small enough. So at least for a dense set of initial values this differential equation has no local solution.

This shows also, that the theorem of Frobenius is wrong, in the following sense: The vector field $x \mapsto T(x)$ generates a 1-dimensional subbundle E of the tangent bundle on the open subset $s \setminus \{0\}$. It is involutive since it is 1-dimensional. But through points representing finite sequences there exist no local integral submanifolds (M with $TM = E|M$). Namely, if c were a smooth nonconstant curve with $c'(t) = f(t).T(c(t))$ for some smooth function f, then $x(t) := c(h(t))$ would satisfy $x'(t) = T(x(t))$, where h is a solution of $h'(t) = 1/f(h(t))$.

As next example consider $E := \mathbb{R}^\mathbb{N}$ and the continuous linear operator $T : E \to E$ given by $T(x_0, x_1, \dots) := (x_1, x_2, \dots)$. The corresponding differential equation has solutions for every initial value $x(0)$, since the coordinates must satisfy the recusive relations $x_{k+1}(t) = x_k'(t)$ and hence any smooth functions $x_0 : \mathbb{R} \to \mathbb{R}$ gives rise to a solution $x(t) := (x_0^{(k)}(t))_k$ with initial value $x(0) = (x_0^{(k)}(0))_k$. So by Borel's theorem there exist solutions to this equation for any initial value and the difference of any two functions with same initial value is an arbitray infinite flat function. Thus the solutions are far from being unique. Note that $\mathbb{R}^{\mathbb{N}}$ is a topological direct summand in $C^\infty(\mathbb{R}, \mathbb{R})$ via the projection $f \mapsto (f(n))_n$, and hence the same situation occurs in $C^\infty(\mathbb{R}, \mathbb{R})$.

Let now $E := C^\infty(\mathbb{R}, \mathbb{R})$ and consider the continuous linear operator $T : E \to E$ given by $T(x) := x'$. Let $x : \mathbb{R} \to C^\infty(\mathbb{R}, \mathbb{R})$ be a solution of the equation $x'(t) = T(x(t))$. In terms of $\hat{x} : \mathbb{R}^2 \to \mathbb{R}$ this says $\frac{\partial}{\partial t}\hat{x}(t, s) = \frac{\partial}{\partial s}\hat{x}(t, s)$. Hence $r \mapsto \hat{x}(t - r, s + r)$ has vanishing derivative everywhere and so this function is constant, and in particular $x(t)(s) = \hat{x}(t, s) = \hat{x}(0, s + t) = x(0)(s + t)$. Thus we have a smooth solution x uniquely determined by the initial value $x(0) \in C^\infty(\mathbb{R}, \mathbb{R})$ which even describes a flow for the vector field T in the sense of (A.6) below. In general this solution is however not real-analytic, since for any $x(0) \in C^\infty(\mathbb{R}, \mathbb{R})$, which is not real-analytic in a neighborhood of a point s the composite $\mathrm{ev}_s \circ x = x(s + \quad)$ is not real-analytic around 0.

A.5. Manifolds and vector fields

In the sequel we shall use smooth manifolds M modelled on c^∞-open subsets of convenient vector spaces. Since we shall need it we also include some results on vector fields and their flows.

Consider vector fields $X_i \in C^\infty(TM)$ and $Y_i \in \Gamma(TN)$ for $i = 1, 2$, and a smooth mapping $f : M \to N$. If X_i and Y_i are f-related for $i = 1, 2$, i. e. $Tf \circ X_i = Y_i \circ f$, then also $[X_1, X_2]$ and $[Y_1, Y_2]$ are f-related.

In particular if $f : M \to N$ is a local diffeomorphism (so $(T_x f)^{-1}$ makes sense for each $x \in M$), then for $Y \in \Gamma(TN)$ a vector field $f^*Y \in \Gamma(TM)$ is defined by $(f^*Y)(x) = (T_x f)^{-1}.Y(f(x))$. The linear mapping $f^* : \Gamma(TN) \to \Gamma(TM)$ is then a Lie algebra homomorphism.

A.6. The flow of a vector field

Let $X \in \Gamma(TM)$ be a vector field. A *local flow* Fl^X for X is a smooth mapping $\mathrm{Fl}^X : M \times \mathbb{R} \supset U \to M$ defined on a c^∞-open neighborhood U of $M \times 0$ such that

1. $\frac{d}{dt}\mathrm{Fl}^X_t(x) = X(\mathrm{Fl}^X_t(x))$.
2. $\mathrm{Fl}^X_0(x) = x$ for all $x \in M$.
3. $U \cap (\{x\} \times \mathbb{R})$ is a connected open interval.

4. $\mathrm{Fl}^X_{t+s} = \mathrm{Fl}^X_t \circ \mathrm{Fl}^X_s$ holds in the following sense. If the right hand side exists then also the left hand side exists and we have equality. Moreover: If Fl^X_s exists, then the existence of both sides is equivalent and they are equal.

Let $X \in \Gamma(TM)$ be a vector field which admits a local flow Fl^X_t. Then for each integral curve c of X we have $c(t) = \mathrm{Fl}^X_t(c(0))$, thus there exists a unique maximal flow. Furthermore, X is Fl^X_t-related to itself, i. e. $T(\mathrm{Fl}^X_t) \circ X = X \circ \mathrm{Fl}^X_t$.

Let $X \in \Gamma(TM)$ and $Y \in \Gamma(TN)$ be f-related vector fields for a smooth mapping $f : M \to N$ which have local flows Fl^X and Fl^Y. Then we have $f \circ \mathrm{Fl}^X_t = \mathrm{Fl}^Y_t \circ f$, whenever both sides are defined.

Moreover, if f is a diffeomorphism we have $\mathrm{Fl}^{f^*Y}_t = f^{-1} \circ \mathrm{Fl}^Y_t \circ f$ in the following sense: If one side exists then also the other side exists, and they are equal.

For $f = Id_M$ this implies that if there exists a flow then there exists a unique maximal flow Fl^X_t.

A.7. The Lie derivative

There are situations where we do not know that the flow of X exists but where we will be able to produce the following assumption: Suppose that $\varphi : \mathbb{R} \times M \supset U \to M$ is a smooth mapping such that $(t, x) \mapsto (t, \varphi(t, x) = \varphi_t(x))$ is a diffeomorphism $U \to V$, where U and V are open neighborhoods of $\{0\} \times M$ in $\mathbb{R} \times M$, and such that $\varphi_0 = \mathrm{Id}_M$ and $\partial_t \varphi_t = X \in \Gamma(TM)$. Then again $\partial_t|_0 (\varphi_t)^* f = \partial_t|_0 f \circ \varphi_t = df \circ X = X(f)$.

In this situation we have for $Y \in \Gamma(TM)$, and for a k-form $\omega \in \Omega^k(M)$:

$$\partial_t|_0 (\varphi_t)^* Y = [X, Y],$$
$$\partial_t|_0 (\varphi_t)^* \omega = \mathcal{L}_X \omega.$$

Appendix B. Regular infinite dimensional Lie groups

B.1. Lie groups

A *Lie group* G is a smooth manifold modelled on c^∞-open subsets of a convenient vector space, and a group such that the multiplication $\mu : G \times G \to G$ and the inversion $\nu : G \to G$ are smooth. We shall use the following notation:
$\mu : G \times G \to G$, multiplication, $\mu(x, y) = x.y$.
$\mu_a : G \to G$, left translation, $\mu_a(x) = a.x$.

$\mu^a : G \to G$, right translation, $\mu^a(x) = x.a$.
$\nu : G \to G$, inversion, $\nu(x) = x^{-1}$.
$e \in G$, the unit element.

The tangent mapping $T_{(a,b)}\mu : T_aG \times T_bG \to T_{ab}G$ is given by

$$T_{(a,b)}\mu.(X_a, Y_b) = T_a(\mu^b).X_a + T_b(\mu_a).Y_b$$

and $T_a\nu : T_aG \to T_{a^{-1}}G$ is given by

$$T_a\nu = -T_e(\mu^{a^{-1}}).T_a(\mu_{a^{-1}}) = -T_e(\mu_{a^{-1}}).T_a(\mu^{a^{-1}}).$$

B.2. Invariant vector fields and Lie algebras

Let G be a (real) Lie group. A vector field ξ on G is called *left invariant*, if $\mu_a^*\xi = \xi$ for all $a \in G$, where $\mu_a^*\xi = T(\mu_{a^{-1}}) \circ \xi \circ \mu_a$. Since we have $\mu_a^*[\xi, \eta] = [\mu_a^*\xi, \mu_a^*\eta]$, the space $\mathfrak{X}_L(G)$ of all left invariant vector fields on G is closed under the Lie bracket, so it is a sub Lie algebra of $\mathfrak{X}(G)$. Any left invariant vector field ξ is uniquely determined by $\xi(e) \in T_eG$, since $\xi(a) = T_e(\mu_a).\xi(e)$. Thus the Lie algebra $\mathfrak{X}_L(G)$ of left invariant vector fields is linearly isomorphic to T_eG, and on T_eG the Lie bracket on $\mathfrak{X}_L(G)$ induces a Lie algebra structure, whose bracket is again denoted by $[\quad,\quad]$. This Lie algebra will be denoted as usual by $\mathfrak{g}$, sometimes by $\mathrm{Lie}(G)$.

We will also give a name to the isomorphism with the space of left invariant vector fields: $L : \mathfrak{g} \to \mathfrak{X}_L(G)$, $X \mapsto L_X$, where $L_X(a) = T_e\mu_a.X$. Thus $[X, Y] = [L_X, L_Y](e)$.

Similarly a vector field η on G is called *right invariant*, if $(\mu^a)^*\eta = \eta$ for all $a \in G$. If ξ is left invariant, then $\nu^*\xi$ is right invariant. The right invariant vector fields form a sub Lie algebra $\mathfrak{X}_R(G)$ of $\mathfrak{X}(G)$, which is again linearly isomorphic to T_eG and induces the negative of the Lie algebra structure on T_eG. We will denote by $R : \mathfrak{g} = T_eG \to \mathfrak{X}_R(G)$ the isomorphism discussed, which is given by $R_X(a) = T_e(\mu^a).X$.

If L_X is a left invariant vector field and R_Y is a right invariant vector field, then $[L_X, R_Y] = 0$. So if the flows of L_X and R_Y exist, they commute.

Let $\varphi : G \to H$ be a smooth homomorphism of Lie groups. Then $\varphi' := T_e\varphi : \mathfrak{g} = T_eG \to \mathfrak{h} = T_eH$ is a Lie algebra homomorphism.

B.3. One parameter subgroups

Let G be a Lie group with Lie algebra $\mathfrak{g}$. A *one parameter subgroup* of G is a Lie group homomorphism $\alpha : (\mathbb{R}, +) \to G$, i.e. a smooth curve α in G with $\alpha(s+t) = \alpha(s).\alpha(t)$, and hence $\alpha(0) = e$.

Note that a smooth mapping $\beta : (-\varepsilon, \varepsilon) \to G$ satisfying $\beta(t)\beta(s) = \beta(t+s)$ for $|t|$, $|s|$, $|t+s| < \varepsilon$ is the restriction of a one parameter subgroup. Namely, choose $0 < t_0 < \varepsilon/2$. Any $t \in \mathbb{R}$ can be uniquely written as $t = N.t_0 + t'$ for $0 \le t' < t_0$ and $N \in \mathbb{Z}$. Put $\alpha(t) = \beta(t_0)^N \beta(t')$. The required properties are easy to check.

Let $\alpha : \mathbb{R} \to G$ be a smooth curve with $\alpha(0) = e$. Let $X \in \mathfrak{g}$. Then the following assertions are equivalent.

1. *α is a one parameter subgroup with $X = \partial_t \alpha(t)$.*
2. *$\alpha(t)$ is an integral curve of the left invariant vector field L_X, and also an integral curve of the right invariant vector field R_X.*
3. *$\mathrm{Fl}^{L_X}(t, x) := x.\alpha(t)$ (or $\mathrm{Fl}^{L_X}_t = \mu^{\alpha(t)}$) is the (unique by (A.6)) global flow of L_X in the sense of (A.6).*
4. *$\mathrm{Fl}^{R_X}(t, x) := \alpha(t).x$ (or $\mathrm{Fl}^{R_X}_t = \mu_{\alpha(t)}$) is the (unique) global flow of R_X.*

Moreover, each of these properties determines α uniquely.

B.4. Exponential mapping

Let G be a Lie group with Lie algebra $\mathfrak{g}$. We say that G admits an *exponential mapping* if there exists a smooth mapping $\exp : \mathfrak{g} \to G$ such that $t \mapsto \exp(tX)$ is the (unique by (B.3)) 1-parameter subgroup with tangent vector X at 0. Then we have by (B.3)

1. $\mathrm{Fl}^{L_X}(t, x) = x. \exp(tX)$.
2. $\mathrm{Fl}^{R_X}(t, x) = \exp(tX).x$.
3. $\exp(0) = e$ and $T_0 \exp = Id : T_0\mathfrak{g} = \mathfrak{g} \to T_eG = \mathfrak{g}$ since $T_0 \exp .X = \partial_t|_0 \exp(0 + t.X) = \partial_t|_0 \mathrm{Fl}^{L_X}(t, e) = X$.
4. Let $\varphi : G \to H$ be a smooth homomorphism between Lie groups admitting exponential mappings. Then the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\varphi'"] \arrow[d, "\exp^G"'] & \mathfrak{h} \arrow[d, "\exp^H"] \\ G \arrow[r, "\varphi"] & H \end{tikzcd}$$

commutes, since $t \mapsto \varphi(\exp^G(tX))$ is a one parameter subgroup of H and $\partial_t|_0 \varphi(\exp^G tX) = \varphi'(X)$, so $\varphi(\exp^G tX) = \exp^H(t\varphi'(X))$.

We shall strengthen this notion in (B.9) below and call it a 'regular Fréchet Lie groups'.

If G admits an exponential mapping, it follows from (B.4).(3) that exp is a diffeomorphism from a neighborhood of 0 in $\mathfrak{g}$ onto a neighborhood of e in G, if a suitable inverse function theorem is applicable. This is true for example for smooth Banach Lie groups, also for gauge groups, but it is wrong for diffeomorphism groups.

If E is a Banach space, then in the Banach Lie group $GL(E)$ of all bounded linear automorphisms of E the exponential mapping is given by the von Neumann series $\exp(X) = \sum_{i=0}^{\infty} \frac{1}{i!} X^i$.

If G is connected with exponential mapping and $U \subset \mathfrak{g}$ is open with $0 \in U$, then one may ask whether the group generated by $\exp(U)$ equals G. Note that this is a normal subgroup. So if G is simple, the answer is yes. This is true for connected components of diffeomorphism groups and many of their important subgroups.

B.5. The adjoint representation

Let G be a Lie group with Lie algebra $\mathfrak{g}$. For $a \in G$ we define $\operatorname{conj}_a : G \to G$ by $\operatorname{conj}_a(x) = axa^{-1}$. It is called the *conjugation* or the *inner automorphism* by $a \in G$. This defines a smooth action of G on itself by automorphisms.

The adjoint representation $\operatorname{Ad} : G \to GL(\mathfrak{g}) \subset L(\mathfrak{g}, \mathfrak{g})$ is given by $\operatorname{Ad}(a) = (\operatorname{conj}_a)' = T_e(\operatorname{conj}_a) : \mathfrak{g} \to \mathfrak{g}$ for $a \in G$. By (B.2) $\operatorname{Ad}(a)$ is a Lie algebra homomorphism. By (B.1) we have $\operatorname{Ad}(a) = T_e(\operatorname{conj}_a) = T_a(\mu^{a^{-1}}).T_e(\mu_a) = T_{a^{-1}}(\mu_a).T_e(\mu^{a^{-1}})$.

Finally we define the (lower case) *adjoint representation* of the Lie algebra $\mathfrak{g}$, $\operatorname{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) := L(\mathfrak{g}, \mathfrak{g})$, by $\operatorname{ad} := \operatorname{Ad}' = T_e \operatorname{Ad}$.

We shall also use the *right Maurer-Cartan form* $\kappa^r \in \Omega^1(G, \mathfrak{g})$, given by $\kappa^r_g = T_g(\mu^{g^{-1}}) : T_g G \to \mathfrak{g}$; similarly the *left Maurer-Cartan form* $\kappa^l \in \Omega^1(G, \mathfrak{g})$ is given by $\kappa^l_g = T_g(\mu_{g^{-1}}) : T_g G \to \mathfrak{g}$.

1. $L_X(a) = R_{\operatorname{Ad}(a)X}(a)$ *for* $X \in \mathfrak{g}$ *and* $a \in G$.
2. $\operatorname{ad}(X)Y = [X, Y]$ *for* $X, Y \in \mathfrak{g}$.
3. $d \operatorname{Ad} = (\operatorname{ad} \circ \kappa^r).\operatorname{Ad} = \operatorname{Ad}.(\operatorname{ad} \circ \kappa^l) : TG \to L(\mathfrak{g}, \mathfrak{g})$.

B.6. Right actions

Let $r : M \times G \to M$ be a right action, so $\check{r} : G \to \operatorname{Diff}(M)$ is a group anti-homomorphism. We will use the following notation: $r^a : M \to M$ and $r_x : G \to M$, given by $r_x(a) = r^a(x) = r(x, a) = x.a$.

For any $X \in \mathfrak{g}$ we define the *fundamental vector field* $\zeta_X = \zeta_X^M \in \mathfrak{X}(M)$ by $\zeta_X(x) = T_e(r_x).X = T_{(x,e)}r.(0_x, X)$.

In this situation the following assertions hold:

1. $\zeta : \mathfrak{g} \to \mathfrak{X}(M)$ *is a Lie algebra homomorphism.*
2. $T_x(r^a).\zeta_X(x) = \zeta_{Ad(a^{-1})X}(x.a)$.
3. $0_M \times L_X \in \mathfrak{X}(M \times G)$ *is* r*-related to* $\zeta_X \in \mathfrak{X}(M)$.

B.7. The right and left logarithmic derivatives

Let M be a manifold and let $f : M \to G$ be a smooth mapping into a Lie group G with Lie algebra $\mathfrak{g}$. We define the mapping $\delta^r f : TM \to \mathfrak{g}$ by the formula

$$\delta^r f(\xi_x) := T_{f(x)}(\mu^{f(x)^{-1}}).T_x f.\xi_x \text{ for } \xi_x \in T_x M.$$

Then $\delta^r f$ is a $\mathfrak{g}$-valued 1-form on M, $\delta^r f \in \Omega^1(M;\mathfrak{g})$. We call $\delta^r f$ the *right logarithmic derivative* of f, since for $f : \mathbb{R} \to (\mathbb{R}^+, \cdot)$ we have $\delta^r f(x).1 = \frac{f'(x)}{f(x)} = (\log \circ f)'(x)$.

Similarly the *left logarithmic derivative* $\delta^l f \in \Omega^1(M, \mathfrak{g})$ of a smooth mapping $f : M \to G$ is given by

$$\delta^l f.\xi_x = T_{f(x)}(\mu_{f(x)^{-1}}).T_x f.\xi_x.$$

Let $f, g : M \to G$ *be smooth. Then the Leibniz rule holds:*

$$\delta^r(f.g)(x) = \delta^r f(x) + \mathrm{Ad}(f(x)).\delta^r g(x).$$

Moreover, the differential form $\delta^r f \in \Omega^1(M;\mathfrak{g})$ *satifies the 'left Maurer-Cartan equation' (left because it stems from the left action of* G *on itself)*

$$d\delta^r f(\xi,\eta) - [\delta^r f(\xi), \delta^r f(\eta)]^{\mathfrak{g}} = 0,$$
$$\textit{or} \quad d\delta^r f - \frac{1}{2}[\delta^r f, \delta^r f]^{\mathfrak{g}}_{\wedge} = 0,$$

where $\xi, \eta \in T_x M$*, and where for* $\varphi \in \Omega^p(M;\mathfrak{g}), \psi \in \Omega^q(M;\mathfrak{g})$ *one puts*

$$[\varphi,\psi]^{\mathfrak{g}}_{\wedge}(\xi_1,\dots,\xi_{p+q}) := \frac{1}{p!q!}\sum_{\sigma} \mathrm{sign}(\sigma)[\varphi(\xi_{\sigma 1},\dots),\psi(\xi_{\sigma(p+1)},\dots)]^{\mathfrak{g}}.$$

For the left logarithmic derivative the corresponding Leibniz rule is uglier, and it satisfies the 'right Maurer Cartan equation':

$$\delta^l(fg)(x) = \delta^l g(x) + Ad(g(x)^{-1})\delta^l f(x),$$
$$d\delta^l f + \frac{1}{2}[\delta^l f, \delta^l f]^{\mathfrak{g}}_{\wedge} = 0.$$

For 'regular Lie groups' a converse to this statement holds, see [30], 40.2. The proof of this result in infinite dimensions uses principal bundle geometry for the trivial principal bundle $\mathrm{pr}_1 : M \times G \to M$ with right principal action. Then the submanifolds $\{(x, f(x).g) : x \in M\}$ for $g \in G$ form a foliation of $M \times G$ whose tangent distribution is complementary to the vertical bundle $M \times TG \subseteq T(M \times G)$ and is invariant under the principal right G-action. So it is the horizontal distribution of a principal connection on $M \times G \to G$. Thus this principal connection has vanishing curvature which translates into the result for the right logarithmic derivative.

B.8

Let G be a Lie group with Lie algebra $\mathfrak{g}$. For a closed interval $I \subset \mathbb{R}$ and for $X \in C^\infty(I, \mathfrak{g})$ we consider the ordinary differential equation

$$\begin{cases} g(t_0) &= e \\ \partial_t g(t) &= T_e(\mu^{g(t)})X(t) = R_{X(t)}(g(t)), \qquad \text{or } \kappa^r(\partial_t g(t)) = X(t), \end{cases} \tag{1}$$

for local smooth curves g in G, where $t_0 \in I$.

(2) Local solution curves g of the differential equation (1) *are unique.*

(3) If for fixed X the differential equation (1) *has a local solution near each $t_0 \in I$, then it has also a global solution $g \in C^\infty(I, G)$.*

(4) If for all $X \in C^\infty(I, \mathfrak{g})$ the differential equation (1) *has a local solution near one fixed $t_0 \in I$, then it has also a global solution $g \in C^\infty(I, G)$ for each X. Moreover, if the local solutions near t_0 depend smoothly on the vector fields X then so does the global solution.*

(5) The curve $t \mapsto g(t)^{-1}$ is the unique local smooth curve h in G which satifies

$$\begin{cases} h(t_0) = e \\ \partial_t h(t) = T_e(\mu_{h(t)})(-X(t)) = L_{-X(t)}(h(t)), \qquad \textit{or } \kappa^l(\partial_t h(t)) = -X(t). \end{cases}$$

B.9. Regular Lie groups

If for each $X \in C^\infty(\mathbb{R}, \mathfrak{g})$ there exists $g \in C^\infty(\mathbb{R}, G)$ satisfying

$$\begin{cases} g(0) = e, \\ \partial_t g(t) = T_e(\mu^{g(t)})X(t) = R_{X(t)}(g(t)), \\ \quad \text{or } \kappa^r(\partial_t g(t)) = \delta^r g(\partial_t) = X(t), \end{cases} \tag{1}$$

then we write

$$\begin{aligned}\operatorname{evol}^r_G(X) &= \operatorname{evol}_G(X) := g(1),\\ \operatorname{Evol}^r_G(X)(t) &:= \operatorname{evol}_G(s\mapsto tX(ts)) = g(t),\end{aligned}$$

and call it the *right evolution* of the curve X in G. By lemma (B.8) the solution of the differential equation (1) is unique, and for global existence it is sufficient that it has a local solution. Then

$$\operatorname{Evol}^r_G : C^\infty(\mathbb{R},\mathfrak{g}) \to \{g\in C^\infty(\mathbb{R},G) : g(0)=e\}$$

is bijective with inverse the right logarithmic derivative δ^r.

The Lie group G is called a *regular Lie group* if $\operatorname{evol}^r : C^\infty(\mathbb{R},\mathfrak{g}) \to G$ exists and is smooth.

We also write

$$\begin{aligned}\operatorname{evol}^l_G(X) &= \operatorname{evol}_G(X) := h(1),\\ \operatorname{Evol}^l_G(X)(t) &:= \operatorname{evol}^l_G(s\mapsto tX(ts)) = h(t),\end{aligned}$$

if h is the (unique) solution of

$$\begin{cases} h(0)=e\\ \partial_t h(t) = T_e(\mu_{h(t)})(X(t)) = L_{X(t)}(h(t)),\\ \quad\text{or } \kappa^l(\partial_t h(t)) = \delta^l h(\partial_t) = X(t).\end{cases} \tag{2}$$

Clearly $\operatorname{evol}^l : C^\infty(\mathbb{R},\mathfrak{g}) \to G$ exists and is also smooth if evol^r does, since we have $\operatorname{evol}^l(X) = \operatorname{evol}^r(-X)^{-1}$ by lemma (B.8).

Let us collect some easily seen properties of the evolution mappings. If $f\in C^\infty(\mathbb{R},\mathbb{R})$, then we have

$$\begin{aligned}\operatorname{Evol}^r(X)(f(t)) &= \operatorname{Evol}^r(f'.(X\circ f))(t).\operatorname{Evol}^r(X)(f(0)),\\ \operatorname{Evol}^l(X)(f(t)) &= \operatorname{Evol}^l(X)(f(0)).\operatorname{Evol}^l(f'.(X\circ f))(t).\end{aligned}$$

If $\varphi : G\to H$ is a smooth homomorphism between regular Lie groups then the diagram

$$\begin{tikzcd} C^\infty(\mathbb{R},\mathfrak{g}) \arrow[r, "\varphi'_*"] \arrow[d, "\operatorname{evol}_G"'] & C^\infty(\mathbb{R},\mathfrak{h}) \arrow[d, "\operatorname{evol}_H"] \\ G \arrow[r, "\varphi"] & H \end{tikzcd}$$

commutes, since $\partial_t\varphi(g(t)) = T\varphi.T(\mu^{g(t)}).X(t) = T(\mu^{\varphi(g(t))}).\varphi'.X(t)$.

Note that each regular Lie group admits an exponential mapping, namely the restriction of evol^r to the constant curves $\mathbb{R}\to\mathfrak{g}$. A Lie group is regular if and only if its universal covering group is regular.

Up to now the following statement holds:

All known Lie groups are regular.

Any Banach Lie group is regular since we may consider the time dependent right invariant vector field $R_{X(t)}$ on G and its integral curve $g(t)$ starting at e, which exists and depends smoothly on (a further parameter in) X. In particular finite dimensional Lie groups are regular.

For diffeomorphism groups the evolution operator is just integration of time dependent vector fields with compact support.

B.10. Extensions of Lie groups

Let H and K be Lie groups. A Lie group G is called a smooth *extension* of H with kernel K if we have a short exact sequence of groups

$$\{e\} \to K \xrightarrow{i} G \xrightarrow{p} H \to \{e\}, \tag{1}$$

such that i and p are smooth and one of the following two equivalent conditions is satisfied:

2 p admits a local smooth section s near e (equivalently near any point), and i is initial (i. e. any f into K is smooth if and only if $i \circ f$ is smooth).
1. i admits a local smooth retraction r near e (equivalently near any point), and p is final (i. e. f from H is smooth if and only if $f \circ p$ is smooth).

Of course by $s(p(x))i(r(x)) = x$ the two conditions are equivalent, and then G is locally diffeomorphic to $K \times H$ via (r, p) with local inverse $(i \circ \mathrm{pr}_1).(s \circ \mathrm{pr}_2)$.

Not every smooth exact sequence of Lie groups admits local sections as required in (2). Let for example K be a closed linear subspace in a convenient vector space G which is not a direct summand, and let H be G/K. Then the tangent mapping at 0 of a local smooth splitting would make K a direct summand.

Let $\{e\} \to K \xrightarrow{i} G \xrightarrow{p} H \to \{e\}$ be a smooth extension of Lie groups. Then G is regular if and only if both K and H are regular.

B.11. Subgroups of regular Lie groups

Let G and K be Lie groups, let G be regular and let $i : K \to G$ be a smooth homomorphism which is initial (see (B.10)) with $T_e i = i' : \mathfrak{k} \to \mathfrak{g}$ injective. We suspect that K is then regular, but we know a proof for this only under the following assumption. There is an open neighborhood $U \subset G$ of e and a smooth mapping $p : U \to E$ into a convenient vector space E such that $p^{-1}(0) = K \cap U$ and p constant on left cosets $Kg \cap U$.

References

1. V. Arnold and B. Khesin. *Topological methods in fluid dynamics.* Applied Math. Sciences. Springer-Verlag, 1998.
2. V. I. Arnold. *Geometrical methods in the theory of ordinary differential equations.* Springer-Verlag, New York, xi+334 pp, 1983. Fundamental Principles of Mathematical Sciences, 250.
3. V.I. Arnold. An a priori estimate in the theory of hydrodynamic stability. *Izvestia Vyssh. Uchebn. Zaved. Matematicka*, 54,5 (1966), 3–5. Russian.
4. V.I. Arnold. Sur la géometrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier*, 16 (1966), 319–361.
5. V.I. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, New York, Heidelberg, 1978. Graduate Texts in Math. 60.
6. E. Binz. Two natural metrics and their covariant derivatives on a manifold of embeddings. *Monatsh. Math.*, 89 (1980), 275–288.
7. E. Binz and H.R. Fischer. The manifold of embeddings of a closed manifold. Proc. Differential geometric methods in theoretical physics, Clausthal 1978. Springer Lecture Notes in Physics 139. 1981.
8. R. Bott. On the characteristic classes of groups of diffeomorphisms. *Enseign. Math.*, 23, 1977.
9. Jean-Luc Brylinski. *Loops spaces, characteristic classes, and geometric quantization.* Birkhäuser, Boston, Basel, Berlin, 1993. Progress in Math. 107.
10. J. Burgers. A mathematical model illustrating the theory of turbulence. *Adv. Appl. Mech.*, 1 (1948), 171–199.
11. R. Camassa and D. Holm. An integrable shallow water equation with peaked solutions. *Phys. Rev. Lett.*, 71 (1993), 1661–1664.
12. A. Constantin, T. Kappeler, B. Kolev, and P. Topalov. On geodesic exponential maps of the Virasoro group. Preprint 13-2004, University of Zürich
13. A. Constantin and B. Kolev. Geodesic flow on the diffeomorphism group of the circle. *Comm. Math. Helv.*, 78 (2003), 787–804.
14. K. Deimling. *Ordinary differential equations in Banach spaces.* Lecture Notes in Mathematics, Vol. 596. Springer-Verlag, Berlin-New York, 1977.
15. J. Eichhorn. The manifold structure of maps between open manifolds. Ann. Global Anal. Geom. 11 (1993), no. 3, 253–300.
16. J. Eichhorn. Diffeomorphism groups on noncompact manifolds. Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 234 (1996), Differ. Geom. Gruppy Li i Mekh. 15-1, 41–64, 262; translation in J. Math. Sci. (New York) 94 (1999), no. 2, 1162–1176
17. J. Eichhorn and J. Fricke. The module structure theorem for Sobolev spaces on open manifolds. Math. Nachr. 194 (1998), 35–47.
18. D.B.A. Epstein. The simplicity of certain groups of homeomorphisms. *Compositio Math.*, 22 (1970), 165–173.
19. C.F. Faà di Bruno. Note sur une nouvelle formule du calcul différentielle. *Quart. J. Math.*, 1 (1855), 359–360.
20. A. Frölicher and A. Kriegl. Linear spaces and differentiation theory. Pure and Applied Mathematics, J. Wiley, Chichester 1988.
21. D. Fuks, *Cohomology of infinite dimensional Lie algebras.* Nauka, Moscow, 1984 (Russian). Transl. English, Contemporary Soviet Mathematics. Consultants Bureau (Plenum Press), New York, 1986.
22. I.M. Gelfand and I.Y. Dorfman. Hamiltonian operators and the algebraic structures connected with them. *Funct. Anal. Appl.*, 13 (1979), 13–30.
23. L. Hörmander. *The Analysis of linear partial differential operators I, II, III, IV.* Springer-Verlag, Berlin, 1983, 1983, 1985, 1985. Grundlehren 256, 257, 274, 275.

24. Gelfand I.M. and Fuchs D.B. The cohomologies of the Lie algebra of the vector fields in a circle *Funkts. Anal. Prilozh.*, 2 (1968), 92-93.
25. Gerd Kainz. A note on the manifold of immersions and its Riemannian curvature. *Monatshefte für Mathematik*, 98 (1984), 211–217.
26. B Khesin and P.W. Michor. *The flow completion of Burgers' equation*, pages 17–26. De Gruyter, Berlin, 2004. IRMA Lectures in Mathematics and Theoretical Physics 5.
27. A.A. Kirillov. The orbits of the group of diffeomorphisms of the circle, and local Lie superalgebras. *Funct. Anal. Appl.*, 15 (1981), 135–136.
28. S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry. Vol. I. Vol. II.* J. Wiley-Intersci-ence, 1963, 1969.
29. S. Kouranbaeva. The Camassa-holm equation on as a geodesic flow on the diffeomorphism group. *J. Math. Phys.*, 40 (1999), 857–868.
30. A. Kriegl and P. W. Michor. A convenient setting for real analytic mappings. *Acta Mathematica*, 165 (1990), 105–159.
31. A. Kriegl and P. W. Michor. *The Convenient Setting for Global Analysis.* AMS, Providence, 1997. Surveys and Monographs 53.
32. A. Kriegl and P. W. Michor. Regular infinite dimensional Lie groups. *J. Lie Theory (http://ww w.emis.de/journals/JLT)*, 7,1, 1997. ESI Preprint 200.
33. A. Lasota and J.A. Yorke. The generic property of existence of solutions of differential equations in Banach space. *J. Differential Equations*, 13 (1973), 1–12.
34. J Marsden and T. Ratiu. *Introduction to mechanics and symmetry.* Springer-Verlag, New York, Berlin, Heidelberg, 1994.
35. John N. Mather. Commutators of diffeomorphisms I, II. *Comm. Math. Helv.*, 49 (1974), 512–528.
36. P. W. Michor. Manifolds of smooth maps III: The principal bundle of embeddings of a non compact smooth manifold. *Cahiers Topol. Geo. Diff.*, 21 (1980), 325–337.
37. P. W. Michor. Manifolds of differentiable mappings. Shiva Mathematics Series, 3. Shiva Publishing Ltd., Nantwich, 1980.
38. P. W. Michor. Gauge theory for diffeomorphism groups. Proceedings of the Conference on Differential Geometric Methods in Theoretical Physics Como 1987, K. Bleuler, M. Werner (eds.), 345–371. Kluwer, Dordrecht, 1988.
39. P. W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *J. Eur. Math. Soc. (JEMS)* 8 (2006), 1-48, arXiv:math.DG/0312384.
40. P. W. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. Documenta Math. 10 (2005), 217–245, arXiv:math.DG/0409303.
41. P.W. Michor and T. Ratiu. Geometry of the Virasoro-bott group. *J. Lie Theory*, 8 (1998), 293–309.
42. G. Misiołek. Conjugate points in the Bott-Virasoro group and the KdV equation. *Proc. Amer. Math. Soc.* 125 (1997), 935–940.
43. G. Misiołek. Classical solutions of the periodic Camassa-holm equation. *GAFA*, 12 (2002), 1080–1104.
44. G. Misiołek A shallow water equation as a geodesic flow on the Bott-virasoro group. *J. Geom. Phys.*, 24 (1998), 203–208.
45. J.P. Ortega and T. Ratiu. *Momentum maps and Hamiltonian reduction.* Birkhäuser, Boston, 2004. Progress in Mathematics 222.
46. V.Y. Ovsienko and B.A. Khesin. Korteweg–de Vries superequations as an Euler equation. *Funct. Anal. Appl.*, 21 (1987), 329–331.
47. A. Pressley and G. Segal. *Loop groups.* Oxford University Press, 1986. Oxford Mathematical Monographs.
48. T. Ratiu. The motion of the n-dimensional rigid body. *Indiana Univ. Math. J.*, 29 (1980), 609–629.
49. G. Segal. The geometry of the Kdv equation. *Internat. J. Modern Phys. A* , 6 (1991), 2859–2869.

50. W. Thurston. Foliations and groups of diffeomorphisms. *Bull. AMS*, 80 (1974), 304–307.
51. A. Weinstein. Symplectic manifolds and their Lagrangian manifolds. *Advances in Math.*, 6 (1971), 329–345.