Geometry of diffeomorphism groups and of shape spaces

Peter W. Michor

Fakultät für Mathematik der Universität Wien, Nordbergstrasse 15, A-1090 Wien, Austria.

Erwin Schrödinger Institut für Mathematische Physik, Boltzmanngasse 9, A-1090 Wien, Austria.

peter.michor@univie.ac.at

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 lecture course will be held again in WS2008/09 and SS 2009. The material here is fluid, since I write on it constantly. Do not print too much of it since it will change a lot.

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 [**42**] and [**50**]. 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 [**77**], it was fully proved in [**13**] for compact M and for general M in [**49**]. See also [**50**], section 13 and [**42**]. 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 [**52**].

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 [**12**] if $N = \mathbb{R}^{\dim M+1}$ (endowed with the standard inner product) and in [**35**] for the general case. In [**55**] 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

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

$$= -\tfrac{1}{2}\int_a^b\int_{\mathbb{R}}(2f_{tt}f_sf_x + 2f_tf_sf_{tx} + 2f_tf_{tx}f_s)dxdt$$
$$= -\int_a^b\int_{\mathbb{R}}\left(f_{tt} + 2\frac{f_tf_{tx}}{f_x}\right)f_sf_xdxdt,$$

so that the geodesic equation with its initial data is:

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

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

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

For vector fields X, Y on $\mathrm{Emb}(\mathbb{R},\mathbb{R})$ the covariant derivative is given by the expression $\nabla_X^{\mathrm{Emb}}Y = dY(X) - \Gamma(X,Y)$. 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

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

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

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

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 [**17**], [**6**], or [**36**].

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

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

Therefore 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

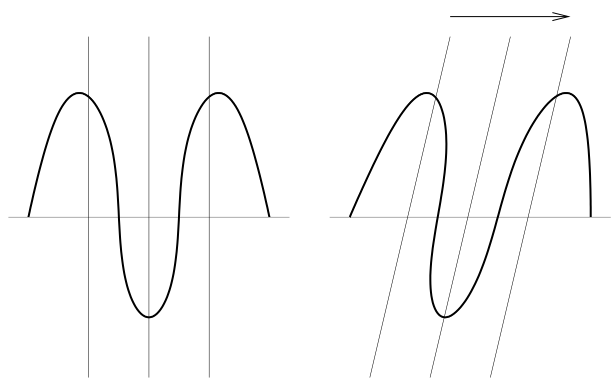


Figure 1. The characteristic flow of the inviscid Burgers' equation tilts the plane.

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 $x(0,z) = z \in \mathbb{R}$.

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.

Background material

2. 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 [**28**] and [**42**]. 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 [**42**] and this is the main reference for the whole section. We list results in the order in which one can prove them, without proofs for which we refer to [**42**]. This should explain how to use these results. Later we also explain the fundamentals about regular infinite dimensional Lie groups.

2.1. 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.

2.2. 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.*

2.3. Theorem. [**28**], 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 [**28**], 4.1.23.

2.4. 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 $T(x_0, x_1, x_2, \dots) := (0, 1^2 x_1, 2^2 x_2, 3^2 x_3, \dots)$. 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 2.6 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 $\mathrm{ev}_s \circ x = x(s+\ \ )$ is not real-analytic around 0.

2.5. Manifolds and vector fields. In the sequel we shall use smooth manifolds M modelled on c^∞-open subsets of convenient vector spaces. Tangent vectors in T_xM have to be defined as equivalence classes of smooth curves c M with $c(0) = x$; two curves c_1, c_2 are equivalent if $(u \circ c_1)'(0) = (u \circ c_2)'(0)$ for one (equivalently, any) chart $u : U \to V$ with $x \in U$. The tangent bundle is then again a smooth manifold; the chart changes of its atlas are just the tangent mappings of the the chart changes of an atlas of M.

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

2.6. 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.

2.7. 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 $\partial_t|_0 \varphi_t = X \in \Gamma(TM)$. 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 $\omega \in \Omega^k(M)$:

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

3. Regular infinite dimensional Lie groups

3.1. 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

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

3.2. 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.

3.3. 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 $X = \partial_t|_0 \alpha(t)$.*

(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 2.6) global flow of L_X in the sense of 2.6.*

(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.

3.4. 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 3.3) 1-parameter subgroup with tangent vector X at 0. Then we have by 3.3

(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 $\varphi(\exp^G tX) = \exp^H(t\varphi'(X))$.

Proof. (1) $\Longrightarrow$ (4) We have $\frac{d}{dt} x.\alpha(t) = \frac{d}{ds}|_0 x.\alpha(t+s) = \frac{d}{ds}|_0 x.\alpha(t).\alpha(s) = \frac{d}{ds}|_0 \mu_{x.\alpha(t)} \alpha(s) = T_e(\mu_{x.\alpha(t)}).\frac{d}{ds}|_0 \alpha(s) = T_e(\mu_{x.\alpha(t)}).X = L_X(x.\alpha(t))$. By uniqueness of solutions we get $x.\alpha(t) = \mathrm{Fl}^{L_X}(t, x)$.

(4) $\Longrightarrow$ (2) This is clear.

(2) $\implies$ (1) We have

$$\begin{aligned}\tfrac{d}{ds}\alpha(t)\alpha(s) &= \tfrac{d}{ds}(\mu_{\alpha(t)}\alpha(s)) = T(\mu_{\alpha(t)})\tfrac{d}{ds}\alpha(s)\\ &= T(\mu_{\alpha(t)})L_X(\alpha(s)) = L_X(\alpha(t)\alpha(s))\end{aligned}$$

and $\alpha(t)\alpha(0) = \alpha(t)$. So we get $\alpha(t)\alpha(s) = \mathrm{Fl}^{L_X}(s, \alpha(t)) = \mathrm{Fl}^{L_X}_s \mathrm{Fl}^{L_X}_t(e) = \mathrm{Fl}^{L_X}(t+s, e) = \alpha(t+s)$.

(4) $\iff$ (1) We have $\mathrm{Fl}^{\nu^*\xi}_t = \nu^{-1} \circ \mathrm{Fl}^{\xi}_t \circ \nu$ by 2.6. Therefore we have

$$\begin{aligned}(\mathrm{Fl}^{R_X}_t(x^{-1}))^{-1} &= (\nu \circ \mathrm{Fl}^{R_X}_t \circ \nu)(x) = \mathrm{Fl}^{\nu^* R_X}_t(x)\\ &= \mathrm{Fl}^{L_X}_{-t}(x) = x.\alpha(-t).\end{aligned}$$

So $\mathrm{Fl}^{R_X}_t(x^{-1}) = \alpha(t).x^{-1}$, and $\mathrm{Fl}^{R_X}_t(y) = \alpha(t).y$.

$\implies$ (3) $\implies$ (1) can be shown in a similar way. □

We shall strengthen this notion in 3.11 below and call it a ‘regular Fréchet Lie groups’.

If G admits an exponential mapping, it follows from 3.4.3 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.

3.5. Analysis on Lie groups. Let be an infinite dimensional Lie group admitting an exponential map (for simplicity's sake) and let V be a convenient vector space. For $f \in C^\infty(G, V)$ we have $df \in \Omega^1(G; V)$, a 1-form on G with values in V. We define the *right trivialized derivative* $\delta f = \delta^r f : G \to L(\mathfrak{g}, V)$ of f by

$$\delta^r f(x).X := df.T_e(\mu^x).X = (R_X f)(x) \text{ for } x \in G, X \in \mathfrak{g}.$$

Lemma. [**51**]

(1) *For $f \in C^\infty(G, \mathbb{R})$ and $g \in C^\infty(G, V)$ we have $\delta(f.g) = f.\delta^r g + \delta f \otimes g$, where we use $\mathfrak{g}^* \otimes V \to L(\mathfrak{g}, V)$.*

(2) *For $f \in C^\infty(G,V)$ we have*

$$\delta\delta f(x)(X,Y) - \delta^r\delta f(x)(Y,X) = \delta f(x)(-[X,Y]).$$

(3) *Fundamental theorem of calculus: For $f \in C^\infty(G,V)$, $x \in G$, $X \in \mathfrak{g}$ we have*

$$f(\exp(X).x) - f(x) = \Big(\int_0^1 \delta f(\exp(tX).x)\,dt\Big)(X).$$

(4) *Taylor expansion with remainder For $f \in C^\infty(G,V)$, $x \in G$, $X \in \mathfrak{g}$ we have*

$$\begin{aligned} f(\exp(X).x) = \sum_{j=0}^{N} \frac{1}{j!}\delta^j f(x)(X^j) + \\ + \int_0^1 \frac{(1-t)^N}{N!}\delta^{N+1} f(\exp(tX).x)\,dt\,(X^{N+1}). \end{aligned}$$

(5) *For $f \in C^\infty(G,V)$ and $x \in G$ the formal Taylor series*

$$\operatorname{Tay}_x f = \sum_{j=0}^{\infty} \frac{1}{j!}\delta^j f(x) : \bigotimes \mathfrak{g} \to \mathbb{R}$$

factors to a linear functional on the universal envelopping algebra of the Lie algebra $(\mathfrak{g}. -[\quad,\quad])$: $\mathcal{U}(\mathfrak{g}, -[\quad,\quad]) \to \mathbb{R}$. If for $A \in \mathcal{U}(\mathfrak{g})$ we denote by L_A the associated left invariant differential operator on G, we have $\langle A, \operatorname{Tay}_x f\rangle = (L_A f)(x)$

Proof. (1) This is mainly notation:

$$\begin{aligned} \delta(f.g)(x)(X) &= d(f.g).T(\mu^x).X = (df.T(\mu^x).X)g(x) + f(x).dg.T(\mu^x).X \\ &= (f.\delta g + \delta f \otimes g)(x)(X). \end{aligned}$$

(2) We compute

$$\begin{aligned} \delta^2 f(x)(X,Y) - \delta^2 f(x)(Y,X) &= ((R_X R_Y - R_Y R_X)f)(x) = (R_{-[X,Y]}f)(x) \\ &= \delta f(x)(-[X,Y]). \end{aligned}$$

(3) By 3.3 we have

$$\begin{aligned} \partial_t f(\exp(tX).x) &= df.T(\mu^x).\partial_t \exp(tX) = df.T(\mu^x).T(\mu^{\exp(tX)}).X \\ &= \delta f(\exp(tX).x)(X) \end{aligned}$$

and thus

$$f(\exp(X).x) - f(x) = \int_0^1 \delta f(\exp(tX).x)\,dt\,(X).$$

(4) By iteration we get

$$f(\exp(X).x) - f(x) = \int_0^1 1.\delta f(\exp(tX).x)(X)\,dt$$

$$= (t-1).\delta f(\exp(tX).x)(X)|_{t=0}^{t=1} - \int_0^1 (t-1).\delta^2 f(\exp(tX).x)(X,X)\,dt$$
$$= \delta f(x)(X) + \frac{-(1-t)^2}{2}.\delta^2 f(\exp(tX).x)(X^2)|_{t=0}^{t=1} +$$
$$+ \int_0^1 \frac{(1-t)^2}{2}.\delta^3 f(\exp(tX).x)(X^3)\,dt = \dots$$
$$= \sum k = 0^N \frac{1}{k!}\delta^k f(x)(X^k) + \int_0^1 \frac{(1-t)^N}{N!}.\delta^{N+1} f(\exp(tX).x)(X^{X+1})\,dt$$

(5) is a consequence of (4) and of (2). □

3.6. Vector fields and differential forms. Let G be an infinite dimensional Lie group. For $f \in C^\infty(G,\mathfrak{g})$ we get a smooth vector field $R_f \in \mathfrak{X}(G)$ by $R_f(x) := T_e(\mu^x).f(x)$. This describes a bibounded linear isomorphism $R : C^\infty(G,\mathfrak{g}) \to \mathfrak{X}(G)$. If $h \in C^\infty(G,V)$ then we have $R_f h(x) = dh(R_f(x)) = dh.T_e(\mu^x).f(x) = \delta^r h(x).f(x)$, for which we write shortly $R_f h = \delta^r h.f = \delta h.f$.

For $g \in C^\infty(G,\Lambda^k\mathfrak{g}^*)$ we get a k-form $R_g \in \Omega^k(G)$ by $(R_g)_x = g(x) \circ \Lambda^k T_x(\mu^{x^{-1}})$. This gives an isomorphism $R : C^\infty(G, L^k_{\text{skew}}(\mathfrak{g};\mathbb{R})) \to \Omega(G)$.

Proposition. [**51**]

(1) *For $f,g \in C^\infty(G,\mathfrak{g})$ we have*
$$\begin{aligned}[R_f,R_g]_{\mathfrak{X}(G)} &= R_{K(f,g)} \quad \textit{where}\\ K(f,g)(x) &= -[f(x),g(x)]_{\mathfrak{g}} + \delta^r g(x).f(x) - \delta^r f(x).g(x) \quad \textit{or}\\ K(f,g) &= -[f,g]_{\mathfrak{g}} + \delta g.f - \delta f.g.\end{aligned}$$

(2) *For $g \in C^\infty(G, L^k_{skew}(\mathfrak{g};\mathbb{R}))$ and $f_i \in C^\infty(G,\mathfrak{g})$ we have*
$$R_g(R_{f_1},\dots,R_{f_k}) = g.(f_1,\dots,f_k).$$

(3) *For $g \in C^\infty(G, L^k_{skew}(\mathfrak{g};\mathbb{R}))$ the exterior derivative is given by*
$$d(R_g) = R_{\delta^\wedge g - \partial^{\mathfrak{g}}\circ g},$$
where $\delta^\wedge g : G \to L^{k+1}_{skew}(\mathfrak{g};\mathbb{R})$ is given by
$$\delta^\wedge g(x)(X_0,\dots,X_k) = \sum_{i=0}^k (-1)^i \delta g(x)(X_i)(X_0,\dots,\widehat{X_i},\dots,X_k),$$
and where $\partial^{\mathfrak{g}}$ is the Chevalley differential on $L^*_{skew}(\mathfrak{g};\mathbb{R})$.

(4) *For $g \in C^\infty(G, L^k_{skew}(\mathfrak{g};\mathbb{R}))$ and $f \in C^\infty(G,\mathfrak{g})$ the Lie derivative is given by*
$$\mathcal{L}_{R_f} R_g = R_{\mathcal{L}^{\mathfrak{g}}_f \circ g + \mathcal{L}^\delta_f g},$$

where

$$(\mathcal{L}^{\mathfrak g}_f g)(x)(X_1,\dots,X_k) = -\sum_i (-1)^i g(x)([f(x),X_i]_{\mathfrak g}, X_1,\dots,\widehat{X_i},\dots,X_k)$$

$$\begin{aligned}(\mathcal{L}^{\delta}_f g)(x)(X_1,\dots,X_k) &= \delta g(x)(f(x))(X_1,\dots,X_k)+\\ &\quad+\sum_i (-1)^i g(x)(\delta f(x)(X_i), X_1,\dots,\widehat{X_i},\dots,X_k)\end{aligned}$$

Proof. (1) For $h\in C^\infty(G,\mathbb{R})$ we have

$$\begin{aligned}(R_g\,h)(x) &= \delta h(x)(g(x))\\ (R_f\,R_g\,h) &= \delta\big(\delta h(\quad)(g(x))\big)(x)(f(x)) + \delta h(x)(\delta g(x)(f(x)))\\ &= \delta^2 h(x)(f(x),g(x)) + \delta h(x).\delta g(x).f(x)\\ [R_f,R_g](h) &= R_f.R_g.h - R_g.R_f.h\\ &= \delta^2 h.(f,g) + \delta h.\delta g.f - \delta^2 h(g,f) - \delta h.\delta f.g\\ &= \delta h.\big(-[f,g]_{\mathfrak g} + \delta g.f - \delta f.g\big), \qquad \text{by 3.5.2}\end{aligned}$$

(2) is obvious. (3) It suffices to evaluate the exterior differential dR_g on right invariant vector fields R_{X_i} where $f_0=X_0,\dots,f_k=X_k\in\mathfrak g$. So we have

$$(dR_g)_x(R_{X_0},\dots,R_{X_k}) = \sum_{i=0}^k (-1)^i R_{X_i}(R_g(R_{X_0},\dots,\widehat{R_{X_i}},\dots,R_{X_k}))(x)$$

$$+\sum_{i<j}(-1)^{i+j}R_g([R_{X_i},R_{X_j}],R_{X_0},\dots,\widehat{R_{X_i}},\dots,\widehat{R_{X_j}},\dots,R_{X_k})(x)$$

$$=\sum_{i=0}^k (-1)^i \delta^r g(x)(X_i)(X_0,\dots,\widehat{X_i},\dots,X_k))$$

$$+\sum_{i<j}(-1)^{i+j}g(x)(-[X_i,X_j],X_0,\dots,\widehat{X_i},\dots,\widehat{X_j},\dots,X_k).$$

(4) can be proved similarly and is left for the reader. It will not be used. □

3.7. 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. Recall the left and right invariant vector field mappings $L,R:\mathfrak g\to\mathfrak X(G)$ from 3.2, given by $L_X(g)=T_e(\mu_g).X$ and $R_X=T_e(\mu^g).X$, respectively. They are related by $L_X(g)=R_{\operatorname{Ad}(g)X}(g)$.

The adjoint representation $\mathrm{Ad} : G \to GL(\mathfrak{g}) \subset L(\mathfrak{g},\mathfrak{g})$ is given by $\mathrm{Ad}(a) = (\mathrm{conj}_a)' = T_e(\mathrm{conj}_a) : \mathfrak{g} \to \mathfrak{g}$ for $a \in G$. By 3.2 $\mathrm{Ad}(a)$ is a Lie algebra homomorphism. By 3.1 we have $\mathrm{Ad}(a) = T_e(\mathrm{conj}_a) = T_a(\mu^{a^{-1}}).T_e(\mu_a) = T_{a^{-1}}(\mu_a).T_e(\mu^{a^{-1}})$.

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

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

Proposition.

(1) $L_X(a) = R_{\mathrm{Ad}(a)X}(a)$ *for* $X \in \mathfrak{g}$ *and* $a \in G$.
(2) $\mathrm{ad}(X)Y = [X,Y]$ *for* $X, Y \in \mathfrak{g}$.
(3) $d\,\mathrm{Ad} = (\mathrm{ad} \circ \kappa^r).\mathrm{Ad} = \mathrm{Ad}.(\mathrm{ad} \circ \kappa^l) : TG \to L(\mathfrak{g},\mathfrak{g})$.
(4) κ^r *satisfies the left Maurer-Cartan equation* $d\kappa - \frac{1}{2}[\kappa,\kappa]^\wedge_\mathfrak{g} = 0$.
(5) κ^l *satisfies the right Maurer-Cartan equation* $d\kappa + \frac{1}{2}[\kappa,\kappa]^\wedge_\mathfrak{g} = 0$.

Here $[\ ,\ ]^\wedge_\mathfrak{g}$ denotes the wedge product of $\mathfrak{g}$-valued forms on G induced by the Lie bracket. Note that $\frac{1}{2}[\kappa,\kappa]_\wedge(\xi,\eta) = [\kappa(\xi),\kappa(\eta)]$ for $\xi, \eta \in T_xG$.

Proof. (1) is obvious. To prove (2) we will apply 3.5.2 to $f(x) = X$ and $g(x) = \mathrm{Ad}(x).Y$. We have $L_Y = R_g$ by (1), and $[R_f, L_Y] = [R_X, L_Y] = 0$ by 3.2. So by 3.6.1 we have

$$\begin{aligned} 0 &= [R_X, L_Y](x) = [R_f, R_g](x) \\ &= R(-[X, \mathrm{Ad}.Y]_\mathfrak{g} + \delta(\mathrm{Ad}(\quad).Y).X - 0)(x) \\ [X,Y] &= [X, \mathrm{Ad}(e)Y] = \delta^r(\mathrm{Ad}(\quad).Y)(e).X \\ &= d(\mathrm{Ad}(\quad).X)(e).Y = \mathrm{ad}(X)Y. \end{aligned}$$

(3) We compute

$$\begin{aligned} d\,\mathrm{Ad}(T\mu_g.X) &= \frac{d}{dt}|_0 \mathrm{Ad}(g.\exp(tX)) = \mathrm{Ad}(g).\mathrm{ad}(X) \\ &= \mathrm{Ad}(g).\mathrm{ad}(\kappa^l(T\mu_g.X)), \\ d\,\mathrm{Ad}(T\mu^g.X) &= \frac{d}{dt}|_0 \mathrm{Ad}(\exp(tX).g) = \mathrm{ad}(X).\mathrm{Ad}(g) \\ &= \mathrm{ad}(\kappa^r(T\mu^g.X)).\mathrm{Ad}(g). \end{aligned}$$

(4) We evaluate $d\kappa^r$ on right invariant vector fields R_X, R_Y for $X, Y \in \mathfrak{g}$.

$$\begin{aligned} (d\kappa^r)(R_x, R_Y) &= R_X(k^r(R_Y)) - R_Y(\kappa^r(R_X)) - \kappa^r([R_X, R_Y]) \\ &= R_X(Y) - R_Y(X) + [X,Y] = 0 - 0 + [\kappa^r(R_X), \kappa^r(R_Y)]. \end{aligned}$$

(5) can be proved similarly. $\square$

3.8. 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* $\zeta_X \in \mathfrak{X}(M)$.

3.9. 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

$$\begin{aligned}\delta^r f(\xi_x) :&= T_{f(x)}(\mu^{f(x)^{-1}}).T_x f.\xi_x = \kappa^r_{f(x)}(T_x f.\xi_x)\\ &= (f^*\kappa^r)(\xi_x) \text{ for } \xi_x \in T_x M.\end{aligned}$$

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

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

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

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

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)

$$d\delta^r f(\xi,\eta) - [\delta^r f(\xi), \delta^r f(\eta)]^{\mathfrak{g}} = 0,$$

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

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

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

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 [**42**], 40.2. This result has a geometric interpretation in 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.

Proof. For the Leibniz rule we compute for $\xi_x \in T_xM$, using 3.1,

$$\begin{aligned}
\delta^r(f.g)(\xi_x) &= T_{f(x).g(x)}(\mu^{(f(x).g(x))^{-1}}).T(\mu \circ (f,g)).\xi_x \\
&= T_{f(x).g(x)}(\mu^{(f(x).g(x))^{-1}}).T\mu.(T_xf.\xi_x, T_xg.\xi_x) \\
&= T(\mu^{f(x)^{-1}}).T(\mu^{g(x)^{-1}}).\big(T(\mu_{f(x)}).T_xg.\xi_x) + T(\mu^{g(x)}).T_xf.\xi_x\big) \\
&= T(\mu^{f(x)^{-1}}).T(\mu_{f(x)}).T(\mu^{g(x)^{-1}}).T_xg.\xi_x) + T(\mu^{f(x)^{-1}}).T_xf.\xi_x \\
&= \mathrm{Ad}(f(x)).\delta^r g(\xi_x) + \delta^r f(\xi_x).
\end{aligned}$$

For the Maurer-Cartan equation we use 3.6.4:

$$\begin{aligned}
d(\delta^r f) &= d(f^*\kappa^r) = f^*(d\kappa^r) = f^*(\tfrac{1}{2}[\kappa^r, \kappa^r]^\wedge_{\mathfrak{g}}) = \tfrac{1}{2}[f^*\kappa^r, f^*\kappa^r]^\wedge_{\mathfrak{g}} \\
&= \tfrac{1}{2}[\delta^r f, \delta^r f]^\wedge_{\mathfrak{g}}
\end{aligned}$$

For the left Maurer-Cartan form the proof is analogous. □

3.10. 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

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

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*

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

3.11. 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

$$\mathrm{evol}^r_G(X) = \mathrm{evol}_G(X) := g(1),$$
$$\mathrm{Evol}^r_G(X)(t) := \mathrm{evol}_G(s \mapsto tX(ts)) = g(t),$$

and call it the *right evolution* of the curve X in G. By lemma 3.10 the solution of the differential equation (1) is unique, and for global existence it is sufficient that it has a local solution. Then

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

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

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

We also write

$$\mathrm{evol}^l_G(X) = \mathrm{evol}_G(X) := h(1),$$
$$\mathrm{Evol}^l_G(X)(t) := \mathrm{evol}^l_G(s \mapsto tX(ts)) = h(t),$$

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 $\mathrm{evol}^l : C^\infty(\mathbb{R}, \mathfrak{g}) \to G$ exists and is also smooth if evol^r does, since we have $\mathrm{evol}^l(X) = \mathrm{evol}^r(-X)^{-1}$ by lemma 3.10.

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

$$\mathrm{Evol}^r(X)(f(t)) = \mathrm{Evol}^r(f'.(X \circ f))(t).\, \mathrm{Evol}^r(X)(f(0)),$$
$$\mathrm{Evol}^l(X)(f(t)) = \mathrm{Evol}^l(X)(f(0)).\, \mathrm{Evol}^l(f'.(X \circ f))(t).$$

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

$$\begin{array}{ccc} C^\infty(\mathbb{R},\mathfrak{g}) & \xrightarrow{\varphi'_*} & C^\infty(\mathbb{R},\mathfrak{h}) \\ {\scriptstyle \mathrm{evol}_G}\downarrow & & \downarrow{\scriptstyle \mathrm{evol}_H} \\ G & \xrightarrow{\quad\varphi\quad} & H \end{array}$$

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.

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

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

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).

(3) 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.

3.13. 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 3.12) 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 $Kg \cap U$.

4. Weak symplectic manifolds

4.1. 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 $i(\text{grad}^\omega f)\omega = df$ 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!).

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

A linear lift $j : \mathfrak{g} \to C^\infty(M,\mathbb{R})$ of ζ with $\text{grad}^\omega \circ j = \zeta$ 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 $\bar{\jmath} : \mathfrak{g} \times \mathfrak{g} \to H^0(M)$, given by $(i \circ \bar{\jmath})(X,Y) = \{j(X), j(Y)\} - j([X,Y])$, vanishes in the Lie algebra cohomology $H^2(\mathfrak{g}, H^0(M))$, for if $\bar{\jmath} = \delta\alpha$ 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* $J : M \to \mathfrak{g}' = L(\mathfrak{g}, \mathbb{R})$ to it, which is given by $J(x)(X) = \chi(X)(x)$ for $x \in M$ and $X \in \mathfrak{g}$. It is G-equivariant for a suitably chosen (in general affine) action of G on $\mathfrak{g}'$.

4.2. We now want to carry over to infinite dimensional manifolds the procedure of subsection 4.1. 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 the following three conditions holds:

(1) ω is closed, $d\omega = 0$.

(2) The associated vector bundle homomorphism $\check{\omega} : TM \to T^*M$ is injective.

(3) The gradient of ω with respect to itself exists and is smooth; this can be expressed most easily in charts, so let M be open in a convenient vector space E. Then for $x \in M$ and $X, Y, Z \in T_xM = E$ we have $d\omega(x)(X)(Y, Z) = \omega(\Omega_x(Y, Z), X) = \omega(\tilde{\Omega}_x(X, Y), Z)$ for smooth $\Omega, \tilde{\Omega} : M \times E \times E \to E$ which are bilinear in $E \times E$.

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.

4.3. Every cotangent bundle T^*Q, viewed as a manifold, carries a canonical weak symplectic structure $\omega_Q \in \Omega^2(T^*Q)$, which is defined as follows. Let $\pi_Q^* : T^*Q \to Q$ be the projection. Then the *Liouville form* $\theta_Q \in \Omega^1(T^*Q)$ is given by $\theta_Q(X) = \langle \pi_{T^*Q}(X), T(\pi_Q^*)(X)\rangle$ for $X \in T(T^*Q)$, where $\langle\ ,\ \rangle$ denotes the duality pairing $T^*Q \times_Q TQ \to \mathbb{R}$. Then the symplectic structure on T^*Q is given by $\omega_Q = -d\theta_Q$, 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^*Q is strong if and only if all model spaces of the manifold Q are reflexive and Hilbert spaces.

4.4. 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 *ω-smooth cotangent space* with respect to the symplectic structure ω of M at x in view of the embedding of test functions into distributions. The convenient structure on $T_x^\omega M$ is the one from T_xM. 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.

Note that only for strong symplectic structures the mapping $\check{\omega}_x : T_xM \to T_x^*M$ is a diffeomorphism onto $T_x^\omega M$ with the structure induces from T_x^*M.

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

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

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

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

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

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

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

4.6. 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) *$f \in C_\omega^\infty(E,\mathbb{R})$.*

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

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

4.7. 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 4.6 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 4.6.1 or 4.6.2 translates to a local differential condition describing the functions in $C^\infty_\omega(M,\mathbb{R})$.

4.8. Theorem. *Let (M,ω) be a weak symplectic manifold. The Hamiltonian mapping* $\operatorname{grad}^\omega : C^\infty_\omega(M,\mathbb{R}) \to \mathfrak{X}(M,\omega)$*, which is given by*

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

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 equip $C^\infty_\omega(M,\mathbb{R})$ with the initial structure with respect to the the two following mappings:

$$C^\infty_\omega(M,\mathbb{R}) \xrightarrow{\subset} C^\infty(M,\mathbb{R}), \qquad C^\infty_\omega(M,\mathbb{R}) \xrightarrow[\mathfrak{X}]{\operatorname{grad}^\omega} (M).$$

Then the Poisson bracket is bounded bilinear on $C^\infty_\omega(M,\mathbb{R})$.
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 4.6, that the Hamiltonian mapping $\operatorname{grad}^\omega$ is well defined and has values in $\mathfrak{X}(M,\omega)$, since by [**42**], 34.18.6 we have

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

By [**42**], 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 E equipped with a (non-constant) weak symplectic structure. So let $f,\ g\in C^\infty_\omega(M,\mathbb{R})$ and $X,Y,Z\in E$ then $\{f,g\}(x)=dg(x)(\operatorname{grad}^\omega f(x))$, and thus

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

We have $\operatorname{grad}^\omega f\in\mathfrak{X}(M,\omega)$ and for any $X\in\mathfrak{X}(M,\omega), Y\in\mathfrak{X}(M), y\in E$ the condition $\mathcal{L}_X\omega=0$ implies, using 4.2.3,

$$\begin{aligned}
0 &= (\mathcal{L}_X\omega)(Y,y) = (d\omega(X))(Y,y) - \omega([X,Y],y) - \omega(Y,[X,y])\\
&= \omega(\tilde\Omega(X,Y),y) - \omega([X,Y],y) + \omega(Y,dX(y_2)).
\end{aligned}$$

Again by 4.2.3 we have

$$\begin{aligned}
&d(\omega(\operatorname{grad}^\omega g,y)(\operatorname{grad}^\omega f) =\\
&= d\omega(\operatorname{grad}^\omega f)(\operatorname{grad}^\omega g,y) + \omega(d(\operatorname{grad}^\omega g)(\operatorname{grad}^\omega f),y)\\
&= \omega(\tilde\Omega(\operatorname{grad}^\omega f,\operatorname{grad}^\omega g),y) + \omega(d(\operatorname{grad}^\omega g)(\operatorname{grad}^\omega f),y)
\end{aligned}$$

Collecting all terms we get

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

So 4.6.2 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 4.6.1 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 [**42**], 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. □

4.9. 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 4.8:

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

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 $j : \mathfrak{g} \to (C^\infty_\omega(M), \{\quad,\quad\})$ 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

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

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

4.10. 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*

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

such that

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

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:

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

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

4.11. 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*

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

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.

4.12. 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. □

5. Right invariant weak Riemannian metrics on Lie groups

5.1. 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

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

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

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

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

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

For a variation $g(s, t)$ with fixed endpoints we have then, using the left Maurer-Cartan equation $d\kappa^r - \frac{1}{2}[\kappa^r, \kappa^r]$ from 3.7.4 and integration by parts,

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

where $\operatorname{ad}((g^*\kappa^r)(\partial_t))^\top : \mathfrak{g} \to \mathfrak{g}$ is the adjoint of $\operatorname{ad}((g^*\kappa^r)(\partial_t))$ 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 $u(t) := g^*\kappa^r(\partial_t) = T_{g(t)}(\mu^{g(t)^{-1}})\cdot g'(t)$, the geodesic equation has the expression:

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

This is, of course, just the Euler-Poincaré equation for right invariant systems using the Lagrangian given by the kinetic energy (see [**45**], section 13).

5.2. 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, $(\pi_G, \kappa^r) : TG \to G \times \mathfrak{g}$ 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

(1) $$[R_X, R_Y] = R(-[X, Y]_\mathfrak{g} + dY \cdot R_X - dX \cdot R_Y),$$

$$R^{-1}[R_X, R_Y] = -[X, Y]_{\mathfrak{g}} + R_X(Y) - R_Y(X),$$
$$\gamma_x(R_X(x), R_Y(x)) = \gamma(X(x), Y(x))\,,\; x \in G.$$

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 5.1.1 exists and is given for any $X, Y \in C^\infty(G, \mathfrak{g})$ in terms of the isomorphism R by*

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

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

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

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

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

5.3. 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

$$\begin{aligned}
\mathcal{R}(X, Y) &= [\nabla_X, \nabla_Y] - \nabla_{-[X,Y]_{\mathfrak{g}} + R_X(Y) - R_Y(X)}\\
&= [R_X + \tfrac{1}{2}\operatorname{ad}(X)^\top + \tfrac{1}{2}\alpha(X) - \tfrac{1}{2}\operatorname{ad}(X), R_Y + \tfrac{1}{2}\operatorname{ad}(Y)^\top + \tfrac{1}{2}\alpha(Y) - \tfrac{1}{2}\operatorname{ad}(Y)]\\
&\quad - R_{-[X,Y]_{\mathfrak{g}} + R_X(Y) - R_Y(X)} - \tfrac{1}{2}\operatorname{ad}(-[X, Y]_{\mathfrak{g}} + R_X(Y) - R_Y(X))^\top\\
&\quad - \tfrac{1}{2}\alpha(-[X, Y]_{\mathfrak{g}} + R_X(Y) - R_Y(X)) + \tfrac{1}{2}\operatorname{ad}(-[X, Y]_{\mathfrak{g}} + R_X(Y) - R_Y(X))\\
&= -\tfrac{1}{4}[\operatorname{ad}(X)^\top + \operatorname{ad}(X), \operatorname{ad}(Y)^\top + \operatorname{ad}(Y)]\\
&\quad + \tfrac{1}{4}[\operatorname{ad}(X)^\top - \operatorname{ad}(X), \alpha(Y)] + \tfrac{1}{4}[\alpha(X), \operatorname{ad}(Y)^\top - \operatorname{ad}(Y)]\\
&\quad + \tfrac{1}{4}[\alpha(X), \alpha(Y)] + \tfrac{1}{2}\alpha([X, Y]_{\mathfrak{g}}).
\end{aligned}$$

5.4 The sectional curvature. The obstruction against $\mathrm{ad}(\mathfrak{g})$-invariance of the inner product $\gamma = \langle\ \ ,\ \ \rangle$ on $\mathfrak{g}$ is $\gamma(\mathrm{ad}(X)Y, Z) + \gamma(Y, \mathrm{ad}(X)Z) = \gamma((\mathrm{ad}(X)^\top + \mathrm{ad}(X))Y, Z)$; thus the operator

$$\beta(X) := \mathrm{ad}(X)^\top + \mathrm{ad}(X) : \mathfrak{g} \to \mathfrak{g} \tag{1}$$

plays a distinguished role. By definition, it is self adjoint with respect to γ, i.e., $\gamma(\beta(X)Y, Z) = \gamma(Y, \beta(X)Z)$. For the Soboloev metrics on $\mathfrak{g} = \mathfrak{X}(M)$ it will turn out to be a tensor field.

In terms of β, the curvature 5.3.2 looks like

$$\begin{aligned} \mathcal{R}(X,Y) = &-\tfrac{1}{4}[\beta(X), \beta(Y)] + \tfrac{1}{4}[\beta(X), \alpha(Y)] + \tfrac{1}{4}[\alpha(X), \beta(Y)] \\ &- \tfrac{1}{2}[\mathrm{ad}(X), \alpha(Y)] - \tfrac{1}{2}[\alpha(X), \mathrm{ad}(Y)] \\ &+ \tfrac{1}{4}[\alpha(X), \alpha(Y)] + \tfrac{1}{2}\alpha([X,Y]_{\mathfrak{g}}). \end{aligned} \tag{2}$$

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

$$\begin{aligned} 4\gamma(\mathcal{R}(X,Y)X,Y) &= -\gamma(\beta(X)\beta(Y)X, Y) + \gamma(\beta(Y)\beta(X)X, Y) \\ &\quad + \gamma(\beta(X)\,\mathrm{ad}(X)^\top Y, Y) - \gamma(\mathrm{ad}(\beta(X)X)^\top Y, Y) \\ &\quad + \gamma(\mathrm{ad}(\beta(Y)X)^\top X, Y) - \gamma(\beta(Y)\,\mathrm{ad}(X)^\top X, Y) \\ &\quad - 2\gamma(\mathrm{ad}(X)\,\mathrm{ad}(X)^\top Y, Y) + 0 \\ &\quad - 2\gamma(\mathrm{ad}(\mathrm{ad}(Y)X)^\top X, Y) + 2\gamma(\mathrm{ad}(Y)\,\mathrm{ad}(X)^\top X, Y) \\ &\quad + \gamma(\mathrm{ad}(\mathrm{ad}(X)^\top Y)^\top X, Y) - \gamma(\mathrm{ad}(\mathrm{ad}(X)^\top X)^\top Y, Y) \\ &\quad + 2\gamma(\mathrm{ad}(X)^\top\,\mathrm{ad}(X)Y, Y) \\ &= -\gamma(\beta(Y)X, \beta(X)Y) + \gamma(\beta(X)X, \beta(Y)Y) \\ &\quad + \gamma(\mathrm{ad}(X)^\top Y, \beta(X)Y) - \gamma(Y, \mathrm{ad}(\beta(X)X)Y) \\ &\quad + \gamma(X, \mathrm{ad}(\beta(Y)X)Y) - \gamma(\beta(X)X, \beta(Y)Y) \\ &\quad - 2\gamma(\mathrm{ad}(X)^\top Y, \mathrm{ad}(X)^\top Y) \\ &\quad - 2\gamma(X, \mathrm{ad}(\mathrm{ad}(Y)X)Y) + 2\gamma(\mathrm{ad}(X)^\top X, \mathrm{ad}(Y)^\top Y) \\ &\quad + \gamma(X, \mathrm{ad}(\mathrm{ad}(X)^\top Y)Y) - \gamma(Y, \mathrm{ad}(\mathrm{ad}(X)^\top X)Y) \\ &\quad + 2\gamma(\mathrm{ad}(X)Y, \mathrm{ad}(X)Y) \\ &= -\gamma(\beta(Y)X, \beta(X)Y) + \gamma(\beta(X)X, \beta(Y)Y) \\ &\quad + \gamma(\mathrm{ad}(X)^\top Y, \beta(X)Y) + \gamma(\mathrm{ad}(Y)^\top Y, \beta(X)X) \\ &\quad - \gamma(\mathrm{ad}(Y)^\top X, \beta(Y)X) - \gamma(\beta(X)X, \beta(Y)Y) \\ &\quad - 2\gamma(\mathrm{ad}(X)^\top Y, \mathrm{ad}(X)^\top Y) \\ &\quad + 2\gamma(\mathrm{ad}(Y)^\top X, \mathrm{ad}(Y)X) + 2\gamma(\beta(X)X, \beta(Y)Y) \end{aligned}$$

$$
\begin{aligned}
&\quad - \gamma(\mathrm{ad}(Y)^\top X, \mathrm{ad}(X)^\top Y) + \gamma(\beta(Y)Y, \beta(X)X) \\
&\quad + 2\gamma(\mathrm{ad}(X)Y, \mathrm{ad}(X)Y) \\
&= -\gamma(\beta(Y)X, \beta(X)Y) + 4\gamma(\beta(X)X, \beta(Y)Y) \\
&\quad + \gamma(\mathrm{ad}(X)^\top Y, \beta(X)Y) \\
&\quad - \gamma(\mathrm{ad}(Y)^\top X, \beta(Y)X) \\
&\quad - 2\gamma(\mathrm{ad}(X)^\top Y, \mathrm{ad}(X)^\top Y) \\
&\quad - 2\gamma(\mathrm{ad}(Y)^\top X, \mathrm{ad}(X)Y) \\
&\quad - \gamma(\mathrm{ad}(Y)^\top X, \mathrm{ad}(X)^\top Y) \\
&\quad - 2\gamma(\mathrm{ad}(Y)X, \mathrm{ad}(X)Y) \\
&= -\gamma(\beta(Y)X, \beta(X)Y) + 4\gamma(\beta(X)X, \beta(Y)Y) \\
&\quad + \gamma(\beta(X)Y, \beta(X)Y) - \gamma(\mathrm{ad}(X)Y, \beta(X)Y) \\
&\quad - \gamma(\beta(Y)X, \beta(Y)X) + \gamma(\mathrm{ad}(Y)X, \beta(Y)X) \\
&\quad - 2\gamma(\beta(X)Y, \beta(X)Y) + 2\gamma(\beta(X)Y, \mathrm{ad}(X)Y) \\
&\quad + 2\gamma(\mathrm{ad}(X)Y, \beta(X)Y) - 2\gamma(\mathrm{ad}(X)Y, \mathrm{ad}(X)Y) \\
&\quad - \gamma(\beta(Y)X, \beta(X)Y) + \gamma(\mathrm{ad}(Y)X, \beta(X)Y) \\
&\quad - \gamma(\beta(Y)X, \mathrm{ad}(X)Y) - \gamma(\mathrm{ad}(Y)X, \mathrm{ad}(X)Y) \\
&= -2\gamma(\beta(Y)X, \beta(X)Y) + 4\gamma(\beta(X)X, \beta(Y)Y) \\
&\quad - \gamma(\beta(X)Y, \beta(X)Y) - \gamma(\beta(Y)X, \beta(Y)X) \\
&\quad + 2\gamma(\beta(X)Y, \mathrm{ad}(X)Y) \\
&\quad - 2\gamma(\beta(Y)X, \mathrm{ad}(X)Y) \\
&\quad - \gamma(\mathrm{ad}(X)Y, \mathrm{ad}(X)Y) \\
&= -\gamma\big(\beta(X)Y + \beta(Y)X, \beta(X)Y + \beta(Y)X\big) \\
&\quad + 4\gamma(\beta(X)X, \beta(Y)Y) \\
&\quad + 2\gamma(\beta(X)Y - \beta(Y)X, \mathrm{ad}(X)Y) \\
&\quad - \gamma(\mathrm{ad}(X)Y, \mathrm{ad}(X)Y) \\
&= -\gamma\big(\beta(X)Y - \beta(Y)X, \beta(X)Y - \beta(Y)X\big) - 4\gamma(\beta(X)Y, \beta(Y)X) \\
&\quad + 4\gamma(\beta(X)X, \beta(Y)Y) \\
&\quad + 2\gamma(\beta(X)Y - \beta(Y)X, \mathrm{ad}(X)Y) \\
&\quad - \gamma(\mathrm{ad}(X)Y, \mathrm{ad}(X)Y) \\
&= -\gamma\big(\beta(X)Y - \beta(Y)X, \beta(X)Y - \beta(Y)X\big) \\
&\quad + 2\gamma(\beta(X)Y - \beta(Y)X, \mathrm{ad}(X)Y) - \gamma(\mathrm{ad}(X)Y, \mathrm{ad}(X)Y) \\
&\quad - 4\gamma(\beta(X)Y, \beta(Y)X) + 4\gamma(\beta(X)X, \beta(Y)Y)
\end{aligned}
$$

$$
\begin{aligned}
&= -\big\|\beta(X)Y - \beta(Y)X - \operatorname{ad}(X)Y\big\|_\gamma^2 \\
&\quad - 4\gamma(Y, \beta(X)\beta(Y)X) + 4\gamma(\beta(Y)\beta(X)X, Y) \\
&= -\big\|\beta(X)Y - \beta(Y)X - \operatorname{ad}(X)Y\big\|_\gamma^2 - 4\gamma([\beta(X), \beta(Y)]X, Y)
\end{aligned}
$$

If one prefers Arnold's original formula using only $\operatorname{ad}(X)^\top Y$ instead of $\beta(X)Y$, this becomes

$$
\begin{aligned}
4\gamma(\mathcal{R}(X,Y)X,Y) &= -\big\|\beta(X)Y - \beta(Y)X - \operatorname{ad}(X)Y\big\|_\gamma^2 \\
&\quad - 4\gamma(\beta(Y)X, \beta(X)Y) + 4\gamma(\beta(X)X, \beta(Y)Y) \\
&= -\big\|\operatorname{ad}(X)^\top Y - \operatorname{ad}(Y)^\top X + \operatorname{ad}(X)Y\big\|_\gamma^2 \\
&\quad - 4\gamma(\operatorname{ad}(Y)^\top X + \operatorname{ad}(Y)X, \operatorname{ad}(X)^\top Y + \operatorname{ad}(X)Y) \\
&\quad + 4\gamma(\operatorname{ad}(X)^\top X, \operatorname{ad}(Y)^\top Y)
\end{aligned}
$$

From [**?**, 3.3] we have

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

Putting $Z = X$ and $U = Y$ we get

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

which agrees with the above.

5.5. 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 5.3.1, this identity implies

$$
u_{st} = u_{ts} = \partial_s u_t = -\partial_s(\operatorname{ad}(u)^\top u) = -\operatorname{ad}(u_s)^\top u - \operatorname{ad}(u)^\top u_s
$$

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

Finally we get the Jacobi equation as

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

$$
\text{(1)}\qquad y_{tt} = [\operatorname{ad}(y)^\top + \operatorname{ad}(y), \operatorname{ad}(u)^\top] u - \operatorname{ad}(u)^\top y_t - \alpha(u) y_t + \operatorname{ad}(u) y_t\,.
$$

5.6. 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 5.5.1. First note that from 5.2.2 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:

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

In the second line of the last expression we use

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

and similar forms for the other two terms to get:

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

where in the last line we also used $\mathrm{ad}(u)u = 0$. We now compute the curvature term using 5.3.2:

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

Summing up we get

$$\begin{aligned}
\nabla_{\partial_t}\nabla_{\partial_t} y + \mathcal{R}(y,u)u = y_{tt} &+ \mathrm{ad}(u)^\top y_t + \alpha(u)y_t - \mathrm{ad}(u)y_t \\
&- \tfrac{1}{2}[\mathrm{ad}(y)^\top + \mathrm{ad}(y), \mathrm{ad}(u)^\top]u \\
&+ \tfrac{1}{2}[\alpha(u), \mathrm{ad}(y)]u + \tfrac{1}{2}\,\mathrm{ad}(u)^\top\,\mathrm{ad}(y)u\,.
\end{aligned}$$

Finally we need the following computation using 5.3.1:

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

Inserting we get the desired result:

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

5.7. 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 3.11 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 2.4 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. [**38**], II, p.70. It is given by the following expression which is constant in time t:

$$\begin{aligned}
\omega(y,z) :&= \langle y, \nabla_{\partial_t} z\rangle - \langle \nabla_{\partial_t} y, z\rangle \\
&= \langle y, z_t + \tfrac12 \operatorname{ad}(u)^\top z + \tfrac12 \alpha(u)z - \tfrac12 \operatorname{ad}(u)z\rangle \\
&\qquad - \langle y_t + \tfrac12 \operatorname{ad}(u)^\top y + \tfrac12 \alpha(u)y - \tfrac12 \operatorname{ad}(u)y, z\rangle \\
&= \langle y, z_t\rangle - \langle y_t, z\rangle + \langle [u,y], z\rangle - \langle y, [u,z]\rangle - \langle [y,z], u\rangle \\
&= \langle y, z_t - \operatorname{ad}(u)z + \tfrac12\alpha(u)z\rangle - \langle y_t - \operatorname{ad}(u)y + \tfrac12 \alpha(u)y, z\rangle.
\end{aligned}$$

It is worth while to check directly from the Jacobi field equation 5.5.1 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.

6. The Hamiltonian approach

6.1. 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,

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

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

$$\begin{aligned}
\theta_{G\times\mathfrak{g}^*}(\xi_g, \alpha, \beta) &:= (((\pi, \kappa^r)^{-1})^*\theta_G)_{(g,\alpha)}(\xi_g, \alpha, \beta)\\
&= \theta_G(T_{(g,\alpha)}(\pi, \kappa^r)^{-1}(\xi_g, \alpha, \beta))\\
&= \Big\langle \pi_{T^*G}(T_{(g,\alpha)}(\pi, \kappa^r)^{-1}(\xi_g, \alpha, \beta)), T(\pi)(T_{(g,\alpha)}(\pi, \kappa^r)^{-1}(\xi_g, \alpha, \beta))\Big\rangle\\
&= \Big\langle (\pi, \kappa^r)^{-1}(\pi_G, \pi_{\mathfrak{g}^*})(\xi_g, \alpha, \beta), T(\pi \circ (\pi, \kappa^r)^{-1})(\xi_g, \alpha, \beta))\Big\rangle\\
&= \Big\langle (\pi, \kappa^r)^{-1}(g, \alpha), T(\mathrm{pr}_1)(\xi_g, \alpha, \beta))\Big\rangle = \Big\langle T_g(\mu^{g^{-1}})^*\alpha, \xi_g\Big\rangle\\
&= \langle \alpha, T_g(\mu^{g^{-1}})\xi_g\rangle = \langle \alpha, \kappa^r(\xi_g)\rangle \qquad (2)
\end{aligned}$$

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

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

$$(3) \qquad = \langle \beta_2, X_1\rangle - \langle \beta_1, X_2\rangle - \langle \alpha, [X_1, X_2]\rangle$$

6.2. 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 $\mathfrak{g}$,

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

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:

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

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:

$$\begin{aligned}
&\gamma : G\times\mathfrak{g}\to G\times\mathfrak{g}^*, (g,X)\mapsto (g,\gamma(X))\\
&(\gamma^*\omega)_{(g,X)}((T(\mu^g).X_1, X, Y_1), (T(\mu^g)X_2, X, Y_2))\\
&\qquad = \omega_{(g,\gamma(X))}((T(\mu^g).X_1, \gamma(X), \gamma(Y_1)), (T(\mu^g)X_2, \gamma(X), \gamma(Y_2)))\\
&\qquad = \langle \gamma(Y_2), X_1\rangle - \langle\gamma(Y_1), X_2\rangle - \langle\gamma(X), [X_1, X_2]\rangle\\
(1)\quad &\qquad = \gamma(Y_2, X_1) - \gamma(Y_1, X_2) - \gamma(X, [X_1, X_2])
\end{aligned}$$

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:

$$\begin{aligned}
&(\gamma^*\omega)_{(g,X)}\Big(\operatorname{grad}^{\gamma^*\omega}(f)(g,X), (T(\mu^g)X_2, X, Y_2)\Big) = df(T(\mu^g)X_2; X, Y_2)\\
&\quad = d_1f(g,X)(T(\mu^g)X_2) + d_2f(g,X)(Y_2)\\
&\quad = \gamma(\kappa^r(\operatorname{grad}^\gamma_1(f)(g,X)), X_2) + \gamma(\operatorname{grad}^\gamma_2(f)(g,X), Y_2)\\
&\quad = \gamma(X_1, Y_2) + \gamma(-Y_1 - \operatorname{ad}(X_1)^\top X, X_2) \qquad \text{by (1).}
\end{aligned}$$

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

(2)

$$\begin{aligned}
&\operatorname{grad}^{\gamma^*\omega}(f)(g,X) =\\
&\big(T(\mu^g)\operatorname{grad}^\gamma_2(f)(g,X), X, -\operatorname{ad}(\operatorname{grad}^\gamma_2(f)(g,X))^\top X - \kappa^r(\operatorname{grad}^\gamma_1(f)(g,X))\big)
\end{aligned}$$

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

$$(3)\qquad \operatorname{grad}^{\gamma^*\omega}(\tfrac12\|\ \ \|^2_\gamma)(g,X) = (T(\mu^g)X; X, -\operatorname{ad}(X)^\top X)$$

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

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

which reproduces the geodesic equation from 5.1.

6.3. The momentum mapping. Under the assumptions of 6.2, 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:

$$T(\mu^g): TG \to TG,$$
$$(\pi,\kappa^r)T(\mu^g)T(\mu^h)X = (\pi,\kappa^r)T(\mu^{hg})X = (h.g, X), \quad (h,X) \mapsto (hg, X)$$
$$\zeta_X^{G\times\mathfrak{g}}(h,Y) = \partial_t|_0(h.\exp(tX), Y) = (T(\mu_h)X, 0_Y) \in TG \times T\mathfrak{g} \tag{1}$$

Consider now the diagram from 4.1 in the case of the weak symplectic manifold $(M = G\times\mathfrak{g}, \gamma^*\omega)$:

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

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

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

Thus the momentum mapping is

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

$$\begin{aligned}
&J(h,Y) = \gamma(\mathrm{Ad}(h)^\top Y) \in \mathfrak{g}*\\
\bar J &:= \gamma^{-1}\circ J : G\times\mathfrak{g}\to\mathfrak{g},\\
(2)\qquad &\bar J(h,Y) = \mathrm{Ad}(h)^\top Y \in \mathfrak{g}.
\end{aligned}$$

(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,

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

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

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

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

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

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

$$\iff 0 = 0 - \gamma(Y_1, Z) - \gamma(Y, [X_1, Z]) = -\gamma(Y_1 + \operatorname{ad}(X_1)^\top Y, Z) \ \forall\, Z \in \mathfrak{g}$$
$$\iff Y_1 = -\operatorname{ad}(X_1)^\top Y.$$

(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 $\gamma(\bar J, X) = \langle J, X\rangle = j(X)$.

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

6.4. 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 6.2.4. 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 3.7.3 we have

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

6.5. Symplectic reduction to transposed adjoint orbits. Under the assumptions of 6.2 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:

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

(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 6.2.1 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

$$T_{(h,\operatorname{Ad}(h^{-1})^\top X)} \bar J^{-1}(X) = \{\partial_t|_0(\exp(tZ).h, \operatorname{Ad}((\exp(tZ).h)^{-1})^\top X) : Z \in \mathfrak{g}\}$$

$$= \{(T(\mu^h)Z, -\operatorname{ad}(Z)^\top \operatorname{Ad}(h^{-1})^\top X) : Z \in \mathfrak{g}\} \subset T_h G \times \mathfrak{g}$$
$$= \{(T(\mu^h)Z, -\operatorname{ad}(Z)^\top Y : Z \in \mathfrak{g}\} \subset T_h G \times \mathfrak{g}.$$

Since the infinitesimal action is $\zeta_Z^{G\times\mathfrak{g}}(h,Y) = (T(\mu_h)Z, Y, 0)$ we consider one of the tangent vectors in $T_{(h,Y)}\bar{J}^{-1}(X)$ in the form

$$(T(\mu^h)\operatorname{Ad}(h)Z_1, Y, -\operatorname{Ad}(h^{-1})^\top \operatorname{ad}(Z_1)^\top X),$$

where we used

$$\begin{aligned}
&\gamma(\operatorname{ad}(\operatorname{Ad}(h)Z_1)^\top \operatorname{Ad}(h^{-1})^\top X, U) = \gamma(\operatorname{Ad}(h^{-1})^\top X, [\operatorname{Ad}(h)Z_1, U]) \\
&= \gamma(X, \operatorname{Ad}(h^{-1})[\operatorname{Ad}(h)Z_1, U]) = \gamma(X, [Z_1, \operatorname{Ad}(h^{-1})U]) \\
&= \gamma(\operatorname{ad}(Z_1)^\top X, \operatorname{Ad}(h^{-1})U) = \gamma(\operatorname{Ad}(h^{-1})^\top \operatorname{ad}(Z_1)^\top X, U).
\end{aligned}$$

The second tangent vector $(T(\mu^h)Z, Y, -\operatorname{ad}(Z)^\top \operatorname{Ad}(h^{-1})^\top X)$ we take in standard form in $T_{(h,Y)}\bar{J}^{-1}(X)$. From 6.2.1, we get

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

and if this is the case then the first tangent vector is of the form

$$\begin{aligned}
(T(\mu^h)\operatorname{Ad}(h)Z_1, Y, -\operatorname{Ad}(h^{-1})^\top \operatorname{ad}(Z_1)^\top X) &= (T(\mu_h)Z_1, Y, 0) \\
&= \zeta_{Z_1}^{G\times\mathfrak{g}}(h, Y) \qquad \square
\end{aligned}$$

(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 transposed adjoint orbit* $\operatorname{Ad}(G)^\top X \subset \mathfrak{g}$ *with symplectic form the pullback via* $\gamma : \mathfrak{g} \to \mathfrak{g}^*$ *of the Kostant Kirillov Souriou form*

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

which is given by

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

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

$$\langle \gamma \operatorname{ad}(Y)^\top Z, U\rangle = \gamma(\operatorname{ad}(Y)^\top Z, U) = \gamma(Z, \operatorname{ad}(Y)U) = \\ = \langle \gamma(Z), \operatorname{ad}(Y)U\rangle = \langle \operatorname{ad}(Y)^* \gamma(Z), U\rangle.$$

The quotient space is $\bar{J}^{-1}(X)/G_X = \{(h.G_X, \operatorname{Ad}(h^{-1})^\top X) : h \in G\} \cong \operatorname{Ad}(G)^\top X \cong G/G_X$. 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: By 6.2.1 we have for a point $(h, Z = \operatorname{Ad}(h^{-1})^\top X) \in \bar{J}^{-1}(X)$ and $(T(\mu^h)Y_i, -\operatorname{ad}(Y_i)^\top Z) \in T_{(h,Z)}\bar{J}^{-1}(X)$ that

$$\begin{aligned}
&(\gamma^*\omega)_{(h,Z)}((T(\mu^h).Y_1, -\operatorname{ad}(Y_1)^\top Z), (T(\mu^h)Y_2, -\operatorname{ad}(Y_1)^\top Y)) \\
&\qquad = \gamma(-\operatorname{ad}(Y_2)^\top Z, Y_1) - \gamma(-\operatorname{ad}(Y_1)^\top Z, Y_2) - \gamma(Z, [Y_1, Y_2]) \\
&\qquad = \gamma(Z, [Y_1, Y_2]) = \omega_Z(\operatorname{ad}(Y_1)^\top Z, \operatorname{ad}(Y_2)^\top Z) \quad \square
\end{aligned}$$

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

Geometry of right invariant metrics on diffeomorphism groups

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

This section is based on [**55**].

7.1 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 [**61**], section 38. This is proved in [**42**], 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 [**26**], [**73**], [**46**]. 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 $X = h \circ \varphi^{-1}, Y = k \circ \psi^{-1} \in \mathfrak{X}_c(N)$:

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

7.2. Geodesics and sectional curvature for γ^0 on $\operatorname{Diff}(N)$**.** According to 5.1, 5.3, or 6.4, 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.

7.3 Example: n-dimensional analog of Burgers' equation. For $(N,g) = (\mathbb{R}^n, \text{can})$ or $((S^1)^n, \text{can})$ we have:

$$(\operatorname{ad}(X)Y)^k = \sum_i ((\partial_i X^k)Y^i - X^i(\partial_i Y^k))$$

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

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.

7.4. 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

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

Since ψ is a diffeomorphism with compact support, the two suprema are bounded. Thus $\psi^{-1}\circ\varphi_1\circ\psi \in \mathrm{Diff}_0(N)^{E=0}$.

(7) *We claim that* $\mathrm{Diff}_0(N)^{E=0}$ *is a non-trivial subgroup.* In view of the simplicity of $\mathrm{Diff}_0(N)$ mentioned in 7.1 this concludes the proof.

It remains to find a non-trivial diffeomorphism in $\mathrm{Diff}_0(N)^{E=0}$. 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:

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

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

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

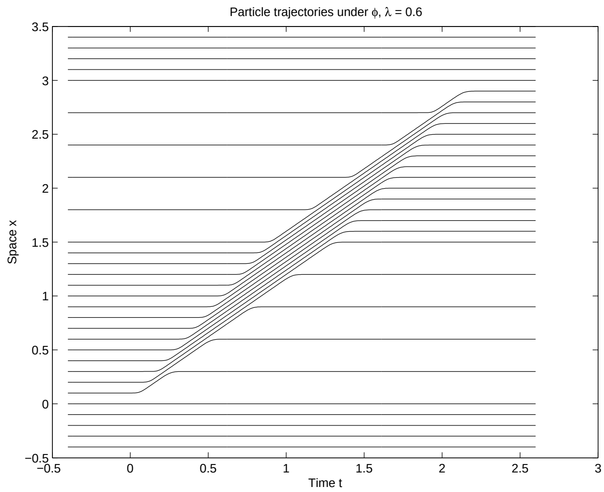
and the whole second term is bounded by $2\varepsilon(t_1-t_0)$. Thus the length is $O(\varepsilon)$.

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 $\operatorname{Diff}_0(N)^{E=0}$ 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 $\operatorname{Diff}_0(N)^{E=0}$ again. □



Remark. Theorem 7.4 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 7.4.7 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.

7.5. 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:

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

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

Locally, on orientable pieces of N, we have:

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

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

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

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

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. □

7.6. Geodesics and sectional curvature for G^A on $\operatorname{Diff}(\mathbb{R})$. We consider the groups $\operatorname{Diff}_c(\mathbb{R})$ or $\operatorname{Diff}(S^1)$ 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:

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

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 9.3.4. Note that here geodesic distance is a well defined metric describing the topology.

8. The regular Lie group of rapidly decreasing diffeomorphisms

8.1. Lemma. *For smooth functions of one variable we have:*

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

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 [**27**], the only beatified mathematician.

Proof. We compose the Taylor expansions of

$$f(g(x)+h):\quad j^\infty_{g(x)}f(h) = \sum_{m\ge 0}\frac{f^{(m)}(g(x))}{m!}h^m,$$

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

Or we use the multinomial expansion

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

to get

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

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

8.2. 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 $n\in\mathbb{N}_{\geq 0}$, 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:

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

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 [**42**], 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 $x \in \mathbb{R}$.

8.3. 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 2.3 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 $c^k(t) = \partial_t^k f(t, \quad)$. □

8.4. 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 8.1 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 $t \mapsto \mathrm{Id} + f(t, \quad), \mathrm{Id} + g(t, \quad)$ are in $C^\infty(\mathbb{R}, \mathrm{Diff}_{\mathcal{S}}(\mathbb{R})_0)$ which means that the functions $f, g \in C^\infty(\mathbb{R}^2, \mathbb{R})$ satisfy the conditions of lemma 8.2. 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 8.1 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 8.3 for f and g. Thus the multiplication respects smooth curves and is smooth.

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

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

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 $\mathrm{Diff}_{\mathcal{S}}(\mathbb{R})$ 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 8.3, 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 8.3 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

$$
\begin{aligned}
\partial_t^n \partial_x^m (f(t, x+g(t,x))) &= (\partial^{(n,m)} f)(t, x+g(t,x))(1+\partial_x g(t,x))^m + \\
&\quad + \sum_{\substack{k_1 \leq n \\ k_2 \leq m+n}} (\partial^{(k_1,k_2)} f)(t, x+g(t,x)) a_{k_1,k_2}(t,x),
\end{aligned}
$$

where $a_{k_1,k_2}(t,x) = O(\frac{1}{(1+|x|^2)^k})$ uniformly in x and locally in t. Again $1 + \partial_x g(t,x) \geq \varepsilon > 0$, 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 8.3. So the inversion is smooth and $\mathrm{Diff}_S(\mathbb{R})$ is a Lie group.

We claim that $\operatorname{Diff}_{\mathcal{S}}(\mathbb{R})$ 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 8.3. The evolution of this time dependent vector field is the function given by the ODE

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

We have to show that f satisfies 8.3. For $0 \le t \le C$ we consider

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

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. □

8.5. 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 8.2:

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

We shall use the norm

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

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

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

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

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

8.6. 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 $\mathcal{F} \circ A_k \circ \mathcal{F}^{-1} : (c_n) \mapsto ((1 + (2\pi n)^2 + \cdots + (2\pi n)^{2k})\, c_n)$ 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 . □

8.7. Sobolev inequality. *We have bounded linear embeddings* $(0 < \alpha \leq 1)$*:*

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

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:

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

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

$$e^{ix\xi} - e^{iy\xi} = (x-y)\int_0^1 i\xi e^{i(y+t(x-y))\xi} dt,$$

$$\left|e^{ix\xi} - e^{iy\xi}\right| \le |x-y|.|\xi|$$

to obtain

$$\begin{aligned}
&2\pi\left|\frac{u(x)-u(y)}{(x-y)^\alpha}\right| \le \int \left|\frac{e^{ix\xi} - e^{iy\xi}}{x-y}\right|^\alpha . \left|e^{ix\xi} - e^{iy\xi}\right|^{1-\alpha} |\hat u(\xi)|\,d\xi \\
&\le 2\int |\hat u(\xi)|(1+|\xi|)^n \frac{|\xi|^\alpha}{(1+|\xi|)^n}\,d\xi \\
&\le 2\left(\int |\hat u(\xi)|^2(1+|\xi|)^{2n} d\xi\right)^{\frac12}\left(\int \frac{|\xi|^{2\alpha}}{(1+|\xi|)^{2n}}\,d\xi\right)^{\frac12} = C_1\|u\|_{H^n}
\end{aligned}$$

where C_1 depends only on $n-\alpha > \frac12$. □

8.8. Banach algebra property. *If $n > \frac12$ 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.

More is true: If $n > \frac12$ and $r \ge -n$ then pointwise multiplication $\mathcal S(\mathbb R)\times\mathcal S(\mathbb R)\to\mathcal S(\mathbb R)$ extends to a bounded bilinear mapping $H^n(\mathbb R)\times H^r(\mathbb R)\to H^{\min\{n,r\}}(\mathbb R)$.

See [**25**] 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 $f^{(l)}\in C^0_b(\mathbb R)$ 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.

For the extended statement we have to use another proof. We will check the Sobolev H^s-norm for $f.g$ where $s=\min\{n,r\}$. We have

$$\begin{aligned}
\widehat{f.g}(\xi)(1+|\xi|)^s &= (\hat f * \hat g)(\xi)(1+|\xi|)^s \\
&= \int \hat f(\xi-\eta)\hat g(\eta)\,d\eta\cdot(1+|\xi|)^s. \\
\frac{(1+|\xi|)^s}{(1+|\xi-\eta|)^n(1+|\eta|)^r} &= O\Big(\frac{(1+|\xi|)^s}{(1+|\xi|)^n}\Big) \quad \text{since } r\ge -n \\
&= O(1) \quad \text{if } s\le n. \text{ Thus we have}
\end{aligned}$$

$$\begin{aligned}\|\widehat{f.g}(1+|\quad|)^s\|_{L^2} &\le C\|(\hat f(1+|\quad|)^n)*(\hat g(1+|\quad|)^r)\|_{L^2}\\ &\le C\|\hat f(1+|\quad|)^n\|_{L^1}\cdot\|\hat g^2(1+|\quad|)^{2r}\|_{L^2}\\ &= C\|f\|^2_{H^n}\cdot\|g\|^2_{H^r},\end{aligned}$$

??? where we used that $L^1(\mathbb{R})$ is a Banach algebra for multiplication. $\square$

8.9. Differentiability of composition. *If $n>0$ then composition $\mathcal{S}(\mathbb{R})\times(\mathrm{Id}_{\mathbb{R}}+\mathcal{S}(\mathbb{R}))\to\mathcal{S}(\mathbb{R})$ 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 8.1 we have

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

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:

8.10. 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 $\|f - f(\quad - t.g)\|_{HC^n} \geq \|f - f(\quad - t)\|_{L^2} = \sqrt{2}$. $\square$

8.11. The topological group $\mathrm{Diff}(\mathbb{R})$**.** 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 8.9. So we can use the implicit function theorem for the equation $f \circ f^{-1} = \mathrm{Id}$. $\square$

8.12 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. $\square$

8.13 Proposition. *For the operator* $A_k = \sum_{i=0}^k (-1)^i \partial_x^{2i}$ *and for* $n \geq 2k$*, 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. □

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

9.1. 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 [**17**]:

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

Using the above relations and the general curvature formula 5.3.2, we get

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

Sectional curvature is non-negative and unbounded:

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

Let us check invariance of the momentum mapping $\bar J$ from 6.3:

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

$$\bar J(g,X) = \operatorname{sign}(g_x).(g_x)^2(X\circ g). \qquad (8)$$

Along a geodesic $t\mapsto g(t,\quad)$, according to 5 and 6.3, the momentum

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

This is what we found in 1.3 by chance.

9.2. 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 5.5.1, which in our case becomes:

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

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 $C^\infty_c(\mathbb{R},\mathbb{R})^2$, respectively. The weak symplectic structure on it is given by 5.7:

$$\begin{aligned}\omega(y,z) &= \langle y, z_t - \tfrac12 u_x z + 2u z_x\rangle - \langle y_t - \tfrac12 u_x y + 2u y_x, z\rangle\\ &= \int_{S^1 \text{or } \mathbb{R}} (y z_t - y_t z + 2u(y z_x - y_x z))\,dx. \qquad (2)\end{aligned}$$

9.3. 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

$$\gamma(X,Y) = \langle X,Y\rangle_k = \sum_{i=0}^k \int (\partial_x^i X)(\partial_x^i Y)\,dx = \int A_k(X)(Y)\,dx$$

$$= \int X A_k(Y)\,dx, \qquad \text{where} \qquad A_k = \sum_{i=0}^k (-1)^i \partial_x^{2i} \tag{1}$$

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 8.5. 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

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

which in turn gives rise to

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

Thus the geodesic equation is

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

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

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

which is the dispersionfree version of the *Camassa-Holm equation*, see [**18**], [**63**], [**40**]. We met it already in 7.6, and will meet the full equation in 10.7. Let us check the invariant momentum mapping from 6.3.2:

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

$$= \operatorname{sign}(g')\Big\langle A_k^{-1}\Big((g')^2(A_k(X)\circ g)\Big), Y\Big\rangle_k$$

(5)

$$\bar J(g,X) = \operatorname{sign}(g_x).A_k^{-1}\Big((g_x)^2(A_k(X)\circ g)\Big).$$

Along a geodesic $t\mapsto g(t,\quad)$, by 3 and 6.3, the expressions

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

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

9.4. Theorem. *Let $k\geq 1$. There exists a HC^{2k+1}-open neighborhood V of $(\mathrm{Id},0)$ in $\operatorname{Diff}(S^1)\times\mathfrak{X}(S^1)$ such that for each $(g_0,u_0)\in V$ there exists a unique C^3 geodesic $g\in C^3((-2,2),\operatorname{Diff}(S^1))$ 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 $\operatorname{Diff}(S^1)$ by $\operatorname{Diff}_{\mathcal S(\mathbb R)}$ and $\mathfrak X(S^1)$ by $\mathfrak X_{\mathcal S}(\mathbb R) = \mathcal S(\mathbb R)\partial_x$.

This result is stated in [**21**], and also this proof follows essentially [**21**]. But there is a mistake in [**21**], p 795, where the authors assume that composition and inversion on $H^n(S^1)$ are smooth. This is wrong. One needs to use 8.12 and 8.13. The mistake was corrected in [**20**], for the more general case of the Virasoro group.

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

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

$$=: -uu_x + A_k^{-1}C_k(u),$$

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

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

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

Now consider the topological group and Banach manifold DiffHC^n described in 8.11.

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

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

Therefore,

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

By 8.12 and 8.13 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 $E_k : \operatorname{DiffHC}^n \times HC^n \to \operatorname{DiffHC}^n \times HC^{n-2k}$ is strongly C^1. This can be proved similarly, again using 8.12 and 8.13.

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 exist 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 $\mathrm{Fl}^n_t(g_0, v_0) = (g(t), v(t))$ 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

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

and consider the neighborhood $V_n := W_{2k+1} \cap \operatorname{DiffHC}^n \times HC^n$ 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 $n \geq 2k+1$. 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 $\operatorname{DiffHC}^{2k+1} \times HC^{2k+1}$, 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 9.3.6, the expression

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

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 8.1.

$$\begin{aligned}
u_x &= (g_{tx} \circ g^{-1})(g^{-1})_x = \frac{g_{tx}}{g_x} \circ g^{-1} \\
\partial_x^2 u &= \Big(\frac{\partial_x^2 g_t}{g_x^2} - g_{tx}\frac{\partial_x^2 g}{g_x^3}\Big) \circ g^{-1} \\
\partial_x(g^{-1}) &= \frac{1}{g_x} \circ g^{-1} \\
\partial_x^2(g^{-1}) \circ g &= -\frac{\partial_x^2 g}{g_x^3}
\end{aligned}$$

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

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 $t\in(-2,2)$:

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

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

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

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. □

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

10.1. 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 8. 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

$$c : \mathrm{Diff} \times \mathrm{Diff} \to \mathbb{R}$$
$$c(\varphi, \psi) := \frac{1}{2}\int \log(\varphi \circ \psi)' d\log \psi' = \frac{1}{2}\int \log(\varphi' \circ \psi) d\log \psi'$$

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:

$$\begin{aligned}
\int \log(\varphi \circ \psi)' d\log \psi' &= \int \log((\varphi' \circ \psi)\psi') d\log \psi' = \\
&= \int \log(\varphi' \circ \psi) d\log \psi' + \int \log(\psi') d\log \psi',
\end{aligned}$$

$$\int \log(\psi') d\log \psi' = \tfrac{1}{2}\int d\ \log(\psi')^2 = 0.$$

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

For the cocycle condition we add the following terms:

$$2c(\varphi_2, \varphi_3) = \int \log(\varphi_2' \circ \varphi_3) d\log \varphi_3'$$

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

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

$$= \int \log(\varphi_1' \circ \varphi_2) d \log \varphi_2' + \int \log(\varphi_1' \circ \varphi_2 \circ \varphi_3) d \log \varphi_3'$$

$$-2c(\varphi_1, \varphi_2) = - \int \log(\varphi_1' \circ \varphi_2) d \log \varphi_2' \qquad \square$$

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

10.2. The Virasoro Lie algebra. Let us compute the Lie algebra of the two versions of the the Virasoro-Bott group. Consider $\mathbb{R} \times_c \mathrm{Diff}$, where again Diff denotes any one of the groups $\mathrm{DiffHC}^+(S^1)$, $\mathrm{Diff}(\mathbb{R})_0$, or $\mathrm{Diff}_{\mathcal{S}}(\mathbb{R})$. So let $\varphi, \psi : \mathbb{R} \to \mathrm{Diff}$ with $\varphi(0) = \psi(0) = \mathrm{Id}$ 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:

$$\mathrm{Ad}\begin{pmatrix} \varphi(t) \\ \alpha(t) \end{pmatrix} \begin{pmatrix} Y \\ \beta'(0) \end{pmatrix} = \partial_s|_0 \begin{pmatrix} \varphi(t) \\ \alpha(t) \end{pmatrix} \begin{pmatrix} \psi(s) \\ \beta(s) \end{pmatrix} \begin{pmatrix} \varphi(t)^{-1} \\ -\alpha(t) \end{pmatrix}$$

$$= \partial_s|_0 \begin{pmatrix} \varphi(t) \circ \psi(s) \circ \varphi(t)^{-1} \\ \alpha(t) + \beta(s) + c(\varphi(t), \psi(s)) - \alpha(t) + c(\varphi(t) \circ \psi(s), \varphi(t)^{-1}) \end{pmatrix}$$

$$(1) \qquad = \begin{pmatrix} \varphi(t)_* Y = \mathrm{Ad}(\varphi(t)) Y \\ \beta_t(0) + \partial_s|_0 c(\varphi(t), \psi(s)) + \partial_s|_0 c(\varphi(t) \circ \psi(s), \varphi(t)^{-1}) \end{pmatrix}$$

$$\left[\begin{pmatrix} X \\ \alpha'(0) \end{pmatrix}, \begin{pmatrix} Y \\ \beta'(0) \end{pmatrix} \right] =$$

$$= \partial_t|0 \begin{pmatrix} \varphi(t)_* Y = \mathrm{Ad}(\varphi(t)) Y \\ \beta'(0) + \partial_s|_0 c(\varphi(t), \psi(s)) + \partial_s|_0 c(\varphi(t) \circ \psi(s), \varphi(t)^{-1}) \end{pmatrix}$$

$$(2) \qquad = \begin{pmatrix} -[X, Y] \\ \partial_t|_0 \partial_s|_0 c(\varphi(t), \psi(s)) + \partial_t|_0 \partial_s|_0 c(\varphi(t) \circ \psi(s), \varphi(t)^{-1}) \end{pmatrix}$$

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

$$2\partial_s|_0 c(\varphi(t), \psi(s)) = \partial_s|_0 \int \log(\varphi(t)_x \circ \psi(s))\, d \log(\psi(s)_x)$$

$$= \int_x \frac{(\varphi(t)_{xx}\circ\psi(0))Y}{\varphi(t)_x\circ\psi(0)}\,d\log(\underbrace{\psi(0)'}_{=1}) + \int \log(\varphi(t)_x)\,dY_x$$
$$= \int \log(\varphi(t)_x)Y_{xx}\,dx$$
$$2\partial_t|_0\partial_s|_0 c(\varphi(t),\psi(s)) = \partial_t|_0 \int \log(\varphi(t)_x)Y_{xx}\,dx = \int \frac{X_xY_{xx}}{\varphi(0)_x}dx = \int X_xY_{xx}dx.$$

For the second term we first check:

$$(\varphi^{-1})_x = \frac{1}{\varphi_x\circ\varphi^{-1}}, \qquad (\varphi^{-1})_{xx} = -\frac{\varphi_{xx}\circ\varphi^{-1}}{(\varphi_x\circ\varphi^{-1})^3},$$
$$\varphi^{-1}(x) = y, \qquad \frac{1}{\varphi_x\circ\varphi^{-1}}dx = dy$$
$$d\log((\varphi^{-1})_x) = -\frac{\varphi_{xx}\circ\varphi^{-1}}{(\varphi_x\circ\varphi^{-1})^2}dx = -\frac{\varphi_{xx}}{\varphi_x}dy$$

and continue to compute

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

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

$$\omega(X,Y) = \omega(X)Y = \int X'dY' = \int X'Y''dx = \tfrac12\int \det\begin{pmatrix}X' & Y'\\ X'' & Y''\end{pmatrix}dx,$$

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 [**29**]), 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:

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

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

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

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:

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

10.3. H^0-Geodesics on the Virasoro-Bott groups. We shall use the L^2-inner product on $\mathbb{R}\times_\omega \mathfrak{g}$, where $\mathfrak{g} = \mathfrak{X}(S^1), \mathfrak{X}_c(\mathbb{R}), \mathcal{S}(\mathbb{R})\partial_x$:

$$(1)\qquad \left\langle \begin{pmatrix} X \\ a \end{pmatrix}, \begin{pmatrix} Y \\ b \end{pmatrix} \right\rangle_0 := \int XY\,dx + ab.$$

Integrating by parts we get

$$\begin{aligned}
\left\langle \operatorname{ad}\begin{pmatrix}X\\a\end{pmatrix}\begin{pmatrix}Y\\b\end{pmatrix}, \begin{pmatrix}Z\\c\end{pmatrix}\right\rangle_0 &= \left\langle \begin{pmatrix}X'Y-XY'\\ \omega(X,Y)\end{pmatrix}, \begin{pmatrix}Z\\c\end{pmatrix}\right\rangle_0\\
&= \int (X'YZ - XY'Z + cX'Y'')\,dx\\
&= \int (2X'Z + XZ' + cX''')Y\,dx\\
&= \left\langle \begin{pmatrix}Y\\b\end{pmatrix}, \operatorname{ad}\begin{pmatrix}X\\a\end{pmatrix}^\top \begin{pmatrix}Z\\c\end{pmatrix}\right\rangle_0, \quad \text{where}\\
\operatorname{ad}\begin{pmatrix}X\\a\end{pmatrix}^\top\begin{pmatrix}Z\\c\end{pmatrix} &= \begin{pmatrix}2X'Z+XZ'+cX'''\\0\end{pmatrix}.
\end{aligned}$$

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

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

Formula 5.1.2 gives the H^0 geodesic equation on the Virasoro-Bott group:

$$\text{(2)}\qquad \begin{pmatrix}u_t\\a_t\end{pmatrix} = -\operatorname{ad}\begin{pmatrix}u\\a\end{pmatrix}^\top\begin{pmatrix}u\\a\end{pmatrix} = \begin{pmatrix}-3u_xu - au_{xxx}\\0\end{pmatrix} \qquad \text{where}$$

$$\begin{aligned}
\begin{pmatrix}u(t)\\a(t)\end{pmatrix} &= \partial_s\begin{pmatrix}\varphi(s)\\ \alpha(s)\end{pmatrix}.\begin{pmatrix}\varphi(t)^{-1}\\ -\alpha(t)\end{pmatrix}\Big|_{s=t}\\
&= \partial_s\begin{pmatrix}\varphi(s)\circ\varphi(t)^{-1}\\ \alpha(s)-\alpha(t)+c(\varphi(s),\varphi(t)^{-1})\end{pmatrix}\Big|_{s=t}\\
&= \begin{pmatrix}\varphi_t\circ\varphi^{-1}\\ \alpha_t - \int \frac{\varphi_{tx}\varphi_{xx}}{2\varphi_x^2}dx\end{pmatrix}
\end{aligned}$$

since we have

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

$$= \int \frac{\varphi_t(t)' \circ \varphi(t)^{-1}}{\varphi(t)' \circ \varphi(t)^{-1}} \left(-\frac{\varphi(t)'' \circ \varphi(t)^{-1}}{(\varphi(t)' \circ \varphi(t)^{-1})^2} \right) dx \qquad \text{by 10.2}$$
$$= -\int \frac{\varphi_t' \varphi''}{(\varphi')^2} dy = -\int \frac{\varphi_{tx}\varphi_{xx}}{\varphi_x^2} dx.$$

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 [**45**], section 13.8.

Let us compute the invariant momentum mapping from 6.3.2. First we need the transpose of the adjoint action 10.2.5:

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

Thus the invariant momentum mapping 6.3.2 turns out as

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

Along a geodesic $t \mapsto g(t, \quad) = \begin{pmatrix}\varphi(t, \quad)\\ \alpha(t)\end{pmatrix}$, according to 3 and 6.3, the momentum

(5)

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

is constant in t.

10.4. The curvature. The computation of the curvature at the identity element has been done independently by [**56**] and Misiolek [**61**]. 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 $\operatorname{ad}\binom{X}{a}^{\top}$, $\alpha\binom{X}{a}$, and $\operatorname{ad}\binom{X}{a}$ etc. given above into formula 5.3.2 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

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

which coincides with formula (2.3) in Misiolek [**61**]. This in turn leads to the following expression for the sectional curvature

$$\left\langle 4\mathcal{R}\left(\binom{X_1}{a_1},\binom{X_2}{a_2}\right)\binom{X_1}{a_1},\binom{X_2}{a_2}\right\rangle_0=$$

$$=\int\Big(4(X_1X_2''-X_1''X_2)X_1X_2+8(X_1X_2'-X_1'X_2)X_1'X_2$$
$$+2(a_1X_2^{(4)}-a_2X_1^{(4)})X_1X_2+10(a_1X_2'''-a_2X_1''')X_1'X_2$$
$$+18(a_1X_2''-a_2X_1'')X_1''X_2$$

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

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.

10.5. 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 5.5.1 which in our case looks as follows.

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

which leads to

$$
\begin{aligned}
(1)\qquad y_{tt} &= -u(4y_{tx} + 3uy_{xx} + ay_{xxxx}) - u_x(2y_t + 2ay_{xxx})\\
&\qquad - u_{xxx}(b_t + \omega(y, u) - 3ay_x) - ay_{txxx},\\
(2)\qquad b_{tt} &= \omega(u, y_t) + \omega(y, 3u_xu) + \omega(y, au_{xxx}).
\end{aligned}
$$

Equation 2 is equivalent to:

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

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 $(2')$:

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

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:

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

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 [**61**] 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.

10.6. 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 5.7:

$$\omega\left(\begin{pmatrix} y \\ b \end{pmatrix}, \begin{pmatrix} z \\ c \end{pmatrix}\right) = \left\langle \begin{pmatrix} y \\ b \end{pmatrix}, \begin{pmatrix} z_t \\ c_t \end{pmatrix} \right\rangle_0 - \left\langle \begin{pmatrix} y_t \\ b_t \end{pmatrix}, \begin{pmatrix} z \\ c \end{pmatrix} \right\rangle_0 + \left\langle \left[\begin{pmatrix} u \\ a \end{pmatrix}, \begin{pmatrix} y \\ b \end{pmatrix}\right], \begin{pmatrix} z \\ c \end{pmatrix} \right\rangle_0$$

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

where the constant C_1 relates to c as B_1 does to b, see 10.5.3 and 10.5.4.

10.7. 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.

$$
\begin{aligned}
(1) \qquad \left\langle \begin{pmatrix} X \\ a \end{pmatrix}, \begin{pmatrix} Y \\ b \end{pmatrix} \right\rangle_k &:= \int (XY + X'Y' + \dots + X^{(k)} Y^{(k)})\, dx + ab \\
&= \int A_k(X) Y\, dx + ab = \int X A_k(Y)\, dx + ab, \\
&\quad \text{where } A_k = \sum_{i=0}^k (-1)^i \partial_x^{2i} \text{ as in 9.3.1.}
\end{aligned}
$$

Integrating by parts we get

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

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

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

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

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

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

Formula 5.1.2 gives the geodesic equation on the Virasoro-Bott group:

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

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

as in 10.3.2 Thus a is a constant in time and the geodesic equation contains the equation from the Korteweg-de Vries hierarchy:

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

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*, [**21**], [**49**]. See 9.3.4 for the dispersionfree version.

Let us compute the invariant momentum mapping from 6.3.2. First we need the transpose of the adjoint action 10.2.5:

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

$$\operatorname{Ad}\begin{pmatrix}\varphi\\ \alpha\end{pmatrix}^{\top}\begin{pmatrix}Y\\ b\end{pmatrix}=\begin{pmatrix}A_k^{-1}((A_k(Y)\circ\varphi)(\varphi')^2+bS(\varphi))\\ b\end{pmatrix}$$

Thus the invariant momentum mapping 6.3.2 turns out as

(5)

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

Along a geodesic $t\mapsto g(t,\ \ )=\binom{\varphi(t,\ \ )}{\alpha(t)}$, according to 4 and 6.3, 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.

10.8. Theorem. [**20**] *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 $\operatorname{Diff}(S^1)$, $\mathfrak{X}(S^1)$, $\mathrm{DiffHC}^n(S^1)$, $HC^n(S^1)$, or $\operatorname{Diff}_{\mathcal{S}}(\mathbb{R})$, $\mathfrak{X}_{\mathcal{S}}(\mathbb{R})$, $\mathrm{DiffHC}^n(\mathbb{R})$, $HC^n(\mathbb{R})$, respectively.

Proof. For $u\in HC^n$, $n\ge 2k+1$, we have as in the proof of 9.4

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

where $B_k: HC^n\to HC^{n-2k}$ is a bounded quadratic operator. Recall from 10.7.4 that we have to solve (where a is a real constant)

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

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

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

Note that here we need $2k \geq 3$. In [**62**] this result was obtained for $k \geq 3/2$. We put

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

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

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

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

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

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

and by 8.12 and 8.13 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 $E_k : \mathrm{DiffHC}^n \times HC^n \to \mathrm{DiffHC}^n \times HC^{n-2k}$ 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 $\mathrm{DiffHC}^n \times HC^n$. 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 $\mathrm{DiffHC}^n \times HC^n$ 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

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

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

Claim. For any initial point $(g_0, v_0) \in V_n$ the solution $\mathrm{Fl}^n_t(g_0, v_0) = (g(t), v(t))$ 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 $n \geq 2k+1$. 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 9.3.6, the expression

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

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 8.1:

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

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 $t \in (-2, 2)$:

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

$$P_k(g, v) = \frac{Q_k(g, \partial_x g, \dots, \partial_x^{2k-1} g, v, \partial_x v, \dots, \partial_x^{2k-1} v)}{g_x^2}$$

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

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

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. $\square$

Geometry of shape spaces of plane curves

11. The manifold of immersed closed curves

11.1. Conventions. It is often convenient to use the identification $\mathbb{R}^2 \cong \mathbb{C}$, giving us:

$$\bar{x}y = \langle x, y\rangle + i\det(x, y), \quad \det(x, y) = \langle ix, y\rangle.$$

We shall use the following spaces of C^∞ (smooth) diffeomorphisms and curves, and we give the shorthand and the full name:

$\mathrm{Diff}(S^1)$, the regular Lie group (section 3) of all diffeomorphisms $S^1 \to S^1$ with its connected components $\mathrm{Diff}^+(S^1)$ of orientation preserving diffeomorphisms and $\mathrm{Diff}^-(S^1)$ of orientation reversing diffeomorphisms.

$\mathrm{Diff}_1(S^1)$, the subgroup of diffeomorphisms fixing $1 \in S^1$. We have diffeomorphically $\mathrm{Diff}(S^1) = \mathrm{Diff}_1(S^1) \times S^1 = \mathrm{Diff}_1^+(S^1) \times (S^1 \rtimes \mathbb{Z}_2)$.

$\mathrm{Emb} = \mathrm{Emb}(S^1, \mathbb{R}^2)$, the manifold of all smooth embeddings $S^1 \to \mathbb{R}^2$. Its tangent bundle is given by $T\,\mathrm{Emb}(S^1, \mathbb{R}^2) = \mathrm{Emb}(S^1, \mathbb{R}^2) \times C^\infty(S^1, \mathbb{R}^2)$.

$\mathrm{Imm} = \mathrm{Imm}(S^1, \mathbb{R}^2)$, the manifold of all smooth immersions $S^1 \to \mathbb{R}^2$. Its tangent bundle is given by $T\,\mathrm{Imm}(S^1, \mathbb{R}^2) = \mathrm{Imm}(S^1, \mathbb{R}^2) \times C^\infty(S^1, \mathbb{R}^2)$.

$\mathrm{Imm}_f = \mathrm{Imm}_f(S^1, \mathbb{R}^2)$, the manifold of all smooth free immersions $S^1 \to \mathbb{R}^2$, i.e., those with trivial isotropy group for the right action of $\mathrm{Diff}(S^1)$ on $\mathrm{Imm}(S^1, \mathbb{R}^2)$.

$B_e = B_e(S^1, \mathbb{R}^2) = \operatorname{Emb}(S^1, \mathbb{R}^2)/\operatorname{Diff}(S^1)$, the manifold of 1-dimensional connected submanifolds of $\mathbb{R}^2$, see 11.3.

$B_i = B_i(S^1, \mathbb{R}^2) = \operatorname{Imm}(S^1, \mathbb{R}^2)/\operatorname{Diff}(S^1)$, an infinite dimensional 'orbifold'; its points are, roughly speaking, smooth curves with crossings and multiplicities, see 11.5.

$B_{i,f} = B_{i,f}(S^1, \mathbb{R}^2) = \operatorname{Imm}_f(S^1, \mathbb{R}^2)/\operatorname{Diff}(S^1)$, a manifold, the base of a principal fiber bundle, see 11.4.3.

We want to avoid referring to a path in our infinite dimensional spaces like Imm or B_e as a curve, because it is then a 'curve of curves' and confusion arises when you refer to a curve. So we will always talk of *paths* in the infinite dimensional spaces, not curves. Curves will be in $\mathbb{R}^2$. Moreover, if $t \mapsto (\theta \mapsto c(t, \theta))$ is a path, its t-th curve will be denoted by $c(t) = c(t, \quad)$. By c_t we shall denote the derivative $\partial_t c$, and $c_\theta = \partial_\theta c$.

11.2. Length and curvature on $\operatorname{Imm}(S^1, \mathbb{R}^2)$**.** The volume form on S^1 induced by c is given by

$$\text{(1)} \qquad \operatorname{vol} : \operatorname{Emb}(S^1, \mathbb{R}^2) \to \Omega^1(S^1), \qquad \operatorname{vol}(c) = |c_\theta|\, d\theta$$

and its derivative is

$$\text{(2)} \qquad d\operatorname{vol}(c)(h) = \frac{\langle h_\theta, c_\theta\rangle}{|c_\theta|} d\theta.$$

We shall also use the *normal unit field*

$$n_c = i\frac{c_\theta}{|c_\theta|}.$$

The length function is given by

$$\text{(3)} \qquad \ell : \operatorname{Imm}(S^1, \mathbb{R}^2) \to \mathbb{R}, \qquad \ell(c) = \int_{S^1} |c_\theta|\, d\theta$$

and its differential is

$$\text{(4)} \qquad \begin{aligned} d\ell(c)(h) &= \int_{S^1} \frac{\langle h_\theta, c_\theta\rangle}{|c_\theta|} d\theta = -\int_{S^1} \Big\langle h, \frac{c_{\theta\theta}}{|c_\theta|} - \frac{\langle c_{\theta\theta}, c_\theta\rangle}{|c_\theta|^3} c_\theta \Big\rangle d\theta \\ &= -\int_{S^1} \langle h, \kappa(c).ic_\theta\rangle d\theta = -\int_{S^1} \langle h, n_c\rangle \kappa(c) \operatorname{vol}(c) \end{aligned}$$

The curvature mapping is given by

$$\text{(5)} \qquad \kappa : \operatorname{Imm}(S^1, \mathbb{R}^2) \to C^\infty(S^1, \mathbb{R}), \qquad \kappa(c) = \frac{\det(c_\theta, c_{\theta\theta})}{|c_\theta|^3} = \frac{\langle ic_\theta, c_{\theta\theta}\rangle}{|c_\theta|^3}$$

and is equivariant so that $\kappa(c \circ f) = \pm\kappa(c) \circ f$ for $f \in \operatorname{Diff}^{\pm}(S^1)$. Its derivative is given by

$$\text{(6)} \qquad d\kappa(c)(h) = \frac{\langle ih_\theta, c_{\theta\theta}\rangle}{|c_\theta|^3} + \frac{\langle ic_\theta, h_{\theta\theta}\rangle}{|c_\theta|^3} - 3\kappa(c)\frac{\langle h_\theta, c_\theta\rangle}{|c_\theta|^2}.$$

With some work, this can be shown to equal:

$$d\kappa(c)(h) = \frac{\langle h, c_\theta\rangle}{|c_\theta|^2}\kappa_\theta + \frac{\langle h, ic_\theta\rangle}{|c_\theta|}\kappa^2 + \frac{1}{|c_\theta|}\Big(\frac{1}{|c_\theta|}\big(\frac{\langle h, ic_\theta\rangle}{|c_\theta|}\big)_\theta\Big)_\theta. \tag{7}$$

To verify this, note that both the left and right hand side are equivariant with respect to $\mathrm{Diff}(S^1)$, hence it suffices to check it for constant speed parametrizations, i.e. $|c_\theta|$ is constant and $c_{\theta\theta} = \kappa|c_\theta|ic_\theta$. By linearity, it is enough to take the 2 cases $h = aic_\theta$ and $h = bc_\theta$. Substituting these into formulas (6) and (7), the result is straightforward.

11.3. The principal bundle of embeddings $\mathrm{Emb}(S^1, \mathbb{R}^2)$**.** We recall some basic results whose proof can be found in [**42**]:

(A) The set $\mathrm{Emb}(S^1, \mathbb{R}^2)$ *of all smooth embeddings* $S^1 \to \mathbb{R}^2$ *is an open subset of the Fréchet space* $C^\infty(S^1, \mathbb{R}^2)$ *of all smooth mappings* $S^1 \to \mathbb{R}^2$ *with the* C^∞*-topology. It is the total space of a smooth principal bundle* $\pi : \mathrm{Emb}(S^1, \mathbb{R}^2) \to B_e(S^1, \mathbb{R}^2)$ *with structure group* $\mathrm{Diff}(S^1)$*, the smooth regular Lie group group of all diffeomorphisms of* S^1*, whose base* $B_e(S^1, \mathbb{R}^2)$ *is the smooth Fréchet manifold of all submanifolds of* $\mathbb{R}^2$ *of type* S^1*, i.e., the smooth manifold of all simple closed curves in* $\mathbb{R}^2$*.* ([**42**], 44.1)

(B) This principal bundle admits a smooth principal connection described by the horizontal bundle whose fiber $\mathcal{N}_c$ *over* c *consists of all vector fields* h *along* c *such that* $\langle h, c_\theta\rangle = 0$*. The parallel transport for this connection exists and is smooth.* ([**42**], 39.1 and 43.1)

See 11.4.3 for a sketch of proof of the first part in a slightly more general situation. Here we want to sketch the use of the second part. Suppose that $t \mapsto (\theta \mapsto c(t, \theta))$ is a path in $\mathrm{Emb}(S^1, \mathbb{R}^2)$. Then $\pi \circ c$ is a smooth path in $B_e(S^1, \mathbb{R}^2)$. Parallel transport over it with initial value $c(0, \cdot)$ is a now a path f in $\mathrm{Emb}(S^1, \mathbb{R}^2)$ which is horizontal, i.e., we have $\langle f_t, f_\theta\rangle = 0$. This argument will play an important role below. In 11.5 below we will prove this property for general immersions.

11.4. Free immersions. The manifold $\mathrm{Imm}(S^1, \mathbb{R}^2)$ of all immersions $S^1 \to \mathbb{R}^2$ is an open set in the manifold $C^\infty(S^1, \mathbb{R}^2)$ and thus itself a smooth manifold. An immersion $c : S^1 \to \mathbb{R}^2$ is called *free* if $\mathrm{Diff}(S^1)$ acts freely on it, i.e., $c \circ \varphi = c$ for $\varphi \in \mathrm{Diff}(S^1)$ implies $\varphi = \mathrm{Id}$. We have the following results:

(1) *If* $\varphi \in \mathrm{Diff}(S^1)$ *has a fixed point and if* $c \circ \varphi = c$ *for some immersion* c *then* $\varphi = \mathrm{Id}$. This is ([**19**], 1.3).

(2) *If for* $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$ *there is a point* $x \in c(S^1)$ *with only one preimage then* c *is a free immersion.* This is ([**19**], 1.4). There exist free immersions without such points: Consider a figure eight consisting of two touching ovals,

and map S^1 to this by first transversing the upper oval 3 times and then the lower oval 2 times. This is a free immersion.

(3) **The manifold $B_{i,f}(S^1, \mathbb{R}^2)$.** ([**19**], 1.5) *The set* $\mathrm{Imm}_f(S^1, \mathbb{R}^2)$ *of all free immersions is open in* $C^\infty(S^1, \mathbb{R}^2)$ *and thus a smooth submanifold. The projection*

$$\pi : \mathrm{Imm}_f(S^1, \mathbb{R}^2) \to \frac{\mathrm{Imm}_f(S^1, \mathbb{R}^2)}{\mathrm{Diff}(S^1)} =: B_{i,f}(S^1, \mathbb{R}^2)$$

onto a Hausdorff smooth manifold is a smooth principal fibration with structure group $\mathrm{Diff}(S^1)$*. By* ([**42**], 39.1 and 43.1) *this fibration admits a smooth principal connection described by the horizontal bundle with fiber $\mathcal{N}_c$ consisting of all vector fields h along c such that $\langle h, c_\theta \rangle = 0$. This connection admits a smooth parallel transport over each each smooth curve in the base manifold.*

We might view $\mathrm{Imm}_f(S^1, \mathbb{R}^2)$ as the nonlinear Stiefel manifold of parametrized curves in $\mathbb{R}^2$ and consequently $B_{i,f}(S^1, \mathbb{R}^2)$ as the nonlinear Grassmannian of unparametrized simple closed curves.

Sketch of proof. See also [**19**] for a slightly different proof with more details. For $c \in \mathrm{Imm}_f(S^1, \mathbb{R}^2)$ and $s = (s_1, s_2) \in \mathcal{V}(c) \subset C^\infty(S^1, \mathbb{R} \times S^1)$ consider

$$\varphi_c(s) : S^1 \to \mathbb{R}^2, \qquad \varphi_c(s)(\theta) = c(s_2(\theta)) + s_1(s_2(\theta)).n_c(s_2(\theta))$$

where $\mathcal{V}(c)$ is a C^∞-open neighborhood of $(0, \mathrm{Id}_{S^1})$ in $C^\infty(S^1, \mathbb{R} \times S^1)$ chosen in such a way that:

- $s_2 \in \mathrm{Diff}(S^1)$ for each $s \in \mathcal{V}(c)$.
- $\varphi_c(s)$ is a free immersion for each $s \in \mathcal{V}(c)$.
- For $(s_1, s_2) \in \mathcal{V}(c)$ and $\alpha \in \mathrm{Diff}(S^1)$ we have $(s_1, s_2 \circ \alpha) \in \mathcal{V}(c)$.

Obviously $\varphi_c(s_1, s_2) \circ \alpha = \varphi_c(s_1, s_2 \circ \alpha)$ and s_2 is uniquely determined by $\varphi_c(s_1, s_2)$ since this is a free immersion. Thus the inverse of φ_c is a smooth chart for the manifold $\mathrm{Imm}_f(S^1, \mathbb{R}^2)$. Moreover, we consider the mapping (which will be important later)

$$\begin{aligned}
&\psi_c : C^\infty(S^1, (-\varepsilon, \varepsilon)) \to \mathrm{Imm}_f(S^1, \mathbb{R}^2), \qquad \mathcal{Q}(c) := \psi_c(C^\infty(S^1, (-\varepsilon, \varepsilon)))\\
&\psi_c(f)(\theta) = c(\theta) + f(\theta) n_c(\theta) = \varphi_c(f, \mathrm{Id}_{S^1})(\theta),\\
&\pi \circ \psi : C^\infty(S^1, (-\varepsilon, \varepsilon)) \to B_{i,f}(S^1, \mathbb{R}^2),
\end{aligned}$$

where ε is small. Then (an open subset of) $\mathcal{V}(c)$ splits diffeomorphically into

$$C^\infty(S^1, (-\varepsilon, \varepsilon)) \times \mathrm{Diff}\, S^1$$

and thus its image under φ_c splits into $\mathcal{Q}(c) \times \mathrm{Diff}(S^1)$. So the inverse of $\pi \circ \psi_c$ is a smooth chart for $B_{i,f}(S^1, \mathbb{R}^2)$. That the chart changes induced by the mappings φ_c and ψ_c contructed here are smooth is shown by writing them in terms of compositions and projections only and applying the setting of [**42**]. $\square$

11.5. Non free immersions. Any immersion is proper since S^1 is compact and thus by ([**19**], 2.1) the orbit space $B_i(S^1,\mathbb{R}^2) = \operatorname{Imm}(S^1,\mathbb{R}^2)/\operatorname{Diff}(S^1)$ is Hausdorff. Moreover, by ([**19**], 3.1 and 3.2) for any immersion c the isotropy group $\operatorname{Diff}(S^1)_c$ is a finite cyclic group which acts as group of covering transformations for a finite covering $q_c : S^1 \to S^1$ such that c factors over q_c to a free immersion $\bar{c} : S^1 \to \mathbb{R}^2$ with $\bar{c} \circ q_c = c$. Thus the subgroup $\operatorname{Diff}_1(S^1)$ of all diffeomorphisms φ fixing $1 \in S^1$ acts freely on $\operatorname{Imm}(S^1,\mathbb{R}^2)$. Moreover, for each $c \in \operatorname{Imm}$ the submanifold $\mathcal{Q}(c)$ from the proof of 11.4.3 (dropping the freeness assumption) is a slice in a strong sense:

- $\mathcal{Q}(c)$ is invariant under the isotropy group $\operatorname{Diff}(S^1)_c$.
- If $\mathcal{Q}(c)\circ\varphi \cap \mathcal{Q}(c) \neq \emptyset$ for $\varphi \in \operatorname{Diff}(S^1)$ then φ is already in the isotropy group $\varphi \in \operatorname{Diff}(S^1)_c$.
- $\mathcal{Q}(c)\circ\operatorname{Diff}(S^1)$ is an invariant open neigbourhood of the orbit $c\circ\operatorname{Diff}(S^1)$ in $\operatorname{Imm}(S^1,\mathbb{R}^2)$ which admits a smooth retraction r onto the orbit. The fiber $r^{-1}(c\circ\varphi)$ equals $\mathcal{Q}(c\circ\varphi)$.

Note that also the action

$$\operatorname{Imm}(S^1,\mathbb{R}^2)\times\operatorname{Diff}(S^1) \to \operatorname{Imm}(S^1,\mathbb{R}^2)\times\operatorname{Imm}(S^1,\mathbb{R}^2), \qquad (c,\varphi) \mapsto (c, c\circ\varphi)$$

is proper so that all assumptions and conclusions of Palais' slice theorem [**66**] hold. This results show that the orbit space $B_i(S^1,\mathbb{R}^2)$ has only very simple singularities of the type of a cone $\mathbb{C}/\{e^{2\pi k/n} : 0 \le k < n\}$ times a Fréchet space. We may call the space $B_i(S^1,\mathbb{R}^2)$ an infinite dimensional *orbifold*. The projection $\pi : \operatorname{Imm}(S^1,\mathbb{R}^2) \to B_i(S^1,\mathbb{R}^2) = \operatorname{Imm}(S^1,\mathbb{R}^2)/\operatorname{Diff}(S^1)$ is a submersion off the singular points and has only mild singularities at the singular strata. The normal bundle $\mathcal{N}_c$ mentioned in 11.3 is well defined and is a smooth vector subbundle of the tangent bundle. We do not have a principal bundle and thus no principal connections, but we can prove the main consequence, the existence of horizontal paths, directly:

Proposition. *For any smooth path c in $\operatorname{Imm}(S^1,\mathbb{R}^2)$ there exists a smooth path φ in $\operatorname{Diff}(S^1)$ with $\varphi(0,\quad) = \operatorname{Id}_{S^1}$ depending smoothly on c such that the path e given by $e(t,\theta) = c(t,\varphi(t,\theta))$ is horizontal: $e_t \perp e_\theta$.*

Proof. Let us write $e = c\circ\varphi$ for $e(t,\theta) = c(t,\varphi(t,\theta))$, etc. We look for φ as the integral curve of a time dependent vector field $\xi(t,\theta)$ on S^1, given by $\varphi_t = \xi\circ\varphi$. We want the following expression to vanish:

$$\langle \partial_t(c\circ\varphi), \partial_\theta(c\circ\varphi)\rangle = \langle c_t\circ\varphi + (c_\theta\circ\varphi)\,\varphi_t, (c_\theta\circ\varphi)\,\varphi_\theta\rangle$$

$$\begin{aligned}&= (\langle c_t, c_\theta\rangle \circ \varphi)\, \varphi_\theta + (\langle c_\theta, c_\theta\rangle \circ \varphi)\, \varphi_\theta\, \varphi_t \\ &= \big((\langle c_t, c_\theta\rangle + \langle c_\theta, c_\theta\rangle\, \xi) \circ \varphi\big)\, \varphi_\theta.\end{aligned}$$

Using the time dependent vector field $\xi = -\frac{\langle c_t, c_\theta\rangle}{|c_\theta|^2}$ and its flow φ achieves this. $\square$

11.6. The manifold of immersions with constant speed. Let $\mathrm{Imm}_a(S^1, \mathbb{R}^2)$ be the space of all immersions $c : S^1 \to \mathbb{R}^2$ which are parametrized by scaled arc length, so that $|c_\theta|$ is constant.

Proposition. *The space* $\mathrm{Imm}_a(S^1, \mathbb{R}^2)$ *is a smooth manifold. There is a diffeomorphism* $\mathrm{Imm}(S^1, \mathbb{R}^2) = \mathrm{Imm}_a(S^1, \mathbb{R}^2) \times \mathrm{Diff}_1^+(S^1)$ *which respects the splitting* $\mathrm{Diff}(S^1) = \mathrm{Diff}_1^+(S^1) \ltimes (S^1 \ltimes \mathbb{Z}_2)$*. There is a smooth action of the rotation and reflection group* $S^1 \ltimes \mathbb{Z}_2$ *on* $\mathrm{Imm}_a(S^1, \mathbb{R}^2)$ *with orbit space* $\mathrm{Imm}_a(S^1, \mathbb{R}^2)/(S^1 \ltimes \mathbb{Z}_2) = B_i(S^1, \mathbb{R}^2)$*.*

Proof. For $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$ we put

$$\begin{aligned}&\sigma_c \in \mathrm{Diff}_1(S^1), &&\sigma_c(\theta) = \exp\Big(\frac{2\pi i \int_1^\theta |c'(u)| du}{\int_{S^1} |c'(u)| du}\Big)\\ &\alpha : \mathrm{Imm}(S^1, \mathbb{R}^2) \to \mathrm{Imm}_a(S^1, \mathbb{R}^2), &&\alpha(c)(\theta) := c(\sigma_c^{-1}(\theta)).\end{aligned}$$

By the fundamentals of manifolds of mappings 2 the mapping α is smooth from $\mathrm{Imm}(S^1, \mathbb{R}^2)$ into itself and we have $\alpha \circ \alpha = \alpha$.

Now we show that $\mathrm{Imm}_a(S^1, \mathbb{R}^2)$ is a manifold. We use the notation from the proof of 11.4.3 with the freeness assumption dropped. For $c \in \mathrm{Imm}_a(S^1, \mathbb{R}^2)$ we use the following mapping as the inverse of a chart:

$$\begin{aligned}C^\infty(S^1, (-\varepsilon, \varepsilon)) \times S^1 \to \bigcup_{\theta \in S^1} \mathcal{Q}(c(\quad +\theta)) \xrightarrow{\alpha} \mathrm{Imm}_a(S^1, \mathbb{R}^2),\\ (f, \theta) \mapsto \psi_{c(\quad +\theta)}(f(\quad +\theta)) \mapsto \alpha(\psi_{c(\quad +\theta)}(f(\quad +\theta)))\end{aligned}$$

The chart changes are smooth: If for $(f_i, \theta_i) \in C^\infty(S^1, (-\varepsilon, \varepsilon)) \times S^1$ we have $\alpha(\psi_{c_1(\quad +\theta_1)}(f_1(\quad +\theta_1))) = \alpha(\psi_{c_2(\quad +\theta_2)}(f_2(\quad +\theta_2)))$ then the initial points agree and both curves are equally oriented so that $c_1(\theta + \theta_1) + f_1(\theta + \theta_1) n_{c_1}(\theta + \theta_1) = c_2(\varphi(\theta) + \theta_2) + f_2(\varphi(\theta) + \theta_2) n_{c_2}(\varphi(\theta) + \theta_2)$ for all θ. From this one can express (f_2, θ_2) smoothly in terms of (f_1, θ_1).

For the latter assertion one has to show that a smooth path through e_1 in $\mathcal{Q}(c_1)$ is mapped to a smooth path in $\mathrm{Diff}_1(S^1)$. This follows from the finite dimensional implicit function theorem. The mapping α is now smooth into $\mathrm{Imm}_a(S^1, \mathbb{R}^2)$ and the diffeomorphism $\mathrm{Imm}(S^1, \mathbb{R}^2) \to \mathrm{Imm}_a(S^1, \mathbb{R}^2) \times \mathrm{Diff}_1(S^1)$ is given by $c \mapsto (\alpha(c), \sigma_c)$ with inverse $(e, \varphi) \mapsto e \circ \varphi^{-1}$. Only the group $S^1 \ltimes \mathbb{Z}_2$ of rotations and reflections of S^1 then still acts on $\mathrm{Imm}_a(S^1, \mathbb{R}^2)$ with orbit space $B_i(S^1, \mathbb{R}^2)$. The rest is clear. $\square$

11.7. Tangent space, length, curvature, and Frenet-Serret formulas on $\operatorname{Imm}_a(S^1,\mathbb{R}^2)$**.** A smooth curve $t\mapsto c(\ \ ,t)\in\operatorname{Imm}(S^1,\mathbb{R}^2)$ lies in $\operatorname{Imm}_a(S^1,\mathbb{R}^2)$ if and only if $|\partial_\theta c|^2=|c_\theta|^2$ is constant in θ, i.e., $\partial_\theta|c_\theta|^2=2\langle c_\theta,c_{\theta\theta}\rangle=0$. Thus $h=\partial_t|_0 c\in T_c\operatorname{Imm}(S^1,\mathbb{R}^2)=C^\infty(S^1,\mathbb{R}^2)$ is tangent to $\operatorname{Imm}_a(S^1,\mathbb{R}^2)$ at the foot point c if and only if $\langle h_\theta,c_{\theta\theta}\rangle+\langle h_{\theta\theta},c_\theta\rangle=\langle h_\theta,c_\theta\rangle_\theta=0$, i.e., $\langle h_\theta,c_\theta\rangle$ is constant in θ. For $c\in\operatorname{Imm}_a(S^1,\mathbb{R}^2)$ the volume form is constant in θ since $|c_\theta|=\ell(c)/2\pi$. Thus for the curvature we have

$$\kappa:\operatorname{Imm}_a(S^1,\mathbb{R}^2)\to C^\infty(S^1,\mathbb{R}),\quad \kappa(c)=\Big(\frac{2\pi}{\ell(c)}\Big)^3\det(c_\theta,c_{\theta\theta})=\Big(\frac{2\pi}{\ell(c)}\Big)^3\langle ic_\theta,c_{\theta\theta}\rangle$$

and for the derivative of the length function we get

$$d\ell(c)(h)=\int_{S^1}\frac{\langle h_\theta,c_\theta\rangle}{|c_\theta|}d\theta=\frac{(2\pi)^2}{\ell(c)}\langle h_\theta(1),c_\theta(1)\rangle.$$

Since $c_{\theta\theta}$ is orthogonal to c_θ we have (Frenet formulas)

$$\begin{aligned}
c_{\theta\theta}&=\Big(\frac{2\pi}{\ell(c)}\Big)^2\langle ic_\theta,c_{\theta\theta}\rangle ic_\theta=\frac{\ell(c)}{2\pi}\kappa(c)ic_\theta,\\
c_{\theta\theta\theta}&=\frac{\ell(c)}{2\pi}\kappa(c)_\theta\, ic_\theta+\frac{\ell(c)}{2\pi}\kappa(c)ic_{\theta\theta}=\frac{\ell(c)}{2\pi}\kappa(c)_\theta\, ic_\theta-\Big(\frac{\ell(c)}{2\pi}\Big)^2\kappa(c)^2c_\theta.
\end{aligned}$$

The derivative of the curvature thus becomes:

$$d\kappa(c)(h)=-2\Big(\frac{2\pi}{\ell(c)}\Big)^2\langle h_\theta,c_\theta\rangle\kappa(c)+\Big(\frac{2\pi}{\ell(c)}\Big)^3\langle ic_\theta,h_{\theta\theta}\rangle\big).$$

11.8. Horizontality on $\operatorname{Imm}_a(S^1,\mathbb{R}^2)$**.** Let us denote by $\operatorname{Imm}_{a,f}(S^1,\mathbb{R}^2)$ the splitting submanifold of Imm consisting of all constant speed free immersions. From 11.6 and 11.4.3 we conclude that the projection $\operatorname{Imm}_{a,f}(S^1,\mathbb{R}^2)\to B_f(S^1,\mathbb{R}^2)$ is principal fiber bundle with structure group $S^1\ltimes\mathbb{Z}_2$, and it is a reduction of the principal fibration $\operatorname{Imm}_f\to B_f$. The principal connection described in 11.4.3 is not compatible with this reduction. But we can easily find some principal connections. The one we will use is described by the horizontal bundle with fiber $\mathcal{N}_a,c$ consisting of all vector fields h along c such that $\langle h_\theta,c_\theta\rangle_\theta=0$ (tangent to Imm_a) and $\langle h(1),c_\theta(1)\rangle=0$ for $1\in S^1$ (horizontality). This connection admits a smooth parallel transport; but we can even do better, beyond the principal bundle, in the following proposition whose proof is similar and simpler than that of proposition 11.5.

Proposition. *For any smooth path c in $\operatorname{Imm}_a(S^1,\mathbb{R}^2)$ there exists a smooth curve φ_c in S^1 with $\varphi_c(0)=1$ depending smoothly on c such that the path e given by $e(t,\theta)=c(t,\varphi_c(t)\theta)$ is horizontal: $e_t(1)\perp e_\theta(1)$.* ☐

11.9. The degree of immersions. Recall that the degree of an immersion $c : S^1 \to \mathbb{R}^2$ is the winding number with respect to 0 of the tangent $c' : S^1 \to \mathbb{R}^2$. Since this is invariant under isotopies of immersions, the manifold $\operatorname{Imm}(S^1, \mathbb{R}^2)$ decomposes into the disjoint union of the open submanifolds $\operatorname{Imm}^k(S^1, \mathbb{R}^2)$ for $k \in \mathbb{Z}$ according to the degree k. We shall also need the space $\operatorname{Imm}^k_a(S^1, \mathbb{R}^2)$ of all immersions of degree k with constant speed.

11.10. Theorem.

(1) *The manifold* $\operatorname{Imm}^k(S^1, \mathbb{R}^2)$ *of immersed curves of degree* k *contains the subspace* $\operatorname{Imm}^k_a(S^1, \mathbb{R}^2)$ *as smooth strong deformation retract.*

(2) *For* $k \neq 0$ *the manifold* $\operatorname{Imm}^k_a(S^1, \mathbb{R}^2)$ *of immersed constant speed curves of degree* k *contains* S^1 *as a strong smooth deformation retract.*

(3) *For* $k \neq 0$ *the manifold* $B^k_i(S^1, \mathbb{R}^2) := \operatorname{Imm}^k(S^1, \mathbb{R}^2)/\operatorname{Diff}^+(S^1)$ *is contractible.*

Note that for $k \neq 0$ Imm^k is invariant under the action of the group $\operatorname{Diff}^+(S^1)$ of orientation preserving diffeomorphism only, and that any orientation reversing diffeomorphism maps Imm^k to Imm^{-k}.

The nontrivial S^1 in Imm^k appears in 2 ways: (a) by rotating each curve around $c(0)$ so that $c'(0)$ rotates. And (b) also by acting $S^1 \ni \beta \mapsto (c(\theta) \mapsto c(\beta\theta))$. The two corresponding elements a and b in the fundamental group are then related by $a^k = b$ which explains our failure to describe the topological type of B^0_i.

Proof. (1) is a consequence of 11.6 since $\operatorname{Diff}^+_1(S^1)$ is contractible.

The general proof is inspired by the proof of the Whitney-Graustein theorem, [**78**], [**33**], [**34**]. We shall view curves here as 2π-periodic plane-valued functions. For any curve c we consider its *center of mass*

$$C(c) = \operatorname{Center}(c) := \frac{1}{\ell(c)} \int_0^{2\pi} c(u)\, |c'(u)|\, du \in \mathbb{R}^2$$

which is invariant under $\operatorname{Diff}(S^1)$. We shall also use $\alpha(c) = c'(0)/|c'(0)|$.

The case $k \neq 0$. We first embedd S^1 into $\operatorname{Imm}(S^1, \mathbb{R}^2)$ in the following way. For $\alpha \in S^1 \subset \mathbb{C} = \mathbb{R}^2$ and $k \neq 0$ we but $e_\alpha(\theta) = \alpha.e^{ik\theta}/ik$, a circle of radius $1/|k|$ transversed k-times in the direction indicated by the sign of k. Note that we have $\operatorname{Center}(e_\alpha) = 0$ and $e'_\alpha(0) = \alpha$.

Since the isotopies to be constructed later will destroy the property of having constant speed, we shall first construct a smooth deformation retraction $A : [0,1] \times \operatorname{Imm}^k \to \operatorname{Imm}^k_{1,0}$ onto the subspace $\operatorname{Imm}^k_{1,0}$ of unit speed degree $k \neq 0$ curves with center 0.

Let $c : \mathbb{R} \to \mathbb{R}^2$ be an arbitrary constant speed immersion of degree k, period 2π, and length $\ell(c)$. Let $s_c(v) = \int_0^v |c'(u)|\, du$ be the arc-length function of c

and put

$$A(c,t,u) = \Big(1 - t + t\frac{2\pi}{\ell(c)}\Big).\Big(c\big((1-t)u + t.s_c^{-1}(\tfrac{\ell(c)}{2\pi}u)\big) - t.C(c)\Big).$$

Then A_c is an isotopy between c and $c_1 := A(c,1,\quad)$ depending smoothly on c. The immersion c_1 has unit speed, length 2π, and $\mathrm{Center}(c_1) = 0$. Moreover, for the winding number w_0 around 0 we have:

$$w_0(c_1'|_{[0,2\pi]}) = \deg(c_1) = \deg(c) = k = \deg(e_{\alpha(c)}) = w_0(e'_{\alpha(c)}|_{[0,2\pi]}).$$

Thus Imm^k contains the space $\mathrm{Imm}^k_{1,0}$ of unit speed immersions with center of mass 0 and degree k as smooth strong deformation retract.

For $c \in \mathrm{Imm}^k_{1,0}$ a unit speed immersion with center 0 we now construct an isotopy $t \mapsto H^1(c,t,\quad)$ between c and a suitable curve e_α. It will destroy the unit speed property, however. For $d\arg = \frac{-xdy+ydx}{\sqrt{x^2+y^2}}$ we put:

$$\begin{aligned}
\varphi_c(u) &:= \int_{c'|[0,u]} d\arg, \qquad\qquad \text{so that } c'(u) = c'(0)\,e^{i\varphi_c(u)},\\
\alpha(c) &:= \frac{1}{2\pi}\int_0^{2\pi}(\varphi_c(v) - kv)dv,\\
\psi_c(t,u) &:= (1-t)\varphi_c(u) + t(ku + \alpha(c)),\\
h(c,t,u) &:= \int_0^u e^{i\psi_c(t,v)}dv - \frac{u}{2\pi}\int_0^{2\pi} e^{i\psi_c(t,v)}dv,\\
H^1(c,t,u) &:= c'(0)\Big(h(c,t,u) - \mathrm{Center}(h(c,t,\quad)\Big)
\end{aligned}$$

Then $H^1(c,t,u)$ is smooth in all variables, 2π-periodic in u, with center of mass at 0, $H^1(1,c,u)$ equals one the e_α's, and $H^1(0,c,u) = c(u)$. But $H^1(c,t,\quad)$ is, however, no longer of unit speed in general. And we still have to show that $t \mapsto h(c,t,\quad)$ (and consequently H^1) is an isotopy.

$$\partial_u h(c,t,u) = e^{i\psi_c(t,u)} - \frac{1}{2\pi}\int_0^{2\pi} e^{i\psi_c(t,v)}dv,$$

$$\Big|\frac{1}{2\pi}\int_0^{2\pi} e^{i\psi_c(t,v)}dv\Big| \le 1. \tag{4}$$

If the last inequality is strict we have $\partial_u h(t,u) \neq 0$ so that h is an isotopy. If we have equality then $\psi_c(t,v)$ is constant in v which leads to a contradiction as follows: If $k \neq 0$ then $\psi_c(t,2\pi) - \psi_c(t,0) = 2\pi k$ so it cannot be constant for any t.

Let us finally check how this construction depends on the choice of the base point $c(0)$. We have:

$$\begin{aligned}
&\varphi_{c(\beta+\quad)}(u) = \varphi_c(\beta+u) - \varphi_c(\beta),\\
&\alpha(c(\beta+\quad)) = \alpha(c) + k\beta - \varphi_c(\beta),
\end{aligned}$$

$$\begin{aligned}
\psi_{c(\beta+\ \ )}(t,u) &= \psi_c(t,u+\beta) - \varphi_c(\beta),\\
h(c(\beta+\ \ ),t,u) &= e^{-i\varphi_c(\beta)}(h(c,t,\beta+u) - h(c,t,\beta)),\\
H^1(c(\beta+\ \ ),t,u) &= H^1(c,t,\beta+u).
\end{aligned}$$

Let us now deform H^1 back into $\mathrm{Imm}^k_{1,0}$. For $c \in \mathrm{Imm}^k_{1,0}$ we consider

$$\begin{aligned}
H^2(c,t,u) &:= A(1,H^1(c,t,\ \ ),u),\\
H^3(c,t,u) &:= H^2(c,t,u+\varphi_{H^2(c)}(t)),
\end{aligned}$$

where the φ_f for a unit speed path f is from proposition 11.8, so that $H^3(c)$ is a horizontal path of unit speed curves of length 2π, (i.e., $\partial_t H^3(c,t,0)\perp\partial_u|_0 H^3(c,t,u)$). The isotopy A reacts in a complicated way to rotations of the parameter, but we have $A(c(\beta+\ \ ),1,u) = A(c,1,\frac{2\pi}{\ell(c)}s_c(\beta)+u)$. Thus $H^3(c(\ \ +\beta),t,u) = H^3(c,t,u+\beta)$, so H^3 is equivariant under the rotation group $S^1 \subset \mathrm{Diff}(S^1)$. For $k \neq 0$ we get an equivariant smooth strong deformation retract within $\mathrm{Imm}^k_{1,0}$ onto the subset $\{e_\alpha : \alpha \in S^1\} \subset \mathrm{Imm}^k_{1,0}$ which is invariant under the rotation group $S^1 \subset \mathrm{Diff}(S^1)$. It factors to a smooth contraction on B^k_i. This proves assertions (2) and (3) for $k \neq 0$. $\square$

11.11. The homotopy type of degree 0 immersions. This subsection follows [**39**]. Here we treat the homotopy type of the shape space $B^0(S^1,\mathbb{R}^2)$ of curves of degree 0. In part (1) we first give simple argument which shows that $\pi_1(B^0(S^1,\mathbb{R}^2))$ contains $\mathbb{Z}$. In section (2) we give a more involved proof that $\mathrm{Imm}^0(S^1,\mathbb{R}^2)$ is homotopy equivalent to S^1. In section (3) we show that factoring out $\mathrm{Diff}^+(S^1)$ gives a fibration with homotopically trivially embedded fiber, and then the homotopy sequence shows that $\pi_1(B^{0,+}(S^1,\mathbb{R}^2)) = \mathbb{Z}$, $\pi_2(B^{0,+}(S^1,\mathbb{R}^2)) = \mathbb{Z}$, and $\pi_k(B^{0,+}(S^1,\mathbb{R}^2)) = 0$ for $k > 2$. Factoring out the larger group $\mathrm{Diff}(S^1)$ gives a two-sheeted covering and the final result.

(1) **Proposition.** *The space* $\mathrm{Imm}^0(S^1,\mathbb{R}^2)/\mathrm{Diff}^+(S^1)$ *is not contractible.*

Proof. We shall view a curve $c \in \mathrm{Imm}(S^1,\mathbb{R}^2)$ as a 2π-periodic plane valued function. A smooth function $a = a(c,\ \ ) : \mathbb{R} \to \mathbb{R}$ is called an argument of a curve c if

$$\frac{c'(\theta)}{|c'(\theta)|} = \exp(i\,a(\theta));$$

it is unique up to addition of an inter multiple of 2π. If the curve c has degree k then $a(\theta+2\pi)-a(\theta) = 2k\pi$. Thus, a curve c is in $\mathrm{Imm}^0(S^1,\mathbb{R}^2)$ if and only if some (any) argument of c is 2π-periodic. For a curve $c \in \mathrm{Imm}^0(S^1,\mathbb{R}^2)$, we define the *average argument* $\alpha(c) \in S^1$ by

$$\alpha(c) = \exp\left(\frac{i}{l(c)}\int_0^{2\pi} a(c,\theta)|c'(\theta)|d\theta\right),$$

which does not depend on the choice of $a(c,\quad)$ and defines a well-defined smooth mapping $\alpha : \mathrm{Imm}^0(S^1,\mathbb{R}^2) \to S^1$. Also, since any argument a of a degree 0 curve is 2π-periodic, $\alpha(c)$ is invariant under the action of $\mathrm{Diff}^+(S^1)$. So we can view α as a map

$$\alpha\colon B^{0,+}(S^1,\mathbb{R}^2) = \mathrm{Imm}^0(S^1,\mathbb{R}^2)/\operatorname{Diff}^+(S^1) \to S^1.$$

For $\varphi \in S^1 \subset \mathbb{C} = \mathbb{R}^2$, the rotation map $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ act on $B^{0,+}(S^1,\mathbb{R}^2)$ and obviously

$$\alpha(\varphi.c) = \varphi.\alpha(c).$$

So choosing a free orbit $S^1.C$ for the rotation action of S^1 on $B^{0,+}(S^1,\mathbb{R}^2)$, the composition

$$S^1.C \hookrightarrow \mathrm{Imm}^0(S^1,\mathbb{R}^2)/\operatorname{Diff}^+(S^1) \xrightarrow{\alpha} S^1$$

equals the identity on S^1, thus $\pi_1(S^1) = \mathbb{Z} \subset \pi_1(B^{0,+}(S^1,\mathbb{R}^2))$.

Moreover, $\alpha(c(-\quad)) = -\alpha(c)$ implies that α factors as follows, where the vertical arrows are 2-sheeted coverings:

$$\begin{tikzcd}
B^{0,+}(S^1,\mathbb{R}^2) \arrow[r, equal] & \mathrm{Imm}^0(S^1,\mathbb{R}^2)/\operatorname{Diff}^+(S^1) \arrow[r, "\alpha"] \arrow[d, "2"'] & S^1 \arrow[d, "2"] \\
B^{0}(S^1,\mathbb{R}^2) \arrow[r, equal] & \mathrm{Imm}^0(S^1,\mathbb{R}^2)/\operatorname{Diff}(S^1) \arrow[r, "\bar\alpha"] & S^1
\end{tikzcd}$$

Thus we also get in a similar way $\pi_0(S^1) = \mathbb{Z} \subset \pi_0(B^0(S^1,\mathbb{R}^2))$. □

(2) **Proposition (Homotopy of $\mathrm{Imm}^0(S^1,\mathbb{R}^2)$).** *The space $\mathrm{Imm}^0(S^1,\mathbb{R}^2)$ of degree 0 immersions in the plane is homotopy equivalent to S^1.*

Proof. This will follow from (3)– (6)below. □

(3)Let $\mathrm{Imm}^{0,*}(S^1,\mathbb{R}^2) := \{c \in \mathrm{Imm}^0(S^1,\mathbb{R}^2);\ c(0) = 0\}$. Clearly we have $\mathrm{Imm}^0(S^1,\mathbb{R}^2) \cong \mathrm{Imm}^{0,*}(S^1,\mathbb{R}^2) \times \mathbb{R}^2$ and $\mathrm{Imm}^0(S^1,\mathbb{R}^2) \sim \mathrm{Imm}^{0,*}(S^1,\mathbb{R}^2)$, where $\cong$ denotes homoeomorphism and $\sim$ homotopy equivalent. Let us define a map

$$\Phi\colon \mathrm{Imm}^{0,*} \to C^\infty(S^1,\mathbb{R}_+) \times C^\infty(S^1,S^1)$$

$$\Phi(c)(\theta) = \left(|c_\theta(\theta)|, \frac{c_\theta(\theta)}{|c_\theta(\theta)|}\right) =: (v(\theta), e(\theta)).$$

The map Φ is injective. For $(v,e) = \Phi(c)$, the winding number of e equals the degree 0 of c and thus $\int_0^{2\pi} v.e\, d\theta = 0$.

Lemma. *The length of the image of e is greater than π.*

Proof. If not, there exists a number $r \in \mathbb{R}$ such that

$$\exp(ir) \in \mathrm{Im}(e) \subset \exp(i[r - \pi/2, r + \pi/2]).$$

Then, $\langle \exp(ir), e(\theta)\rangle$ is nonnegative for any θ and strictly positive for some θ. Therefore $\int_0^{2\pi}\langle \exp(ir), v.e\rangle\, d\theta > 0$. This contradicts

$$\int_0^{2\pi}\langle \exp(ir), v.e\rangle\, d\theta = \left\langle \exp(ir), \int_0^{2\pi} v.e\, d\theta\right\rangle = \langle \exp(ir), 0\rangle = 0. \quad \square$$

(4)Let us define the set

$$C^{\infty,0}_{>\pi}(S^1, S^1) = \{e \in C^\infty(S^1, S^1);\ \deg(e) = 0,\ \mathrm{length}(\mathrm{Im}(e)) > \pi\}$$

and consider the map

$$\mathrm{pr}_2 \circ \Phi\colon\ \mathrm{Imm}^{0,*}(S^1, S^1) \to C^{\infty,0}_{>\pi}(S^1, S^1),$$

where pr_2 denotes the second projection.

Lemma. *The map* $\mathrm{pr}_2 \circ \Phi\colon \mathrm{Imm}^{0,*}(S^1, S^1) \to C^{\infty,0}_{>\pi}(S^1, S^1)$, *is surjective, has contractible fibers, admits a global smooth section, and is a homotopy equivalence.*

Proof. For a map $e \in C^{\infty,0}_{>\pi}(S^1, S^1)$, there exist points θ_1, θ_2, θ_3 such that $0 \in \mathrm{int}([e(\theta_1), e(\theta_2), e(\theta_3)])$, where $[\cdot,\cdot,\cdot]$ denotes the convex hull of three points. Let $v_1 \in C^\infty(S^1, \mathbb{R}_{>0})$ be a map such that $\int_0^{2\pi} v_1 d\theta = 1$ and $v_1(\theta)$ is close to 0 if θ is not close to θ_1. Then $\int_0^{2\pi} v_1.e\, d\theta$ is close to $e(\theta_1)$. We also define v_2 and v_3 similarly, so that

$$0 \in \mathrm{int}\left(\left[\int_0^{2\pi} v_1.e\, d\theta, \int_0^{2\pi} v_2.e\, d\theta, \int_0^{2\pi} v_3.e\, d\theta\right]\right).$$

Therefore there exist positive numbers a_1, a_2, a_3 with

$$a_1\int_0^{2\pi} v_1.e\, d\theta + a_2\int_0^{2\pi} v_2.e\, d\theta + a_3\int_0^{2\pi} v_3.e\, d\theta = 0.$$

Define c by

$$c(\theta) = \int_0^{\theta} (a_1 v_1(u) + a_2 v_2(u) + a_3 v_3(u)) e(u)\, du.$$

Then c is in $\mathrm{Imm}^{0,*}$ and $(\mathrm{pr}_2 \circ \Phi)(c) = e$, which means that $\mathrm{pr}_2 \circ \Phi$ is surjective.

We next show that for any $e \in C^{\infty,0}_{>\pi}(S^1, S^1)$, the inverse image $(\mathrm{pr}_2 \circ \Phi)^{-1}(e)$ is contractible. Namely, let $V(e) \subset C^\infty(S^1, \mathbb{R}_+)$ be given by

$$V(e) = \left\{ v \in C^\infty(S^1, \mathbb{R}_+);\ \int_0^{2\pi} v.e\, d\theta = 0\right\},$$

an open convex subset of the linear subspace $\{v \in C^\infty(S^1, \mathbb{R});\ \int_0^{2\pi} v.e\, d\theta = 0\} \subset C^\infty(S^1, \mathbb{R})$. Thus $V(e)$ is contractible for each e. Moreover, $V(e)$ is homeomorphic to $(\mathrm{pr}_2 \circ \Phi)^{-1}(e)$ by the map $\mathrm{pr}_1 \circ \Phi\colon (p_2 \circ \Phi)^{-1}(e) \to V(e)$.

For fixed $\theta_1, \theta_2, \theta_3$ the construction above works for each $e \in C^{\infty,0}_{>\pi}(S^1, S^1)$ for which 0 is contained in the interior of the convex hull of $e(\theta_1), e(\theta_2), e(\theta_3)$; these e form an open set in $C^{\infty,0}_{>\pi}(S^1, S^1)$ on which we get a continuous (even smooth) section of $p_2 \circ \Phi$. Open sets like that cover $C^{\infty,0}_{>\pi}(S^1, S^1)$. So we get smooth local sections whose domains cover the base. Since the base is open in a nuclear Fréchet space, it is smoothly paracompact (see [**42**], 16.10) we can use convexity of all fibers and a smooth partition of unity on the base $C^{\infty,0}_{>\pi}(S^1, S^1)$ to construct a global smooth section s.

Finally, since all fibers are convex, there is a smooth strong fiber preserving deformation retraction of $\mathrm{Imm}^{0,*}(S^1, S^1)$ onto the image the global section s. $\square$

(5)To study the topology of $C^{\infty,0}_{>\pi}(S^1, S^1)$, we introduce the set of 2π-periodic functions

$$C^{\infty,p}(\mathbb{R}, \mathbb{R}) = \{c \in C^\infty(\mathbb{R}, \mathbb{R});\ c(\theta + 2\pi) = c(\theta)\}.$$

For $c \in C^{\infty,p}(\mathbb{R}, \mathbb{R})$ let $\mathrm{Var}(c) = \max c - \min c$ and let $\mathrm{Ave}(c) = \frac{1}{2\pi}\int_0^{2\pi} c\, d\theta$. For $k \geq 0$, $C^{\infty,p}_{>k}(\mathbb{R}, \mathbb{R}) = \{c \in C^{\infty,p}(\mathbb{R}, \mathbb{R});\ \mathrm{Var}(c) > k\}$. Define a diffeomorphism $g\colon C^{\infty,p}_{>0}(\mathbb{R}, \mathbb{R}) \to C^{\infty,p}_{>\pi}(\mathbb{R}, \mathbb{R})$ by

$$g(c) = \frac{\mathrm{Var}(c) + \pi}{\mathrm{Var}(c)}(c - \mathrm{Ave}(c)) + \mathrm{Ave}(c).$$

The diffeomorphism g satisfies $g(c(\quad +2n\pi)) = g(c)(\quad +2n\pi)$, thus induces the diffeomorphism

$$\tilde{g}\colon C^{\infty,0}_{>0}(S^1, S^1) \to C^{\infty,0}_{>\pi}(S^1, S^1),$$

where $C^{\infty,0}_{>0}(S^1, S^1)$ denotes the set of nonconstant smooth maps of degree 0 in $C^\infty(S^1, S^1)$.

(6)We consider now the evaluation ev_1 at $1 \in S^1$ whose fiber at $1 \in S^1$ is the smooth manifold of based smooth loops of degree 0 in S^1, with the constant loop 1 deleted:

$$C^{\infty,0}((S^1, 1), (S^1, 1)) \setminus \{1\} \hookrightarrow C^{\infty,0}_{>0}(S^1, S^1) \xrightarrow{\mathrm{ev}_1} S^1$$

Lemma. *The map* $\mathrm{ev}_1 : C^{\infty,0}_{>0}(S^1, S^1) \to S^1$ *is a smooth trivial fibration with a global section and smoothly contractible fibers. Moreover, it is a homotopy equivalence.*

Proof. A smooth section $s : S^1 \to C^{\infty,0}_{>0}(S^1, S^1)$ of ev_1 is given by $s(\varphi)(\theta) = \varphi.\exp(i\,\mathrm{Im}(\theta))$. The fiber of ev_1 over φ is the space $C^{\infty,0}((S^1, 1), (S^1, \varphi)) \setminus \{\varphi\}$ consisting of all non-constant smooth loops of degree 0 mapping 1 to φ, which is diffeomorphic to the fiber $C^{\infty,0}((S^1, 1), (S^1, 1)) \setminus \{1\}$ via multiplication by φ.

It remains to show that the fiber $C^{\infty,0}((S^1,1),(S^1,1))\setminus\{1\}$ is contractible. Via lifting to the univeral cover, $C^{\infty,0}((S^1,1),(S^1,1))$ is diffeomorphic to the space $\{f\in C^{\infty,p}(\mathbb{R},\mathbb{R}) : f(0)=0\}$ of periodic functions mapping 0 to 0. Via Fourier expansion $f(t)=\sum_{n\in\mathbb{Z}} a_n \exp(int)$ this is isomorpic to the space of all rapidly decreasing complex sequences $(a_k)_{k\in\mathbb{Z}}$ with $\overline{a_k}=a_{-k}$ and $\sum_k a_k = 0$. This space is isomorphic to the space $\mathfrak{s}$ of rapidly decreasing sequences $(b_n)_{n\geq 1}$ by $a_n=b_n$ for $n\geq 1$, $a_{-n}=\bar b_n$, and $a_0=2\operatorname{Re}(\sum_{n\geq 1} b_n)$. Now we have to show that this is still contractible if we remove the constant sequence 0. Then it is homotopy equivalent to its intersection with the sphere in ℓ^2, i.e., to the space $S:=\{b\in\mathfrak{s} : \sum_{n\geq 1} b_n^2 = 1\}$. But this is contractible by a standard argument which is explained on page 513 of [**42**] for the space of finite sequences. Namely, consider the homotopy $A : \mathfrak{s}\times[0,1]\to\mathfrak{s}$ through isometries which is given by $A_0=\mathrm{Id}$ and by

$$\begin{aligned}A_t(b_1,b_2,b_3,\dots) = (&b_1,\dots,b_{n-2},b_{n-1}\cos\theta_n(t), b_{n-1}\sin\theta_n(t),\\ &b_n\cos\theta_n(t), b_n\sin\theta_n(t), b_{n+1}\cos\theta_n(t), b_{n+1}\sin\theta_n(t),\dots)\end{aligned}$$

for $\frac{1}{n+1}\leq t\leq \frac{1}{n}$, where $\theta_n(t)=\varphi(n((n+1)t-1))\frac{\pi}{2}$ for a fixed smooth function $\varphi:\mathbb{R}\to\mathbb{R}$ which is 0 on $(-\infty,0]$, grows monotonely to 1 in $[0,1]$, and equals 1 on $[1,\infty)$. The mapping A is Lipschitz continuous for each seminorm $\|b\|_k=\sup\{|b_n|n^k : n\geq 1\}$ of $\mathfrak{s}$ with constant 2^k, and is isometric for ℓ^2. Then $A_{1/2}(b_1,b_2,\dots)=(b_1,0,b_2,0,\dots)$ is in $\mathfrak{s}_{\text{odd}}$, and on the other hand $A_1(b_1,b_2,\dots)=(0,b_1,0,b_2,0,\dots)$ is in $\mathfrak{s}_{\text{even}}$. This is a variant of a homotopy constructed by [**68**]. Now $A_t|S$ for $0\leq t\leq 1/2$ is a homotopy on S between the identity and $A_{1/2}(S)\subset\mathfrak{s}_{\text{odd}}$. The latter set is contractible, for example in a stereographic chart. $\square$

(7)If we put together all mappings constructed above we get the following commutative diagram where we indicate isomorphism $\cong$, homotopy equivalence $\sim$, or 2-sheeted covering 2, and a free orbit $S^1.c$ for the rotation action on Imm^0:

$$\begin{tikzcd}
& S^1 \arrow[r, "="] & S^1 \arrow[r, "2"] & S^1 \\
S^1.c \arrow[ur, "\cong"] \arrow[r, "\subset"] & \mathrm{Imm}^0 \arrow[u, "\alpha"'] \arrow[r, two heads] \arrow[d, "\sim"', "\tilde g^{-1}\circ\mathrm{pr}_2\circ\Phi"] & B^{0,+} \arrow[u, "\alpha"] \arrow[r, "2"] & B^0 \arrow[u, "\bar\alpha"] \\
S^1 \arrow[u, "="] & C^{\infty,0}_{>0} \arrow[l, "\mathrm{ev}_1", "\sim"'] & &
\end{tikzcd}$$

(8) **Proposition.** *The mapping* $\mathrm{Imm}^0(S^1,\mathbb{R}^2)\to B^{0,+}(S^1,\mathbb{R}^2)$ *is a (Serre) fibration.*

Proof. First we replace $\operatorname{Imm}^0(S^1,\mathbb{R}^2)$ by the subset $\operatorname{Imm}^0_a(S^1,\mathbb{R}^2)$ consisting of all immersions which are parametrized by scaled arc-length which is a strong deformation retract, see [**54**], 2.6. The normalizer of the $\operatorname{Diff}^+(S^1)$-action on it is just the action of S^1 which shifts the initial point. We have to show that for any compactly generated space P and a homotopy $h : [0,1]\times P \to B^{0,+}$ whose initial value $h(0,\quad)$ admits a continuous lift there exists a continuous lift of the whole homotopy:

$$\begin{tikzcd}
\{0\}\times P \arrow[r, "H(0,\quad)"] \arrow[d, "\subset"] & \operatorname{Imm}^0_a(S^1,\mathbb{R}^2) \arrow[d] \\
{[0,1]\times P} \arrow[r, "h"] \arrow[ur, dashed, "H"] & B^{0,+}(S^1,\mathbb{R}^2)
\end{tikzcd}$$

To get the lift H we just have to specify the initial point coherently from $H(0,p)(1)$ over $[0,1]\ni t\mapsto h(t,p)$.

For that we need a description of the elements in $B^{0,+}(S^1,\mathbb{R}^2)$. A point C in it can be described by the following data:

For some n and $i=1,\dots,n$, there are open sets $U_i=U_i(C)\subseteq\mathbb{R}^2$, smooth functions $f_i=f_i(C):U_i\to\mathbb{R}$ such that $f_i^{-1}(0)=:C_i$ is a component C_i of C with $\operatorname{grad}(f_i)$ is a unit vector field with flow lines unit speed straight lines passing orthogonally through C_i in such a way that for $x\in C_i$ the frame consisting of $\operatorname{grad}(f_i)(x)$ and the unit tangent to C_i at x is positively oriented. The unparameterized smooth oriented 1-manifolds $C_1,C_2,\dots,C_n$ (in that order) describe C. Note that there is a choice for the U_i and their cyclic order, but then the f_i are unique.

For every $p\in P$ the initial point $H(0,p)(1)$ lies in some component $h(0,p)_i$ of $h(0,p)$, and we may move it orthogonally along $\operatorname{grad}(f_i(h(t,p)))$ to get a coherent choice of initial points. This takes care of the lift H. □

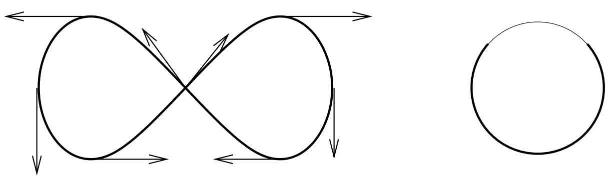
(9) **Lemma.** *The fiber* $\operatorname{Diff}^+(S^1)$ *maps homotopically trivial into the fibration* $\operatorname{Imm}^0(S^1,\mathbb{R}^2)\to B^{0,+}(S^1,\mathbb{R}^2)$.

Proof. As in the proof of (4)we consider the space $\operatorname{Imm}^0_a(S^1,\mathbb{R}^2)$ of degree 0 immersions with constant speed parametrizations. Let c be the unit speed parameterized horizontal figure eight, and consider the diagram where

$c_*(f) = c \circ f$:

$$\begin{tikzcd}
S^1 \arrow[r, "c_*"] \arrow[d, "\sim"', "\subset"] & \mathrm{Imm}_a^0 \arrow[d, "\sim"', "\subset"] \arrow[dr, two heads] & \\
\mathrm{Diff}^+(S^1) \arrow[r, "c_*"] & \mathrm{Imm}^0 \arrow[r, two heads] \arrow[d, "\sim"', "\tilde g^{-1} \circ \mathrm{pr}_2 \circ \Phi"] & B^{0,+} \\
S^1 & C^{\infty,0}_{>0} \arrow[l, "\mathrm{ev}_1", "\sim"'] &
\end{tikzcd}$$

We have to show that the mapping from the upper left S^1 to the lower left S^1 is nullhomotopic. It is essentially (suppressing $\tilde g^{-1}$) given by $\beta \mapsto \frac{c'(\beta)}{|c'(\beta)|}$. From the figure



we see that this mapping covers everything below the northern polar region twice and avoids the northern polar region, so it is nullhomotopic. $\square$

(10) **Corollary.** *We have the following homotopy groups:*

$$\begin{aligned}
&\pi_1(B^{0,+}(S^1,\mathbb{R}^2)) = \mathbb{Z}, && \pi_1(B^0(S^1,\mathbb{R}^2)) = \mathbb{Z}, \\
&\pi_2(B^{0,+}(S^1,\mathbb{R}^2)) = \mathbb{Z}, && \pi_2(B^0(S^1,\mathbb{R}^2)) = \mathbb{Z}, \\
&\pi_k(B^{0,+}(S^1,\mathbb{R}^2)) = 0, && \pi_k(B^0(S^1,\mathbb{R}^2)) = 0 \quad \textit{for } k > 2.
\end{aligned}$$

Proof. By (1) we have the long exact homotopy sequence

$$\dots \to \pi_k(S^1) \xrightarrow{0} \pi_k(\mathrm{Imm}_a^0) \to \pi_k(B^{0,+}) \to \pi_{k-1}(S^1) \to \dots$$

and by section (2) the space Imm_a^0 is homotopy equivalent to S^1. This gives the homotopy groups of $B^{0,+}(S^1,\mathbb{R}^2))$. Since $B^{0,+}(S^1,\mathbb{R}^2)) \to B^0(S^1,\mathbb{R}^2))$ is a two-sheeted covering, we can also read of the homotopy groups of $B^{0,+}(S^1,\mathbb{R}^2))$. $\square$

11.12 Bigger spaces of 'immersed' curves. We want to introduce a larger space containing $B_i(S^1,\mathbb{R}^2)$, which is complete in a suitable metric. This will serve as an ambient space which will contain the completion of $B_i(S^1,\mathbb{R}^2)$. Let $\mathrm{Cont}(S^1,\mathbb{R}^2)$ be the space of all *continuous* functions $c : S^1 \to \mathbb{R}^2$. Instead of a group operation and its associated orbit space, we introduce an equivalence relation on $\mathrm{Cont}(S^1, \mathbb{R}^2)$. Define a subset $R \subset S^1 \times S^1$ to be a *monotone correspondence* if it is the image of a map

$$x \to (h(x) \mod 2\pi, k(x) \mod 2\pi), \quad \text{where}$$
$$h, k : \mathbb{R} \to \mathbb{R} \quad \text{are monotone non-decreasing continuous functions such that}$$
$$h(x+2\pi) \equiv h(x) + 2\pi, k(x+2\pi) \equiv k(x) + 2\pi.$$

In words, this is an orientation preserving homeomorphism from S^1 to S^1 which is allowed to have intervals where one or the other variable remains constant while the other continues to increase. (These correspondences arise naturally in computer vision in comparing the images seen by the right and left eyes, see [**11**].) Then we define the equivalence relation on $\mathrm{Cont}(S^1, \mathbb{R}^2)$ by $c \sim d$ if and only if there is a monotone correspondence R such that for all $\theta, \varphi \in R, c(\theta) = d(\varphi)$. It is easily seen that any non-constant $c \in \mathrm{Cont}(S^1, \mathbb{R}^2)$ is equivalent to an c_1 which is not constant on any intervals in S^1 and that for such c_1's and d_1's, the equivalence relation amounts to $c_1 \circ h \equiv d_1$ for some homeomorphism h of S^1. Let $B_i^{cont}(S^1, \mathbb{R}^2)$ be the quotient space by this equivalence relation. We call these *Fréchet curves.*

The quotient metric on $B_i^{cont}(S^1, \mathbb{R}^2)$ is called the *Fréchet* metric, a variant of the *Hausdorff* metric mentioned in the Introduction, both being L^∞ type metrics. Namely, define

$$\begin{aligned} d_\infty(c, d) &= \inf_{\text{monotone corresp.} R} \left(\sup_{(\theta,\varphi)\in R} |c(\theta) - d(\varphi)| \right) \\ &= \inf_{\text{homeomorph.} h: S^1 \to S^1} \|c \circ h - d\|_\infty. \end{aligned}$$

It is straightforward to check that this makes $B_i^{cont}(S^1, \mathbb{R}^2)$ into a complete metric space.

Another very natural space is the subset $B_i^{lip}(S^1, \mathbb{R}^2) \subset B_i^{cont}(S^1, \mathbb{R}^2)$ given by the non-constant *Lipschitz* maps $c : S^1 \to \mathbb{R}^2$. The great virtue of Lipschitz maps is that their images are rectifiable curves and thus each of them is equivalent to a map d in which θ is proportional to arclength, as in the previous section. More precisely, if c is Lipschitz, then c_θ exists almost everywhere and is bounded and we can reparametrize by:

$$h(\theta) = \int_0^\theta |c_\theta| d\theta \Big/ \int_0^{2\pi} |c_\theta| d\theta,$$

obtaining an equivalent d for which $|d_\theta| \equiv L/2\pi$. This d will be unique up to rotations, i.e. the action of S^1 in the previous section.

This subspace of rectifiable Fréchet curves is the subject of a nice compactness theorem due to Hilbert, namely that the set of all such curves in a closed bounded subset of $\mathbb{R}^2$ and whose length is bounded is compact in the Fréchet metric. This can be seen as follows: we can lift all such curves to specific Lipschitz maps c whose Lipschitz constants are bounded. This set is an equicontinuous set of functions by the bound on the Lipschitz constant. By the Ascoli-Arzela theorem the topology of pointwise convergence equals then the topology of uniform convergence on S^1. So this set is a closed subset in a product of S^1 copies of a large ball in $\mathbb{R}^2$; this product is compact. The Fréchet metric is coarser than the uniform metric, so our set is also compact.

12. Metrics on spaces of curves

12.1 Need for invariance under reparametrization. The pointwise metric on the space of immersions $\mathrm{Imm}(S^1, \mathbb{R}^2)$ is given by

$$G_c(h,k) := \int_{S^1} \langle h(\theta), k(\theta)\rangle d\theta.$$

This Riemannian metric is not invariant under reparameterizations of the variable θ and thus does not induce a sensible metric on the quotient space $B_i(S^1, \mathbb{R}^2)$. Indeed, it induces the zero metric since *for any two curves* $C_0, C_1 \in B_i(S^1, \mathbb{R}^2)$ *the infimum of the arc lengths of curves in* $\mathrm{Imm}(S^1, \mathbb{R}^2)$ *which connect embeddings* $c_0, c_1 \in \mathrm{Imm}(S^1, \mathbb{R}^2)$ *with* $\pi(c_i) = C_i$ *turns out to be zero.* To see this, take any c_0 in the $\mathrm{Diff}(S^1)$-orbit over C_0. Take the following variation $c(\theta, t)$ of c_0: for θ outside a small neighborhood U of length ε of 1 in S^1, $c(\theta, t) = c_0(\theta)$. If $\theta \in U$, then the variation for $t \in [0, 1/2]$ moves the small part of c_0 so that $c(\theta, 1/2)$ for θ in U takes off C_0, goes to C_1, traverses nearly all of C_1, and returns to C_0. Now in the orbit through $c(\cdot, 1/2)$, reparameterize in such a way that the new curve is diligently traversing C_1 for $\theta \notin U$, and for $\theta \in U$ it travels back to C_0, runs along C_0, and comes back to C_1. This reparametrized curve is then varied for $t \in [1/2, 1]$ in such a way, that the part for $\theta \in U$ is moved towards C_2. It is clear that the length of both variations is bounded by a constant (depending on the distance between C_0 and C_1 and the lengths of both C_0 and C_1) times ε.

12.2. The simplest Riemannian metric on B_i. Let $h, k \in C^\infty(S^1, \mathbb{R}^2)$ be two tangent vectors with foot point $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$. The induced volume form is $\mathrm{vol}(c) = \langle \partial_\theta c, \partial_\theta c\rangle^{1/2}\, d\theta = |c_\theta| d\theta$. We consider first the simple H^0 weak Riemannian metric on $\mathrm{Imm}(S^1, \mathbb{R}^2)$:

$$G_c(h,k) := \int_{S^1} \langle h(\theta), k(\theta)\rangle |c'(\theta)|\, d\theta \tag{1}$$

which is invariant under $\mathrm{Diff}(S^1)$. This makes the map $\pi : \mathrm{Imm}(S^1, \mathbb{R}^2) \to B_i(S^1, \mathbb{R}^2)$ into a *Riemannian submersion* (off the singularities of $B_i(S^1, \mathbb{R}^2)$) which is very convenient. We call this the H^0*-metric.*

Now we can determine the bundle $\mathcal{N} \to \operatorname{Imm}(S^1, \mathbb{R}^2)$ of tangent vectors which are normal to the $\operatorname{Diff}(S^1)$-orbits. The tangent vectors to the orbits are $T_c(c \circ \operatorname{Diff}(S^1)) = \{g.c_\theta : g \in C^\infty(S^1, \mathbb{R})\}$. Inserting this for k into the expression (1) of the metric we see that

$$\begin{aligned} \mathcal{N}_c &= \{h \in C^\infty(S^1, \mathbb{R}^2) : \langle h, c_\theta\rangle = 0\} \\ &= \{aic_\theta \in C^\infty(S^1, \mathbb{R}^2) : a \in C^\infty(S^1, \mathbb{R})\} \\ &= \{bn_c \in C^\infty(S^1, \mathbb{R}^2) : b \in C^\infty(S^1, \mathbb{R})\}, \end{aligned} \tag{2}$$

where n_c is the normal unit field along c.

A tangent vector $h \in T_c \operatorname{Imm}(S^1, \mathbb{R}^2) = C^\infty(S^1, \mathbb{R}^2)$ has an orthonormal decomposition

$$\begin{aligned} h &= h^\top + h^\perp \in T_c(c \circ \operatorname{Diff}^+(S^1)) \oplus \mathcal{N}_c \qquad \text{where} \\ h^\top &= \frac{\langle h, c_\theta\rangle}{|c_\theta|^2} c_\theta \in T_c(c \circ \operatorname{Diff}^+(S^1)), \\ h^\perp &= \frac{\langle h, ic_\theta\rangle}{|c_\theta|^2} ic_\theta \in \mathcal{N}_c, \end{aligned} \tag{3}$$

into smooth tangential and normal components.

Since the Riemannian metric G on $\operatorname{Imm}(S^1, \mathbb{R}^2)$ is invariant under the action of $\operatorname{Diff}(S^1)$ it induces a metric on the quotient $B_i(S^1, \mathbb{R}^2)$ as follows. For any $C_0, C_1 \in B_i$, consider all liftings $c_0, c_1 \in \operatorname{Imm}$ such that $\pi(c_0) = C_0, \pi(c_1) = C_1$ and all smooth curves $t \mapsto (\theta \mapsto c(t, \theta))$ in $\operatorname{Imm}(S^1, \mathbb{R}^2)$ with $c(0, \cdot) = c_0$ and $c(1, \cdot) = c_1$. Since the metric G is invariant under the action of $\operatorname{Diff}(S^1)$ the arc-length of the curve $t \mapsto \pi(c(t, \cdot))$ in $B_i(S^1, \mathbb{R}^2)$ is given by

$$\begin{aligned} L_G^{\text{hor}}(c) &:= L_G(\pi(c(t, \cdot))) = \int_0^1 \sqrt{G_{\pi(c)}(T_c\pi.c_t, T_c\pi.c_t)}\, dt = \int_0^1 \sqrt{G_c(c_t^\perp, c_t^\perp)}\, dt \\ &= \int_0^1 \Big(\int_{S^1} \Big\langle \frac{\langle c_t, ic_\theta\rangle}{|c_\theta|^2} ic_\theta, \frac{\langle c_t, ic_\theta\rangle}{|c_\theta|^2} ic_\theta \Big\rangle |c_\theta|\, d\theta \Big)^{\frac{1}{2}} dt \\ &= \int_0^1 \Big(\int_{S^1} \langle c_t, n_c\rangle^2 |c_\theta|\, d\theta \Big)^{\frac{1}{2}} dt \\ &= \int_0^1 \Big(\int_{S^1} \langle c_t, ic_\theta\rangle^2 \frac{d\theta}{|c_\theta|} \Big)^{\frac{1}{2}} dt \end{aligned} \tag{4}$$

The metric on $B_i(S^1, \mathbb{R}^2)$ is defined by taking the infimum of this over all paths c (and all lifts c_0, c_1):

$$\operatorname{dist}_G^{B_i}(C_1, C_2) = \inf_c L_G^{\text{hor}}(c).$$

Unfortunately, we will see below that this metric is too weak: the distance that it defines turns out to be identically zero! For this reason, we will mostly study in this paper a family of stronger metrics. These are obtained by the most minimal change in G. We want to preserve two simple properties of the metric: that it is local and that it has no derivatives in it. The standard way to strengthen the metric is go from an H^0 metric to an H^1 metric. But when we work out the natural H^1 metric, picking out those terms which are local and do not involve derivatives leads us to our chosen metric.

We consider next the H^1 weak Riemannian metric on $\mathrm{Imm}(S^1,\mathbb{R}^2)$:

$$G^1_c(h,k) := \int_{S^1} \big(\langle h(\theta),k(\theta)\rangle + A\frac{\langle h_\theta,k_\theta\rangle}{|c_\theta|^2}\big)|c_\theta|\,d\theta. \tag{5}$$

which is invariant under $\mathrm{Diff}(S^1)$. Thus $\pi : \mathrm{Imm}(S^1,\mathbb{R}^2) \to B_i(S^1,\mathbb{R}^2)$ is again a *Riemannian submersion* off the singularities of $B_i(S^1,\mathbb{R}^2)$. We call this the H^1*-metric* on B_i.

To understand this metric better, we assume $h = k = a\frac{ic_\theta}{|c_\theta|} + b\frac{c_\theta}{|c_\theta|}$. Moreover, for any function $f(\theta)$, we write $f_s = \frac{f_\theta}{|c_\theta|}$ for the derivative with respect to arc length. Then:

$$h_s = \frac{h_\theta}{|c_\theta|} = (aic_s + bc_s)_s = (a_s + \kappa b)ic_s + (b_s - \kappa a)c_s.$$

Therefore:

$$\begin{aligned}
G^1_c(h,h) &= \int_{S^1} \big(a^2 + b^2 + A(a_s+\kappa b)^2 + A(b_s-\kappa a)^2\big)ds \\
&= \int_{S^1} \big(a^2(1+A\kappa^2) + Aa_s^2\big) + 2A\kappa(a_s b - b_s a) + \big(b^2(1+A\kappa^2) + Ab_s^2\big)ds
\end{aligned}$$

Letting T_1 and T_2 be the differential operators $T_1 = I + A\kappa^2 - A(\frac{d}{ds})^2$, $T_2 = A(\kappa_s + 2\kappa\frac{d}{ds})$, then integrating by parts on S^1, we get:

$$G^1_c(h,h) = \int_{S^1} \big(T_1(a).a + 2T_2(a).b + T_1(b).b\big)ds.$$

Note that T_1 is a positive definite self-adjoint operator on functions on c, hence it has an inverse given by a Green's function which we write T_1^{-1}. Completing the square and using that T_1 is self-adjoint, we simplify the metric to:

$$G^1_c(h,h) = \int_c \Big(T_1(a).a - T_1^{-1}(T_2(a)).T_2(a) + T_1\big(b + T_1^{-1}(T_2(a))\big).\big(b + T_1^{-1}(T_2(a))\big)\Big)ds.$$

If we fix a and minimize this in b, we get the bundle $\mathcal{N}^1 \to \mathrm{Imm}(S^1,\mathbb{R}^2)$ of tangent vectors which are G^1-normal to the $\mathrm{Diff}(S^1)$-orbits. In other words:

$$\mathcal{N}^1_c = \{h \in C^\infty(S^1,\mathbb{R}^2) : h = aic_s + bc_s, b = -T_1^{-1}(T_2(a))\}$$

and on horizontal vectors of this type:

$$G^1_c(h,h) = \int_c \big((1+A\kappa^2)a^2 + Aa_s^2\big)ds - \int_c T_1^{-1}(T_2(a)).T_2(a)ds.$$

If we drop terms involving a_s, say because we assume $|a_s|$ is small, then what remains is just the integral of $(1+A\kappa^2)a^2$ plus the integral of $T_1^{-1}(\kappa_s a)\kappa_s a$. The second is a non-local regular integral operator, so dropping this we are left with the main metric of this paper:

$$G^A_c(h,h) = \int_c (1+A\kappa^2)a^2 ds, h = aic_s$$

which we call the H^0_κ-*metric with curvature weight* A. For further reference, on $\mathrm{Imm}(S^1,\mathbb{R}^2)$, for a constant $A \geq 0$, it is given by

$$G^A_c(h,k) := \int_{S^1} (1 + A\kappa_c(\theta)^2)\langle h(\theta), k(\theta)\rangle |c'(\theta)|\, d\theta \tag{6}$$

which is again invariant under $\mathrm{Diff}(S^1)$. Thus $\pi : \mathrm{Imm}(S^1,\mathbb{R}^2) \to B_i(S^1,\mathbb{R}^2)$ is again a *Riemannian submersion* off the singularities. Note that for this metric (6), the bundle $\mathcal{N} \subset T\,\mathrm{Imm}(S^1,\mathbb{R}^2)$ is the same as for $A = 0$, as described in (2). The arc-length of a curve $t \mapsto \pi(c(t,\cdot))$ in $B_i(S^1,\mathbb{R}^2)$ is given by the analogon of (4)

$$\begin{aligned} L^{\mathrm{hor}}_{G^A}(c) := L_{G^A}(\pi(c(t,\cdot))) &= \int_0^1 \sqrt{G^A_{\pi(c)}(T_c\pi.c_t, T_c\pi.c_t)}\, dt = \int_0^1 \sqrt{G^A_c(c_t^\perp, c_t^\perp)}\, dt \\ &= \int_0^1 \Big(\int_{S^1} (1+A\kappa_c^2)\langle c_t, n_c\rangle^2 |c_\theta|\, d\theta\Big)^{\frac{1}{2}} dt \\ &= \int_0^1 \Big(\int_{S^1} (1+A\kappa_c^2)\langle c_t, ic_\theta\rangle^2 \frac{d\theta}{|c_\theta|}\Big)^{\frac{1}{2}} dt \end{aligned} \tag{7}$$

The metric on $B_i(S^1,\mathbb{R}^2)$ is defined by taking the infimum of this over all paths c (and all lifts c_0, c_1):

$$\mathrm{dist}^{B_i}_{G^A}(C_1, C_2) = \inf_c L^{\mathrm{hor}}_{G^A}(c).$$

Note that if a path $\pi(c)$ in $B_i(S^1,\mathbb{R}^2)$ is given, then one can choose its lift to a path c in $\mathrm{Imm}(S^1,\mathbb{R}^2)$ to have various good properties. Firstly, we can choose the lift $c(0,\quad)$ of the inital curve to have a parametrization of constant speed, i.e. if its length is ℓ, then $|c_\theta|(\theta,0) = \ell/2\pi$ for all $\theta \in S^1$. Secondly, we can make the tangent vector to c everywhere horizontal, i.e. $< c_t, c_\theta >\equiv 0$, by 11.5. Thirdly, we can reparametrize the coordinate t on the path of length L so that the path is traversed at constant speed, i.e.

$$\int_{S^1} (1+A\kappa_c^2) < c_t, ic_\theta >^2 d\theta/|c_\theta| \equiv L^2, \text{ for all } 0 \leq t \leq 1.$$

12.3. A Lipschitz bound for arc length in G^A. We apply the Cauchy-Schwarz inequality to the derivative 11.2.4 of the length function along a path $t \mapsto c(t, \quad)$:

$$\begin{aligned}
\partial_t \ell(c) = d\ell(c)(c_t) &= -\int_{S^1} \kappa(c)\langle c_t, n_c\rangle |c_\theta| d\theta \le \Big| \int_{S^1} \kappa(c)\langle c_t, n_c\rangle |c_\theta| d\theta \Big| \\
&\le \Big(\int_{S^1} 1^2 |c_\theta| d\theta \Big)^{\frac{1}{2}} \Big(\int_{S^1} \kappa(c)^2 \langle c_t, n_c\rangle^2 |c_\theta| d\theta \Big)^{\frac{1}{2}} \\
&\le \ell(c)^{\frac{1}{2}} \frac{1}{\sqrt{A}} \Big(\int_{S^1} (1 + A\kappa(c)^2) \langle c_t, n_c\rangle^2 |c_\theta| d\theta \Big)^{\frac{1}{2}}
\end{aligned}$$

Thus

$$\partial_t(\sqrt{\ell(c)}) = \frac{\partial_t \ell(c)}{2\sqrt{\ell(c)}} \le \frac{1}{2\sqrt{A}} \Big(\int_{S^1} (1 + A\kappa(c)^2) \langle c_t, n_c\rangle^2 |c_\theta| d\theta \Big)^{\frac{1}{2}}$$

and by using (12.2.7) we get

$$\begin{aligned}
\sqrt{\ell(c_1)} - \sqrt{\ell(c_0)} &= \int_0^1 \partial_t(\sqrt{\ell(c)})\, dt \\
&\le \frac{1}{2\sqrt{A}} \int_0^1 \Big(\int_{S^1} (1 + A\kappa(c)^2) \langle c_t, n_c\rangle^2 |c_\theta| d\theta \Big)^{\frac{1}{2}} dt \\
&= \frac{1}{2\sqrt{A}} L^{\text{hor}}_{G^A}(c). \qquad (1)
\end{aligned}$$

If we take the infimum over all paths connecting c_0 with the $\operatorname{Diff}(S^1)$-orbit through c_1 we get:

Lipschitz continuity of $\sqrt{\ell} : B_i(S^1, \mathbb{R}^2) \to \mathbb{R}_{\ge 0}$. *For C_0 and C_1 in $B_i(S^1, \mathbb{R}^2) = \operatorname{Imm}(S^1, \mathbb{R}^2)/\operatorname{Diff}(S^1)$ we have for $A > 0$:*

$$\sqrt{\ell(C_1)} - \sqrt{\ell(C_0)} \le \frac{1}{2\sqrt{A}} \operatorname{dist}^{B_i}_{G^A}(C_1, C_2). \qquad (2)$$

12.4. Bounding the area swept by a path in B_i. Secondly, we want to bound the area swept out by a path starting from C_0 to reach any curve C_1 nearby in our metric. First we use the Cauchy-Schwarz inequality in the Hilbert space $L^2(S^1, |c_\theta(t,\theta)| d\theta)$ to get

$$\begin{aligned}
&\int_{S^1} 1.|c_t(t,\theta)||c_\theta(t,\theta)| d\theta = \langle 1, |c_t|\rangle_{L^2} \le \\
&\qquad \le \|1\|_{L^2} \|c_t\|_{L^2} = \Big(\int_{S^1} |c_\theta(t,\theta)|\, d\theta \Big)^{\frac{1}{2}} \Big(\int_{S^1} |c_t(t,\theta)|^2 |c_\theta(t,\theta)|\, d\theta \Big)^{\frac{1}{2}}.
\end{aligned}$$

Now we assume that the variation $c(t, \theta)$ is horizontal, so that $\langle c_t, c_\theta\rangle = 0$. Then $L_{G^A}(c) = L^{\text{hor}}_{G^A}(c)$. We use this inequality and then the intermediate value theorem of integral calculus to obtain

$$\begin{aligned}
L^{\mathrm{hor}}_{G^A}(c) = L_{G^A}(c) &= \int_0^1 \sqrt{G^A_c(c_t,c_t)}\,dt \\
&= \int_0^1 \Big(\int_{S^1} (1+A\kappa(c)^2)|c_t(t,\theta)|^2|c_\theta(t,\theta)|\,d\theta\Big)^{\frac12} dt \\
&\ge \int_0^1 \Big(\int_{S^1} |c_t(t,\theta)|^2|c_\theta(t,\theta)|\,d\theta\Big)^{\frac12} dt \\
&\ge \int_0^1 \Big(\int_{S^1} |c_\theta(t,\theta)|\,d\theta\Big)^{-\frac12} \int_{S^1} |c_t(t,\theta)||c_\theta(t,\theta)|\,d\theta\,dt \\
&= \Big(\int_{S^1} |c_\theta(t_0,\theta)|\,d\theta\Big)^{-\frac12} \int_0^1\int_{S^1} |c_t(t,\theta)||c_\theta(t,\theta)|\,d\theta\,dt \\
&\qquad \text{for some intermediate value } 0\le t_0\le 1, \\
&= \frac{1}{\sqrt{\ell(c(t_0,\cdot))}} \int_{[0,1]\times S^1} |\det dc(t,\theta)|\,d\theta\,dt.
\end{aligned}$$

Area swept out bound. *If c is any path from C_0 to C_1, then*

$$\begin{pmatrix}\textit{area of the region swept}\\ \textit{out by the variation } c\end{pmatrix} \le \max_t \sqrt{\ell(c(t,\cdot))}\cdot L^{hor}_{G^A}(c). \tag{1}$$

This result enables us to compare the double cover $B_i^{\mathrm{or}}(S^1,\mathbb{R}^2)$ of our metric space $B_i(S^,\mathbb{R}^2)$ consisting of oriented unparametrized curves to the fundamental space of geometric measure theory. Note that there is a map h_1 from B_i^{or} to the space of 1-currents $\mathcal{D}_1'$ given by:

$$< h_1(c \bmod \mathrm{Diff}^+(S^1)),\omega >= \int_{S^1} c^*\omega, \quad c\in \mathrm{Imm}(S^1,\mathbb{R}^2).$$

The image $h_1(C)$ is, in fact, closed. For any C, define the integer-valued measurable function w_C on $\mathbb{R}^2$ by:

$$w_C((x,y)) = \text{winding number of } C \text{ around } (x,y).$$

Then it is easy to see that, as currents, $h_1(C)=\partial(w_C dxdy)$, hence $\partial h_1(C) = 0$.

Although h_1 is obviously injective on the space B_e, it is not injective on B_i as illustrated in Figure 1 below. The image of this mapping lies in the basic subset $\mathcal{I}_{1,c}\subset\mathcal{D}_1'$ of closed *integral* currents, namely those which are both closed and countable sums of currents defined by Lipschitz mappings $c_i:[0,1]\to\mathbb{R}^2$ of finite total length. Integral currents carry what is called the *flat* metric, which, for closed 1-currents, reduces (by the isoperimetric

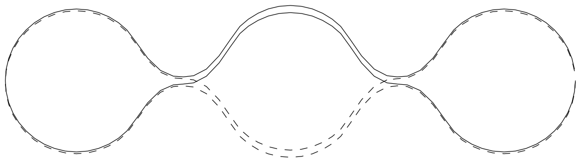


Figure 2. Two distinct immersions of S^1 in the plane whose underlying currents are equal. One curve is solid, the other dashed.

inequality) to the area distance

$$d^\flat(C_1, C_2) = \iint_{\mathbb{R}^2} |w_{C_1} - w_{C_2}| dxdy. \tag{2}$$

To connect this with our 'area swept out bound', note that if we have any path c in $\mathrm{Imm}(S^1, \mathbb{R}^2)$ joining C_1 and C_2, this path defines a 2-current $w(c)$ such that $\partial w(c) = h_1(C_1) - h_1(C_2)$ and

$$\int_{\mathbb{R}^2} |w(c)|\, dxdy \leq \int_0^1 \int_{S^1} |\det c| d\theta dt$$

which is what we are calling the area swept out. But $\partial(w_{C_1} - w_{C_2}) = h_1(C_1) - h_1(C_2)$ too, so $w(c) = w_{C_1} - w_{C_2}$. Thus

$$d^\flat(C_1, C_2) \leq \min_{\text{all paths } c \text{ joining } C_1, C_2} \big[\text{area swept out by } c\big] \tag{3}$$

Finally, we recall the fundamental compactness result of geometric measure theory in this simple case: the space of integral 1-currents of bounded length is compact in the flat metric. This implies that our 'area swept out bound' above has the Corollary:

Corollary.

(4) *If $\{C_n\}$ is any Cauchy sequence in B_i for the metric dist_{G^A}, then $\{h_1(C_n)\}$ is a Cauchy sequence in $\mathcal{I}_{1,c}$ on which length is bounded.*

(5) *Hence h_1 extends to a continuous map from the completion $\overline{B_i}$ of B_i in the metric G^A to $\mathcal{I}_{1,c}$.*

12.5. Bounding how far curves move in small paths in B_i. We want to bound the maximum distance a curve C_0 can move on any path whose length is small in G^A metric. Fix the initial curve C_0 and let ℓ be its length. The result is:

Maximum distance bound. *Let $\epsilon < \min\{2\sqrt{A\ell}, \ell^{3/2}\}/8$ and consider $\eta = 4(\ell^{3/4}A^{-1/4} + \ell^{1/4})\sqrt{\epsilon}$. Then for any path c starting at C_0 whose length is ϵ, the final curve lies in the tubular neighborhood of C_0 of width η. More* *precisely, if we choose the path $c(t,\theta)$ to be horizontal, then* $\max_\theta |c(0,\theta) - c(1,\theta)| < \eta$.

Proof. For all of this proof, we assume the path in B_i has been lifted to a horizontal path $c \in \mathrm{Imm}(S^1, \mathbb{R}^2)$ with $|c_\theta|(\theta, 0) \equiv \ell/2\pi$, so that $\langle c_t, c_\theta \rangle \equiv 0$, and also $\int_{S^1}(1 + A\kappa_c^2)|c_t|^2|c_\theta|\, d\theta \equiv \epsilon^2$. The first step in the proof is to refine the Lipschitz bound on the length of a curve to a local estimate. Note that by horizontality

$$\frac{\partial}{\partial t}\sqrt{|c_\theta|} = \frac{< c_{\theta t}, c_\theta >}{2|c_\theta|^{3/2}} = -\frac{< c_t, c_{\theta\theta} >}{2|c_\theta|^{3/2}} = -\frac{< c_t, ic_\theta >}{2|c_\theta|}\kappa_c\, |c_\theta|^{1/2} = \mp\tfrac{1}{2}\kappa_c\, |c_t||c_\theta|^{1/2}$$

hence

$$\int_{S^1}\Big(\frac{\partial}{\partial t}\sqrt{|c_\theta|}\Big)^2 ds \le \frac{\epsilon^2}{4A}.$$

Now we make the key definition:

$$\widetilde{|c_\theta|}(t,\theta) = \min_{0\le t_1 \le t} |c_\theta|(t_1, \theta).$$

Note that the t-derivative of $\widetilde{|c_\theta|}$ is either 0 or equal to that of $|c_\theta|$ and is ≤ 0. Thus we have:

$$\begin{aligned}
\int_{S^1}\Big(\sqrt{\frac{\ell}{2\pi}} - \sqrt{\widetilde{|c_\theta|}(1,\theta)}\Big)d\theta &\le \int_0^1\int_{S^1} -\frac{\partial}{\partial t}\sqrt{\widetilde{|c_\theta|}}\,d\theta dt \\
&\le \int_0^1\int_{S^1}\Big|\frac{\partial}{\partial t}\sqrt{|c_\theta|}\Big|d\theta dt \\
&\le \int_0^1\Big(\int_{S^1} d\theta\Big)^{1/2}\cdot\Big(\int_{S^1}\Big|\frac{\partial}{\partial t}\sqrt{|c_\theta|}\Big|^2 d\theta\Big)^{1/2} dt \\
&\le \sqrt{2\pi}\cdot\frac{\epsilon}{2\sqrt{A}}.
\end{aligned}$$

To make use of this inequality, let $E = \big\{\theta : \widetilde{|c_\theta|}(1,\theta) \le \big(1-(A\ell)^{-1/4}\sqrt{\varepsilon}\big)\ell/2\pi\big\}$. Our assumption on ε gives $(A\ell)^{-1/4}\sqrt{\varepsilon} < 1/2$, hence on $S^1 \setminus E$ we have $\widetilde{|c_\theta|} > \ell/4\pi$. On E we have also $(\widetilde{|c_\theta|})^{1/2} \le (1-(A\ell)^{-1/4}\sqrt{\varepsilon}/2)\sqrt{\ell/2\pi}$. Combining this with the previous inequality, we get (where $\mu(E)$is the measure of E):

$$\mu(E)\frac{1}{2\sqrt{2\pi}}\Big(\frac{\ell}{A}\Big)^{1/4}\sqrt{\varepsilon} \le \sqrt{2\pi}\cdot\frac{\varepsilon}{2\sqrt{A}}, \quad \text{hence} \quad \mu(E) \le 2\pi\frac{\sqrt{\varepsilon}}{(A\ell)^{1/4}} < \pi.$$

We now use the lower bound on $|c_\theta|$ on $S^1 - E$ to control $c(1,\theta) - c(0,\theta)$:

$$\begin{aligned}
\int_{S^1-E}|c(1,\theta) - c(0,\theta)|\, d\theta &\le \int_0^1\int_{S^1-E}|c_t|\, d\theta dt \\
&\le \sqrt{2\pi}\cdot\int_0^1\Big(\int_{S^1-E}|c_t|^2 d\theta\Big)^{1/2} dt
\end{aligned}$$

$$\leq \frac{\sqrt{2\pi}}{\sqrt{\frac{\ell}{4\pi}}} \int_0^1 \Big(\int_{S^1 - E} |c_t|^2 |c_\theta| d\theta \Big)^{1/2} dt \leq \frac{2\sqrt{2}\pi}{\sqrt{\ell}} \cdot \varepsilon$$

Again, introduce a small exceptional set $F = \{\theta \mid \theta \notin E \text{ and } |c(1,\theta) - c(0,\theta)| \geq \ell^{1/4}\sqrt{\varepsilon}\}$. By the inequality above, we get:

$$\mu(F) \cdot \ell^{1/4}\sqrt{\varepsilon} \leq \frac{2\sqrt{2}\pi\varepsilon}{\sqrt{\ell}}, \quad \text{hence } \mu(F) \leq \frac{2\sqrt{2}\pi\sqrt{\varepsilon}}{\ell^{3/4}} < \pi.$$

The last inequality follows from the second assumption on ε. Knowing $\mu(E)$ and $\mu(F)$ gives us the lengths $|c(0,E)|$ and $|c(0,F)|$ in $\mathbb{R}^2$. But we need the lengths $|c(1,E)|$ and $|c(1,F)|$ too. We get these using the fact that the whole length of C_1 can't be too large, by 12.3:

$$\sqrt{|C_1|} \leq \sqrt{\ell} + \frac{\varepsilon}{2\sqrt{A}}, \quad \text{hence}$$
$$|C_1| \leq \ell + 2\varepsilon\sqrt{\frac{\ell}{A}} \leq \ell + \sqrt{\varepsilon} \cdot \frac{\ell^{3/4}}{A^{1/4}}.$$

On $S^1 \setminus E$ we have $\widetilde{|c_\theta|} > (1 - (A\ell)^{-1/4}\sqrt{\varepsilon})\ell/2\pi$, thus we get

$$\begin{aligned}
|c(1, E \cup F)| &= |C_1| - |c(1, S^1 \setminus (E \cup F))| \\
&\leq \ell + \sqrt{\varepsilon} \cdot \frac{\ell^{3/4}}{A^{1/4}} - \Big(1 - \frac{\sqrt{\varepsilon}}{(A\ell)^{1/4}}\Big) \frac{\ell}{2\pi} \Big(2\pi - \mu(E \cup F)\Big) \\
&\leq \sqrt{\varepsilon} \cdot \Big(3\frac{\ell^{3/4}}{A^{1/4}} + \sqrt{2}\ell^{1/4}\Big)
\end{aligned}$$

Finally, we can get from $c(0,\theta)$ to $c(1,\theta)$ by going via $c(0,\theta')$ and $c(1,\theta')$ where $\theta' \in S^1 \setminus (E \cup F) \neq \emptyset$. Thus

$$\begin{aligned}
\max_\theta |c(0,\theta) - c(1,\theta)| &\leq |c(0, E \cup F)| + \ell^{1/4}\sqrt{\varepsilon} + |c(E \cup F, 1)| \\
&\leq 4(\ell^{3/4}A^{-1/4} + \ell^{1/4})\sqrt{\varepsilon} \quad \square
\end{aligned}$$

Combining this bound with the Lipschitz continuity of the square root of arc length, we get:

12.6. Corollary. *For any $A > 0$, the map from $B_i(S^1, \mathbb{R}^2)$ in the* $\operatorname{dist}_{G^A}$ *metric to the space $B_i^{cont}(S^1, \mathbb{R}^2)$ in the Fréchet metric is continuous, and, in fact, uniformly continuous on every subset where the length ℓ is bounded. In particular,* $\operatorname{dist}_{G^A}$ *is a separating metric on $B_i(S^1, \mathbb{R}^2)$. Moreover, the completion $\overline{B_i}(S^1, \mathbb{R}^2)$ of $B_i(S^1, \mathbb{R}^2)$ in this metric can be identified with a subset of $B_i^{lip}(S^1, \mathbb{R}^2)$.*

If we iterate this bound, then we get the following:

12.7. Corollary. *Consider all paths in B_i joining curves C_0 and C_1. Let L be the length of such a path in the dist_{G^A} metric and let $\ell_{\min}, \ell_{\max}$ be the minimum and maximum of the arc lengths of the curves in this path. Then there are parametrizations c_0, c_1 of C_0 and C_1 such that:*

$$\max_\theta |c_0(\theta) - c_1(\theta)| \leq 50 \max(LF^*, \sqrt{\ell_{\max} LF^*}), \text{ where}$$

$$F^* = \max\Big(\frac{1}{\sqrt{\ell_{\min}}}, \sqrt{\frac{\ell_{\max}}{A}}\Big).$$

To prove this, you need only break up the path into a minimum number of pieces for which the maximum distance bound 12.5 holds and add together the estimates for each piece. We will only sketch this proof which is straightforward. The constant 50 is just what comes out without attempting to optimize the bound. The second option for bound, $50\sqrt{\ell_{\max} LF^*}$ is just a rephrasing of the bound already in the theorem for short paths. If the path is too long to satisfy the condition of the theorem, we break the path at intermediate curves C_i of length ℓ_i such that each begins a subpath with length $\varepsilon_i = \min(\sqrt{A\ell_i}, \ell_i^{3/2})/8$ and which don't overlap for more than 2:1. Thus $\sum_i \varepsilon_i \leq 2L$. Then apply the maximum distance bound 12.5 to each piece, letting η_i be the bound on how far points move in this subpath *or any parts thereof* and verify:

$$\eta_i \leq 2\sqrt{2}\ell_i \leq 16\sqrt{2}\varepsilon_i F^*,$$

from which we get what we need by summing over i.

12.8. A final Corollary shows that if we parametrize any path appropriately, we get explicit equicontinuous continuity bounds on the parametrization depending only on $L, \ell_{\max}$ and $\ell_{\min}$. This is a step towards establishing the existence of weak geodesics. The idea is this: instead of the horizontal parametrization $\langle c_t, c_\theta \rangle \equiv 0$, we parametrize each curve at constant speed $|c_\theta| \equiv \ell(t)/2\pi$ where $\ell(t)$ is the length of the t^{th} curve and ask only that $\langle c_t, c_\theta\rangle(0,t) \equiv 0$ for some base point $0 \in [0, 2\pi]$, see 11.8. Then we get:

Corollary. *If a path $c(t,\theta), 0 \leq t \leq 1$ satisfies*

$$\begin{aligned} &|c_\theta(\theta, t)| \equiv \ell(t)/2\pi && \text{for all } \theta, t \\ &\langle c_t, c_\theta \rangle (0, t) \equiv 0 && \text{for all } t \text{ and} \\ &\int_{C_t} (1 + A\kappa_{C_t}^2) |\langle c_t, i c_\theta\rangle|^2 d\theta / |c_\theta| \equiv L^2 && \text{for all } t, \end{aligned}$$

then

$$|c(t_1, \theta_1) - c(t_2, \theta_2)| \leq \frac{\ell_{\max}}{2\pi} |\theta_1 - \theta_2| + 7(\ell_{\max}^{3/4}/A^{1/4} + \ell_{\max}^{1/4})\sqrt{L(t_1 - t_2)}$$

whenever $|t_1 - t_2| \leq \min(2\sqrt{A\ell_{\min}}, \ell_{\min}^{3/2})/(8L)$.

Proof. We need to compare the constant speed parametrization here with the horizontal parametrization – call it c^* – used in the maximum distance bound 12.5. Under the horizontal parametrization, let the point (t_1, θ_1) on C_{t_1} correspond to (t_2, θ_1^*) on C_{t_2}, i.e. $c(t_2, \theta_1^*) = c^*(t_2, \theta_1)$. Let $C = (\ell_{\max}^{3/4}/A^{1/4} + \ell_{\max}^{1/4})$. Then we know from 12.5 that

$$|c(t_1, \theta_1) - c(t_2, \theta_1^*)| \leq 4C\sqrt{L(t_1 - t_2)}.$$

To compare θ_1 and θ_1^*, we use the properties of the set E in the proof of 12.5 to estimate:

$$\begin{aligned}
\frac{(\theta_1^* - \theta_1)\ell_2}{2\pi} &= \int_0^{\theta_1} |c_\theta^*(t_2, \varphi)| d\varphi - \frac{\theta_1 \ell_2}{2\pi} \\
&\geq \Big(1 - \frac{\sqrt{L(t_1 - t_2)}}{(A\ell_1)^{1/4}}\Big)(\theta_1 - \mu(E))\frac{\ell_1}{2\pi} - \frac{\theta_1 \ell_2}{2\pi} \\
&\geq -2\ell_1 \frac{\sqrt{L(t_1 - t_2)}}{(A\ell_1)^{1/4}} - |\ell_1 - \ell_2| \text{ and similarly} \\
\frac{((2\pi - \theta_1^*) - (2\pi - \theta_1))\ell_2}{2\pi} &= \int_{\theta_1}^{2\pi} |c_\theta^*(t_2, \varphi)| d\varphi - \frac{(2\pi - \theta_1)\ell_2}{2\pi} \\
&\geq -2\ell_1 \frac{\sqrt{L(t_1 - t_2)}}{(A\ell_1)^{1/4}} - |\ell_1 - \ell_2|
\end{aligned}$$

Combining these and using the Lipschitz property of length, we get:

$$\begin{aligned}
\frac{|\theta_1^* - \theta_1|\ell_2}{2\pi} &\leq 2C\sqrt{L(t_1 - t_2)} + 2|\sqrt{\ell_1} - \sqrt{\ell_2}|\sqrt{\ell_{\max}} \\
&\leq 2C\sqrt{L(t_1 - t_2)} + \sqrt{\ell_{\max}} \frac{L(t_1 - t_2)}{\sqrt{A}} \leq \frac{5}{2} C\sqrt{L(t_1 - t_2)}
\end{aligned}$$

Thus, finally:

$$\begin{aligned}
|c(t_1, \theta_1) - c(t_2, \theta_2)| &\leq |c(t_1, \theta_1) - c(t_2, \theta_1^*)| + \\
&\quad + |c(t_2, \theta_1^*) - c(t_2, \theta_1)| + |c(t_2, \theta_1) - c(t_2, \theta_2)| \\
&\leq 4C\sqrt{L(t_1 - t_2)} + \frac{5}{2} C\sqrt{L(t_1 - t_2)} + \frac{\ell_{\max}}{2\pi}|\theta_1 - \theta_2|. \quad \square
\end{aligned}$$

12.9. One might also ask whether the maximum distance bound 12.5 can be strengthened to assert that the 1-jets of such curves C must be close to the 1-jets of C_0. The answer is NO, as is easily seen from looking a small wavelet-type perturbations of C_0. Specifically, calculate the length of the path: $c(t, \theta) = c_0(\theta) + t \cdot af(\theta/a) \cdot i(c_0)_\theta(\theta), 0 \leq t \leq 1$ where $f(x)$ is an *arbitrary* C^2 function with compact support and a is very small. We claim the length of this path is $O(\sqrt{a})$, while the 1-jet at the point $\theta = 0$ of the final curve of the path approaches $(1 + if'0)(c_0)_\theta(0)$.

We sketch the proof, which is straightforward. Let $C_{a,t}$ be the curves on this path. Then $\sup|c_t| = O(a)$, $\sup|\kappa_{C_{a,t}}| = O(1/a)$, $A \le |c_\theta| \le B$ for suitable $A, B > 0$ and $\ell(\text{support}(c_t)) = O(a)$. Then the integral $\int_{S^1}(1 + A\kappa_c^2)(c_t, ic_\theta)^2\frac{d\theta}{|c_\theta|}$ breaks up into 2 pieces, the first being $O(a^2)$, the second being $O(1)$ and the integral vanishing outside an interval of length $O(a)$. Thus the total distance is $O(\sqrt{a})$.

12.10. The H^0-distance on $B_i(S^1,\mathbb{R}^2)$ vanishes. Let $c_0, c_1 \in \operatorname{Imm}(S^1,\mathbb{R}^2)$ be two immersions, and suppose that $t \mapsto (\theta \mapsto c(t,\theta))$ is a smooth curve in $\operatorname{Imm}(S^1,\mathbb{R}^2)$ with $c(0,\cdot) = c_0$ and $c(1,\cdot) = c_1$.

The arc-length for the H^0-metric of the curve $t \mapsto \pi(c(t,\cdot))$ in $B_i(S^1,\mathbb{R}^2)$ is given by 12.2.7 as

$$(1)\qquad L^{\text{hor}}_{G^0}(c) = \int_0^1 \Big(\int_{S^1} \langle c_t, ic_\theta\rangle^2 \frac{d\theta}{|c_\theta|}\Big)^{\frac{1}{2}} dt$$

Theorem. *For $c_0, c_1 \in \operatorname{Imm}(S^1,\mathbb{R}^2)$ there exists always a path $t \mapsto c(t,\cdot)$ with $c(0,\cdot) = c_0$ and $\pi(c(1,\cdot)) = \pi(c_1)$ such that $L^{hor}_{G^0}(c)$ is arbitrarily small.*

Heuristically, the reason for this is that if the curve is made to zig-zag wildly, say with teeth at an angle α, then the length of the curve goes up by a factor $1/\cos(\alpha)$ but the *normal* component of the motion of the curve goes down by the factor $\cos(\alpha)$ – and this normal component is squared, hence it dominates.

Proof. Take a path $c(t,\theta)$ in $\operatorname{Imm}(S^1,\mathbb{R}^2)$ from c_0 to c_1 and make it horizontal using 11.5 so that that $\langle c_t, c_\theta\rangle = 0$; this forces a reparametrization on c_1.

Now let us view c as a smooth mapping $c : [0,1]\times[0,1] \to \mathbb{R}^2$. We shall use the piecewise linear reparameterization $(\varphi(t,\theta),\theta)$ of the square shown above, which for $0 \le t \le 1/2$ deforms the straight line into a zig-zag of height 1 and period $n/2$ connecting the two end-curves, and then removes the teeth for $1/2 \le t \le 1$. In detail: Let $\tilde c(t,\theta) = c(\varphi(t,\theta),\theta)$ where

$$\varphi(t,\theta) = \begin{cases} 2t(2n\theta - 2k) & \text{for } 0 \le t \le 1/2,\ \frac{2k}{2n} \le \theta \le \frac{2k+1}{2n} \\ 2t(2k+2-2n\theta) & \text{for } 0 \le t \le 1/2,\ \frac{2k+1}{2n} \le \theta \le \frac{2k+2}{2n} \\ 2t-1+2(1-t)(2n\theta-2k) & \text{for } 1/2 \le t \le 1,\ \frac{2k}{2n} \le \theta \le \frac{2k+1}{2n} \\ 2t-1+2(1-t)(2k+2-2n\theta) & \text{for } 1/2 \le t \le 1,\ \frac{2k+1}{2n} \le \theta \le \frac{2k+2}{2n}. \end{cases}$$

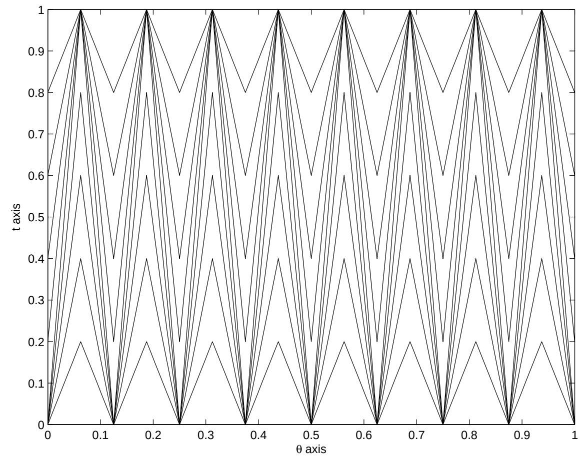


Figure 3. The reparametrization of a path of curves used to make its length arbitrarily small.

Then we get $\tilde{c}_\theta = \varphi_\theta . c_t + c_\theta$ and $\tilde{c}_t = \varphi_t . c_t$ where

$$\varphi_\theta = \begin{cases} +4nt \\ -4nt \\ +4n(1-t) \\ -4n(1-t) \end{cases}, \qquad \varphi_t = \begin{cases} 4n\theta - 4k \\ 4k + 4 - 4n\theta \\ 2 - 4n\theta + 4k \\ -(2 - 4n\theta + 4k) \end{cases}.$$

Also, $\langle c_t, c_\theta \rangle = 0$ implies $\langle \tilde{c}_t, i\tilde{c}_\theta \rangle = \varphi_t . |c_t| . |c_\theta|$ and $|\tilde{c}_\theta| = |c_\theta| \sqrt{1 + \varphi_\theta^2 (|c_t|/|c_\theta|)^2}$. Thus

$$L^{\text{hor}}(\tilde{c}) = \int_0^1 \Big(\int_0^1 \langle \tilde{c}_t, i\tilde{c}_\theta \rangle^2 \frac{d\theta}{|\tilde{c}_\theta|} \Big)^{\frac{1}{2}} dt = \int_0^1 \Big(\int_0^1 \frac{\varphi_t^2 |c_t|^2 |c_\theta|}{\sqrt{1 + \varphi_\theta^2 (\frac{|c_t|}{|c_\theta|})^2}} d\theta \Big)^{\frac{1}{2}} dt =$$

$$= \int_0^{\frac{1}{2}} \Bigg(\sum_{k=0}^{n-1} \Big(\int_{\frac{2k}{2n}}^{\frac{2k+1}{2n}} \frac{(4n\theta - 4k)^2 |c_t(\varphi,\theta)|^2 |c_\theta(\varphi,\theta)|}{\sqrt{1 + (4nt)^2 (\frac{|c_t(\varphi,\theta)|}{|c_\theta(\varphi,\theta)|})^2}} d\theta +$$

$$+ \int_{\frac{2k+1}{2n}}^{\frac{2k+2}{2n}} \frac{(4k + 4 - 4n\theta)^2 |c_t(\varphi,\theta)|^2 |c_\theta(\varphi,\theta)|}{\sqrt{1 + (4nt)^2 (\frac{|c_t(\varphi,\theta)|}{|c_\theta(\varphi,\theta)|})^2}} d\theta \Big) \Bigg)^{\frac{1}{2}} dt +$$

$$+ \int_{\frac{1}{2}}^1 \Bigg(\sum_{k=0}^{n-1} \Big(\int_{\frac{2k}{2n}}^{\frac{2k+1}{2n}} \frac{(2 - 4n\theta + 4k)^2 |c_t(\varphi,\theta)|^2 |c_\theta(\varphi,\theta)|}{\sqrt{1 + (4n)^2 (1-t)^2 (\frac{|c_t(\varphi,\theta)|}{|c_\theta(\varphi,\theta)|})^2}} d\theta +$$

$$+ \int_{\frac{2k+1}{2n}}^{\frac{2k+2}{2n}} \frac{(2 - 4n\theta + 4k)^2 |c_t(\varphi,\theta)|^2 |c_\theta(\varphi,\theta)|}{\sqrt{1 + (4n)^2 (1-t)^2 (\frac{|c_t(\varphi,\theta)|}{|c_\theta(\varphi,\theta)|})^2}} d\theta \Big) \Bigg)^{\frac{1}{2}} dt$$

The function $|c_\theta(\varphi,\theta)|$ is uniformly bounded above and away from 0, and $|c_t(\varphi,\theta)|$ is uniformly bounded. Thus we may estimate

$$\sum_{k=0}^{n-1}\int_{\frac{2k}{2n}}^{\frac{2k+1}{2n}} \frac{(4n\theta-4k)^2|c_t(\varphi,\theta)|^2|c_\theta(\varphi,\theta)|}{\sqrt{1+(4nt)^2(\frac{|c_t(\varphi,\theta)|}{|c_\theta(\varphi,\theta)|})^2}}d\theta$$
$$\leq O(1)\sum_{k=0}^{n-1}\int_0^{\frac{1}{2n}} \frac{4n^2\theta^2|c_t(\varphi(t,\frac{2k}{2n}+\theta),\frac{2k}{2n}+\theta)|^2}{\sqrt{1+(4nt)^2|c_t(\varphi(t,\frac{2k}{2n}+\theta),\frac{2k}{2n}+\theta)|^2}}d\theta$$

We estimate as follows. Fix $\varepsilon>0$. First we split of the integral $\int_{t=0}^{\varepsilon}$ which is $O(\varepsilon)$ uniformly in n; so for the rest we have $t\geq\varepsilon$. The last sum of integrals is now estimated as follows: Consider first the set of all θ such that $|c_t(\varphi(t,\frac{2k}{2n}+\theta),\frac{2k}{2n}+\theta)|<\varepsilon$ which is a countable disjoint union of open intervals. There we get the estimate $O(1).n.4n^2.\varepsilon^2(\theta^3/3)|_{\theta=0}^{\theta=1/2n}=O(\varepsilon)$, uniformly in n. On the complementary set of all θ where $|c_t(\varphi(t,\frac{2k}{2n}+\theta),\frac{2k}{2n}+\theta)|\geq\varepsilon$ we use also $t\geq\varepsilon$ and estimate by $O(1).n.4n^2.\frac{1}{4n\varepsilon^2}.(\theta^3/3)|_{\theta=0}^{\theta=1/2n}=O(\frac{1}{\varepsilon^2 n})$. The other sums of integrals can be estimated similarly, thus $L^{\text{hor}}(\tilde c)$ goes to 0 for $n\to\infty$. It is clear that one can approximate φ by a smooth function without changing the estimates essentially. $\square$

12.11. Non-smooth curves in the completion of B_i. We have seen in 12.6 that the completion of B_i in the metric G^A lies in the space of Lipschitz maps $c: S^1\to\mathbb{R}^2$ mod monotone correspondences, that is, rectifiable Fréchet immersed curves. But how big is it really? We cannot answer this, but we show, in this section, that certain non-smooth curves are in the completion. To be precise, if c is rectifiable, then we can assume c is parametrized at constant speed $|c_\theta|\equiv L/2\pi$ where L is the length of the curve. Therefore $c_\theta=(L/2\pi)e^{i\alpha(\theta)}$ for some measurable function $\alpha(\theta)$ giving the orientation of the tangent line at almost every point. We will say that a rectifiable curve c is *1-BV* if the function α is of bounded variation. Note that this means that the derivative of α exists as a finite signed measure, hence the curvature of c – which is $(2\pi/L)\alpha'$ – is also a finite signed measure. In particular, there are a countable set of 'vertices' on such a curve, points where α has a discontinuity and the measure giving its curvature has an atomic component. Note that α has left and right limits everywhere and vertices can be assigned angles, namely $\alpha_+(\theta)-\alpha_-(\theta)$.

Theorem. *All 1-BV rectifiable curves are in the completion of B_i with respect to the metric G^A.*

Proof. This is proven using the following lemma:

Lemma. *Let $c(t,\theta)$, $0 < t \leq 1$ be an open path of smooth curves $c(t)$ and let $\alpha(t,\theta) = \arg(c_\theta(\theta,t))$. Assume that*

1. *the length of all curves $c(t)$ is bounded by C_1,*
2. *$|c_t| \leq C_2$, for all (t,θ),*
3. *For all t, the total variation in θ of $\alpha(\theta,t)$ is bounded by C_3 and*
4. *the curvature of $c(t)$ satisfies $|\kappa_{c(t)}(\theta,t)| \leq C_4/t$ for all θ.*

Then the length of this path is bounded by $C_2(\sqrt{C_1} + 2\sqrt{AC_3C_4})$.

To prove the lemma, let s_t be arc length on $c(t)$ and estimate the integral:

$$\begin{aligned}\int_{c(t)} (1 + A\kappa(c(t))(t,\theta)^2)\langle c_t, \frac{ic_\theta}{|c_\theta|}\rangle^2 |c_\theta| d\theta &\leq C_2^2\big(C_1 + A\int_{c(t)} \kappa_{c(t)}^2 ds_t\big)\\ &= C_2^2\big(C_1 + A\int_{c(t)} \kappa_{c(t)}\frac{d\alpha}{ds_t} ds_t\big)\\ &\leq C_2^2\big(C_1 + A\frac{C_4}{t}C_3\big).\end{aligned}$$

Taking the square root of both sides and integrating from 0 to 1, we get the result.

We apply this lemma to the simplest possible smoothing of a 1-BV rectifiable curve c_0:

$$c(t,\theta) = \frac{1}{\sqrt{2\pi}t}\int_{\mathbb{R}} c_0(\theta-\varphi)e^{-\varphi^2/2t^2}d\varphi = \frac{1}{\sqrt{2\pi}t}\int_{\mathbb{R}} c_0(\varphi)e^{-(\theta-\varphi)^2/2t^2}d\varphi, 0 < t \leq 1.$$

Note that t is the standard deviation of the Gaussian, *not* the variance. We assume c_0 has a constant speed parametrization and $c_0' = (L/2\pi)e^{i\alpha}$ as above, where α' is a finite signed measure. Thus:

$$\begin{aligned}c_\theta &= \frac{L}{(2\pi)^{3/2}t}\int_{\mathbb{R}} e^{i\alpha(\theta-\varphi)-\varphi^2/2t^2}d\varphi\\ c_{\theta\theta} &= \frac{iL}{(2\pi)^{3/2}t}\int_{\mathbb{R}} e^{i\alpha(\varphi)-(\theta-\varphi)^2/2t^2}\alpha'(d\varphi)\end{aligned}$$

Moreover, using the second expression for the convolution and the heat equation for the Gaussian, we see that $c_t = tc_{\theta\theta}$. We now estimate:

$$\begin{aligned}|c_\theta| &\leq L/2\pi, \quad \text{hence length}(C_t) \leq L\\ |c_{\theta\theta}| &\leq \frac{L}{(2\pi)^{3/2}t}\int_{S^1}\sum_n e^{-(\theta-\varphi-nL)^2/2t^2}|\alpha'|(d\varphi)\\ &\leq \sup_x\big(\sum_n e^{-(x-nL)^2/2t^2}\big)\frac{L\cdot\mathrm{Var}(\arg(c_0'))}{(2\pi)^{3/2}t} = O(1/t),\end{aligned}$$

$$\int_{S^1}|c_{\theta\theta}|d\theta \leq \frac{L}{2\pi}\Big(\int_{\mathbb{R}}\frac{1}{\sqrt{2\pi}t}e^{-\theta^2/2t^2}d\theta\Big)\Big(\int_{S^1}|\alpha'(d\varphi)|\Big) = \frac{L}{2\pi}\mathrm{Var}(\arg(c_0'))$$

$$|c_t| = t|c_{\theta\theta}| = O(1).$$

To finish the proof, all we need to do is get a lower bound on $|c_\theta|$. However, $|c_\theta|$ can be very small if the curve c_0 has corners with small angles. In fact, c_0 can even double back on itself, giving a 'corner' with angle π. We need to treat this as a special case. When all the vertex angles of c_0 are less than π, we can get a lower bound for $|c_\theta|$ as follows. We start with the estimate:

$$\begin{aligned}|c_\theta(\theta)| &= \Big|\frac{1}{\sqrt{2\pi}t}\int_{\mathbb{R}} e^{i\alpha(\theta-\varphi)-\varphi^2/2t^2}d\varphi/e^{i\alpha(\theta)}\Big|\\ &\geq \Big|\frac{1}{\sqrt{2\pi}t}\int_{\mathbb{R}} \cos(\alpha(\theta-\varphi)-\alpha(\theta))e^{-\varphi^2/2t^2}d\varphi\Big|\end{aligned}$$

We break up the integral over $\mathbb{R}$ into 3 intervals $(-\infty,\theta-\delta/2],[\theta-\delta/2,\theta+\delta/2],[\theta+\delta/2,+\infty)$ for a suitable δ. If t is sufficiently small, the integral of the Gaussian over the first and third intervals goes uniformly to 0 and, on the middle interval, goes to 1. Thus it suffices to estimate the cos in the middle interval. We use a remark on BV functions:

Lemma. *For any BV function $f(x)$ and any $C > 0$, there is a $\delta > 0$ such that on every interval I of length less than δ, either $f|_I$ has a single jump of size $\geq C$ or $\max(f|_I) - \min(f|_I) \leq C$.*

In fact, let $C - \varepsilon$ be the size of the largest jump in f less than C and break up the domain of f into intervals J_i on each of which the variation of f is less than $\epsilon/2$, big jumps being on their boundaries. If δ is less than the minimum of the lengths of the J_i, we get what we want.

Now let $\pi - \beta$ be the largest vertex angle of the curve c_0. Using the last lemma, choose a δ so that on every interval I in the θ-line of length less than δ, either I contains a single vertex with exterior angle $\geq \beta/3$ or $\max\alpha|_I - \min\alpha|_I \leq \beta/3$. Now if there is no vertex in $[\theta-\delta/2,\theta+\delta/2]$, then $|\alpha(\theta-\varphi)-\alpha(\theta)| \leq \beta/3$ on this interval and our lower bound is:

$$|c_\theta(\theta)| \geq \cos(\beta/3) - o(t).$$

On the other hand , if there is such a vertex, say at $\bar\theta$, then α varies by at most $\beta/3$ in $[\theta-\delta/2,\bar\theta)$, jumps by at most $\pi-\beta$ at $\bar\theta$ and then varies by at most $\beta/3$ on $(\bar\theta,\theta+\delta/2]$. Assume $\theta < \bar\theta$ (the case $\theta > \bar\theta$ is similar). Then:

$$\cos(\alpha(\theta-\varphi)-\alpha(\theta)) \geq \begin{cases}\cos(\beta/3), & \text{if } \varphi \in (\theta-\bar\theta,\theta+\delta/2]\\ \cos(\pi-\beta+\beta/3) = -\cos(2\beta/3), & \text{if } \varphi\in[\theta-\delta/2,\theta-\bar\theta)\end{cases}$$

Thus:

$$|c_\theta(\theta)| \geq \tfrac{1}{2}(\cos(\beta/3)-\cos(2\beta/3)) - o(t).$$

hence, if t is sufficiently small, we get a uniform lower bound on $|c_\theta|$. Since $|\kappa_{C_t}| \leq |c_{\theta\theta}|/|c_\theta|^2$, we get the required upper bound both on $|\kappa_{C_t}|$ and on the variation of α_{C_t}, i.e. $\int_{S^1} |\kappa_{C_t}|$ and all the requirements of the lemma are satisfied.

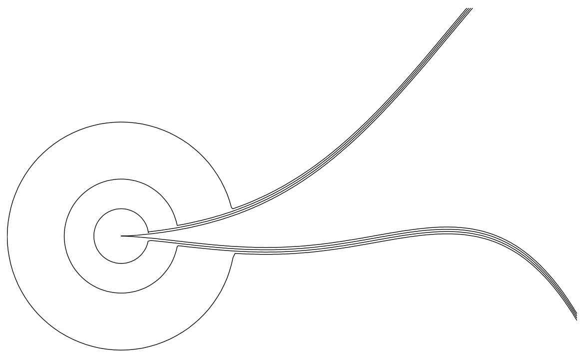


Figure 4. Approximating $1-BV$ curves with zero angle vertices by curves with positive angle vertices.

If c_0 has a vertex with angle π, we need to add an extra argument. c_0 certainly has at most a finite number of such vertices and we can construct a new curve by drawing a circle of radius t around each of these vertices and letting $c_0^{(t)}$ be the curve which follows c_0 until it hits one of these circles and then replaces the vertex with a circuit around the circle: see Figure 3. Each of the curves $c_0^{(t)}$ is in the completion of B_i by the previous argument and the path formed by the $c_0^{(t)}$'s also has finite length, hence c_0 is in the completion. We omit the details which are straightforward.

12.12. The energy of a path as 'anisotropic area' of its graph in $\mathbb{R}^3$. Consider a path $t \mapsto c(t,\cdot)$ in the manifold $\operatorname{Imm}(S^1,\mathbb{R}^2)$. It projects to a path $\pi\circ c$ in $B_i(S^1,\mathbb{R}^2)$ whose energy is

$$\begin{aligned}
E_{G^A}(\pi\circ c) &= \tfrac12 \int_a^b G^A_{\pi(c)}(T_c\pi.c_t, T_c\pi.c_t)\,dt \\
&= \tfrac12 \int_a^b G^A_c(c_t^\perp, c_t^\perp)\,dt = \tfrac12\int_a^b\int_{S^1}(1+A\kappa(c)^2)\langle c_t^\perp, c_t^\perp\rangle |c_\theta|\,d\theta\,dt \\
&= \tfrac12 \int_a^b\int_{S^1}(1+A\kappa(c)^2)\left\langle \frac{\langle c_t, ic_\theta\rangle}{|c_\theta|^2} ic_\theta, \frac{\langle c_t, ic_\theta\rangle}{|c_\theta|^2} ic_\theta\right\rangle |c_\theta|\,d\theta\,dt \\
&= \tfrac12 \int_a^b\int_{S^1}(1+A\kappa(c)^2)\langle c_t, ic_\theta\rangle^2 \frac{d\theta}{|c_\theta|}\,d\theta dt \qquad (1)
\end{aligned}$$

If the path c is horizontal, i.e., it satisfies $\langle c_t, c_\theta\rangle = 0$. Then $\langle c_t, ic_\theta\rangle = |c_t|.|c_\theta|$ and we have

$$E^{\text{hor}}_{G^A}(c) = \tfrac12\int_a^b\int_{S^1}(1+A\kappa(c)^2)|c_t|^2|c_\theta|\,d\theta\,dt, \qquad \langle c_t, c_\theta\rangle = 0 \qquad (2)$$

which is just the usual energy of c.
Let $c(t,\theta) = (x(t,\theta), y(t,\theta))$ be still horizontal and consider the graph

$$\Phi(t,\theta) = (t, x(t,\theta), y(t,\theta)) \in \mathbb{R}^3.$$

We also have $|x_t y_\theta - x_\theta y_t| = |\det(c_t, c_\theta)| = |c_t|.|c_\theta|$ and for the vector product $\Phi_t \times \Phi_\theta = (x_t y_\theta - x_\theta y_t, -y_\theta, x_\theta)$, so we get
$|\Phi_t \times \Phi_\theta|^2 = (x_t y_\theta - x_\theta y_t)^2 + y_\theta^2 + x_\theta^2 = (x_\theta^2 + y_\theta^2)(x_t^2 + y_t^2 + 1) = |c_\theta|^2(|c_t|^2 + 1)$.
We express now $E^{\text{hor}}(c)$ as an integral over the immersed surface $S \subset \mathbb{R}^3$ parameterized by Φ in terms of the surface area $d\mu_S = |\Phi_t \times \Phi_\theta|\, d\theta\, dt$ as follows:

$$\begin{aligned}
E^{\text{hor}}_{G^A}(c) &= \tfrac{1}{2}\int_a^b \int_{S^1} (1 + A\kappa(c)^2) \frac{|c_t|^2 |c_\theta|}{|\Phi_t \times \Phi_\theta|} |\Phi_t \times \Phi_\theta|\, d\theta\, dt \\
&= \tfrac{1}{2}\int_{[a,b]\times S^1} (1 + A\kappa(c)^2) \frac{|c_t|^2}{\sqrt{|c_t|^2 + 1}}\, d\mu_S
\end{aligned}$$

Next we want to express the integrand as as a function γ of the unit normal $n_S = (\Phi_t \times \Phi_\theta)/|\Phi_t \times \Phi_\theta|$. Let $e_0 = (1,0,0)$, then the absolute value of the t-component n_S^0 of the unit normal n_S is

$$|n_S^0| := |\langle e_0, n_S\rangle| = \frac{|c_t|}{\sqrt{|c_t|^2 + 1}}, \qquad \text{and} \qquad \frac{|c_t|^2}{\sqrt{|c_t|^2 + 1}} = \frac{|n_S^0|^2}{\sqrt{1 - |n_S^0|^2}}.$$

Thus for horizontal c (i.e., with $c_t \perp c_\theta$) we have

Horizontal energy as anisotropic area.

$$(3) \qquad \boxed{E^{hor}_{G^A}(c) = \tfrac{1}{2}\int_{[a,b]\times S^1} (1 + A\kappa(c)^2) \frac{|n_S^0|^2}{\sqrt{1 - |n_S^0|^2}}\, d\mu_S}$$

Here the final expression is only in terms of the surface S and does not depend on the curve c being horizontal. This anisotropic area functional has to be minimized in order to prove that geodesics exists between arbitrary curves (of the same degree) in $B_i(S^1, \mathbb{R}^2)$. Thus we are led to the

Question. *For immersions $c_0, c_1 : S^1 \to \mathbb{R}^2$ does there exist an immersed surface $S = (\mathrm{ins}_{[0,1]}, c) : [0,1] \times S^1 \to \mathbb{R} \times \mathbb{R}^2$ such that the functional* (3) *is critical at S?*

A first step is:

Bounding the area. *For any path $[a,b] \ni t \mapsto c(t,\ \ )$ the area of the graph surface $S = S(c)$ is bounded as follows:*

$$(4) \qquad \mathit{Area}(S) = \int_{[a,b]\times S^1} d\mu_S \le 2E^{hor}_{G^A}(c) + \max_t \ell(c(t,\ \ ))(b-a)$$

Proof. Writing the unit normal $n_S = (n_S^0, n_S^1, n_S^2) \in S^2$ according to the coordinates (t, x, y) we have

$$|n_S^1| + |n_S^2| + \frac{|n_S^0|^2}{\sqrt{1 - |n_S^0|^2}} \geq |n_S^1|^2 + |n_S^2|^2 + |n_S^0|^2 = 1$$

Since $|n_S^1| d\mu_S$ is the area element of the projection of S onto the (t, y)-plane we have

$$\begin{aligned}\operatorname{Area}(S) &= \int_{[a,b]\times S^1} d\mu_S \leq \int_{[a,b]\times S^1} (1 + A\kappa(c)^2)\Big(|n_S^1| + |n_S^2| + \frac{|n_S^0|^2}{\sqrt{1-|n_S^0|^2}}\Big) d\mu_S \\ &\leq 2E^{\text{hor}}_{G^A}(c) + \max_t \ell(c(t, \quad))(b - a). \quad \square\end{aligned}$$

13. Geodesic equations and sectional curvatures

13.1. Geodesics on $\operatorname{Imm}(S^1, \mathbb{R}^2)$**.** The energy of a curve $t \mapsto c(t, \cdot)$ in the space $\operatorname{Imm}(S^1, \mathbb{R}^2)$ is

$$E_{G^A}(c) = \tfrac{1}{2} \int_a^b \int_{S^1} (1 + A\kappa_c^2)\langle c_t, c_t\rangle |c_\theta|\, d\theta\, dt.$$

By calculating its first variation, we get the equation for a geodesic:

Geodesic Equation.
(1)

$$\boxed{\Big((1 + A\kappa^2)|c_\theta| \cdot c_t\Big)_t = \Big(\frac{-1 + A\kappa^2}{2} \cdot \frac{|c_t|^2}{|c_\theta|} \cdot c_\theta + A\frac{(\kappa|c_t|^2)_\theta}{|c_\theta|^2} \cdot ic_\theta\Big)_\theta.}$$

Proof. From 11.2 we have

$$\kappa(c)_s = \frac{\langle ic_{s\theta}, c_{\theta\theta}\rangle}{|c_\theta|^3} + \frac{\langle ic_\theta, c_{s\theta\theta}\rangle}{|c_\theta|^3} - 3\kappa\frac{\langle c_{s\theta}, c_\theta\rangle}{|c_\theta|^2}.$$

and

$$\begin{aligned}c_{\theta\theta} &= \frac{\langle c_{\theta\theta}, c_\theta\rangle}{|c_\theta|^2} c_\theta + \frac{\langle c_{\theta\theta}, ic_\theta\rangle}{|c_\theta|^2} ic_\theta \\ &= \frac{|c_\theta|_\theta}{|c_\theta|} c_\theta + \kappa(c)|c_\theta| ic_\theta.\end{aligned}$$

Now we compute

$$\begin{aligned}&\partial_s|_0 E(c) = \tfrac{1}{2}\partial_s|_0 \int_a^b \int_{S^1} (1 + A\kappa^2)\langle c_t, c_t\rangle |c_\theta|\, d\theta\, dt \\ &= \int_a^b \int_{S^1} \Big(A\kappa\kappa_s|c_\theta||c_t|^2 + (1 + A\kappa^2)\langle c_{st}, c_t\rangle |c_\theta| + \frac{1 + A\kappa^2}{2}|c_t|^2 \frac{\langle c_{s\theta}, c_\theta\rangle}{|c_\theta|}\Big)\, d\theta dt\end{aligned}$$

$$
\begin{aligned}
&= \int_a^b \int_{S^1} \Big(A\kappa \langle ic_{s\theta}, c_{\theta\theta}\rangle \frac{|c_t|^2}{|c_\theta|^2} + A\kappa\langle ic_\theta, c_{s\theta\theta}\rangle \frac{|c_t|^2}{|c_\theta|^2} - 3A\kappa^2 \langle c_{s\theta}, c_\theta\rangle \frac{|c_t|^2}{|c_\theta|} \\
&\qquad\qquad - \Big\langle c_s, \big((1+A\kappa^2)|c_\theta| c_t\big)_t + \big(\frac{1+A\kappa^2}{2}\frac{|c_t|^2}{|c_\theta|} c_\theta\big)_\theta \Big\rangle \Big)\, d\theta dt \\
&= \int_a^b \int_{S^1} \Big(\Big\langle c_s, A\Big(\kappa \frac{|c_t|^2}{|c_\theta|^2} ic_{\theta\theta}\Big)_\theta \Big\rangle + \Big\langle c_s, A\Big(\kappa\frac{|c_t|^2}{|c_\theta|^2} ic_\theta\Big)_{\theta\theta}\Big\rangle + \Big\langle c_s, 3A\Big(\kappa^2 \frac{|c_t|^2}{|c_\theta|} c_\theta\Big)_\theta \Big\rangle \\
&\qquad\qquad - \Big\langle c_s, \big((1+A\kappa^2)|c_\theta| c_t\big)_t + \big(\frac{1+A\kappa^2}{2}\frac{|c_t|^2}{|c_\theta|} c_\theta\big)_\theta \Big\rangle \Big)\, d\theta dt \\
&= \int_a^b \int_{S^1} \Big\langle c_s, -\big((1+A\kappa^2)|c_\theta| c_t\big)_t + F_\theta \Big\rangle\, d\theta dt
\end{aligned}
$$

where

$$
\begin{aligned}
F = A\kappa \frac{|c_t|^2}{|c_\theta|^2} ic_{\theta\theta} + A(\kappa |c_t|^2)_\theta \frac{ic_\theta}{|c_\theta|^2} - 2A\kappa |c_t|^2 \frac{|c_\theta|_\theta ic_\theta}{|c_\theta|^3} + A\kappa \frac{|c_t|^2}{|c_\theta|^2} ic_{\theta\theta} + \\
+ 3A\kappa^2 \frac{|c_t|^2}{|c_\theta|} c_\theta - \frac{1+A\kappa^2}{2}\frac{|c_t|^2}{|c_\theta|} c_\theta
\end{aligned}
$$

Substituting the expression for $c_{\theta\theta}$ and simplifying, this reduces to

$$
F = \frac{-1+A\kappa^2}{2}\frac{|c_t|^2}{|c_\theta|} c_\theta + A(\kappa|c_t|^2)_\theta \frac{ic_\theta}{|c_\theta|^2}
$$

which gives the required formula for geodesics.

Putting $A = 0$ in 13.1.1 we get the geodesic equation for the H^0-metric on $\mathrm{Imm}(S^1, \mathbb{R}^2)$

$$
\big(|c_\theta| c_t\big)_t = -\frac{1}{2}\big(\frac{|c_t|^2 c_\theta}{|c_\theta|}\big)_\theta \tag{2}
$$

13.2 Geodesics on $B_i(S^1, \mathbb{R}^2)$. We may also restrict to geodesics which are perpendicular to the orbits of $\mathrm{Diff}(S^1)$, i.e. $\langle c_t, c_\theta\rangle \equiv 0$, obtaining the geodesics in the quotient space $B_i(S^1, \mathbb{R}^2)$. To write this in the simplest way, we introduce the 'velocity' a by setting $c_t = iac_\theta/|c_\theta|$ (so that $|c_t|^2 = a^2$). When we substitute this into the above geodesic equation, the equation splits into a multiple of c_θ and a multiple of ic_θ. The former vanishes identically and the latter gives:

$$
\begin{aligned}
\big((1+A\kappa^2)|c_\theta| a\big)_t \frac{ic_\theta}{|c_\theta|} &= \frac{-1+A\kappa^2}{2} a^2 \big(\frac{c_\theta}{|c_\theta|}\big)_\theta + A\big(\frac{(\kappa a^2)_\theta}{|c_\theta|}\big)_\theta \frac{ic_\theta}{|c_\theta|}, \qquad \text{or} \\
\big((1+A\kappa^2)|c_\theta| a\big)_t &= \frac{-1+A\kappa^2}{2} \kappa |c_\theta| a^2 + A\big(\frac{(\kappa a^2)_\theta}{|c_\theta|}\big)_\theta.
\end{aligned}
$$

If we use derivatives with respect to arclength instead of θ and write these with the subscript s, so that $f_s = f_\theta/|c_\theta|$, this simplifies. We need:

$$|c_\theta|_t = \frac{\langle c_\theta, c_{t\theta}\rangle}{|c_\theta|} = -\frac{\langle c_{\theta\theta}, c_t\rangle}{|c_\theta|} = -a\frac{\langle c_{\theta\theta}, ic_\theta\rangle}{|c_\theta|^2} = -a\kappa|c_\theta|$$

which gives us a simple form for the equation for geodesics on $B_i(S^1,\mathbb{R}^2)$:

$$\big((1+A\kappa^2)a\big)_t = \frac{1+3A\kappa^2}{2}\kappa a^2 + A(\kappa a^2)_{ss}. \tag{1}$$

Finally, we may expand the t-derivatives on the left hand side, using the formula $\kappa_t = a\kappa^2 + a_{ss}$ noted in 11.2.7; we also collect all constraint equations that we chose along the way:

$$\boxed{\begin{aligned} 0 = \langle c_t, c_s\rangle,\quad c_t = aic_s,\quad \kappa = \langle c_{ss}, ic_s\rangle \\ a_t = \frac{\frac12\kappa a^2 + A\big(a^2(\kappa_{ss} - \frac12\kappa^3) + 4\kappa_s a a_s + 2\kappa a_s^2\big)}{1+A\kappa^2}. \end{aligned}} \tag{2}$$

Handle this with care: Going to unit speed parametrization (so that f_s is really a holonomic partial derivative) destroys the first constraint 'horizontality'. This should be seen as a gauge fixing.

13.3. Geodesics on $B_i(S^1,\mathbb{R}^2)$ for $A = 0$. Let us now set $A = 0$. We keep looking at horizontal geodesics, so that $\langle c_t, c_\theta\rangle = 0$ and $c_t = iac_\theta/|c_\theta|$ for $a \in C^\infty(S^1)$. We use the functions a, $s = |c_\theta|$, and κ. We use equations from 13.2 but we do not use the anholonomic derivative:

(1)

$$s_t = -a\kappa s, \qquad a_t = \tfrac12\kappa a^2, \qquad \kappa_t = a\kappa^2 + \frac1s\Big(\frac{a_\theta}{s}\Big)_\theta = a\kappa^2 + \frac{a_{\theta\theta}}{s^2} - \frac{a_\theta s_\theta}{s^3}.$$

We may assume that $s|_{t=0}$ is constant. Let $v(\theta) = a(0,\theta)$ be the initial value for a. Then from equations (1) we get

$$\frac{s_t}{s} = -a\kappa = -2\frac{a_t}{a} \qquad \Longrightarrow \qquad \log(sa^2)_t = 0$$

so that sa^2 is constant in t,

$$s(t,\theta)a(t,\theta)^2 = s(0,\theta)a(0,\theta)^2 = v(\theta)^2, \tag{2}$$

a smooth family of conserved quantities along the geodesic. This leads to the substitutions

$$s = \frac{v^2}{a^2}, \qquad \kappa = 2\frac{a_t}{a^2}$$

which transform the last equation (1) to

$$a_{tt} - 4\frac{a_t^2}{a} - \frac{a^6 a_{\theta\theta}}{2v^4} + \frac{a^6 a_\theta v_\theta}{v^5} - \frac{a^5 a_\theta^2}{v^4} = 0, \quad a(0,\theta) = v(\theta), \tag{3}$$

a nonlinear hyperbolic second order equation. Note that (2) implies that wherever $v = 0$ then also $a = 0$ for all t. For that reason, let us transform equation (3) into a less singular form by substituting $a = vb$. Note that $b = 1/\sqrt{s}$. The outcome is

$$(4) \qquad (b^{-3})_{tt} = -\frac{v^2}{2}(b^3)_{\theta\theta} - 2vv_\theta(b^3)_\theta - \frac{3vv_{\theta\theta}}{2}b^3, \qquad b(0,\theta) = 1.$$

13.4. The induced metric on $B_{i,f}(S^1,\mathbb{R}^2)$ in a chart. We also want to compute the curvature of $B_i(S^1,\mathbb{R}^2)$ in this metric. For this, we need second derivatives and the most convenient way to calculate these seems to be to use a local chart. Consider the smooth principal bundle $\pi : \mathrm{Imm}_f(S^1,\mathbb{R}^2) \to B_{i,f}(S^1,\mathbb{R}^2)$ with structure group $\mathrm{Diff}(S^1)$ described in 11.4.3. We shall describe the metric in the following chart near $C \in B_{i,f}(S^1,\mathbb{R}^2)$: Let $c \in \mathrm{Imm}_f(S^1,\mathbb{R}^2)$ be parametrized by arclength with $\pi(c) = C$ of length L, with unit normal n_c. We assume that the parameter θ runs in the scaled circle S^1_L below. As in the proof of 11.4.3 we consider the mapping

$$\begin{aligned}
&\psi : C^\infty(S^1_L,(-\varepsilon,\varepsilon)) \to \mathrm{Imm}_f(S^1_L,\mathbb{R}^2), \qquad \mathcal{Q}(c) := \psi(C^\infty(S^1_L,(-\varepsilon,\varepsilon)))\\
&\psi(f)(\theta) = c(\theta) + f(\theta)n_c(\theta) = c(\theta) + f(\theta)ic'(\theta),\\
&\pi\circ\psi : C^\infty(S^1_L,(-\varepsilon,\varepsilon)) \to B_{i,f}(S^1,\mathbb{R}^2),
\end{aligned}$$

where ε is so small that $\psi(f)$ is an embedding for each f. By 11.4.3 the mapping $(\pi\circ\psi)^{-1}$ is a smooth chart on $B_{i,f}(S^1_L,\mathbb{R}^2)$. Note that:

$$\begin{aligned}
\psi(f)' &= c' + f'ic' + fic'' = (1 - f\kappa_c)c' + f'ic'\\
\psi(f)'' &= c'' + f''ic' + 2f'ic'' + fic''' = -(2f'\kappa_c + f\kappa_c')c' + (\kappa_c + f'' - f\kappa_c^2)ic'\\
n_{\psi(f)} &= \frac{1}{\sqrt{(1-f\kappa_c)^2 + f'^2}}\Big((1-f\kappa_c)ic' - f'c'\Big),\\
T_f\psi.h &= h.ic' \quad \in C^\infty(S^1,\mathbb{R}^2) = T_{\psi(f)}\mathrm{Imm}_f(S^1_L,\mathbb{R}^2)\\
&= \frac{h(1-f\kappa_c)}{\sqrt{(1-f\kappa_c)^2+f'^2}}n_{\psi(f)} + \frac{hf'}{(1-f\kappa_c)^2+f'^2}\psi(f)',\\
(T_f\psi.h)^\perp &= \frac{h(1-f\kappa_c)}{\sqrt{(1-f\kappa_c)^2+f'^2}}n_{\psi(f)} \qquad \in \mathcal{N}_{\psi(f)},\\
\kappa_{\psi(f)} &= \frac{1}{((1-f\kappa_c)^2+f'^2)^{3/2}}\langle i\psi(f)',\psi(f)''\rangle\\
&= \frac{\kappa_c + f'' - 2f\kappa_c^2 - ff''\kappa_c + f^2\kappa_c^3 + 2f'^2\kappa_c + ff'\kappa_c'}{((1-f\kappa_c)^2+f'^2)^{3/2}}
\end{aligned}$$

Let G^A denote also the induced metric on $B_{i,f}(S^1_L,\mathbb{R}^2)$. Since π is a Riemannian submersion, $T_{\psi(f)}\pi : (\mathcal{N}_{\psi(f)}, G^A_{\psi(f)}) \to (B_{i,f}(S^1_L,\mathbb{R}^2), G^A_{\pi(\psi(f))})$ is an isometry. Then we compute for $f \in C^\infty(S^1_L, (-\varepsilon, \varepsilon))$ and $h, k \in C^\infty(S^1_L, \mathbb{R})$

$$((\pi \circ \psi)^* G^A)_f(h,k) = G^A_{\pi(\psi(f))}\Big(T_f(\pi \circ \psi)h, T_f(\pi \circ \psi)k\Big)$$
$$= G^A_{\psi(f)}\Big((T_f\psi.h)^\perp, (T_f\psi.k)^\perp\Big)$$
$$= \int_{S^1_L} (1 + A\kappa^2_{\psi(f)})\Big\langle (T_f\psi.h)^\perp, (T_f\psi.k)^\perp \Big\rangle |\psi(f)'|\, d\theta$$
$$= \int_{S^1_L} (1 + A\kappa^2_{\psi(f)}) \frac{hk(1 - f\kappa_c)^2}{\sqrt{(1 - f\kappa_c)^2 + f'^2}}\, d\theta$$

This is the expression from which we have to compute the geodesic equation in the chart on $B_{i,f}(S^1_L, \mathbb{R}^2)$.

13.5. Computing the Christoffel symbols in $B_{i,f}(S^1_L, \mathbb{R}^2)$ at $C = \pi(c)$. We have to compute second derivatives in f of the expression of the metric in 13.2. For that we expand the two main contributing expressions in f to order 2, where we put $\kappa = \kappa_c$.

$$\kappa_{\psi(f)} =$$
$$= (1 - 2f\kappa + f^2\kappa^2 + f'^2)^{-3/2}(\kappa + f'' - 2f\kappa^2 - ff''\kappa + f^2\kappa^3 + 2f'^2\kappa + ff'\kappa')$$
$$= \kappa + (f'' + f\kappa^2) + (f^2\kappa^3 + \tfrac{1}{2}f'^2\kappa + ff'\kappa' + 2ff''\kappa) + O(f^3)$$
$$(1 - f\kappa)^2(1 - 2f\kappa + f^2\kappa^2 + f'^2)^{-1/2} = 1 - f\kappa - \tfrac{1}{2}f'^2 + O(f^3)$$

Thus

$$(1 + A\kappa^2_{\psi(f)}) \frac{(1 - f\kappa_c)^2}{\sqrt{(1 - f\kappa_c)^2 + f'^2}} = 1 + A\kappa^2 + 2Af''\kappa + Af\kappa^3 - f\kappa -$$
$$- \tfrac{1}{2}f'^2 + Af^2\kappa^4 + A\tfrac{1}{2}f'^2\kappa^2 + 2Aff'\kappa\kappa' + Af''^2 + 4Aff''\kappa^2$$

and finally

(1) $$G^A_f(h,k) = ((\pi \circ \psi)^* G^A)_f(h,k) =$$
$$= \int_{S^1_L} hk\Big((1 + A\kappa^2) + (2Af''\kappa + Af\kappa^3 - f\kappa) + -\tfrac{1}{2}f'^2$$
$$+ A(4ff''\kappa^2 + f^2\kappa^4 + \tfrac{1}{2}f'^2\kappa^2 + 2ff'\kappa\kappa' + f''^2) + O(f^3)\Big)\, d\theta.$$

We differentiate the metric

$$dG^A(f)(l)(h,k) = \int_{S^1_L} hk\Big(2Al''\kappa + (A\kappa^3 - \kappa)l + 4Alf''\kappa^2 + 4Afl''\kappa^2 +$$
$$+ 2Afl\kappa^4 + (A\kappa^2 - 1)f'l' + 2Alf'\kappa\kappa' + 2Afl'\kappa\kappa' + 2Af''l'' + O(f^2)\Big)\, d\theta$$

and compute the Christoffel symbol

$$
\begin{aligned}
&- 2G^A_f(\Gamma_f(h,k),l) = -dG^A(f)(l)(h,k) + dG^A(f)(h)(k,l) + dG^A(f)(k)(l,h)\\
&= \int_{S^1_L} l\Big((A\kappa^3 - \kappa + 2A\kappa\kappa' f' + 4A\kappa^2 f'' + 2A\kappa^4 f)kh\\
&\qquad\qquad + (2A\kappa + 4A\kappa^2 f + 2Af'')(h''k + hk'')\\
&\qquad\qquad + (A\kappa^2 f' - f' + 2A\kappa\kappa' f)(h'k + hk') + O(f^2)\Big)\, d\theta\\
&\quad - \int_{S^1_L} \Big(l'(A\kappa^2 f' hk - f' hk + 2A\kappa\kappa' f hk)\\
&\qquad\qquad + l''(2A\kappa hk + 4A\kappa^2 f hk + 2Af'' hk) + O(f^2)\Big)\, d\theta\\
&= \int_{S^1_L} l\Big((A\kappa^3 - \kappa - 2A\kappa'')hk - 4A\kappa'(h'k + hk') - 4A\kappa h'k'\\
&\qquad\qquad + (-2Af^{(4)} - f'' + 2A\kappa^4 f - 6A\kappa'^2 f - 6A\kappa\kappa'' f - 10A\kappa\kappa' f' + A\kappa^2 f'')hk\\
&\qquad\qquad - (2f' + 4Af''' + 12A\kappa\kappa' f + 6A\kappa^2 f')(h'k + hk')\\
&\qquad\qquad - 2(4A\kappa^2 f + 2Af'')h'k' + O(f^2)\Big)\, d\theta.
\end{aligned}
$$

Thus

$$
\begin{aligned}
&G^A_f(\Gamma_f(h,k),l) =\\
&= \int_{S^1_L} l\Big((\tfrac{1}{2}\kappa - \tfrac{1}{2}A\kappa^3 + A\kappa'')hk + 2A\kappa'(h'k + hk') + 2A\kappa h'k'\\
&\qquad\qquad - (-Af^{(4)} - \tfrac{1}{2}f'' + A\kappa^4 f - 3A\kappa'^2 f - 3A\kappa\kappa'' f - 5A\kappa\kappa' f' + \tfrac{1}{2}A\kappa^2 f'')hk\\
&\qquad\qquad + (f' + 2Af''' + 6A\kappa\kappa' f + 3A\kappa^2 f')(h'k + hk')\\
&\qquad\qquad + (4A\kappa^2 f + 2Af'')h'k' + O(f^2)\Big)\, d\theta.
\end{aligned}
$$

At the center of the chart, for $f = 0$, we get

$$
\begin{aligned}
&G^A_0(\Gamma_0(h,k),l) =\\
&= \int_{S^1_L} l\Big((\tfrac{1}{2}\kappa - \tfrac{1}{2}A\kappa^3 + A\kappa'')hk + 2A\kappa'(h'k + hk') + 2A\kappa h'k'\Big)\, d\theta\\
&= \int_{S^1_L} l\Big(\frac{(\frac{1}{2}\kappa - \frac{1}{2}A\kappa^3 + A\kappa'')hk + 2A\kappa'(h'k + hk') + 2A\kappa h'k'}{1 + A\kappa^2}\Big)(1 + A\kappa^2)\, d\theta\\
&= G^A_0\Big(\frac{(\frac{1}{2}\kappa - \frac{1}{2}A\kappa^3 + A\kappa'')hk + 2A\kappa'(h'k + hk') + 2A\kappa h'k'}{1 + A\kappa^2}, l\Big)
\end{aligned}
$$

so that

$$
(2) \qquad \Gamma_0(h,k) = \frac{(\frac{1}{2}\kappa - \frac{1}{2}A\kappa^3 + A\kappa'')hk + 2A\kappa'(h'k + hk') + 2A\kappa h'k'}{1 + A\kappa^2}.
$$

Letting $h = k = f_t$, this leads to the geodesic equation, valid at $f = 0$:

$$f_{tt} = \frac{(\frac{1}{2}\kappa - \frac{1}{2}A\kappa^3 + A\kappa'')f_t^2 + 4A\kappa' f_t f_t' + 2A\kappa(f_t')^2}{1 + A\kappa^2}.$$

If we substitute a for f_t and a_t for f_{tt}, this is the same as the previous geodesic equation derived in 13.2 by variational methods. There is a subtle point here, however: why is it ok to identify the second derivatives a_t and f_{tt} with each other? To check this let $c(\theta) + (ta_1(\theta) + \frac{t^2}{2}a_2(\theta))ic'(\theta)$ be a 2-jet in our chart. Then if we reparametrize the nearby curves by substituting $\theta - \frac{t^2}{2}a_1a_1'$ for θ, letting

$$\begin{aligned} c(t,\theta) &= c(\theta - \tfrac{t^2}{2}a_1a_1') + (ta_1(\theta - \tfrac{t^2}{2}a_1a_1') + \frac{t^2}{2}a_2(\theta - \tfrac{t^2}{2}a_1a_1'))ic(\theta - \tfrac{t^2}{2}a_1a_1')' \\ &\equiv c(\theta) - (\tfrac{t^2}{2}a_1a_1')c'(\theta) + (ta_1(\theta) + \frac{t^2}{2}a_2(\theta))ic'(\theta) \mod t^3 \end{aligned}$$

then $\langle c', c_t\rangle \equiv 0 \mod t^2$, hence this 2-jet is horizontal and $\langle c_{tt}, ic'\rangle \equiv a_2 \mod t$ as required.

13.6. Computation of the sectional curvature in $B_{i,f}(S^1_L, \mathbb{R}^2)$ at C. We now go further. We use the following formula which is valid in a chart:

(1)

$$\begin{aligned} &2R_f(m,h,m,h) = 2G^A_f(R_f(m,h)m,h) = \\ &= -2d^2G^A(f)(m,h)(h,m) + d^2G^A(f)(m,m)(h,h) + d^2G^A(f)(h,h)(m,m) \\ &\quad - 2G^A(\Gamma(h,m),\Gamma(m,h)) + 2G^A(\Gamma(m,m),\Gamma(h,h)) \end{aligned}$$

The sectional curvature at the two-dimensional subspace $P_f(m,h)$ of the tangent space which is spanned by m and h is then given by:

$$k_f(P(m,h)) = -\frac{G^A_f(R(m,h)m,h)}{\|m\|^2\|h\|^2 - G^A_f(m,h)^2}. \tag{2}$$

We compute this directly for $f = 0$. From the expansion up to order 2 of $G^A_f(h,k)$ in 13.5.1 we get:

$$\begin{aligned} &\frac{1}{2!}d^2G^A(0)(m,l)(h,k) = \int_{S^1_L} hk\Big(-\tfrac{1}{2}m'l' + \\ &\quad + A\Big(2(ml'' + m''l)\kappa^2 + ml\kappa^4 + \tfrac{1}{2}m'l'\kappa^2 + (ml' + m'l)\kappa\kappa' + m''l''\Big)\Big)\,d\theta \end{aligned} \tag{3}$$

Thus we have:

$$\begin{aligned} &-d^2G^A(0)(m,h)(h,m) + \tfrac{1}{2}d^2G^A(0)(m,m)(h,h) + \tfrac{1}{2}d^2G^A(0)(h,h)(m,m) = \\ &= -2\int_{S^1_L} hm\Big(-\tfrac{1}{2}m'h' + \end{aligned}$$

$$+ A\Big(2(mh'' + m''h)\kappa^2 + mh\kappa^4 + \tfrac{1}{2}m'h'\kappa^2 + (mh' + m'h)\kappa\kappa' + m''h''\Big)\Big)\, d\theta$$

$$+ \int_{S^1_L} hh \left(-\tfrac{1}{2}m'^2 + A\Big(4mm''\kappa^2 + m^2\kappa^4 + \tfrac{1}{2}m'^2\kappa^2 + 2mm'\kappa\kappa' + m''^2\Big)\right) d\theta$$

$$+ \int_{S^1_L} mm \left(-\tfrac{1}{2}h'h' + A\Big(4hh''\kappa^2 + hh\kappa^4 + \tfrac{1}{2}h'h'\kappa^2 + 2hh'\kappa\kappa' + h''h''\Big)\right) d\theta$$

$$= \int_{S^1_L} \Big(\tfrac{1}{2}(A\kappa^2 - 1)(mh' - m'h)^2 + A(mh'' - m''h)^2\Big)\, d\theta.$$

For the second part of the curvature we have

$$- G_0(\Gamma_0(h,m), \Gamma_0(m,h)) + G_0(\Gamma_0(m,m), \Gamma_0(h,h)) =$$

$$= \int_{S^1_L} -\Big((\tfrac{1}{2}\kappa - \tfrac{1}{2}A\kappa^3 + A\kappa'')hm + 2A\kappa'(h'm + m'h) + 2A\kappa h'm'\Big)^2 \frac{d\theta}{1 + A\kappa^2}$$

$$+ \int_{S^1_L} \Big((\tfrac{1}{2}\kappa - \tfrac{1}{2}A\kappa^3 + A\kappa'')m^2 + 4A\kappa' mm' + 2A\kappa m'^2\Big)$$

$$\Big((\tfrac{1}{2}\kappa - \tfrac{1}{2}A\kappa^3 + A\kappa'')h^2 + 4A\kappa' hh' + 2A\kappa h'^2\Big)\frac{d\theta}{1 + A\kappa^2}$$

$$= \int_{S^1_L} \Big((A\kappa^2 - A^2\kappa^4 + 2A^2\kappa\kappa'' - 4A^2\kappa'^2)(mh' - m'h)^2\Big)\frac{d\theta}{1 + A\kappa^2}$$

Thus we get

$$R_0(m,h,m,h) = G_0^A(R_0(m,h)m, h) =$$

$$= \int_{S^1_L} \Big(\tfrac{1}{2}(A\kappa^2 - 1)(mh' - m'h)^2 + A(mh'' - m''h)^2\Big)\, d\theta$$

$$+ \int_{S^1_L} \Big((A\kappa^2 - A^2\kappa^4 + 2A^2\kappa\kappa'' - 4A^2\kappa'^2)(mh' - m'h)^2\Big)\frac{d\theta}{1 + A\kappa^2}$$

Letting $W = mh' - hm'$ be the Wronskian of m and h and simplifying, we have:
(4)

$$\boxed{\begin{aligned} &R_0(m,h,m,h) = \\ &\quad = \int_{S^1_L} \Big(\frac{-(A\kappa^2 - 1)^2 + 4A^2\kappa\kappa'' - 8A^2\kappa'^2}{2(1 + A\kappa^2)}\Big) W^2 d\theta + \int_{S^1_L} AW'^2 d\theta \end{aligned}}$$

What does this formula say? First of all, if $\mathrm{supp}(m) \cap \mathrm{supp}(h) = \emptyset$, the sectional curvature in the plane spanned by m and h is 0. Secondly, we can divide the curve c into two parts:

$$c_A^+ = \text{ set of points where } \kappa\kappa'' < 2(\kappa')^2 + \big(\tfrac{A^{-1} - \kappa^2}{2}\big)^2$$

$$c_A^- = \text{ set of points where } \kappa\kappa'' > 2(\kappa')^2 + \big(\tfrac{A^{-1} - \kappa^2}{2}\big)^2.$$

Note that if A is sufficiently small, $c_A^- = \emptyset$ and even if A is large, c_A^- need not be non-empty. But if $\operatorname{supp}(m), \operatorname{supp}(h) \subset c_A^-$, the sectional curvature is always negative. The interesting case is when $\operatorname{supp}(m), \operatorname{supp}(h) \subset c_A^+$. We may introduce the self-adjoint differential operator on $L^2(S^1)$:

$$Sf = f'' + \frac{(A\kappa^2 - 1)^2 - 4A^2\kappa\kappa'' + 8A^2\kappa'^2}{2A(1 + A\kappa^2)} f$$

so that $R = -A\langle SW, W\rangle$. The eigenvalues of S tend to $-\infty$, hence S has a finite number of positive eigenvalues. If we take, for example, $m = 1$ and h such that h' is in the span of the positive eigenvalues, the corresponding sectional curvature will be positive. In general, the condition that the sectional curvature be positive is that the Wronskian W have a sufficiently large component in the positive eigenspace of S. The special case where c is the unit circle may clarify the picture: then

$$Sf = f'' + \frac{(A - 1)^2}{2A(1 + A)} f$$

and the eigenfunctions are linear combinations of sine's and cosine's. It is easy to see that for any A, a plane spanned by m and h of pure frequencies k and l will have positive curvature if and only if k and l are sufficiently near each other (asymptotically $|k - l| < |A - 1|/\sqrt{A + a^2}$), hence 'beat' at a low frequency.

13.7. The sectional curvature for the induced H^0-metric on $B_{i,f}(S_L^1, \mathbb{R}^2)$ in a chart. In the setting of 13.2 we have for $f \in C^\infty(S_L^1, (-\varepsilon, \varepsilon))$ and $h, k \in C^\infty(S_L^1, \mathbb{R})$

$$\begin{aligned}
(1) \qquad G_f^0(h, k) &= ((\pi \circ \psi)^* G^0)_f(h, k) = G^0_{\pi(\psi(f))}\Big(T_f(\pi \circ \psi)h, T_f(\pi \circ \psi)k\Big) \\
&= G^0_{\psi(f)}\Big((T_f\psi.h)^\perp, (T_f\psi.k)^\perp\Big) \\
&= \int_{S_L^1} \frac{hk(1 - f\kappa_c)^2}{\sqrt{(1 - f\kappa_c)^2 + f'^2}}\, d\theta
\end{aligned}$$

At the center of the chart described in 13.4, i.e., for $f = 0$, the Christoffel symbol 13.5.2 for $A = 0$ becomes

$$(2) \qquad \Gamma_0(h, k) = \tfrac{1}{2}\kappa_c hk$$

The curvature 13.6.4 at $f = 0$ for $A = 0$ becomes

$$\begin{aligned}
&R_0(m, h, m, h) = G_0(R_0(m, h)m, h) = \\
(3) \qquad &= -\tfrac{1}{2}\int_{S_L^1} (h'm - hm')^2\, d\theta = -\tfrac{1}{2}\int_{S_L^1} W(m, h)^2\, d\theta
\end{aligned}$$

and the sectional curvature $k_0(P(m,h))$ from 13.5.2 for $A=0$ and $f=0$ is non-negative.

In the full chart 13.2, starting from the metric 13.6.1, we managed to compute the full geodesic equation not just for $f=0$ but for general f, so long as $A=0$. The outcome is

$$\begin{aligned}\Gamma_f(h,h) &= \frac{\kappa_c h^2}{1-f\kappa_c} + \frac{-\frac{1}{2}\kappa_c(1-f\kappa_c)h^2 + (\frac{1}{2}h^2 f'' + 2hh'f')}{\big((1-f\kappa_c)^2+f'^2\big)} \\ &\qquad - \frac{\kappa_c h^2 f'^2}{(1-f\kappa_c)\big((1-f\kappa_c)^2+f'^2\big)} + \frac{\frac{3}{2}\kappa_c(1-f\kappa_c)h^2f'^2 - \frac{3}{2}h^2f'^2f''}{\big((1-f\kappa_c)^2+f'^2\big)^2}.\end{aligned} \tag{4}$$

The geodesic equation is thus

$$\begin{aligned}f_{tt} &= -\frac{\kappa_c f_t^2}{1-f\kappa_c} - \frac{-\frac{1}{2}\kappa_c(1-f\kappa_c)f_t^2 + (\frac{1}{2}f_t^2 f_{\theta\theta} + 2f_t f_{t\theta} f_\theta)}{\big((1-f\kappa_c)^2+f_\theta{}^2\big)} \\ &\qquad + \frac{\kappa_c f_t^2 f_t h^2}{(1-f\kappa_c)\big((1-f\kappa_c)^2+f_\theta{}^2\big)} - \frac{\frac{3}{2}\kappa_c(1-f\kappa_c)f_t^2f_\theta{}^2 - \frac{3}{2}f_t^2f_\theta{}^2f_{\theta\theta}}{\big((1-f\kappa_c)^2+f_\theta{}^2\big)^2}.\end{aligned} \tag{5}$$

For $A>0$ we were unable to get the analogous result.

14. Examples and numerical results

14.1. The geodesics running through concentric circles. The simplest possible geodesic in B_i is given by the set of all circles with common center. Let C_r be the circle of radius r with center the origin. Consider the path of such circles $C_{r(t)}$ given by the parametrization $c(t,\theta)=r(t)e^{i\theta}$, where $r(t)$ is a smooth increasing function $r:[0,1]\to\mathbb{R}_{>0}$. Then $\kappa_c(t,\theta)=1/r(t)$. If we vary r then the horizontal energy and the variation of this curve are

$$\begin{aligned}E^{\text{hor}}_{G^A}(c) &= \frac{1}{2}\int_0^1\int_{S^1}\big(1+A/r^2\big)r_t^2 r\,d\theta\,dt \\ \partial_s|_{s=0}E^{\text{hor}}_{G^A}(c) &= \int_0^1\int_{S^1}\Big(1+\frac{A}{r^2}\Big)r_s\Big(-r_{tt}-\frac{(1-A/r^2)}{2(r+A/r)}r_t^2\Big)r\,d\theta\,dt\end{aligned}$$

so that c is a geodesic if and only if

$$r_{tt} + \frac{(1-A/r^2)}{2(r+A/r)}r_t^2 = 0. \tag{1}$$

Also the geodesic equation 13.1.1 reduces to (1) for c of this form.

The solution of (1) can be written in terms of the inverse of a complete elliptic integral of the second kind. More important is to look at what happens for small and large r. As $r\to 0$, the ODE reduces to:

$$r_{tt} - \frac{r_t^2}{2r} = 0$$

whose general solution is $r(t) = C(t-t_0)^2$ for some contants C, t_0. In other words, at one end, the path ends in finite time with the circles imploding at their common center. Note that $r' \to 0$ as $r \to 0$ but not fast enough to prevent the collapse. On the other hand, as $r \to \infty$, the ODE becomes:

$$r_{tt} + \frac{r_t^2}{2r} = 0$$

whose general solution is $r(t) = C(t-t_0)^{2/3}$ for some constants C, t_0. Thus at the other end of the geodesic, the circles expand forever but with decreasing speed.

An interesting point is that this geodesic has conjugate points on it, so that it is a extremal path but not a local minimum for length over all intervals. This is a concrete reflection of the collapse of the metric when $A = 0$. To work this out, take any $f(\theta)$ such that $\int_0^{2\pi} f d\theta = 0$ and any function $a(t)$. Then $X = f(\theta)a(t)\partial/\partial r$ is a vector field along the geodesic, i.e. a family of tangent vectors to B_e at each circle $C_{r(t)}$ normal to the tangent to the geodesic. Its length is easily seen to be:

$$\|X\|^2_{C_{r(t)}} = \Big(r(t) + \frac{A}{r(t)}\Big)a(t)^2 \int_0^{2\pi} f(\theta)^2 d\theta.$$

We need to work out its covariant derivative:

$$\nabla_{\frac{d}{dt}}(X) = f(\theta)a_t\frac{\partial}{\partial r} + \Gamma_{C_r}\big(r_t\frac{\partial}{\partial r}, f(\theta)a\frac{\partial}{\partial r}\big).$$

Using a formula for the Christoffel symbol which we get from 13.2.2 by polarizing, and noting that $\kappa \equiv 1/r, \kappa_s \equiv 0$, we get:

$$\begin{aligned}\nabla_{\frac{d}{dt}}(X) &= f(\theta)a_t\frac{\partial}{\partial r} + f(\theta)ar_t\Big(\frac{1-A/r^2}{2(r+A/r)}\Big)\frac{\partial}{\partial r}\\ &= f(\theta)\big(r+A/r\big)^{-1/2}\Big(\big(r+A/r\big)^{1/2}a\Big)_t\frac{\partial}{\partial r}.\end{aligned}$$

(This formula also follows from the fact that the vectors $(r+A/r)^{-1/2}\partial/\partial r$ have length independent of t, hence covariant derivative zero.) Jacobi's equation is therefore:

$$(2) \qquad f(\theta)\big(r+A/r\big)^{-1/2}\Big(\big(r+A/r\big)^{1/2}a\Big)_{tt}\frac{\partial}{\partial r} + R\big(X, r_t\frac{\partial}{\partial r}\big)\big(r_t\frac{\partial}{\partial r}\big) = 0,$$

where R is the curvature tensor. For later purposes, it is convenient to write this eqation using r as the independent variable along the geodesic rather than t and think of a as a function of r. Note that for any function b along the geodesic, $b_t = b_r r_t$ and

$$b_{tt} = b_{rr}r_t^2 + b_r r_{tt} = \big(b_{rr} - \frac{(1-A/r^2)}{2(r+A/r)}b_r\big)r_t^2.$$

Then a somewhat lengthy bit of algebra shows that:

$$\left(r+A/r\right)^{-\frac{1}{2}}\left(\left(r+A/r\right)^{\frac{1}{2}}a\right)_{tt}=\left(r+A/r\right)^{-\frac{1}{4}}\left(\left(r+A/r\right)^{\frac{1}{4}}a\right)_{rr}r_t^2+F(r)ar_t^2,$$

$$\text{where}\qquad F(r)=-\frac{5}{16}\Big(\frac{1-A/r^2}{r+A/r}\Big)^2+\frac{A}{2r^3(r+A/r)}.$$

To work out the structure of R in this case, use the fact that the circles C_r and the vector field $\partial/\partial r$ are invariant under rotations. This means that the map $f\mapsto R(\partial/\partial r, f\partial/\partial r)(\partial/\partial r)$ has the two properties: it commutes with rotations and it is symmetric. The only such maps are diagonal in the Fourier basis, i.e. there are real constants λ_n such that

$$R\Big(\partial/\partial r,\begin{cases}\cos(n\theta)\partial/\partial r\\ \sin(n\theta)\partial/\partial r\end{cases}\Big)(\partial/\partial r)=\lambda_n\begin{cases}\cos(n\theta)\partial/\partial r\\ \sin(n\theta)\partial/\partial r\end{cases}.$$

To evaluate λ_n, we take the inner product with $\cos(n\theta)$ (or $\sin(n\theta)$ and use our calculation of $R_0(m,h,m,h)$ in section 13.6 to show:

$$\begin{aligned}\Big\langle R\big(\frac{\partial}{\partial r},\cos(n\theta)\frac{\partial}{\partial r}\big)\big(\frac{\partial}{\partial r}\big),\cos(n\theta)\frac{\partial}{\partial r}\Big\rangle&=R_0\Big(\frac{\partial}{\partial r},\cos(n\theta)\frac{\partial}{\partial r},\frac{\partial}{\partial r},\cos(n\theta)\frac{\partial}{\partial r}\Big)\\&=\int_0^{2\pi}\Big(-\frac{(1-A/r^2)^2}{2(1+A/r^2)}W^2+AW'^2\Big)rd\theta\end{aligned}$$

where

$$W=1.\frac{d}{ds}\cos(n\theta)=-n\frac{\sin(n\theta)}{r}\qquad\text{and}\qquad W'=\frac{d}{ds}W=-n^2\frac{\cos(n\theta)}{r^2}.$$

Simplifying, this gives:

$$\begin{aligned}\lambda_n\|\cos(n\theta)\frac{\partial}{\partial r}\|^2&=\int_0^{2\pi}\Big(-\frac{(1-A/r^2)^2}{2(r+A/r)}n^2\sin^2(n\theta)+\frac{A}{r^3}n^4\cos^2(n\theta)\Big)d\theta\\&=-\frac{(1-A/r^2)^2}{2(r+A/r)}n^2\pi+\frac{A}{r^3}n^4\pi\end{aligned}$$

hence

$$\lambda_n=-\frac{(1-A/r^2)^2}{2(r+A/r)^2}\cdot n^2+\frac{A}{r^3(r+A/r)}\cdot n^4.$$

Thus for $X=\cos(n\theta)a_n(t)\partial/\partial r$, if we combine everything, Jacobi's equation reads:

$$\begin{aligned}&\left(r+A/r\right)^{-\frac{1}{4}}\left(\left(r+A/r\right)^{\frac{1}{4}}a_n\right)_{rr}=\\&\qquad=\Big(-\frac{(1-A/r^2)^2}{2(r+A/r)^2}(n^2-\tfrac{5}{8})+\frac{A}{r^3(r+A/r)}(n^4-\tfrac{1}{2})\Big)a_n.\end{aligned}\tag{3}$$

Calling the right hand side the *potential* of Jacobi's equation, we can check that for each n, the potential is positive for small r, negative for large r and it has one zero, approximately at $\sqrt{2An}$ for large n. Thus, for small r, these

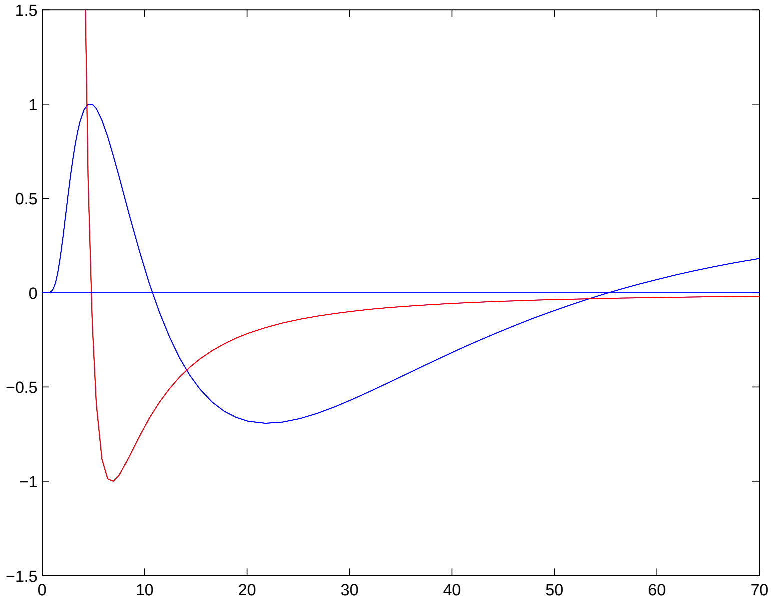


Figure 5. The potential in the Jacobi ODE and its solution for an infinitesimal triangular perturbation of the circles in the geodesic of concentric circles. Note the first conjugate point at $10.77\sqrt{A}$.

perturbations diverge from the geodesic of circles. For large r, if we write $b_n = \big(r + A/r\big)^{1/4} a_n$, then Jacobi's equation approaches:

$$(b_n)_{rr} \approx -\frac{n^2 - 0.625}{2r^2} b_n.$$

This is solved by $b_n = cx^\lambda + c' x^{\lambda'}$ where λ, λ' are solutions of $\lambda^2 - \lambda = -(n^2 - 0.625)/2$. For $n = 1$, λ, λ' are real and b_n has no zeros; but for $n > 1$, λ, λ' have an imaginary part, say $i\gamma_n$, and

$$b_n \approx \sqrt{r}\big(c\cos(\gamma_n \log(r)) + c' \sin(\gamma \log(r))\big)$$

with infinitely many zeros.

Figure 4 shows the solution for $n = 3$ which approaches 0 as $r \to 0$. The first zero of this solution is about $10.77\sqrt{A}$, making it a conjugate point of $r = 0$. For other n, the first such conjugate point appears to be bigger, so we conclude: on any segment $0 < r_1 < r_2 < 10.77\sqrt{A}$, the geodesic of circles is locally (and presumably globally) minimizing.

14.2. The geodesic connecting two distant curves. For any two distant curves C_1, C_2, one can construct paths from one to the other by (a) changing C_1 to some auxiliary curve D near C_1, (b) translating D without modifying it to a point near C_2 and (c) changing the translated curve D to

C_2. If C_1 and C_2 are very far from each other, the energy of the translation will dominate the energy required to modify them both to D. Thus we expect that a geodesic between distant curves will aymptotically utilize a curve D which is optimized for least energy translation. To find such curves D, heuristically we may argue that it should be a curve such that the path given by all its translates in a fixed direction is a geodesic.

Such geodesics can be found as special cases of the general geodesic. We fix $e = (1,0)$ as the direction of translation and assume that the path $\{D + te\}$ is a geodesic. We need to express this geodesic up to order $O(t^2)$ in the chart used in section 13.4. Let $c(s)$ be arc length parametrization of D and $\theta(s)$ be the orientation of D at point $c(s)$, i.e. $c_s = \cos(\theta) + i\sin(\theta)$. Then a little calculation shows that if we reparametrize nearby curves via $\tilde{s} = s - \langle e, c_s\rangle t$, then the path of translates in direction e is just:

$$\begin{aligned} c(\tilde{s}) + te &= c(s) + \big(t\langle e, ic_s\rangle + \frac{t^2}{2}\langle e, c_s\rangle^2\kappa + O(t^3)\big)ic_s \\ &= c(s) + \big(-\sin(\theta(s))t + \frac{t^2}{2}\cos^2(\theta(s)\kappa) + O(t^3)\big)ic_s. \end{aligned}$$

Thus, in the notation of 13.2, $a = -\sin(\theta)$, hence $a_s = -\cos(\theta)\kappa$ and, moreover, $a_t = \cos^2(\theta)\kappa$. Substituting this in the geodesic formula 13.2.1, we get

$$\begin{aligned} &(1 + A\kappa^2)\cos^2(\theta)\kappa = \\ &= \frac{\kappa\sin^2(\theta)}{2} + A\big((\kappa_{ss} - \frac{\kappa^3}{2})\sin^2(\theta) + 4\cos(\theta)\sin(\theta)\kappa\kappa_s + 2\kappa^3\cos^2(\theta)\big). \end{aligned}$$

Since $\kappa = \theta_s$, this becomes, after some manipulation, a singular third order equation for $\theta(s)$:

$$\theta_{sss} = 4\cot(\theta)\theta_s\theta_{ss} + (\tfrac{1}{2} - \cot^2(\theta))\theta_s(\theta_s^2 - \tfrac{1}{A}).$$

One solution of this equation is $\theta(s) \equiv \frac{1}{\sqrt{A}}$, i.e. a circle of radius $\sqrt{A}$. In fact, this seems to be the only simple closed curve which solves this equation. However, if we drop smoothness, a weak solution of this equation is given by the C^1, piecewise C^2-curve made up of 2 semi-circles of radius $\sqrt{A}$ joined by 2 straight line segments parallel to the vector e and separated by the distance $2\sqrt{A}$ (as in figure 5). Note that such 'cigar'-shaped curves can be made with line segments of any length.

A numerical approach to minimize $E^{\text{hor}}_{G^1}(c)$ for variations c with initial and end curves circles at a certain distance produced the two such geodesics shown in Figure 5. Note that the middle curve is indeed close to such a 'cigar'-shape. However, the width of this shape is somewhat greater than $2\sqrt{A}$: this is presumably because the endcurves of this path are not sufficiently far apart. Thus experiments as well as the theory suggest strongly

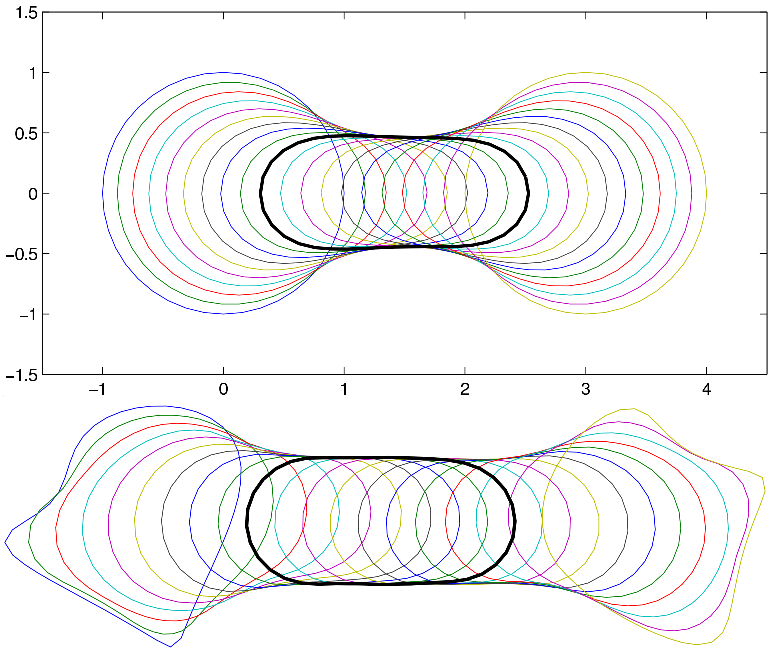


Figure 6. On the top, the geodesic joining circles of radius 1 at distance 3 apart with $A = .1$ (using 20 time samples and a 40-gon for the circle). On the bottom, the geodesic joining 2 'random' shapes of size about 1 at distance 5 apart with $A = .25$ (using 20 time samples and a 48-gon approximation for all curves). In both cases the middle curve which is highlighted.

that geodesics joining any two curves sufficiently far apart compared to their size asymptotically approach a constant 'cigar'-shaped C^1-intermediate curve made up of 2 semi-circles of radius $\sqrt{A}$ and 2 parallel line segments. We conjecture that this is true.

14.3 The growth of a 'bump' on a straight line, when $A = 0$. We have seen above that the geodesic spray is locally well-defined when $A = 0$. To understand this spray and see whether it appears to have global solutions, we take that the initial curve contains a segment with curvature identically zero, i.e. contains a line segment, and that the initial velocity a is set to a smooth function with compact support contained in this segment. For simplicity, we take the velocity a to be a cubic B-spline, i.e. a piecewise cubic which is C^2 with 5 non-C^3 knots approximating a Gaussian blip. The result of integrating is shown in Figure 6. Note several things: first, where the curvature is zero, the curve moves with constant velocity if we follow the orthogonal trajectories. Secondly, where the curve is moving opposite to its curvature (like an expanding circle, the part in the middle), it is

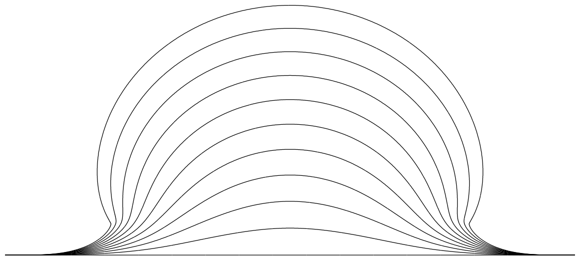


Figure 7. The forward integration of the geodesic equation when $A = 0$, starting from a straight line in the direction given by a smooth bump-like vector field. Note that two corner like singularities with curvature going to ∞ are about to form.

deccelerating; but where it is moving with its curvature (like a contracting circle, the parts on the 2 ends), it is accelerating. This acceleration in the 2 ends, creates higher and higher curvature until a corner forms. In the figure, the simulation is stopped just before the curvature explodes. In the middle, the curve appears to be getting more and more circular. As the corners form, the curve is approaching the boundary of our space. Perhaps, with the right entropy condition, one can prolong the solution past the corners with a suitable piecewise C^1-curve.

Although this calculation assumes $A = 0$, one will find very similar geodesics when A is much smaller than $1/\kappa^2, 1/(\kappa_s \log(a)_s)$ and κ/κ_{ss}, so that the dominant terms in the geodesic equation are those without an A. In other words, geodesics between large smooth curves are basically the same as those with $A = 0$.

14.4 Several geodesic triangles in B_e. We have examined dilations, translations and the evolution of blips. We look next at rotations. To get a pure rotational situation, we consider ellipses centered at $(0, 0)$ with the same eccentricity 3 and maximum radius 1, but differently oriented. We take 3 such ellipses, with orientations differing by 60° and 120° degrees. Joining each pair by a geodesic, we get a triangle in B_e.

We wanted to examine whether along the geodesic joining 2 such ellipses (a) one ellipse rotates into the other or (b) the initial ellipse shrinks towards a circle, while the final ellipse grows, independently of one another. It turns out that, depending on the value of A, both can happen. Note that we get similar geodesics by either changing A or making the ellipses smaller or larger with A held fixed. For each A, we get an absolute distance scale with unit $1/\sqrt{A}$ and, if the ellipses are bigger than this, (b) dominates, while, if smaller, (a) dominates.

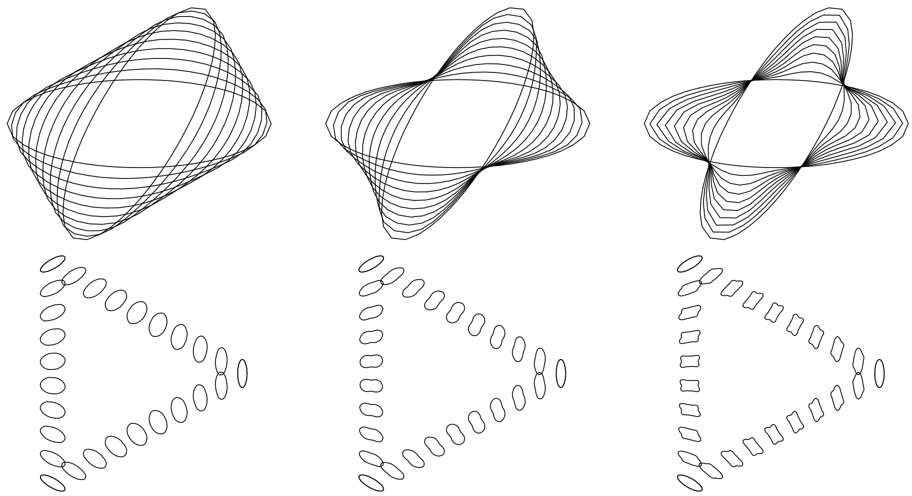


Figure 8. Top Row: Geodesics in three metrics joining the same two ellipses. The ellipses have eccentricity 3, the same center and are at 60° degree angles to each other. At left, $A = 1$; in middle $A = 0.1$; on right $A = 0.01$. Bottom Row: Geodesic triangles in B_e formed by joining three ellipses at angles 0, 60 and 120 degrees, for the same three values of A. Here the intermediate shapes are just rotated versions of the geodesic in the top row but are laid out on a plane triangle for visualization purposes.

The results are shown in Figure 7. We have taken the three values $A = 1, 0.1$ and 0.01. For each value, on the top, we show the geodesic joining 2 of the ellipses as a sequence of curves in their common ambient $\mathbb{R}^2$. Below this, we show the triple of geodesics as a triangle, by displaying the intermediate curves as small shapes along lines joining the ellipses. This Euclidean triangle is being used purely for display, to indicate that the computed structure is a triangle in B_e. Note that for $A = 1$, the intermediate shapes are very close to ellipses, whose axes are rotating; while for $A = 0.01$, the bulges in one ellipse shrink while those of the other grow.

We can also compute the angles in B_e between the sides of this triangle. They work out to be 34° when $A = 1$, i.e. the angle sum for the triangle is 102°, much less than π radians, showing strong negative sectional curvature in the plane containing this triangle. But if $A = 0.1$ or 0.01, the angle is 77° and 69° respectively, giving more than π radians in the triangle. Thus the sectional curvature is positive for such small values of A.

14.5 Notes on the numerical simulations. All simulations in this paper were carried out in MatLab. The forward integration for the geodesic equation for $A = 0$ was carried out by the simplest possible finite difference scheme. This seems very stable and reliable. Solving for the geodesics was done using the MatLab minimization routine `fminunc` using both its medium and large scale modes. This, however, was quite unstable due to discretization artifacts. A general path between two curves was represented by a matrix of points in $\mathbb{R}^2$, approximating each curve by a polygon and sampling the path discretely. The difficulty is that when the polygons have very acute angles, the discretization tends to be highly inaccurate because of the high curvature localized at one vertex. Initially, in order to minimize the number of variables in the problem, we tried to use small numbers of samples and higher order accurate discrete approximations to the derivatives. In all these attempts, the discrete approximation "cheated" by finding minima to the energy of the path with polygons with very small angles. The only way we got around this was to use first order accurate expressions for the derivatives and relatively large numbers of samples (e.g. 48 points on each curve, 20 samples along the geodesic, hence $2 \times 20 \times 48 = 1920$ variables in the expression for the energy.

Another problem is that the energy only depends on the path of unparametrized curves and is independent of the parametrization. To solve this, we added a term to the energy which is minimized by constant speed parametrizations. This still leaves a possibly wandering basepoint, and we added ϵ times another term which asked that all points on each curve should move as normally as possible. In practice, if the initialization was reasonable, this term was not needed. The final discrete energy that was minimized was this. Let $x_{i,j}$ be the i^{th} sample point on the j^{th} curve C_j. For each (i,j), estimate the sum of the squared curvature of C_j plus the squared acceleration of the parametrization by:

$$k(i,j) = \tfrac{1}{2}\Big(\frac{1}{\|x_{i-1,j} - x_{i,j}\|^4} + \frac{1}{\|x_{i,j} - x_{i+1,j}\|^4}\Big) \cdot \|x_{i-1,j} - 2x_{i,j} + x_{i+1,j}\|^2.$$

(The harmonic mean of the segment lengths is used here to further force the parametrization to be uniform.) Then, for each (i,j), the *four* triangles $t = \{a = (i,j), b = (i \pm 1, j), c = (i, j \pm 1)\}$ around (i,j) are considered and the energy is taken to be:

$$\sum_{i,j,t}\Big(\frac{\langle (x_a - x_b), (x_a - x_c)^\perp\rangle^2 + \epsilon\langle (x_a - x_b), (x_a - x_c)\rangle^2}{\|x_a - x_b\|}\Big)(1 + Ak(a)).$$

We make no guarantees about the accuracy of this simulation! The results, however, seem to be stable and reasonable.

15. The Hamiltonian approach

In our previous papers, we have derived the geodesic equation in our various metrics by setting the first variation of the energy of a path equal to 0. Alternately, the geodesic equation is the Hamiltonian flow associated to the first fundamental form (i.e. the length-squared function given by the metric on the tangent bundle). The Hamiltonian approach also provides a mechanism for converting symmetries of the underlying Riemannian manifold into conserved quantities, the momenta. We first need to be quite formal and lay out the basic definitions, esp. distinguishing between the tangent and cotangent bundles rather carefully: The former consists of smooth vector fields along immersions whereas the latter is comprised of 1-currents along immersions. Because of this we work on the tangent bundle and we pull back the symplectic form from the contantent bundle to $T\operatorname{Imm}(S^1,\mathbb{R}^2)$. We use the basics of symplectic geometry and momentum mappings on cotangent bundles in infinite dimensions, and we explain each step. See [**53**], section 2, for a detailed exposition in similar notation as used here.

15.1. The setting. Consider as above the smooth Fréchet manifold $\operatorname{Imm}(S^1,\mathbb{R}^2)$ of all immersions $S^1 \to \mathbb{R}^2$ which is an open subset of $C^\infty(S^1,\mathbb{R}^2)$. The tangent bundle is $T\operatorname{Imm}(S^1,\mathbb{R}^2) = \operatorname{Imm}(S^1,\mathbb{R}^2) \times C^\infty(S^1,\mathbb{R}^2)$, and the cotangent bundle is $T^*\operatorname{Imm}(S^1,\mathbb{R}^2) = \operatorname{Imm}(S^1,\mathbb{R}^2) \times \mathcal{D}(S^1)^2$ where the second factor consists of pairs of periodic distributions.

We consider smooth Riemannian metrics on $\operatorname{Imm}(S^1,\mathbb{R}^2)$, i.e., smooth mappings

$$\begin{aligned}
&G : \operatorname{Imm}(S^1,\mathbb{R}^2) \times C^\infty(S^1,\mathbb{R}^2) \times C^\infty(S^1,\mathbb{R}^2) \to \mathbb{R}\\
&(c,h,k) \mapsto G_c(h,k), \qquad \text{bilinear in } h,k\\
&G_c(h,h) > 0 \qquad \text{for } h \neq 0.
\end{aligned}$$

Each such metric is *weak* in the sense that G_c, viewed as bounded linear mapping

$$\begin{aligned}
G_c : T_c\operatorname{Imm}(S^1,\mathbb{R}^2) = C^\infty(S^1,\mathbb{R}^2) &\to T_c^*\operatorname{Imm}(S^1,\mathbb{R}^2) = \mathcal{D}(S^1)^2\\
G : T\operatorname{Imm}(S^1,\mathbb{R}^2) &\to T^*\operatorname{Imm}(S^1,\mathbb{R}^2)\\
G(c,h) &= (c, G_c(h,\ .\ ))
\end{aligned}$$

is injective, but can never be surjective. We shall need also its tangent mapping

$$TG : T(T\operatorname{Imm}(S^1,\mathbb{R}^2)) \to T(T^*\operatorname{Imm}(S^1,\mathbb{R}^2))$$

We write a tangent vector to $T\operatorname{Imm}(S^1,\mathbb{R}^2)$ in the form $(c,h;k,\ell)$ where $(c,h) \in T\operatorname{Imm}(S^1,\mathbb{R}^2)$ is its foot point, k is its vector component in the $\operatorname{Imm}(S^1,\mathbb{R}^2)$-direction and where ℓ is its component in the $C^\infty(S^1,\mathbb{R}^2)$-direction. Then TG is given by

$$TG(c,h;k,\ell) = (c, G_c(h,\ .\ ); k, D_{(c,k)}G_c(h,\ .\ ) + G_c(\ell,\ .\ ))$$

Moreover, if $X = (c,h;k,\ell)$ then we will write $X_1 = k$ for its first vector component and $X_2 = \ell$ for the second vector component. Note that only these smooth functions on $\mathrm{Imm}(S^1,\mathbb{R}^2)$ whose derivative lies in the image of G in the cotangent bundle have G-gradients. This requirement has only to be satisfied for the first derivative, for the higher ones it follows (see [**42**]). We shall denote by $C^\infty_G(\mathrm{Imm}(S^1,\mathbb{R}^2))$ the space of such smooth functions.

We shall always assume that G is invariant under the reparametrization group $\mathrm{Diff}(S^1)$, hence each such metric induces a Riemann-metric on the quotient space $B_i(S^1,\mathbb{R}^2)=\mathrm{Imm}(S^1,\mathbb{R}^2)/\mathrm{Diff}(S^1)$.

In the sequel we shall further assume that that *the weak Riemannian metric G itself admits G-gradients with respect to the variable c in the following sense:*

$$\boxed{D_{c,m}G_c(h,k)=G_c(m,H_c(h,k))=G_c(K_c(m,h),k)}\quad\text{where}$$
$$\begin{aligned}H,K:\mathrm{Imm}(S^1,\mathbb{R}^2)\times C^\infty(S^1,\mathbb{R}^2)\times C^\infty(S^1,\mathbb{R}^2)&\to C^\infty(S^1,\mathbb{R}^2)\\ (c,h,k)&\mapsto H_c(h,k),K_c(h,k)\end{aligned}$$

are smooth and bilinear in h,k.

Note that H and K could be expressed in (abstract) index notation as $g_{ij,k}g^{kl}$ and $g_{ij,k}g^{il}$. We will check and compute these gradients for several concrete metrics below.

15.2. The fundamental symplectic form on $T\,\mathrm{Imm}(S^1,\mathbb{R}^2)$ induced by a weak Riemannian metric. The basis of Hamiltonian theory is the natural 1-form on the cotangent bundle $T^*\,\mathrm{Imm}(S^1,\mathbb{R}^2)$ given by:

$$\begin{aligned}\Theta:T(T^*\,\mathrm{Imm}(S^1,\mathbb{R}^2))=\mathrm{Imm}(S^1,\mathbb{R}^2)\times\mathcal{D}(S^1)^2\times C^\infty(S^1,\mathbb{R}^2)\times\mathcal{D}(S^1)^2&\to\mathbb{R}\\ (c,\alpha;h,\beta)&\mapsto\langle\alpha,h\rangle.\end{aligned}$$

The pullback via the mapping $G:T\,\mathrm{Imm}(S^1,\mathbb{R}^2)\to T^*\,\mathrm{Imm}(S^1,\mathbb{R}^2)$ of the 1-form Θ is then:

$$(G^*\Theta)_{(c,h)}(c,h;k,\ell)=G_c(h,k).$$

Thus the symplectic form $\omega=-dG^*\Theta$ on $T\,\mathrm{Imm}(S^1,\mathbb{R}^2)$ can be computed as follows, where we use the constant vector fields $(c,h)\mapsto(c,h;k,\ell)$:

$$\begin{aligned}\omega_{(c,h)}((k_1,\ell_1),(k_2,\ell_2))&=-d(G^*\Theta)((k_1,\ell_1),(k_2,\ell_2))|_{(c,h)}\\ &=-D_{(c,k_1)}G_c(h,k_2)-G_c(\ell_1,k_2)+D_{(c,k_2)}G_c(h,k_1)+G_c(\ell_2,k_1)\\ &=G_c\big(k_2,H_c(h,k_1)-K_c(k_1,h)\big)+G_c(\ell_2,k_1)-G_c(\ell_1,k_2)\end{aligned}\tag{1}$$

15.3. The Hamiltonian vector field mapping. Here we compute the Hamiltonian vectorfield $\mathrm{grad}^\omega(f)$ associated to a smooth function f on the tangent space $T\,\mathrm{Imm}(S^1,\mathbb{R}^2)$, that is $f\in C^\infty_G(\mathrm{Imm}(S^1,\mathbb{R}^2)\times C^\infty(S^1,\mathbb{R}^2))$ assuming that it has smooth G-gradients in both factors. See [**42**], section 48. Using the explicit formulas in 15.2, we have:

$$\omega_{(c,h)}\left(\mathrm{grad}^\omega(f)(c,h),(k,\ell)\right)=\omega_{(c,h)}\left((\mathrm{grad}^\omega_1(f)(c,h),\mathrm{grad}^\omega_2(f)(c,h)),(k,\ell)\right)=$$

$$= G_c\big(k, H_c\big(h, \operatorname{grad}_1^\omega(f)(c,h)\big)\big) - G_c(K_c(\operatorname{grad}_1^\omega(f)(c,h),h),k)$$
$$+ G_c(\ell, \operatorname{grad}_1^\omega(f)(c,h)) - G_c(\operatorname{grad}_2^\omega(f)(c,h),k)$$

On the other hand, by the definition of the ω-gradient we have

$$\omega_{(c,h)}\left(\operatorname{grad}^\omega(f)(c,h),(k,\ell)\right) = df(c,h)(k,\ell) = D_{(c,k)}f(c,h) + D_{(h,\ell)}f(c,h)$$
$$= G_c(\operatorname{grad}_1^G(f)(c,h),k) + G_c(\operatorname{grad}_2^G(f)(c,h),\ell)$$

and we get the expression of the Hamiltonian vectorfield:

$$\boxed{\begin{aligned}&\operatorname{grad}_1^\omega(f)(c,h) = \operatorname{grad}_2^G(f)(c,h)\\ &\operatorname{grad}_2^\omega(f)(c,h) = -\operatorname{grad}_1^G(f)(c,h) + H_c\big(h, \operatorname{grad}_2^G(f)(c,h)\big) - K_c(\operatorname{grad}_2^G(f)(c,h),h)\end{aligned}}$$

Note that for a smooth function f on $T\operatorname{Imm}(S^1,\mathbb{R}^2)$ the ω-gradient exists if and only if both G-gradients exist.

15.4. The geodesic equation. The geodesic flow is defined by a vector field on $T\operatorname{Imm}(S^1,\mathbb{R}^2)$. One way to define this vector field is as the Hamiltonian vector field of the energy function

$$E(c,h) = \frac{1}{2}G_c(h,h), \qquad E : \operatorname{Imm}(S^1,\mathbb{R}^2)\times C^\infty(S^1,\mathbb{R}^2)\to\mathbb{R}.$$

The two partial G-gradients are:

$$\begin{aligned}G_c(\operatorname{grad}_2^G(E)(c,h),\ell) &= d_2E(c,h)(\ell) = G_c(h,\ell)\\ \operatorname{grad}_2^G(E)(c,h) &= h\\ G_c(\operatorname{grad}_1^G(E)(c,h),k) &= d_1E(c,h)(k) = \tfrac{1}{2}D_{(c,k)}G_c(h,h)\\ &= \tfrac{1}{2}G_c(k,H_c(h,h))\\ \operatorname{grad}_1^G(E)(c,h) &= \tfrac{1}{2}H_c(h,h).\end{aligned}$$

Thus the geodesic vector field is

$$\begin{aligned}\operatorname{grad}_1^\omega(E)(c,h) &= h\\ \operatorname{grad}_2^\omega(E)(c,h) &= \tfrac{1}{2}H_c(h,h) - K_c(h,h)\end{aligned}$$

and the geodesic equation becomes:

$$\begin{cases} c_t &= h\\ h_t &= \tfrac{1}{2}H_c(h,h) - K_c(h,h)\end{cases} \qquad\text{or}\qquad \boxed{c_{tt} = \tfrac{1}{2}H_c(c_t,c_t) - K_c(c_t,c_t)}$$

This is nothing but the usual formula for the geodesic flow using the Christoffel symbols expanded out using the first derivatives of the metric tensor.

15.5. The momentum mapping for a G-isometric group action. We consider now a (possibly infinite dimensional regular) Lie group with Lie algebra $\mathfrak{g}$ with a right action $g \mapsto r^g$ by isometries on $\mathrm{Imm}(S^1, \mathbb{R}^2)$. If $\mathfrak{X}(\mathrm{Imm}(S^1, \mathbb{R}^2))$ denotes the set of vector fields on $\mathrm{Imm}(S^1, \mathbb{R}^2)$, we can specify this action by the fundamental vector field mapping $\zeta : \mathfrak{g} \to \mathfrak{X}(\mathrm{Imm}(S^1, \mathbb{R}^2))$, which will be a bounded Lie algebra homomorphism. The fundamental vector field $\zeta_X, X \in \mathfrak{g}$ is the infinitesimal action in the sense:

$$\zeta_X(c) = \partial_t|_0 r^{\exp(tX)}(c).$$

We also consider the tangent prolongation of this action on $T\,\mathrm{Imm}(S^1, \mathbb{R}^2)$ where the fundamental vector field is given by

$$\zeta_X^{T\,\mathrm{Imm}} : (c,h) \mapsto (c,h; \zeta_X(c), D_{(c,h)}(\zeta_X)(c) =: \zeta'_X(c,h))$$

The basic assumption is that the action is by isometries,

$$G_c(h,k) = ((r^g)^* G)_c(h,k) = G_{r^g(c)}(T_c(r^g)h, T_c(r^g)k).$$

Differentiating this equation at $g = e$ in the direction $X \in \mathfrak{g}$ we get

$$0 = D_{(c,\zeta_X(c))} G_c(h,k) + G_c(\zeta'_X(c,h),k) + G_c(h, \zeta'_X(c,k)) \tag{1}$$

The key to the Hamiltonian approach is to define the group action by Hamiltonian flows. To do this, we define the *momentum map* $j : \mathfrak{g} \to C^\infty_G(T\,\mathrm{Imm}(S^1, \mathbb{R}^2), \mathbb{R})$ by:

$$\boxed{j_X(c,h) = G_c(\zeta_X(c), h).}$$

Equivalently, since this map is linear, it is often written as a map

$$\mathcal{J} : T\,\mathrm{Imm}(S^1, \mathbb{R}^2) \to \mathfrak{g}', \qquad \langle \mathcal{J}(c,h), X \rangle = j_X(c,h).$$

The main property of the momentum map is that it fits into the following commutative diagram and is a homomorphism of Lie algebras:

$$\begin{tikzcd}
H^0(T\,\mathrm{Imm}) \arrow[r, "i"] & C^\infty_G(T\,\mathrm{Imm}, \mathbb{R}) \arrow[rr, "\mathrm{grad}^\omega"] & & \mathfrak{X}(T\,\mathrm{Imm}, \omega) \arrow[r] & H^1(T\,\mathrm{Imm}) \\
& & \mathfrak{g} \arrow[ul, "j"] \arrow[ur, "\zeta^{T\,\mathrm{Imm}}"'] & &
\end{tikzcd}$$

where $\mathfrak{X}(T\,\mathrm{Imm}, \omega)$ is the space of vector fields on $T\,\mathrm{Imm}$ whose flow leaves ω fixed. We need to check that:

$$\zeta_X(c) = \mathrm{grad}_1^\omega(j_X)(c,h) = \mathrm{grad}_2^G(j_X)(c,h)$$
$$\zeta'_X(c,h) = \mathrm{grad}_2^\omega(j_X)(c,h) = -\,\mathrm{grad}_1^G(j_X)(c,h) + H_c(h, \zeta_X(c)) - K_c(\zeta_X(c), h)$$

The first equation is obvious. To verify the second equation, we take its inner product with some k and use:

$$\begin{aligned}
G\big(k, \mathrm{grad}_1^G(j_X)(c,h)\big) &= D_{(c,k)} j_X(c,h) = D_{(c,k)} G_c(\zeta_X(c), h) + G_c(\zeta'_X(c,k), h) \\
&= G_c(k, H_c(\zeta_X(c), h)) + G_c(\zeta'_X(c,k), h).
\end{aligned}$$

Combining this with (1), the second equation follows. Let us check that it is also a homomorphism of Lie algebras using the Poisson bracket:

$$\begin{aligned}\{j_X, j_Y\}(c,h) &= dj_Y(c,h)(\operatorname{grad}^\omega_1(j_X)(c,h), \operatorname{grad}^\omega_2(j_X)(c,h)) \\ &= dj_Y(c,h)(\zeta_X(c), \zeta'_X(c,h)) \\ &= D_{(c,\zeta_X(c))}G_c(\zeta_Y(c),h) + G_c(\zeta'_Y(c,\zeta_X(x)),h) + G_c(\zeta_Y(c), \zeta'_X(c,h)) \\ &= G_c(\zeta'_Y(c,\zeta_X(c)) - \zeta'_X(c,\zeta_Y(c)),h) \qquad \text{by (1)} \\ &= G_c([\zeta_X,\zeta_Y](c),h) = G_c(\zeta_{[X,Y]}(c),h) = j_{[X,Y]}(c).\end{aligned}$$

Note also that $\mathcal{J}$ is equivariant for the group action, by the following arguments: For g in the Lie group let r^g be the right action on $\operatorname{Imm}(S^1,\mathbb{R}^2)$, then $T(r^g)\circ\zeta_X\circ(r^g)^{-1} = \zeta_{\operatorname{Ad}(g^{-1})X}$. Since r^g is an isometry the mapping $T(r^g)$ is a symplectomorphism for ω, thus $\operatorname{grad}^\omega$ is equivariant. Thus $j_X\circ T(r^g) = j_{\operatorname{Ad}(g)X}$ plus a possible constant which we can rule out since $j_X(c,h)$ is linear in h.

By Emmy Noether's theorem, along any geodesic $t\mapsto c(t,\ .\ )$ this momentum mapping is constant, thus for any $X\in\mathfrak{g}$ we have

$$\boxed{\langle \mathcal{J}(c,c_t), X\rangle = j_X(c,c_t) = G_c(\zeta_X(c),c_t) \quad \text{is constant in } t.}$$

We can apply this construction to the following group actions on $\operatorname{Imm}(S^1,\mathbb{R}^2)$.

- The smooth right action of the group $\operatorname{Diff}(S^1)$ on $\operatorname{Imm}(S^1,\mathbb{R}^2)$, given by composition from the right: $c\mapsto c\circ\varphi$ for $\varphi\in\operatorname{Diff}(S^1)$. For $X\in\mathfrak{X}(S^1)$ the fundamental vector field is then given by

  $$\zeta^{\mathrm{Diff}}_X(c) = \zeta_X(c) = \partial_t|_0(c\circ \mathrm{Fl}^X_t) = c_\theta.X$$

  where Fl^X_t denotes the flow of X. The *reparametrization momentum*, for any vector field X on S^1 is thus:

  $$j_X(c,h) = G_c(c_\theta.X,h).$$

  Assuming the metric is reparametrization invariant, it follows that on any geodesic $c(\theta,t)$, the expression $G_c(c_\theta.X,c_t)$ is constant for all X.
- The left action of the Euclidean motion group $M(2)=\mathbb{R}^2\rtimes SO(2)$ on $\operatorname{Imm}(S^1,\mathbb{R}^2)$ given by $c\mapsto e^{aJ}c+B$ for $(B,e^{aJ})\in\mathbb{R}^2\times SO(2)$. The fundamental vector field mapping is

  $$\zeta_{(B,a)}(c) = aJc+B$$

  The *linear momentum* is thus $G_c(B,h), B\in\mathbb{R}^2$ and if the metric is translation invariant, $G_c(B,c_t)$ will be constant along geodesics. The *angular momentum* is similarly $G_c(Jc,h)$ and if the metric is rotation invariant, then $G_c(Jc,c_t)$ will be constant along geodesics.

- The action of the scaling group of $\mathbb{R}$ given by $c \mapsto e^r c$, with fundamental vector field $\zeta_a(c) = a.c$. If the metric is scale invariant, then the *scaling momentum* $G_c(c, c_t)$ will also be invariant along geodesics.

15.6. Metrics and momenta on the group of diffeomorphisms. Very similar things happen when we consider metrics on the group $\mathrm{Diff}(\mathbb{R}^2)$. As above, the tangent space to $\mathrm{Diff}(\mathbb{R}^2)$ at the identity is the vector space of vector fields $\mathfrak{X}(\mathbb{R}^2)$ on $\mathbb{R}^2$ and we can identify $T\,\mathrm{Diff}(\mathbb{R}^2)$ with the product $\mathrm{Diff}(\mathbb{R}^2) \times \mathfrak{X}(\mathbb{R}^2)$ using right multiplication in the group to identify the tangent at a point φ with that at the identity. The definition of this product decomposition means that right multiplication by ψ carries (φ, X) to $(\varphi \circ \psi, X)$. As usual, suppose that conjugation $\varphi \mapsto \psi \circ \varphi \circ \psi^{-1}$ has the derivative at the identity given by the linear operator Ad_ψ on the Lie algebra $\mathfrak{X}(\mathbb{R}^2)$. It is easy to calculate the explicit formula for Ad:

$$\mathrm{Ad}_\psi(X) = (D\psi \cdot X) \circ \psi^{-1}.$$

Then left multiplication by ψ on $\mathrm{Diff}(\mathbb{R}^2) \times \mathfrak{X}(\mathbb{R}^2)$ is given by $(\varphi, X) \mapsto (\psi \circ \varphi, \mathrm{Ad}_\psi(X))$. We now want to carry over the ideas of 15.5 replacing the space $\mathrm{Imm}(S^1, \mathbb{R}^2)$ by $\mathrm{Diff}(\mathbb{R}^2)$ and the group action there by the right action of $\mathrm{Diff}(\mathbb{R}^2)$ on itself. The Lie algebra $\mathfrak{g}$ is therefore $\mathfrak{X}(\mathbb{R}^2)$ and the fundamental vector field $\zeta_X(c)$ is now the vector field with value

$$\zeta_X(\varphi) = \partial_t|_0(\varphi \mapsto \varphi \circ \exp(tX) \circ \varphi^{-1}) = \mathrm{Ad}_\varphi(X)$$

at the point φ. We now assume we have a positive definite inner product $G(X, Y)$ on the Lie algebra $\mathfrak{X}(\mathbb{R}^2)$ and that we use right translation to extend it to a Riemannian metric on the full group $\mathrm{Diff}(\mathbb{R}^2)$. This metric being, by definition, invariant under the right group action, we have the setting for momentum. The theory of the last section tells us to define the momentum mapping by:

$$j_X(\varphi, Y) = G(\zeta_X(\varphi), Y).$$

Noether's theorem tells us that if $\varphi(t)$ is a geodesic in $\mathrm{Diff}(\mathbb{R}^2)$ for this metric, then this momentum will be constant along the lift of this geodesic to the tangent space. The lift of $\varphi(t)$, in the product decomposition of the tangent space is the curve:

$$t \mapsto (\varphi(t), \partial_t(\varphi) \circ \varphi^{-1}(t))$$

hence the theorem tells us that:

$$G(\mathrm{Ad}_{\varphi(t)}(X), \partial_t(\varphi) \circ \varphi^{-1}(t)) = \text{constant}$$

for all X. If we further assume that Ad has an adjoint with respect to G:

$$G(\mathrm{Ad}_\varphi(X), Y) \equiv G(X, \mathrm{Ad}^*_\varphi(Y))$$

then this invariance of momentum simplifies to:

$$\boxed{\mathrm{Ad}^*_{\varphi(t)}\big(\partial_t(\varphi)\circ\varphi^{-1}(t)\big) = \text{constant}}$$

This is a very strong invariance and it encodes an integrated form of the geodesic equations for the group.

16. Almost local Riemannian metrics

16.1. The general almost local metric G^Φ. We have introduced above the Φ-metrics:

$$G^\Phi_c(h,k) := \int_{S^1}\Phi(\ell_c,\kappa_c(\theta))\langle h(\theta),k(\theta)\rangle ds.$$

Since $\ell(c)$ is an integral operator the integrand is not a local operator, but the nonlocality is very mild. We call it *almost local*. The metric G^Φ is invariant under the reparametization group $\mathrm{Diff}(S^1)$ and under the Euclidean motion group. Note (see [**54**], 2.2) that

$$\begin{aligned}
D_{(c,h)}\ell_c &= \int_{S^1}\frac{\langle h_\theta,c_\theta\rangle}{|c_\theta|}d\theta = \int_{S^1}\langle D_s(h),v\rangle ds\\
&= -\int_{S^1}\langle h,D_s(v)\rangle ds = -\int_{S^1}\kappa(c)\langle h,n\rangle ds\\
D_{(c,h)}\kappa_c &= \frac{\langle Jh_\theta,c_{\theta\theta}\rangle}{|c_\theta|^3}+\frac{\langle Jc_\theta,h_{\theta\theta}\rangle}{|c_\theta|^3}-3\kappa(c)\frac{\langle h_\theta,c_\theta\rangle}{|c_\theta|^2}\\
&= \langle D_s^2(h),n\rangle - 2\kappa\langle D_s(h),v\rangle.
\end{aligned}$$

We compute the G^Φ-gradients of $c\mapsto G^\Phi_c(h,k)$:

$$\begin{aligned}
D_{(c,m)}G^\Phi_c(h,k) = \int_{S^1}\Big(&\partial_1\Phi(\ell,\kappa).D_{(c,m)}\ell_c.\langle h,k\rangle + \partial_2\Phi(\ell,\kappa).D_{(c,m)}\kappa_c.\langle h,k\rangle\\
&+\Phi(\ell,\kappa).\langle h,k\rangle\langle .D_s(m),v\rangle\Big)\,ds
\end{aligned}$$

$$\begin{aligned}
&= -\int_{S^1}\kappa_c\langle m,n\rangle ds\cdot\int_{S^1}\partial_1\Phi(\ell,\kappa)\langle h,k\rangle ds\\
&\quad+\int_{S^1}\Big(\partial_2\Phi(\ell,\kappa)(\langle D_s^2(m),n\rangle-2\kappa\langle D_s(m),v\rangle)+\Phi(\ell,\kappa)\langle D_s(m),v\rangle\Big)\langle h,k\rangle ds\\
&= \int_{S^1}\Phi(\ell,\kappa)\Big\langle m\ ,\ \frac{1}{\Phi(\ell,\kappa)}\Big(-\kappa_c\Big(\int\partial_1\Phi(\ell,\kappa)\langle h,k\rangle ds\Big)n + D_s^2\Big(\partial_2\Phi(\ell,\kappa)\langle h,k\rangle n\Big)\\
&\qquad + 2D_s\Big(\partial_2\Phi(\ell,\kappa)\kappa\langle h,k\rangle v\Big) - D_s\Big(\Phi(\ell,\kappa)\langle h,k\rangle v\Big)\Big)\quad\Big\rangle ds
\end{aligned}$$

According to 15.1 we should rewrite this as

$$D_{(c,m)}G^\Phi_c(h,k) = G^\Phi_c(K^\Phi_c(m,h),k) = G^\Phi_c\big(m,H^\Phi_c(h,k)\big),$$

where the two G^Φ-gradients K^Φ and H^Φ of $c \mapsto G_c^\Phi(h,k)$ are given by:

$$\begin{aligned}
K_c^\Phi(m,h) &= -\Big(\int_{S^1} \kappa_c \langle m, n\rangle ds\Big)\frac{\partial_1\Phi(\ell,\kappa)}{\Phi(\ell,\kappa)}h \\
&\quad + \frac{\partial_2\Phi(\ell,\kappa)}{\Phi(\ell,\kappa)}\Big(\langle D_s^2(m), n\rangle - 2\kappa\langle D_s(m), v\rangle\Big)h + \langle D_s(m), v\rangle h \\
H_c^\Phi(h,k) &= \frac{1}{\Phi(\ell,\kappa)}\bigg(- \Big(\kappa_c \int \partial_1\Phi(\ell,\kappa)\langle h,k\rangle ds\Big)n + D_s^2\Big(\partial_2\Phi(\ell,\kappa)\langle h,k\rangle n\Big) + \\
&\qquad + 2D_s\Big(\partial_2\Phi(\ell,\kappa)\kappa\langle h,k\rangle v\Big) - D_s\Big(\Phi(\ell,\kappa)\langle h,k\rangle v\Big)\bigg)
\end{aligned}$$

By substitution into the general formula of 15.4, this gives the geodesic equation for G^Φ, but in a form which doesn't seem very revealing, hence we omit it. Below we shall give the equation for the special case of horizontal geodesics, i.e. geodesics in B_i.

16.2. Conserved momenta for G^Φ. According to 15.5 the momentum mappings for the reparametrization, translation and rotation group actions are conserved along any geodesic $t \mapsto c(t, \ .\ )$:

$\Phi(\ell_c,\kappa_c)\langle v, c_t\rangle \lvert c_\theta\rvert^2 \ \in \mathfrak{X}(S^1)$	reparametrization momentum
$\int_{S^1} \Phi(\ell_c,\kappa_c) c_t ds \ \in \mathbb{R}^2$	linear momentum
$\int_{S^1} \Phi(\ell_c,\kappa_c)\langle Jc, c_t\rangle ds \in \mathbb{R}$	angular momentum

Note that setting the reparametrization momentum to 0 and doing symplectic reduction there amounts exactly to investigating the quotient space $B_i(S^1,\mathbb{R}^2) = \mathrm{Imm}(S^1,\mathbb{R}^2)/\operatorname{Diff}(S^1)$ and using horizontal geodesics for doing so; a horizontal geodesic is one for which $\langle v, c_t\rangle = 0$; or equivalently it is G^Φ-normal to the $\operatorname{Diff}(S^1)$-orbits. If it is normal at one time it is normal forever (since the reparametrization momentum is conserved). This was the approach taken in [**54**].

16.3. Horizontality for G^Φ. The tangent vectors to the $\operatorname{Diff}(S^1)$ orbit through c are $T_c(c \circ \operatorname{Diff}(S^1)) = \{X.c_\theta : X \in C^\infty(S^1,\mathbb{R})\}$. Thus the bundle of horizontal vectors is

$$\begin{aligned}
\mathcal{N}_c &= \{h \in C^\infty(S^1,\mathbb{R}^2) : \langle h, v\rangle = 0\} \\
&= \{a.n \in C^\infty(S^1,\mathbb{R}^2) : a \in C^\infty(S^1,\mathbb{R})\}
\end{aligned}$$

A tangent vector $h \in T_c\mathrm{Imm}(S^1,\mathbb{R}^2) = C^\infty(S^1,\mathbb{R}^2)$ has an orthonormal decomposition

$$h = h^\top + h^\perp \in T_c(c \circ \operatorname{Diff}^+(S^1)) \oplus \mathcal{N}_c \qquad \text{where}$$

$$
\begin{aligned}
h^\top &= \langle h, v\rangle v \in T_c(c \circ \mathrm{Diff}^+(S^1)),\\
h^\perp &= \langle h, n\rangle n \in \mathcal{N}_c,
\end{aligned}
$$

into smooth tangential and normal components, independent of the choice of $\Phi(\ell, \kappa)$. For the following result the proof given in [**54**], 2.5 works without any change:

Lemma. *For any smooth path c in $\mathrm{Imm}(S^1, \mathbb{R}^2)$ there exists a smooth path φ in $\mathrm{Diff}(S^1)$ with $\varphi(0, \ .\ ) = \mathrm{Id}_{S^1}$ depending smoothly on c such that the path e given by $e(t, \theta) = c(t, \varphi(t, \theta))$ is horizontal: $e_t \perp e_\theta$.* □

Consider a path $t \mapsto c(\cdot, t)$ in the manifold $\mathrm{Imm}(S^1, \mathbb{R}^2)$. It projects to a path $\pi \circ c$ in $B_i(S^1, \mathbb{R}^2)$ whose energy is called the *horizontal energy* of c:

$$
\begin{aligned}
E^{\mathrm{hor}}_{G^\Phi}(c) &= E_{G^\Phi}(\pi \circ c) = \tfrac{1}{2} \int_a^b G^\Phi_{\pi(c)}(T_c\pi.c_t, T_c\pi.c_t)\, dt\\
&= \tfrac{1}{2} \int_a^b G^\Phi_c(c_t^\perp, c_t^\perp)\, dt = \tfrac{1}{2} \int_a^b \int_{S^1} \Phi(\ell_c, \kappa_c) \langle c_t^\perp, c_t^\perp\rangle ds\, dt
\end{aligned}
$$

$$
\boxed{E^{\mathrm{hor}}_{G^\Phi}(c) = \tfrac{1}{2} \int_a^b \int_{S^1} \Phi(\ell_c, \kappa_c) \langle c_t, n\rangle^2\, d\theta dt}
$$

For a horizontal path this is just the usual energy. As in [**54**], 3.12 we can express $E^{\mathrm{hor}}(c)$ as an integral over the graph S of c, the immersed surface $S \subset \mathbb{R}^3$ parameterized by $(t, \theta) \mapsto (t, c(t, \theta))$, in terms of the surface area $d\mu_S = |\Phi_t \times \Phi_\theta|\, d\theta\, dt$ and the unit normal $n_S = (n_S^0, n_S^1, n_S^2)$ of S:

$$
E^{\mathrm{hor}}_{G^\Phi}(c) = \tfrac{1}{2} \int_{[a,b] \times S^1} \Phi(\ell_c, \kappa_c) \frac{|n_S^0|^2}{\sqrt{1 - |n_S^0|^2}}\, d\mu_S
$$

Here the final expression is only in terms of the surface S and its fibration over the time axis, and is valid for any path c. This anisotropic area functional has to be minimized in order to prove that geodesics exists between arbitrary curves (of the same degree) in $B_i(S^1, \mathbb{R}^2)$.

16.4. The horizontal geodesic equation. Let $c(\theta, t)$ be a horizontal geodesic for the metric G^Φ. Then $c_t(\theta, t) = a(\theta, t).n(\theta, t)$. Denote the integral of a function over the curve with respect to arclength by a bar. Then the geodesic equation for horizontal geodesics is:

$$
\boxed{\begin{aligned}
a_t = \frac{-1}{2\Phi}\Big(&\left(-\kappa\Phi + \kappa^2 \partial_2 \Phi\right) a^2 - D_s^2\left(\partial_2\Phi \cdot a^2\right) + 2\partial_2\Phi \cdot a D_s^2(a)\\
&- 2\partial_1\Phi \cdot \overline{(\kappa a)} \cdot a + \overline{(\partial_1 \Phi \cdot a^2)} \cdot \kappa\Big)
\end{aligned}}
$$

This comes immediately from the formulas for H and K in the metric G^Φ when you substitute $m = h = k = a.n$ and consider only the n-part. We obtain in this case:

$$\begin{aligned}
\Phi\cdot\langle K,n\rangle &= -\overline{(\kappa a)}.\partial_1\Phi.a + \partial_2\Phi.D_s^2(a).a + \partial_2\Phi.\kappa^2a^2 - \Phi\kappa a^2\\
\Phi\cdot\langle H,n\rangle &= -\overline{(\partial_1\Phi a^2)}.\kappa + D_s^2(\partial_2\Phi.a^2) + \partial_2\Phi.\kappa^2a^2 - \Phi\kappa a^2.
\end{aligned}$$

and the geodesic formula follows by substitution.

16.5. Curvature on $B_{i,f}(S^1,\mathbb{R}^2)$ for G^Φ. We compute the curvature of $B_i(S^1,\mathbb{R}^2)$ in the general almost local metric G^Φ. We proceed as in [**54**], 2.4.3. We use the following chart near $C\in B_i(S^1,\mathbb{R}^2)$. Let $c\in \operatorname{Imm}_f(S^1,\mathbb{R}^2)$ be parametrized by arclength with $\pi(c)=C$ of length L, with unit normal n_c. We assume that the parameter θ runs in the scaled circle S^1_L below.

$$\begin{aligned}
&\psi: C^\infty(S^1_L,(-\varepsilon,\varepsilon))\to \operatorname{Imm}_f(S^1_L,\mathbb{R}^2), \qquad \mathcal{Q}(c):=\psi(C^\infty(S^1_L,(-\varepsilon,\varepsilon)))\\
&\psi(f)(\theta)=c(\theta)+f(\theta)n_c(\theta)=c(\theta)+f(\theta)ic'(\theta),\\
&\pi\circ\psi: C^\infty(S^1_L,(-\varepsilon,\varepsilon))\to B_{i,f}(S^1,\mathbb{R}^2),
\end{aligned}$$

where ε is so small that $\psi(f)$ is an embedding for each f. We have (see [**54**], 2.4.3)

$$\begin{aligned}
\psi(f)' &= c'+f'ic'+fic'' = (1-f\kappa_c)c'+f'ic'\\
\psi(f)'' &= c''+f''ic'+2f'ic''+fic''' = -(2f'\kappa_c+f\kappa_c')c'+(\kappa_c+f''-f\kappa_c^2)ic'\\
n_{\psi(f)} &= \frac{1}{\sqrt{(1-f\kappa_c)^2+f'^2}}\Big((1-f\kappa_c)ic'-f'c'\Big),\\
T_f\psi.h &= h.ic' \qquad \in C^\infty(S^1,\mathbb{R}^2)=T_{\psi(f)}\operatorname{Imm}_f(S^1_L,\mathbb{R}^2)\\
&= \frac{h(1-f\kappa_c)}{\sqrt{(1-f\kappa_c)^2+f'^2}}n_{\psi(f)} + \frac{hf'}{(1-f\kappa_c)^2+f'^2}\psi(f)',\\
(T_f\psi.h)^\perp &= \frac{h(1-f\kappa_c)}{\sqrt{(1-f\kappa_c)^2+f'^2}}n_{\psi(f)} \qquad \in\mathcal{N}_{\psi(f)},\\
\kappa_{\psi(f)} &= \frac{1}{((1-f\kappa_c)^2+f'^2)^{3/2}}\langle i\psi(f)',\psi(f)''\rangle\\
&= \kappa_c+(f''+f\kappa_c^2)+(f^2\kappa_c^3+\tfrac12 f'^2\kappa_c+ff'\kappa_c'+2ff''\kappa_c)+O(f^3)\\
\ell(\psi(f)) &= \int_{S^1_L}|\psi(f)|\,d\theta = \int_{S^1_L}(1-2f\kappa_c+f^2\kappa_c^2+f'^2)^{1/2}\,d\theta\\
&= \int_{S^1_L}\Big(1-f\kappa_c+\frac{f'^2}{2}+O(f^3)\Big)\,d\theta = L-\overline{f\kappa_c}+\tfrac12\overline{f'^2}+O(f^3)
\end{aligned}$$

where we use the shorthand $\overline{g} = \int_{S^1_L} g(\theta)\,d\theta = \int_{S^1_L} g(\theta)\,ds$. Let G^Φ denote also the induced metric on $B_{i,f}(S^1_L, \mathbb{R}^2)$. Since π is a Riemannian submersion, for $f \in C^\infty(S^1_L, (-\varepsilon, \varepsilon))$ and $h, k \in C^\infty(S^1_L, \mathbb{R})$ we have

$$((\pi\circ\psi)^* G^\Phi)_f(h,k) = G^\Phi_{\pi(\psi(f))}\Big(T_f(\pi\circ\psi)h, T_f(\pi\circ\psi)k\Big)$$
$$= G^\Phi_{\psi(f)}\Big((T_f\psi.h)^\perp, (T_f\psi.k)^\perp\Big) = \int_{S^1_L} \Phi(\ell(\psi(f)), \kappa_{\psi(f)}) \frac{hk(1-f\kappa_c)^2}{\sqrt{(1-f\kappa_c)^2 + f'^2}}\,d\theta$$

We have to compute second derivatives in f of this. For that we expand the main contributing expressions in f to order 2:

$$\begin{aligned}
&(1-f\kappa)^2(1-2f\kappa+f^2\kappa^2+f'^2)^{-1/2} = 1 - f\kappa - \tfrac12 f'^2 + O(f^3)\\
&\Phi(\ell,\kappa) = \Phi(L,\kappa_c) + \partial_1\Phi(L,\kappa_c)(\ell - L) + \partial_2\Phi(L,\kappa_c)(\kappa-\kappa_c)\\
&\quad + \partial_1\partial_2\Phi(L,\kappa_c)(\ell-L)(\kappa-\kappa_c)\\
&\quad + \frac{\partial_1^2\Phi(L,\kappa_c)}{2}(\ell-L)^2 + \frac{\partial_2^2\Phi(L,\kappa_c)}{2}(\kappa-\kappa_c)^2 + O(3)
\end{aligned}$$

We simplify notation as $\kappa = \kappa_c$, $\Phi = \Phi(L,\kappa_c)$, $((\pi\circ\psi)^*G^\Phi)_f = G^\Phi_f$ etc. and expand the metric:

$$\begin{aligned}
G^\Phi_f(h,k) = \int_{S^1_L} hk\bigg(&\Phi - \partial_1\Phi.\overline{f\kappa} + \partial_2\Phi.(f''+f\kappa^2) - \Phi.f\kappa\\
&+ \tfrac12\partial_1\Phi.\overline{f'^2} + \partial_2\Phi.(f^2\kappa^3 + \tfrac12 f'^2\kappa + ff'\kappa' + 2ff''\kappa)\\
&- \partial_1\partial_2\Phi.\overline{f\kappa}(f''+f\kappa^2) + \frac{\partial_1^2\Phi}{2}(\overline{f\kappa})^2 + \frac{\partial_2^2\Phi}{2}(f''+f\kappa^2)^2\\
&+ \partial_1\Phi.f\kappa.\overline{f\kappa} - \partial_2\Phi.f\kappa.(f''+f\kappa^2) - \Phi.\tfrac12 f'^2\bigg)d\theta + O(f^3)
\end{aligned}$$

Note that $G^\varphi_0(h,k) = \int_{S^1_L} hk\Phi\,d\theta$. We to differentiate the metric and compute the Christoffel symbol at the center $f = 0$

$$\begin{aligned}
&-2G^A_0(\Gamma_0(h,k), l) = -dG^A(0)(l)(h,k) + dG^A(0)(h)(k,l) + dG^A(0)(k)(l,h)\\
&= \int_{S^1_L}\Big(-\partial_1\Phi.\overline{h\kappa}.kl - \partial_1\Phi.h.\overline{k\kappa}.l + \partial_1\Phi.hk\int l\kappa\,d\theta_1 - \partial_2\Phi''.hkl\\
&\qquad - 2\partial_2\Phi'.h'kl - 2\partial_2\Phi'.hk'l - 2\partial_2\Phi.h'k'l + \partial_2\Phi.hkl\kappa^2 - \Phi.hkl\kappa\Big)d\theta
\end{aligned}$$

Thus

$$\begin{aligned}
\Gamma_0(h,k) = \frac{1}{2\Phi}\Big(&\partial_1\Phi.(\overline{h\kappa}.k + h.\overline{k\kappa}) - \kappa\overline{\partial_1\Phi.hk}\\
&+ \partial_2\Phi''.hk + 2\partial_2\Phi'.h'k + 2\partial_2\Phi' hk' + 2\partial_2\Phi.h'k'
\end{aligned}$$

$$- \partial_2\Phi.hk\kappa^2 + \Phi.hk\kappa\Big)$$

Letting $h = k = f_t = a$, this leads to the geodesic equation from 16.4. For the sectional curvature we use the following formula which is valid in a chart:

$$\begin{aligned}
2R_f(m,h,m,h) &= 2G^A_f(R_f(m,h)m,h) = \\
&= -2d^2G^A(f)(m,h)(h,m) + d^2G^A(f)(m,m)(h,h) + d^2G^A(f)(h,h)(m,m) \\
&\quad - 2G^A(\Gamma(h,m),\Gamma(m,h)) + 2G^A(\Gamma(m,m),\Gamma(h,h))
\end{aligned}$$

The sectional curvature at the two-dimensional subspace $P_f(m,h)$ of the tangent space which is spanned by m and h is then given by:

$$k_f(P_f(m,h)) = -\frac{G^\Phi_f(R(m,h)m,h)}{\|m\|^2\|h\|^2 - G^\Phi_f(m,h)^2}.$$

We compute this directly for $f = 0$, using the expansion up to order 2 of $G^A_f(h,k)$ and the Christoffels. We let $W(\theta_1,\theta_2) = h(\theta_1)m(\theta_2) - h(\theta_2)m(\theta_1)$ so that its second derivative $\partial_2 W(\theta_1,\theta_1) = W_2(\theta_1,\theta_1) = h(\theta_1)m'(\theta_1) - h'(\theta_1)m(\theta_1)$ is the Wronskian of h and m. Then we have our final result for the main expression in the horizontal sectional curvature, where we use $\int = \int_{S^1_L}$, $\overline{g} = \int_{S^1_L} g\,ds$, and $\Phi_1 = \partial_1\Phi$ etc. Also recall that the base curve is parametrized by arc-length.

$$\boxed{\begin{aligned}
&R^\Phi_0(m,h,m,h) = G^\Phi_0(R_0(m,h)m,h) = \\
&= \int \Big(\kappa.\Phi_2 - \frac{\Phi}{2} + \frac{\Phi_2.\Phi_2'' - 2(\Phi_2')^2 - (\Phi_2\kappa)^2}{2\Phi}\Big)(\theta_1)W_2(\theta_1,\theta_1)^2\,d\theta_1 \\
&+ \int \frac{\Phi_{22}(\theta_1)}{2} W_{22}(\theta_1,\theta_1)^2\,d\theta_1 \\
&+ \iint \Big(\frac{\Phi_1'\Phi_2}{\Phi} - \frac{\Phi_1\Phi_2\Phi_1'}{\Phi^2}\Big)(\theta_1)W_2(\theta_1,\theta_1)\int W(\theta_1,\theta_2)\kappa(\theta_2)\,d\theta_2\,d\theta_1 \\
&+ \iint \Big(\frac{\Phi_1\Phi_2}{\Phi} - \Phi_{12}\Big)(\theta_1)W_{22}(\theta_1,\theta_1)\int W(\theta_1,\theta_2)\kappa(\theta_2)\,d\theta_2\,d\theta_1 \\
&+ \iint \frac{\Phi_1(\theta_1)}{2}\Big(1 - \frac{\Phi_2.\kappa}{\Phi}(\theta_2)\Big)W_1(\theta_1,\theta_2)^2\,d\theta_2\,d\theta_1 \\
&+ \iint \Big(\frac{\Phi_2.\kappa^3 - \Phi_2''.\kappa}{4\Phi} - \frac{\kappa^2}{4} + \Big(\frac{\Phi_2'.\kappa}{2\Phi}\Big)' + \overline{\Big(\frac{\kappa^2}{8\Phi}\Big).\Phi_1}\Big)(\theta_1)\Phi_1(\theta_2)W(\theta_1,\theta_2)^2\,d\theta_2\,d\theta_1 \\
&+ \iiint \Big(\frac{\Phi_{11}}{2} - \frac{\Phi_1^2}{4\Phi}\Big)(\theta_1) - \Phi_1(\theta_1)\frac{\Phi_1}{2\Phi}(\theta_2)\Big) \\
&\qquad \kappa(\theta_2)\kappa(\theta_3)W(\theta_1,\theta_2)W(\theta_1,\theta_3)\,d\theta_2\,d\theta_1\,d\theta_3
\end{aligned}}$$

16.6. Special case: the metric G^A. If we choose $\Phi(\ell_c, \kappa_c) = 1 + A\kappa_c^2$ then we obtain the metric used in [**54**], given by

$$G_c^A(h,k) = \int_{S^1} (1 + A\kappa_c(\theta)^2)\langle h(\theta), k(\theta)\rangle ds.$$

As shown in our earlier paper, $\sqrt{\ell}$ is Lipschitz in this metric and the metric dominates the Frechet metric.

The horizontal geodesic equation for the G^A-metric reduces to

$$a_t = \frac{-\frac{1}{2}\kappa_c a^2 + A\left(a^2(-D_s^2(\kappa_c) + \frac{1}{2}\kappa_c^3) - 4D_s(\kappa_c)aD_s(a) - 2\kappa_c D_s(a)^2\right)}{1 + A\kappa_c^2}$$

as found in [**54**],4.2. Along a geodesic $t \mapsto c(t,\ .\ )$ we have the following conserved quantities:

$$(1 + A\kappa_c^2)\langle v, c_t\rangle|c_\theta|^2 \in \mathfrak{X}(S^1) \qquad \text{reparametrization momentum}$$

$$\int_{S^1} (1 + A\kappa_c^2)c_t ds \in \mathbb{R}^2 \qquad \text{linear momentum}$$

$$\int_{S^1} (1 + A\kappa_c^2)\langle Jc, c_t\rangle ds \in \mathbb{R} \qquad \text{angular momentum}$$

For $\Phi(\ell,\kappa) = 1 + A\kappa^2$ we have $\partial_1\Phi = 0$, $\partial_2\Phi = 2A\kappa$, $\partial_2^2\Phi = 2A$, and the general curvature formula in 16.5 for the horizontal curvature specializes to the formula in [**54**], 4.6.4:

$$R_0^\Phi(m,h,m,h) = \int \Big(-\frac{(1 - A\kappa^2)^2 - 4A^2\kappa\kappa'' + 8A^2\kappa'^2}{2(1 + A\kappa^2)} W_2^2 + AW_{22}^2 \Big) d\theta.$$

16.7. Special case: the conformal metrics. We put $\Phi(\ell(c), \kappa(c)) = \Phi(\ell(c))$ and obtain the metric proposed by Menucci and Yezzi and, for Φ linear, independently by Shah [**71**]:

$$G_c^\Phi(h,k) = \Phi(\ell_c)\int_{S^1} \langle h,k\rangle ds = \Phi(\ell_c)G_c^0(h,k).$$

All these metrics are conformally equivalent to the basic L^2-metric G^0. As they show, the infimum of path lengths in this metric is positive so long as Φ satifies an inequality $\Phi(\ell) \geq C.\ell$ for some $C > 0$. This follows, as in [**54**], 3.4, by the inequality on area swept out by the curves in a horizontal path $c_t = a.n$:

$$\int |a|.ds \leq \left(\int a^2.ds\right)^{1/2} \cdot \ell^{1/2} \leq \left(\frac{\ell}{\Phi(\ell)}\right)^{1/2} \cdot (G^\Phi(a,a))^{1/2}$$

$$\text{Area swept out} \leq \max_t \left(\frac{\ell_{c(t,\cdot)}}{\Phi(\ell_{c(t,\cdot)})}\right)^{1/2} \cdot \big(G^\Phi\text{-path length}\big) \leq \frac{G^\Phi\text{-path length}}{\sqrt{C}}.$$

The horizontal geodesic equation reduces to:

$$a_t = \frac{\kappa}{2}a^2 - \frac{\partial_1\Phi}{\Phi}\cdot\left(\tfrac{1}{2}\left(\int a^2.ds\right)\kappa - \left(\int \kappa.a.ds\right)a\right)$$

If we change variables and write $b(s,t) = \Phi(\ell(t)).a(s,t)$, then this equation simplifies to:

$$\boxed{b_t = \frac{\kappa}{2\Phi}\left(b^2 - \frac{\partial_1\Phi}{\Phi}\int b^2\right)}$$

Along a geodesic $t \mapsto c(t,\ .\ )$ we have the following conserved quantities:

$$\begin{aligned}
&\Phi(\ell_c)\langle v, c_t\rangle|c'(\theta)|^2\ \in \mathfrak{X}(S^1) &&\text{reparametrization momentum}\\
&\Phi(\ell_c)\int_{S^1} c_t ds\ \in \mathbb{R}^2 &&\text{linear momentum}\\
&\Phi(\ell_c)\int_{S^1}\langle Jc, c_t\rangle ds \in \mathbb{R} &&\text{angular momentum}
\end{aligned}$$

For the conformal metrics, sectional curvature has been computed by Shah [**71**] using the method of local charts from [**54**]. We specialize formula 16.5 to the case that $\Phi(\ell,\kappa) = \Phi(\ell)$ is independent of κ. Then $\partial_2\Phi = 0$. We also assume that h, m are orthonormal so that $\Phi\overline{h^2} = \Phi\overline{m^2} = 1$ and $\Phi\overline{hm} = 0$. Then the the sectional curvature at the two-dimensional subspace $P_0(m,h)$ of the tangent space which is spanned by m and h is then given by:

$$\begin{aligned}
k_0(P_0(m,h)) &= -\frac{G_0^\Phi(R_0(m,h)m,h)}{\|m\|^2\|h\|^2 - G_0^\Phi(m,h)^2} = \\
&= \tfrac{1}{2}\Phi.\overline{W(h,m)^2} + \frac{\partial_1\Phi}{4\Phi}.(\overline{m^2\kappa^2} + \overline{h^2\kappa^2}) + \frac{3(\partial_1\Phi)^2 - 2\Phi\partial_1^2\Phi}{4\Phi^2}(\overline{h\kappa}^2 + \overline{m\kappa}^2)\\
&\quad - \frac{\partial_1\Phi}{2\Phi}(\overline{m'^2} + \overline{h'^2}) - \frac{(\partial_1\Phi)^2}{4\Phi^3}\overline{\kappa^2}
\end{aligned}$$

which is the same as the equation (11) in [**71**]. Note that the first line is positive while the last line is negative. The first term is the curvature term for the H^0-metric. The key point about this formula is how many positive terms it has. This makes it very hard to get smooth geodesics in this metric. For example, in the case where $\Phi(\ell) = c.\ell$, the analysis of Shah [**71**] proves that the infimum of G^Φ path length between two embedded curves C and D is exactly the area of the symmetric difference of their interiors: $\mathrm{Area}(\mathrm{Int}(C)\Delta\,\mathrm{Int}(D))$, but that this length is realized by a smooth path if and only if C and D can be connected by 'grassfire', i.e. a family in which the length $|c_t(\theta,t)| \equiv 1$.

16.8. Special case: the smooth scale invariant metric G^{SI}. Choosing the function $\Phi(\ell,\kappa) = \ell^{-3} + A\frac{\kappa^2}{\ell}$ we obtain the metric:

$$G_c^{SI}(h,k) = \int_{S^1} \Big(\frac{1}{\ell_c^3} + A\frac{\kappa_c^2}{\ell_c}\Big)\langle h,k\rangle ds.$$

The beauty of this metric is that (a) it is scale invariant and (b) $\log(\ell)$ is Lipschitz, hence the infimum of path lengths is always positive. Scale invariance is clear: changing c, h, k to $\lambda\cdot c, \lambda\cdot h, \lambda\cdot k$ changes ℓ to $\lambda\cdot\ell$ and κ to κ/λ so the λ's in G^{SI} cancel out. To see the second fact, take a horizontal path $c_t = a\cdot n$, $0\le t\le 1$, and abbreviate the lengths of the curves in this path, $\ell_{c(t,\cdot)}$, to $\ell(t)$. Then we have:

$$\begin{aligned}
\frac{\partial \log\ell(t)}{\partial t} &= \frac{1}{\ell(t)}\int_{S^1}\kappa_{c(t,\cdot)}(\theta)\cdot a(\theta,t)ds, \quad \text{hence}\\
\left|\frac{\partial \log\ell(t)}{\partial t}\right| &= \left(\frac{\int \kappa^2a^2ds}{\ell(t)}\right)^{1/2}\cdot\left(\frac{\int 1\cdot ds}{\ell(t)}\right)^{1/2}\\
&\le \frac{1}{\sqrt{A}}\big(G^{SI}(a,a)\big)^{1/2}, \quad \text{hence}\\
|\log(\ell(1)) - \log(\ell(0))| &\le \text{SI-path length}/\sqrt{A}.
\end{aligned}$$

Thus in a path whose length in this metric is K, the lengths of the individual curves can increase or decrease at most by a factor $e^{K/\sqrt{A}}$. Now use the same argument as above to control the area swept out by such a path:

$$\begin{aligned}
\int |a|ds &\le \Big(\int a^2ds\Big)^{1/2}\cdot\Big(\int 1\cdot ds\Big)^{1/2}\\
&\le \big(\ell^3 G^{SI}(a,a)\big)^{1/2}\cdot\ell^{1/2} = \ell^2\cdot G^{SI}(a,a)^{1/2}, \quad \text{hence}\\
\text{Area-swept-out} &\le e^{K/\sqrt{A}}\ell(0)^2\cdot K
\end{aligned}$$

which verifies the second fact. We can readily calculate the geodesic equation for horizontal geodesics in this metric as another special case of the equation for G^Φ:

$$\begin{aligned}
a_t = \frac{-1}{1+A(\ell\kappa)^2}\bigg(&\big(-1+A(\ell\kappa)^2\big)\frac{\kappa a^2}{2} - A\ell^2D_s^2(\kappa)a^2 - 2A\ell^2\kappa D_s(a)^2\\
&- 4A\ell^2D_s(\kappa)aD_s(a) + \big(3+A(\ell\kappa)^2\big)\,\overline{(a\kappa)}\cdot a - \frac{3}{2}\overline{(a^2)}\cdot\kappa - \frac{A\ell^2}{2}\overline{(\kappa a)^2}\cdot\kappa\bigg)
\end{aligned}$$

where the "overline" stands now for the *average* of a function over the curve, i.e. $\int\cdots ds/\ell$. Since this metric is scale invariant, there are now *four* conserved quantities, instead of three:

$$\Phi(\ell,\kappa)\langle v,c_t\rangle|c'(\theta)|^2 \ \in \mathfrak{X}(S^1) \qquad \text{reparametrization momentum}$$

$$\int_{S^1} \Phi(\ell,\kappa) c_t ds \ \in \mathbb{R}^2 \qquad \text{linear momentum}$$
$$\int_{S^1} \Phi(\ell,\kappa) \langle Jc, c_t\rangle ds \in \mathbb{R} \qquad \text{angular momentum}$$
$$\int_{S^1} \Phi(\ell,\kappa) \langle c, c_t\rangle ds \in \mathbb{R} \qquad \text{scaling momentum}$$

It would be very interesting to compute and compare geodesics in these special metrics.

16.9. The Wasserstein metric and a related G^Φ-metric. The Wasserstein metric (also known as the Monge-Kantorovich metric) is a metric between probability measures on a common metric space, see [**3**], and [**4**] for more details. It has been studied for many years globally and is defined as follows: let μ and ν be 2 probability measures on a metric space (X, d). Consider all measures ρ on $X \times X$ whose marginals under the 2 projections are μ and ν. Then:

$$d_{\text{wass}}(\mu,\nu) = \inf_{\rho: p_{1,*}(\rho)=\mu, p_{2,*}(\rho)=\nu} \iint_{X\times X} d(x,y) d\rho(x,y).$$

It was discovered only recently by Benamou and Brenier [**15**] that, if $X = \mathbb{R}^n$, this is, in fact, path length for a Riemannian metric on the space of probability measures $\mathcal{P}$. In their theory, the tangent space at μ to the space of probability measures and the infinitesimal metric are defined by:

$$T_{\mu,\mathcal{P}} = \left\{ \text{vector fields } h = \nabla f \text{ completed in the norm } \int |h|^2 d\mu \right\}$$

where the tangent h to a family $t \mapsto \mu(t)$ is defined by the identity:

$$\frac{\partial \mu}{\partial t} + \operatorname{div}(h.\mu) = 0.$$

In our case, we want to assign to an immersion c the scaled arc length measure $\mu_c = ds/\ell$. This maps B_i to $\mathcal{P}$. The claim is that the pull-back of the Wasserstein metric by this map is intermediate between $G^{\ell^{-1}}$ and G^{Φ_W}, where

$$\Phi_W(\ell,\kappa) = \ell^{-1} + \tfrac{1}{12}\ell\kappa^2.$$

This is not hard to work out.

(1) Because we are mod-ing out by vector fields of norm 0, the vector field h is defined only along the curve c and its norm is $\ell^{-1}.\int \|h\|^2 ds$.

(2) If we split $h = av + bn$, then the condition that $h = \nabla f$ means that $\int a.ds = 0$ and the norm is $\ell^{-1}.\int (a^2 + b^2) ds$.

(3) But moving c infinitesimally by h, scaled arc length parametrization of c must still be scaled arc length. Let $c(\cdot,t) = c + th$. Then this means $|c_\theta|_t = \text{cnst.}\,|c_\theta|$ at $t = 0$. Since $|c_\theta|_t = \langle c_{t\theta}, c_\theta\rangle/|c_\theta|$, this condition is the same as $\langle D_s(av + bn), v\rangle = \text{cnst.}$, or $D_s a - b\kappa_c = \text{cnst.}$.

(4) Combining the last 2 conditions on b, we get a formula for a in terms of b, namely $a = K * (b\kappa_c)$, where we convolve with respect to arc length using the kernel $K(x) = \text{sign}(x)/2 - x/\ell, -\ell \le x \le \ell$.

(5) Finally, since $|K * f|(x) \le |K|.|f| = \sqrt{\ell/12}|f|$ for all f, it follows that

$$\ell^{-1}.\int b^2 ds \le \ell^{-1}.\int (a^2 + b^2)ds \le \ell^{-1}.\int (b^2 + \frac{(\ell\kappa)^2}{12}.b^2)ds$$

which sandwiches the Wasserstein norm between $G^{\ell^{-1}}$ and G^{Φ_W} for $\Phi_W = \ell^{-1}.(1 + (\ell\kappa)^2/12)$.

17. Immersion-Sobolev metrics

17.1. The $G^{\text{imm},n}$-metric. We note first that the differential operator $D_s = \frac{\partial_\theta}{|c_\theta|}$ is anti self-adjoint for the metric G^0, i.e., for all $h, k \in C^\infty(S^1.\mathbb{R}^2)$ we have

$$\int_{S^1} \Big\langle D_s(h), k\Big\rangle ds = \int_{S^1} \Big\langle h, -D_s(k)\Big\rangle ds$$

We can define a Sobolev-type weak Riemannian metric[1] on $\text{Imm}(S^1, \mathbb{R}^2)$ which is invariant under the action of $\text{Diff}(S^1)$ by:

$$G_c^{\text{imm},n}(h,k) = \int_{S^1} \left(\langle h,k\rangle + A.\langle D_s^n h, D_s^n k\rangle\right).ds \tag{1}$$

$$= \int_{S^1} \langle L_n(h), k\rangle ds \qquad \text{where}$$

$$L_n(h) \text{ or } L_{n,c}(h) = (I + (-1)^n A.D_s^{2n})(h) \tag{2}$$

The interesting special case $n = 1$ and $A \to \infty$ has been studied by Trouvé and Younes in [**74, 79**] and by Mio, Srivastava and Joshi in [**59, 60**]. In this case, the metric reduces to:

$$G_c^{\text{imm},1,\infty}(h,k) = \int_{S^1} \langle D_s(h), D_s(k)\rangle.ds$$

which ignores translations, i.e. it is a metric on $\text{Imm}(S^1, \mathbb{R}^2)$ modulo translations. Now identify $\mathbb{R}^2$ with $\mathbb{C}$, so that this space embeds as follows:

$$\text{Imm}(S^1, \mathbb{R}^2)/\text{transl.} \hookrightarrow C^\infty(S^1, \mathbb{C})$$
$$c \longmapsto c_\theta.$$

[1] There are other choices for the higher order terms, e.g. summing all the intermediate derivatives with or without binomial coefficients. These metrics are all equivalent and the one we use leads to the simplest equations.

Then Trouvé and Younes use the new shape space coordinates $Z(\theta) = \sqrt{c_\theta(\theta)}$ and Mio et al use the coordinates $\Phi(\theta) = \log(c_\theta(\theta))$ – with *complex* square roots and logs. Both of these unfortunately require the introduction of a discontinuity, but this will drop out when you minimize path length with respect to reparametrizations. The wonderful fact about $Z(\theta)$ is that in a family $Z(t,\theta)$, we find:

$$Z_t = \frac{c_{t,\theta}}{2\sqrt{c_\theta}}, \quad \text{so} \int_{S^1} |Z_t|^2 d\theta = \tfrac{1}{4}\int \frac{|c_{t,\theta}|^2}{|c_\theta|^2}|c_\theta| d\theta = \tfrac{1}{4}\int |D_s(c_t)|^2 ds$$

so the metric becomes a *constant* metric on the vector space of functions Z. With Φ, one has $\int |\Phi_t|^2 ds = \int |D_s(c_t)|^2 ds$, which is simple but not quite so nice. One can expect a very explicit representation of the space of curves in this metric.

Returning to the general case, for each fixed c of length ℓ, the differential operator $L_{n,c}$ is simply the constant coefficient ordinary differential operator $f \mapsto f + (-1)^n A.f^{(2n)}$ on the s-line modulo $\ell.\mathbb{Z}$. Thus its Green's function is a linear combination of the exponentials $\exp(\lambda.x)$, where λ are the roots of $1 + (-1)^n A.\lambda^{2n} = 0$. A simple verification gives its Green's function (which we will not use below):

$$F_n(x) = \frac{1}{2n} \cdot \sum_{\lambda^{2n} = (-1)^{n+1}/A} \frac{\lambda}{1 - e^{\lambda \ell}} e^{\lambda x}, \quad 0 \le x \le \ell.$$

This means that the dual metric $\check{G}_c^{\mathrm{imm},n} = (G_c^{\mathrm{imm},n})^{-1}$ on the *smooth cotangent space* $C^\infty(S^1,\mathbb{R}^2) \cong G_c^0(T_c \operatorname{Imm}(S^1,\mathbb{R}^2)) \subset T_c^* \operatorname{Imm}(S^1,\mathbb{R}^2) \cong \mathcal{D}(S^1)^2$ is given by the integral operator L^{-1} which is convolution by F_n with respect to arc length s:

$$\check{G}_c^{\mathrm{imm},n}(h,k) = \iint_{S^1 \times S^1} F_n(s_1 - s_2).\langle h(s_1), k(s_2)\rangle.ds_1.ds_2.$$

17.2. Geodesics in the $G^{\mathrm{imm},n}$-metric. Differentiating the operator $D_s = \frac{1}{|c_\theta|}\partial_\theta$ with respect to c in the direction m we get $-\frac{\langle m_\theta, c_\theta\rangle}{|c_\theta|^3}\partial_\theta$, or $-\langle D_s m, v\rangle D_s$. Thus differentiating the big operator $L_{n,c}$ with respect to c in the direction m, we get:

$$D_{(c,m)} L_{n,c}(h) = (-1)^{n+1} A. \sum_{j=0}^{2n-1} D_s^j \langle D_s(m), v\rangle D_s^{2n-j}(h) \tag{3}$$

Thus we have

$$D_{(c,m)} G_c^{\mathrm{imm},n}(h,k) =$$

$$= A.\int_{S^1} (-1)^{n+1} \sum_{j=0}^{2n-1} \Big\langle D_s^j \langle D_s m, v\rangle D_s^{2n-j}(h), k\Big\rangle ds + \int_{S^1} \langle L_n(h), k\rangle \langle D_s m, v\rangle ds$$

$$= A.\int_{S^1}\sum_{j=1}^{2n-1}(-1)^{n+j+1}\Big\langle\langle D_s m, v\rangle D_s^{2n-j}(h), D_s^j k\Big\rangle ds + \int_{S^1}\langle h,k\rangle\langle D_s m, v\rangle ds$$

$$= \int_{S^1}\Big\langle m, A.\sum_{j=1}^{2n-1}(-1)^{n+j}D_s\Big(\langle D_s^{2n-j}h, D_s^j k\rangle v\Big) - D_s(\langle h,k\rangle v)\Big\rangle ds$$

According to 15.1 we should rewrite this as

$$D_{(c,m)}G_c^{\text{imm},n}(h,k) = G_c^{\text{imm},n}(K_c^n(m,h),k) = G_c^{\text{imm},n}\big(m, H_c^n(h,k)\big),$$

and thus we find the two versions K^n and H^n of the G^n-gradient of $c \mapsto G_c^{\text{imm},n}(h,k)$ are given by:

$$(4)\quad K_c^n(m,h) = L_n^{-1}\Big((-1)^{n+1}A.\sum_{j=1}^{2n-1}D_s^j\langle D_s m, v\rangle D_s^{2n-j}(h) + \langle D_s m, v\rangle h\Big)$$

and by

$$H_c^n(h,k) = L_n^{-1}\Big(A.\sum_{j=1}^{2n-1}(-1)^{n+j}D_s\Big(\langle D_s^{2n-j}h, D_s^j k\rangle v\Big) - D_s(\langle h,k\rangle v)\Big)$$

$$= L_n^{-1}\Big(A.\sum_{j=1}^{2n-1}(-1)^{n+j}\langle D_s^{2n-j+1}h, D_s^j k\rangle v + A.\sum_{j=2}^{2i}(-1)^{n+j-1}\langle D_s^{2n-j+1}h, D_s^j k\rangle v$$

$$+ A.\sum_{j=1}^{2n-1}(-1)^{n+j}\langle D_s^{2n-j}h, D_s^j k\rangle\kappa_c n - \langle D_s h, k\rangle v - \langle h, D_s k\rangle v - \langle h,k\rangle\kappa_c n\Big)$$

$$= L_n^{-1}\Big(- \langle L_n(h), D_s k\rangle v - \langle D_s h, L_n(k)\rangle v - \langle h,k\rangle\kappa(c) n$$

$$(5)\qquad + A.\sum_{j=1}^{2n-1}(-1)^{n+j}\langle D_s^{2n-j}h, D_s^j k\rangle\kappa(c) n\Big)$$

since $D_s\big(v\big) = \kappa(c)n$. By 15.4 the geodesic equation for the metric G^n is

$$c_{tt} = \tfrac{1}{2}H_c^n(c_t,c_t) - K_c^n(c_t,c_t).$$

We expand it to get:

$$\boxed{\begin{aligned} L_n(c_{tt}) &= -\langle L_n(c_t), D_s(c_t)\rangle v - \frac{|c_t|^2\kappa(c)}{2}n - \langle D_s(c_t), v\rangle c_t \\ &\quad + \frac{A}{2}.\sum_{j=1}^{2n-1}(-1)^{n+j}\langle D_s^{2n-j}c_t, D_s^j c_t\rangle\kappa(c)n \\ &\quad + (-1)^n A.\sum_{j=1}^{2n-1} D_s^j\big(\langle D_s(c_t), v\rangle D_s^{2n-j}(c_t)\big)\end{aligned}} \tag{6}$$

From (3) we see that

$$(L_n(c_t))_t - L_n(c_{tt}) = dL_n(c)(c_t)(c_t) = (-1)^{n+1}A.\sum_{j=0}^{2n-1} D_s^j\langle D_s(c_t), v\rangle D_s^{2n-j}(c_t).$$

so that a more compact form of the geodesic equation of the metric G^n is:

$$\boxed{\begin{aligned} (L_n(c_t))_t &= -\langle L_n(c_t), D_s(c_t)\rangle v - \frac{|c_t|^2\kappa(c)}{2}n - \langle D_s(c_t), v\rangle L_n c_t \\ &\quad + \frac{A}{2}.\sum_{j=1}^{2n-1}(-1)^{n+j}\langle D_s^{2n-j}c_t, D_s^j c_t\rangle\kappa(c)n\end{aligned}} \tag{7}$$

For $n = 0$ this agrees with [**54**], 4.1.2.

17.3. Existence of geodesics. Theorem. *Let $n \geq 1$. For each $k \geq 2n+1$ the geodesic equation 17.2* (6) *has unique local solutions in the Sobolev space of H^k-immersions. The solutions depend C^∞ on t and on the initial conditions $c(0,\ .\ )$ and $c_t(0,\ .\ )$. The domain of existence (in t) is uniform in k and thus this also holds in* $\mathrm{Imm}(S^1,\mathbb{R}^2)$.

Proof. We consider the geodesic equation as the flow equation of a smooth (C^∞) vector field on the H^2-open set $U^k \times H^k(S^1,\mathbb{R}^2)$ in the Sobolev space $H^k(S^1,\mathbb{R}^2)\times H^k(S^1,\mathbb{R}^2)$ where $U^k = \{c \in H^k : |c_\theta| > 0\} \subset H^k$ is H^2-open. To see that this works we will use the following facts: By the Sobolev inequality we have a bounded linear embedding $H^k(S^1,\mathbb{R}^2) \subset C^m(S^1,\mathbb{R}^2)$ if $k > m + \frac{1}{2}$. The Sobolev space $H^k(S^1,\mathbb{R})$ is a Banach algebra under pointwise multiplication if $k > \frac{1}{2}$. For any fixed smooth mapping f the mapping $u \mapsto f\circ u$ is smooth $H^k \to H^k$ if $k > 0$. The mapping $(c,u) \mapsto L_{n,c}u$ is smooth $U \times H^k \to H^{k-2n}$ and is a bibounded linear isomorphism $H^k \to H^{k-2n}$ for fixed c. This can be seen as follows (see 17.6 below): It is true if c is parametrized by arclength (look at it in the space of Fourier coefficients). The index is invariant under continuous deformations of elliptic operators of fixed degree, so the index of $L_{n,c}$ is zero in general. But $L_{n,c}$ is self-adjoint positive, so it is injective with vanishing index, thus surjective. By the open mapping theorem it is then bibounded. Moreover $(c,w) \mapsto L_{n,c}^{-1}(w)$ is smooth $U^k \times H^{k-2n} \to H^k$ (by the inverse function theorem on Banach spaces). The mapping $(c,f) \mapsto D_s f = \frac{1}{|c_\theta|}\partial_\theta f$ is smooth $H^k \times H^m \supset U \times H^m \to H^{m-1}$ for $k \geq m$, and is linear in f. Let us write $D_c f = D_s f$ just for the remainder of this proof to stress the dependence on c. We have $v = D_c c$ and $n = J D_c c$. The mapping $c \mapsto \kappa(c)$ is smooth on the H^2-open set $\{c : |c_\theta| > 0\} \subset H^k$ into H^{k-2}. Keeping all this in mind we now write the geodesic equation as follows:

$$
\begin{aligned}
c_t &= u =: X_1(c,u) \\
u_t &= L_{n,c}^{-1}\Big(- \langle L_{n,c}(u), D_c(u)\rangle D_c(c) - \frac{|c_t|^2\kappa(c)}{2} J D_c(c) - \langle D_c(u), D_c c\rangle u \\
&\qquad + \frac{A}{2}. \sum_{j=1}^{2n-1} (-1)^{n+j} \langle D_c^{2n-j} u, D_c^j u\rangle \kappa(c) J D_c(c) \\
&\qquad + (-1)^n A. \sum_{j=1}^{2n-1} D_c^j \left(\langle D_c(u), D_c(c)\rangle D_c^{2n-j}(u)\right) \Big) \\
&=: X_2(c,u)
\end{aligned}
$$

Now a term by term investigation of this shows that the expression in the brackets is smooth $U^k \times H^k \to H^{k-2n}$ since $k - 2n \geq 1 > \frac{1}{2}$. The operator $L_{n,c}^{-1}$ then takes it smoothly back to H^k. So the vector field $X = (X_1, X_2)$ is smooth on $U^k \times H^k$. Thus the flow Fl^k exists on H^k and is smooth in t and the initial conditions for fixed k.

Now we consider smooth initial conditions $c_0 = c(0,\ .\ )$ and $u_0 = c_t(0,\ .\ ) = u(0,\ .\ )$ in $C^\infty(S^1, \mathbb{R}^2)$. Suppose the trajectory $\mathrm{Fl}_t^k(c_0, u_0)$ of X through these intial conditions in H^k maximally exists for $t \in (-a_k, b_k)$, and the trajectory $\mathrm{Fl}_t^{k+1}(c_0, u_0)$ in H^{k+1} maximally exists for $t \in (-a_{k+1}, b_{k+1})$ with $b_{k+1} < b_k$. By uniqueness we have $\mathrm{Fl}_t^{k+1}(c_0,u_0) = \mathrm{Fl}_t^k(c_0,u_0)$ for $t \in (-a_{k+1,} b_{k+1})$. We now apply ∂_θ to the equation $u_t = X_2(c,u) = L_{n,c}^{-1}(\dots)$, note that the commutator $[\partial_\theta, L_{n,c}^{-1}]$ is a pseudo differential operator of order $-2n$ again, and write $w = \partial_\theta u$. We obtain $w_t = \partial_\theta u_t = L_{n,c}^{-1}\partial_\theta(\dots) + [\partial_\theta, L_{n,c}^{-1}](\dots)$. In the term $\partial_\theta(\dots)$ we consider now only the terms $\partial_\theta^{2n+1} u$ and rename them $\partial_\theta^{2n} w$. Then we get an equation $w_t(t,\theta) = \tilde{X}_2(t, w(t,\theta))$ which is inhomogeneous bounded linear in $w \in H^k$ with coefficients bounded linear operators on H^k which are C^∞ functions of $c, u \in H^k$. These we already know on the intervall $(-a_k, b_k)$. This equation therefore has a solution $w(t,\ .\ )$ for all t for which the coefficients exists, thus for all $t \in (a_k, b_k)$. The limit $\lim_{t \nearrow b_{k+1}} w(t,\ .\ )$ exists in H^k and by continuity it equals $\partial_\theta u$ in H^k at $t = b_{k+1}$. Thus the H^{k+1}-flow was not maximal and can be continued.

So $(-a_{k+1}, b_{k+1}) = (a_k, b_k)$. We can iterate this and conclude that the flow of X exists in $\bigcap_{m\geq k} H^m = C^\infty$. $\square$

17.4. The conserved momenta of $G^{\text{imm},n}$. According to 15.5 the following momenta are preserved along any geodesic $t \mapsto c(t,\ .\ )$:

$\langle c_\theta, L_{n,c}(c_t)\rangle \lvert c_\theta(\theta)\rvert \ \in \mathfrak{X}(S^1)$	reparametrization momentum
$\int_{S^1} L_{n,c}(c_t)\, ds = \int_{S^1} c_t\, ds \ \in \mathbb{R}^2$	linear momentum
$\int_{S^1} \langle Jc, L_{n,c}(c_t)\rangle\, ds \in \mathbb{R}$	angular momentum

17.5. Horizontality for $G^{\text{imm},n}$. $h \in T_c \operatorname{Imm}(S^1, \mathbb{R}^2)$ is $G_c^{\text{imm},n}$-orthogonal to the $\operatorname{Diff}(S^1)$-orbit through c if and only if

$$0 = G_c^{\text{imm},n}(h, \zeta_X(c)) = G_c^{\text{imm},n}(h, c_\theta.X) = \int_{S^1} X.\langle L_{n,c}(h), c_\theta\rangle\, ds$$

for all $X \in \mathfrak{X}(S^1)$. So the $G^{\text{imm},n}$-normal bundle is given by

$$\mathcal{N}_c^n = \{h \in C^\infty(S, \mathbb{R}^2) : \langle L_{n,c}(h), v\rangle = 0\}.$$

The $G^{\text{imm},n}$-orthonormal projection $T_c \operatorname{Imm} \to \mathcal{N}_c^n$, denoted by $h \mapsto h^\perp = h^{\perp, G^n}$ and the complementary projection $h \mapsto h^\top \in T_c(c \circ \operatorname{Diff}(S^1))$ are determined as follows:

$$h^\top = X(h).v \qquad \text{where } \langle L_{n,c}(h), v\rangle = \langle L_{n,c}(X(h).v), v\rangle$$

Thus we are led to consider the linear differential operators associated to $L_{n.c}$

$$\begin{aligned}
&L_c^\top, L_c^\perp : C^\infty(S^1) \to C^\infty(S^1),\\
&L_c^\top(f) = \langle L_{n,c}(f.v), v\rangle = \langle L_{n,c}(f.n), n\rangle,\\
&L_c^\perp(f) = \langle L_{n,c}(f.v), n\rangle = -\langle L_{n,c}(f.n), v\rangle.
\end{aligned}$$

The operator $L_c^\top$ is of order $2n$ and also unbounded, self-adjoint and positive on $L^2(S^1, \lvert c_\theta\rvert\, d\theta)$ since

$$\begin{aligned}
\int_{S^1} L_c^\top(f) g ds &= \int_{S^1} \langle L_{n,c}(fv), v\rangle g ds\\
&= \int_{S^1} \langle fv, L_{n,c}(gv)\rangle ds = \int_{S^1} f L_c^\top(g) ds,\\
\int_{S^1} L_c^\top(f) f ds &= \int_{S^1} \langle fv, L_{n,c}(fv)\rangle ds \quad > 0 \text{ if } f \neq 0.
\end{aligned}$$

In particular, $L_c^\top$ is injective. $L_c^\perp$, on the other hand is of order $2n-1$ and a similar argument shows it is skew-adjoint. For example, if $n=1$, then one finds that:

$$\begin{aligned} L_c^\top &= -A.D_s^2 + (1+A.\kappa^2).I \\ L_c^\perp &= -2A.\kappa.D_s - A.D_s(\kappa).I \end{aligned}$$

17.6. Lemma. *The operator $L_c^\top : C^\infty(S^1) \to C^\infty(S^1)$ is invertible.*

Proof. This is because its index vanishes, by the following argument: The index is invariant under continuous deformations of elliptic operators of degree $2n$. The operator

$$L_c^\top(f) = (-1)^n \frac{A}{|c_\theta|^{2n}} \partial_\theta^{2n}(f) + \text{ lower order terms}$$

is homotopic to $(1+(-1)^n\partial_\theta^{2n})(f)$ and thus has the same index which is zero since the operator $1+(-1)^n\partial_\theta^{2n}$ is invertible. This can be seen by expanding in Fourier series where the latter operator is given by $(\hat{f}(m)) \mapsto ((1+m^{2n})\hat{f}(m))$, a linear isomorphism of the space of rapidly decreasing sequences. Since $L_c^\top$ is injective, it is also surjective. $\square$

To go back and forth between the 'natural' horizontal space of vector fields $a.n$ and the $G^{\text{imm},n}$-horizontal vector fields $\{h \mid \langle Lh, v\rangle = 0\}$, we only need to use these operators and the inverse of $L^\top$. Thus, given a, we want to find b and f such that $L(an+bv) = fn$, so that $an+bv$ is $G^{\text{imm},n}$-horizontal. But this implies that

$$L^\perp(a) = -\langle L(an), v\rangle = \langle L(bv), v\rangle = L^\top(b).$$

Thus if we define the operator $C_c : C^\infty(S^1) \to C^\infty(S^1)$ by

$$C_c := (L_c^\top)^{-1} \circ L_c^\perp,$$

we get a pseudo-differential operator of order -1 (which is an integral operator), so that $a.n + C(a).v$ is always $G^{\text{imm},n}$-horizontal. In particular, the restriction of the metric $G^{\text{imm},n}$ to horizontal vector fields $h_i = a_i.n + b_i.v$ can be computed like this:

$$\begin{aligned} G_c^{\text{imm},n}(h_1,h_2) &= \int_{S^1} \langle Lh_1, h_2\rangle.ds \\ &= \int_{S^1} \langle L(a_1.n + b_1.v), n\rangle.a_2.ds \\ &= \int_{S^1} \Big(L^\top(a_1) + L^\perp(b_1)\Big).a_2.ds \\ &= \int_{S^1} \Big(L^\top + L^\perp \circ C\Big)\, a_1.a_2.ds. \end{aligned}$$

Thus the metric restricted to horizontal vector fields is given by the pseudo differential operator $L^{\mathrm{red}} = L^\top + L^\perp \circ (L^\top)^{-1} \circ L^\perp$. On the quotient space B_i, if we identify its tangent space at C with the space of normal vector fields $a.n$, then:

$$\boxed{G_C^{imm,n}(a_1, a_2) = \int_C (L^\top + L^\perp \circ (L^\top)^{-1} \circ L^\perp) a_1 \cdot a_2 \cdot ds}$$

Now, although this operator may be hard to analyze, its inverse, the metric on the cotangent space to B_i, is simple. The tangent space to B_i at a curve C is canonically the quotient of that of $\mathrm{Imm}(S^1, \mathbb{R}^2)$ at a parametrization c of C, modulo the subspace of multiples of v. Hence the cotangent space to B_i at C injects into that of $\mathrm{Imm}(S^1, \mathbb{R}^2)$ at c with image the linear functionals that vanish on v. In terms of the dual basis $\check{v}, \check{n}$, these are multiples of $\check{n}$. On the smooth cotangent space $C^\infty(S^1, \mathbb{R}^2) \cong G_c^0(T_c \mathrm{Imm}(S^1, \mathbb{R}^2)) \subset T_c^* \mathrm{Imm}(S^1, \mathbb{R}^2) \cong \mathcal{D}(S^1)^2$ the dual metric is given by convolution with the elementary kernel K_n which is a simple sum of exponentials. Thus we need only restrict this kernel to multiples $a(s).\check{n}_c(s)$ to obtain the dual metric on B_i. The result is that:

$$\check{G}_c^n(a_1, a_2) = \iint_{S^1 \times S^1} K_n(s_1 - s_2).\langle n_c(s_1), n_c(s_2)\rangle.a_1(s_1).a_2(s_2).ds_1 ds_2.$$

17.7. Horizontal geodesics. The normal bundle $\mathcal{N}_c$ mentioned in 17.5 is well defined and is a smooth vector subbundle of the tangent bundle. But $\mathrm{Imm}(S^1, \mathbb{R}^2) \to B_i(S^1, \mathbb{R}^2) = \mathrm{Imm} / \mathrm{Diff}(S^1)$ is *not* a principal bundle and thus there are no principal connections, but we can prove the main consequence, the existence of horizontal paths, directly:

Proposition. *For any smooth path c in $\mathrm{Imm}(S^1, \mathbb{R}^2)$ there exists a smooth path φ in $\mathrm{Diff}(S^1)$ with $\varphi(0, \ .\ ) = \mathrm{Id}_{S^1}$ depending smoothly on c such that the path e given by $e(t, \theta) = c(t, \varphi(t, \theta))$ is horizontal: $\langle L_{n,e}(e_t), e_\theta \rangle = 0$.*

Proof. Writing D_c instead of D_s we note that $D_{c\circ\varphi}(f \circ \varphi) = \frac{(f_\theta \circ \varphi)\varphi_\theta}{|c_\theta \circ \varphi|.|\varphi_\theta|} = (D_c(f)) \circ \varphi$ for $\varphi \in \mathrm{Diff}^+(S^1)$. So we have $L_{n,c\circ\varphi}(f \circ \varphi) = (L_{n,c} f) \circ \varphi$.
Let us write $e = c \circ \varphi$ for $e(t, \theta) = c(t, \varphi(t, \theta))$, etc. We look for φ as the integral curve of a time dependent vector field $\xi(t, \theta)$ on S^1, given by $\varphi_t = \xi \circ \varphi$. We want the following expression to vanish:

$$\begin{aligned}
\langle L_{n,c\circ\varphi}(\partial_t(c \circ \varphi)), \partial_\theta(c \circ \varphi)\rangle &= \langle L_{n,c\circ\varphi}(c_t \circ \varphi + (c_\theta \circ \varphi)\, \varphi_t), (c_\theta \circ \varphi)\, \varphi_\theta \rangle \\
&= \langle L_{n,c}(c_t) \circ \varphi + L_{n,c}(c_\theta.\xi) \circ \varphi, c_\theta \circ \varphi \rangle \varphi_\theta \\
&= \big((\langle L_{n,c}(c_t), c_\theta\rangle + \langle L_{n,c}(\xi.c_\theta), c_\theta\rangle) \circ \varphi \big)\, \varphi_\theta.
\end{aligned}$$

Using the time dependent vector field $\xi = -\frac{1}{|c_\theta|}(L_c^\top)^{-1}(\langle L_{n,c}(c_t), v\rangle)$ and its flow φ achieves this. $\square$

If we write

$$c_t = na + vb = \big(n, v\big)\begin{pmatrix} a \\ b \end{pmatrix}$$

then we can expand the condition for horizontality as follows:

$$\begin{aligned}
D_s(c_t) &= \big(D_s a + \kappa(c)b\big)n + \big(D_s b - \kappa(c)a\big)v. \\
&= (n, v)\begin{pmatrix} D_s & \kappa \\ -\kappa & D_s \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \\
L_n^c(c_t) &= c_t + (-1)^n A(n, v)\begin{pmatrix} D_s & \kappa \\ -\kappa & D_s \end{pmatrix}^{2n}\begin{pmatrix} a \\ b \end{pmatrix} \\
&= c_t + (-1)^n A(n, v)\begin{pmatrix} D_s^2 - \kappa^2 & D_s\kappa + \kappa D_s \\ -D_s\kappa - \kappa D_s & D_s^2 - \kappa^2 \end{pmatrix}^n\begin{pmatrix} a \\ b \end{pmatrix}
\end{aligned}$$

so that horizontality becomes

$$0 = \langle L_{n,c}(c_t), v\rangle = \langle c_t, v\rangle + (-1)^n A(0,1)\begin{pmatrix} D_s^2 - \kappa^2 & D_s\kappa + \kappa D_s \\ -D_s\kappa - \kappa D_s & D_s^2 - \kappa^2 \end{pmatrix}^n\begin{pmatrix} a \\ b \end{pmatrix}$$

We may specialize the general geodesic equation to horizontal paths and then take the v and n parts of the geodesic equation. For a horizontal path we may write $L_{n,c}(c_t) = \tilde{a}n$ for $\tilde{a}(t,\theta) = \langle L_{n,c}(c_t), n\rangle$. The v part of the equation turns out to vanish identically and then n part gives us (because n_t is a multiple of v):

$$\boxed{\tilde{a}_t = -\frac{|c_t|^2\kappa(c)}{2} - \langle D_s c_t, v\rangle\tilde{a} + \frac{A\kappa(c)}{2}\sum_{j=1}^{2n-1}(-1)^{n+j}\langle D_s^{2n-j}c_t, D_s^j c_t\rangle}$$

Note that applying 17.3 with horizontal initial vectors gives us local existence and uniqueness for solutions of this horizontal geodesic equation.

17.8. A Lipschitz bound for arclength in $G^{\mathrm{imm},n}$. We apply the inequality of Cauchy-Schwarz to the derivative of the length function $\ell(c) = \int |c_\theta| d\theta$ along a path $t \mapsto c(t,\ .\ )$:

$$\begin{aligned}
\partial_t \ell(c) = d\ell(c)(c_t) &= \int_{S^1}\frac{\langle c_{t\theta}, c_\theta\rangle}{|c_\theta|}\,d\theta = \int_{S^1}\langle D_s(c_t), v\rangle\,ds \\
&\le \Big(\int_{S^1}|D_s(c_t)|^2 ds\Big)^{\frac12}\cdot\Big(\int_{S^1}1^2 ds\Big)^{\frac12} \le \sqrt{\ell(c)}\frac{1}{A}\|c_t\|_{G^1}, \\
&\le \sqrt{\ell(c)}C(A,n)\|c_t\|_{G^n}, \\
\partial_t\sqrt{\ell(c)} = \frac{\partial_t\ell(c)}{2\sqrt{\ell(c)}} &\le \frac{C(A,n)}{2}\|c_t\|_{G^n}.
\end{aligned}$$

Thus we get

$$\begin{aligned}|\sqrt{\ell(c(1,\ .\ ))}-\sqrt{\ell(c(0,\ .\ ))}| &\le \int_0^1 |\partial_t\sqrt{\ell(c)}|\,dt \le \frac{C(A,n)}{2}\int_0^1 \|c_t\|_{G^n}\,dt\\ &= \frac{C(A,n)}{2}L_{G^n}(c).\end{aligned}$$

Taking the infinimum of this over all paths $t\mapsto c(t,\ .\ )$ from c_0 to c_1 we see that for $n\ge 1$ we have the Lipschitz estimate:

$$|\sqrt{\ell(c_1)}-\sqrt{\ell(c_0)}| \le \frac{1}{2}\operatorname{dist}^{\mathrm{Imm}}_{G^n}(c_1,c_0)$$

Since we have $L^{\mathrm{hor}}_{G^n}(c)\le L_{G^n}(c)$ with equality for horizontal curves we also have:

$$\boxed{\textit{If } n\ge 1,\quad |\sqrt{\ell(C_1)}-\sqrt{\ell(C_0)}| \le \frac{1}{2}\operatorname{dist}^{B_i}_{G^n}(C_1,C_0)}$$

17.9. Scale invariant immersion Sobolev metrics. Let us mention in passing that we may use the length of the curve to modify the immersion Sobolev metric so that it becomes scale invariant:

$$\begin{aligned}G^{\mathrm{imm,scal},n}_c(h,k) &= \int_{S^1}\big(\ell(c)^{-3}\langle h,k\rangle + \ell(c)^{2n-3}A\langle D^n_s(h),D^n_s(k)\rangle\big)\,ds\\ &= \int_{S^1}\big\langle(\ell(c)^{-3}+(-1)^n\ell(c)^{2n-3}AD^{2n}_s)h,k\big\rangle\,ds\end{aligned}$$

This metric can easily be analyzed using the methods described above. In particular we note that the geodesic equation on $\mathrm{Imm}(S^1,\mathbb{R}^2)$ for this metric is built in a similar way than that for $G^{\mathrm{imm},n}$ and that the existence theorem in 17.3 holds for it. Note the conserved momenta along a geodesic $t\mapsto c(t,\ .\ )$ are:

$$\begin{aligned}&\frac{1}{\ell(c)^3}\int_{S^1} c_t\,ds + (-1)^n\ell(c)^{2n-3}A\int_{S^1} D^{2n}_s(c_t)\,ds\\ &\quad = \frac{1}{\ell(c)^3}\int_{S^1} c_t\,ds \in\mathbb{R}^2 && \text{linear momentum}\\ &\frac{1}{\ell(c)^3}\int_{S^1}\langle Jc,c_t\rangle\,ds + (-1)^n\ell(c)^{2n-3}A\int_{S^1}\langle Jc,D^{2n}_s(c_t)\rangle\,ds && \text{angular momentum}\\ &\frac{1}{\ell(c)^3}\int_{S^1}\langle c,c_t\rangle\,ds + (-1)^n\ell(c)^{2n-3}A\int_{S^1}\langle c,D^{2n}_s(c_t)\rangle\,ds && \text{scaling momentum}\end{aligned}$$

As in the work of Trouvé and Younes [**74, 79**], we may consider the following variant.

$$\begin{aligned}G^{\mathrm{imm,scal},n,\infty}_c(h,k) &= \lim_{A\to\infty}\frac{1}{A}\int_{S^1}\big\langle(\ell(c)^{-3}+(-1)^n\ell(c)^{2n-3}AD^{2n}_s)h,k\big\rangle\,ds\\ &= (-1)^n\ell(c)^{2n-3}\int_{S^1}\big\langle D^{2n}_s h,k\big\rangle\,ds\end{aligned}$$

It is degenerate with kernel the constant tangent vectors. The interesting fact is that the scaling momentum for $G^{\text{imm,scal},1,\infty}$ is given by

$$-\frac{1}{\ell(c)}\int_{S^1}\langle c, D_s^2(c_t)\rangle ds = \partial_t \log \ell(c).$$

18. Sobolev metrics on $\mathrm{Diff}(\mathbb{R}^2)$ and on its quotients

18.1. The metric on $\mathrm{Diff}(\mathbb{R}^2)$. We consider the regular Lie group $\mathrm{Diff}(\mathbb{R}^2)$ which is either the group $\mathrm{Diff}_c(\mathbb{R}^2)$ of all diffeomorphisms with compact supports of $\mathbb{R}^2$ or the group $\mathrm{Diff}_{\mathcal{S}}(\mathbb{R}^2)$ of all diffeomorphisms which decrease rapidly to the identity. The Lie algebra is $\mathfrak{X}(\mathbb{R}^2)$, by which we denote either the Lie algebra $\mathfrak{X}_c(\mathbb{R}^2)$ of vector fields with compact support or the Lie algebra $\mathfrak{X}_{\mathcal{S}}(\mathbb{R}^2)$ of rapidly decreasing vector fields, with the negative of the usual Lie bracket. For any $n \geq 0$, we equip $\mathrm{Diff}(\mathbb{R}^2)$ with the right invariant weak Riemannian metric $G^{\mathrm{Diff},n}$given by the Sobolev H^n-inner product on $\mathfrak{X}_c(\mathbb{R}^2)$.

$$\begin{aligned}
H^n(X,Y) &= \sum_{\substack{0\leq i,j\leq n\\ i+j\leq n}} \frac{A^{i+j}n!}{i!j!(n-i-j)!}\int_{\mathbb{R}^2}\langle \partial^i_{x^1}\partial^j_{x^2}X, \partial^i_{x^1}\partial^j_{x^2}Y\rangle\,dx\\
&= \sum_{\substack{0\leq i,j\leq n\\ i+j\leq n}} (-A)^{i+j}\frac{n!}{i!j!(n-i-j)!}\int_{\mathbb{R}^2}\langle \partial^{2i}_{x^1}\partial^{2j}_{x^2}X, Y\rangle\,dx\\
&= \int_{\mathbb{R}^2}\langle LX,Y\rangle\,dx \qquad \text{where}\\
L = L_{A,n} &= (1-A\Delta)^n, \qquad \Delta = \partial^2_{x^1}+\partial^2_{x^2}.
\end{aligned}$$

(We will write out the full subscript of L only where it helps clarify the meaning.) The completion of $\mathfrak{X}_c(\mathbb{R}^2)$ is the Sobolev space $H^n(\mathbb{R}^2)^2$. With the usual L^2-inner product we can identify the dual of $H^n(\mathbb{R}^2)^2$ with $H^{-n}(\mathbb{R}^2)^2$ (in the space of tempered distributions). Note that the operator $L : H^n(\mathbb{R}^2)^2 \to H^{-n}(\mathbb{R}^2)^2$ is a bounded linear operator. On $L^2(\mathbb{R}^2)$ the operator L is unbounded selfadjoint and positive. In terms of Fourier transform we have $\widehat{L_{A,n}u}(\xi) = (1 + A|\xi|^2)^n\hat{u}$. Let $F_{A,n}$ in the space of tempered distributions $\mathcal{S}'(\mathbb{R}^2)$ be the fundamental solution (or Green's function: note that we use the letter 'F' for 'fundamental' because 'G' has been used as the metric) of $L_{A,n}$ satisfying $L_{A,n}(F_{A,n}) = \delta_0$ which is given by

$$F_{A,n}(x) = \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle}\frac{1}{(1+A|\xi|^2)^n}\,d\xi.$$

The functions $F_{A,n}$ are given by the classical modified Bessel functions K_r (in the notation, e.g., of Abramowitz and Stegun [**1**] or of Matlab) by the formula:

$$F_{A,n}(x) = \frac{1}{2^n \pi (n-1)! A} \cdot \left(\frac{|x|}{\sqrt{A}}\right)^{n-1} K_{n-1}\left(\frac{|x|}{\sqrt{A}}\right).$$

and it satisfies $(L^{-1}u)(x) = \int_{\mathbb{R}^2} F(x-y)u(y)\,dy$ for each tempered distribution u. The function $F_{A,n}$ is C^{n-1} except that $F_{A,1}$ has a log-pole at zero. At infinity, $F_{A,n}(x)$ is asymptotically a constant times $x^{n-3/2}e^{-x}$: these facts plus much much more can be found in [**1**].

18.2. Strong conservation of momentum and 'EPDiff'. What is the form of the conservation of momentum for a geodesic $\varphi(t)$ in this metric, that is to say, a flow $x \mapsto \varphi(x,t)$ on $\mathbb{R}^2$? We need to work out Ad^*_φ first. Using the definition, we see:

$$\begin{aligned}\int_{\mathbb{R}^2} \langle LX, \mathrm{Ad}^*_\varphi(Y)\rangle &:= \int_{\mathbb{R}^2} \langle L\,\mathrm{Ad}_\varphi(X), Y\rangle = \int_{\mathbb{R}^2} \langle (d\varphi.X)\circ\varphi^{-1}, LY\rangle \\ &= \int_{\mathbb{R}^2} \det(d\varphi)\langle d\varphi.X, LY\circ\varphi\rangle = \int_{\mathbb{R}^2} \langle X, \det(d\varphi).d\varphi^T.(LY\circ\varphi)\rangle\end{aligned}$$

hence:

$$\mathrm{Ad}^*_\varphi(Y) = L^{-1}\left(\det(d\varphi).d\varphi^T.(LY\circ\varphi)\right).$$

Now the conservation of momentum for geodesics $\varphi(t)$ of right invariant metrics on groups says that:

$$L^{-1}\left(\det(d\varphi)(t).d\varphi(t)^t.\left(L(\frac{\partial\varphi}{\partial t}\circ\varphi^{-1})\circ\varphi\right)\right)$$

is independent of t. This can be put in a much more transparent form. First, L doesn't depend on t, so we cross out the outer L^{-1}. Now let $v(t) = \frac{\partial\varphi}{\partial t}\circ\varphi^{-1} \in \mathfrak{X}(\mathbb{R}^2)$ be the tangent vector to the geodesic. Let $u(t) = Lv(t)$, so that:

$$\det(d\varphi)(t).d\varphi(t)^t.(u(t)\circ\varphi(t))$$

is independent of t. We should *not* think of $u(t)$ as a vector field on $\mathbb{R}^2$: this is because we want $\langle u, v\rangle$ to make invariant sense in any coordinates whatsoever. This means we should think of u as expanding to the differential form:

$$\omega(t) = (u_1.dx^1 + u_2.dx^2)\otimes\mu$$

where $\mu = dx^1\wedge dx^2$, the area form. But then:

$$\varphi(t)^*(\omega(t)) = \langle d\varphi^t.(u\circ\varphi(t)), dx\rangle \otimes \det(d\varphi)(t).\mu$$

so conservation of momentum says simply:

$$\boxed{\varphi(t)^*\omega(t) \text{ is independent of } t}$$

This motivates calling $\omega(t)$ the momentum of the geodesic flow. As we mentioned above, conservation of momentum for a Riemannian metric on a group is very strong and is an integrated form of the geodesic equation. To see this, we need only take the differential form of this conservation law. $v(t)$ is the infinitesimal flow, so the infinitesimal form of the conservation is:

$$\frac{\partial}{\partial t}\omega(t) + \mathcal{L}_{v(t)}(\omega(t)) = 0$$

where $\mathcal{L}_{v(t)}$ is the Lie derivative. We can expand this term by term:

$$\begin{aligned}
\mathcal{L}_{v(t)}(u_i) &= \sum_j v^j.\frac{\partial u_i}{\partial x^j} \\
\mathcal{L}_{v(t)}(dx^i) &= dv^i = \sum_j \frac{\partial v^i}{\partial x^j}.dx^j \\
\mathcal{L}_{v(t)}(\mu) &= \operatorname{div} v(t)\mu \\
\mathcal{L}_{v(t)}(\omega(t)) &= \left(\sum_{i,j}\left(v^j.\frac{\partial u_i}{\partial x^j}.dx^i + u^j.\frac{\partial v^j}{\partial x^i}.dx^i\right) + \operatorname{div} v.\sum_i u_i dx^i\right) \otimes \mu.
\end{aligned}$$

The resulting differential equation for geodesics has been named *EPDiff*:

$$\boxed{\begin{aligned}
&v = \frac{\partial \varphi}{\partial t} \circ \varphi^{-1}, \qquad u = L(v) \\
&\frac{\partial u_i}{\partial t} + \sum_j \left(v^j.\frac{\partial u_i}{\partial x^j} + u^j.\frac{\partial v^j}{\partial x^i}\right) + \operatorname{div} v.u_i = 0.
\end{aligned}}$$

Note that this is a special case of the general equation of Arnold: $\partial_t u = -\operatorname{ad}(u)^* u$ for geodesics on any Lie group in any right (or left) invariant metric. The name 'EPDiff' was coined by Holm and Marsden and stands for 'Euler-Poincaré', although it takes a leap of faith to see it in the reference they give to Poincaré.

18.3. The quotient metric on $\operatorname{Emb}(S^1,\mathbb{R}^2)$**.** We now consider the quotient mapping $\operatorname{Diff}(\mathbb{R}^2) \to \operatorname{Emb}(S^1,\mathbb{R}^2)$ given by $\varphi \mapsto \varphi \circ i$. Since this identifies $\operatorname{Emb}(S^1,\mathbb{R}^2)$ with a right coset space of $\operatorname{Diff}(\mathbb{R}^2)$, and since the metric G^n_{diff} is right invariant, we can put a quotient metric on $\operatorname{Emb}(S^1,\mathbb{R}^2)$ for which this map is a Riemannian submersion. Our next step is to identify this metric. Let $\varphi \in \operatorname{Diff}(\mathbb{R}^2)$ and let $c = \varphi \circ i \in \operatorname{Emb}(S^1,\mathbb{R}^2)$. The fibre of this map through φ is the coset

$$\varphi.\operatorname{Diff}^0(S^1,\mathbb{R}^2) = \{\psi \mid \psi \circ c \equiv c\}.\varphi.$$

whose tangent space is (the right translate by φ of) the vector space of vector fields $X \in \mathfrak{X}(\mathbb{R}^2)$ with $X \circ c \equiv 0$. This is the vertical subspace. Thus the horizontal subspace is

$$\left\{ Y \,\middle|\, \int_{\mathbb{R}^2} \langle LY, X\rangle dx = 0, \text{ if } X \circ c \equiv 0 \right\}.$$

If we want $Y \in \mathfrak{X}(\mathbb{R}^2)$ then the horizontal subspace is 0. But we can also search for Y in a bigger space of vector fields on $\mathbb{R}^2$. What we need is that $LY = c_*(p(\theta).ds)$, where p is a function from S^1 to $\mathbb{R}^2$ and ds is arc-length measure supported on C. To make $c_*(p(\theta).ds)$ pair with smooth vector fields $\mathfrak{X}(\mathbb{R}^2)$ in a coordinate invariant way, we should interpret the values of p as 1-forms. Solving for Y, we have:

$$Y(x) = \int_{S^1} F(x - c(\theta)).p(\theta)ds$$

(where, to make Y a vector field, the values of p are now interpreted as vectors, using the standard metric on $\mathbb{R}^2$ to convert 1-forms to vectors). Because F is not C^∞, we have a case here where the horizontal subspace is not given by C^∞ vector fields. However, we can still identify the set of vector fields in this horizontal subspace which map bijectively to the C^∞ tangent space to $\mathrm{Emb}(S^1, \mathbb{R}^2)$ at c. Mapped to $T_c \mathrm{Emb}(S^1, \mathbb{R}^2)$, the above Y goes to:

$$\begin{aligned}
(Y \circ c)(\theta) &= \int_{S^1} F(c(\theta) - c(\theta_1)).p(\theta_1).|c'(\theta_1)|d\theta_1 \\
&=: (F_c * p)(\theta) \quad \text{where} \qquad (1)\\
F_c(\theta_1, \theta_2) = F(c(\theta_1) - c(\theta_2)) &= \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{i\langle c(\theta_1) - c(\theta_2), \xi\rangle} \frac{1}{(1 + A|\xi|^2)^n} d\xi.
\end{aligned}$$

Note that here, convolution on S^1 uses the metric $L^2(S^1, |c'(\theta|d\theta)$ and it defines a self-adjoint operator for this Hilbert space. Moreover, it is covariant with respect to change in parametrization:

$$F_{c\circ\varphi} * (f \circ \varphi) = (F_c * f) \circ \varphi.$$

What are the properties of the kernel F_c? From the properties of F, we see that F_c is C^{n-1} kernel (except for log poles at the diagonal when $n = 1$). It is also a pseudo-differential operator of order $-2n + 1$ on S^1. To see that let us assume for the moment that each function of θ is a periodic function on $\mathbb{R}$. Then

$$\begin{aligned}
c(\theta_1) - c(\theta_2) &= \int_0^1 c_\theta(\theta_2 + t(\theta_1 - \theta_2))dt.(\theta_1 - \theta_2) =: \tilde{c}(\theta_1, \theta_2)(\theta_1 - \theta_2) \\
F_c(\theta_1, \theta_2) &= \frac{1}{2\pi} \int_{\mathbb{R}^2} e^{i(\theta_1 - \theta_2)\langle \tilde{c}(\theta_1, \theta_2), \xi\rangle} \frac{1}{(1 + A|\xi|^2)^n} d\xi \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} e^{i(\theta_1 - \theta_2)\eta_1} \Big(\int_{\mathbb{R}} \frac{|\tilde{c}(\theta_1, \theta_2)|^{-2}}{(1 + \frac{A}{|\tilde{c}(\theta_1,\theta_2)|^2}(|\eta_1|^2 + |\eta_2|^2))^n} d\eta_2 \Big) d\eta_1
\end{aligned}$$

$$=: \frac{1}{2\pi}\int_{\mathbb{R}} e^{i(\theta_1-\theta_2)\eta_1}\tilde{F}_c(\theta_1,\theta_2,\eta_1)\,d\eta_1$$

where we changed variables as $\eta_1 = \langle \tilde{c}(\theta_1,\theta_2),\xi\rangle$ and $\eta_2 = \langle J\tilde{c}(\theta_1,\theta_2),\xi\rangle$. So we see that $F_c(\theta_1,\theta_2)$ is an elliptic pseudo differential operator kernel of degree $-2n+1$ (the loss comes from integrating with respect to η_2). The symbol $\tilde{F}_c$ is real and positive, so the operator $p \mapsto F_c * p$ is self-adjoint and positive. Thus it is injective, and by an index argument similar to the one in 17.6 it is invertible. The inverse operator to the integral operator F_c is a pseudo-differential operator L_c of order $2n-1$ given by the distribution kernel $L_c(\theta,\theta_1)$ which satisfies

$$L_c * F_c * f = F_c * L_c * f = f$$

$$L_{c\circ\varphi} * (h\circ\varphi) = ((L_c * h)\circ\varphi) \quad \text{for all } \varphi\in \operatorname{Diff}^+(S^1) \tag{2}$$

If we write $h = Y\circ c$, then we want to express the horizontal lift Y in terms of h and write Y_h for it. The set of all these Y_h spans the horizontal subspace which maps isomorphically to $T_c\operatorname{Emb}(S^1,\mathbb{R}^2)$. Now:

$$h = Y\circ c = (F * (c_*(p.ds)))\circ c = F_c * p.$$

Therefore, using the inverse operator, we get $p = L_c * h$ and:

$$Y_h = F * (c_*(p.ds)) = F * (c_*((L_c * h).ds)) \quad \text{or}$$

$$Y_h(x) = \int_{S^1} F(x - c(\theta))\int_{S^1} L_c(\theta,\theta_1)h(\theta_1)|c'(\theta_1)|d\theta_1|c'(\theta)|d\theta$$

and $LY_h = c_*((L_c * h).ds)$. Thus we can finally write down the quotient metric

$$\begin{aligned} G_c^{\mathrm{diff},n}(h,k) &= \int_{\mathbb{R}^2}\langle LY_h, Y_k\rangle dx \\ &= \int_{S^1}\Big\langle L_c * h(\theta), \int_{S^1} F(c(\theta)-c(\theta_1))\int_{S^1} L_c(\theta_1,\theta_2)k(\theta_2)ds_2\,ds_1\Big\rangle ds \\ &= \int_{S^1}\langle L_c * h(\theta), k(\theta)\rangle ds = \iint_{S^1\times S^1} L_c(\theta,\theta_1)\langle h(\theta_1),k(\theta)\rangle ds_1\,ds. \end{aligned} \tag{3}$$

The dual metric on the associated smooth cotangent space $L_c * C^\infty(S^1,\mathbb{R}^2)$ is similarly:

$$\check{G}_c^{\mathrm{diff},n}(p,q) = \iint_{S^1\times S^1} F_c(\theta,\theta_1)\langle p(\theta_1),q(\theta)\rangle ds_1\,ds.$$

18.4. The geodesic equation on $\operatorname{Emb}(S^1,\mathbb{R}^2)$ via conservation of momentum. A quite convincing but not rigorous derivation of this equation can be given using the fact that under a submersion, geodesics on the quotient space are the projections of those geodesics on the total space which are horizontal at one and hence every point. In our case, the geodesics on $\mathrm{Diff}(\mathbb{R}^2)$ can be characterized by the strong conservation of momentum we found above: $\varphi(t)^*\omega(t)$ is independent of t. If $X(t)$ is the tangent vector to the geodesic, i.e. the velocity $X(t) = \partial_t\varphi \circ \varphi^{-1}(t)$, then $\omega(t)$ is just $LX(t) = c_*(p(\theta,t).ds) = c_*(p(\theta,t).|c_\theta(\theta,t)|.d\theta)$ considered as a measure valued 1-form instead of a vector field.

When we pass to the quotient $\mathrm{Emb}(S^1,\mathbb{R}^2)$, a horizontal geodesic of diffeomorphisms $\varphi(t)$ with $\varphi(0) =$ identity gives a geodesic path of embeddings $c(\theta,t) = \varphi(t) \circ c(0,\theta)$. For these geodesic equations, it will be most convenient to take as the momentum the 1-form $\tilde{p}(\theta,t) = p(\theta,t).|c_\theta(\theta,t)|$, the measure factor $d\theta$ being constant along the flow. We must take the velocity to be the horizontal vector field $X(t) = F * c(\cdot,t)_*(\tilde{p}(\theta,t).d\theta)$. For this to be the velocity of the path of maps c, we must have $c_t(\theta,t) = X(c(\theta),t)$ because the global vector field X must extend c_t. To pair $\tilde{p}$ and c_t, we regard $\tilde{p}$ as a 1-form along c (the area factor having been replaced by the measure $d\theta$ supported on C). The geodesic equation must be the differential form of the conservation equation:

$$\boxed{\varphi(t)^*\tilde{p}(\cdot,t) \text{ is independent of } t.}$$

More explicitly, if d_x stands for differentiating with respect to the spatial coordinates x, y, then this means:

$$d_x\varphi(t)^T|_{c(\theta,t)}\tilde{p}(\theta,t) = \text{ cnst.}$$

We differentiate this with respect to t, using the identity:

$$\partial_t d_x\varphi(t) = d_x(\varphi_t(t)) = d_x(X \circ \varphi(t)) = (d_x(X) \circ \varphi(t)) \cdot d_x\varphi(t),$$

we get

$$0 = d_x\varphi(t)^T \cdot \left((d_x(X)^T \circ c(\theta,t)) \cdot \tilde{p}(\theta,t) + \tilde{p}_t(\theta,t)\right).$$

Writing this out and putting the discussion together, we get the following form for the geodesic equation on $\mathrm{Emb}(S^1,\mathbb{R}^2)$:

$$\begin{aligned} c_t(\theta,t) &= X(t) \circ c(\theta,t) \\ \tilde{p}_t(\theta,t) &= -\operatorname{grad} X^t(c(\theta,t),t) \cdot \tilde{p}(\theta,t) \\ X(t) &= F * c(\cdot,t)_* \left(\tilde{p}(\theta,t).d\theta\right) \end{aligned}$$

Note that X is a vector field on the plane: these are not closed equations if we restrict X to the curves. The gradient of X requires that we know the normal derivative of X to the curves. Alternatively, we may introduce a second *vector-valued* kernel on S^1 depending on c by:

$$F'_c(\theta_1,\theta_2) = \operatorname{grad} F(c(\theta_1) - c(\theta_2)).$$

Then the geodesic equations may be written:

$$\boxed{\begin{aligned} &c_t(\theta,t) = (F_c * \tilde{p})(\theta,t) \\ &\tilde{p}_t(\theta,t) = -\langle \tilde{p}(\theta,t), (F_c' * \tilde{p})(\theta,t)\rangle. \end{aligned}}$$

where, in the second formula, the dot product is between the two $\tilde{p}$'s and the vector value is given by F_c'.

The problem with this approach is that we need to enlarge the space $\mathrm{Diff}(\mathbb{R}^2)$ to include diffeomorphisms which are not C^∞ along some C^∞ curve but have a mild singularity normal to the curve. Then we would have to develop differential geometry and the theory of geodesics on this space, etc. It seems more straightforward to outline the direct derivation of the above geodesic equation, along the lines used above.

18.5. The geodesic equation on $\mathrm{Emb}(S^1, \mathbb{R}^2)$, direct approach. The space of invertible pseudo differential operators on a compact manifold is a regular Lie group (see [**2**]), so we can use the usual formula $d(A^{-1}) = -A^{-1}.dA.A^{-1}$ for computing the derivative of L_c with respect to c. Note that we have a simple expression for $D_{c,h}F_c$, namely

$$D_{c,h}F_c(\theta_1,\theta_2) = dF(c(\theta_1)-c(\theta_2))(h(\theta_1)-h(\theta_2)) = \langle F_c'(\theta_1,\theta_2), h(\theta_1)-h(\theta_2)\rangle$$

hence

$$\begin{aligned} D_{c,\ell}L_c(\theta_1,\theta_2) &= -\int_{(S^1)^2} L_c(\theta_1,\theta_3)\, D_{c,h}F_c(\theta_3,\theta_4)\, L_c(\theta_4,\theta_2)\, d\theta_3\, d\theta_4 \\ &= -\int_{(S^1)^2} L_c(\theta_1,\theta_3)\, \langle (F_c'(\theta_3,\theta_4), \ell(\theta_3)\rangle\, L_c(\theta_4,\theta_2)\, d\theta_3\, d\theta_4 \\ &\quad + \int_{(S^1)^2} L_c(\theta_1,\theta_4)\, \langle (F_c'(\theta_4,\theta_3), \ell(\theta_3)\rangle\, L_c(\theta_3,\theta_2)\, d\theta_3\, d\theta_4 \end{aligned}$$

We can now differentiate the metric where $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ is the variable on $(S^1)^n$:

$$\begin{aligned} D_{c,\ell}G^{\mathrm{diff},n}(h,k) &= \int_{(S^1)^2} D_{c,\ell}L_c(\theta_1,\theta_2)\langle h(\theta_2), k(\theta_1)\rangle d\theta \\ &= \int_{(S^1)^4} \Big\langle -L_c(\theta_1,\theta_3)F_c'(\theta_3,\theta_4)L_c(\theta_4,\theta_2) \\ &\qquad + L_c(\theta_1,\theta_4)F_c'(\theta_4,\theta_3)\, L_c(\theta_3,\theta_2), \ell(\theta_3)\Big\rangle \langle h(\theta_2), k(\theta_1)\rangle\, d\theta \end{aligned}$$

We have to write this in the form

$$D_{c,\ell}G_c^{\mathrm{diff},n}(h,k) = G_c^{\mathrm{diff},n}(\ell, H_c(h,k)) = G_c^{\mathrm{diff},n}(K_c(\ell,h),k)$$

For H_c we use $\delta(\theta_5-\theta_3)=\int L_c(\theta_5,\theta_6)F_c(\theta_6,\theta_3)d\theta_6=(L_c*F_c)(\theta_5,\theta_3)$ as follows:

$$\begin{aligned}D_{c,\ell}G^{\mathrm{diff},n}(h,k)=\int_{(S^1)^6}&L_c(\theta_5,\theta_6)\Big\langle \ell(\theta_5),\Big(-L_c(\theta_1,\theta_3)F_c'(\theta_3,\theta_4)L_c(\theta_4,\theta_2)\\&+L_c(\theta_1,\theta_4)F_c'(\theta_4,\theta_3)\,L_c(\theta_3,\theta_2)\Big)F_c(\theta_6,\theta_3)\langle h(\theta_2),k(\theta_1)\rangle\Big\rangle\,d\theta\end{aligned}$$

Thus

$$\begin{aligned}H_c(h,k)(\theta_0)=\int_{(S^1)^4}&\Big(-L_c(\theta_1,\theta_3)F_c'(\theta_3,\theta_4)L_c(\theta_4,\theta_2)\\&+L_c(\theta_1,\theta_4)F_c'(\theta_4,\theta_3)\,L_c(\theta_3,\theta_2)\Big)F_c(\theta_0,\theta_3)\langle h(\theta_2),k(\theta_1)\rangle\,d\theta\end{aligned}$$

Similarly we get

$$\begin{aligned}D_{c,\ell}G^{\mathrm{diff},n}(h,k)=\int_{(S^1)^6}&L_c(\theta_6,\theta_5)\Big\langle F_c(\theta_1,\theta_6)\Big\langle -L_c(\theta_1,\theta_3)F_c'(\theta_3,\theta_4)L_c(\theta_4,\theta_2)\\&+L_c(\theta_1,\theta_4)F_c'(\theta_4,\theta_3)\,L_c(\theta_3,\theta_2),\ell(\theta_3)\Big\rangle h(\theta_2),k(\theta_5)\Big\rangle\,d\theta\end{aligned}$$

so that

$$\begin{aligned}&K_c(\ell,h)(\theta_0)=\\&=\int_{(S^1)^2}\Big(-\langle F_c'(\theta_0,\theta_1)L_c(\theta_1,\theta_2),\ell(\theta_0)\rangle+\langle F_c'(\theta_0,\theta_1)\,L_c(\theta_1,\theta_2),\ell(\theta_1)\rangle\Big)h(\theta_2)\,d\theta\end{aligned}$$

By 15.4 the geodesic equation is given by

$$c_{tt}(\theta_0)=\tfrac12 H_c(c_t,c_t)(\theta_0)-K_c(c_t,c_t)(\theta_0)$$

Let us rewrite the geodesic equation in terms of L_c*c_t. We have (suppressing the variable t and collecting all terms)

$$\begin{aligned}(L_c*c_t)_t(\theta_0)&=\int_{S^1}D_{c,c_t}L_c(\theta_0,\theta_1)c_t(\theta_1)\,d\theta_1+L_c*c_{tt}\\&=\tfrac12\int_{S^1}\Big(F_c'(\theta_1,\theta_0)-F_c'(\theta_0,\theta_1)\Big)\langle L_c*c_t(\theta_0),L_c*c_t(\theta_1)\rangle\,d\theta_1\end{aligned}$$

Since the kernel F is an even function we get the same geodesic equation as above for the momentum $\tilde p(\theta,t)=L_c*c_t=p(\theta,t).|c_\theta|$:

$$\tilde p_t(\theta_0)=-\int_{S^1}F_c'(\theta_0,\theta_1)\langle\tilde p(\theta_0),\tilde p(\theta_1)\rangle\,d\theta_1 \tag{1}$$

18.6. Existence of geodesics. Theorem. *Let $n\ge 1$. For each $k>2n-\frac12$ the geodesic equation 18.5 (1) has unique local solutions in the Sobolev space of H^k-embeddings. The solutions are C^∞ in t and in the initial conditions $c(0,\ .\ )$ and $c_t(0,\ .\ )$. The domain of existence (in t) is uniform in k and thus this also holds in $\mathrm{Emb}(S^1,\mathbb{R}^2)$.*

An even stronger theorem, proving *global* existence on the level of H^k-diffeomorphisms on $\mathbb{R}^2$, has been proved by [**73, 75, 76**].

Proof. Let $c \in H^k$. We begin by checking that F'_c is a pseudo differential operator kernel of order $-2n+2$ as we did for F_c in 18.3.

$$
\begin{aligned}
c(\theta_1) - c(\theta_2) &=: \tilde c(\theta_1,\theta_2)(\theta_1-\theta_2)\\
\operatorname{grad} F(x) &= \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i\langle x,\xi\rangle}\frac{J\xi}{(1+A|\xi|^2)^n}\,d\xi\\
F'_c(\theta_1,\theta_2) &= \frac{1}{2\pi}\int_{\mathbb{R}^2} e^{i(\theta_1-\theta_2)\langle \tilde c(\theta_1,\theta_2),\xi\rangle}\frac{J\xi}{(1+A|\xi|^2)^n}\,d\xi\\
&= \frac{1}{2\pi}\int_{\mathbb{R}} e^{i(\theta_1-\theta_2)\eta_1}\Big(\int_{\mathbb{R}}\frac{|\tilde c(\theta_1,\theta_2)|^{-3}.J\eta}{(1+\frac{A}{|\tilde c(\theta_1,\theta_2)|^2}(|\eta_1|^2+|\eta_2|^2))^n}\,d\eta_2\Big)\,d\eta_1\\
&=: \frac{1}{2\pi}\int_{\mathbb{R}} e^{i(\theta_1-\theta_2)\eta_1}\tilde F_c(\theta_1,\theta_2,\eta_1)\,d\eta_1
\end{aligned}
$$

where we changed variables as $\eta_1 = \langle \tilde c(\theta_1,\theta_2),\xi\rangle$ and $\eta_2 = \langle J\tilde c(\theta_1,\theta_2),\xi\rangle$. So we see that $F'_c(\theta_1,\theta_2)$ is an elliptic pseudo differential operator kernel of degree $-2n+2$ (the loss of 1 comes from integrating with respect to η_2). We write the geodesic equation in the following way:

$$
\begin{aligned}
c_t &= F_c * q =: Y_1(c,q)\\
q_t &= \langle q, F'_c * q\rangle = \int F'_c(\ .\ ,\theta)\langle q(\theta), q(\ .\ )\rangle\,d\theta =: Y_2(c,q)
\end{aligned}
$$

We start with $c \in H^k$ where $k > 2n-\frac12$, in the H^2-open set $U^k := \{c : |c_\theta| > 0\} \subset H^k$. Then $q = L_c * c_t \in H^{k-2n+1}$ and $F'_c * q \in H^{k-1} \subset H^{k-2n+1}$. By the Banach algebra property of the Sobolev space H^{k-2n+1} the expression (with missuse of notation) $Y_2(c,q) = \langle q, F'_c * q\rangle \in H^{k-2n+1}$. Since the kernel F is not smooth only at 0, all appearing pseudo differential operators kernels are C^∞ off the diagonal, thus are smooth mappings in c with values in the space of operators between the relevant Sobolev spaces. Let us make this more precise. We claim that $c \mapsto F'_c * (\ .\ ) \in L(H^k, H^{k+2n-2})$ is C^∞. Since the Sobolev spaces are convenient, we can (a) use the smooth uniform boundedness theorem [**42**], 5.18, so that it suffices to check that for each fixed $q \in H^k$ the mapping $c \mapsto F'_c * q$ is smooth into H^{k+2n-2}. Moreover, by [**42**], 2.14 it suffices (b) to check that this is weakly smooth: Using the L^2-duality between H^{k+2n-2} and $H^{-k-2n+2}$ it suffices to check, that for each $p \in H^{-k-2n+2}$ the expression

$$
\begin{aligned}
&\int p(\theta_1)(F'_c * q)(\theta_1)\,d\theta_1 =\\
&= \iint \frac{p(\theta_1)}{2\pi}\int_{\mathbb{R}} e^{i(\theta_1-\theta_2)\eta_1}\Big(\int_{\mathbb{R}}\frac{|\tilde c(\theta_1,\theta_2)|^{-3}.J\eta}{(1+\frac{A}{|\tilde c(\theta_1,\theta_2)|^2}(|\eta_1|^2+|\eta_2|^2))^n}\,d\eta_2\Big)\,d\eta_1 q(\theta_2)\,d\theta_1\,d\theta_2
\end{aligned}
$$

is a smooth mapping $\mathrm{Emb}(S^1,\mathbb{R}^2) \to \mathbb{R}^2$. For that we may assume that c depends on a further smooth variable s. Convergence of this integral depends on the highest order term in the asymptotic expansion in η, which does not change if we differentiate with respect to s.

Thus the geodesic equation is the flow equation of a smooth vector field $Y = (Y_1, Y_2)$ on $U^k \times H^{k-2n+1}$. We thus have local existence and uniqueness of the flow Fl^k on $U^k \times H^{k-2n+1}$.

Now we consider smooth initial conditions $c_0 = c(0,\ .\ )$ and $q_0 = q(0,\ .\ ) = (L_c{*}c_t)(0,\ .\ )$ in $C^\infty(S^1,\mathbb{R}^2)$. Suppose the trajectory $\mathrm{Fl}^k_t(c_0,q_0)$ of Y through these intial conditions in $U^k \times H^{k+1-2n}$ maximally exists for $t \in (-a_k, b_k)$, and the trajectory $\mathrm{Fl}^{k+1}_t(c_0,u_0)$ in $U^{k+1} \times H^{k+2-2n}$ maximally exists for $t \in (-a_{k+1}, b_{k+1})$ with $b_{k+1} < b_k$. By uniqueness we have $\mathrm{Fl}^{k+1}_t(c_0,u_0) = \mathrm{Fl}^k_t(c_0,u_0)$ for $t \in (-a_{k+1,}b_{k+1})$. We now apply ∂_θ to the equation $q_t = Y_2(c,q)$, note that the commutator $q \mapsto [F'_c, \partial_\theta] * q = \partial_t h(F'_c * q) - F'_c * (\partial_\theta q)$ is a pseudo differential operator of order $-2n+2$ again, and obtain

$$\begin{aligned} \partial_\theta q_t = \int [F'_c, \partial_\theta](\ .\ ,\theta)\langle q(\theta), q(\ .\ )\rangle\, d\theta + \int F'_c(\ .\ ,\theta)\langle \partial_\theta q(\theta), q(\ .\ )\rangle\, d\theta \\ + \int F'_c(\ .\ ,\theta)\langle q(\theta), \partial_\theta q(\ .\ )\rangle\, d\theta \end{aligned}$$

which is an inhomogeneous linear equation for $w = \partial_\theta q$ in $U^k \times H^{k+1-2n}$. By the variation of constant method one sees that the solution w exists in H^k for as long as (c,q) exists in $U^k \times H^{k+1-2n}$, i.e., for all $t \in (-a_k, b_k)$. By continuity we can conclude that $w = \partial_\theta q$ is the derivative in H^{k+2-2n} for $t = b_{k+1}$, and thus the domain of definition was not maximal. Iterating this argument we can conclude that the solution (c,q) lies in C^∞ for $t \in (-a_k, b_k)$. □

18.7. Horizontality for $G^{\mathrm{diff},n}$. The tangent vector $h \in T_c\,\mathrm{Emb}(S^1,\mathbb{R}^2)$ is $G^{\mathrm{diff},n}_c$-orthogonal to the $\mathrm{Diff}(S^1)$-orbit through c if and only if

$$0 = G^{\mathrm{diff},n}_c(h, \zeta_X(c)) = \int_{(S^1)^2} L_c(\theta_1,\theta_2)\langle h(\theta_2), c_\theta(\theta_1)\rangle X(\theta_1)\, ds_1\, ds_2$$

for all $X \in \mathfrak{X}(S^1)$. So the $G^{\mathrm{diff},n}$-normal bundle is given by

$$\mathcal{N}^{\mathrm{diff},n}_c = \{h \in C^\infty(S^1,\mathbb{R}^2) : \langle L_c * h, v\rangle = 0\}.$$

Working exactly as in section 4, we want to split any tangent vector into vertical and horizontal parts as $h = h^\top + h^\perp$ where $h^\top = X(h).v$ for $X(h) \in \mathfrak{X}(S^1)$ and where $h^\perp$ is horizontal, $\langle L_c * h^\perp, v\rangle = 0$. Then $\langle L_c * h, v\rangle = \langle L_c * (X(h)v), v\rangle$ and we are led to consider the following operators:

$$\begin{aligned} &L^\top_c, L^\perp_c : C^\infty(S^1) \to C^\infty(S^1), \\ &L^\top_c(f) = \langle L_c * (f.v), v\rangle = \langle L_c * (f.n), n\rangle, \end{aligned}$$

$$L_c^\perp(f) = \langle L_c * (f.v), n\rangle = -\langle L_c * (f.n), v\rangle.$$

The pseudo differential operator $L_c^\top$ is unbounded, selfadjoint and positive on $L^2(S^1, d\theta)$ since we have

$$\int_{S^1} L_c^\top(f).f\, d\theta = \int_{(S^1)^2} \langle L_c(\theta_1,\theta_2)f(\theta_2)v(\theta_2), f(\theta_1).v(\theta_1)\rangle\, d\theta = \|f.v\|^2_{G^{\mathrm{diff},n}} > 0.$$

Thus $L_c^\top$ is injective and by an index argument as in 17.6 the operator $L_c^\top$ is invertible. Moreover, the operator $L_c^\perp$ is skew-adjoint. To go back and forth between the natural horizontal space of vector fields $a.n$ and the $G^{\mathrm{diff},n}$-horizontal vectors, we have to find b such that $L_c * (a.n + b.v) = f.n$ for some f. But then

$$L_c^\perp(a) = -\langle L_c{*}(a.n), v\rangle = \langle L_c{*}(b.v), v\rangle = L_c^\top(b) \quad \text{thus} \quad b = (L_c^\top)^{-1}L_c^\perp(a).$$

Thus $a.n + (L_c^\top)^{-1}L_c^\perp(a).v$ is always $G^{\mathrm{diff},n}$-horizontal and is the horizontal projection of $a.n + b.v$ for any b.

Proposition. *For any smooth path c in $\mathrm{Imm}(S^1,\mathbb{R}^2)$ there exists a smooth path φ in $\mathrm{Diff}(S^1)$ with $\varphi(0,\ .\ ) = \mathrm{Id}_{S^1}$ depending smoothly on c such that the path e given by $e(t,\theta) = c(t,\varphi(t,\theta))$ is $G^{\mathrm{diff},n}$-horizontal: $\langle L_c{*}e_t, e_\theta\rangle = 0$.*

Proof. Let us write $e = c\circ\varphi$ for $e(t,\theta) = c(t,\varphi(t,\theta))$, etc. We look for φ as the integral curve of a time dependent vector field $\xi(t,\theta)$ on S^1, given by $\varphi_t = \xi\circ\varphi$. We want the following expression to vanish. In its computation the equivariance of L_c under $\varphi \in \mathrm{Diff}^+(S^1)$ from 18.3(2) will play an important role.

$$\begin{aligned}
&\big\langle L_{c\circ\varphi} * (\partial_t(c\circ\varphi)), \partial_\theta(c\circ\varphi)\big\rangle = \langle L_{c\circ\varphi} * (c_t\circ\varphi + (c_\theta\circ\varphi)\,\varphi_t), (c_\theta\circ\varphi)\,\varphi_\theta\rangle\\
&= \big\langle ((L_c * c_t)\circ\varphi) \ + ((L_c * (c_\theta.\xi))\circ\varphi), (c_\theta\circ\varphi)\varphi_\theta\big\rangle\\
&= \big((\langle L_c * c_t, c_\theta\rangle + \langle L_c * (\xi.c_\theta), c_\theta\rangle)\circ\varphi\big)\,\varphi_\theta.
\end{aligned}$$

Using the time dependent vector field $\xi = -(L_c^\top)^{-1}\langle L_c * c_t, c_\theta\rangle$ and its flow φ achieves this. $\square$

To write the quotient metric on B_e, we want to lift normal vector fields $a.n$ to a curve C to horizontal vector fields on $\mathrm{Emb}(S^1,\mathbb{R}^2)$. Substituting $h = a.n + (L_c^\top)^{-1}L_c^\perp(a).v,\ k = b.n + (L^\top)^{-1}L^\perp(b).v$ in 18.3(3), we get as above:

$$G_C^{\mathrm{diff},n}(a,b) = \int_C \Big(L_c^\top + L_c^\perp(L_c^\top)^{-1}L_c^\perp\Big)(a).b ds.$$

The dual metric on the cotangent space is just the restriction of the dual metric on $\mathrm{Emb}(S^1,\mathbb{R}^2)$ to the cotangent space to B_e and is much simpler. We simply set $p = f.n,\ q = g.n$ and get:

$$\check{G}_C^{\mathrm{diff},n}(f,g) = \iint_{C^2} F(x(s) - x(s_1)).\langle n(s), n(s_1)\rangle.f(s)g(s_1).dsds_1$$

where $x(s) \in \mathbb{R}^2$ stands for the point in the plane with arc length coordinate s and F is the Bessel kernel. Since these are dual inner products, we find that the two operators, (a) convolution with the kernel $F(x(s) - x(s_1)).\langle n(s), n(s_1)\rangle$ and (b) $L_c^\top + L_c^\perp (L_c^\top)^{-1} L_c^\perp$ are inverses of each other.

18.8. The geodesic equation on B_e via conservation of momentum. The simplest way to find the geodesic equation on B_e is again to specialize the general rule $\varphi(t)^*\omega(t) =$ cnst. to the horizontal geodesics. Now horizontal in the present context, that is for B_e, requires more of the momentum $\omega(t)$. As well as being given by $c_*(p(s).ds)$, we require the 1-form p to kill the tangent vectors v to the curve. If we identify 1-forms and vectors using the Euclidean metric, then we may say simply $p(s) = a(s).n$, where a is a scalar function on C. But note that if you take the momentum as $c_*(a(s)n(s)ds)$ and integrate it against a vector field X, then you find:

$$\langle X, c_*(a(s)n(s)ds)\rangle = \int_C a(s)\langle X, n(s)\rangle ds = \int_C a(s).i_X(dx \wedge dy)$$

where i_X is the 'interior product' or contraction with X taking a 2-form to a 1-form. Noting that 1-forms can be integrated along curves without using any metric, we see that the 2-form along c defined by $\{a(s).(dx \wedge dy)_{c(s)}\}$ can be naturally paired with vector fields so it defines a canonical measure valued 1-form. Therefore, the momentum for horizontal geodesics can be identified with this 2-form.

If $\varphi(x,t)$ is a horizontal geodesic in $\mathrm{Diff}(\mathbb{R}^2)$, then the curves $C_t = \mathrm{image}(c(\cdot,t))$ are given by $C_t = \varphi(C_0, t)$ and the momentum is given by $a(\theta,t).(dx \wedge dy)$, where $c(\theta,t)$ parametrizes the curves C_t. Note that in order to differentiate a with respect to t, we need to assign parameters on the curves C_t simultaneously. We do this in the same way we did for almost local metrics: assume c_θ is a multiple of the normal vector n_C. But $\theta_0 \mapsto \varphi(\theta_0, t)$ gives a second map from C_0 to C_t: in terms of θ, assume this is $\theta = \bar\varphi(\theta_0, t)$. Then the conservation of momentum means simply:

$$a(\bar\varphi(\theta_0,t),t).\det(D_x\varphi)(c(\theta_0,0),t) \text{ is independent of } t.$$

Let X be the global vector field giving this geodesic, so that $\varphi_t = X \circ \varphi$. Note that $\bar\varphi_t = (\langle X \circ c, v\rangle/|c_\theta|) \circ \bar\varphi$. Using this fact, we can differentiate the displayed identity. Recalling the definition of the flow from its momentum and the identifying $T_C B_e$ with normal vector fields along C, we get the full equations for the geodesic:

$$\begin{aligned} C_t &= \langle X, n\rangle \cdot n \\ a_t &= -\langle X, v\rangle D_s(a) - \mathrm{div}(X).a \\ X &= F * c_*(a(s)n(s)ds) \end{aligned}$$

Note, as in the geodesic equations in 18.4, that we must use F to extend X to the whole plane. In this case, we only need (a) the normal component of X along C, (b) its tangential component along C and (c) the divergence of X along C. These are obtained by convolving $a(s)$ with the kernels (which we give now in terms of arc-length):

$$\begin{aligned} F_c^{nn}(s_1,s_2) &= F(c(s_1)-c(s_2))\langle n(s_1),n(s_2)\rangle ds_2 \\ F_c^{vn}(s_1,s_2) &= F(c(s_1)-c(s_2))\langle v(s_1),n(s_2)\rangle ds_2 \\ F_c^{\mathrm{div}}(s_1,s_2) &= \langle \operatorname{grad} F(c(s_1)-c(s_2)), n(s_1)\rangle ds_2 \end{aligned}$$

Then the geodesic equations become:

$$\begin{aligned} C_t &= (F_c^{nn} * a).n \\ a_t &= -(F_c^{vn} * a)D_s(a) - (F_c^{\mathrm{div}} * a).a \end{aligned}$$

Alternately, we may specialize the geodesic equation in 18.4 to horizontal paths. Then the v part vanishes identically and the n part gives the last equation above. We omit this calculation.

19. Examples

19.1. The Geodesic of concentric circles. All the metrics that we have studied are invariant under the motion group, thus the 1-dimensional submanifold of B_e consisting of all concentric circles centered at the origin is the fixed point set of the group of rotations around the center. Therefore it is a geodesic in all our metrics. It is given by the map $c(t,\theta) = r(t)e^{i\theta}$. Then $c_\theta = ire^{i\theta}$, $v_c = ie^{i\theta}$, $n_c = -e^{i\theta}$, $\ell(c) = 2\pi r(t)$, $\kappa_c = \frac{1}{r}$ and $c_t = r_t e^{i\theta} = -r_t.n_c$.

The parametrization $r(t)$ can be determined by requiring constant speed σ, i.e. if the metric is $G(h,k)$, then we require $G_c(c_t,c_t) = r_t^2 G_c(n_c,n_c) = \sigma^2$, which leads to $\sqrt{G_c(c_t,c_t)}\,dt = \pm\sqrt{G_c(n_c,n_c)}\,dr$. To determine when the geodesic is complete as $r \to 0$ and $r \to \infty$, we merely need to look at its length which is given by:

$$\int_0^\infty \sqrt{G_c(c_t,c_t)}dt = \int_0^\infty \sqrt{G_c(n_c,n_c)}dr,$$

and we need to ask whether this integral converges or diverges at its two limits. Let's consider this case by case.

The metric G^Φ: The geodesic is determined by the equation:

$$G^\Phi(c_t,c_t) = 2\pi r \cdot \Phi\left(2\pi r(t), \frac{1}{r(t)}\right) \cdot r_t^2 = \sigma^2.$$

Differentiating this with respect to t leads to the geodesic equation in the standard form $r_{tt} = r_t^2 f(r)$. It is easily checked that all three invariant momentum mappings vanish: the reparameterization, linear and angular momentum.

Theorem. *If $\Phi(2\pi r, 1/r) \approx C_1 r^a$ (resp. $C_2 r^b$) as $r \to 0$ (resp. ∞), then the geodesic of concentric circles is complete for $r \to 0$ if and only $a \leq -3$ and is complete for $r \to \infty$ if and only if $b \geq -3$. In particular, for $\varphi = \ell^k$, we find $k = a = b$ and the geodesic is given by $r(t) = cnst.t^{2/(k+3}$. For the scale invariant case $\Phi(\ell,\kappa) = \frac{4\pi^2}{\ell^3} + \frac{\kappa^2}{\ell}$, we find $a = b = -3$, the geodesic is given by $r(t) = e^{\sqrt{2}\sigma t}$ and is complete. Moreover, in this case, the scaling momentum $\frac{2r_t}{r}$ is constant in t along the geodesic.*

The proof is straightforward.

The metric $G^{\mathbf{imm},n}$ Recall from 17.1 the operator $L_{n,c} = I + (-1)^n A.D_s^{2n}$. For $c(t,\theta) = r(t)e^{i\theta}$ 19.1 we have

$$L_{n,c}(c_t) = \big(1 + (-1)^n \frac{A}{r^{2n}}\partial_\theta^{2n}\big)(r_t e^{i\theta}) = r_t\big(1 + \frac{A}{r^{2n}}\big)e^{i\theta}$$

which is still normal to c_θ. So $t \mapsto c(t,\ .\ )$ is a horizontal path for any choice of $r(t)$. Thus its speed is the square root of:

$$G^{\mathrm{imm},n}(c_t, c_t) = 2\pi r \cdot \left(1 + \frac{A}{r^{2n}}\right) \cdot r_t^2 = \sigma^2.$$

For $n = 1$ this is the same as the identity for the metric with $\Phi(\ell,\kappa) = 1 + A\kappa^2$ which was computed in [**54**], 5.1. An explanation of this phenomenon is in [**54**], 3.2.

Theorem. *The geodesic of concentric cirles is complete in the $G^{\mathrm{imm},n}$ metric if $n \geq 2$. For $n = 1$, it is incomplete as $r \to 0$ but complete if $r \to \infty$.*

The metric $G^{\mathrm{diff},n}$ To evaluate the norm of a path of concentric circles, we now need to find the vector field X on $\mathbb{R}^2$ gotten by convolving the Bessel kernel with the unit normal vector field along a circle. Using circular symmetry, we find that:

$$\begin{aligned}
&X(x,y) = f(r)\left(\frac{x}{r}, \frac{y}{r}\right) r_t \\
&\left(I - A(\partial_{rr} + \frac{1}{r}\partial_r)\right)^n f = 0 \text{ except on the circle } r = r_0 \\
&f \in C^{2n-2} \text{ everywhere}, f(r_0) = 1
\end{aligned}$$

For $n = 1$, we can solve this and the result is the vector field on $\mathbb{R}^2$given by the Bessel functions I_1 and K_1:

$$X(x,y) = \begin{cases} \frac{I_1(r/\sqrt{A})}{I_1(r_0/\sqrt{A})} \text{ if } r \leq r_0 \\ \frac{K_1(r/\sqrt{A})}{K_1(r_0/\sqrt{A})} \text{ if } r \geq r_0 \end{cases}$$

Using the fact that the Wronskian of I_1, K_1 is $1/r$, we find:

$$G^{\text{diff},1}(r_t n, r_t n) = \int \langle (I - A\triangle)X, X\rangle r_t^2 = \frac{2\pi r_t^2}{K_1(r/\sqrt{A}).I_1(r/\sqrt{A})}.$$

Using the asymptotic laws for Bessel functions, one finds that the geodesic of concentric circles has finite length to $r = 0$ but infinite length to $r = \infty$.

For $n > 1$, it gets harder to solve for X. But lower bounds are not hard:

$$\begin{aligned} G^{\text{diff},n}(n,n) &= \inf_{X, \langle X,n\rangle \equiv 1 \text{ on } C_r} \int \langle (I - A\triangle)^n X, X\rangle \\ &\geq A^n. \inf_{\left(\substack{X,\langle X,n\rangle\equiv 1 \text{ on } C_r \\ X\to 0, \text{ when } x\to\infty}\right)} \int \langle \triangle^n(X), X\rangle \underset{\text{def}}{=} M(r) \end{aligned}$$

Then $M(r)$ scales with r: $M(r) = M(1)/r^{2n-2}$, hence the length of the path when the radius shrinks to 0 is bounded below by $\int_0 dr/r^{n-1}$ which is infinite if $n > 1$. On the other hand, the metric $G^{\text{diff},n}$ dominates the metric $G^{\text{diff},1}$ so the length of the path when the radius grows to infinity is always infinite. Thus:

Theorem. *The geodesic of concentric cirles is complete in the $G^{\text{diff},n}$ metric if $n \geq 2$. For $n = 1$, it is incomplete as $r \to 0$ but complete if $r \to \infty$.*

19.2. Unit balls in five metrics at a 'cigar'-like shape. It is useful to get a sense of how our various metrics differ. One way to do this is to take one simple shape C and examine the unit balls in the tangent space $T_C B_e$ for various metrics. All of our metrics (except the simple L^2 metric) involve a constant A whose dimension is length-squared. We take as our base shape C a strip of length L, where $L \gg \sqrt{A}$, and width w, where $w \ll \sqrt{A}$. We round the two ends with semi-circles, as shown in on the top in figure 1.

As functions of a normal vector field $a.n$ along C, the metrics we want to compare are:

(1) $G_C^A(a,a) = \int_C (1 + A\kappa^2)a^2.ds$

(2) $G_C^{\text{imm},1}(a,a) = \inf_b \int_C \Big(|a\cdot n + b\cdot v)|^2 + A|D_s(a\cdot n + b\cdot v)|^2\Big) ds$

(3) $G_C^{\text{diff},1}(a,a) = \frac{1}{\sqrt{A}} \inf_{\left(\substack{\mathbb{R}^2\text{-vec.flds.}X \\ \langle X,n\rangle = a}\right)} \iint_{\mathbb{R}^2} \Big(|X|^2 + A|DX|^2\Big)\, dx\, dy,$

(4) $G_C^{\text{diff},2}(a,a) = \frac{1}{\sqrt{A}} \inf_{\left(\substack{\mathbb{R}^2\text{-vec.flds.}X \\ \langle X,n\rangle = a}\right)} \iint_{\mathbb{R}^2} \Big(|X|^2 + 2A|DX|^2 + A^2|D^2X|^2\Big)\, dx\, dy$

The term $\frac{1}{\sqrt{A}}$ in the last 2 metrics is put there so that the double integrals have the same 'dimension' as the single integrals. In this way, all the metrics will be comparable.

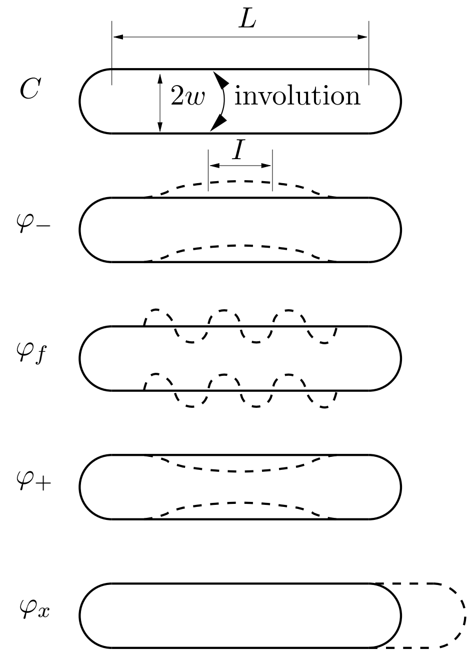


Figure 9. The cigarlike shape and their deformations

To compare the 4 metrics, we don't take all normal vector fields $a.n$ along C. Note that C has an involution, which flips the top and bottom edges. Thus we have even normal vector fields and odd normal vector fields. Examples are shown in figure 1. We will consider two even and two odd normal vector field, described below, and normalize each of them so that $\int_C a^2 ds = 1$. They are also shown in the figure. They involve some interval I along the long axis of the shape of length $\lambda \gg w$. The interval determines a part I_t of the top part of C and I_b of the bottom.

(1) Let $a \equiv +1/\sqrt{2\lambda}$ along I_t and $a \equiv -1/\sqrt{2\lambda}$ along I_b, a zero elsewhere except we smooth it at the endpoints of I. Call this odd vector field φ_-.

(2) Fix a high frequency f and, on the same intervals, let $a(x) = \pm \sin(f.x)/\sqrt{\lambda}$. Call this odd vector field φ_f.

(3) The third vector field is even and is defined by $a(x) = \sqrt{\frac{2}{\pi w}}\langle n, \frac{\partial}{\partial x}\rangle$ at the right end of the curve, being zero along top, bottom and left end. Call this φ_x. The factor in front normalizes this vector field so that its L^2 norm is 1.

(4) Finally, we define another even vector field by $a = +1/\sqrt{2\lambda}$ on both I_t and I_b, zero elsewhere except for being smoothed at the ends of I. Call this φ_+.

The following table shows the approximate leading term in the norm of each of these normal vector fields a in each of the metrics. By approximate, we mean the exact norm is bounded above and below by the entry times a constant depending only on C and by leading term, we mean we ignore terms in the small ratios $w/\sqrt{A}$, $\lambda/\sqrt{A}$:

function	G^A	$G^{\text{imm},1}$	$G^{\text{diff},1}$	$G^{\text{diff},2}$
φ_-	1	1	1	1
φ_f	1	$(\sqrt{A}f)^2$	$\sqrt{A}f$	$(\sqrt{A}f)^3$
φ_+	1	1	$\sqrt{A}/w$	$(\sqrt{A}/w)^3$
φ_x	A/w^2	$\sqrt{A}/w$	$\frac{\sqrt{A}}{w\log(A/w^2)}$	$\sqrt{A}/w$

Thus, for instance:

$$\begin{aligned} G^A(\varphi_x, \varphi_x) &= \frac{2}{\pi w} \int_{\text{right end}} (1 + A\kappa^2)\langle n, \tfrac{\partial}{\partial x}\rangle^2 ds \\ &= \tfrac{2}{\pi w}(1 + Aw^{-2})\ell(\text{right end})\text{Ave}(\langle n, \tfrac{\partial}{\partial x}\rangle^2) \\ &= (1 + Aw^{-2}) \approx A/w^2 \end{aligned}$$

The values of all the other entries under G^A are clear because $\kappa \equiv 0$ in their support.

To estimate the other entries, we need to estimate the horizontal lift, i.e., the functions b or v. To estimate the norms for $G^{\text{imm},1}$, we take $b = 0$ in all cases except φ_x and then get

$$G_C^{\text{imm},1}(a, a) = \|a\|^2_{H^1_A}$$

the first Sobolev norm. For $a = \varphi_f$, we simplify this, replacing the full norm by the leading term $A(D_s(a))^2$ and working this out. To compute $G^{\text{imm},1}(\varphi_x, \varphi_x)$, let $k = \sqrt{\frac{2}{\pi w}}$ be the normalizing factor and lift $a.n$ along the right end of C to the $\mathbb{R}^2$ vector field $k.\frac{\partial}{\partial x}$. This adds a tangential component which we taper to zero on the top and bottom of C like $k.e^{-x/\sqrt{A}}$. This gives the estimate in the table.

Finally, consider the 2 metrics $G^{\text{diff},k}$, $k = 1, 2$. For these, we need to lift the normal vector fields along C to vector fields on all of $\mathbb{R}^2$. For the two odd vector fields $f = \varphi_-$ and $f = \varphi_f$, we take v to be constant along the small vertical lines inside C and zero in the extended strip $-w \le y \le w, x \notin I$ and we define v outside $-w \le y \le w$ by:

$$v(x, y + w) = v(x, -y - w) = F(x, y)\frac{\partial}{\partial y},$$

$$\widehat{F}(\xi,\eta) = \frac{k\sqrt{A}(1+A\xi^2)^{k-1/2}.\hat{f}(\xi)}{\pi(1+A(\xi^2+\eta^2))^k}$$

We check the following:

(a) $\Big((I-A\triangle)^k F\Big)^\wedge = \frac{k}{\pi}\sqrt{A}(1+A\xi^2)^{k-1/2}\hat{f}(\xi)$ is indep. of η hence support$(I-A\triangle)^k F) \subset \{y=0\}$

(b) $\displaystyle\int \hat{F} d\eta = \hat{f}$, hence $F|_{y=0} = f$.

Thus:

$$\begin{aligned} G_C^{\mathrm{diff},k}(f,f) &\approx \tfrac{1}{\sqrt{A}} \iint_{\mathbb{R}^2} \langle (I-A\triangle)^k F, F\rangle dxdy \\ &= \tfrac{k}{\pi}\int (1+A\xi^2)^{k-1/2}|\hat{f}(\xi)|^2 d\xi = \tfrac{k}{\pi}\|f\|^2_{H_A^{k-1/2}} \end{aligned}$$

The leading term in the k^{th} Sobolev norm of φ_f is $(\sqrt{A}f)^{2k}$, which gives these entries in the table.

To estimate $G_C^{\mathrm{diff},1}(\varphi_+,\varphi_+)$, we define v by extending φ_+ linearly across the vertical line segments $-w \le y \le w, x \in I$, i.e. to $\varphi_+(x)y/w$. This gives the leading term now, as the derivative there is φ_+/w. In fact for any odd vector field a of L^2-norm 1 and for which the derivatives are sufficiently small compared to w, the norm has the same leading term:

$$G_C^{\mathrm{diff},1}(a,a) \approx \sqrt{A}/w.$$

To estimate $G_C^{\mathrm{diff},2}(\varphi_+,\varphi_+)$, we need a smoother extension across the interior of C. We can take $\varphi_+(x).\frac{3}{2}(\frac{y}{w} - \frac{1}{3}(\frac{y}{w})^3)$. Computing the square integral of the second derivative, we get the table entry $G_C^{\mathrm{diff},2}(a,a) \approx (\sqrt{A}/w)^3$.

To estimate $G_C^{\mathrm{diff},k+1}(\varphi_x,\varphi_x)$, we now take v to be

$$v = c(k,A,w)\left[\left(\tfrac{|x|}{\sqrt{A}}\right)^k K_k\left(\tfrac{|x|}{\sqrt{A}}\right) * \chi_D\right]\frac{\partial}{\partial x}.$$

where D is the disk of radius w containing the arc making up the right hand end of C, and where $c(k,A,w)$ is a constant to be specified later. The function $\frac{1}{2\pi k!A}\left(\frac{|x|}{\sqrt{A}}\right)^k K_k\left(\frac{|x|}{\sqrt{A}}\right)$ is the fundamental solution of $(I-A\triangle)^{k+1}$ and is C^1 for $k>0$ but with a log pole at 0 for $k=0$. Thus:

$$(I-A\triangle)^{k+1}v = 2\pi k! A c(k,A,w)\chi_D.\frac{\partial}{\partial x}$$

while, up to upper and lower bounds depending only on k, the restriction of v to the disk D itself is equal to $\log(\sqrt{A}/w)c(0,A,w)w^2$ if $k=0$ and simply $c(k,A,w)w^2$ for $k>0$. By symmetry v is also constant on the boundary of D and thus v extends φ_x if we take $c(0,A,w) = c_0/\log(\sqrt{A}/w)w^{5/2}$ if

$k = 0$ and $c(k, A, w) = c_k / w^{5/2}$ if $k > 0$ (constants c_k depending only on k). Computing the H^k-norm of v, we get the last table entries.

Summarizing, we can say that the large norm of φ_x is what characterizes G^A; the large norms of φ_+ characterize G^{diff}; and the rate of growth in frequency of the norm of φ_f distinguishes all 4 norms.

Bibliography

[1] *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* Edited by Milton Abramowitz and Irene A. Stegun. Reprint of the 1972 edition. Dover Publications, Inc., New York, 1992

[2] M. Adams, T. Ratiu, R. Schmid. A Lie group structure for pseudo differential operators. *Math. Ann.* 273 (1986), 529–551.

[3] L. Ambrosio, N. Gigli, G. Savaré. Gradient flows with metric and differentiable structures, and applications to the Wasserstein space. *Rend. Mat. Acc. Lincei* s.9, 15 (2004), 327–343.

[4] L. Ambrosio, N. Gigli, G. Savaré. *Gradient Flows In Metric Spaces and in the Space of Probability Measures.* Lectures in Mathematics. ETH Zürich. Birkhäuser, Basel, 2005.

[5] V. Arnold and B. Khesin. *Topological methods in fluid dynamics.* Applied Math. Sciences. Springer-Verlag, 1998.

[6] 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.

[7] V.I. Arnold. An a priori estimate in the theory of hydrodynamic stability. *Izvestia Vyssh. Uchebn. Zaved. Matematicka*, 54,5 (1966), 3–5. Russian.

[8] 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.

[9] V.I. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, New York, Heidelberg, 1978. Graduate Texts in Math. 60.

[10] M.F. Beg, M.I. Miller, A. Trouve, L. Younes. Computing large deformation metric mappings via geodesic flows of diffeomorphisms. *International Journal of Computer Vision* 61,2 (2005), 139-157.

[11] P. Belhumeur. A Bayesian Approach to Binocular stereopsis. *Int. J. of Computer Vision*, 19:236-260, 1996.

[12] E. Binz. Two natural metrics and their covariant derivatives on a manifold of embeddings. *Monatsh. Math.*, 89 (1980), 275–288.

[13] 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.

[14] R. Bott. On the characteristic classes of groups of diffeomorphisms. *Enseign. Math.*, 23, 1977.

[15] J.-D. Benamou, Y. Brenier. A computational fluid mechanics solution to the Monge Kantorovich mass transfer problem. *Numer. Math.* 84 (2000), 375–393

[16] Jean-Luc Brylinski. *Loops spaces, characteristic classes, and geometric quantization.* Birkhäuser, Boston, Basel, Berlin, 1993. Progress in Math. 107.

[17] J. Burgers. A mathematical model illustrating the theory of turbulence. *Adv. Appl. Mech.*, 1 (1948), 171–199.

[18] R. Camassa and D. Holm. An integrable shallow water equation with peaked solutions. *Phys. Rev. Lett.*, 71 (1993), 1661–1664.

[19] Vicente Cervera, Francisca Mascaro, and Peter W. Michor. The orbit structure of the action of the diffeomorphism group on the space of immersions. *Diff. Geom. Appl.*, 1:391–401, 1991.

[20] A. Constantin, T. Kappeler, B. Kolev, and P. Topalov. On geodesic exponential maps of the Virasoro group. Preprint 13-2004, University of Zürich

[21] A. Constantin and B. Kolev. Geodesic flow on the diffeomorphism group of the circle. *Comm. Math. Helv.*, 78 (2003), 787–804.

[22] K. Deimling. *Ordinary differential equations in Banach spaces.* Lecture Notes in Mathematics, Vol. 596. Springer-Verlag, Berlin-New York, 1977.

[23] J. Eichhorn. The manifold structure of maps between open manifolds. Ann. Global Anal. Geom. 11 (1993), no. 3, 253–300.

[24] 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

[25] J. Eichhorn and J. Fricke. The module structure theorem for Sobolev spaces on open manifolds. Math. Nachr. 194 (1998), 35–47.

[26] D.B.A. Epstein. The simplicity of certain groups of homeomorphisms. *Compositio Math.*, 22 (1970), 165–173.

[27] C.F. Faà di Bruno. Note sur une nouvelle formule du calcul différentielle. *Quart. J. Math.*, 1 (1855), 359–360.

[28] A. Frölicher and A. Kriegl. Linear spaces and differentiation theory. Pure and Applied Mathematics, J. Wiley, Chichester 1988.

[29] D. Fuks, *Cohomology of infinite dimensional Lie algebras.* Nauka, Moscow, 1984 (Russian). Transl. English, Contemporary Soviet Mathematics. Consultants Bureau (Plenum Press), New York, 1986.

[30] I.M. Gelfand and I.Y. Dorfman. Hamiltonian operators and the algebraic structures connected with them. *Funct. Anal. Appl.*, 13 (1979), 13–30.

[31] 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.

[32] 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.

[33] W.C. Graustein. A new form of the four vertex theorem. *Mh. Math. Physics*, 43:381–384, 1936.

[34] W.C. Graustein. Extensions of the four vertex theorem. *Trans. AMS*, 41:9–23, 1937.

[35] Gerd Kainz. A note on the manifold of immersions and its Riemannian curvature. *Monatshefte für Mathematik*, 98 (1984), 211–217.

[36] 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.

[37] A.A. Kirillov. The orbits of the group of diffeomorphisms of the circle, and local Lie superalgebras. *Funct. Anal. Appl.*, 15 (1981), 135–136.

[38] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry. Vol. I. Vol. II.* J. Wiley-Intersci-ence, 1963, 1969.

[39] H. Kodama and P. W. Michor. The homotopy type of the space of degree 0 immersed curves. *Revista Matemtica Complutense* 19 (2006), no. 1, 227-234.

[40] S. Kouranbaeva. The Camassa-holm equation on as a geodesic flow on the diffeomorphism group. *J. Math. Phys.*, 40 (1999), 857–868.

[41] KM90 A. Kriegl and P. W. Michor. A convenient setting for real analytic mappings. *Acta Mathematica*, 165 (1990), 105–159.

[42] A. Kriegl and P. W. Michor. *The Convenient Setting for Global Analysis.* AMS, Providence, 1997. Surveys and Monographs 53.

[43] A. Kriegl and P. W. Michor. Regular infinite dimensional Lie groups. *J. Lie Theory (http://ww w.emis.de/journals/JLT)*, 7,1, 1997.

[44] 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.

[45] J Marsden and T. Ratiu. *Introduction to mechanics and symmetry.* Springer-Verlag, New York, Berlin, Heidelberg, 1994.

[46] John N. Mather. Commutators of diffeomorphisms I, II. *Comm. Math. Helv.*, 49 (1974), 512–528.

[47] A. Mennucci, A. Yezzi. Metrics in the space of curves. arXiv:math.DG/0412454

[48] A. Mennucci, A. Yezzi, G. Sundaramoorthi. Sobolev–type metrics in the space of curves. arXiv:math.DG/0605017.

[49] 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.

[50] P. W. Michor. Manifolds of differentiable mappings. Shiva Mathematics Series, 3. Shiva Publishing Ltd., Nantwich, 1980.

[51] Peter W. Michor. The cohomology of the diffeomorphism group is a Gelfand-fuks cohomology. *Suppl. Rendiconti del Circolo Matematico di Palermo, Serie II,*, 14 (1987) 235– 246

[52] 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.

[53] P. W. Michor. Some Geometric Evolution Equations Arising as Geodesic Equations on Groups of Diffeomorphism, Including the Hamiltonian Approach. In: Phase Space Analysis of Partial Differential Equations. Progress in Non Linear Differential Equations and Their Applications, Vol. 69. Birkhäuser Boston, 2006.

[54] 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.

[55] 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.

[56] P.W. Michor and T. Ratiu. Geometry of the Virasoro-bott group. *J. Lie Theory*, 8 (1998), 293–309.

[57] M.I. Miller and L. Younes. Group actions, homeomorphisms, and matching: A general framework. *International Journal of Computer Vision* 41, No 1/2 (2001), 61-84.

[58] M.I. Miller and A. Trouve and L. Younes On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of biomedical Engineering* 4 (2002), 375-405.

[59] W. Mio and A. Srivastava. Elastic-String Models for Representation and Analysis of Planar Shapes. *Proceedings of the IEEE Computer Society International Conference on Computer Vision and Pattern Recognition (CVPR)*, Washington, DC, 2004.

[60] W. Mio, A. Srivastava and S. Joshi. On the Shape of Plane Elastic Curves. *preprint*, Florida State University (Math/Stat/CS departments), 2005.

[61] G. Misiołek. Conjugate points in the Bott-Virasoro group and the KdV equation. *Proc. Amer. Math. Soc.* 125 (1997), 935–940.

[62] G. Misiołek. Classical solutions of the periodic Camassa-holm equation. *GAFA*, 12 (2002), 1080–1104.

[63] G. Misiołek A shallow water equation as a geodesic flow on the Bott-virasoro group. *J. Geom. Phys.*, 24 (1998), 203–208.

[64] J.P. Ortega and T. Ratiu. *Momentum maps and Hamiltonian reduction*. Birkhäuser, Boston, 2004. Progress in Mathematics 222.

[65] V.Y. Ovsienko and B.A. Khesin. Korteweg–de Vries superequations as an Euler equation. *Funct. Anal. Appl.*, 21 (1987), 329–331.

[66] R. Palais. On the existence of slices for actions of non-compact Lie groups *Ann. of Math.* (2) 73:295–323, 1961.

[67] A. Pressley and G. Segal. *Loop groups*. Oxford University Press, 1986. Oxford Mathematical Monographs.

[68] T. R. Ramadas. On the space of maps inducing isomorphic connections. *Ann. Inst. Fourier* 32 (1982), 263–276.

[69] T. Ratiu. The motion of the n-dimensional rigid body. *Indiana Univ. Math. J.*, 29 (1980), 609–629.

[70] G. Segal. The geometry of the Kdv equation. *Internat. J. Modern Phys. A* , 6 (1991), 2859–2869.

[71] J. Shah. H^0-type Riemannian metrics on the space of planar curves. arXiv:math.DG/0510192

[72] W. Thurston. Foliations and groups of diffeomorphisms. *Bull. AMS*, 80 (1974), 304–307.

[73] A. Trouvé. Infinite dimensional group actions and pattern recognition. Technical report, DMI, Ecole Normale Superieure, 1995.

[74] A. Trouvé and L. Younes. Diffeomorphic Matching in 1D. *Proc. ECCV 2000*, D. Vernon editor, 2000.

[75] A. Trouvé and L. Younes. Local Analysis on a Shape Manifold. preprint, Univ. de Paris 13, 2002. available at the URL ”citeseer.ist.psu.edu/608827.html”.

[76] A. Trouv and L. Younes. Local Geometry of Deformable Templates *SIAM Journal of Mathematical Analysis* 37,1 (2005), 17–59.

[77] A. Weinstein. Symplectic manifolds and their Lagrangian manifolds. *Advances in Math.*, 6 (1971), 329–345.

[78] H. Whitney. Regular families of curves I, II. *Proc. Nat. Acad. Sci USA*, 18:275–278, 340–342, 1932.

[79] L. Younes. Computable elastic distances between shapes. *SIAM J. Appl. Math.* 58 (1998), 565–586.