Theorie kritischer Punkte für Symmetrien mit Fixpunkten

Critical Point Theory for Symmetries with Fixed Points

Dissertation zur Erlangung des
Doktorgrades (Dr. rer. nat.)
des Fachbereiches Mathematik und
Informatik, Physik, Geographie
der Justus-Liebig-Universität Gießen

Vorgelegt von
Stefan Born
Gießen
2010

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis“ niedergelegt sind, eingehalten.

Contents

1. Introduction

1.1. A class of variational problems with symmetries

Since Poincaré the investigation of periodical orbits has played an essential role in understanding dynamical systems. The calculus of variations allows to characterise periodical solutions of Hamiltonian or Lagrangian systems as critical points of a continuously differentiable action functional f on an appropriate Banach manifold X.

If the functional satisfies the Palais-Smale condition, sublevel sets

$$X^a := f^{-1}(]-\infty, a]) \quad \text{and} \quad X^b := f^{-1}(]-\infty, b])$$

are homotopy equivalent, unless there is a critical value in $]a, b]$. Therefore any homotopy invariant can be used to detect critical values and critical sets and to describe the topology change that occurs at the critical levels, which may lead to estimates for the 'size' of the critical set.

For non-degenerate critical points with finite index the rank of (co)homology groups yields useful invariants. This leads to Morse theory, which estimates the number of critical points (from below) by the Betti numbers of X.

If we deal with possibly non-degenerate critical points, we need different invariants, namely category, cup-length or, in the case of a G-space, a G-index like the Krasnosel'skii genus or the length. A Lusternik-Schnirelman type theory estimates the number of critical points (from below) by any of these invariants.

Of course this is a rather naive sketch of the idea, as the methods often will not compare the overall homotopy types of the sublevel sets, but rather measure topology changes in an indirect way, for example by a min-max characterisation of critical values.

Let us consider a simple example, the torus T^n, for which the sum of the Betti numbers is 2^n and the Lusternik-Schnirelman category is $n + 1$. Hence a differentiable function on the torus T^n has at least 2^n critical points, if they are nondegenerate, and at least $n + 1$ critical points without that restriction. For a

Hamiltonian system on T^*T^n that is of a certain quadratic form on the fibres, Conley and Zehnder [CZ83] were the first to prove the same estimates for the number of T-periodic solutions, as the problem can be reduced to critical point theory of a functional on finite dimensional vector bundles over T^n, which have the same Betti numbers and cup-length as T^n. In a similar way Conley and Zehnder proved, that a Hamiltonian system on T^{2n} has at least 2^{2n} nondegenerate and at least $2n + 1$ possibly degenerate T-periodic solutions. (For similar results by different methods cf. for instance [Rab88], [Fel92]).

The group $G = \mathbb{Z}_2$ operates on T^n with 2^n fixed points, which induces an action with 2^n fixed points on T^*T^n. A T-periodic Hamiltonian system on T^*M which is symmetric with respect to the G-operation, has 2^n constant solutions with values precisely the fixed points. The same is valid for a symmetric Lagrangian system on T^n. Thus the above mentioned multiplicity results will not provide any interesting solutions beyond the trivial ones.

One would like to apply multiplicity results for symmetric variational problems to find more solutions. Unfortunately, these results cannot be used out of the box, as they rely on G-index theories, which fail to give nice results in the presence of fixed points for the G-action, because these fixed points usually have infinite G-index.

Consider, for instance, the Borel cohomology of a $\mathbb{Z}_2$-space X

$$H^*_G(X) = H^*(X \times_G EG; \mathbb{Z}_2) \ .$$

As we can pull back $H^*(BG; \mathbb{Z}_2)$ via $p : X \times_G EG \to BG$, we have an $H^*(BG; \mathbb{Z}_2)$-module structure on $H^*_G(X)$. $H^*(BG; \mathbb{Z}_2)$ is a polynomial algebra over $\mathbb{Z}_2$ with one generator ω in degree 1. The length of a $\mathbb{Z}_2$-space X (as defined in Example 4.4, [Bar93]) is

$$\ell(X) := \min\{k \geq 0 \mid \omega^k \, 1_X = 0\} \ .$$

The Borel cohomology of a fixed point is isomorphic to $H^*(BG)$ via p^*, it is a free $H^*(BG, \mathbb{Z}_2)$-algebra, and $H^*(BG) = \mathbb{Z}_2[\omega]$ contains arbitrary large nonvanishing products, so $\ell(pt) = \infty$. For the purpose of critical point theory it is a sorry sight that a rich product structure, which might otherwise detect critical points, can be swallowed by a single fixed point. However, the length only makes use of the $H^*(BG)$-module structure, for which all fixed points behave in the same way. More information is contained in $H^*_G(X)$.

This lies at the heart of an observation by Bartsch and Wang [BW97a],[BW97c], [BW97b]. Let us sketch their idea.

i) For a fixed point $p \in T^n$ there is a cohomology class $\mu \in h^n_G(T^n, T^n \backslash \{p\})$, whose restriction to $\{p\}$ is $\omega^n \, 1_p$ and whose restriction to each of the other fixed points is zero. Now suppose that f is defined on a G-vector bundle $\pi : T^n \times V \to T^n$ such that the zero section $\sigma : T^n \to T^n \times V$ is equivariant. Let us denote $T^n \times V$ by X. The restriction of $\pi^*(\mu)$ to $\sigma(p)$ is $\omega^n \, 1_{\sigma(p)}$, and the restriction to $\sigma(q)$ is zero for any other fixed point $q \in (T^n)^G$. This class *recognises* the fixed point it stems from.

ii) For a k-dimensional representation $G \times \mathbb{R}^k \to \mathbb{R}^k$ without trivial summands the restriction

$$H^*_G(D^k) \to H^*_G(S^{k-1})$$

is surjective in all degrees and an isomorphism in degrees $\leq k-1$. If there is an equivariant map

$$(D^k, S^{k-1}) \to (X, X^{f(\sigma(p))-\epsilon}) \ ,$$

that maps 0 to $\sigma(p)$, we conclude that the restriction of $\omega^i \pi^*(\mu)$ to $X^{f(\sigma(p))-\epsilon}$ is not zero if $i + n \leq k - 1$.

iii) Again for a k-dimensional representation, the restriction

$$H^*_G(D^k, S^{k-1}) \to H^*_G(D^k)$$

is injective in all degrees. Therefore a relative class in $H^*_G(D^k, S^{k-1})$ is zero, if it restricts to zero in $H^*_G(D^k)$. Suppose the class $\omega^i \pi^*(\mu)$ vanishes at level c, that is,

$$\omega^i \pi^*(\mu)|_{X^{c+\epsilon}} \neq 0 \quad \text{and} \quad \omega^i \pi^*(\mu)|_{X^{c-\epsilon}} = 0$$

for all $\epsilon > 0$. Suppose furthermore that a fixed point $x \neq \sigma(p)$ at level c is non-degenerate. Then a relative class $\xi \in H^*_G(X^{c+\epsilon}, X^{c-\epsilon})$ must restrict to zero on $H^*_G(U, U^{c-\epsilon})$ for a small ball U with centre x.

By (i) and (ii) they construct $(k - n)$ classes that are nonzero on $X^{f(\sigma(p))-\epsilon}$. By (iii) they conclude that the critical set at vanishing levels for these classes must contain some non-fixed points. A multiplicity result can be obtained, once it is proved that some of these classes eventually vanish, that is, below some level c. If f is bounded below, all classes eventually vanish, there must be at least $k - n$ G-orbits of nontrivial critical points. If, however, f behaves at infinity like a non-degenerate quadratic form of index n^- on the fibres, all classes with degree $\geq n + n^-$ eventually vanish, so we find $k - (n + n^-)$ orbits of nontrivial critical points.

In this way Bartsch and Wang obtain a series of theorems for Lagrangian and Hamiltonian system on a torus T^{2n} or a cotangent bundle T^*T^n. If a constant solution that corresponds to a fixed point of T^n has a Maslov (or Conley-Zehnder) index $i > n$, the difference $i - n$ is a lower bound for the number of G-orbits of nontrivial solutions.

The requirements of this theory place it between Morse theory and Lusternik-Schnirelman theory. The trivial solutions below $\sigma(p)$ have to be *non-degenerate*, in order to enable a symmetrical critical point theory for *degenerate* critical points.

1.2. The scope of this thesis

1.2.1. Critical point theory for symmetries with fixed points

The thesis gives partial answers to a series of questions which are raised by the work of Bartsch and Wang.

- If there are *less* fixed points below $\sigma(p)$, the theory should predict *more* nontrivial critical points. After all, if there is *no* fixed point below $\sigma(p)$, we can use Lusternik-Schnirelman theory for a G-index to detect critical points. In the above setting, for a functional bounded below, this would yield k orbits of critical points, not only $k - n$. How many or which fixed points may lie below $\sigma(p)$, in order to obtain an estimate between k and $k - n$?
- Can we make this construction work for manifolds other than T^n?
- For which groups does such a theory work?

In order to address the first two issues it was necessary to study Borel cohomology of a G-manifold with the goal to understand how cohomology classes restrict to the fixed point set. For an isolated fixed point p and a closed and open subset F of the fixed point set X^G we defined the quantity

$$\sigma(x, F)$$

as the minimal degree of a Borel cohomology class μ that separates x and F, in the sense that $\mu|_x \neq 0$ and $\mu|_C = 0$ for every component C of F. *Please note, that the difference $n - \sigma(x, F)$ measures the improvement with respect to Bartsch and Wang.*

As above, it is true for a G-manifold with a semifree (SF) or cohomologically semifree (CSF) G-action, that an isolated G-fixed point can be separated from $X^G - \{p\}$ by a class of degree n (thus yielding the estimate of $k - n$ fixed points in our example). This corresponds to the algebraic fact that $H^d_G(M) \to H^d_G(M^G)$ is surjective in degree $d \geq n - \dim G$. We can prescribe classes at the fixed points and extend them to a global class that restricts to these classes.

In degrees $d < n - \dim G$ the restriction is not surjective in general, so we would either like to characterise the image of the restriction map, or at least to have some sufficient criteria for membership in the image. We dealt with both cases.

i) (Section 2.1.2) For manifolds that are totally nonhomologous to zero (TNHZ), the more ambitious goal of understanding the image of the restriction map can be pursued. In Section 2.1.2 we restate and reproof a well-known criterion for a G-manifold to be (TNHZ) (under certain assumptions on $H^*_G(BG)$). The criterion is precisely that the sum of the Betti-numbers of M is equal to the sum of the Betti numbers of M^G. In the case of isolated fixed points this is precisely the situation we encountered, where the number of solutions predicted by Morse theory is exhausted by trivial solutions.

   A central and perhaps most interesting part of the thesis is Example 2.1.34, where we consider a class of (TNHZ) spaces with isolated fixed points, which comprises the torus with 2^n fixed points. For this class we construct a cohomology extension of the fibre that allows to describe precisely the restriction of each Borel cohomology class to the fixed point set. The 'separation problem' is boiled down to an algebraic problem. Given p and F we can enumerate all classes of degree d and check whether any of them separates p from F.

   For the special case of the torus with $\mathbb{Z}_2$-action we tried to get beyond enumerating. The algebraic problem can be reshaped as a problem of algebraic geometry over the field $\mathbb{Z}_2$. The fixed points can be identified with the elements of the vector space $V := (\mathbb{Z}_2)^n$, and the Borel cohomology $H^*_G(M)$ in degrees $0, \ldots, n$ can be identified with the ring of square free polynomials

$$\mathbb{Z}_2[X_1, \ldots, X_n]/(X_i^2 = X_i, i = 1, \ldots, n) \ .$$

   p and F can be separated in degree d if and only if there is a squarefree polynomial P of degree d with with $P(p) = 1$ and $P(x) = 0$ for all $x \in F$.

   For $d = 1$ this translates to whether F is contained in an affine hyperplane that does not contain p. For $d = 2$, whether F is contained in a quadric that does not contain p, and so on. Of course, a quadric over finite fields is not quite what we are used to from quadrics over $\mathbb{R}$ and $\mathbb{C}$.

   What can we say about $\sigma(p, F)$, if we are only given the cardinality $N := |F|$? We define the quantity

$$\sigma(N) := \max\{\sigma(p, F) \mid |F| = N, p \notin F\}$$

We prove a series of estimates of $\sigma(N)$ which are not optimal. A few examples: $\sigma(1..2) \leq 1$, $\sigma(3..5) \leq 2$, $\sigma(N) \leq [N/2] + 1$, $\sigma(2^n - 2) \leq n - 1$.

ii) (Section 2.1.3) Few G-manifolds can be expected to be (TNHZ). We found a geometrical criterion guaranteeing that p can be separated from F in degree d.

p is contained in a G-invariant submanifold with codimension d, which does not intersect F and whose normal bundle is orientable for Borel cohomology.

The G-Thom class of such a submanifold provides a separating class for x and F, if $H^*(BG)$ has unbounded cohomology (Proposition 2.1.40).

Thom classes enter the picture once more when it comes to generalising the third part of Bartsch's and Wang's idea. A cohomology class that is zero when restricted to a an isolated non-degenerate fixed point x, must be zero on $(U, U^{c-\epsilon})$ for a distinguished neighbourhood U of x, thus rendering this fixed point 'invisible' for the Lusternik-Schnirelman argument.

A component of the fixed point set, which consists of critical points, and which is normally non-degenerate with respect to f, such that the subbundle of the normal bundle defined by the negative eigenspaces of the Hessian is H_G^*-orientable, becomes invisible in the same way, if $H^*(BG)$ contains a free polynomial generator (Proposition 2.1.46).

In Section 2.2 we deduce abstract critical point theorems that generalise Bartsch and Wang to arbitrary manifolds and improve their estimates in the case of the torus by the difference $n - \sigma(p, F)$.

The requirements on the groups, however, leave few choices. As we need a free polynomial generator and (SF) or (CSF), the general purpose theorems (Theorem 2.2.28 and Theorem 2.2.31) are restricted to $G = S^1$ and $G = \mathbb{Z}_p$ (p prime), for which these conditions are immediate. Similar theorems should be valid under the conditions (SF) or (CSF), if $H_G^*(BG)$ contains a free polynomial generator, yet I couldn't think of any relevant examples.

1.2.2. Lagrangian systems

Lagrangian systems were meant as a mere example for the abstract theorems. But in the process of writing down the preliminaries to the application of these theorems, I started working on details and minor contradictions I found in the literature. For example, even in most recent publications there are occasional errors as to the differentiability of the Lagrangian action functional for a non-classical Lagrangian. (The action functional for a Lagrangian that is not quadratic on the fibres of TM, but satisfies quadratic growth conditions, is not twice Fréchet differentiable.) It might have been wiser to restrict the application to classical Lagrangians and quote the properties of the action functional from the literature. As it is, Chapter 3 consists mainly of well-known facts about the action functional with some new proofs. Those versed in Lagrangian systems may skip most of it and jump to the multiplicity result in Section 3.3.

As the functional is not C^2 and the abstract theorems require non-degenerate critical points, a certain apparatus was necessary. The definition of a topologically non-degenerate critical point x in Section 2.2 is based on the existence of *distinguished* neighbourhoods U of x, such that $(U, U^{f(x)-\epsilon})$ has the homotopy type of (D^k, S^{k-1}). Although our functional is not C^2, it is possible to define a non-degenerate second derivative in a weaker sense and a smooth pseudo-gradient field, which allows to construct nice distinguished neighbourhoods (here we follow Ghoussoub and Chang [CG96]). The construction of the smooth pseudo-gradient fields is based on a recent idea by Abbondandolo and Schwarz [AS09].

The exposition follows the view that a Lagrangian system on a compact Riemannian manifold is a geometric object. If the Lagrangian is the kinetic energy, the solutions are geodesics. For a general Lagrangian the solutions are geodesics perturbed by forces. This view led to the attempt to give intrinsic versions of properties and proofs, without charts, wherever it was feasible. This is one of the reasons why this text makes a case for Hadamard differentiability in the calculus of variations. Gâteaux differentiability is defined with respect to a chart only. It is a reasonable intermediary step in proofs of Fréchet differentiability, but is not a property of functions on a manifold. On the contrary, Gâteaux differentiability with respect to *any* chart is a geometrical property. Intrinsic definitions of differentiability would try to avoid charts and to define the derivative by means of curves that represent tangent vectors. The property 'Gâteaux differentiability with respect to every chart' can be established in this intrinsic way. Hadamard differentiability is a stronger type of differentiability that can be given in an intrinsical way. The gradient of our functional happens to be differentiable in the sense of Hadamard, which implies that it is Fréchet differentiable on any compactly embedded submanifold. In the end we didn't need this property, but hopefully it may be of some use for other problems.

The application of our abstract critical point theorems to Lagrangian systems on compact manifolds in Theorem 3.3.2 generalises Theorem 3.13. of [BW97b] and improves the result in the case of a torus.

Ich danke meinem Betreuer Thomas Bartsch für seine Unterstützung und seine Geduld.
Außerdem danke ich den Kollegen und den Freunden, die auf ganz verschiedene Weise zur Entstehung dieser Arbeit beigetrage haben, Lars Schmidt-Thieme, Barbara Schmidt-Thieme, Mike Scherfner, Matthias Plaue, Dirk Ferus, Bernhard Lani-Wayda, Benjamin Düsberg, Jana Sierk, Anke Eisenkopf, Simone Rajilic, Dunjo Rajilic und Jella Hauß. Albrecht Beutelspacher danke ich für ein Gespräch über endliche Körper.
Besonders danke ich meinen Eltern.

2. Symmetric critical point theory

2.1. Equivariant cohomology

2.1.1. Preliminaries and definitions

Remark 2.1.1 Let us make a preliminary remark for the information of the reader. The exposition of this subsection contains mainly well-known facts as presented in [AP93],[tD87],[Bre72]. It is, however, intended to be a self-contained exposition and contains special versions of the theorems adapted for our purpose, which could not be found in the literature. The exposition is notably different from [tD87]. We do not derive injectivity or isomorphy of $h^*_G(X) \to h^*_G(X^G)$ in certain degrees from the typical localization results $S^{-1}h^*_G(X) \to S^{-1}h^*_G(X^G)$, but rather prove them directly. We preferred this approach, as we need the degree information and would otherwise have to recover it from the localizations. Another aspect of this exposition is its attempt to make sure that the theorems are valid for the category of spaces we have to consider in the applications. The theorems to be found in the literature often consider classes of spaces different from the ones we needed. The precise assumptions of our "equivariant Whitehead theorem" proposition 2.1.4 could be found nowhere. $\diamond$

G will always be a compact Lie group. A frequent class of Lie groups for our purposes is $\mathcal{G} := \{\mathbb{Z}_2, \mathbb{Z}_p \text{ } (p \text{ odd}), S^1\}$.

Let $\pi : EG \to BG$ a model for the universal numerable principal G-bundle, e. g. the Milnor model (s. Husemoller [Hus93]). The total space EG is contractible and BG is path-connected and has the homotopy type of a CW-complex. We will sometimes consider models for $EG \to BG$ with special properties, but in general any numerable principal G-bundle $E \to B$ with contractible E can serve as a model, s. Dold [Dol63].

h^* be a **general cohomology theory** on the category of pairs of topological spaces with values in R-modules for some commutative ring with unit R.

That means, it has the excision, homotopy and exactness properties. If the dimension axiom ($h^k = 0$ for $k < 0$) is fulfilled, we have an **ordinary cohomology theory**. When we use ordinary cohomology, we will usually take Čech-cohomology.

The corresponding reduced cohomology theory will be denoted by $\tilde{h}^*$.

If h^* is a cohomology theory with values in k-vector spaces for some field k, we say that a **Künneth theorem** is valid for h^* *under some condition (C) on spaces* (X, A) *and* (Y, B), if

$$h^*(X \times Y, X \times B \cup A \times Y) \cong h^*(X, A) \otimes_k h^*(Y, B) \ .$$

If not mentioned otherwise we suppose:

(C) $h^q(X)$ is finitely generated for each q or $h^q(Y)$ is finitely generated for each q.

Thus singular cohomology satisfies a Künneth theorem under condition (C). For sheaf cohomology with compact supports (with values in k-vector spaces), Čech-cohomology on pairs of compact Hausdorff spaces, or Alexander-Spanier cohomology on pairs (X, A) of locally compact spaces with $A \subset X$ closed, we have a Künneth theorem.

For any left G-space X we can apply the Borel construction

Definition 2.1.2

$$X_G := EG \times_G X = (EG \times G)/\text{diagonal action of} \ \ G \ .$$

which defines a functor from the category of (pairs of) G-spaces to the category of (pairs of) topological spaces.

Now we obtain Borel-(h^*, G)-cohomology as a functor from the category of (pairs of) G-spaces to the category of abelian groups

Definition 2.1.3 *Let* (X, A)*,* $A \subset X$ *be a pair of* G*-spaces.*

$$h^*_G(X) := h^*(X_G) \ ,$$
$$h^*_G(X, A) := h^*(X_G, A_G) \ .$$

Borel-cohomology has the properties of a *cohomology theory* in the category of G-spaces. We give a list of the essential properties that is neither complete nor optimal. Let (X, A), (X', A') be pairs of G-spaces.

i) (HTP) A G-invariant continuous map $f : (X, A) \to (X', A')$ induces a homomorphism $h^*_G(X', A') \to h^*_G(X, A)$ that depends only on the G-homotopy class of f.

ii) (SEQ) There is a sequence of natural transformations

$$\delta^n : h^n_G(A) \to h^{n+1}_G(X, A) ,$$

such that we have a long exact sequence

$$\cdots \to h^n_G(X, A) \to h^n_G(X) \to h^n_G(A) \xrightarrow{\delta^n} h^{n+1}_G(X, A) \to \cdots .$$

iii) (EXC) Is U an open G-invariant subset of A s. th. $\overline{U} \subset \mathrm{int}(A)$, the inclusion $j : (X - U, A - U) \to (X, A)$ induces an *excision* isomorphism

$$h^*_G(X, A) \to h^*_G(X - U, A - U)$$

iv) (MV) For a triad $(X; X_0, X_1)$ of G-spaces, X_0 and X_1 open in X, $X = X_0 \cup X_1$ we have a family of natural transformations $\delta^n : h^n_G(X_0 \cap X_1) \to h^{n+1}_G(X)$ and an exact Mayer-Vietoris-Sequence

$$\cdots \to h^n_G(X) \to h^n_G(X_0) \oplus h^n_G(X_1) \to h^n_G(X_0 \cap X_1) \xrightarrow{\delta^n} h^{n+1}_G(X) \to \cdots .$$

If h^* satisfies the strong excision property (as Čech or Alexander Spanier cohomologier on paracompact Hausdorff spaces), h^*_G will satisfy a strong excision property

(SEXC) Let (X, A) and (Y, B) be pairs of G-spaces, with X and Y paracompact Hausdorff and A and B closed. Let $f : (X, A) \to (Y, B)$ a closed continuous G-equivariant map such that f induces a one-to-one map of $X - A$ onto $Y - B$. Then the induced map

$$f^* : h^*_G(Y, B) \to h^*_G(X, A)$$

is an isomorphism.

Often it is useful to approximate EG by finite CW complexes and to replace a space by a weakly homotopy equivalent space. We have the following:

Proposition 2.1.4 *([Ful05]) Let G be a compact Lie group and $E \to B$ a G-principal bundle with n-connected E, that is $\pi_i(E) = 0$ for $0 < i \leq n$. Let X be a path-connected G-space, and suppose E is a G-CW complex. Suppose furthermore there is a path-connected G-space X' and a continuous map $f : X' \to X$ such that $\pi_i(f) : \pi_i(X') \to \pi(X)$ is an isomorphism for $i = 0 \dots n$ and an epimorphism for $i = n + 1$.*

Then there is a canonical isomorphism for singular cohomology H^* with coefficients in some commutative ring R

$$H^i(EG \times_G X; R) \to H^i(E \times_G X'; R) ,$$

if $i \leq n$ (and an epimorphism for $i = n+1$).

If the spaces X, X' are locally G-contractible paracompact Hausdorff spaces, $E \times_G X$ and $EG \times_G X$ are locally contractible and their singular cohomology is isomorphic to Čech cohomology (or Alexander-Spanier cohomology) h^* with coefficients R, thus

$$h^i(EG \times_G X) \to h^i(E \times_G X') ,$$

is an isomorphism for $i \leq n$ (and an epimorphism for $i = n+1$).

Proof: The classifying map from $E \to E/G$ to the universal bundle defines a bundle map

$$\begin{tikzcd} E \times X' \arrow[r] \arrow[d] & EG \times X \arrow[d] \\ E \times_G X' \arrow[r] & EG \times_G X \end{tikzcd} ,$$

which induces a homorphism of the long exact homotopy sequences that correspond to the two bundles.

$$\begin{tikzcd} \pi_i(G) \arrow[r] \arrow[d] & \pi_i(E \times X) \arrow[r] \arrow[d] & \pi_i(E \times_G X) \arrow[r] \arrow[d] & \pi_{i-1}(G) \arrow[d] \\ \pi_i(G) \arrow[r] & \pi_i(EG \times X) \arrow[r] & \pi_i(EG \times_G X) \arrow[r] & \pi_{i-1}(G) \end{tikzcd}$$

As EG is contractible, $\pi_i(X)$ and $\pi_i(EG \times X)$ are isomorphic for all i via the inclusion of X and the projection on X. Now for $i \leq n$ we have $\pi_i(E \times X') \cong \pi_i(X') \cong \pi_i(X)$, and the composition

$$\pi_i(E \times X') \to \pi_i(EG \times X) \to \pi_i(X)$$

is an isomorphism, thus $\pi_i(E \times X) \cong \pi_i(EG \times X)$. The five lemma yields an isomorphism $\pi_i(E \times_G X') \to \pi_i(EG \times_G X)$ for $i \leq n$. (N.B. that the five lemma is valid for non-abelian groups as well, which is needed for the first bits of the sequence). For $i = n+1$ the homomorphism $\pi_i(E \times X') \to \pi_i(X')$ and hence $\pi_i(E \times X') \to \pi_i(EG \times X)$ is surjective. The surjectivity of $\pi_{n+1}(E \times_G X') \to \pi_{n+1}(EG \times_G X)$ follows once more from the five lemma. By the Whitehead theorem ([Spa66], Theorem 7.5.9) this induces an isomorphism in singular homology with integral coefficients in degree $i \leq n$ and an epimorphism for $i = n+1$. By the universal coefficient theorem for cohomology and the five lemma, we obtain an isomorphism in singular cohomology with coefficients R for $i \leq n$ (s. Corollary 3.4. of [Hat02]) and a monomorphism for $i = n+1$. (The epimorphism $H_{n+1}(E \times_G X') \to H_{n+1}(EG \times_G X)$ induces a monomorphism of $Hom(\cdot, R)$.)

The G-CW complex E is locally G-contractible. We can chose a model of EG, which is a G-CW complex. If X is locally G-contractible and paracompact, the product of X with a G-CW complex is paracompact ([Mor63], Theorem 1), and the quotients are locally contractible and paracompact (as G is compact, the projection to the quotient is closed and we can apply e. g. [Dug73], VIII, 2.6). Hence Čech-cohomology (or Alexander-Spanier cohomology) with coefficients in R is isomorphic to singular cohomology with coefficients in R. □

Remark 2.1.5 This is essentially a *Whitehead theorem* for cohomology with coefficients in R. There are various versions of such theorems in the literature, but just not the right one for our purpose, so we took the pains to follow the rather standard arguments. ◇

Remark 2.1.6 One might be tempted to reformulate the proposition as a statement about the inverse limit

$$\lim_{\leftarrow} h^*(EG^m \times_G X) \text{ for } m \to \infty \,.$$

This limit, however, is *not* isomorphic to $h^*_G(X)$, although there *is* an isomorphism in each degree. We obtain the direct product $\Pi_n h^n_G$ instead of the direct sum. For $G = \mathbb{Z}/2$ and X a point the inverse limit is the ring of infinite power series in one variable of degree 1, whereas the equivariant cohomology in this case is the polynomial ring in one variable. ◇

Remark 2.1.7 G-CW complexes are locally G-contractible. According to [Ill83] finite dimensional manifolds with a smooth G-action admit a G-CW decomposition.

The most general condition for X to be locally G-contractible is, as far as we know, that X is a G-ANR (for the category of metric G-spaces).

For our applications we will have to consider G-Banach manifolds or Hilbert manifolds X. Whenever the G-action $G \times X \to X$ is of class $C^{2,1}$, it is locally smooth, which means that there are linear tubes around each orbit. We conclude that X, which is metric, hence paracompact, is a G-ANR. (s. e.g. [CP91], Corollary B.5. and the remark.) Let us consider the most relevant examples for our purposes.

1. Suppose we have finite dimensional C^∞-manifolds M and N with $\dim M = m$ and $k - \frac{m}{p} > 0$. By the Whitney embedding theorem there is a C^∞-embedding of N into some $\mathbb{R}^\nu$. Then, the topological space

$$W^{k,p}(M,N) := W^{k,p}(M,\mathbb{R}^\nu) \cap C^0(M,N)$$

is well-defined and a closed submanifold of the Banach space $W^{k,p}(M,\mathbb{R}^\nu)$. For a G-manifold N the embedding can be constructed as a G-equivariant embedding into a G-representation $V \cong \mathbb{R}^\nu$ (Mostow, [Mos57]). The smooth G-operation $G \times V \to V$ induces a smooth G-operation on $W^{k,p}(M,N)$ via $(g \cdot f)(x) := g \cdot f(x)$.

2. If, however, we have a G-operation on M and define a G-operation on the Sobolev manifold $W^{k,p}(M,N)$ by $(g \cdot f)(x) := f(g^{-1} \cdot x)$, we encounter a more difficult situation. For a finite group G and compact M this G-operation is smooth. For an infinite compact Lie group the map $G \times W^{k,p} \to W^{k,p}$ is not even differentiable, in general. However, for fixed $g \in G$ the map $f \mapsto g \cdot f$ *is* smooth. $\diamond$

We will be interested in the equivariant Čech-cohomology of closed subsets of G-ANRs as well, which are not locally G-contractible in general. In this case, to obtain similar statements about Čech cohomology, we cannot use the Whitehead theorem, as singular and Čech cohomology behave behave quite differently. A different proof, however, yields the following proposition.

Proposition 2.1.8 *For any $i \in \mathbb{N}$, there is an $m = m(i) \in N$, such that the following is true:*

If EG^m is the m-G-skeleton of a G-CW model of the classifying space EG and X is a paracompact Hausdorff space. there is a canonical isomorphism for Čech-cohomology h^* with coefficients in some commutative ring R

$$h^i(EG \times_G X) \to h^i(EG^m \times_G X)\ .$$

The result is true for any cohomology theory h^* such that there is an N with $h^q(X) \cong 0$ for $q < N$.

We postpone the proof until we have introduced the Leray-Serre spectral sequence.

The last propositions allow us to calculate the ordinary equivariant cohomology of a G-space up to dimension $n-1$ by using any free n-connected G-space E as approximation of EG. These approximating spaces can be chosen to be compact. From this fact we derive nice continuity (or *tautness*) properties of equivariant Čech-cohomology.

Proposition 2.1.9 *(s. [tD87], Ex. III, (3.15)) Suppose X is hereditarily paracompact Hausdorff G-space (e.g. a G-ANR) and h^* is Čech-cohomology with some field coefficients and $n \in \mathbb{N}$.*

Then, for any locally closed G-set $A \subset X$

$$\lim_{\rightarrow} h^n_G(V) \cong h^n_G(A) \ ,$$

where we take the direct limit over all neighbourhoods $V \in \Lambda$ for some cofinal system Λ of G-neighbourhoods of A, for example all G-neighbourhoods, all open G-neighbourhoods or all closed G-neighbourhoods.

In particular this implies:

i) If $\alpha \in h^n_G(A)$ and U is a neighbourhood of A there is a neighbourhood $V \subset U$ of A, $V \in \Lambda$ and $\tilde{\alpha} \in h^n_G(V)$ such that $\tilde{\alpha}|_A = \alpha$.

ii) If $h^n_G(A)$ is finitely generated as a $h^*_G(pt)$-module, there is a neighbourhood $V \subset U$ of A, $V \in \Lambda$, such that the restriction

$$h^n_G(V) \to h^n_G(A)$$

is surjective.

iii) If $\alpha \in h^n_G(U)$ and $\alpha|_A = 0$ then there is a neighbourhood V of A with $V \subset U$, $V \in \Lambda$ and $\alpha|_V = 0$.

Proof: There is a CW approximation $EG^m \to EG^{m+1} \to \cdots EG$ of EG by finite and hence compact G-CW complexes EG^m such that EG^m is m-connected, so if we chose $m >> n$, Proposition 2.1.8 allows us to calculate $h^n_G(A)$ as $h^n(EG^m \times_G A)$.

As $EG^m \times_G A$ is tautly embedded in $EG^m \times_G X$ for h^n, if we let vary open neigbourhoods $\tilde{U}$ of $EG^m \times_G A$ we have

$$\lim_{\rightarrow} h^n(\tilde{U}) \cong h^n(EG^m \times_G A) \ .$$

For any such $\tilde{U}$, the preimage $p^{-1}(\tilde{U}) \subset EG^m \times X$ is an open G-invariant neighbourhood of $EG^m \times A$ in $EG^m \times X$. As EG^m is compact, the *tube lemma* yields an open neighbourhood V of A in X, such that $\tilde{U} \supset EG^m \times_G V \supset EG^m \times_G A$. Thus

$$h^n_G(A) \cong \lim_{\rightarrow} h^n(\tilde{U}) \cong \lim_{\rightarrow} h^n(EG^m \times_G V) \cong \lim_{\rightarrow} h^n_G(V) \ ,$$

where V varies over the open neighbourhoods of A. In the last isomorphism we applied again Proposition 2.1.8 for $V \subset X$, which is paracompact as X is hereditarily paracompact.

The system of all neighbourhoods and the system of all open neighbourhoods are cofinal anyway. As X is paracompact there is a closed neighbourhood inside each open neighbourhood, thus the system of all closed neighbourhoods is cofinal as well.

The other statements follow from the definition of the direct limit. For the second statement: All homomorphisms are $h^*(BG)$-module homomorphisms, thus it suffices to find pre-images of the generators in order to prove surjectivity. $\square$

If we know a bit more about A we have a better result for *any* generalised cohomology theory h^*:

Proposition 2.1.10 *Let G be a compact Lie group and $A \subset X$ be G-spaces. Suppose A is a G-neighbourhood deformation retract. Then there is a G-neighbourhood U of A in X such that the restriction*

$$h_G^*(U) \to h_G^*(A)$$

is an isomorphism for any cohomology theory h^*.

Proof: Chose a neighbourhood $U \supset A$ with a G-deformation retraction, that is a continuous map $r : [0,1] \times U \to U$ with $r(0,\cdot) = id$, $r(1,U) = A$, $r(t,a) = a$ for all $a \in A$, G-equivariant in the second variable. This defines a G-deformation retraction

$$\tilde{r} : [0,1] \times EG \times U \to EG \times U, (t,e,x) \mapsto (e, r(x,t))$$

to $EG \times A$ and a deformation retraction

$$r_G : [0,1] \times EG \times_G U \to EG \times_G U\ .$$

of the orbit space $EG \times_G U$ to $EG \times_G A$. Thus, the latter spaces are homotopy equivalent and the restriction induces an isomorphism in any cohomology theory. $\square$

The map

$$p : X_G \to BG,\ \ [(e,x)] \mapsto [e] \tag{2.1}$$

is the projection of a numerable fibre bundle with fibre X associated to the universal bundle $EG \to BG$ via the G-operation of X and hence a fibration. If $A \subset X$ we obtain $A_G \subset X_G$ and (X_G, A_G) is a pair of bundles over BG.

The projection p induces an $h_G^*(pt) = h^*(BG)$-module structure on $h_G^*(X)$ and $h_G^*(X,A)$, if h^* is a cohomology theory with a product (that is, if it is functor to the category of rings or R-algebrae.)

Proposition 2.1.11 *If the G-action has a fixed point $\{x\} \overset{i_x}{\subset} X$, the composition*

$$h^*(BG) \xrightarrow{p^*} h^*(X_G) \quad \cong \quad h^*_G(X) \xrightarrow{i_x^*} h^*_G(x) \quad \cong \quad h^*(BG)$$

is an isomorphism, hence p^* is injective. □

If $j_X : X \to X_G$ is the injection of a fibre, we obtain a homomorphism $j^* : h^*_G(X) \cong h^*(X_G) \overset{j_X^*}{\to} h^*(X)$ from the equivariant to the non-equivariant h^*-cohomology of X.

Definition 2.1.12 *We call the pair of G-spaces (X, A)* **totally nonhomologous to zero (TNHZ)** *in (X_G, A_G) -with respect to h^*, if $j^*_{(X,A)} : h^*_G(X, A) \to h^*(X, A)$ is surjective.*

Definition 2.1.13 *Let (X, A) be a pair of G-spaces and h^* a cohomology theory with R-modules as values. A* **cohomology extension of the fibre (CEF)** *is an R-module homomorphism*

$$t : h^*(X, A) \to h^*_G(X, A)$$

such that $j^*_{(X,A)} \circ t$ is an isomorphism.

Proposition 2.1.14 *Let h^* a cohomology theory with values in the category of R-modules, where R is a commutative ring with unit. Let (X, A) be a pair of G-spaces. Suppose $h^*(X, A)$ is a free R-module.*

Then the following statements are equivalent:

1. *(X, A) is TNHZ.*
2. *(X, A) has a CEF.*

Proof: A homomorphism to a free R-module has a right inverse if, and only if, it is surjective. □

Let h^* be a cohomology theory with values in R-algebrae. We denote the algebra product by $\cup$. Then, a cohomology extension of the fibre $t : h^*(X, A) \to h^*_G(X, A)$ defines a $h^*(BG)$-module homomorphism

$$\begin{aligned} \Phi_t : h^*(BG) \otimes_R h^*(X, A) &\to h^*_G(X, A) \\ \alpha \otimes \beta &\mapsto p^*(\alpha) \cup t(\beta) \ . \end{aligned}$$

Proposition 2.1.15 *(Leray-Hirsch theorem, [tD87], III. 1.14) Suppose $h^*(X, A)$ is a finitely generated free R-module and a Künneth theorem is valid for h^*. Suppose we have a cohomology extension of the fibre as above. Then the homomorphism Φ_t is an isomorphism.*

Proof: Tom Dieck's argument uses the conditions mentioned above. □

Proposition 2.1.16 *Let h^* be a cohomology theory with values in k-vector spaces, for which a Künneth theorem is valid.*

If the G-operation on a pair of spaces (X, A) is trivial and $h^*(X, A)$ is finitely generated as a k-vector space, $h^*_G(X, A)$ is a free $h^*(BG)$-module and isomorphic to the $h^*(BG)$-module

$$h^*(BG) \otimes_k h^*(X, A) \ .$$

The isomorphism is an isomorphism of k-algebrae. In particular (X, A) is (TNHZ).

Proof: Under the assumptions $(X_G, A_G) \cong (X, A) \times BG$. The Künneth theorem implies the statement. □

Only for the special and somewhat extreme case (TNHZ), the $h^*(BG)$-module structure can be described by a tensor product. In general we have to consider spectral sequences. As BG is path-connected, we have a Leray-Serre-Atiyah-Hirzebruch spectral sequence associated to p. There are different versions with different sets of technical assumptions. For our purpose it will be sufficient to have the following:

Theorem 2.1.17 *Let $p : E \to B$ be a fibration with fibre F such that the base B is a path-connected CW-complex. Let h^* be an additive cohomology theory.*

Assume that either B is finite dimensional or h^* is bounded below on the fibre (i.e. $h^q(F) \cong 0$ for all $q < N$ for some N).

Then there exists a spectral sequence

$$H^p(B; \mathcal{F}^q) \cong E_2^{p,q} \Rightarrow h^{p+q}(E) \ ,$$

where $\mathcal{F}^q$ is the sheaf given by $\mathcal{F}^q(U) \cong h^q(p^{-1}(U))$ with stalks $h^q(F)$, and $H^*(\cdot, \mathcal{F}^q)$ is the corresponding sheaf-cohomology.

and a relative version thereof:

Theorem 2.1.18 *Let $p : E \to B$ be a fibration with fibre F and $E_0 \subset E$ such that $p|_{E_0} : E_0 \to B$ is a fibration with fibre F_0. Assume that B is a path-connected CW complex. Let h^* be an additive cohomology theory.*

Assume that either B is finite dimensional or h^* is bounded below on the fibre (i.e. $h^q(F) \cong 0$ for all $q < N$ for some N).

Then there exists a spectral sequence of

$$H^p(B; \mathcal{F}^q) \cong E_2^{p,q} \Rightarrow h^{p+q}(E, E_0) \ ,$$

where $\mathcal{F}^q$ is the sheaf of local coefficients given by $\mathcal{F}^q(U) \cong h^q(p^{-1}(U), (p|_{E_0})^{-1}(U))$ with stalks $h^q(F)$, and $H^*(\cdot, \mathcal{F}^q)$ is the corresponding sheaf-cohomology.

Remark 2.1.19 1. As every CW-complex is homotopy equivalent to a simplicial complex, which has a *good cover* in the terminology of Bott and Tu ([BT82]), their proof of the Leray-Serre spectral sequence applies to both cases.

On the other hand George W. Whitehead's [Whi78] theorem XIII, 4.9* implies the first version under the additional assumption that the fibre F is path-connected. Theorem 9.34 of Davis and Kirk [DK01] is exactly the version we need, but unfortunately it is given without proof.

2. Here, by the assumption that B is a CW complex, hence locally contractible, the corresponding sheaves $\mathcal{F}^q$ are locally constant and B is homologically locally connected. By Bredon ([Bre67], III.1. and exercise (8)), the singular cohomology of B with coefficients in $\mathcal{F}^q$ and sheaf cohomology of B with respect to the sheaf $\mathcal{F}^q$ are isomorphic. By Hatcher ([Hat02], section 3.H.) this singular cohomology can be identified with the homology of the complex

$$Hom_{\mathbb{Z}[\pi_1(B)]}(C_*(\tilde{B}), h^q(F)) ,$$

where $\tilde{B}$ is the universal cover of B. $\pi_1(B)$ operates on $\tilde{B}$ by deck transformations and on F by the transport of the injection of a fibre.

In the special case of $H^*(BG, \mathcal{F}^q)$ for a finite group, G we obtain the chain complex

$$Hom_{\mathbb{Z}[G]}(C_*(EG), h^q(F)) .$$

$\diamond$

We can now give the proof of Proposition 2.1.8.

Proof: We have a map in the category of fibre bundles

$$\begin{tikzcd} EG^m \times_G X \arrow[r] \arrow[d, "p_m"'] & EG \times_G X \arrow[d, "p"] \\ BG^m \arrow[r] & BG \end{tikzcd} ,$$

which induces a restriction map of Leray-Serre spectral sequences

$$E_2^{p,q} \cong H^p(BG; \mathcal{F}^q) \to H^p(BG^m, \mathcal{F}^q|_{BG^m}) \cong \tilde{E}_2^{p,q} ,$$

where $\mathcal{F}$ is the sheaf given by $\mathcal{F}^q(U) \cong h^q(p^{-1}(U))$. As BG^m is the m-skeleton of BG, the restriction $E_2^{p,q} \to \tilde{E}_2^{p,q}$ is an isomorphism for $p < m$. This induces isomorphisms between $E_3^{p,q}$ and $\tilde{E}_3^{p,q}$ for $p < m-2$, between $E_4^{p,q}$ and $\tilde{E}_4^{p,q}$ for $p < m-5$, ... By induction we get an isomorphism between $E_r^{p,q}$ and $\tilde{E}_r^{p,q}$ for $p < m - r(r-1)/2 + 1$. If m is large enough there will be an isomorphism

$$E_{i+2}^{p,q} \to \tilde{E}_{i+2}^{p,q}$$

for $p + q = i$. These groups survive to E_∞, thus we have isomorphic quotients of filtrations of $h^i(EG \times_G X)$ and $h^i(EG^m \times_G X)$, which implies an isomorphism of the cohomologies for field coefficients.[1] □

Proposition 2.1.20 *Let G be a compact Lie group. If X is a free G-space such that $X \to X/G$ is a numerable principal G-bundle, $EG \times_G X \to X/G$ is a numerable fibre bundle with contractible fibre EG, hence a homotopy equivalence. We have*

$$h_G^*(X) \cong h^*(X/G).$$

The condition on X is fulfilled, if X is a G-CW complex or a (paracompact) differentiable G-manifold.

Proof: (s. [tD87], I, 8.18 (iii)) A trivialization of $X \to X/G$ over U/G yields a homeomorphism $U \to G \times U/G$ over U/G. Now

$$EG \times_G U \cong EG \times_G (G \times U/G) \cong EG \times U/G$$

over U/G. Now, $EG \times_G X \to X/G$ is a numerable fibre bundle, as it can be trivialised over a numerable cover of X/G. By Dold [Dol63], Corollary 3.2. the fibre bundle projection is a homotopy equivalence (actually shrinkable).

Both G-CW complexes and G-manifolds are completely regular, hence the projection $X \to X/G$ is a locally trivial G-principal bundle ([Bre72], II. 5.8.)

For X a G-CW complex X/G is a CW-complex, hence locally contractible and paracompact, thus it has numerable category and the G-principal bundle $X \to X/G$ allows local trivializations over a numerable covering.

If X is a differentiable G-manifold, there are *slices* for the G-operation, which define local trivializations and give X/G the structure of an $n-$dim G-dimensional differentiable manifold. As X is paracompact and $X \to X/G$ is closed, the orbit space $X/$ is paracompact and $X \to X/G$ is a numerable principal bundle. □

[1] I suppose that a better proof would allow $m = i + 1$, as this would agree with Proposition 2.1.4 for locally G-contractible X. Bredon suggests such a proof, but I don't see how to do it, unless the spectral sequence collapses at the E_2-term, which is the case for X totally nonhomologous to zero. As I do not need a sharp bound for m, I refrain from bothering.

Corollary 2.1.21 *Let G be a compact Lie group of dimension d. Suppose that X is a free G-space and either a G-CW-complex with cells $D^i \times G$, $i \leq n - d$, or a differentiable G-manifold of dimension n.*

Then X/G has dimension $n - d$ (as a CW complex or manifold, respectively).

If h^* is an ordinary cohomology theory this implies

$$h_G^k(X) = 0 \text{ for } k > n - d \ .$$

If, moreover, the cohomology ring $h_G^*(pt)$ is unbounded in degree, $h_G^*(X)$ consists of torsion elements with respect to the $h_G^*(pt)$-module structure, that is for all $u \in h_G^*(X)$ there is an $\alpha \in h_G^*(pt) \cong h^*(BG)$ such that $p^*(\alpha) \cup u = 0$.

Proof: By proposition 2.1.20 $h_G^*(X) \cong h^*(X/G)$. For $* > n - d$ and ordinary cohomology this is the trivial group. Hence all elements in $h_G^*(pt)$ of higher degree are annihilated, wenn pulled back to X. □

For more general spaces there are different notions of dimension. The most useful for our purpose is *cohomological dimension*.

Definition 2.1.22 *If R is a ring the cohomological dimension $\mathrm{cd}_R(X)$ of a space X is defined as the supremum of all integers n such that there is a sheaf $\mathcal{F}$ of R-modules with $H^n(X, \mathcal{F}) \ncong 0$.*

By Bredon [Bre67], II, 15.8. cohomological dimension is not increasing when passing to locally closed subspaces, and it is locally defined in the following sense: If each point in x admits a locally closed neighbourhood N with $\mathrm{cd}_k N \leq n$, then $\mathrm{cd}_k X \leq n$.

The following proposition implies the dimension claims of 2.1.20:

Proposition 2.1.23 *Let X be a paracompact Hausdorff space, on which the compact Lie group G operates and let k a field. Then $\mathrm{cd}_k(X/G) \leq \mathrm{cd}_k(X) - d$, where $d = \min_{x \in X} \mathrm{cd}_k(G/G_x)$. If the G-operation has constant orbit type, we have equality.*

Proof: First we prove the last statement. X is completely regular, thus for an operation with constant orbit type the projection $X \to X/G$ is a locally trivial fibre bundle. Thus every point in X has a closed neighbourhood that is homemorphic to $A \times B$, with A a closed subset of X/G and B a closed subset of G/G_x for any $x \in X$.

Now the local definition of cd_k and the Künneth theorem (k is a field) imply

$$\mathrm{cd}_k(X) = \mathrm{cd}_k(X/G) + \mathrm{cd}_k(G/G_x) \ .$$

The rest of the proof is a slightly modified version of Quillen's proof ([Qui71], I, Proposition A.11.). Since the closed subgroups of G satisfy the descending chain condition we may proceed by induction and suppose that the claim is true for all proper closed sugroups of G. The quotient map $X \to X/G$ induces a homeomorphism on X^G. Thus, if X^G is nonempty, $\mathrm{cd}_k X^G/G = \mathrm{cd}_k X^G \leq \mathrm{cd}_k X =: n$. By Quillen's proposition $A.8.$ it is now sufficient to prove that $\mathrm{cd}_k(A) \leq n - d$ with $d = \min_{x \in X} \mathrm{cd}_k(G/G_x)$ for all closed subspaces A of X/G that are disjoint from X^G/G. For that we just have to prove that every $y \in Y/G$, $Y := X - X^G$ has a closed neighbourhood with the required property. Now by the existence of slices in completely regular spaces (a paracompact Hausdorff space is completely regular), y has a closed neighbourhood N s.th. $q^{-1}(N) \cong G \times_H S$ for a closed subset $S \subset Y$, $H = G_x$, $q(x) = y$ and $N = S/H$. By the induction hypothesis

$$\mathrm{cd}_k(N) \leq \mathrm{cd}_k(S) - \min_{s \in S} \mathrm{cd}_k(H/H_s).$$

On the other hand, as H operates freely on $G \times S$

$$\mathrm{cd}_k(Y) \geq \mathrm{cd}_k(q^{-1}(N)) = \mathrm{cd}_k(G \times S) - \mathrm{cd}_k(H) = cd_k(G) + \mathrm{cd}_k(S) - \mathrm{cd}_k(H) \ .$$

The last identity is again a consequence of the Künneth theorem, as k is a field. As $H \to H/H_s$ is a submersion of compact manifolds we have

$$\mathrm{cd}_k(H/H_s) = \mathrm{cd}_k(H) - \mathrm{cd}_k(H_S) \ .$$

Now, putting this together we obtain for all $[(g,s)] \in q^{-1}(N)$

$$\mathrm{cd}_k(N) \leq \mathrm{cd}_k(S) - \min_{s \in S}(\mathrm{cd}_k(H) - \mathrm{cd}_k(H_s)) \leq \mathrm{cd}_k(Y) - \min_{s \in S}(\mathrm{cd}_k(G) - \mathrm{cd}_k(H_s)) \ .$$

The isotropy group of $[(1,s)] \in G \times_H S$ with respect to the G-operation is H_s. Thus

$$\mathrm{cd}_k(N) \leq \mathrm{cd}_k(Y) - \min_{s \in S} \mathrm{cd}_k(G/G_{[(1,s)]}) = \mathrm{cd}_k(Y) - \min_{y \in Y} \mathrm{cd}_k(G/G_y) \ .$$

□

Proposition 2.1.24 *Let G be a compact Lie group, X an n-dimensional G-manifold with fixed point set F. h^* is an additive cohomology theory with values in k-vector spaces for some field k.*

Suppose one of the two sets of conditions:

1. *(SF, semifree) the G-action is semifree, $d := \dim G$.*

2. *(CSF, cohomologically semifree) $h^0_G(G/G_x) \cong k$ and $h^s_G(G/G_x) \cong 0$ for all $s > 0$ and all $x \notin F$, $d := \min_{x \notin X^G} \dim(G/G_x) - 1$.*

Then the restriction to the fixed point set

$$h^k_G(X) \to h^k_G(X^G)$$

is an isomorphism for $k > n - d$ and an epimorphism for $k = n - d$.

Remark 2.1.25 1. Please note that the definition of d is not the same in the two cases, if the action is (SF) and - a fortiori - (CSF). The first set of assumptions provides the stronger result.

2. Below we give a few examples of actions and cohomology theories that are (SF) or (CSF). We *do not* try to give a general method to construct a cohomology theory for which all orbits but the fixed points are "invisible" and which thus makes the action (CSF). This would be done by localizing a given cohomology theory with respect to all cohomology classes which restrict to zero on some of the non constant orbits. Under some additional assumptions on the number of orbit types ("finitely many orbit types" or "finitely many connected orbit types") we could prove that restriction to the fixed point set is an isomorphism for such a theory. Usually we would lose the degree information by localization, which, however, can be recovered in some cases. In particular this can be done for the cases most relevant for our purposes ($G = S^1$ or $G = \mathbb{Z}_p$). We may remark, that in these cases no assumptions on the number of orbit types would be needed. (A $\mathbb{Z}_p$-action has only two orbit types, anyway. Any S^1-action has finitely many connected orbit types which allows to derive a localization theorem for $\mathbb{Q}$-coefficients.)

3. Actions of $\mathbb{Z}^p$, p prime are semifree, anyway.

4. Actions of $G = S^1$ are cohomologically semifree for ordinary cohomology with coefficients in any field of characteristic 0: All *proper* subgroups G_x of S^1 are finite cyclic of order m. Hence $G/G_x \cong S^1$ and $(G/G_x)_G = EG \times_G (G/G_x) = BG_x$ and $\rho : EG \to BG_x$ is a covering space with m leaves. Thus there is a transfer map $\tau : h^*(EG) \to h^*(BG_x)$ such that $\tau \circ \rho^*$ is multiplication by m. As it factors over $h^*(EG) \cong 0$ for $* > 0$ we obtain (CSF) in characteristic 0. $\diamond$

Proof: For $G = S^1$ $(char(k) = 0)$ or $G = \mathbb{Z}/m$ this is stated in ([tD87], III. 4.9) as a consequence of a localization theorem.

In general the fixed point set $F := X^G$ is a closed submanifold of X.

By the tubular neighbourhood theorem (s. Bredon [Bre72], VI. 2.2), there is an open G-invariant neighbourhood U of F diffeomorphic to a G-vector bundle ξ over F via an equivariant diffeomorphism $\phi : E(\xi) \to U$.

There is a G-invariant inner product on ξ (s. [Bre72], VI. 2.1), hence the disc bundle $D(\xi) = \{\xi_p | \langle \xi_p, \xi_p \rangle < 1\}$ and its closure $\overline{D(\xi)}$ map diffeomorphically onto neighbourhoods V and $\bar{V}$ of F. All three spaces U, V and $\bar{V}$ are G-homotopy equivalent to F, as the total spaces $E(\xi)$, $D(\xi)$ and $\overline{D(\xi)}$ can be retracted to the zero section F by the radial deformation retraction

$$f_\lambda(\xi_p) = (1 - \lambda)\xi_p,$$

which is equivariant.

Summing up these results, we have found open G-neighbourhoods U and V of X^G with $U \supset \bar{V} \supset V \supset X^G$ and $U \simeq_G V \simeq_G X^G$.

Now we cover X by the open sets U and $U' := X - \bar{V}$ and obtain the Mayer-Vietoris sequence:

$$h_G^{s-1}(U \cap U') \longrightarrow h_G^s(X) \longrightarrow h_G^s(U) \oplus h_G^s(U') \longrightarrow h_G^s(U \cap U')$$

Under condition (SF) $U \cap U'$ and U' are free and non closed n-dimensional G-manifolds, by proposition 2.1.21 we have $h_G^s(U \cap U') \cong h^s((U \cap U')/G)$ and $h_G^s(U') \cong h^s(U'/G)$. Both groups are trivial for $s \geq n - \dim G$. as these quotient manifolds are not closed. The statements follow immediately from the Mayer-Vietoris sequence.

Under condition (CSF) we consider the Leray spectral sequence for the map $f : U_G \to U/G$ (which, in general, is *no fibre bundle* and *no fibration*). The E_2-term is

$$E_2^{p,q} = H^p(U/G, \mathcal{F}) ,$$

the sheaf cohomology of U with respect to the Leray sheaf of f, that is the sheafification $\mathcal{F}$ of $V \mapsto h^q(f^{-1}(V))$. The tubular neighbourhood theorem yields for every orbit Gx a G-invariant neighbourhood V' in U that can be G-deformation retracted to Gx. V'_G projects to an open neighbourhood V of x in U/G, s. th. $h^q(f^{-1}(V)) \cong h^q(f^{-1}(x)) \cong h^q(BG_x)$. By assumption this is the zero object for $q > 0$ and isomorphic to k for $q = 0$.

By proposition 2.1.23, $cd_k(U/G) = cd_k(U) - \tilde{d}$ with $\tilde{d} := \min_{x \in U} cd_k(G/G_x)$, we obtain that $h_G^s(U) \cong 0$ for $s > n - \tilde{d}$ or, equivalently, for $s \geq n - d$, $d := \tilde{d} - 1$. □

Corollary 2.1.26 *Let G be a compact Lie group, X an n-dimensional G-manifold and A a G-invariant subspace. Let the further assumptions of 2.1.24 be satisfied. Then, the restriction*

$$h_G^*(X, A) \to h_G^*(X^G, A^G)$$

is an isomorphism in degrees $* > n - d + 1$.

Proof: We apply the five lemma to the restriction homomorphism from the long exact sequence

$$\cdots \to h_G^s(X, A) \to h_G^s(X) \to h_G^s(A) \to h_G^{s+1}(X, A) \to \cdots$$

to the long exact sequence

$$\cdots \to h_G^s(X^G, A^G) \to h_G^s(X^G) \to h_G^s(A^G) \to h_G^{s+1}(X^G, A^G) \to \cdots .$$

By the above proposition 2.1.24 $h_G^s(X) \to h_G^s(X^G)$ and $h_G^s(A) \to h_G^s(A^G)$ are isomorphisms for $s > n - d$. The five lemma yields the claim. □

Corollary 2.1.27 *Suppose X is an n-dimensional G-manifold, h^* a cohomology theory, the assumptions of Proposition 2.1.24 are satisfied, and $h^*(X)$ is a finite dimensional k-vector space. Suppose furthermore that $h^*(BG)$ is finitely generated as a k-algebra and has an element $\omega \in h^d(BG)$, $d > 0$ such that multiplication with ω defines an isomorphism $h^i(BG) \to h^{i+d}(BG)$ for all $i \geq 0$.*

If X is (TNHZ), the restriction $h_G^*(X) \to h_G^*(X^G)$ is injective.

Proof: By the Leray-Hirsch theorem (proposition 2.1.15) and proposition 2.1.16 the restriction is an $h^*(BG)$-module homorphism of free $h^*(BG)$-modules

$$r : h^*(BG) \otimes_k h^*(X) \to h^*(BG) \otimes_k h^*(X^G) .$$

For all generators u of the k-algebra $h^*(BG)$ and all $v \in h^*(X)$ we have $r(u \otimes_k v) \neq 0$. We argue indirectly: If it were zero the image of $u \cup \omega^k \otimes_k v$ would be zero, which contradicts proposition 2.1.24. □

Remark 2.1.28 For (TNHZ) G-manifolds X we thus get a description of $h_G^*(X)$ as a sub-algebra of $h_G^*(X^G)$.

Even without the condition (TNHZ) we have seen that in degrees $\geq n$ all cohomology classes in $h_G^*(X^G)$ can be extended to cohomology classes in $h_G^*(X)$. In particularly any class carried by a component of the fixed point set can be extended to a class in $h_G^*(X)$ that is zero on all the other fixed point components and will in this way "recognise" the fixed point component it stems from.

In general, this is not the case below degree n. The corollary, however, tells us that for (TNHZ) there are still as many as $\dim h_G^d(X)$ different classes carried by the fixed point set. Thus, even if it is not possible to extend a class from a component of the fixed point set to X in the way sketched above, we can "often" extend it to a class that vanishes on some of the fixed point components.

For non (TNHZ) G-manifolds this may still happen, although there may be in general much less classes available.

This observation lies at the heart of the critical point theorems of the next section. For the $d = n$ it was first conceived by Bartsch and Wang ([BW97a], [BW97c],[BW97b]).

Before we can make use of it in degrees $< n$ be, we must find means to obtain more information on the possible extensions, which will be done subsequently. $\diamond$

Definition 2.1.29 *Let X be a G-space.*

We say that a cohomology class $\alpha \in h_G^*(X)$ **recognises** *a clopen subset $A \subset X^G$ if $\alpha|_C \neq 0$ for every component C of A and $\alpha|_C = 0$ for every component of X^G not contained in A. In this case, we write* $\mathrm{rec}\ (\alpha) := A$.

For clopen subsets A and B of X^G we say that A **is separated from** *B* **by a cohomology class** *$\alpha \in h_G^d(X)$, if $A \subset rec(\alpha)$ and $B \cap \mathrm{rec}\ (\alpha) = \emptyset$.*

For two clopen subsets A and B of X^G we set

$$\sigma(A, B) := \min\{d \in \mathbb{N} | A \textit{ is separated from } B \\ \textit{by a cohomology class of degree } d\} .$$

Remark 2.1.30 Please note, that $\sigma(A, B)$ is not symmetric in its arguments, unless we consider cohomology with $\mathbb{Z}_2$-coefficients.

In any case we have $\sigma(A, B) \leq n$. $\diamond$

2.1.2. More on (TNHZ)

It is useful to have verifiable criteria for (X, A) to be $(TNHZ)$ in (X_G, A_G). The spectral sequences yield such criteria which only need certain relations between Betti numbers. The following proposition is well-known (cf. [AP93],[Bre72],[tD87]), but usually stated for special cases only. We have tried to make explicit which conditions are needed and rewritten the proof.

Let us fix some notation before stating the proposition. Let $\mathbb{K}$ be a field, and let h^* be an ordinary additive cohomology with field coefficients $\mathbb{K}$. Let $H^*(\cdot, \mathcal{F})$ singular cohomology with values in a locally constant sheaf $\mathcal{F}$.

If G_0 is the component of 1 in G, we have $\pi_1(BG) \cong G/G_0$. If V is a G/G_0-representation $\mathbb{K}$ vectorspace and a base point of BG is given, we can construct a unique local coefficient system on BG (s. [Spa66], p. 58). We denote the sheaf on BG that correponds to this local coefficient system (s. [Spa66], p. 360)) by $\mathcal{F}(V)$ (dropping the choice of the basepoint in the notation).

Proposition 2.1.31 *Let X be a smooth n-dimensional G-manifold with finitely generated cohomology $h^*(X)$, which meets the conditions of proposition 2.1.24. Now suppose*

(P) $h^*(BG)$ is a finitely generated k-algebra and has at least one generator ω of degree $d > 0$ such that multiplication by ω defines an isomorphism $h^k(BG) \to h^{k+d}(BG)$ for all $k \geq 0$. For any finite dimensional G/G_0-representation $\mathbb{K}$-vector space V and $m \in \mathbb{Z}$ the following implication is valid:

$$\begin{aligned} &\dim H^*(BG, \mathcal{F}(V)) = \dim h^*(BG) \dim V \quad \textit{for all} \quad s > m \\ \Leftrightarrow\ &V^{G/G_0} \cong V\ . \end{aligned}$$

Then, X is (TNHZ), iff

$$\sum_{i=0}^{n} \dim h^i(X) = \sum_{i=0}^{n} \dim h^n(X^G)\ . \tag{2.2}$$

Remark 2.1.32 By remark 2.1.19 $H^*(BG, \mathcal{F}(V))$ can be considered either as singular cohomology with local coefficients or as cohomology with respect to a locally constant sheaf. It can be calculated as the homology of the cochain complex $\mathrm{Hom}_{\mathbb{Z}[G/G_0]}(C_*(\widetilde{BG}), V)$, where the singular complex $C_*(\widetilde{BG})$ can be replaced by the cell complex of the universal cover of any CW-model for BG. For a finite group we have $EG \simeq \widetilde{BG}$, and $C_*(\widetilde{BG})$ can be replaced by a any free resolution of $\mathbb{Z}$ over $\mathbb{Z}[G]$.

Condition (P) can be easily verfied for some *periodic* chain complexes $C_*(\widetilde{BG})$, as for $G = \mathbb{Z}_p$ and $\mathbb{K} = \mathbb{Z}_p$, p prime, and for $G = S^1$ or $G = S^3 = Sp(1)$ and any field $\mathbb{K}$. The G-action gives $h^*(X)$ the structure of a $\mathbb{K}[G/G_0]$-module. For connected groups this structure is trivial, anyway. For $G = \mathbb{Z}_p$ the coboundaries in $\mathrm{Hom}_{\mathbb{Z}[G/G_0]}(C_*(\widetilde{BG}), V)$ are zero if and only if V is a trivial $\mathbb{K}[G/G_0]$-module.

Corollary 2.1.33 *Let X be a smooth n-dimensional G-manifold with finitely generated ordinary cohomology $h^*(X)$ with values in $\mathbb{K}$-vector spaces. Suppose*

($G = \mathbb{Z}_p$ and $\mathbb{K} = \mathbb{Z}_p$) or ($G = S^1$ and $\mathbb{K} = \mathbb{Q}$) or ($G = S^3$ and the G-action satisfies (CSF)). We have the following equivalence:

X is $TNHZ$ iff (2.2) is satisfied.

Proof:(Proposition 2.1.31)

First we have to recall a few facts concerning the fixed point set. By proposition 2.1.16,

$$h^*_G(X^G) \cong h^*(X^G) \otimes_{\mathbb{K}} h^*(BG)$$

as $h^*(BG)$-modules and k-algebrae. Proposition 2.1.24 implies that $h^s_G(X) \to h^s_G(X^G)$ is an isomorphism for degree $s > n$.

Now assume that X is (TNHZ). By the Leray-Hirsch theorem (proposition 2.1.15) we have

$$h^*_G(X) \cong h^*(X) \otimes_k h^*(BG)$$

as $h^*(BG)$-modules. Now, the isomorphism $h^s_G(X) \to h^s_G(X^G)$ yields

$$\sum_{i=0}^{s} \dim h^i(X) \dim h^{s-i}(BG) = \sum_{i=0}^{s} \dim h^i(X^G) \dim h^{s-i}(BG) \tag{2.3}$$

Multiplication by ω induces an isomorphism $h^k(BG) \cong h^{k+d}(BG)$ for each $k \geq 0$. If we now sum equation (2.3) over $s \in \{s_0, \ldots, s_0 + (d-1)\}$ and divide by $0 \neq t := \sum_{i=0}^{d-1} \dim h^i(BG)$, we obtain equation (2.2). (For $d > 1$, if $h^k(BG) \cong 0$ for $k \not\equiv 0 \bmod d$ the argument gives actually more precise information, namely that the sums of the dimensions of $h^k(X)$ and $h^k(X^G)$ for $k \equiv s \bmod d$ are equal for all s.)

The second part of the proof relies on the fact, that condition (2.2) implies the collapse of the spectral sequence at E_2 in high total degree, and condition (P) then allows to conclude that E_2 is a tensor product and collapses everywhere.

Assume (2.2) is satisfied.

The Leray-Serre spectral sequence (Theorem 2.1.17) for $X_G \to BG$ implies that

$$\dim h^s(X_G) \leq \sum_{i=0}^{s} \dim(h^i(X)) \dim(h^{s-i}(BG)) \; . \tag{2.4}$$

Suppose for the moment, that we have equality in (2.4). for all $s \geq 0$. This implies that the spectral sequence collapses at the E_2-term and that $E_2^{0,*}(X_G \to$

$BG) \cong H^0(BG, \mathcal{F}(h^*(X))) \cong h^*(X)$. $H^0(BG, \mathcal{F}(h^*(X))$ can be identified with the set of G/G_0-invariant elements of $h^*(X)$.

Hence the local coefficient system is constant $\cong h^*(X)$, and the inclusion of one fibre $X \to \{p\}$ in $X_G \to BG$ induces an isomorphism

$$E_2^{0,*}(X_G \to BG) \cong H^0(BG, h^*(X)) \to H^0(\{p\}, h^*(X)) \cong E_2^{0,*}(X \to \{p\}) \ .$$

Since both spectral sequences collapse at the E_2-term and $E_2^{0,*}(X \to \{p\}) \cong h^*(X)$, $h^*(X_G) \to h^*(X)$ has to be surjective. This is precisely the condition (TNHZ).

It remains to prove equality in (2.4) for any $s \geq 0$. As $h^s_G(X) \to h^s_G(X^G)$ is an isomorphism for $s > n$, we have

$$\dim h^s(X_G) = \sum_{i=0}^{s} \dim h^i(X^G) \dim h_G^{s-i}(pt) \ . \tag{2.5}$$

Now we sum again over both sides of the equation

$$\sum_{j=0}^{d-1} \dim h^{s+j}(X_G) = (\sum_{i=0}^{n} h^i(X^G))t \tag{2.6}$$

for $s > n$, and, by (2.2),

$$\sum_{j=0}^{d-1} \dim h^{s+j}(X_G) = (\sum_{i=0}^{n} h^i(X))t \ . \tag{2.7}$$

Summing in the same way over inequality (2.4), we get that it must be in fact an equation for $s > n$. It remains to settle the question for $s \leq n$.

Consider again the Leray-Serre spectral sequence for $X_G \to BG$. (For the following part of the proof, cf. [tD87], III. 4.16.).

For any s, the groups $E_\infty^{s-i,i}$ are quotients of successive terms in a filtration of $h^s_G(X)$, hence

$$\dim h^s_G(X) = \sum_{i=0}^{s} \dim E_\infty^{s-i,i} \ .$$

The group $E_\infty^{s-i,i}$ is a subquotient of $E_2^{s-i,i}$, thus

$$\dim h^s_G(X) \leq \sum_{i=0}^{s} \dim E_2^{s-i,i} = \sum_{i=0}^{s} \dim H^{s-i}(BG, \mathcal{F}(h^i(X))).$$

We have shown already, that we have equality for $s > n$. By condition (P), this implies that the the G/G_0-operation on $h^*(X)$ is trivial and the sheaves (and the corresponding local coefficient systems) $\mathcal{F}(h^i(X))$ are actually constant. Thus,

$$E_2^{s-i,i} \cong h^{s-i}(BG) \otimes_{\mathbb{K}} h^i(X)$$

for all $s \geq 0$. Furthermore, the spectral sequence collapses for $s > n$ at E_2.

We are done if we can prove that for all $u \in h^p(BG)$ and $v \in h^q(X)$ the element $u \otimes_{\mathbb{K}} v \in E_2^{p,q}$ is mapped to zero by d_2. We chose an $i \geq 0$ such that the total degree of $(u \cup \omega^i) \otimes_{\mathbb{K}} v$ is greater than n. Hence

$$\begin{aligned} 0 &= d_2((u \cup \omega^i) \otimes_{\mathbb{K}} v) \\ &= d_2(u \otimes_{\mathbb{K}} v) \cdot (\omega^i \otimes_{\mathbb{K}} 1_X) + (-1)^{p+q}(u \otimes_K v) \cdot d_2(\omega^i \otimes_{\mathbb{K}} 1_X) \\ &= d_2(u \otimes_{\mathbb{K}} v) \cdot (\omega^i \otimes_{\mathbb{K}} 1_X) \ . \end{aligned}$$

As multiplication with ω defines isomorphisms $h^k(BG) \to h^{k+d}(BG)$, we conclude that $d_2(u \otimes_{\mathbb{K}} v) = 0$: The spectral sequence collapses at E_2, $E_2 \cong E_\infty$.

The tensor product structure of $E_2 \cong E_\infty$ allows to give a cohomology extension of the fibre: X is (TNHZ). □

Example 2.1.34 *We will describe a class of (TNHZ) spaces for $G = \mathbb{Z}_2, \mathbb{Z}_p, S^1$, that contains the motivating example of the torus with $\mathbb{Z}_2$-action. For this class we obtain a complete description of $h_G^*(X)$ as a graded $\mathbb{K}$-algebra and $h^*(BG)$-module. Furthermore we can describe, whether classes in $h_G^*(X)$ restrict to non-zero classes on certain fixed points. The computation or estimation of the minimal separation degree $\sigma(x, F)$ for a fixed point x and a set F of fixed points reduces to an algebraic task.*

Let h^* be some ordinary cohomology with $\mathbb{K}$-coefficients.

(1) $G = \mathbb{Z}_2 = \{1, g\}$ acts on $S^{k_i} \subset \mathbb{R}^{k_i+1}$ by $g \cdot (x_1, \ldots, x_{k_i+1}) = (-x_1, \ldots, -x_{k_i}, x_{k_i+1})$ with two fixed points.

(2) For $p > 2$, p prime $G = \mathbb{Z}_p = \{1, g, g^2, \ldots, g^{p-1}\}$ acts on $S^{k_i} \subset \mathbb{R}^{k_i+1} \cong \mathbb{C}^{\frac{k_i}{2}} \times \mathbb{R}$ for k_i even by

$$g \cdot (z_1, \ldots, z_{\frac{k_i}{2}}, x_{k_i+1}) = (\zeta_p z_1, \ldots, \zeta_p z_{\frac{k_i}{2}}, x_{k_i+1}) \ , \ \zeta_p = e^{i\frac{2\pi}{p}}$$

with two fixed points - "north pole" $N_i = \{\nu_i\} = \{(0, \cdots, 0, 1)\}$ and "south pole" $S_i = \{\sigma_i\} = \{(0, \cdots, 0, -1)\}$.

(3) $G = S^1 \subset \mathbb{C}$ acts on $S^{k_i} \subset \mathbb{R}^{k_i+1} \cong \mathbb{C}^{\frac{k_i}{2}} \times \mathbb{R}$ for k_i even by

$$z \cdot (z_1, \ldots, z_{\frac{k_i}{2}}, x_{k_i+1}) = (zz_1, \ldots, zz_{\frac{k_i}{2}}, x_{k_i+1})$$

with two fixed points

In any case the product manifold $X := \prod_{i=1}^{n} S^{k_i}$ has cohomology with field coefficients that is isomorphic as a vector space to $\mathbb{K}^{2^n}$. The action on this manifold has exactly 2^n fixed points. Thus, by Proposition 2.1.31 and Remark 2.1.32, is (TNHZ) for ordinary cohomology with coefficients $\mathbb{K} = \mathbb{Z}_p$ (for $G = \mathbb{Z}_p$) or $\mathbb{K} = \mathbb{R}$ (for $G = S^1$).

Let us now exploit this fact to obtain a description of $h^*_G(X)$ and its relations to the fixed points. By corollary 2.1.27 the restriction $r : h^*_G(X) \to h^*_G(X^G)$ embeds $h^*_G(X)$ in $h^*_G(X^G)$ as a $\mathbb{K}$-algebra and an $h^*(BG)$-module. Henceforth denote X^G by F. In any of the three cases above the dimension of $h^s_G(pt) = h^s(BG)$ is at most 1 for each s, and there is a periodicity generator $\omega \in h^d_G(pt)$ (with $d = 1$ in case (1) and $d = 2$ in cases (2) and (3)).

($\mathbf{n = 1}$)

First, we construct a cohomology extension $t : h^*(X) \to h^*_G(X)$ of the fibre for $n = 1$, i.e. $X = S^k$. We start with an arbitrary cohomology extension of the fibre (which exists, as X is (TNHZ)). In degree zero t must be the inverse of the isomorphism $h^0(X_G) \to h^0(X)$ induced by the inclusion of a fibre $i_X : X \to X_G$. Thus, as $i^*_X(1_{X_G}) = 1_X$ we have $t(1_X) = 1_{X_G}$ and

$$r(t(1_X)) = r(1_{X_G}) = 1_{F_G} = 1_{N_G} + 1_{S_G} \ .$$

Now let $\phi_X \in h^k(S^k)$ be the orientation class of S^k. By the Leray-Hirsch theorem (proposition 2.1.15) we know that $h^*_G(X)$ is freely generated by $t(1_X)$ and $t(\phi_X)$ as an $h^*_G(pt)$-module. By proposition 2.1.24

$$h^k_G(X) \to h^k_G(F) \cong \mathbb{K}u_k 1_{N_G} \oplus \mathbb{K}u_k 1_{S_G} \ ,$$

is an isomorphism ($u_k = \omega^{k/d}$ is a generator of $h^k_G(pt)$, n.b. that k is even in cases (2) and (3)).

Thus $t(\phi_X)$ and $u_k t(1_X)$ must be linearly independent over $\mathbb{K}$. Moreover the class $i^*_X(u_k t(1_X))$ has to be zero. (Otherwise we could define a cohomology extension of the fibre $\tilde{t}$ by $\tilde{t}(\phi_X) = \lambda u_k t(1_X)$ with some $\lambda \in \mathbb{K}$, and $\tilde{t}(1_X) = 1_{X_G}$. The linear dependence of $u_k\tilde{t}(1_X)$ and $\tilde{t}(\phi_X)$) would contradict the Leray-Hirsch theorem, proposition 2.1.15). Thus, for all $c \in \mathbb{K}$ $t_c(\phi_X) := t(\phi_X) + c u_k t(1_X)$ and $t_c(1_X) = t(1_X)$ defines a cohomology extension of the fibre. By an appropriate choice of c we can achieve $r(t_c(\phi_X)) = \lambda u_k 1_{N_G}$ with some $0 \neq \lambda \in \mathbb{K}$.

If we set $s := \lambda^{-1} t_c(\phi_X)$ we verify that by the above embedding in $h^*_G(F)$

$$h^*_G(X) \cong h^*_G(pt)[s]/(s^2 = s\omega^{k/d})$$

as a (graded) $\mathbb{K}$-algebra and a $h^*(BG)$-module. (All relations can be checked on $h^*_G(F)$.) The generator s restricts to $\omega^{k/d}1_{N_G}$ on F, whereas 1 restricts to $1_{N_G} + 1_{S_G}$.

(n arbitrary)

Let π_i the projection of X on the i-th factor $X_i = S^{k_i}$. By the construction above we can chose $\tilde{s}_i \in h^{k_i}_G$, which restricts to a multiple of the orientation class of S^{k_i} under $S^{k_i} \hookrightarrow S^{k-1}_G$, and with $r(\tilde{s}_i) = u_{k_i}1_{N_G}$. We set $s_i := \pi_i^*(\tilde{s}_i)$. and note that

$$s_i^2 = \pi_i^*(\tilde{s}_i^2) = \pi_i^*(\tilde{s}_i\omega^{k_i/d}) = s_i\omega^{k_i/d} \ .$$

We do this for each $i \in \{1, \ldots, n\}$.

If we write $F = F_1 \times \cdots \times F_n$, $F_i = N_i \cup S_i$, we have the commutative diagram of G-spaces

$$\begin{tikzcd} F \arrow[r, "i_F"] \arrow[d, "\pi_i|_F"'] & X \arrow[d, "\pi_i"'] \\ F_i \arrow[r, "i_{F_i}"] & X_i \end{tikzcd}$$

and hence

$$\begin{aligned} r(s_i) = i_F^*(s_i) &= i_F^*\pi_i^*\tilde{s}_i \\ &= (\pi_i|_F)^* i_{F_i}^*\tilde{s}_i \\ &= (\pi_i|_F)^*(\omega^{k_i/d}1_{N_i}) \\ &= \sum_{x=(x_1,\ldots,x_n)\in F,\ x_i=\nu_i} \omega^{k_i/d}1_x \ . \end{aligned}$$

This means, s_i *recognises* all fixed points that project to the north pole in X_i in the sense of remark 2.1.28 and definition 2.1.29.

It ensues, that the product $s_{i_1} \cdots s_{i_l}$ recognises all fixed points that project to the north pole in x_{i_j} for $j \in \{1, \ldots, l\}$, and that the 2^n different products are linearly independent over $h^*(BG)$.

By the above, all elements of $h^*_G(X)$ are $h^*(BG)$-linear combinations of such products, and in each case we can calculate the set of fixed points recognised.

We also see, that the square free products are independent over $\mathbb{K}$ in each degree and over $h^*(BG)$ as elements of the $h^*(BG)$-module $h^*_G(X)$. Furthermore, the r-image of the d-fold squarefree products generates a sub vector space of dimension $\binom{n}{d}$ in $h^d_G(F)$, and the $h^*(BG)$-span of all products of dimension $\leq d$ has dimension $\binom{\sum_{s=0}^{d} n}{d}$ in degree d, which is equal to $\sum_{s=0}^{d} h^d(X)$. As $h^*_G(X)$ is freely generated as a $H^*(BG)$-module by the image of a cohomology extension of the fibre, the dimension of $h^d_G(X)$ is $\binom{\sum_{s=0}^{d} n}{d}$. Hence the products of the s_i generate $h^*_G(X)$ over $h^*(BG)$.

$h^*_G(X)$ is isomorphic as a $\mathbb{K}$-algebra and as a $h^*(BG)$-module to

$$h^*(BG)[s_1, \ldots, s_n]/(s_i^2 = s_i\omega^{k_i/d}, i = 1 \ldots n) .$$

Note, that in cases (1) and (3) we have $h^*(BG) \cong \mathbb{K}[\omega]$ and can rewrite this as

$$\mathbb{K}[s_1, \ldots, s_n, \omega]/(s_i^1 = s_i\omega^{k_i/d}, i = 1 \ldots n) .$$

$(\mathbf{X} = \mathbf{T^n},\ \mathbf{G} = \mathbb{Z}_\mathbf{2})$

In this case we want to make more transparent which subsets of F can occur as

$$\text{rec } (u) = \{x \in F | i_x^* u \neq 0\}$$

for some class $u \in h^*_G(X)$, that is, which subsets precisely can be recognised by some class u.

The fixed point set $F = F_1 \times \cdots \times F_n$ consists of all n-tuples $(x_1, \ldots, x_N)$ with $\{x_i\} = N_i$ or $\{x_i\} = S_i$. If we map S_i to 0 and N_i to 1, this set is bijectively mapped to $(\mathbb{Z}_2)^n$.

We now define a $\mathbb{Z}_2$-algebra homomorphism

$$\begin{aligned}\Phi :\ & h^*_G(X) \cong h^*(BG)[s_1, \ldots, s_n]/(s_i^2 = s_i\omega^{k_i}, i = 1 \ldots n) \\ & \qquad \to \mathbb{Z}_2[X_1, \ldots, X_n]/(X_i^2 = X_i, i = 1 \ldots n) ,\end{aligned}$$

such that each $p \in h^*_G(X)$ recognises precisely the fixed points that correspond to the solutions of

$$\Phi(p)(X_1, \ldots, X_n) = 1 \text{ in } (\mathbb{Z}_2)^n .$$

We define Φ by $\Phi(s_i) := X_i$ and $\Phi(\omega) := 1$ for a generator ω of the polynomial ring $h^*_G(pt)$. The solutions of $\Phi(s_i) = 1$ correspond to the points recognised by s_i, and the map Φ behaves as required under products and sums:

p_1p_2 recognises the intersection of the sets recognised by p_1 and p_2, $p_1 + p_2$ recognises precisely the points that are recognised by either p_1 or p_2, not both. The solutions of $\Phi(p_1)(X_1, \ldots, X_n) = 1$ and $\Phi(p_2)(X_1, \ldots, X_n) = 1$ are related in the same way to the solutions of $\Phi(p_1 p_2)(X_1, \ldots, X_n) = 1$ or $\Phi(p_1 + p_2)(X_1, \ldots, X_n) = 1$, respectively.

As a set, we can identify $\mathbb{Z}_2[X_1, \ldots, X_N]/(X_i^2 = X_i, i = 1 \ldots n)$ with the subset $M \subset \mathbb{Z}_2[X_1, \ldots, X_n]$ of all square-free polynomials. Each equivalence class contains exactly one member of M. If $q = \sum_j \Pi_{k=1}^{d_j} X_{i_{jk}} \in M$ is a polynomial of degree d, we define a map $\Psi : M \to h_G^*(X)$ by setting

$$\Psi(\sum_j \Pi_{k=1}^{d_j} X_{i_{jk}}) = \sum_j \omega^{d-d_j} \Pi_{k=1}^{d_j} s_{i_{jk}} \in h_G^*(X) .$$

Hence $\Phi(\Psi(q)) = q$, that means Φ is surjective. We note that the degree of $\Psi(q)$ is equal to the degree of q,

The solution set of an equation in $\mathbb{Z}_2[X_1, \ldots, X_n]/(X_i^2 = X_i)$ is equal to the solution set of its representative in M, as $x = x^2$ in $\mathbb{Z}_2$. Thus solution sets of square free polynomial equations in $\mathbb{Z}_2[X_1, \ldots, X_n]$ correspond bijectively to the possible subsets of the fixed point set recognised by some cohomology class in $h_G^*(X)$.

For our applications, given a subset A of F and a fixed point $x \notin A$, it will be useful to decide, whether $\{x\}$ can be separated from A by some cohomology class $u \in h_G^d(X)$ (as in definition 2.1.29). This translates into the question whether for some $x \in \mathbb{Z}_2^n$ and some subset $A \subset \mathbb{Z}_2^n$ there is a square free polynomial $p \in \mathbb{Z}_2[X_1, \ldots, X_n]$ of degree d such that $p(x) = 1$ and $p(y) = 0$ for all $y \in A$. We will likewise say that p separates x from A.

Obviously $x = (x_1, \ldots, x_n)$ can be separated from any A ($x \notin A$) by the polynomial $p = p_1 \cdots p_n$ with $p_i = X_i$ if $x_i = 1$, and $p_i = 1 - X_i$ if $x_i = 0$. p has degree n. (Please note, that we know that already, as $h_G^n(X) \to h_G^n(F)$ is an isomorphism by proposition 2.1.24 and corollary 2.1.27) The interesting case is in degrees $< n$.

The equation $m(x_1, \ldots, x_n) = 1$ for a monomial m of degree $d < n$ defines an affine subspace of $\mathbb{Z}_2^n$ of dimension $n - d$. Inversely, if A is contained in the complement of an affine subspace B of $\mathbb{Z}_2^n$ with $\dim B = n - d$ and $x \in B$, there is an affine isomorphism $I : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ such that $I(y) \in \{1\}^d \times \mathbb{Z}_2^{n-d}$ iff $y \in B$ Now the monomial $m := X_1 \cdots X_d$ satisfies $m(I(x)) = 1$ and $m(y) = 0$ for all $y \in I(A)$. Hence $m \circ I$ is a polynomial of degree d with the required properties. Please note that if $A = \mathbb{Z}_2^n - B$ we have $|A| = 2^n - 2^{n-d}$. By [LN83], theorem 1.6.11 $m(y) = 0$ has at least 2^{n-d} solutions. Hence $2^n - 2^{n-d}$ is the maximal cardinality of all A such that x can be separated from A by a polynomial of degree d. This means that

$$|A| > 2^n - 2^{n-d} \;\Rightarrow\; \sigma(x, A) > d .$$

Now we are looking for upper estimates on $\sigma(x, A)$ that depend only on the cardinality $N := |A|$ of A.

We distinguish different cases for the degree d and try to find some upper bounds on N that guarantee that $\sigma(x, A) \leq d$ for any $|A| = N$.

degree n Any $x \notin A$ can be separated from A by a polynomial of degree n for any N. This was already proved above:

$$\sigma(x, A) \leq n \; .$$

degree $n-1$ $x \notin A$ can be separated from A by a polynomial of degree $n-1$, iff $N \leq 2^n - 2$. This is equivalent to the existence of an affine 1-dimensional subspace in $\mathbb{Z}_2^n - A$ containing x. Any two points form a 1-dimensional affine subspace.

$$N \leq 2^n - 2 \Rightarrow \; \sigma(x, A) \leq n - 1 \; .$$

degree $d < n-1$ For any two points in A there is an affine subspace of dimension $n-1$ ("hyperplane") that contains x and does not contain the two points. For $N \leq 2d$ the intersection of d such subspaces is an affine subspace S of dimension $\geq n - d$ such that $x \in S$ and $A \cap S = \emptyset$. By the above we can separate x from A by a polynomial of degree $\leq d$.

$$N \leq 2d \Rightarrow \sigma(x, A) \leq d$$

or, equivalently

$$\sigma(x, A) \leq [(N+1)/2] \; .$$

This result is not optimal for $d > 1$.

In order to obtain better estimates for $\sigma(x, A)$ that depend on $|A|$ only we define

$$\sigma(N) := \max\{\sigma(x, A) | \; x \notin A, \; |A| = N\} \; .$$

If $x \notin A$, $k = [\log_2(N)]$ we can find an affine hyperplane of $\mathbb{Z}_2^n$ that contains at least $k+1$ points of A and does not contain $\{x\}$.

(Proof: Suppose we have l different points $x_1, \ldots, x_l$ of A, $l \leq k$, such that the affine subspace defined by these points doesn't meet $\{x\}$. Together with x these points define an affine subspace S of dimension at most l, which contains at most 2^l points. Thus $|S \cap A| = |(S \backslash \{x\}) \cap A| \leq 2^l - 1 < 2^k \leq |A|$, there is an $x_{l+1} \in A \backslash S$. The affine subspace defined by $\{x_1, \ldots, x_{l+1}\}$ does not contain $\{x\}$. In this way we can inductively find a subset of A with $k+1$ elements, such that $\{x\}$ is not a member of the affine subspace defined by these points.)

Thus
$$\sigma(N) \leq \sigma(N - [\log_2(N)] - 1) + 1 .$$
This recursion formula yields better estimates than $\sigma(N) \leq [(N+1)/2]$. We mention the estimates for the first values of N:

$$\begin{aligned} \sigma(1 \dots 2) &\leq 1 \\ \sigma(3 \dots 5) &\leq 2 \\ \sigma(6 \dots 9) &\leq 3 \\ \sigma(10 \dots 13) &\leq 4 \end{aligned}$$

We do not know an explicit formula $\sigma(N) \leq f(N)$ for the estimates obtainable in this way. ◇

Remark 2.1.35 In Example 2.1.34 the G-equivariant ordinary cohomology for $G = \mathbb{Z}_p$ can be calculated explicitly by an algebraic Borel construction as in [AP93]. ◇

2.1.3. Thom classes of invariant submanifolds

In the last section the problem of finding homology classes that separate two sets of fixed points could be reduced to an algebraic problem in some particular situations guaranteeing (TNHZ). In general, however, the problem is better understood as a geometrical problem.

Let us suppose in the following that $M \subset X$ is a G-invariant closed submanifold of the G-manifold X. Let us suppose that $\dim X = n$ and $\dim M = m$. Then, by [Bre72], VI.2.2., M has an open tubular neighbourhood U, that is a smooth G-vector bundle $\xi : \nu(M) \to M$ (actually the normal bundle with respect to a G-invariant Riemannian metric) and a G-homeomorphism $\phi : \nu(M) \to U$ such that the restriction of ϕ to the zero section is the inclusion of M in X. $\nu(M) \times EG \to M \times EG$ is a G-equivariant vector bundle over a free G-space and the quotient $\nu(M)_G \to M_G$ is a vector bundle with the same fibre.

Definition 2.1.36 *We say, that $\nu(M)_G$* **is orientable with respect to the ordinary cohomology theory** *h^*, if it has a* **Thom class***, that is, a class $\mu \in h^{n-m}(\nu(M)_G, \nu(M)_G - M_G)$, that, for all $x \in M_G$, restricts to a generator of $h^{n-m}(\nu(M)_G|_x, \nu(M)_G|_x - 0)$.*

The class in $h_G^{n-m}(X, X - M) \cong h^{n-m}(X_G, X_G - M_G)$ that is mapped to μ by the excision isomorphism will be called the **normal G-Thom class of M in X***, and we will say that M* **is normally G-orientable***.*

If $\nu(M)_G$ is orientable, we have a Thom-isomorphism

$$\theta : h_G^*(M) \to h_G^{*+n-m}(\nu(M), \nu(M) - M) \cong h_G^{*+n-m}(X, X - M)$$

defined by the the cup-product with μ.

Remark 2.1.37 In general the orientiability of $\nu(M)_G$ is a subtle question. In some special cases, however, we can easily guarantee orientability:

i) If h^* is ordinary cohomology with $\mathbb{Z}_2$-coefficients, $\nu(M)_G$ is h^*-orientable anyway.

ii) If $\nu(M)$ has a complex structure that is compatible with the G-operation, $\nu(M)_G \to M_G$ will inherit a complex structure and hence be a orientable vector bundle (in the ordinary sense). Thus it is orientable for any ordinary cohomology.

iii) If $G = \mathbb{Z}_p$, $h^* = H^*(\cdot, \mathbb{Z}_p)$, p odd or $G = S^1$ and M is an isolated G-manifold of fixed points, it is normally G-orientable.

   We see this, as follows. M has the equivariant normal bundle $\nu(M)$, the fibres of $\nu(M)$ are representations of G with no trivial subrepresentations. This uniquely defines a complex structure on the fibres of $\nu(M)$, which is compatible with the G-operation. By the last remark we conclude that M is normally G-orientable.

Proposition 2.1.38 *Let G be a compact Lie group that acts smoothly on an n-dimensional manifold X, $M \subset X$ a G-invariant m-dimensional closed submanifold and h^* an ordinary cohomology theory with coefficients R, such that M is normally orientable and the G-action is (CSF). Suppose that there is a neighbourhood U of M in X such that all fixed points of U are contained in M. Let*

$$\mu \in h_G^*(X, X - M)$$

be the normal Thom class of M in X.

Suppose that $h_G^*(pt)$ has unbounded cohomology. Then $\mu|_X \in h_G^*(X)$ separates every component of the fixed point set of M from the fixed point set of $X - M$.

Proof: We may assume that an open tubular neighbourhood of M is contained in U. Let $\tilde{\mu}$ be the Thom class of $\nu(M)_G$ that corresponds to μ under the excision isomorphism.

Let $C \subset M$ be a component of the fixed point set of M. Then, by proposition 2.1.16, $h_G^*(C) \cong h^*(C) \otimes_R h_G^*(pt)$. By assumption $h_G^*(pt)$ is unbounded, hence

$h_G^*(C)$ is unbounded, in particular there is a non zero class $\alpha \in h_G^N(C) = h^N(C_G)$ for some $N > n$.

The normal bundle $\nu(M)$ restricts to a vectorbundle $\xi : E \to C$ over C. The Thom class $\tilde{\mu} \in h^{n-m}(\nu(M)_G, \nu(M)_G - M_G)$ restricts to a generator of $h^{n-m}(F, F - 0)$ on every fibre. Hence $\hat{\mu} := \tilde{\mu}|_{(E_G, E_G - C_G)}$ is a Thom class for ξ, thus $0 \neq \hat{\mu} \cup \alpha \in h^{n-m+N}(E_G, E_G - C_G) \cong h_G^{n-m+N}(E, E - C)$. By proposition 2.1.24 and the assumption (CSF) the restriction to the fixed point set is an isomorphism above degree n. Thus

$$0 \neq (\hat{\mu} \cup \alpha)|_C = \hat{\mu}|_C \cup \alpha = \tilde{\mu}|_C \cup \alpha \in h_G^*(C) \; .$$

As the excision isomorphism $h_G^*(X, X - M) \to h_G^*(\nu(M), \nu(M) - M)$ is induced by the restriction, we have

$$\mu|_C = \tilde{\mu}|_C \neq 0 \; .$$

On the other hand, if C is a component of the fixed point set of X disjoint from M, the restriction of μ to C factors through

$$0 \cong h_G^*(X - M, X - M) \; .$$

We have thus verified that μ separates the fixed point set of M from the fixed point set of $X - M$. □

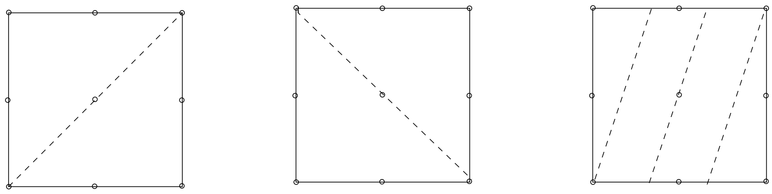
Remark 2.1.39 The proof actually proves a bit more, namely that $\mu \cup \alpha|_C \neq 0$ on components C of the fixed point set for arbitrary cohomology classes α of sufficiently high degree. In the cases $G = \mathbb{Z}_2$, $G = \mathbb{Z}_p$, p odd, or $G = S^1$ this implies that the restriction of μ to any component of the fixed point set must in fact be a power of the free polynomial generator of $h_G^*(pt)$, and hence also the restriction to any fixed point within.

For $G = \mathbb{Z}_p$, p odd, $h_G^*(pt)$ is an algebra generated by one free generator and one alternating generator, and the statement is related to the fact that $\mathbb{Z}_p$-representations without trivial summands are even-dimensional. ◇

Corollary 2.1.40 *Under the assumptions of the proposition above, we have the estimate*

$$\sigma(M^G, (X - M)^G) \leq n - m \; .$$

Remark 2.1.41 The corollary allows to obtain once more the estimates of example 2.1.34. Just note for example, that a subset of the fixed point set of T^n corresponding to an m-dimensional affine subspace of $\mathbb{Z}_2^n$ is actually contained in an m-dimensional invariant submanifold of T^n. However, there may be many different submanifolds containing the same subset of the fixed point set, all yielding different Thom classes, but the same restriction to the fixed point set.



Three different G-submanifolds of T^2 containing the same two fixed points. $\diamond$

Remark 2.1.42 Please note, that these Thom classes that separate M^G from $(X-M)^G$ satisfy a stronger property, namely their restriction to $X-M$ is zero. As we may eventually need the stronger property, we define a variant of σ. $\diamond$

Definition 2.1.43 *Let A and B be two disjoint G-invariant subsets of the G-space X. We define*

$$\sigma_T(A,B) = \min\{n \geq 0 | \exists \alpha \in h^n_G(X),\ A^G \subset \text{rec }(\alpha) \textit{ and } \alpha|_B = 0\}\ .$$

Remark 2.1.44 If $A, B \subset X^G$, $\sigma_T(A,B) = \sigma(A,B)$. $\diamond$

Proposition 2.1.45 *Under the assumptions of proposition 2.1.38 we have*

$$\sigma_T(M, X-M) \leq n-m$$

$\square$

The normal Thom class of an invariant submanifold is not zero, when restricted to the fixed points in this submanifold. This implies under certain assumption that for a (TNHZ) submanifold M the cohomology $h^*_G(X, X-M)$ is embedded in $h^*_G(M^G)$. We will need this later for $M \subset X^G$.

Proposition 2.1.46 *Let the assumptions of proposition 2.1.38 hold. Assume furthermore that M is a (TNHZ) submanifold of X, $h^*(M)$ is a finitely generated k-vector space, $h^*(BG)$ is finitely generated as a k-algebra and has an element $\omega \in h^d(BG)$, $d > 0$ such that the product with ω is an isomorphism. Then the inclusion $M \to (X, X-M)$ induces an injective homomorphism in h^*_G.*

Proof: By proposition 2.1.38 the Thom class $\mu \in h^*_G(X, X-M)$ restricts to a non zero class on M^G, hence $\mu|_M$ is non zero as well. Let U be a tubular G-invariant neighbourhood of M in X with $U^G = M^G$ and $\pi : U \to M$ the normal bundle projection. By the Thom isomorphism, every nonzero element of $h^*_G(X, X-M) \cong h^*_G(U, U-M)$ has the form $\mu|_{(U,U-M)} \cup \pi^*(\alpha)$ for some $0 \neq \alpha \in h^*_G(M)$. By the Leray-Hirsch theorem (proposition 2.1.15) $h^*_G(M)$ is a free $h^*(BG)$-module and thus $\omega^s\alpha \neq 0$ for all $s \geq 0$, and by the Thom isomorphism $0 \neq \mu|_{(U,U-M)} \cup \pi^*(\omega^s\alpha)$. By corollary 2.1.26 the homomorphism $h^*_G(U, U-M) \to h^*_G(M^G)$ is an ismorphism above some degree. Thus $0 \neq \mu|_{M^G} \cup (\omega^s\alpha)|_{M^G}$ and $0 \neq \mu|_{M^G} \cup \alpha|_{M^G}$, and the injectivity is proved. $\square$

Remark 2.1.47 This implies that under the above assumptions $h_G^*(X, X - M) \to h_G^*(X)$ is injective. In the terminology of Atiyah and Bott [AB82] the relative classes are **self-completing**. The above proposition is partially implied by a dualization of proposition 1.9. of Atiyah and Bott if M is a manifold of fixed points. Our conditions guarantee that the equivariant Euler class (which is the restriction of the equivariant Thom class to the zero section) is not a zero divisor. $\diamond$

2.2. Critical point theory on a Finsler manifold

2.2.1. General setting

In this section X will be a $C^{1,1}$ Finsler G-manifold and $f : X \to \mathbb{R}$ a G-invariant C^1-functional.

For $c \in \mathbb{R}$ we denote the sublevel sets by

$$f^c := \{x \in X |\ f(x) \leq c\}$$

and the critical set on the level c by

$$K_c := \{x \in X |\ Df(x) = 0\ ,\ f(x) = c\}\ .$$

and the overall critical set by K. For a subset $A \subset X$ we also write

$$A^c := A \cap f^c\ ,$$

when there is no doubt, which functional f we mean.

For such functionals we have an *(equivariant) deformation lemma* as follows:

Theorem 2.2.1 *[Str90] Suppose X is a complete $C^{1,1}$-Finsler manifold and $f \in C^1(X)$ satisfies the (PS) condition on level β for some $\beta \in \mathbb{R}$. Let $\bar{\epsilon} > 0$ be given and let N be a closed neighbourhood of K_β. Then there exists as number $\epsilon \in]0, \bar{\epsilon}[$ and a continuous 1-parameter family of homeomorphisms $\Phi(\cdot, t)$ of X, $0 \leq t < \infty$, with the properties*

i) $\Phi(u, t) = u$, if $t = 0$ or $Df(u) = 0$ or $|f(u) - \beta| \geq \bar{\epsilon}$;

ii) $f(\Phi(u, t))$ is non-increasing in t for any $u \in X$;

iii) $\Phi(f^{\beta+\epsilon} \backslash N, 1) \subset f^{\beta-\epsilon}$, and $\Phi(f^{\beta+\epsilon}, 1) \subset f^{\beta-\epsilon} \cup N$.

Moreover, Φ has the semigroup property $\Phi(\cdot, s) \circ \Phi(\cdot, r) = \Phi(\cdot, s + t)$, $\forall s, t \geq 0$. If X admits a compact group of symmetries G and if f is G-invariant, Φ can be constructed to be G-equivariant.

We note also a second equivariant deformation lemma:

Theorem 2.2.2 *([Cha93], Theorem 3.2 and 7.2) Let M be a C^2 Finsler manifold with a continuous G-operation. Suppose that $f \in C^1(M, \mathbb{R})$ is G-invariant satisfies the $(PS)_c$ condition for all $c \in [a, b]$ and that a is the only critical value of f in $[a, b[$. Assume that the connected components of K_a are parts of critical G-orbits. Then f^a is a G-equivariant strong deformation retract of $f^b - K_b$.*

The deformation theorems enable us to do equivariant Lusternik-Schnirelman theory on X. We intend to investigate an intermediary situation with non-degenerate critical fixed points and G-orbits of degenerate critical points characterised by the vanishing of certain equivariant cohomology classes constructed above.

Definition 2.2.3 *For a G-manifold X and a cohomology class $\alpha \in h^*_G(X)$ we define*

$$c(\alpha) := \inf\{c \in \mathbb{R} | \alpha|_{X^c} \neq 0\} .$$

It should be understood, that the infimum is $+\infty$, if $\alpha = 0$.

Proposition 2.2.4 *Suppose X is a complete $C^{1,1}$-Finsler G-manifold and $f \in C^1(X)$ a G-invariant function.*

If $\alpha \in h^*_G(X)$ and $c(\alpha) \in \mathbb{R}$, and if f satisfies the Palais-Smale condition on level $c(\alpha)$, then $c(\alpha)$ is a critical value of f.

Proof: Suppose $c := c(\alpha) \in \mathbb{R}$ is a regular value of f. In this case the deformation theorem (Theorem 2.2.1) can be applied with $N = \emptyset$. We have a one-parameter family of G-homeomorphisms $\Phi(\cdot, \cdot)$ with $\Phi(u, 0) = u$ and $\Phi(f^{c+\epsilon}, 1) \subset f^{c-\epsilon}$. By definition we have

$$0 \neq \alpha|_{f^{c+\epsilon}} \text{ and } 0 = \alpha|_{f^{c-\epsilon}} . \tag{2.8}$$

Consider

$$g_t : f^{c+\epsilon} \to f^{c+\epsilon} : \ u \mapsto \Phi(u, t) .$$

By homotopy invariance of our cohomology theory

$$\alpha|_{f^{c+\epsilon}} = g_0^* \alpha|_{f^{c+\epsilon}} = g_1^* \alpha|_{f^{c+\epsilon}} = g_1^* \alpha_{f^{c-\epsilon}} ,$$

which contradicts (2.8). $\square$

Definition 2.2.5 *For a topological space X and a non-empty subspace $A \subset X$ we define the Lusternik-Schnirelman category*

$$\mathrm{cat}_X A := \min\{k \geq 0 | \ A \subset U_1 \cup \cdots \cup U_k, \ \ U_i \textit{ open and contractible in } X\} .$$

Remark 2.2.6 According to Borsuk, a metrisable space is an ANR (absolute neighbourhood retract), iff it is an ANE (absolute neighbourhood extensor) (s. e. g. [HW48], p. 86,). For an ANR/ANE we can use closed sets instead of open sets in the definition of LS category, as the contraction of a closed set in X can be extended to a contraction of an open neighbourhood. For closed subsets of an ANR/ANE, LS category satisfies a certain continuity property, as follows.

Please note that the definition of LS category used by homotopy theorists usually differs from ours by -1, so that in their terms contractible sets have category 0.$\diamond$

Proposition 2.2.7 *([MW89], Lemma 4.7) If A is a closed subspace of an ANR X, there is a closed neighbourhood U of A such that*

$$\mathrm{cat}_X(U) = \mathrm{cat}_X(A) \ .$$

Definition 2.2.8 *Let X be a topological space and $u \in h^*(X)$. We define the* **category weight** *of u as*

$$\mathrm{cwgt}\ (u) := \sup\{k | u|_A = 0 \quad \text{for every closed subspace } A \subset X \text{ with } \mathrm{cat}_X(A) \leq k\} \ .$$

The **strict category weight** *of $u \in h^*(X)$ is defined as*

$$\mathrm{swgt}\ (u) := \min\{\mathrm{cwgt}\ (f^*u)|\ f : Z \to X,\ Z \quad \textit{a normal space}\} \ .$$

Remark 2.2.9 The definition of category weight was originally given by Fadell and Husseini in [Fad92] for ANRs (with respect to separable metric spaces). Category weight, however, is not homotopy invariant. This motivated the equivalent definitions of *essential category weight* by Strom [Str02] and of *strict category weight* by Rudyak [Rud99],[Rud98]. However in their work they usually assume the spaces to be CW complexes, and it is not obvious to which class of spaces their results apply.

As we would like to estimate the category of subsets of paracompact Banach manifolds, which are metrisable and hence normal, the above definition of strict category weight is sufficient. By [FHT01], proposition 27.8., for a normal subset A of a normal topological space X, the condition $\mathrm{cat}_X A \leq m$ is equivalent to the existence of a lifting of the inclusion $A \to X$ to the second Ganea fibration of X. Thus the proof of theorem 3.6. of [Fad92] proves in fact the following proposition. $\diamond$

Proposition 2.2.10 *Let h^* be Čech cohomology with $\mathbb{Z}_p$-coefficients. Let X be a paracompact space and $\omega \in h^1(X)$ and $\beta(\omega) \in h^2(X)$ the image of that class under the Bockstein operator. If $\beta(\omega) \neq 0$ we have*

$$\mathrm{swgt}\ (\beta(\omega)) \geq 2 \ .$$

Proof: We sketch the adaption of Fadell and Husseini's proof to our situation. According to Huber [Hub61] Čech cohomology of paracompact spaces is represented as a homotopy functor, in particular

$$h^1(X) = [X, K(1, \mathbb{Z}_p)] \ ,$$

where $Y := K(1, \mathbb{Z}_p)$ is an Eilenberg-MacLane space. Let i_1 be the distinguished element of $h^1(Y)$ represented by the identity. $0 \neq \omega$ is represented by an essential map $g : X \to Y$, and $g^* i_1 = \omega$. Let $f : Z \to X$ be a continuous map from a normal space Z to X and $i_A : A \to Z$ the inclusion of a closed, thus normal subset A into Z. Chose a basepoint $*$ in A and accordingly in Z and Y such that all maps are pointed. We set $E_A = \mathcal{P}(Z, A) \to A$, the space of paths in Z with initial point $*$ and end point in A. We define as well the spaces $E_Y = \mathcal{P}(Y)$ and $E_Z = \mathcal{P}(Z)$ of paths with initial point $*$ and consider the commutative diagram of fibrations:

$$\begin{tikzcd}
E_A *_A E_A \arrow[r] \arrow[d] & E_Z *_Z E_Z \arrow[r] \arrow[d] & E_Y *_Y E_Y \arrow[d] \\
A \arrow[r, "i_A"] & Z \arrow[r, "g \circ f"] & Y
\end{tikzcd}$$

If $\mathrm{cat}_Z A \leq 2$ the second Ganea fibration $E_A *_A E_A \to A$ has a section. Hence $(g \circ f) \circ i_A$ factors through $E_Y *_Y E_Y \simeq S\Omega Y$.

But ΩY has p contractible components, therefore $S\Omega Y$ is homotopy equivalent to a wedge of circles and $h^q(S\Omega Y) = 0$ for $q > 1$. The cohomology operation β commutes with all these maps and we conclude $\beta\omega = 0$. $\square$

Now we come to a slightly awkward definition of a *topologically non-degenerate critical manifold*. We would have preferred to characterise this property by the Conley index, but we need actually a bit more than the homotopy type of N/L for an index pair (N, L), as we require the map $N_0 \to N$ itself to be equivalent to the injection of a sphere bundle into a disk bundle and that L can be chosen as a sublevel set $N^{c-\epsilon}$. This is, of course, satisfied whenever the Hessian of a C^2-functional on a Hilbert manifold is normally hyperbolic.

Definition 2.2.11 *If X is a paracompact C^1-Finsler manifold and if f is a C^1-function on X, a submanifold $M \subset X$ is called a* **topologically non-degenerate critical submanifold**, *if f is critical on M, $f|_M = c$ is constant and there is a subset W of X^c which is diffeomorphic to a k-dimensional normed vector bundle $E^- \subset \nu(M)$ over M such that inside every neighbourhood of M there is a closed neighbourhood U of M in X with critical set M, and there is $\bar{\epsilon} > 0$ such that for all $0 < \epsilon < \bar{\epsilon}$*

$$(U, U^{c-\epsilon}) \overset{\simeq}{\leftarrow} (U^c, U^{c-\epsilon}) \overset{\simeq}{\leftarrow} (W \cap U^c, W \cap U^{c-\epsilon}) \overset{\simeq}{\rightarrow} (W, W^{c-\epsilon}) \overset{\simeq}{\leftarrow} (DE^-, SE^-)$$

where the first three homotopy equivalences are induced by inclusions. We call such a closed neighbourhood U **distinguished**.

If X is equipped with a G-operation we define accordingly a **topologically equivariantly non-degenerate critical submanifold** *as a submanifold like* *above, which furthermore is G-invariant and require that U can be chosen to be G-invariant, W a G-space and E^- a G-vector bundle. All maps are understood to be G-maps, and all homotopy equivalences G-homotopy equivalences.*

In both cases k is called the **Morse index** $i(M)$ *of the submanifold. If (DE^-, SE^-) is (equivariantly) orientable we call the the non-degenerate critical submanifold* **(equivariantly) orientable in the unstable direction***.*

Remark 2.2.12 If, in the above definition, we fix W and chose a second distinguished neighbourhood V inside a distinguished neighbourhood U, the inclusion $(V, V^{c-\epsilon}) \to (U, U^{c-\epsilon})$ is a homotopy equivalence, hence induces an isomorphism in cohomology.

The definition of a *dynamically isolated critical set* by Chang and Ghoussoub ([CG96]) will help us to construct *distinguished neighbourhoods* and to compare our definition with well known local invariants like the Conley index.

Definition 2.2.13 *([CG96], Def. I.10) Suppose we are given a connected paracompact complete $C^{1,1}$-Finsler manifold X and a C^1-function $f : X \to \mathbb{R}$.*

V is a **pseudo-gradient field** *for f, if it is a locally Lipschitz vector field V on X with*

$$\|V(x)\| \leq A\|Df(x)\| \,, \tag{2.9}$$
$$Df(x)(V(x)) \geq B|Df(x)\|^2 \,, \tag{2.10}$$

for some constants $A, B > 0$ and all $x \in X$. (For the existence of pseudo-gradient fields, s. Remark 2.2.14 below.)

Let $\eta : \mathbb{R} \times X \to X$ be the global flow of the locally Lipschitz vector field

$$V_1(x) := -\min\{\mathrm{dist}(x, K), 1\} \frac{V(x)}{\|V(x)\|_x} \,.$$

For any closed set $O \subset X$ we define

$$\tilde{O} := \bigcup_{t \in \mathbb{R}} \eta(t, O) \,.$$

A subset S of the critical set K is said to be a **dynamically isolated critical set***, if there exists a closed neighbourhood O of S and regular values $\alpha < \beta$ of f such that*

$$O \subset f^{-1}([\alpha, \beta])$$

and

$$\overline{\tilde{O}} \cap K \cap f^{-1}([\alpha, \beta]) = S \ .$$

We shall then say that (O, α, β) *is an* **isolating triplet** *for* S.

Remark 2.2.14 1. Pseudo-gradient fields for a C^1 functional always exist on $X \backslash K$. (In most contexts, e.g. [Str90], pseudo-gradient fields are defined on $X \backslash K$ only.)

Further conditions are needed to guarantee the overall existence, as required by our definition.

In our main application the functional will be $C^{1,1}$, so the gradient field can serve as a pseudo-gradient field. However, it will be essential (Proposition 2.2.20), that we can chose a different pseudo-gradient field in neighbourhoods of certain critical points, in order to prove that they are non-degenerate. This will provide a criterion for non-degeneracy (Corollary 2.2.21) tailored for our needs.

2. Ghoussoub's and Chang's definition of a pseudo-gradient vector field assumes $B = 1$. If our definition is satisfied with a number $B < 1$ we can multiply V by $1/B$ to obtain a vector field that satisfies their definition. In many publications the factors required are $A = 2$ and $B = 1/2$, which is too restrictive for some application. The factors. however, are immaterial to the constructions. $\diamond$

Proposition 2.2.15 *Suppose, as above, we are given a paracompact* $C^{1,1}$*-Finsler manifold* X, *a* C^1*-function* $f : X \to \mathbb{R}$, *and* V, V_1, η *as above. Suppose that* f *satisfies the (PS) condition.*

If $M \subset f^{-1}(c) \subset X$ *is a subset of the critical set* K *of* X *and* N *is a closed neighbourhood of* M *such that* $K \cap N = M$, *then* M *is a* dynamically isolated critical set *and there is an* $\bar{\epsilon} > 0$ *such that* $(O, c - \bar{\epsilon}, c + \bar{\epsilon})$ *with* $O := N \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}])$ *is an isolating triplet for* M.

For all $\epsilon \in]0, \bar{\epsilon}]$ *the inclusion of pairs*

$$\begin{aligned} &(\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]), \tilde{O} \cap f^{-1}(c - \bar{\epsilon})) \\ &\to (\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]), \tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c - \epsilon])) \end{aligned}$$

is a homotopy equivalence.

If $N' \supset N$ *with* $N' \cap K = M$ *and* $\operatorname{dist}(\partial N, \partial N') > 0$, *the set* $\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}])$ *is contained in* N' *for* $\bar{\epsilon}$ *small enough.*

If X *is a* G*-manifold and* f *a* G*-invariant map, all constructions are in the category of* G*-spaces and* G*-maps.*

Proof: As f satisfies (PS), the critical set K_c on level c is compact, and M is a closed and open subset of K_c and compact. Chose some $\delta > 0$. Again by (PS) the set $K' := K \cap f^{-1}([c-\delta, c+\delta])$ is compact. Set

$$\alpha := \operatorname{dist}(N, K' \backslash M) > 0 \ .$$

Let V be the closed $\alpha/2$-neighbourhood of $K'\backslash M$. By its construction $\operatorname{dist}(V, N) \geq \alpha/2$.

As f satisfies the (PS) condition, $\|Df\|$ is bounded below by some $\beta > 0$ on $A := f^{-1}([c-\delta, c+\delta]) \backslash (\mathring{V} \cup \mathring{N})$, and A has a positive distance $\gamma > 0$ to K'. Thus

$$\begin{aligned} -Df(x)(V_1(x)) &= \min\{\operatorname{dist}(x, K), 1\} \frac{Df(x)(V(x))}{\|V(x)\|_x} \\ &\geq \gamma \frac{B\|Df(x)\|^2}{\|V(x)\|} \\ &\geq \frac{\gamma B}{A} \|Df(x)\| \\ &\geq \frac{\beta\gamma B}{A} \ . \end{aligned} \tag{2.11}$$

We chose $0 < \bar{\epsilon} \leq \delta$ so small that

$$T := 2\bar{\epsilon} \frac{A}{\beta\gamma B}$$

satisfies $T < \alpha/2$.

Set $O := N \cap f^{-1}([c-\bar{\epsilon}, c+\bar{\epsilon}])$.

We claim that

$$\eta([-T, T], O) \cap f^{-1}([c-\bar{\epsilon}, c+\bar{\epsilon}]) = \tilde{O} \cap f^{-1}([c-\bar{\epsilon}, c+\bar{\epsilon}]) \ . \tag{2.12}$$

and this set is disjoint from $V \cap f^{-1}([c-\bar{\epsilon}, c+\bar{\epsilon}])$

For the equation (2.12) "$\subset$" is obvious, we only have to prove "$\supset$".

Suppose we have a point $y \in \tilde{O} \cap f^{-1}([c+\bar{\epsilon}, c-\bar{\epsilon}])$, which is not contained in O. Then there is an $x \in \partial O$ an $t \in \mathbb{R}^+$ (or $t \in \mathbb{R}^-$) such that $\eta(t, x) = y$ and $\eta(]0, t[, x) \notin \mathring{O}$ (or $\eta(]t, 0[, x) \notin \mathring{O}$).

If this trajectory $\eta(]0, t[, x)$ (or $\eta(]t, 0[, x)$) does not meet V, we have

$$2\bar{\epsilon} \geq |f(y) - f(x)| \geq \frac{\beta\gamma B}{A} |t| \ ,$$

hence $|t| \leq T$, which proves that y is contained in the left hand side of (2.12).

Now it remains to prove that the trajectory does indeed not meet V. Suppose it does. In this case the length of the trajectory must be $\geq \alpha/2$, which by $\|V_1\| \leq 1$ entails

$$|t| > |t_0| > \alpha/2 > T \; ,$$

if t_0 is the first entry time of the trajectory into V. Then

$$|f(y) - f(x)| \geq |f(\eta(t_0, x)) - f(x)| \geq \frac{\beta\gamma B}{A}|t_0| \geq \frac{\beta\gamma B}{A}T > 2\bar{\epsilon} \; ,$$

a contradiction.

In particular the critical set in

$$\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]) = \overline{\tilde{O}} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}])$$

is precisely M. We conclude that $(O, c - \bar{\epsilon}, c + \bar{\epsilon})$ is an isolating triplet for M.

The flow yields a deformation retraction from $\tilde{O} \cap f^{-1}([c-\bar{\epsilon}, c-\epsilon])$ to $\tilde{O} \cap f^{-1}(c-\bar{\epsilon})$, which induces a homotopy equivalence.

Now it remains to prove the last claim, namely that for $N' \supset N$ with $N' \cap K = M$ and $\text{dist}(\partial N, \partial N') > 0$, the $\bar{\epsilon}$ can be chosen so that $\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]) \subset N'$. If such an N' is given, we set $V := X \backslash \mathring{N}'$. Then V is closed neighbourhood of K'

$$\text{dist}(V, N) = \text{dist}(\partial V, \partial N) = \text{dist}(\partial N', \partial N) \; .$$

If we set $\alpha := 2 \; \text{dist}(V, N)$ the proof remains valid, and by construction $\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]) \cap V = \emptyset$, hence

$$\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]) \subset \mathring{N}' \subset N' \; .$$

$\square$

Remark 2.2.16 For a critical submanifold M on the level c we define the unstable set $W := \{x \in X | \; \lim_{t \to -\infty} \eta(t, x) \in M\}$. If the inclusion

$$(W \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]), W \cap f^{-1}(c - \bar{\epsilon})) \to (\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]), \tilde{O} \cap f^{-1}(c - \bar{\epsilon}))$$

induces a homtopy equivalence and if

$$\Big(W \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}]), W \cap f^{-1}(c - \bar{\epsilon})\Big) \cong (DE^-, SE^-) \; ,$$

then $\tilde{O} \cap f^{-1}([c - \bar{\epsilon}, c + \bar{\epsilon}])$ is a distinguished neighbourhood for M and M is topologically non-degenerate in our sense. $\diamond$

Remark 2.2.17 (X,f,η as in Proposition 2.2.15.)

We want to compare the Morse index in Definition 2.2.11 with the cohomological Morse-Conley index.

Suppose $M \subset K_c$ is a clopen subset of K, and U a closed neighbourhood of M such that $U \cap K = M$ and $f(U)$ is bounded.

As M is compact, U contains an α-neighbourhood of M for some $\alpha > 0$. . If we set $N := \overline{U_{\alpha/2}(M)}$ and $N' := U$, Proposition 2.2.15 yields that M is a dynamically isolated critical set, and $(O, c-\bar{\epsilon}, c+\bar{\epsilon})$ is an isolating triplet, such that

$$O' := \tilde{O} \cap f^{-1}([c-\bar{\epsilon}, c+\bar{\epsilon}]) \subset \mathring{U} \; .$$

We observe that

$$(O' \cap f^{-1}([c-\epsilon, c+\bar{\epsilon}]), O' \cap f^{-1}(c-\epsilon))$$

is a Conley (and Gromoll-Meyer) index pair for the isolating neighbourhood U of M for any $0 < \epsilon \leq \bar{\epsilon}$ (see [CG96], Section III and Theorem IV.3).

Let M be a topologically non-degenerate critical submanifold. Then a distinguished neighbourhood U contains a neighbourhood O' as above, which contains again a distinguished neighourhood U', which contains a neighbourhood O'' as above. Consider the sequence of injections

$$(U, U^{c-\epsilon}) \xleftarrow{k} (O', (O')^{c-\epsilon}) \xleftarrow{j} (U', (U')^{c-\epsilon}) \xleftarrow{i} (O'', (O'')^{c-\epsilon})$$

and the induced homomorphisms

$$h^*(U, U^{c-\epsilon}) \xrightarrow{k^*} h^*(O', (O')^{c-\epsilon}) \xrightarrow{j^*} h^*(U', (U')^{c-\epsilon}) \xrightarrow{i^*} h^*(O'', (O'')^{c-\epsilon}) \; .$$

By Remark 2.2.12 $j^* \circ k^*$ is an isomorphism, hence k^* is injective and j^* is surjective. The proof of Proposition II.2 in [CG96] gives more than the abstract isomorphism between $h^*(O', (O')^{c-\epsilon})$ and $h^*(O'', (O'')^{c-\epsilon})$, it allows to conclude that the ismorphism is induced by the inclusion $j \circ i$. Thus $i^* \circ j^*$ is an isomorphism, j^* is surjective.

If M is a topologically non-degenerate critical point

$$h^*(O', (O')^{c-\epsilon}) \cong h^*(U', (U')^{c-\epsilon}) \cong \begin{cases} R & \text{for} \quad * = k \; , \\ 0 & \text{else} \; . \end{cases}$$

Accordingly, for a topologically non-degenerate manifold M that is orientable in the unstable direction fo h^*, we get

$$h^*(O', (O')^{c-\epsilon}) \cong h^*(U', (U')^{c-\epsilon}) \cong h^{*-k}(M) \; ,$$

where k is the degree of the Thom class. (The last isomorphism is essentially the cup product with this Thom class.) A numerical Morse index can be defined as the degree of the Thom class.

It is the same number as the Morse index in our definition of a topologically non-degenerate submanifold. $\diamond$

Remark 2.2.18 When we refer to the (equivariant) cohomological Morse index of a possibly degenerate isolated critical point p we mean the (equivariant) cohomology $h^*(N, L)$ of a Conley index pair (N, L) for some $p \in U \subset \tilde{O} \cap f^{-1}([\alpha, \beta])$ where (O, α, β) is an isolating triplet for p. By [CG96], Proposition II.2 and Theorem IV.3 this index is well-defined (up to isomorphy). $\diamond$

Remark 2.2.19 We ought to state some well-known criteria which guarantee, that the conditions of Definition 2.2.11 are satisfied.

1. If X is a Hilbert-Riemann manifold, $p \in X$ and $f \in C^2(X, \mathbb{R})$ with $Df(p) = 0$ the second derivative $D^2 f(p) \in L(T_{0_p}(X))$ by means of the Riemannian connection corresponds to a symmetric bilinear map ("the Hessian") $T_pX \times T_pX \to \mathbb{R}$ and this via the Hilbert structure to a self-adjoint linear map $L : T_pX \to T_pX$. If L has a bounded inverse, p is a nondegenerate critical point in our sense and the index of L, which is the Morse index in the ordinary sense, is equal to the Morse index in our definition.

   The proof proceeds thus: After applying the Morse lemma we can consider f to be a non-degenerate quadratic form. The homotopy equivalences are the consequences of deformation retractions, which are well known (s. e.g. the proofs of Chang [Cha93], Theorems 4.1 and 4.4).

2. Assume X is a C^2-Hilbert manifold, $f \in C^2(M, \mathbb{R})$ and $M \subset X$ is a connected submanifold with $Df|_M = 0$ and $f|_M = \text{const.}$ such that there is a neighbourhood U of M with critical set M. Like above $D^2 f(X)$ corresponds to a selfadjoint operator $L_x \in L(T_xX)$. Assume furthermore, that there is an $\epsilon > 0$ with

   $$\sigma(L_x) \cap ([-\epsilon, \epsilon] - \{0\}) = \emptyset.$$

   and $\dim\ \ker L_x = \dim\ M$ for all $x \in M$.

   The normal bundle of M is invariant under the operation of L_x, $x \in M$, hence we obtain a section of operators $x \mapsto \tilde{L}_x$ on the normal bundle. If for all $x \in M$ the spectrum of $\tilde{L}_x$ does not contain zero, $\tilde{L}_x$ has a continuous inverse by the above assumptions. We then can define deformation retractions, fibrewise like above, to prove that M is an isolated non-degenerate critical submanifold in our sense (s. [Cha93], section 7.3).

3. There are different sets of conditions on an isolated critical point p of a C^2-functional on a C^2-Finsler manifold which imply that this point is non-degenerate in our sense (s. e.g. Chang [Cha93], Definition 1.4.3. and the ensuing remark).

   There is a neighbourhood of p, on which TM is trivialised as $U \times E$, E a Banach space, such that there is a hyperbolic operator L (= a continuous operator whose spectrum has positive distance to the imaginary axis) with

   a) $D^2f(p)(Lv, w) = D^2f(p)(v, Lw)\ \forall v, w \in E$,

   b) $D^2f(p)(Lv, v) > 0\ \forall v \in E - \{0\}$,

   c) $Df(x)(Lv) > 0\ \forall x \in U,\ x = p + v$.

   A Morse lemma can be proved like in the Hilbert case, and the claim follows with an adaption of the arguments sketched above. $\diamond$

We will need the topological non-degeneracy of a critical point in a particular situation, where the functional is not C^2 and not even twice Fréchet differentiable at the critical point. Therefore the Morse lemma cannot be used and the easy argument above does not apply. We will see, however, that weaker conditions guarantee the same topological structure. The following proposition is tailored for our needs, we did not bother to find the weakest assumptions under which the conclusion remains valid.

Proposition 2.2.20 *Suppose f is a $C^{1,1}$-functional defined on a separable Hilbert-space H with an isolated critical point at 0 and $f(0) = 0$. Let U be an open r-neighbourhood of 0.*

Suppose that L is a bounded self-adjoint invertible operator, which defines the smooth quadratic form

$$g : H \to \mathbb{R},\ v \mapsto g(v) = \frac{1}{2}\langle v, Lv\rangle \ ,$$

and that the dimension of the negative eigenspace E^- of L is finite ($k \in \mathbb{N}$).

Now suppose that on the set U the gradient ∇g is a pseudo-gradient field for f (Definition 2.2.13).

Then there is an isolating triplet $(O, -\epsilon, \epsilon)$ for f with respect to the global flow of a normalised pseudo-gradient field for f (as in Definition 2.2.13), such that

$$\tilde{O} \cap f^{-1}([-\epsilon, \epsilon]) \subset U$$

and

$$\begin{aligned} &(\tilde{O} \cap f^{-1}([-\epsilon, \epsilon]), \tilde{O} \cap f^{-1}(-\epsilon)) \\ \simeq\ &(W \cap f^{-1}([-\epsilon, \epsilon]), W \cap f^{-1}(-\epsilon)) \\ \simeq\ &(E^-, E^- - 0) \ , \end{aligned}$$

where W is the unstable set of 0 with respect to the flow and the first homotopy equivalence is induced by the inclusion.

0 *is a topologically non-degenerate critical point with Morse index* $\dim\ E^-$ *in the sense of Definition 2.2.11.*

If G operates on H by isometries, f and g are G-invariant and U is a G-neighbourhood, all constructions are in the category of G-spaces and G-maps.

Proof: A preliminary remark: As ∇g is the gradient of a non-degenerate quadratic form with index k, it is immediate that the Conley index of 0 with respect to η has the homotopy type of a k-sphere and that there are index pairs of homotopy type (DE^-, SE^-). But unfortunately it is not obvious that the index pairs can be obtained in the form $(N, N \cap f^{c-\epsilon})$. Therefore we construct such an index pair in a pedestrian way.

Set $U' := \overline{U_{2r/3}(0)}$. As H is paracompact, there is a partition of unity subordinate to the open cover $(H \backslash U') \cup U$, which enables us to glue a pseudo-gradient field for f (on $H \backslash U'$) and ∇g (on U), in order to obtain a pseudo-gradient field $V(x)$ for f, such that

$$V(x) = \nabla g(x) \quad \text{for} \quad x \in U' \ .$$

Now define

$$V_1(X) := -\min\{\mathrm{dist}(x, K), 1\} \frac{V(x)}{\|V(x)\|_x} \ ,$$

and let $\eta(t, x)$ be the global flow of V_1.

The intersection of the unstable manifold of 0 with respect to η,

$$W := \{x \in H \mid \lim_{t \to -\infty} \eta(t, x) = 0\} \ ,$$

with U' is precisely the intersection of the negative eigenspace E^- of L with U'.

The conditions on a pseudo-gradient field imply for $x \in U'$

$$\langle \nabla g(x), \nabla g(x) \rangle \leq S Df(x)(\nabla g(x)) \ , \tag{2.13}$$

for some constant $S \in \mathbb{R}$.

Now set $O := \overline{U_r(0)}$. By proposition 2.2.15 $\{0\}$ is a dynamically isolated set with respect to g and η, and there is an $\epsilon' > 0$ such that $(O, -\epsilon', \epsilon')$ is an isolating triplet for $\{0\}$ with respect to g and η, and $\tilde{O} \cap g^{-1}([-\epsilon', \epsilon']) \subset U'$. Chose such an ϵ'.

For any $0 < \delta \leq r/2$ we set $O_\delta := \overline{U_\delta(0)}$. It is clear from the definition that $(O_\delta, -\epsilon', \epsilon')$ is an isolating triplet for $\{0\}$ with respect to g and η as well and $\tilde{O}_\delta \cap g^{-1}([-\epsilon', \epsilon'] \subset U'$.

It is easy to check that

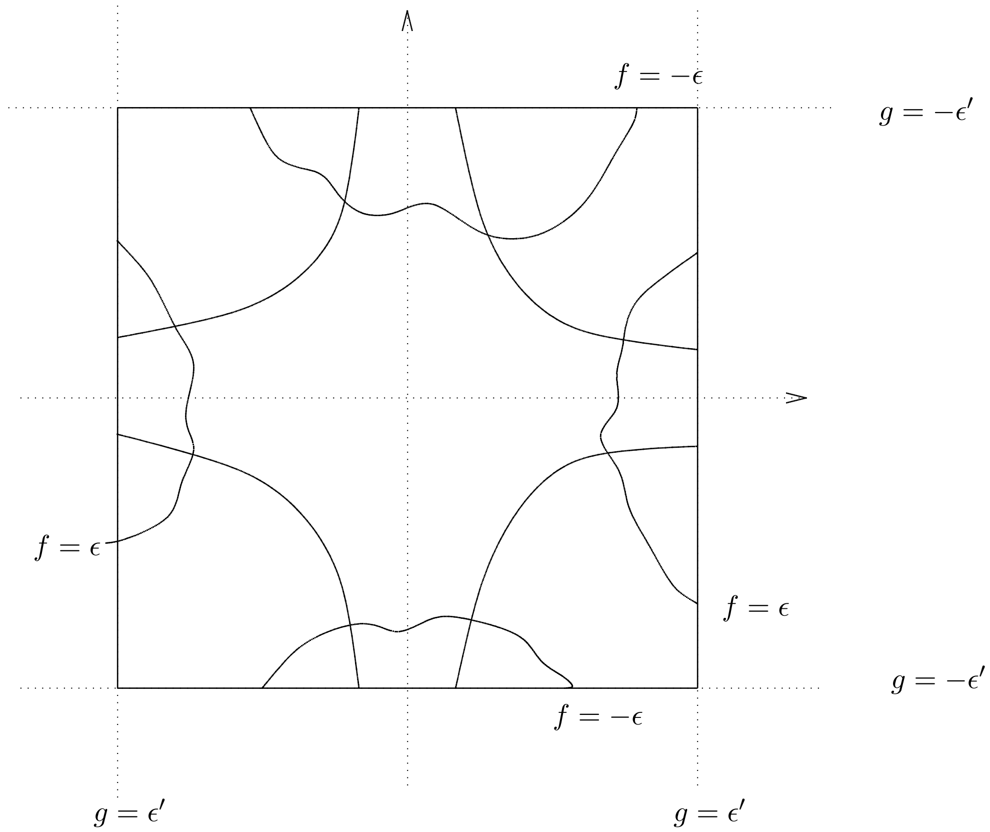
$$\begin{aligned} &(\tilde{O}_\delta \cap g^{-1}([-\epsilon', \epsilon']), \tilde{O}_\delta \cap g^{-1}(-\epsilon')) \\ \simeq\ &(W \cap g^{-1}([-\epsilon', \epsilon']), W \cap g^{-1}(-\epsilon')) \\ \simeq\ &(E^-, E^- - 0) \ . \end{aligned}$$

In particular this is the case for $O = O_{r/2}$. We claim that we can chose $\delta > 0$ and $\epsilon > 0$ such that

i) $f(x) < -\epsilon$ for all $x \in \tilde{O}_\delta \cap g^{-1}(-\epsilon')$,

ii) $f(x) > \epsilon$ for all $x \in \tilde{O}_\delta \cap g^{-1}(\epsilon')$.

This will ensure that we have the following situation.



Thus $\tilde{O}_\delta \cap f^{-1}([-\epsilon, \epsilon])$ will be contained in $\tilde{O}_\delta \cap g^{-1}([-\epsilon', \epsilon])$ and can be chosen as a distinguished neighbourhood.

Proof of the claim: Let $\tilde{\eta}$ be the global flow of $-\nabla g$ (which is well-defined, as ∇g is globally Lipschitz continuous). η and $\tilde{\eta}$ have the same flow lines on U'. Every $x \in \tilde{O}_\delta \cap g^{-1}(-\epsilon')$ is connected to some $y \in U_\delta$ by a trajectory $\tilde{\eta}([0,T], y)$ of the flow and

$$\begin{aligned} g(y) + \epsilon' &= -(g(x) - g(y)) \\ &= -\int_0^T \frac{d}{dt} g(\tilde{\eta}(t,y))dt \\ &= \int_0^T \nabla g(\tilde{\eta}(t,y)), \nabla g(\tilde{\eta}(t,y))\rangle dt \ . \end{aligned}$$

By (2.13) we conclude

$$\begin{aligned} & g(y)+\epsilon' \leq -S\int_0^T Df(\tilde{\eta}(t,y))(\nabla g(\tilde{\eta}(t,y)))dt \\ \Rightarrow \qquad & g(y)+\epsilon' \leq S(f(y)-f(x)) \ . \end{aligned}$$

and

$$-f(x) \geq \frac{1}{S}g(y) - f(y) + \frac{\epsilon'}{S} \ .$$

As f and g are continuous, we can chose δ so small that we can gurantee

$$\frac{1}{S}g(y) - f(y) > -\frac{\epsilon'}{2S} \ ,$$

for all $y \in O_\delta$ so that

$$\epsilon := \frac{\epsilon'}{2S} < -f(x) \ .$$

In the same way we prove the second part of the claim with the same δ.

For a sufficiently small δ we have $|f(y)| < \epsilon$ for all $y \in O_\delta$. Then $(O_\delta, -\epsilon, \epsilon)$ is an isolating triplet for $\{0\}$ with respect to f and η, as $O_\delta \subset f^{-1}([-\epsilon,\epsilon])$ and

$$\{0\} \subset \overline{\tilde{O}}_\delta \cap f^{-1}([-\epsilon,\epsilon]) \subset \overline{\tilde{O}} \cap g^{-1}([-\epsilon',\epsilon']) \subset U' \ .$$

For every

$$x \in A_- := (g^{-1}([-\epsilon',\epsilon'])\backslash f^{-1}([-\epsilon,\infty[)) \cap \tilde{O}_\delta$$

the unique trajectory from O_δ to x has to pass through $f^{-1}(-\epsilon)$, hence there is a $y(x) \in f^{-1}(-\epsilon)$ and a $t(x) < 0$ such that

$$\eta(t(x),x) = y \ .$$

As the flow η is transversal to $f^{-1}(-\epsilon)$, $y(x)$ and $t(x)$ are continuous functions on A_-.

For every

$$x \in A_+ := (g^{-1}([-\epsilon',\epsilon'])\backslash f^{-1}(]-\infty,\epsilon])) \cap \tilde{O}_\delta$$

the unique trajectory from x to O_δ has to pass through $f^{-1}(\epsilon)$, hence there is a $y(x) \in f^{-1}(\epsilon)$ and a $t(x) > 0$ such that

$$\eta(t(x),x) = y(x) \ .$$

As the flow is transversal to $f^{-1}(\epsilon)$, $y(x)$ and $t(x)$ are continuous functions on A_+.

Now we define a deformation retraction from $g^{-1}([-\epsilon', \epsilon'] \cap \tilde{O}_\delta$ to $f^{-1}([-\epsilon, \epsilon]) \cap \tilde{O}_\delta$ by

$$\psi(s, x) = \begin{cases} \eta(s\ t(x), x)\ , & \text{falls} \quad x \in A_+ \cup A_- \\ x & \text{else}\ . \end{cases}$$

As W is invariant under the flow and

$$W \cap U' = E^- \cap U' \subset \tilde{O}_\delta \cap U'\ ,$$

the retraction induces a retraction from

$$g^{-1}([-\epsilon', \epsilon']) \cap W = g^{-1}([-\epsilon', \epsilon']) \cap (E^- \cap U')$$

to

$$f^{-1}([-\epsilon, \epsilon]) \cap W = f^{-1}([-\epsilon, \epsilon]) \cap (E^- \cap U')\ .$$

$f^{-1}([-\epsilon, \epsilon]) \cap \tilde{O}_\delta$ is a distinguished neighbourhood of 0. □

This proposition implies the following local version for Hilbert spaces actually needed in our applications.

Corollary 2.2.21 *Let X be a $C^{1,1}$-Hilbert-manifold modelled on the Hilbert space H and $f : X \to \mathbb{R}$ a $C^{1,1}$-functional. Suppose x is an isolated critical point of f, and there is a chart $X \supset V \overset{\psi}{\to} U \subset H$ with $x \in V$ and $\psi(x) = 0$.*

Suppose that L is a bounded self-adjoint invertible operator, which defines the smooth quadratic form

$$g : H \to \mathbb{R},\ v \mapsto g(v) = \frac{1}{2}\langle v, Lv\rangle\ ,$$

and that the dimension of the negative eigenspace E^- of L is finite ($k \in \mathbb{N}$).

Now suppose that on the set U the gradient ∇g is a pseudo-gradient field for $f \circ \psi^{-1}$.

Then x is a topologically non-degenerate critical point with Morse index $\dim E^-$ *in the sense of Definition 2.2.11.*

Proof: Let $W \subset \overline{W} \subset U$ be a smaller open neigbourhood of 0. By means of a $C^{1,1}$-partition of unity subordinated to the open cover $(H \backslash \overline{W}, U)$ of H, we can define a $C^{1,1}$-functional $\tilde{f} : H \to \mathbb{R}$, which is identical with $f \circ \psi^{-1}$ on W.

Now we can apply Proposition 2.2.20 to $\tilde{f}$ (and W instead of U) and deduce that 0 is a topologically non-degenerate critical point of $\tilde{f}$ with Morse index dim E^-. This a local property, hence x is a topologically non-degenerate critical point of f with the same Morse-index. □

Definition 2.2.22 *Let G be $\mathbb{Z}_p$ or S^1. Let X be a G-Banach manifold and $f \in C^1(X)$.*

For a critical fixed point x with $f(x) = c$ we define $k_G(x)$ to be the supremum of all $k \in \mathbb{N}$ such that for every neighbourhood U of x there is a representation V of G with $V^G = \{0\}$ and $\dim\ V = k$ and an equivariant continuous map $g : D^V \to U^c$ with $g(0) = x$ and $g(S^V) \subset U^{c-\epsilon}$ for some $\epsilon > 0$.

Remark 2.2.23 For $G = \mathbb{Z}_2$, an isolated fixed point x and a neighbourhood U with $U^G = \{x\}$ this invariant is in fact a special case of *the G-capacity* of U^c as treated e.g. in [CM00].

In general the definition of G-capacity uses joins $G * \cdots * G$ instead of spheres.$\diamond$

Proposition 2.2.24 *Let G be $\mathbb{Z}_p$ or S^1. Let X be a C^2-Hilbert G-manifold and $f \in C^2(X)$.*

Let x be a fixed point and a critical point of f with Morse index $\mu(x)$. Then

$$k_G(x) \geq \mu(x)\ .$$

Proof: In the nondegenerate case the argument is as follows: By the equivariant Morse lemma there is a local chart to a neighbourhood of 0 in a G-Hilbert space H, in which f is represented by a quadratic form. Now we can embed an arbitrarily small disk D^k in E^- for $k \leq \dim E^- = \mu(x)$.

In the degenerate case we consider any local G-equivariant chart $\psi : U \to V$ with $x \in U \subset X$ and $V \subset H$, and set $\tilde{f} := f \circ \psi^{-1} : V \to \mathbb{R}$. The second derivative of $\tilde{f}$ at 0 is represented by the symmetric operator $\tilde{f}''(0)$ with Morse index $\mu(x)$. Let $E :=$ Ker $(\tilde{f}''(x))$ and $F := E^\perp$. The restriction $g := \tilde{f}|_{V\cap F}$ has a non-degenerate critical point at 0, and

$$g''(0) = {}_F\backslash \tilde{f}''(0)/_F$$

has a negative eigenspace of dimension $\mu(x)$. We now apply the equivariant Morse lemma to g and obtain the required embedding.

Please note that we did *not* need a generalised Morse-lemma for the degenerate case, which would require more assumptions, typically that $f''(x)$ is a Fredholm operator, and would provide a much stronger statement. $\square$

Theorem 2.2.25 *Suppose X a complete $C^{2,1}$-Finsler G-manifold and $f \in C^1(X)$ satisfies the Palais-Smale condition in X^a for some $a \in \mathbb{R}$. Let us assume that h^* is Cech cohomology theory.*

Let $0 \neq \alpha \in h_G^s(X^a)$ *and* $\omega \in h^*(BG)$ *be a cohomology class of strict category weight* w *and* $c := c(\alpha) = c(\alpha \cup p^*\omega) \in \mathbb{R}$ *(with* $p : EG \times_G X \to BG$*). Then we have*

$$p^*\omega|_{K_c} \neq 0$$

and

$$\mathrm{cat}((K_c)_G) \geq w + 1 \ .$$

i) If K_c *is a free* G*-space we have* $\mathrm{cat}(K_c/G) \geq w + 1$*.*

ii) Suppose the restriction of α *to the fixed point set of* $(K_c)^G$ *is zero and the fixed point set* $(K_c)^G$ *consists of non-degenerate critical submanifolds of* X *with* h_G^**-orientable negative bundles,* $h*(K_c^G)$ *is finitely generated as* k*-vector space,* $h^*(BG)$ *is finitely generated as a* k*-algebra, and there is an element* $\omega \in h^d(BG)$ *such that multiplication with* ω *defines isomorphisms* $h^s(BG) \to h^{s+d}(BG)$ *for all* $s \geq 0$*. Then*

$$p^*\omega|_{K_c - K_c^G} \neq 0 \ .$$

Proof: Again we proceed indirectly and suppose that $p^*\omega|_{K_c} = 0$ under the above assumptions. As the G-action is $C^{2,1}$, it is locally smooth by remark 2.1.7, by proposition 2.1.9 we have the continuity property for h_G^*, hence there is a closed G-invariant neighbourhood N of K_c such that $p^*\omega|_N = 0$. We now apply Theorem 2.2.1 with $\bar{\epsilon} = 1$. There is an $\epsilon \in]0, \bar{\epsilon}[$ such that $\Phi(f^{c+\epsilon} \backslash N, 1) \subset f^{c-\epsilon}$ and $\Phi(f^{c+\epsilon,1}, 1) \subset f^{c-\epsilon} \cup N)$. We replace N by $N \cap f^{c-\epsilon}$.

Now $\alpha|_{f^{c+\epsilon}} \neq 0$ and $\alpha|_{f^{c-\epsilon}} = 0$. Via the long exact sequence of the couple $(f^{c+\epsilon}, f^{c-\epsilon})$

$$\cdots \to h_G^*(f^{c+\epsilon}, f^{c-\epsilon}) \xrightarrow{\rho} h_G^*(f^{c+\epsilon}) \to h_G^*(f^{c-\epsilon}) \to \cdots$$

it follows, that $\alpha|_{f^{c+\epsilon}}$ is the restriction of a non-zero relative class $\tilde{\alpha} \in h_G^*(f^{c+\epsilon}, f^{c-\epsilon})$ to $f^{c+\epsilon}$. ρ is a $h_G^*(pt)$-algebra homomorphism. Therefore

$$\rho(\tilde{\alpha} \cup p^*\omega) = \alpha \cup p^*\omega \ .$$

As $\Phi(f^{c+\epsilon}, 1) \subset f^{c-\epsilon} \cup N$ and $\Phi(f^{c+\epsilon} \backslash N, 1) \subset f^{c-\epsilon}$ the isomorphism

$$id = \Phi(\cdot, 1)|^*_{(f^{c+\epsilon}, f^{c-\epsilon})} : h_G^*(f^{c+\epsilon}, f^{c-\epsilon}) \to h_G^*(f^{c+\epsilon}, f^{c-\epsilon})$$

factorises through $h_G^*(f^{c-\epsilon} \cup N, f^{c-\epsilon})$, thus $\tilde{\alpha}|_{(f^{c-\epsilon} \cup N, f^{c-\epsilon})} \neq 0$. By the strong excision property, which is valid for equivariant Čech cohomology, the restriction $h_G^*(f^{c-\epsilon} \cup N, f^{c-\epsilon}) \to h_G^*(N, N \cap f^{c-\epsilon})$ is an isomorphism and

$$0 \neq \tilde{\alpha}|_{(N, N \cap f^{c-\epsilon})} \ .$$

In the same way we obtain that

$$0 \neq (\tilde{\alpha} \cup \ p^*\omega)|_{(N, N \cap f^{c-\epsilon})} \ .$$

A basic property of the cup product allows to write

$$0 \neq \tilde{\alpha}|_{(N,N\cap f^{c-\epsilon})} \ \cup \ \ p^*\omega|_{(N,N\cap f^{c-\epsilon})} = \tilde{\alpha}|_{(N,N\cap f^{c-\epsilon})} \ \cup \ \ p^*\omega|_N \ .$$

By assumption $p^*\omega|_N = 0$, a contradiction.

We have proved $p^*\omega|_{K_c} \neq 0$ in $h^*_G(K_c) \cong h^*((K_c)_G)$. The definition of category weight implies that cat $(K_c)_G \geq w + 1$.

The first of the two other statements is obvious by the homotopy equivalence of $(K_c)_G \simeq K_c/G$ for a free G-operation.

The second statement affords more work:

We proceed indirectly: Assume that $p^*\omega|_{K_c - K_c^G} = 0$, then there is a closed G-neighbourhood U of $K_c - K_c^G$ such that $p^*\omega|_U = 0$.

Now suppose $N = U \cup V$, where V is a distinguished G-neighbourhood of K_c^G (in the sense of Definition 2.2.11; if V is distinguished, so is $V \cap f^{c-\epsilon}$). By the assumptions on K_c^G we know that these neighbourhoods can be chosen disjoint. The Mayer-Vietoris sequence yields

$$h^*_G(U \cup V, (U \cup V)^{c-\epsilon}) \cong h^*_G(U, U^{c-\epsilon}) \oplus h^*_G(V, V^{c-\epsilon}) \ , \tag{2.14}$$

the isomorphism is induced by the restrictions.

By the (PS) condition K_c^G is compact and has only finitely many components K_1,... K_k. We can assume that $V = V_1 \cup \cdots \cup V_k$, where V_i is a neighbourhood of the component K_i such that $(V_i, V_i^{c-\epsilon})$ is G-homotopy equivalent to (DE_i, SE_i) where E_i is a G-vector bundle over K_i, whose dimension is equal to the Morse index of K_i.

As (DE_i, SE_i) is G-homotopy equivalent to $(DE_i, DE_i - 0_{K_i})$, we can apply Proposition 2.1.46 to conclude that the restriction homomorphism $h^*_G(DE_i, SE_i) \to h^*_G(DE_i)$ injective, Thus the restriction homorphism

$$h^*_G(V_i, V_i^{c-\epsilon}) \to h^*_G(V_i)$$

is injective. However, $\tilde{\alpha}$ restricts to 0 in $h^*_G(V)$ by assumption, and hence in $h^*_G(V_i)$. Therefore

$$0 = \tilde{\alpha}|_{(V, V^{c-\epsilon})} \ .$$

Furthermore

$$(\tilde{\alpha} \cup \omega)|_{(U,U^{c-\epsilon})} = \tilde{\alpha}|_{(U,U^{c-\epsilon})} \cup \omega|_U = 0 \ ,$$

as $\omega|_U = 0$ by hypothesis.

By the isomorphism (2.14) $(\tilde{\alpha} \cup \omega)|_{(N,N^{c-\epsilon})} = 0$, a contradiction to the fact proved above.

□

Remark 2.2.26 Unfortunately, in the case $G = \mathbb{Z}_p$, this theorem is not sufficient to prove the existence of at least

$$\mathrm{swgt}(\beta(\omega)^k) + 1 = 2k + 1$$

critical points in $f^{-1}([c(\alpha), c(\alpha \cup p^*(\beta(\omega)^k))])$, where $\beta(\omega)$ is the free generator of degree 2 if $h^*_G(pt)$, which has $\mathrm{swgt}(\beta(\omega)) \geq 2$ according to Proposition 2.2.10. We would like to remark, however, that such a multiplicity result is valid in the absence of fixed points. It can be obtained by means of *relative category*, as in the following remark. ◇

Remark 2.2.27 If $c(\alpha) < c(\alpha \cup p^*\omega) =: c$, we obtain that $0 \neq p^*\omega \cup \tilde{\alpha} \in h^*_G(f^c, f^b)$ for some $b < c(\alpha)$, $\tilde{\alpha} \in h^*_G(f^c, f^b)$.

Now let us suppose that the G-action is free on $f^c - f^b$. Then, for the class of G-spaces $\mathcal{A} = \{G\}$ we have

$$(\mathcal{A}) - \mathrm{cat}(f^c, f^b) \geq w + 1 \ ,$$

with the relative category in the sense of Clapp and Puppe [CP91].

Let us give a proof, which is a modification of the proof of Proposition 4.3 of [CP91]: If $(X_0, X_1, \ldots, X_w)$ is an open G-covering of f^c by $w+1$ open subspaces such that $f^b \subset X_0$ and X_0 is G-deformable into f^b mod f^b, and for $i = 1..w$ each of the inclusions $X_i \to f^c$ factors up to homotopy through G (the only member of $\mathcal{A}$). We can assume that $X_i \cap f^b = \emptyset$ for all $i = 1..w$. So $Y := (X_1 \cup \cdots \cup X_w)_G \simeq X_1/G \cup \ldots \cup X_w/G$, a union of w contractible spaces, that is cat $Y = w$. Therefore $p^*\omega|_Y = 0$ and $p^*\omega$ is the image of some class $\tilde{\omega} \in h^*_G(f^a, Y)$.

By definition of X_0 the class $\tilde{\alpha}$ is mapped to zero under the restriction $(f^c, f^b) \to (X_0, f^b)$, hence it is the image of a class $\gamma \in h^*_G(f^c, X_0)$. But $\tilde{\omega} \cup \gamma \in h^*_G(f^c, Y \cup X_0) = h^*_G(f^c, f^c) \cong 0$, hence $0 = p^*\omega \cup \tilde{\alpha} \in h^*_G(f^c, f^b)$, a contradiction.

For $G = \mathbb{Z}_p$ we can therefore define a version of the cup-length that takes into consideration the category weight. For h^* Čech cohomology with $\mathbb{Z}_p$-coefficients we have $h^*_G(pt) \cong \mathbb{Z}_p[x, y]/(x^2 = 0) \cong \mathbb{Z}_p[y] \otimes \Lambda[x]$. where x has degree one and

$y = \beta(x)$. The non-trivial products in $h^*_G(pt)$ have the form y^k or xy^k with strict category weights $2k$ and $1+2k$. If we define

$$l(X, X') = 1 + \max\{\text{swgt }(\omega) | \omega \in h^*_G(pt),\ \ \exists \gamma \in h^*_G(X, X'),\ p^*\omega \cup \gamma \neq 0\}\ .$$

we obtain like above $l(X, X') \leq \mathcal{A} - \text{cat}(X, X')$. We were not able to prove, however, that this modified cup-length satisfies the usual subadditivity properties of a length, unless under further assumptions, which are slightly weaker than the assumptions mentioned in Remark 4.14 of Bartsch [Bar93]. $\diamond$

2.2.2. Critical point theory over a space

In this subsection X is a $C^{1,1}$-Finsler manifold, $\pi : X \to M$ a surjective map to a finite dimensional manifold M. X and M are G-spaces, $\pi : X \to M$ is G-equivariant. Let us further suppose that $\pi : X \to M$ has a G-equivariant section $\sigma : M \to X$ such that $X^G \subset \sigma(M)$. [2]

Theorem 2.2.28 *Let either be* $(G = \mathbb{Z}_2, h^* = \check{H}^*(\cdot, \mathbb{Z}_2), d = 0)$, *or* $(G = S^1, h^* = \check{H}^*(\cdot, \mathbb{Q}), d = 1)$.

For $f \in C^1(X)$, *assume that* $p \in X^G$ *is a critical point with* $f(p) = b$, *and* F *be the fixed point set strictly below level* b. *Suppose all components of* F *are non-degenerate critical submanifolds.*

Assume furthermore that f *satisfies* (PS) *below level* c, *and that there is some* $a < b$ *and some class* $\mu \in h^{n^-}_G(X, f^a)$ *such that* $\mu|_p \neq 0$.

Then, there are at least

$$(k_G(p) - (\sigma(\pi(p), \pi(F)) + n^-))\frac{1}{d+1}$$

non-fixed G*-orbits of critical points between level* a *and level* b.

The conditions are met in two important cases:

1. f *is bounded below (then chose* $a < \min\ f$, $n^- = 0$*).*

[2] In some applications X will be a vector bundle over M and σ will be the zero section. It can be useful, however, to consider the free loop space $H^{1,2}(S^1, M)$ as a fibration over M, where $\sigma(x)$ is the constant loop $t \mapsto x$.

2. $(X, f^a) \simeq_M (DE^{-\infty}, SE^{-\infty})$ *for a* h_G^**-orientable* n^-*-dimensional vector bundle* $E^{-\infty}$*.*

Remark 2.2.29 In particular, by the estimate 2.2.24, if $\pi : X \to M$ is a Hilbert bundle, p a critical fixed point of Morse index $i(p)$ and $f|_{f^a}$ is equivalent to a quadratic form of index n^- on the fibres with h_G^*-orientable fibres, there are at least

$$(i(p) - (\sigma(\pi(p), \pi(F)) + n^-))\frac{1}{d+1}$$

non-fixed G-orbits of critical points below level b. $\diamond$

Remark 2.2.30 The critical levels c are characterised by the vanishing of certain cohomology classes of degrees

$$m \in \{\sigma(\pi(p), \pi(F)) + n^-, \dots, k_G(p) - 1\} ,$$

which induce non-vanishing classes on $h_G^*(N, N^{c-\epsilon})$ for a neighbourhood N of the non-fixed critical set on level c. If, for $G = \mathbb{Z}_2$, this non-fixed critical set consists of a single G-orbit Gx of isolated points, we can use a neighbourhood as in Proposition 2.2.15 and use a Gromoll-Meyer argument to obtain that

$$\mu(x) \le m \le \mu(x) + \nu(x) .$$

Proof: (of Theorem 2.2.28)

Set $k := \sigma(\pi(p), \pi(F))$, then there is an $\alpha \in h_G^k(M)$ with $\alpha|_{\pi(p)} \neq 0$ and $\alpha|_{\pi(F)} = 0$. By assumption there is a class μ in $h_G^{n^-}(X, f^a)$ such that $(\pi^*\alpha \cup \mu)|_p \neq 0$.

For $m := k_G(p)$ there is a representation disk D^m and an equivariant continuous map $g : D^m \to f^b$ with $g(0) = p$ and $g(S^{m-1}) \subset f^{b-\epsilon}$ for some $\epsilon > 0$.

The basic idea for the mulitplicity result is that for some nonnegative exponents i and a polynomial generator ω of $h_G^*(pt)$, the product $\omega^i \cup (\pi^*\alpha \cup \mu)$ restricts to a non-zero class in $h_G^*(f^{b-\epsilon})$.

We distinguish between the two cases $G = \mathbb{Z}_2$ and $G = S^1$, in order to find the optimal exponents:

$\mathbf{G} = \mathbb{Z}_\mathbf{2}$ Recall that $h^*(BG) \cong \mathbb{Z}_2[\omega]$ with deg $\omega = 1$. The restriction of the class

$$\eta := \omega^{m-1-(k+n^-)} \cup (\pi^*\alpha \cup \mu) \in h_G^{m-1}(X)$$

to p is not zero. Now consider

$$\begin{tikzcd} h_G^{m-1}(D^m) \arrow[d, "r"'] & h_G^{m-1}(X) \arrow[l, "g^*"'] \arrow[d] \\ h_G^{m-1}(S^{m-1}) & h_G^{m-1}(f^{b-\epsilon}) \arrow[l, "g|_{S^{m-1}}^*"'] \end{tikzcd} .$$

As $(g|_0)^*(\eta|_p) \neq 0$ and the inclusion $0 \to D^m$ induces an isomorphism $h_G^{m-1}(D^m) \cong h_G^{m-1}(\{0\})$, the class $g^*\eta$ is not zero either. The restriction map r is injective, therefore we conclude by the commutativity of the diagram that

$$\eta|_{f^{b-\epsilon}} \neq 0 .$$

$\mathbf{G = S^1}$ Recall that $h^*(BG) \cong \mathbb{R}[\omega]$ with deg $\omega = 2$. The restriction of the class

$$\eta := \omega^{(m-2-(k+n^-))/2} \cup (\pi^*\alpha \cup \mu) \in h_G^{m-2}(X)$$

to p is not zero. (m and n^- are even as the dimensions of S^1-representations without trivial summands, and k is even, as it is the degree of a class that does not vanish when restricted to $\pi(p)$.) Now consider

$$\begin{tikzcd} h_G^{m-2}(D^m) \arrow[d, "r"'] & h_G^{m-2}(X) \arrow[l, "g^*"'] \arrow[d] \\ h_G^{m-2}(S^{m-1}) & h_G^{m-2}(f^{b-\epsilon}) \arrow[l, "g|_{S^{m-1}}^*"'] \end{tikzcd} .$$

As $(g|_0)^*(\eta|_p) \neq 0$ and the inclusion $0 \to D^m$ induces an isomorphism $h_G^{m-2}(D^m) \cong h_G^{m-2}(\{0\})$, the class $g^*\eta$ is not zero either. The restriction map r is injective, therefore we conclude by the commutativity of the diagram that

$$\eta|_{f^{b-\epsilon}} \neq 0 .$$

Now, in both cases there is an element $\gamma \in h_G^*(p) \cong h_G^*(pt)$ of degree $m - (d + 1) - (k + n^-)$ such that

$$0 \neq \eta = \gamma \cup (\pi^*\alpha \cup \mu)|_p \in h_G^{m-(d+1)}(p)$$

and $\gamma = \omega^{(m-(d+1)-(k+n^-))/(d+1)}$, where $\omega \in h^{d+1}(BG)$ is a generator of $h^*(BG)$.

F consists only of equivariantly normally orientable submanifolds by remark 2.1.37. This enables us to apply iteratively Theorem 2.2.25, which gives us

$$\frac{1}{d+1}(m-(d+1)-(k+n^-))+1=\frac{1}{d+1}(m-(k+n^-))$$

non-fixed G-orbits of critical points. Either the critical levels $c(\omega^s \cup (\pi^*\alpha \cup \mu))$ are all different for $s=0,\ldots,\frac{1}{d+1}(m-(d+1)-(k+n^-))$ or the critical set on some level has category greater or equal to $\mathrm{swgt}(\omega)+1 \geq 2$, and thus is infinite.

Let us now check, that the conditions are satisfied in the two special cases:

1. f is bounded below, then there is an a such that $f^a=\emptyset$. μ than can be chosen as 1_X. If $\gamma \in h^*_G(M)$, $\pi^*\gamma \cup 1_X = \pi^*\gamma$. As $\pi \circ \sigma = id_M$, we have

$$\gamma|_p = (\sigma^*\pi^*\gamma)|_p = \sigma|^*_p(\pi^*\gamma)|_{\sigma(p)} \ .$$

   Therefore the map $\gamma|_p \mapsto (\pi^*\gamma)|_{\sigma(p)}$ is injective.

2. If $(X, f^a) \simeq_M (DE^{-\infty}, SE^{-\infty})$ for a h^*_G-orientable finite dimensional vector bundle $E^{-\infty}$ over M, there is a Thom-isomorphism

$$h^*_G(M) \to h^{*+n^-}_G(X, f^a) \quad \text{for some } n^- \ ,$$

   given by multiplication with the Thom-class of E^-. By remark 2.1.39 this corresponds to multiplication with a power of the free generator of $h^*_G(pt)$ on the fixed point set.

□

For $G=\mathbb{Z}_p$, p odd, the theorem and its proof are quite similar to the case $G=S^1$ above, however $h^*_G(BG)$ is not a polynomial ring any more. For the multiplicity result, we will only use powers of the free generator ω of

$$h^*_G(pt) \cong \mathbb{Z}_p[x,\omega]/(x^2=0) \cong \mathbb{Z}_p[\omega] \otimes \Lambda[x] \ .$$

Theorem 2.2.31 *Let $G=\mathbb{Z}_p$, $h^*=\check{H}^*(\cdot,\mathbb{Z}_p)$ and $d=1$ ($d+1$ is the degree of the free generator ω of $h^*_G(pt)$.)*

For $f \in C^1(X)$, assume that $p \in X^G$ is a critical point with $f(p)=b$, and F be the fixed point set strictly below level b. Suppose all components of F are non-degenerate critical submanifolds.

Assume furthermore that f satisfies (PS) below level c, and that there is some $a<b$ and some class $\mu \in h^{n^-}_G(X, f^a)$ such that

$$h^*_G(p) \to h^{*+n^-}_G(p),\ \alpha \mapsto \alpha \cup (\mu|_p)$$

is an injective map (which is the case iff the restriction is equal to a power of the free generator.)

Then, there are at least

$$\left[(k_G(p) + d - (\sigma(\pi(p), \pi(F)) + n^-))\frac{1}{d+1}\right]$$

non-fixed G-orbits of critical points between level a and level b. (For a real number s the bracket $[s]$ *means the largest integer not greater than s.)*

The conditions are met in two important cases:

1. f *is bounded below (then chose* $a < \min\ f$, $n^- = 0$*).*

2. $(X, f^a) \simeq_M (DE^{-\infty}, SE^{-\infty})$ *for a* h_G^**-orientable* n^-*-dimensional vector bundle* $E^{-\infty}$*.*

Remark 2.2.32 1. This is obviously *not* an optimal theorem. We have much more information, that remains unused. Above all we know that $\omega = \beta(x)$, where β is the $\mathbb{Z}_p$-Bockstein map, and swgt $(\omega) = 2$.

2. n^- and $k_G(p)$ are even (as dimensions of $\mathbb{Z}_p$-representations without trivial summands). We have to use the 'largest integer not greater than', as we could not rule out the possibility that $\sigma(\pi(p), \pi(F))$ is odd. If we modify the definition of $\sigma(x, F)$ by requiring that the separating classes should restrict to powers of the polynomial generator on x, we always get an even number, which yields the same estimate for the number of orbits of non-fixed critical points as in Theorem 2.2.28. ◇

Proof: The only part of the proof of Theorem 2.2.28 which has to modified, concerns the exponent i of ω such that

$$\eta := \omega^i \cup (\pi^*\alpha \cup \mu)$$

restricts to a non-zero class in $h_G^*(f^{b-\epsilon})$. The largest i, such that is the case, must satisfy

$$\begin{aligned} & (d+1)i + k + n^- \leq m - 1 \\ \Leftrightarrow \quad & i \leq (m - 1 - (k + n^-)) \\ \Leftrightarrow \quad & i \leq [(m - 1 - (k + n^-))/(d+1)] \,. \end{aligned}$$

The maximal i we can chose is $[(m-1-(k+n^-))/(d+1)]$. As above the iterative application of Theorem 2.2.25 yields at least

$$[(m - 1 - (k + n^-))/(d+1)] + 1 = [(m + d - (k + n^-))/(d+1)] \,.$$

G-orbits of non-fixed critical points. □

Remark 2.2.33 Such a Theorem can be established for a semi-free S^3-action and $d = 3$ by an almost identical proof. ◇

Example 2.2.34 We will see how these results allow to improve the results of Bartsch and Wang [BW97b].

Consider the Hamiltonian system

$$-J\dot{z} = H_z(t, z) \ , \tag{2.15}$$

where $J = \begin{pmatrix} 0 & -I_N \\ I_N & 0 \end{pmatrix}$ is the standard symplectic matrix in $\mathbb{R}^{2N \times 2N}$.

(H_1) $H \in C^2(\mathbb{R} \times \mathbb{R}^{2N}, \mathbb{R})$ is 2π-periodic in all variables.

(H_2) H is even in z, that is, $H(t, -z) = H(t, z)$ for all $(t, z) \in \mathbb{R} \times \mathbb{R}^{2N}$.

Here $i_M(\cdot)$ denotes the Maslov index.

We recall Theorem 3.3. of [BW97b] (with slightly changed notation):

Suppose (H_1) and (H_2) hold. Let w_0 be a (possibly degenerate) stationary solution of (2.15) such that the following holds ($\sigma \in \{\pm 1\}$ denotes the sign of $i_M(w_0)$).

(H_3) All trivial solutions with

$$\sigma \int_0^{2\pi} H(t, w)dt > \sigma \int_0^{2\pi} H(t, w_0)$$

are nondegenerate. ◇

Then (2.15) has at least $i_M(w_0) - N$ pairs of nontrivial periodic solutions if $i_M(w_0) > N$ and at least $|i_M(w_0)| - \nu(w_0) - N$ such solutions if $-i_M(w_0) > \nu(w_0) + N$.

The proof in [BW97b] proceeds by a finite dimensional reduction of the Hamiltonian action functional

$$I(z) := \frac{1}{2}(Az, z) - \int_0^{2\pi} H(t, z)dt$$

with $A = -J\frac{d}{dt}$ densely defined on $L^2([0, 2\pi], \mathbb{R}^{2N})$. The reduction yields a functional f on a finite dimensional vector bundle $(E^+ \oplus E^-) \times T^{2N} \to T^{2N}$ with

$\dim E^- = n^-$. The group $G = \mathbb{Z}_2$ operates on this bundle with 2^{2N} fixed points, which are contained in the zero section. Bartsch and Wang prove, that there is an a with

$$f^a \subset ((E^+ \oplus E^-)\backslash E^+) \times T^{2N} .$$

Now as $(E^+ \oplus E^-)\backslash E^+) \times T^{2N}$ is G-homotopy equivalent to $SE^- \times T^{2N}$, there is a Thom class μ_1 in

$$h_G^{n^-}((E^+ \oplus E^-) \times T^{2N}, (E^+ \oplus E^-)\backslash E^+) \times T^{2N}) ,$$

which restricts to a class

$$\mu \in h_G^{n^-}((E^+ \oplus E^-) \times T^{2N}, f^a) .$$

By Remark 2.1.39 $\mu|_{w_0}$ is a power of the generator of $h_G^*(BG)$, so that the cup-product with μ induces an injective map on $h_G^*(w_0)$

Let us consider the case $\sigma = 1$. The trivial solutions w with $\int_0^{2\pi} H(t,w)dt > \int_0^{2\pi} H(t,w_0)$ are precisely those with action below $I(w_0)$. Let $F \subset T^{2N}$ the set of fixed points that correspond to these solutions. As we suppose that these solutions are non-degenerate, we can apply our Theorem 2.2.28 and deduce that there are *at least*

$$k_G(w_0) - (\sigma(w_0,F) + n^-)$$

pairs of 2π*-periodical solutions of (2.15).* $k_G(w_0)$ can be estimated below by the Morse index $i(w_0)$ of w_0, which, by Lemma 3.10. in [BW97b], can be expressed by means of the Maslov index as

$$i(w_0) = (2m+1)N + i_M(w_0) ,$$

where $2mN = \dim E^{\pm}$. Thus

$$\begin{aligned} & k_G(w_0) - (\sigma(w_0,F) + n^-) \\ & \geq i_M(w_0) + (2m+1)N - (\sigma(w_0,F) + 2mN) \\ & = i_M(w_0) + N - \sigma(w_0,F) . \end{aligned}$$

As $\sigma(w_0,F) \leq 2N$, in the worst case we get the number of solution orbits $i_M(w_0) - N$ predicted by Bartsch and Wang. Whenever

$$\sigma(w_0,F) < 2N ,$$

we get *more* solutions. For example (s. Example 2.1.34), if there is one more fixed point above the level $f(w_0)$, we will have

$$\sigma(w_0,F) \leq 2N - 1 .$$

And the estimate

$$\sigma(w_0, F) \leq \left[\frac{|F|+1}{2} \right] ,$$

yields an improvement whenever $|F| \leq 4N - 2$. We do not have an explicit formula for the $\sigma(w_0, F)$, but better estimates by a recursion formula (s. Example 2.1.34). If we happen to know that w_0 is contained in a k-dimensional G-invariant submanifold that does not contain any fixed point below the level $f(w_0)$, then by Corollary 2.1.45

$$\sigma(p, F) \leq 2N - k .$$

Remark 2.2.35 It is interesting to compare results of this type with the information that can be obtained from (equivariant or ordinary) Morse theory. Let us consider only the case $(X, f^a) \simeq_M (DE^-, SE^-)$ of the above Theorem 2.2.28, $G = \mathbb{Z}_2$, M and n-dimensional G-manifold.

If we define the equivariant Poincaré formal power series $P_G(t)$ of (X, f^a) and the equivariant Morse formal power series $M_G(t)$ with respect to the finest Morse decomposition of (X, f^a), we have

$$M_G(t) = P_G(t) + (1+t)Q(t) , \tag{2.16}$$

where $Q(t)$ is a formal power series with non-negative coefficients.

Now suppose that the critical set consists of non-degenerate critical points. In this case we can recover multiplicity results like those of Bartsch and Wang by counting.

The assumption on (X, f^a) implies that

$$P_G(t) = t^{n^-} P_G(M) , \tag{2.17}$$

with $n^- = \dim\ E^-$.

We write $M_G(t) = \sum_{i=0}^{\infty} M_G^i t^i$ and $P_G(t) = \sum_{i=0}^{\infty} P_G^i t^i$. Thus, by (2.17) $P_G^i = \dim h_G^{i-n^-}(M)$, and $M_G^i = \sum_{x \in K} \dim h_G^i(N_1(x), N_2(x))$, where $(N_1(x), N_2(x))$ is an index-pair for x. By the non-degeneracy assumption, this pair is homotopy equivalent to $(D^{\mu(x)}, S^{\mu(x)-1})$. For a fixed point x it is a G-homotopy equivalence, and $(D^{\mu(x)}, S^{\mu(x)-1}) = (D^V, S^V)$ for a G-representation V with $V^G = 0$. We can now express the contribution of the critical fixed points (FP) and the non-fixed critical points (NFP) $M_G^i = M_{FP}^i + M_{NFP}^i$ by their Morse-indices, using the Thom isomorphism: $M_{FP}^i = |\{x \in X^G \mid \mu(x) \geq i\}|$ and $M_{NFP}^i = |\{x \in K \backslash X^G \mid \mu(x) = i\}|$.

Proposition 2.1.24 implies that the number $|F|$ of fixed points equals $\dim h^i_G(M)$ for $i > n$ and is less or equal to $\dim h^i_G(M)$ for $i = n$. Therefore $P^i_G \geq |F|$ for $i \geq n + n^-$.

Any critical fixed point x of index $\mu := \mu(x)$ greater than $n + n^-$ would lead to

$$M^i_{FP} < |F| \leq P^i_G$$

for $i = n + n^-, \dots, \mu - 1$.

We need at least $\mu - (n+n^-)$ G-orbits of non-degenerate non-fixed critical points to compensate this, that is to achieve $M^i_G \geq P^i_G$ as implied by (2.16).

A different counting argument is possible in the non-equivariant case. $\diamond$

3. Symmetric Lagrangian systems

3.1. The geometrical and analytical framework

3.1.1. Riemannian manifolds

In the following M is an n-dimensional Riemannian C^∞-manifold and TM its tangent bundle, $\tau_M : TM \to M$ the bundle projection. Suppose there is a smooth isometric G-operation on M with isolated fixed points. This induces a G-operation on TM all of whose fixed points lie in the zero section 0_M. A local chart (parametrisation) of M induces charts (parametrisations) of TM and TTM via the tangent functor. We mean such charts (parametrisations) whenever we speak of *local charts (parametrisations) of TM or TTM*. In local coordinates we denote elements of TM by (q, v) and elements of TTM by (q, v, u, w), which should be interpreted as

$$T\phi(q, v) \quad \text{or} \quad TT\phi(q, v, u, w)$$

for some parametrisation $\phi : U \to M$, $U \subset \mathbb{R}^n$.

The Riemannian structure determines a Levi-Civita connection

$$\begin{aligned} \nabla : \Gamma(TM) &\to \Gamma(T^*M) \otimes \Gamma(TM) \ , \\ X &\mapsto \nabla X \ . \end{aligned}$$

We set $\nabla_V(X) := \nabla X(V)$ for any vector field $V \in \Gamma(TM)$. If we have a vector field X along a differentiable curve c the connection allows to define another vector field along c, the covariant derivative $\frac{D}{dt}X$ (s. [dC92], Prop. 2.2). For a vector field X along a parametrised surface $(u, v) \mapsto s(u, v)$ we can define $\frac{D}{\partial u}X$ and $\frac{D}{\partial v}X$ accordingly (s. [dC92], Def 3.3).

We should mention, how these constructions appear in local coordinates. Given some local parametrisation $\phi : U \to M$, the partial derivatives of ϕ define n linear independent vector fields $e_1, \dots e_n$ on $\phi(U)$, i.e.

$$e_i|_{\phi(q)} := \frac{\partial \phi}{\partial q_i}|_q \ .$$

Now let x^i, v^i be the components of the vector field X and V with respect to the basis $(e_1, \dots, e_n)$.

$$\nabla_V(X) = \sum_{ij} v^j \frac{\partial x^i}{\partial q_j} e_i + \sum_{ijk} v^i x^j \Gamma^k_{ij} e_k \ , \tag{3.1}$$

where the Christoffel symbols Γ^k_{ij} are defined as smooth functions on the domain of definition of ϕ.

For a vector field $X(t)$ along $c(t)$ (i.e. a differentiable map $X :]a, b[\to TM$ with $\tau_M(X(t)) = c(t)$) we can define components of $X(t)$ along $c(t)$ by $X(t) = \sum_i x^i(t) e_i(c(t))$. We set $\frac{d}{dt} c(t) = \sum_i v^i(t) e_i(c(t))$.

$$\frac{DX}{dt}(t) = \sum_i \dot{x}^i(t) e_i(t) + \sum_{ijk} v^i(t) x^j(t) \Gamma^k_{ij} e_k \ . \tag{3.2}$$

It is sometimes useful to abbreviate the term that is due to the curvature of M by

$$A_q(V, X) := \sum_{ijk} v^i x^j \Gamma^k_{ij} e_k \ , \tag{3.3}$$

which defines *locally* a smooth bilinear form.

The vector bundle $TTM \to TM$ contains the vertical subbundle $V = (T\tau_M)^{-1}(0_M)$. If we have $v, w \in T_qM$ we define the *vertical lift* of w over (q, v) as

$$vl_{(q,v)}(w) := \frac{d}{dt}(q, v + tw)|_{t=0} \in T_{q,v}(TM) \ .$$

The vertical lift defines an ismorphism $T_qM \to V_{(q,v)}$.

In local coordinates we have $vl_{(q,v)}(w) = (q, v, 0, w)$ (compare [AM78], Definition 3.7.5 and the following remark).

On the other hand, for any z in $T_{(q,v)}TM$ we define the *horizontal part* of z to be

$$z_{hor} := T\tau_M(z) \in T_qM \ .$$

In local coordinates we have $(q, v, u, w)_{hor} = (q, u)$.

The connection allows us to define the *vertical part* of $z \in T_{(q,v)}TM$ as well: Let z be represented by a differentiable curve $\gamma : I \to TM$ with $\gamma(0) = (q, v)$ and $\dot\gamma(0) = z$. We can define a path in T_qM by defining $\tilde\gamma(t)$ as the parallel transport of $\gamma(t)$ to T_qM via $\tau\gamma$. Now

$$vl_{(q,v)}(z_{vert}) := \frac{d}{dt}\tilde\gamma(t) \in V_{(q,v)}.$$

As the vertical lift is injective, this defines $z_{vert} \in T_qM$. In local coordinates we have

$$(q, v, u, w)_{vert} = (q, w + A_q(v, u)) .$$

Finally, the *horizontal lift* of $u \in T_qM$ over (q, v) is defined as follows. Let v be represented by a differentiable curve $\gamma : I \to M$ with $\gamma(0) = v$ and $\dot\gamma(0) = v$. Now we define a curve $\tilde\gamma : I \to M$ by the parallel transport of u via γ and set $hl_{(q,v)}(u) := \dot{\tilde\gamma}(0)$. In local coordinates we have $hl_{(q,v)}(u) = (q, v, u, -A_q(v, u))$.

The elements of TTM with vanishing vertical part form the *horizontal subbundle* H over TM. We have

$$TTM \cong H \oplus V \tag{3.4}$$

as bundles over TM. We denote by z_V and z_H the components of any $z \in TTM$ with respect to this splitting. The vertical (horizontal) part construction defines an isomorphism $H \cong TM$ ($V \cong TM$). Please note that $(z_V)_{vert} = z_{vert}$ and $(z_H)_{hor} = z_{hor}$.

A metric on TM can defined in many ways, but for us it is convenient to chose the Sasaki metric: For z, $\tilde z \in T_{(q,v)}TM$ set

$$\langle z, \tilde z\rangle := \langle z_{vert}, \tilde z_{vert}\rangle + \langle z_{hor}, \tilde z_{hor}\rangle .$$

Thus, in the Sasaki metric H and V are orthogonal subbundles.

If $f : TM \to \mathbb{R}$ is differentiable, the differential Df is a section of T^*TM and the gradient ∇f is a vector field on TM.

For $q \in M$ and $w, v \in T_qM$ we set

$$D_Hf(q, v)(w) = Df(q, v)(hl_{q,v}(w)) \quad \text{and} D_Vf(q, v)(w) = Df(q, v)(vl_{q,v}(w)) .$$

We denote the components of ∇f with respect to the splitting (3.4) by $\nabla_H f$ and $\nabla_V f$.

It should be noted that the projections $(\nabla_H f)_{hor} = (\nabla f)_{hor}$ and $(\nabla_V f)_{vert} = (\nabla f)_{vert}$ do *not* depend on the Sasaki metric, as we see by the following argument: Let $(q, v) \in TM$. Then for all $w \in T_qM$ we have

$$Df(q, v)(hl_{(q,v)}(w)) = \langle \nabla_H f(q, v), hl_{(q,v)}(w)\rangle_{(q,v)} = \langle (\nabla_H f(q, v))_{hor}, w\rangle_q , \tag{3.5}$$

and correspondingly for the vertical part. We will continue to write $(\nabla_V f)_{vert}$ instead of $(\nabla f)_{vert}$ and $(\nabla_V f)_{hor}$ instead of $(\nabla f)_{hor}$, as we consider these slightly redundant expressions more suggestive of their meaning.

Please note that

$$\begin{aligned}
&\langle \nabla_V f(q,v)_{vert}, w\rangle_q \\
&= \langle \nabla f(q,v), vl_{q,v}(w)\rangle_{(q,v)} \\
&= Df(q,v)(vl_{(q,v)}(w)) \\
&= D_V f(q,v)(w) \ .
\end{aligned}$$

and correspondingly

$$\langle \nabla_H f(q,v)_{hor}, w\rangle_q = D_H f(q,v)(w) \ .$$

The horizontal and the vertical derivatives behave differently. There is a much easier way to describe the vertical derivative of f, which is not possible for the horizontal derivative: The fibre T_qM has the canonical structure of a normed vector space. Therefore the Fréchet derivative of

$$g : T_qM \to \mathbb{R}, v \mapsto f(q,v)$$

is well defined at any $v \in T_qM$ and

$$\begin{aligned}
Dg(v)(w) &= \frac{d}{ds} g(v+sw)|_{s=0} = \frac{d}{ds} f(q, v+sw)|_{s=0} \\
&= Df(q,v)(vl_{(q,v)}(w)) = D_V f(q,v)(w) \ .
\end{aligned}$$

That means, we could have defined $D_V f(q,v)$ simply as the derivative of $f(q,\cdot)$ at v.

For $f \in C^k$ we can as well define the k-fold vertical derivative $D_V^k f(q,v)$ as the k-linear form on $(T_qM)^k$ which is the the k-th Fréchet derivative of

$$g : T_qM \to \mathbb{R}, v \mapsto f(q,v)) \ .$$

In general, higher derivatives on manifolds are tricky. At critical points $z \in TM$ of f we can define a second derivative in the ordinary sense by

$$D^2 f(z)(u,w) := \frac{d^2}{dsdt} f(\sigma(s,t))|_{s=0,t=0} \ ,$$

where $\sigma(s,t)$ is a parametrised surface with $\sigma(0,0) = z$, $\frac{\partial}{\partial s}\sigma(0,0) = u \in T_zTM$ and $\frac{\partial}{\partial t}\sigma(0,0) = w \in T_zTM$. At critical points this definition does not depend on the choice of σ, at general points it does. By means of the Levi-Civita connection we *can* define the covariant Hessian of f at all points as $\nabla\nabla f$, which is a section of $L(TM, TM)$.

We will mostly assume that

(MC) M is compact.

If M is compact, M can be smoothly, isometrically and equivariantly embedded as a closed submanifold in some G-representation $\mathbb{R}^N$ ([MS80]), thus TM in $\mathbb{R}^{2N}$. This embedding also defines a Riemannian structure on TM, which in fact is not the Sasaki metric. The images of H and V are still orthogonal in this metric, but the projection to the horizontal part does not induce an isometry on the fibres of H.

If we drop the assumption of compactness, such an embedding does no longer have to exist. In this case we just assume that it does:

(ME) M can be smoothly, isometrically and equivariantly embedded as a closed submanifold in some G-representation $\mathbb{R}^N$.

The embedding gives a chart-independent meaning to the notation (q, v) for elements of $TM \subset \mathbb{R}^{2N}$.

3.1.2. Manifolds of Sobolev loops

We will sketch two different ways of defining the Sobolev manifold, on which we intend to do calculus of variations.

1. The *first definition* makes use of the isometric embedding $M \to \mathbb{R}^N$:

   The Sobolev space $H^1(]0,T[, \mathbb{R}^N)$ of absolutely continuous functions with square integrable derivative embeds (compactly) in $C^0([0,T], \mathbb{R}^N)$, it contains the closed (split) subspace of T-periodic functions

   $$H := H^1_T(S^1, \mathbb{R}^N) = H^1(]0,T[, \mathbb{R}^N) \cap \{f \in C^0([0,T], \mathbb{R}^N) |\ f(0) = f(T)\}\ .$$

   We define

   $$\begin{aligned} X := H^1_T(S^1, M) := H^1_T(S^1, \mathbb{R}^N) \\ \cap \{f \in C^0([0,T], \mathbb{R}^N) |\ f(t) \in M \text{ for all } t \in [0,T]\}\ . \end{aligned}$$

   This is a closed subset of the Hilbert space $H^1_T(S^1, \mathbb{R}^N)$, as the evaluation at any t is a continuous map. We define X_0 as the component of X that contains the contractible curves.

In order to see, that X is actually a smooth closed Hilbert-submanifold of the Hilbert space $H^1_T(S^1,\mathbb{R}^N)$, we need a bit more work. Apparently, this is all well known, but we could find no proper reference. Therefore, we sketch a proof, which is based on smooth retractions:

There is an open (tubular) neighbourhood U of M which retracts to M via a C^∞-map $r: U \to M$. If $i_M : M \to U$ is the embedding we have $r \circ i_M = id_M$ and

$$(i_M \circ r) \circ (i_M \circ r) = i_M \circ id_M \circ r = i_M \circ r.$$

Thus, the map $\tilde{r} := i_M \circ r$ is idempotent. The set

$$\tilde{U} := \{\gamma \in H^1_T(S^1,\mathbb{R}^N) | \gamma(t) \in U \text{ for all } t\}$$

is open in $H^1_T(S^1,\mathbb{R}^N)$ (as intersection of the Hilbert space with the open set of all *continuous maps* $S^1 \to U$.) Now we define

$$R : \tilde{U} \to \tilde{U},\ \gamma \mapsto \tilde{r} \circ \gamma\ .$$

This map is smooth (as all derivatives of r are bounded in a compact neighbourhood of $\gamma([0,T])$) and idempotent, its image coincides with X and $R|_X = id_X$.

Now we conclude by Lemma A.1 from the Appendix that X is a Hilbert-submanifold of $H^1_T(S^1,\mathbb{R}^N)$.

At any $\gamma \in X$ the derivative $L_\gamma := DR(\gamma)$ is idempotent as well and defines a splitting

$$H^1_T(S^1,\mathbb{R}^N) \cong \text{Ker } (L_\gamma) \oplus \text{Im } (L_\gamma)\ . \tag{3.6}$$

The tangent space of X at γ is given by

$$T_\gamma X := \{\frac{\partial}{\partial s}\gamma_s|_{s=0}\quad \Big| \gamma_\cdot :]-\epsilon,\epsilon[\to X \subset H^1_T(S^1,\mathbb{R}^N) \text{ differentiable in } s=0 \text{ and } \gamma_0 = \gamma\ \}\ .$$

By the lemma, we have $T_\gamma X = \text{Im } L_\gamma = \text{Ker } (\text{id} - L_\gamma)$, i.e. a closed subspace of the Hilbert space $H^1_T(S^1,\mathbb{R}^N)$, and thus a Hilbert space with the inner product $\langle\cdot,\cdot\rangle_e$ it inherits. This makes $(X,\langle\cdot,\cdot\rangle_e)$ a complete Riemannian manifold and defines an "extrinsical" norm $\|\cdot\|_e$ on its fibres.

2. The *second definition* makes use of local charts. A map $\gamma : I \to M$ is defined to be in $X := H^1_T(S^1, M)$, if for every C^∞-chart $M \supset U \xrightarrow{\phi} V \subset \mathbb{R}^n$ the map $\phi \circ \gamma|_{\gamma^{-1}(U)}$ is in $H^1_{loc}(\gamma^{-1}(U),\mathbb{R}^n)$.

We sketch the construction of local charts via vector bundle neighbourhoods following Palais and Terng ([PT88], 11.1.8-11.1.10):

In order to obtain a chart close to some $\gamma \in X$, we would like to pull back TM via γ. γ however is only continuous, not smooth in general. Therefore we need two steps: *First* we chose a smooth $\tilde{\gamma}$ C^0-close to γ, *then* we pull back TM via $\tilde{\gamma}$. More precisely this is done as follows.

For $\gamma \in X$ we can chose an open neighbourhood V of $\gamma([0,T])$ with compact closure. The minimum m of the point-wise injectivity radius on $\bar{V}$ is positive, hence there is an ϵ with $0 < 2\epsilon < m$ and $B_{2\epsilon}(\gamma(t)) \subset V$ for all $t \in [0,T]$. Now, we can chose a smooth curve $\tilde{\gamma}$ such that

$$\max\ d(\gamma(t), \tilde{\gamma}(t)) < \epsilon.$$

There is a C^0-neighbourhood U of the zero section in $\tilde{\gamma}^*TM$ such that the exponential map induces a diffeomorphism of $U_{(t,\tilde{\gamma}(t))}$ to an ϵ-neighbourhood of $\tilde{\gamma}(t)$ for each $t \in [0,T]$. $\tilde{\gamma}^*TM$ is trivial as a vector bundle over $[0,T]$, hence there is a smooth map $\Psi : \tilde{\gamma}^*TM \to \mathbb{R}^n$, which is a linear map $\Psi_t : \gamma(t)^*TM \to \mathbb{R}^n$ on each fibre. The linear isomorphism $\mathbb{R}^n \to \gamma(0)^*TM = \gamma(T)^*TM \to \mathbb{R}^n$ is represented by an $A \in GL(n,\mathbb{R})$. $GL(n,\mathbb{R})$ has two path components, A is either in the same component of the identity matrix or in the component of $\mathbf{diag}(-1,1,\ldots,1)$. We can therefore modify Ψ_t, so that A is either the identity or $\mathbf{diag}(-1,1,\ldots,1)$.

A section c of U is periodic iff the corresponding function $\tilde{c} := \Psi \circ c : [0,T] \to \mathbb{R}^n$ satisfies the condition $\tilde{c}(T) = A\tilde{c}(0)$.

We are looking for a chart in a neighbourhood of the fixed element $\gamma \in X$. The homotopy class of γ determines the matrix A (either the identity or $\mathbf{diag}(-1,1,\ldots,1)$). We make the following definition (which depends on A and therefore on γ):

$H^1_{T^*}(S^1,\mathbb{R}^n)$ is defined as the set of H^1-functions $c : [0,T] \to \mathbb{R}^n$ that satisfy the condition $\tilde{c}(T) = A\tilde{c}(0)$.

Now for every T-periodic H^1-section $v(t)$ of $U \to [0,T]$ the map $\Phi(v) : t \mapsto \exp_{\tilde{\gamma}(t)}(v(t))$ is an element of X. We can define a chart from a C^0-ϵ-neighbourhood U' of $\tilde{\gamma}$ in X to an open subset V' of $H^1_{T^*}(S^1,\mathbb{R}^n)$:

$$\Theta : U' \to V', \Theta(c)(t) = \Psi((\Phi^{-1}(c))(t)) \ .$$

As $\gamma \in U'$, we have thus given a chart at γ.

The model space in this construction depends on the homotopy class of γ. In two cases, however, we need just one:

a) If M is orientable, the pullback of TM via any closed curve $S^1 \to M$ is orientable as well and thus trivial. (The set of ismorphy classes of n-vector bundles over S^1 maps bijectively to $\pi_0(SO(n)) = 0$ in this case.) Therefore A can be chosen as the identity and we obtain the untwisted space $H^1_T(S^1, \mathbb{R}^n)$ as model space.

b) For the component X_0 of contractible curves the same is valid. If a closed curve in M is contractible, the pullback of TM via this curve is trivial. As above we obtain the model space $H^1_T(S^1, \mathbb{R}^n)$.

The parametrisation of X that corresponds to Θ can be written as

$$\Theta^{-1} : V' \to U', \Xi(v)(t) = \phi(t, v(t)) ,$$

where $\phi(t, \cdot) = \phi_t(\cdot)$ is the family of parametrisations of M defined by

$$\phi_t(w) := \exp_{\tilde{\gamma(t)}}(\Psi_t(w)) .$$

Any family of smooth parametrisations $\phi_t(w)$ defines a local chart of X. Usually it is immaterial whether the parametrisation is derived from the exponential map.

In this second setting X does not inherit a Riemannian structure from an embedding. There is still a "natural" candidate for a Riemannian structure on X.

Again, if γ_s is a differentiable curve in X with $\gamma = \gamma_0$ we have

$$v := \frac{\partial}{\partial s}\gamma_s|_{s=0} \in T_\gamma X.$$

Please note that in local charts of M we have

$$v = \sum_i v^i(t)\frac{\partial}{\partial x_i}$$

with v^i in H^1_{loc}. In particular the v^i are continuous.

We define

$$\langle v, v\rangle_i = \|v\|_i^2 := \int_0^T (\langle v(t), v(t)\rangle_{\gamma(t)} + \langle \frac{D}{dt}v(t), \frac{D}{dt}v(t)\rangle_{\gamma(t)})dt ,$$

where $\frac{D}{dt}$ means the vertical part of $\frac{d}{dt}(\gamma(t), v(t))$ as defined by the Levi-civita connection.

We check that the integral is well-defined: The $v(t)$ are continous, so the first summand is integrable. For the second summand we write in local charts:

$$\frac{D}{dt}(\sum_i v^i(t)\frac{\partial}{\partial x_i}) \tag{3.7}$$

$$= \sum_i \dot{v}^i \frac{\partial}{\partial x_i} + \sum_{ijk} \Gamma^k_{ij} v^i \dot{\gamma}^j \frac{\partial}{\partial x_k} \; . \tag{3.8}$$

As Γ^k_{ij} is bounded along $\gamma(t)$ the whole expression is in L^2, thus the second part of the integral is defined.

The bilinear form corresponding to this quadratic form defines a Hilbert structure on $T\gamma X$.

Remark 3.1.1 If a compact Lie group G operates on M by isometries and $\gamma \in H^1_T(X)$ is a constant curve with value a fixed point p of M, G acts by isometries on $T_\gamma X \cong H^1_T(S^1, \mathbb{R}^n)$, and the local chart of $T_\gamma X$ is G-equivariant. $\diamond$

After losing considerable time wondering about the advantages and differences of the different definitions we would like to point out one fact:

Proposition 3.1.2 *The two Hilbert structures on $T\gamma X$ do* not *coincide, but they define equivalent norms, i.e.*

$$\| \cdot \|_i \leq \| \cdot \|_e \leq C(\gamma) \| \cdot \|_i \; ,$$

where $C(\gamma) = (1 + C_2 \|\dot{\gamma}\|^2_{L^2})^{1/2}$ for some universal constant C_2, if M is compact.

If M is non-compact, for any compact subset $K \subset M$ there is a constant C_2 such that the inequality is valid with $C(\gamma) = (1 + C_2 \|\dot{\gamma}\|^2_{L^2})^{1/2}$ for every γ such that $\gamma([0, T]) \subset K$.

Proof: The second inequality is obvious, as $\frac{D}{dt}v(t)$ can be identified with the orthogonal projection of $\frac{d}{dt}v(t) \in \mathbb{R}^N$ to $T_{\gamma(t)}M$.

We prove the first inequality for a smooth γ. In this case an orthonormal base of $T_{\gamma 0}$ defines by parallel transport n vectorfields $X_1(t), \ldots, X_n(t)$ along γ, which form an orthonormal base at every $t \in [0, T]$.

Now every $v \in T_\gamma X$ has a unique representation as

$$v = \sum v^i X_i \; . \tag{3.9}$$

We define $\tilde{v} = (v_1, \ldots, v_n) : [0,T] \to \mathbb{R}^n$ and have the following equalities. (The standard scalar products in $\mathbb{R}^n$ and $\mathbb{R}^N$ are denoted by $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ and $\langle \cdot, \cdot \rangle_{\mathbb{R}^N}$. The p-norm on $\mathbb{R}^n$ or $\mathbb{R}^N$ is, as usual, $\| \cdot \|_p$.)

$$\langle v(t), v(t) \rangle_{\mathbb{R}^N} = \langle \tilde{v}(t), \tilde{v}(t) \rangle_{\mathbb{R}^n} \tag{3.10}$$

$$\langle \frac{d}{dt} v(t), \frac{d}{dt} v(t) \rangle_{\mathbb{R}^N} = \langle \frac{d}{dt} \tilde{v}(t), \frac{d}{dt} \tilde{v}(t) \rangle_{\mathbb{R}^n} + \langle \sum_i v^i(t) \dot{X}_i(t), \rangle \sum_i v^i(t) \dot{X}_i(t) \rangle_{\mathbb{R}^N} \tag{3.11}$$

$$\langle \frac{D}{dt} v(t), \frac{D}{dt} v(t) \rangle_{\mathbb{R}^N} = \langle \frac{d}{dt} \tilde{v}(t), \frac{d}{dt} \tilde{v}(t) \rangle_{\mathbb{R}^n} \ . \tag{3.12}$$

Furthermore

$$\|\dot{X}_i(t)\|_2 = \|B(\dot{\gamma}(t), X_i(t))\|_2 \le C_1 \|\dot{\gamma}(t)\|_2 \ \|X_i(t)\|_2 \ ,$$

where $B_p(\cdot, \cdot) : T_pM \times T_pM \to T_pM^{\perp}$ is the second fundamental form of the embedding $M \to \mathbb{R}^N$ and C_1 a uniform bound for B_p on the manifold M, if M is compact. If $K \subset M$ is compact and $\gamma([0,T]) \subset K$ we take C_1 to be abound for B_p on K. We conclude that

$$\begin{aligned}
&\langle \sum_i v^i(t) \dot{X}_i(t), \sum_i v^i(t) \dot{X}_i(t) \rangle_{\mathbb{R}^N} \\
&\le (\sum_i |v^i(t)| \|\dot{X}_i(t)\|_2)^2 \\
&\le (\sum_i |v^i(t)| C_1 \|\dot{\gamma}(t)\|_2 \ \|X_i(t)\|_2)^2 \\
&\le C_1^2 \|\dot{\gamma}(t)\|_2^2 \|\tilde{v}(t)\|_1^2 \\
&\le C_1^2 \|\dot{\gamma}(t)\|_2^2 n \|\tilde{v}(t)\|_2^2 \ .
\end{aligned}$$

Together with the equations (3.10), (3.11), (3.12) we obtain for $v \in T_\gamma X$, $\gamma \in X \cap C_T^\infty(S^1, M)$

$$\begin{aligned}
\|v\|_e^2 &\le \|\tilde{v}\|_{L^2}^2 + \|\dot{\tilde{v}}\|_{L^2}^2 + C_1^2 n \|\tilde{v}\|_\infty^2 \int_0^T \|\dot{\gamma}(t)\|_2^2 dt \\
&\le \|\tilde{v}\|_{L^2}^2 + \|\dot{\tilde{v}}\|_{L^2}^2 + C_2 \|\dot{\gamma}\|_{L^2}^2 (\|\tilde{v}\|_{L^2}^2 + \|\dot{\tilde{v}}\|_{L^2}^2) \text{ (Sobolev embedding)} \\
&\le (1 + C_2 \|\dot{\gamma}\|_{L^2}^2) \|v\|_i^2 \ .
\end{aligned}$$

As $\|\cdot\|_i$ and $\|\cdot\|_e$ are continuous as functions of (γ, v) with respect to the topology of TX induced by the embedding $TX \to H^1_T(S^1, \mathbb{R}^N) \oplus H^1_T(S^1, \mathbb{R}^N)$, and the smooth functions are dense in X, the inequalities are valid for every $\gamma \in X$.

□

Two short lemmas will be useful for the subsequent analysis. They generalise Schwarz' or Clairaut's theorem on the commutativity of partial derivatives.

Lemma 3.1.3 *If $[-\epsilon, \epsilon] \to H^1_T(S^1, \mathbb{R}^N)$, $s \mapsto \gamma_s$ is a continuous curve, differentiable at $s = 0$ with $\frac{\partial \gamma_s}{\partial s} = u$, we have for almost all $t \in \mathbb{R}$*

$$\frac{\partial}{\partial s}\frac{\partial}{\partial t}\gamma_s(t)\Big|_{s=0} = \frac{\partial}{\partial t}\frac{\partial}{\partial s}\gamma_s(t)\Big|_{s=0} .$$

Proof: By a slight abuse of notation we suppose that γ_s is a curve of continuous representatives of the respective classes in H^1_T. Now

$$\frac{\partial}{\partial s}(\gamma_s(t+\Delta t) - \gamma_s(t))\Big|_{s=0} = \frac{\partial}{\partial s}\int_t^{t+\Delta t} \frac{\partial}{\partial \tau}\gamma_s(\tau) d\tau\Big|_{s=0} , \tag{3.13}$$

as $\gamma_s(\cdot)$ is absolutely continuous for each s. By assumption on γ_s the difference quotient $\frac{\gamma_s - \gamma_0}{s}$ converges in $H^1_T(S^1, \mathbb{R}^N)$ to u for $s \to 0$. Therefore $\frac{\partial}{\partial t}\frac{\gamma_s - \gamma_0}{s}$ converges in $L^2([0,T], \mathbb{R}^N)$ to $\dot{u}$ for $s \to 0$. Hence

$$\begin{aligned}
(3.13) &= \int_t^{t+\Delta t} \frac{\partial}{\partial s}\frac{\partial}{\partial \tau}\gamma_s(\tau)\Big|_{s=0} d\tau && (3.14)\\
&= \int_t^{t+\Delta t} \frac{\partial}{\partial s}\frac{\partial}{\partial \tau}\gamma_s(t)\Big|_{s=0} d\tau + \int_t^{t+\Delta t} (\frac{\partial}{\partial s}\frac{\partial}{\partial \tau}\gamma_s(\tau) - \frac{\partial}{\partial s}\frac{\partial}{\partial \tau}\gamma_s(t))\Big|_{s=0} d\tau && (3.15)\\
&= \frac{\partial^2}{\partial s \partial t}\gamma_s(t)\Big|_{s=0} \Delta t + \int_t^{t+\Delta t} (\frac{\partial}{\partial s}\frac{\partial}{\partial \tau}\gamma_s(\tau) - \frac{\partial}{\partial s}\frac{\partial}{\partial \tau}\gamma_s(t))\Big|_{s=0} d\tau. && (3.16)
\end{aligned}$$

Almost all $t \in \mathbb{R}$ are Lebesgue points of $\frac{\partial}{\partial s}\frac{\partial}{\partial t}\gamma_s(t)\Big|_{s=0}$, thus at almost all $t \in \mathbb{R}$ the second summand is $o(\Delta t)$. This means that the coefficient of Δt in the first summand is in fact $\frac{\partial}{\partial t}\frac{\partial}{\partial s}\gamma_s(t)\Big|_{s=0}$. □

Lemma 3.1.4 *If $[-\epsilon, \epsilon] \to X$, $s \mapsto \gamma_s$ is a continuous curve, differentiable at $s = 0$ with $\frac{\partial \gamma_s}{\partial s} = u \in T_{\gamma_0}X$, we have for almost all $t \in \mathbb{R}$*

$$\frac{D}{\partial s}\frac{\partial}{\partial t}\gamma_s(t)\bigg|_{s=0} = \frac{D}{\partial t}\frac{\partial}{\partial s}\gamma_s(t)\bigg|_{s=0} .$$

Proof: We use the embeddings $M \to \mathbb{R}^N$ and $X \to H^1_T(S^1, \mathbb{R}^N)$. The curve in X is a curve in $H^1_T(S^1, \mathbb{R}^N)$ with the same properties. By Lemma 3.1.3 the partial derivatives commute at $s = 0$ for almost every t. By projecting these second derivatives to $T_{\gamma_0(t)}M$ the result follows.

We might as well prove the result via local charts of M: Formula (3.2) and Lemma 3.1.3 entail our lemma. □

The completeness of $(X, \langle\cdot,\cdot\rangle_e)$ was a consequence of the closed embedding. As the comparison of the two norms depends on the base point, we have to verify that $(X, \langle\cdot,\cdot\rangle_i)$ is complete as well.

Proposition 3.1.5 *For (MC) or (ME) $(X, \langle\cdot,\cdot\rangle_i)$ is a complete Riemannian manifold.*

Proof: We first prove the fact under the assumption (MC) that M is compact.

The geodesic distance on X defined by $\langle\cdot,\cdot\rangle_i$ (or $\langle\cdot,\cdot\rangle_e$) is denoted by $d_i(\cdot,\cdot)$ (or $d_e(\cdot,\cdot)$). As the estimation of $\|\cdot\|_e$ by $\|\cdot\|_i$ from above involves $\|\dot\gamma\|_{L^2}$ we have to establish a local bound for $\|\dot\gamma\|_{L^2}$ with respect to the metric d_i.

Suppose $d_i(\gamma, \tilde\gamma) < \epsilon$. Then there is a smooth path $[0, 1] \to X$, $s \mapsto \gamma_s$ with $\gamma_0 = \gamma$ and $\gamma_1 = \tilde\gamma$ such that

$$\int_0^1 \|\frac{d}{ds}\gamma_s\|_i ds < 2\epsilon .$$

Now define

$$f(s) := \|\dot\gamma_s\|_{L^2} = \sqrt{\int_0^T \langle\frac{\partial\gamma_s(t)}{\partial t}, \frac{\partial\gamma_s(t)}{\partial t}\rangle dt} .$$

f is a continuous function and assumes its maximum in some $s_0 \in [0,1]$.

$$\begin{aligned}
(f(s_0))^2 &= (f(0))^2 + \int_{s_0}^{0} \frac{d}{ds}(f(s))^2 ds \\
&= (f(0))^2 + \int_{s_0}^{0} 2\int_0^T \langle \frac{D}{\partial s}\frac{\partial}{\partial t}\gamma_s(t), \frac{D}{\partial s}\frac{\partial}{\partial t}\gamma_s(t)\rangle dt ds \\
&= (f(0))^2 + \int_{s_0}^{0} 2\int_0^T \langle \frac{D}{\partial t}\frac{\partial}{\partial s}\gamma_s(t), \frac{D}{\partial s}\frac{\partial}{\partial t}\gamma_s(t)\rangle dt ds \quad \text{(s. Lemma 3.1.4)} \\
&\leq (f(0))^2 + 2\int_{s_0}^{0} \|\frac{d\gamma_s}{ds}\|_i \|\dot{\gamma}_s\|_{L_2} ds \\
&= (f(0))^2 + 2\int_0^1 \|\frac{d\gamma_s}{ds}\|_i f(s) ds \\
&\leq (f(0))^2 + 2f(s_0)\int_0^1 \|\frac{d\gamma_s}{ds}\|_i ds \\
&\leq \|\dot{\gamma}\|_{L_2}^2 + 4\epsilon f(s_0)\ .
\end{aligned}$$

Now this implies the following inequality for $f(s_0)$

$$(f(s_0) - 2\epsilon)^2 \leq \|\dot{\gamma}\|_{L_2}^2 + 4\epsilon^2$$

and thus for all $s \in [0,1]$.

$$\|\dot{\tilde{\gamma}}_s\|_{L^2} \leq f(s_0) \leq 2\epsilon + \sqrt{\|\dot{\gamma}\|_{L_2}^2 + 4\epsilon^2}\ .$$

We derive especially

$$\begin{aligned}
d_e(\gamma,\tilde{\gamma}) &\leq \int_0^1 \|\frac{d\gamma_s}{ds}\|_e ds \\
&\leq \int_0^1 (1 + C_2\|\dot{\gamma_s}\|_{L^2}^2)^{1/2} \|\frac{d\gamma_s}{ds}\|_i ds \\
&\leq \int_0^1 (1 + C_2(2\epsilon + \sqrt{\|\dot{\gamma}\|_{L_2}^2 + 4\epsilon^2})^2)^{1/2} \|\frac{d\gamma_s}{ds}\|_i ds \\
&\leq (1 + C_2(2\epsilon + \sqrt{\|\dot{\gamma}\|_{L_2}^2 + 4\epsilon^2})^2)^{1/2} d_i(\gamma,\tilde{\gamma})\ .
\end{aligned}$$

Suppose γ_n is a Cauchy-sequence with respect to d_i. For $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that $d_i(\gamma_m,\gamma_n) < \epsilon$ for all $m,n \geq N$. Thus

$$\begin{aligned}
d_e(\gamma_m,\gamma_n) &\leq (1 + C_2(2\epsilon + \sqrt{\|\dot{\gamma}_m\|_{L_2}^2 + 4\epsilon^2})^2)^{1/2} d_i(\gamma_m,\gamma_n) \\
&\leq (1 + C_2(2\epsilon + \sqrt{(2\epsilon + \sqrt{\|\dot{\gamma}\|_{L_2}^2 + 4\epsilon^2})^2 + 4\epsilon^2})^2)^{1/2} d_i(\gamma_m,\gamma_n)\ .
\end{aligned}$$

Thus $(\gamma_n)_{n\in\mathbb{N}}$ is a Cauchy sequence with respect to d_e. By completeness it converges in (X, d_e) and hence in (X, d_i).

Now we drop the assumption (MC) and assume (ME) only. For a possibly non-compact manifold M, the constant C_2 can be chosen for $\gamma([0,T]) \subset K$ with some compact $K \subset M$. We establish our result, if we prove, that a Cauchy-sequence with respect to d_i is Cauchy with respect to the C^0-distance defined via $X \hookrightarrow H^1_T(S^1, \mathbb{R}^N) \hookrightarrow C^0_T(S^1, \mathbb{R}^N)$.

First we have to establish a preliminary estimate. For $v \in T_\gamma X$ we define

$$\|v\|_\infty := \sup_{0\le t\le T} \|v(t)\|_2 \ .$$

As v is continuous, this supremum is in fact a maximum, i.e. there is t_1 such that

$$\|v(t_1)\|_2 = \|v\|_\infty \ .$$

Now we have the following estimates (cf. [Kli78], Prop 1.2.1)

$$\begin{aligned}\|v\|_\infty^2 &= \|v(t)\|_2^2 + \int_t^{t_1} \frac{d}{ds}(\|v(s)\|_2^2)ds \\ &\le \|v(t)\|_2^2 + 2\int_0^T \|v(s)\|_2 \|\frac{D}{ds}v(s)\|_2 ds \ .\end{aligned}$$

Now by integrating both sides over t from 0 to T we obtain

$$\begin{aligned}T\|v\|_\infty^2 &\le \|v\|_2^2 + T(\|v\|_2^2 + \int_0^T \|\frac{D}{ds}v(s)\|_2^2\|ds) \\ &\le (T+1)\|v\|_i \ .\end{aligned}$$

So,

$$\|v\|_\infty \le \sqrt{\frac{T+1}{T}}\|v\|_i^2 \ . \tag{3.17}$$

The geodesic distance on X defined by $\langle\cdot,\cdot\rangle_i$ is denoted by $d_i(\cdot,\cdot)$. We compare this to the C^0-distance defined via $X \hookrightarrow H^1_T(S^1,\mathbb{R}^N) \hookrightarrow C^0_T(S^1,\mathbb{R}^N)$. Let $s \mapsto \gamma_s(t)$ a differentiable curve in X with $\gamma_0 = \gamma$ and $\gamma_1 = \tilde\gamma$.

$$\begin{aligned}\|\gamma - \tilde\gamma\|_{C^0} &= \|\gamma(t_0) - \tilde\gamma(t_0)\|_{\mathbb{R}^N} \\ &= \|\int_0^1 \frac{\partial\gamma_s(t_0)}{\partial s} ds\|_2 \\ &\le \int_0^1 \|\frac{\partial\gamma_s(t_0)}{\partial s}\|_2 ds \\ &\le \int_0^1 \sqrt{\frac{T+1}{T}}\|\frac{\partial\gamma_s}{\partial s}\|_i ds \ (\text{s.}(3.17)) \ .\end{aligned}$$

Thus, minimizing over all possible γ_s, we obtain

$$\|\gamma - \tilde{\gamma}\|_{C^0} \leq \sqrt{\frac{T+1}{T}} d_i(\gamma, \tilde{\gamma}) \; .$$

A Cauchy-sequence with respect to d_i is therefore a C^0-Cauchy-sequence, and the statement follows. $\square$

Remark 3.1.6 Both manifold structures of X coincide, as one can prove by extending a chart of the second type to a chart of $H^1_T(S^1, \mathbb{R}^N)$. In specific situations it remains to chose one of the Riemannian structures. The completeness result can be proved by reduction to a chart without referring to the embedding, as in [Kli78].

We should remark that

$$\|v\|_i = \|\tilde{v}\|_{H^1([0,T],\mathbb{R}^n)} \; ,$$

which is one more reason of considering $\|\cdot\|_i$ the more natural Riemannian structure.

For most constructions it does not matter which Riemannian structure we chose, for some, however, it does. Occasionally we define the Hessian of a functional at non-critical points. The Hessian in this case is well defined with respect to a Riemannian metric only. Differentiability itself does *not* depend on the specific metric.

More crucially, the Palais-Smale condition (PS) depends on the metric. The two Riemannian metrics above are not uniformly equivalent, as the second estimate involves a constant that depends on the base point. For the functionals we consider it makes no difference, but we suspect there may be functionals satisfying (PS) for one and not for the other. $\diamond$

3.2. The Lagrangian action functional

The linear G-representation on $\mathbb{R}^N$ induces a linear G-operation on H, and hence a smooth action on X. We suppose X is equipped with the intrinsic Hilbert Riemannian structure.

Classical mechanics derives T-periodic trajectories on a configuration space M as critical points of an action functional

$$S_L : H^1_T(S^1, M) \to \mathbb{R}, \gamma \mapsto \int_0^T L(t, (\gamma(t), \gamma'(t)))dt \; ,$$

with a **Lagrangian** function

$$L : \mathbb{R} \times TM \to \mathbb{R},\ (t, (q, v)) \mapsto L(t, (q, v)).$$

In general L will be T-periodic in t and a Caratheodory funtion, subquadratic on the fibres:

(TP) $L(t, (q, v))$ is T-periodic in t.

(CF) *(Caratheodory function)* L is measurable in t for each $(q, v) \in TM$ and $L(t, \cdot)$ is continuous for almost every $t \in \mathbb{R}$.

(SQ) *(subquadratic)* There is a continuous function $l : M \to \mathbb{R}$ such that

$$L(t, (q, v)) \leq l(q)(1 + \langle v, v\rangle_q) .$$

Lemma 3.2.1 *If the Lagrangian satisfies (TP),(CF) and (SQ) the action functional is well-defined.*

Proof: Assume $\gamma \in H^1_T(S^1, M)$. Then $(\gamma, \dot\gamma) : \mathbb{R} \to TM$ is a measurable function and can thus be approximated a.e. by step functions $s_n : \mathbb{R} \to TM$. For each n the function $L_n : t \mapsto L(t, s_n(t))$ is measurable, and by the continuity in the second argument the sequence (L_n) converges almost everywhere to $L_\gamma : t \mapsto L(t, (\gamma(t), \dot\gamma(t)))$, hence this function is measurable.

As γ is continuous, $\gamma([0, T])$ is compact and $t \mapsto l(\gamma(t))$ assumes a maximum M. We have

$$|L(t, (\gamma(t), \dot\gamma(t)))| \leq M\langle\dot\gamma(t), \dot\gamma(t)\rangle$$

and thus L_γ is integrable and S_L is well-defined. □

The functional is C^1, if we assume additionally, that the Lagrangian is continuously differentiable in the second variable, i.e.

(CD) $L(t, \cdot)$ is continuously differentiable for almost every $t \in \mathbb{R}$.

and that the derivative satisifies

(SQ') $\|\nabla_V L(t, (q, v))_{vert}\|_q \leq l_1(q)(1+\|v\|_q)$ and $\|\nabla_H L(t, (q, v))_{hor}\|_q \leq l_2(q)(1+\|v\|^2_q)$ with continuous functions $l_1, l_2 : M \to \mathbb{R}$.

Please note that

$$\|\nabla_V L(t, (q, v))\|_{(q,v)} = \|(\nabla_V L(t, (q, v)))_{vert}\|_q$$

and

$$\|\nabla_H L(t, (q, v))\|_{(q,v)} = \|(\nabla_H L(t, (q, v)))_{hor}\|_q$$

by definition of the Sasaki metric.

Remark 3.2.2 Recall that by equation (3.5)

$$\begin{aligned}&(\nabla L(t,(q,v)))_{vert} = \nabla_V L(t,(q,v)))_{vert} \quad \text{and}\\ &(\nabla L(t,(q,v)))_{hor} = \nabla_H L(t,(q,v))_{hor}\end{aligned}$$

do *not* depend on the Sasaki metric, thus (SQ') is in fact given in terms of the geometry of M.

One may still prefer a more transparent version of (SQ') in terms of local charts.

We consider local parametrisations $\phi : U \to M$ of M, $T\phi : U \times \mathbb{R}^n \to TM$ of TM and $TT\phi : U \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to TTM$, and we set $\tilde{L}(t,q,v) = L(t, T\phi(q,v))$. $(T\phi)^{-1}$ yields a representation $T_qM \to \mathbb{R}^n$, which induces a dual representation $\psi : T_q^*M \to \mathbb{R}^n$.

Let $G(q)$ the symmetric positive definite $n \times n$ matrix that represents the Riemannian structure of M. The corresponding isomorphism $\Phi_q : T_q^*M \to T_qM$ is then represented by $\mathbb{R}^n \to \mathbb{R}^n,\ l \mapsto G(q)^{-1}l$ and the induced scalar product on T_q^*M is represented by $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},\ (l,l') \mapsto \langle l, G(q)^{-1}l'\rangle$.

We recall how we split TTM over $T\phi(q,v)$, namely $TT\phi(q,v,0,w)$ is *vertical* for every $w \in \mathbb{R}^n$ and $TT\phi(q,v,u,-A_q(v,u))$ is *horizontal* for every $u \in \mathbb{R}^n$, the map $w \mapsto (q,v,0,w)$ represents the vertical lift and $u \mapsto (q,v,u,-A_q(v,u))$ the horizontal lift $T_qM \to T_{(q,v)}TM$, which are both isometries with respect to the Riemannian metrics chosen.

With these notations (SQ') is equivalent to

$$\sqrt{\langle \psi(D_v\tilde{L}(t,q,v)), G(q)^{-1}\psi(D_v\tilde{L}(t,q,v))\rangle} \le l_1(q)(1+\sqrt{\langle v, G(q)v\rangle})$$

and

$$\begin{aligned}&\Big\langle \psi(D_q\tilde{L}(t,q,v) - D_v\tilde{L}(t,q,v)A_q(v,\cdot)),\\ &\qquad\qquad G(q)^{-1}\psi(D_q\tilde{L}(t,q,v) - D_v\tilde{L}(t,q,v)A_q(v,\cdot))\Big\rangle\\ \le\, &l_2(q)(1+\langle v, G(q)v\rangle)\end{aligned}$$

with continuous functions $l_1, l_2 : U \to \mathbb{R}$ for every such chart. Here $\langle\cdot,\cdot\rangle$ denotes the standard scalar product on $\mathbb{R}^n$. As $G(q)$ and A_q depend continuously on q, this is equivalent to the much simpler form

(SQ'$_{loc}$) $\|D_v\tilde{L}(t,q,v)\| \le l_1(q)(1+\|v\|)$ and $\|D_q\tilde{L}(t,q,v)\| \le l_2(q)(1+\|v\|^2)$ with continuous functions $l_1, l_2 : U \to \mathbb{R}$ for every such chart. (Here $\|\cdot\|$ denotes the standard euclidean norm on $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$.) ◇

Under these conditions the functional S_L is C^1, as we will see.

Remark 3.2.3 Before we state this fact as a lemma we would like to consider the question of the right type of differentiability to be used in such variational problems on Hilbert or Banach manifolds.

Most proofs of Fréchet-differentiability prove that the functional is Gâteaux-differentiable with respect to some chart, and that the Gâteaux-derivative at each point is a continuous linear map which depends continuously on the point. However, there are cases, when the latter fails to be the case. Still, the proofs usually do not depend on the charts chosen. Thus they prove much more than differentiability in the sense of Gâteaux, though less than differentiability in the sense of Fréchet.

Gâteaux-differentiability of f with respect to any C^k chart is equivalent to differentiability of $f \circ \gamma$ along *any* C^k-curve γ. A closer inspection of most proofs shows that we can indeed prove differentiability of $f \circ \gamma$ at 0 for any continuous curve γ differentiable in 0. This notion occurs as *quasi-differentiability* in Dieudonné's *Foundations of Modern Analyis* ([Die60], Ch. VIII, 4., Problems 4ff). This notion is *independent of charts* and can be shown to be equivalent to *differentiability in the sense of Hadamard* with respect to any chart. $f : X \to Y$ (X, Y Banach spaces) is differentiable at $x \in X$ in the sense of Hadamard, if there is a continuous linear operator A such that

$$f(x + th) = f(x) + tAh + R(th) ,$$

where $R(th)/t \to 0$ for $t \to 0$ uniformly in $h \in S$, where S is any sequentially compact subset of X. For a finite dimensional Banach space X this is equivalent to Fréchet-differentiability. In general, a quasi-differentiable function on a Banach-manifold has the property that it is differentiable when restricted to any finite dimensional or sequentially compactly embedded submanifold (s. [Yam74], Chapter I).

For our problem this will become relevant when it comes to the second derivative. Smooth Lagrangian functions with 'quadratic growth' along the fibres in general do *not* yield a C^2-functional on our Sobolev manifold X, but a functional with quasi-differentiable Fréchet-derivative. $\diamond$

Proposition 3.2.4 *If L satisfies (TP),(CF),(CD) and (SQ'), the functional S_L is C^1. Its derivative is given by*

$$DS_L(\gamma)(u) = \int_0^T \langle \nabla_V L(t, (\gamma(t), \dot\gamma(t)))_{vert}, \frac{D}{dt}u(t)\rangle + \langle \nabla_H L(t, (\gamma(t), \dot\gamma(t)))_{hor}, u(t)\rangle dt \, . \tag{3.18}$$

Proof: The usual proofs are done in local charts. Here we present a slightly more technical intrinsical proof.

Let γ_s, $s \in I := [-\epsilon, \epsilon]$ be a differentiable curve in X with $\gamma_0 = \gamma$ and $\frac{\partial \gamma_s}{\partial s}|_{s=0} = u$.

Now

$$
\begin{aligned}
f(s,t) :=& \frac{\partial L(t, (\gamma_s(t), \dot{\gamma}_s(t)))}{\partial s} \\
=& D_2 L(t, (\gamma_s(t), \dot{\gamma}_s(t)))(\frac{\partial}{\partial s}(\gamma_s, \dot{\gamma}_s)) \\
=& \langle \nabla_V L(t, (\gamma_s(t), \dot{\gamma}_s(t))), (\frac{\partial}{\partial s}(\gamma_s(t), \dot{\gamma}_s(t))_V \rangle \\
& \qquad \langle \nabla_H L(t, (\gamma_s(t), \dot{\gamma}_s(t))), (\frac{\partial}{\partial s}(\gamma_s(t), \dot{\gamma}_s(t))_H \rangle \\
=& \langle \nabla_V L(t, (\gamma_s(t), \dot{\gamma}_s(t)))_{vert}, (\frac{\partial}{\partial s}(\gamma_s(t), \dot{\gamma}_s(t)))_{vert} \rangle \\
& \qquad + \langle \nabla_H L(t, (\gamma_s(t), \dot{\gamma}_s(t)))_{hor}, (\frac{\partial}{\partial s}(\gamma_s(t), \dot{\gamma}_s(t)))_{hor} \rangle \\
=& \langle \nabla_V L(t, (\gamma_s(t), \dot{\gamma}_s(t)))_{vert}, \frac{D}{\partial s}\frac{\partial}{\partial t}\gamma_s(t) \rangle + \langle \nabla_H L(t, (\gamma_s(t), \dot{\gamma}_s(t)))_{hor}, \frac{\partial}{\partial s}\gamma_s(t) \rangle \\
=& \langle \nabla_V L(t, (\gamma_s(t), \dot{\gamma}_s(t)))_{vert}, \frac{D}{\partial t}\frac{\partial}{\partial s}\gamma_s(t) \rangle + \langle \nabla_H L(t, (\gamma_s(t), \dot{\gamma}_s(t)))_{hor}, \frac{\partial}{\partial s}\gamma_s(t) \rangle
\end{aligned}
$$

for almost all t by Lemma 3.1.4. Hence

$$
\begin{aligned}
& \left| \frac{\partial L(t, (\gamma_s(t), \dot{\gamma}_s(t)))}{\partial s} \right| \\
\leq & l_1(\gamma_s(t)) \|\dot{\gamma}_s(t)\| \; \|\frac{D}{\partial t}\frac{\partial}{\partial s}\gamma_s(t)\| \; + l_2(\gamma_s(t)) \|\dot{\gamma}_s(t)\|^2 \|\frac{\partial}{\partial s}\gamma_s(t)\| =: g(s,t) \ .
\end{aligned}
$$

For all s the function $g(s, \cdot)$ is in $L^1([0,T])$, as $\|\dot{\gamma}_s(t)\|$ and $\|\frac{D}{\partial t}\frac{\partial}{\partial s}\gamma_s(t)\|$ are in $L^2([0,T])$. Hence $f(s, \cdot)$ is in $L^1([0,T])$.

Now let us prove that $\Theta : I \to L^1([0,T])$, $s \mapsto f(s, \cdot)$ is continuous. [1]

For any sequence (s_n) in I with $s_n \to s_0$ there is a subsequence (also denoted by (s_n)) such that

i) $\dot{\gamma}_{s_n}$ converges a. e. to $\dot{\gamma}_{s_0}$ and $\frac{d}{ds}\dot{\gamma}_s|_{s=s_n}$ converges a. e. to $\frac{d}{ds}\dot{\gamma}_s|_{s=s_0}$.

ii) All $\dot{\gamma}_{s_n}$ and $\frac{d}{ds}\dot{\gamma}_s|_{s=s_n}$ are dominated by some $h \in L^2([0,T])$.

This implies that $f(s_n, \cdot)$ converges almost everywhere to $f(s_0, \cdot)$, dominated by some L^1-function. Therefore $f(s_n, \cdot)$ converges to $f(s_0, \cdot)$ in L^1. We have proved

[1] Please note that in the usual proofs of Gâteaux differentiability that use a *linear* curve in a chart, nothing would be to prove. So that may be the reason to prefer charts, after all.

the continuity of Θ. Thus the Riemann integral of $\Theta(s)$ is defined (s. [HP57] §1 3.3 f, it is equal to the Bochner integal). We can (s. [HP57], §1 3.4) represent $(s,t) \mapsto \Theta(s)(t)$ by a function $\theta(s,t)$ measurable on $I \times [0,T]$, and $\int_0^s \Theta(\sigma)d\sigma$ is for almost everywhere equal to

$$t \mapsto \int_0^s \theta(\sigma, t)d\sigma \ .$$

Furthermore

$$L(t, (\gamma_s(t), \dot\gamma_s(t))) - L(t, (\gamma(t), \dot\gamma(t))) = \int_0^s \theta(\sigma, t)d\sigma \ .$$

almost everywhere.

As $\int_0^s \Theta(\sigma)d\sigma$ is differentiable with respect to s we have

$$\frac{1}{s}\int_0^s \Theta(\sigma)d\sigma \overset{s\to 0}{\to} \Theta(0) \ .$$

Thus

$$\frac{1}{s}(S_L(\gamma_s) - S_L(\gamma)) \overset{s\to 0}{\to} \int_0^T \theta(0,t)dt = \int_0^T f(0,t)dt \ ,$$

which means that the map $s \mapsto S_L(\gamma_s)$ is differentiable and

$$\frac{d}{ds}S_L(\gamma_s)|_{s=0} = \int_0^T \frac{\partial}{\partial s}L(t, (\gamma_s(t), \dot\gamma_s(t)))dt \ ,$$

i.e. S_L is differentiable in the sense of Gâteaux with respect to *any* C^1-chart. (Please note that by chosing a continuous γ_s, differentiable in $s = 0$ only, we might have proved quasi-differentiability/Hadamard-differentiability. As we will obtain the stronger Fréchet-differentiability in the sequel, we were satisfied with the easier notion.)

The Gâteaux derivative DS_L at some $\gamma \in X$ and $u \in T_\gamma X$ is given by

$$DS_L(\gamma)(u) = \int_0^T \langle \nabla_V L(t, (\gamma(t), \dot\gamma(t)))_{vert}, \frac{D}{dt}u(t)\rangle \tag{3.19}$$

$$+ \langle \nabla_H L(t, (\gamma(t), \dot\gamma(t)))_{hor}, u(t)\rangle dt \ . \tag{3.20}$$

As $f(\gamma) := (t \mapsto \nabla_V L(t, (\gamma(t), \dot\gamma(t)))_{vert})$ is locally bounded in $L^2([0,T], \mathbb{R}^N)$, for any sequence $\gamma_n \overset{n\to\infty}{\to} \gamma \in X$ there is a subsequence (also denoted γ_n) such that $f(\gamma_n)$ converges almost everywhere to $f(\gamma)$. By dominated convergence we deduce that $f(\gamma_n) \to f(\gamma)$ in $L^2([0,T], \mathbb{R}^N)$. Thus the map f is continuous as a map from X to $L^2([0,T], \mathbb{R}^N)$.

In quite the same way the local L^1-bound on $g(\gamma) := (t \mapsto \nabla_H L(t, (\gamma(t), \dot\gamma(t)))_{hor})$ yields the continuity of g as a map from X to $L^1([0,T], \mathbb{R}^N)$.

Thus we obtain the continuity of the Gâteaux derivative, which entails the Fréchet differentiability. □

Remark 3.2.5 There are essentially two ways of localising the problem. We can subdivide the interval $[0,T]$ in small invervals I_k as in Benci [Ben86], so that $\gamma|_{I_k}$ is contained in some chart. Or we can chose a smooth family of parametrisations along γ like the charts in [PT88] or [AS09].

A smooth family of local parametrisations $\phi : [0,T] \times U \to M$, $(t,q) \mapsto \phi(t,q) = \phi_t(q)$ defines a local parametrisation Ξ of X, which maps any curve $q \in H^1_{T^*}$ with $q(t) \in U$ $\forall t$ to $\Xi(q) \in X$ with

$$\Xi(q)(t) := \phi(t, q(t)) \ .$$

Now

$$\frac{d}{dt}\Xi(q)(t) = \frac{\partial \phi_t}{\partial t}(q(t)) + D_q\phi(t, q(t))q'(t) \ .$$

We get the local representation of S_L as $\tilde S_L := S_L \circ \Xi$ as follows:

$$\begin{aligned}
&\tilde S_L(q)\\
&= \int_0^T L\left(t, (\Xi(q)(t), \frac{d}{dt}\Xi(q)(t))\right) dt\\
&= \int_0^T L\left(t, (\phi(t,q), D_q\phi(t, q'(t)) + \frac{\partial \phi_t}{\partial t}(q(t)))\right) dt\\
&= \int_0^T L\left(t, \left(\phi(t,q), D_q\phi(t,q(t))\left(q'(t) + D_q\phi(t,q(t))^{-1}\left(\frac{\partial \phi_t}{\partial t}(q(t))\right)\right)\right)\right) dt\\
&= \int_0^T L\left(t, T\phi_t\left(q, q'(t) + D_q\phi(t,q(t))^{-1}\left(\frac{\partial \phi_t}{\partial t}(q(t))\right)\right)\right) dt
\end{aligned}$$

We write

$$\hat L(t,q,v) := L\left(t, T\phi_t\left(q, v + D_q\phi(t,q)^{-1}(\frac{\partial \phi_t}{\partial t}(q))\right)\right) \ ,$$

so that

$$\tilde S_L(q) = \int_0^T \hat L(t, q(t), q'(t)) dt \ .$$

Abbondandolo in [AS09] incorrectly deals with $\tilde{L}$ instead of $\hat{L}$. This does not really matter, as the conditions on $\tilde{L}$ correspond to the same form of conditions on $\hat{L}$, see below.

Please note that $\tilde{L}(t,q,v) := L(t, T\phi_t(q,v))$ is related to $\hat{L}$ by

$$\hat{L}(t,q,v) = \tilde{L}(t,q,v+(d\phi_t)^{-1}(\frac{\partial \phi_t}{\partial t}(t,q))) \ . \tag{3.21}$$

The local condition (SQ)'$_{loc}$ for $\tilde{L}$ implies a condition of the same form for $\hat{L}$. Let us check the condition for D_q. In order to simplify the calculation set $g(t,q) := (d\phi_t)^{-1}(\frac{\partial \phi_t}{\partial t}(t,q))$.

$$\begin{aligned}
&\|D_q\hat{L}(t,q,v)\| \\
=&\|D_q\tilde{L}(t,q,v+g(t,q)) + D_v\tilde{L}(t,q,v+g(t,q))D_q g(t,q)\| \\
\le& l_2(q)(1+\|v+g(t,q)\|^2) + l_1(q)(1+\|v+g(t,q)\|)\|D_q g(t,q)\| \\
\le& l_2(q)2(1+\|g(t,q)\|^2)(1+\|v\|^2) + l_1(q)\|D_q g(t,q)\|(1+\|g(t,q)\|)(1+\|v\|) \\
\le& l_2(q)2(1+\|g(t,q)\|^2)(1+\|v\|^2) + l_1(q)\|D_q g(t,q)\|(1+\|g(t,q)\|)2(1+\|v\|^2) \\
\le& \tilde{l}_2(q)(1+\|v\|^2) \ .
\end{aligned}$$

(We used $1+\|a+b\|^2 \le 2(1+\|a\|^2)(1+\|b\|^2)$, $1+\|a+b\| \le (1+\|a\|)(1+\|b\|)$ and $1+\|a\| \le 2(1+\|a\|^2)$). In a similar way we obtain $\|D_v\hat{L}(t,q,v)\| \le \tilde{l}_1(q)(1+\|v\|)$.

A curve γ_s is represented in these local coordinates by a curve $\tilde{\gamma}_s$ in $H^1_{T^*}(\mathbb{R}^n)$ with $\frac{d}{ds}\tilde{\gamma}|_{s=0} = u$ and $\tilde{\gamma}_0 = q$. If we repeat the proof with partial derivatives D_q and D_v instead of ∇_H and ∇_V we obtain that $\tilde{S}_L$ is differentiable in the sense of Fréchet with derivative

$$D\tilde{S}_L|_q(u) = \int_0^L (D_q\hat{L}(t,q(t),\dot{q}(t))(u(t)) + D_v\hat{L}(t,q(t),\dot{q}(t))(\dot{u}(t)))dt \tag{3.22}$$

Now we discuss further conditions that imply higher differentiability - in some sense - of S_L.

(TCD) $L(t,\cdot)$ is twice continuously differentiable for almost every $t \in \mathbb{R}$.

(SQ")$_{loc}$

$$\begin{aligned}
\|D_{qq}\tilde{L}(t,q,v)\| &\le l_1(q)\|v\|^2 \ , \\
\|D_{qv}\tilde{L}(t,q,v)\| &\le l_2(q)\|v\| \ , \\
\|D_{vv}\tilde{L}(t,q,v)\| &\le l_3(q)
\end{aligned}$$

with continuous functions $l_1, l_2, l_3 : U \to \mathbb{R}$ for any smooth family of parametrisation $\phi_t : U \to M$ and $\tilde{L}(t,q,v) := L(t, T\phi_t(q,v))$.

An analogue of Proposition 3.2.4 for the second derivative fails. If we impose (TCD) and (SQ”)$_{loc}$, the functional is not C^2 in general. (We refrain from giving a coordinate-free version of this condition, which might only be given in terms of the covariant Hessian of L, as the ordinary Hessian is well-defined at critical points only.)

The problem stems from the well-known fact that Nemitski-operators to L^∞ are not continuous in general. Yet the claim that the functional is C^2 under these conditions, occurs now and then in the literature, e.g. in [AF07], [AS06], and [Lu09]. Abbondandolo corrected the statement in [AS09], actually proving that the functional is C^2, if and only if it is an at most quadratic polynomial along the fibres. (Such Lagrangians are known als *electromagnetical*).

(Q) *(quadratic)* There are C^2-sections A of T^*M and B of $Hom(TM,TM)$, such that

$$L(t,(q,v)) = A(q)(v) + \langle v, B(q)(v)\rangle \ .$$

or locally

(Q)$_{loc}$ There are C^2-functions $A : U \to \mathbb{R}^n$ and $B : U \to \mathbb{R}^{n\times n}$, such that

$$\tilde{L}(t,q,v) = \langle A(q), v\rangle + \langle v, B(q)v\rangle \ .$$

In both cases we can assume that $B(q)$ is symmetric.

This condition is equivalent to

(Q”) $D_{VV}L(t,(q,v))$ is constant for t, q fixed.

or

(Q”)$_{loc}$ $D_{vv}\tilde{L}(t,q,v)$ is independent of v.

Proposition 3.2.6 *If L satisfies (TP),(CF),(TCD) and (SQ”$_{loc}$), the functional S_L is C^1 and its derivative is locally Lipschitz continuous and differentiable in the sense of Hadamard. With the local expressions of Remark 3.2.5 we have:*

$$\begin{aligned} D(D\tilde{S}_L)(q)(u)(v) &= D^2\tilde{S}_L(q)(u,v) \\ &= \int_0^T [D_{qq}\hat{L}(t,q(t),\dot{q}(t))(u(t),v(t)) \\ &\qquad + D_{vq}\hat{L}(t,q(t),\dot{q}(t))(\dot{u}(t),v(t)) \\ &\qquad + D_{qv}\hat{L}(t,q(t),\dot{q}(t))(u(t),\dot{v}(t)) \\ &\qquad + D_{vv}\hat{L}(t,q(t),\dot{q}(t))(\dot{u}(t),\dot{v}(t))]dt \ . \end{aligned} \tag{3.23}$$

$D^2\tilde{S}_L(q)$ is a symmetric bilinear form on $H^1_{T^*}(S^1,\mathbb{R}^n)^2$.

Whenever $j : Y \to X$ is a compact embedding of a Banach manifold Y, the composition $S_L \circ j$ is C^2.

If L satisfies additionally (Q), S_L is C^2.

Proof: As in Remark 3.2.5 the growth conditions (SQ")$_{loc}$ imply the same form of conditions for $\hat{L}$. Furthermore (SQ")$_{loc}$ entails (SQ')$_{loc}$, as we see by integrating $D_{qv}\tilde{L}$ and $D_{vv}\tilde{L}$ along the fibres. Hence the functional is C^1 and the derivative is given by (3.22).

Let q_s be a continuous curve in the domain of definition U of our local parametrisation, differentiable at $s = 0$ with $\frac{d}{ds}q_s = u \in H^1_{T^*}(S^1,\mathbb{R}^n)$ and $q_0 = q$.

We want to express the difference quotients

$$A_s := \frac{1}{s}(D\tilde{S}_L(q_s) - D\tilde{S}_L q_0) \in (H^1_{T^*}(S^1,\mathbb{R}^n))^* \ ,$$

by means of integrals of the second derivatives of $\hat{L}$. For any $v \in H^1_{T^*}(S^1,\mathbb{R}^n)$ we have

$$\begin{aligned} A_s(v) = \frac{1}{s}\int_0^T &[D_q\hat{L}(t,q_s(t),\dot{q}_s(t))(v(t)) + D_v\hat{L}(t,q_s(t),\dot{q}_s(t))(\dot{v}(t)) \\ &- D_q\hat{L}(t,q(t),\dot{q}(t))(v(t)) - D_v\hat{L}(t,q(t),\dot{q}(t))(\dot{v}(t))]dt \end{aligned}$$

Furthermore

$$\begin{aligned} &D_q\hat{L}(t,q_s(t),\dot{q}_s(t))(v(t)) - D_q\hat{L}(t,q(t),\dot{q}(t))(v(t)) \\ &= \int_0^1 D_{qq}\hat{L}(t,(1-r)q(t)+rq_s(t),(1-r)\dot{q}(t)+r\dot{q}_s(t))(q_s(t)-q(t))(v(t))dr \\ &\quad + \int_0^1 D_{vq}\hat{L}(t,(1-r)q(t)+rq_s(t),(1-r)\dot{q}(t)+r\dot{q}_s(t))(\dot{q}_s(t)-\dot{q}(t))(v(t))dr \end{aligned}$$

and

$$\begin{aligned} &D_v\hat{L}(t,q_s(t),\dot{q}_s(t))(\dot{v}(t)) - D_v\hat{L}(t,q(t),\dot{q}(t))(\dot{v}(t)) \\ &= \int_0^1 D_{qv}\hat{L}(t,(1-r)q(t)+rq_s(t),(1-r)\dot{q}(t)+r\dot{q}_s(t))(q_s(t)-q(t))(\dot{v}(t))dr \\ &\quad + \int_0^1 D_{vv}\hat{L}(t,(1-r)q(t)+rq_s(t),(1-r)\dot{q}(t)+r\dot{q}_s(t))(\dot{q}_s(t)-\dot{q}(t))(\dot{v}(t))dr]dt \ . \end{aligned}$$

As q_s is continuous with respect to the C^0-topology the $q_s(t)$ are contained in a compact subset of our chart of M, on which the functions l_i from condition (SQ")$_{loc}$ are bounded by some constant. Therefore by (SQ")$_{loc}$ almost everywhere

$$|\frac{1}{s}\int_0^1 D_{qq}\hat{L}(t,(1-r)q(t)+rq_s(t),(1-r)\dot{q}(t)+r\dot{q}_s(t))(q_s(t)-q(t))(v(t))dr| \tag{3.24}$$

$$\leq C(\|\dot{q}(t)\|+\|\dot{q}_s(t)\|)^2\left\|\frac{q_s(t)-q(t)}{s}\right\|\|v(t)\|\ , \tag{3.25}$$

$$|\frac{1}{s}\int_0^1 D_{vq}\hat{L}(t,(1-r)q(t)+rq_s(t),(1-r)\dot{q}(t)+r\dot{q}_s(t))(\dot{q}_s(t)-\dot{q}(t))(v(t))dr| \tag{3.26}$$

$$\leq C(\|\dot{q}(t)\|+\|\dot{q}_s(t)\|)\left\|\frac{\dot{q}_s(t)-\dot{q}(t)}{s}\right\|\|v(t)\|\ , \tag{3.27}$$

$$|\frac{1}{s}\int_0^1 D_{qv}\hat{L}(t,(1-r)q(t)+rq_s(t),(1-r)\dot{q}(t)+r\dot{q}_s(t))(q_s(t)-q(t))(\dot{u}(t))| \tag{3.28}$$

$$\leq C\left(\|\dot{q}(t)\|+\|\dot{q}_s(t)\|\left\|\frac{q_s(t)-q(t)}{s}\right\|\right)\|\dot{v}(t)\|\ , \tag{3.29}$$

$$|\frac{1}{s}\int_0^1 D_{vv}\hat{L}(t,(1-r)q(t)+rq_s(t),(1-r)\dot{q}(t)+r\dot{q}_s(t))(\dot{q}_s(t)-\dot{q}(t))(\dot{v}(t))dr| \tag{3.30}$$

$$\leq C\left\|\frac{\dot{q}_s(t)-\dot{q}(t)}{s}\right\|\|\dot{v}(t)\|\ . \tag{3.31}$$

Recall that $\dot{q}_s \to \dot{q}$ in $L^2([0,T],\mathbb{R}^n)$ and $\frac{\dot{q}_s-\dot{q}}{s} \to \dot{u}$ in $L^2([0,T],\mathbb{R}^n)$. Hence for any sequence (s_n) there is a subsequence, also denoted by (s_n), such that

1. there is a is a function $h \in L^2([0,T],\mathbb{R})$, which dominates $t \mapsto \|\dot{q}_{s_n}(t)\|$ and $t \mapsto \frac{\dot{q}_s(t)-\dot{q}(t)}{s}$ almost everywhere,

2. $\dot{q}_{s_n}$ converges almost everywhere to $\dot{q}$ and $\frac{\dot{q}_s(t)-\dot{q}(t)}{s}$ converges almost everywhere to $\dot{u}$.

By Sobolev's embedding $\frac{q_s-q}{s} \to u$ in $C^0_T(S^1,\mathbb{R}^n)$ and $\left\|\frac{q_{s_n}-q}{s}\right\|_{C^0}$ is bounded.

Hence

$$\frac{1}{s_n}\int_0^{s_n} D_{qq}\hat{L}(t,(1-r)q(t)+rq_{s_n}(t),(1-r)\dot{q}(t)+r\dot{q}_{s_n}(t))(q_{s_n}(t)-q(t))dr$$

and

$$\frac{1}{s_n}\int_0^{s_n} D_{vq}\hat{L}(t,(1-r)q(t)+rq_{s_n}(t),(1-r)\dot{q}(t)+r\dot{q}_{s_n}(t))(\dot{q}_{s_n}(t)-\dot{q}(t))$$

are dominated in $L^1([0,T],(\mathbb{R}^n)')$,

$$\frac{1}{s_n}\int_0^{s_n} D_{qv}\hat{L}(t,(1-r)q(t)+rq_{s_n}(t),(1-r)\dot{q}(t)+r\dot{q}_{s_n}(t))(q_{s_n}(t)-q(t))$$

and

$$\frac{1}{s_n}\int_0^{s_n} D_{vv}\hat{L}(t,(1-r)q(t)+rq_{s_n}(t),(1-r)\dot{q}(t)+r\dot{q}_{s_n}(t))(\dot{q}_{s_n}(t)-\dot{q}(t))dr$$

are dominated in $L^2([0,T],(\mathbb{R}^n)')$. By Lebesgue's theorem these sequences converge in L^1 and L^2, respectively.

Thus A_{s_n} converges in $(H^1_{T^*}(S^1,\mathbb{R}^n))'$. As this is the case for a subsequence of any sequence (s_n) that converges to 0, we have proved that $D\tilde{S}_L$ is differentiable in the sense of Hadamard at q, and the derivative is given by equation (3.23). In particular DS_L is differentiable in the sense of Gâteaux with respect to any chart.

By the inequalities (3.25)- (3.31), there is a neighbourhood U of q and a constant $L>0$ such that for all $\tilde{q}\in U$ we have

$$\|D\tilde{S}_L(\tilde{q})-D\tilde{S}_L(q)\|_{(H^1)'}\leq L\|\tilde{q}-q\|_{H^1}\ ,$$

i. e. local Lipschitz continuity of $D\tilde{S}_L$.

By Remark 3.2.3 Hadamard differentiability implies that $DS_L\circ j$ is differentiable in the sense of Fréchet, whenever $j:Y\to X$ is a smooth and sequentially compact embedding of a Banach manifold Y. In particular this is true for the embedding of finite dimensional submanifolds. By applying Schwarz' theorem to embedded surfaces we obtain the symmetry of $D^2\tilde{S}_L$.

In order to prove the stronger statement that $DS_L\circ j$ is continuously Fréchet-differentiable we observe that the bilinear form $D^2\tilde{S}_L(q)(u,v)=B_q(u,v)+C_q(u,v)$ with

$$\begin{aligned}B_q(u,v):=\int_0^T &[D_{qq}\hat{L}(t,q(t),\dot{q}(t))(u(t),v(t))\\ &+D_{vq}\hat{L}(t,q(t),\dot{q}(t))(\dot{u}(t),v(t))\\ &+D_{qv}\hat{L}(t,q(t),\dot{q}(t))(u(t),\dot{v}(t))]dt\end{aligned}$$

and

$$C_q(u,v) := \int_0^T [D_{vv}\hat{L}(t,q(t),\dot{q}(t))(\dot{u}(t),\dot{v}(t))]dt$$

Now we observe two facts:

1. $q \mapsto B_q$ defines a continuous map from $U \subset H^1 := H^1_{T^*}(S^1,\mathbb{R}^n)$ to $\mathrm{Bilin}(H^1,H^1)$ with respect to the norm topology:

   For any sequence $q_n \to q$ in U we can find a subsequence such that $(\dot{q}_n)$ is dominated by some $h \in L^2([0,T],\mathbb{R})$ and $(\dot{q}_n)$ converges almost everywhere to q. We can assume that the $q_n(t)$ are contained in a compact subset of M. We define the norm of $\Phi \in \mathrm{Bilin}(\mathbb{R}^n,\mathbb{R}^n)$ as

   $$\|\Phi\| := \max_{x,y\neq 0} \frac{\|\Phi(x,y)\|}{\|x\|\,\|y\|} \ .$$

   Now

   $$\|D_{qq}\hat{L}(t,q_n(t),\dot{q_n}(t)) - D_{qq}\hat{L}(t,q(t),\dot{q}(t))\|| \leq Ch(t)^2 \ ,$$
   $$\|D_{qv}\hat{L}(t,q_n(t),\dot{q_n}(t)) - D_{qv}\hat{L}(t,q(t),\dot{q}(t))\| \leq Ch(t) \ ,$$
   $$\|D_{vq}\hat{L}(t,q_n(t),\dot{q_n}(t)) - D_{vq}\hat{L}(t,q(t),\dot{q}(t))\| \leq Ch(t)$$

   for some constant $C > 0$.
   Hence $D_{qq}\hat{L}(t,q_n(t),\dot{q}_n(t))$ converges in $L^2([0,T],\mathrm{Bilin}(\mathbb{R}^n,\mathbb{R}^n))$, $D_{qv}\hat{L}(t,q_n(t),\dot{q}_n(t))$ and $D_{vq}\hat{L}(t,q_n(t),\dot{q}_n(t))$ converge in $L^2([0,T],\mathrm{Bilin}(\mathbb{R}^n,\mathbb{R}^n))$, which proves that $B_{q_n} \to B_q$ in $\mathrm{Bilin}(H^1,H^1)$, and we are done.

   Please note that the bilinear form B_q is represented by a compact operator $H^1 \to H^1$, as the embedding $H^1 \to L^2$ and its adjoint (Schauder theorem) are compact.

2. The map $q \mapsto C_q$ is *not continuous* in general, as almost everywhere convergence of an L^∞-dominated sequence does *not* imply its L^∞-convergence. Dominated convergence yields, however, that for any $u \in H^1$ the map $U \to (H^1)^*$, $q \mapsto C_q(u,\cdot)$ is continuous. (Equivalently we can consider the map $U \to L(H^1,H^1)$, $q \mapsto J(C_q(u,\cdot))$, where $J : X' \to X$ is the inverse of the Riesz representation. The last statement means that this map is continuous in the strong operator topology.)

   For a compact C^2-map $j : V \to U$ from some open subset V of a Banach space B, we already know by Hadamard-differentiability that $\tilde{S}_L \circ j$ is twice

Fréchet-differentiable. The second Fréchet-derivative is given by

$$\begin{aligned} D^2(\tilde{S}_L \circ j)(q)(u,v) &= D(D\tilde{S}_L|_{j(q)} Dj|_q)(u)(v) \\ &= D^2\tilde{S}_L|_{j(q)}(Dj|_q(u), Dj|_q(v)) + D\tilde{S}_L|_{j(q)}(D^2 j|_q(u,v)) \; . \end{aligned}$$

The second summand is a bilinear form on $B \times B$ that depends continuously on $q \in V$, anyway. In order to prove that the bilinear form defined by the first summand depends continuously on q, it is sufficient to prove that the form

$$\tilde{C}_q(\cdot,\cdot) := C_{j(q)}(Dj|_q(\cdot), Dj|_q(\cdot)) \in \mathrm{Bilin}(B,B)$$

depends continuously on q.

Now suppose $q_n \to q = q_0$ in B, therefore $j(q_n) \to j(q)$ in $U \subset H^1$. Let $g_n \in L^\infty([0,T])$ be defined by

$$g_n(t) := D_{vv}\hat{L}(t, j(q_n)(t), \frac{d}{dt}(j(q_n)(t))) \; .$$

(g_n) is bounded in L^∞ by some constant $C > 0$. It follows that for any $u, v \in H^1$

$$|C_{j(q_n)}(u,v)| \leq C\|u\|_{H^1}\|v\|_{H^1} \; .$$

Now we argue by contradiction. Suppose $\tilde{C}_{q_n}$ does not converge to $\tilde{C}_q$. Then, there is an $\epsilon > 0$ such that

$$\|\tilde{C}_{q_n} - \tilde{C}_q\| > 2\epsilon \quad \text{for almost all } n \in \mathbb{N} \; . \tag{3.32}$$

Hence there are sequences $u_n \in B$, $v_n \in B$, $\|u_n\| = \|v_n\| = 1$ with $|\tilde{C}_{q_n}(u_n,v_n) - \tilde{C}_q(u_n,v_n)| > \epsilon$ for almost all $n \in \mathbb{N}$. As j is a compact map its derivative Dj_q is compact as well, and thus there is a strictly monotonically increasing sequence $(n_i)_{i\in\mathbb{N}}$, such that $Dj|_q(u_{n_i}) \to a \in H^1$ and $Dj|_q(v_{n_i}) \to b \in H^1$ and there is a constant $D > 0$ with $\|Dj|_q(u_{n_i})\| \leq D$ and $\|Dj|_q(v_{n_i})\| \leq D$.

As $C_{q_n}(a,b) \overset{n\to\infty}{\to} C_q(a,b)$ there is an $i_0 \in \mathbb{N}$ such that

$$|C_{q_{n_i}}(a,b) - C_q(a,b)| < \frac{\epsilon}{3} \; , \; \|Dj|_q(u_{n_i}) - a\| < \frac{\epsilon}{6CD} \; , \; \|Dj|_q(v_{n_i}) - b\| < \frac{\epsilon}{6CD}$$

for all $i \geq i_0$. Therefore

$$\begin{aligned} &|C_{q_{n_i}}(Dj|_{q_{n_i}}(u_{n_i}), Dj|_{q_{n_i}}(v_{n_i})) - C_{q_{n_i}}(a,b)| \\ \leq &|C_{q_{n_i}}(Dj|_{q_{n_i}}(u_{n_i}), Dj|_{q_{n_i}}(v_{n_i})) - C_{q_{n_i}}(Dj|_{q_{n_i}}(u_{n_i}), b)| \\ &\quad + |C_{q_{n_i}}(Dj|_{q_{n_i}}(u_{n_i}), b) - C_{q_{n_i}}(a,b)| \\ \leq &CD\frac{\epsilon}{6CD} + CD\frac{\epsilon}{6CD} = \frac{\epsilon}{3} \end{aligned}$$

and

$$\begin{aligned}
&|C_q(a,b) - C_q(Dj|_q(u_{n_i}), Dj|_q(v_{n_i}))| \\
\le &|C_q(a,b) - C_q(Dj|_q(u_{n_i}), b)| + |C_q(Dj|_q(u_{n_i}), b) - C_q(Dj|_q(u_{n_i}), Dj|_q(v_{n_i}))| \\
\le &CD\frac{\epsilon}{6CD} + CD\frac{\epsilon}{6CD} = \frac{\epsilon}{3} \ .
\end{aligned}$$

We estimate

$$\begin{aligned}
&|\tilde{C}_{q_{n_i}}(u_{n_i}, v_{n_i}) - \tilde{C}_q(u_{n_i}, v_{n_i})| \\
= &|C_{q_{n_i}}(Dj|_{q_{n_i}}(u_{n_i}), Dj|_{q_{n_i}}(v_{n_i})) - C_q(Dj|_q(u_{n_i}), Dj|_q(v_{n_i}))| \\
\le &|C_{q_{n_i}}(Dj|_{q_{n_i}}(u_{n_i}), Dj|_{q_{n_i}}(v_{n_i})) - C_{q_{n_i}}(a,b)| \\
&\quad + |C_{q_{n_i}}(a,b) - C_q(a,b)| + |C_q(a,b) - C_q(Dj|_q(u_{n_i}), Dj|_q(v_{n_i}))| \\
\le &\frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon \ ,
\end{aligned}$$

which contradicts (3.32).

All of the above applies under the stronger condition (Q). Now we prove that under condition (Q) the functional S_L is C^2. It is sufficient to prove that $q \mapsto C_q$ is continuous. Now

$$\begin{aligned}
C_q(u,v) &= \int_0^T D_{vv}\hat{L}(t, q(t), \dot{q}(t))(\dot{u}(t), \dot{v}(t))dt \\
&= \int_0^T \langle \dot{u}(t), M(q(t))\dot{v}(t)\rangle dt \ ,
\end{aligned}$$

as $D_{vv}\tilde{L}(t,q,v) = D_{vv}\hat{L}(t,q,v) = \langle \cdot, M(q)\cdot\rangle$.

Any sequence $q_n \to q$ in H^1 converges in C^0, hence the sequence of matrix valued functions $(t \mapsto M(q_n(t)))_{n\in\mathbb{N}}$ converges in C^0, and C_q converges in $\mathrm{Bilin}(H^1, H^1)$. $\square$

Remark 3.2.7 It is not difficult along these lines to prove that S_L is C^2 if and and only if (Q) is satisfied. Let us give a proof of the *only-if*-part as an alternative to Abbondandolo's proof, which is quite different:

Suppose (Q) is not satisfied. Then there is a $q \in M$ and $v_1 \in T_qM$, $v_2 = 0_q \in T_qM$ such that (in a local chart given by $\phi : U \to M$, $T\phi : U \times \mathbb{R}^n \to TM$) $D_{vv}\hat{L}(t,q,v_1) \neq D_{vv}\hat{L}(t,q,v_2)$. As L is C^2, there are $c_1 > 0$, a vector $u \in \mathbb{R}^n$ and neighbourhoods U_i of (q, v_i) in $U \times \mathbb{R}^n$ such that for all $(q_i, w_i) \in U_i$:

$$|\langle u, (D_{vv}\hat{L}(t,q_1,w_1) - D_{vv}\hat{L}(t,q_2,w_2))u\rangle| > c_1\|u\|^2 \ .$$

Without restriction of generality we can assume $q = 0$ and $\langle u, (D_{vv}\hat{L}(t, q_1, w_1) - D_{vv}\hat{L}(t, q_2, w_2))u\rangle > 0$ and $\|u\| = 1$.

Denote by $\gamma \in H^1$ the constant curve with $\gamma(t) = 0$. It remains to construct a sequence of H^1-curves $\gamma_n \to \gamma$ and $u_n \in H^1$ such that $\|C_{\gamma_n}(u_n, u_n) - C_\gamma(u_n, u_n)\| \geq c_2\|u_n\|^2_{H^1}$ for some constant $c_2 > 0$.

We define (for $n > 1$)

$$f_n : [0, T] \to \mathbb{R},\ x \mapsto f_n(x) = \begin{cases} x, & \text{if } x \leq 1/n \\ \frac{1}{n-1}(1-x), & \text{if } 1/n < x \leq 1 \\ 0, & \text{else} \end{cases}$$

and

$$\gamma_n : [0, T] \to \mathbb{R}^n\ , \quad \gamma_n(t) = f_n(t)v_1$$

Then $\gamma_n \in H^1$, $\gamma_n \to \gamma$ in H^1 and $\dot{\gamma}_n(t) = v_1$ for all $t \in [0, 1/n[$. For n sufficiently large $(\gamma_n(t), \dot{\gamma}_n(t)) \in U_1$ for all $t \in]0, 1/n[$. Now let $\phi : \mathbb{R} \to \mathbb{R}$ a C^∞-function with support $[0, 1]$, and suppose $\int_0^T \phi^2(t)dt = a > 0$ and $\int_0^T (\phi')^2(t)dt = b > 0$. We define

$$u_n : [0, T] \to \mathbb{R}^n,\ t \mapsto u_n(t) = \frac{1}{\sqrt{n}}\phi(nt)u$$

and compute for $n > 1/T$

$$\begin{aligned} \|u_n\|_{H^1} &= \sqrt{\int_{\mathbb{R}} [\frac{1}{n}\phi^2(nt) + n(\phi')^2(nt)]dt} \\ &= \sqrt{\frac{1}{n^2}a + b}\ . \end{aligned}$$

We observe, that this norm converges to $\sqrt{b}$ for $n \to \infty$.

$$\begin{aligned} &C_{\gamma_n}(u_n, u_n) - C_\gamma(u_n, u_n) \\ &= \int_0^T \langle \dot{u}_n(t), (D_{vv}\hat{L}(t, \gamma_n(t), \dot{\gamma}_n(t)) - D_{vv}\hat{L}(t, 0, 0))\dot{u}_n(t)\rangle > dt \\ &\geq \int_0^{1/n} c_1\|\dot{u}_n(t)\|^2 dt \\ &= c_1 b \geq \frac{c_1}{2}\|u_n\|^2_{H^1}; \end{aligned}$$

for n sufficiently large.

The construction gives even more information, because it shows that the second derivative is not continuous at the constant solution $\gamma(t) = q$, if L is not quadratic on the fibre over q. ◇

For regularity we assume that

(REG) L is C^2, and the map

$$\Delta : \mathbb{R} \times TM \to \mathbb{R} \times TM, \quad (t,(q,v)) \mapsto (t,(q,\nabla_V L(t,(q,v))_{vert}))$$

is injective with C^1 inverse, and there are continuous functions $l_4, l_5 : M \to \mathbb{R}^+$ and $c > 0$ such that

$$\|\nabla_V L(t,(q,v))_{vert}\| \geq l_4(q)\|v\| - l_5(q) \ .$$

Please note that Δ is C^{k-1} whenever L is C^k.

(REG) is satisfied under the stronger condition

(CON) *(convex)* L is C^2 and there is a function $l_6 : M \to \mathbb{R}^+$ such that

$$D_{VV}L(t,(q,v))(w,w) \geq l_6(q)\|w\|_q^2 \ .$$

or, equivalently, its local version

(CON)$_{loc}$ L is C^2 and for every smooth chart $M \supset V \overset{\phi^{-1}}{\to} U \subset \mathbb{R}^n$ there is a function $l_6 : U \to \mathbb{R}^+$ such that

$$D_{vv}\tilde{L}(t,q,v)(w,w) \geq l_6(q)\|w\|^2 \ .$$

Please recall that $D_{VV}L$ denotes the second vertical derivative as defined in Section 3.1.1.

We should verify the claim that (CON) implies (REG).

Lemma 3.2.8 *(CON) implies (REG).*

Proof: We recall that

$$\langle \nabla_V L(t,(q,v))_{vert}, w\rangle_q = D_V L(t,(q,v))(w) \ .$$

Suppose (CON) is satisfied. Then for $q \in M$, $v,w \in T_qM$:

$$\begin{aligned}
|D_V L(t,(q,v))(v)| &= |D_v\tilde{L}(t,(q,0))(v) + \int_0^1 D_{VV}L(t,(q,sv))(v,v)ds| \\
&\geq -|D_V L(t,(q,0))(v)| + l_6(q)\|v\|_q^2 \\
&\geq -l_7(q)\|v\|_q + l_6(q)\|v\|_q^2 \ .
\end{aligned}$$

Hence

$$\begin{aligned}\|\nabla_V L(t,(q,v))_{vert}\|_q &= \sup_{w\neq 0}\frac{|D_V L(t,(q,v))(w)|}{\|w\|_q}\\ &\geq -l_7(q)+l_6(q)\|v\| \ .\end{aligned}$$

Thus we obtain (REG) with $l_5 = l_7$ and $l_4 = l_6$.

Injectivity follows from integration along the fibre and the differentiability of the inverse from the inverse function theorem. □

Proposition 3.2.9 *If L satisfies (TP), (SQ') and (REG), the critical points of S_L are precisely the T-periodic C^2 solutions of*

$$\nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor} - \frac{D}{dt}\nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert} = 0 \ . \tag{3.33}$$

In local coordinates as above they correspond to the smooth solutions of

$$D_q\hat L(t,\gamma(t),\dot\gamma(t)) - \frac{d}{dt}D_v\hat L(t,\gamma(t),\dot\gamma(t)) = 0 \ . \tag{3.34}$$

If L is C^k the solutions are C^k.

Proof: A proof in local charts can be found e. g. in [AS09], Prop 2.2.

In the appendix we provide an alternative proof, which largely abstains from charts. For vector fields along an H^1-curve $\gamma \in X$ the covariant derivative is the natural notion of derivative. By developing a few tools (which require charts in their definition) like "covariant primitive" and "covariant integration by parts" we can mimic the usual bootstrap arguments without referring to charts. We give this proof because these tools might be of independent interest. □

Remark 3.2.10 Please note that the condition (REG) allows to obtain the regularity of critical points of S_L and an explicit second order system for the critical points of S_L. (REG) would be satisfied for a $L(t,(q,v)) = h_q(v,v)$ with a semi-Riemannian, e.g. Lorentzian, metric h_q. The finiteness of the index, however, needs a stronger condition, e.g. (CON)$_{loc}$. In the Lorentzian example all critical points have infinite index, the problem is strongly indefinite. ◇

The following proposition and the following lemma are based on Abbondandolo and Schwarz ([AS09]).

Proposition 3.2.11 *Suppose L satisfies (TP), (SQ") and (CON). Then at a critical point γ of S_L the second derivative of $\tilde S_L$ in the sense of 3.2.6 defines* *a bilinear form on $T_\gamma X \times T_\gamma X$ which is represented by a symmetric Fredholm operator H_γ. Index and nullity of H_γ are finite.*

In a local chart

$$X \supset V \to U \subset H^1_{T^*}(S^1, \mathbb{R}^n) =: H^1$$

the second Gâteaux derivative $D^2\tilde{S}_L(q)$ is defined for all $q \in U$. With respect to $H^1_{T^*}(S^1, \mathbb{R}^n)$, it is represented by a self-adjoint Fredholm operator $\tilde{H}_q$, which allows a decomposition into self-adjoint operators

$$\tilde{H}_q = A(q) + K(q) \ ,$$

with the following properties:

i) $K(q)$ is compact for every $q \in U$, and the map $q \to K(q)$ is continuous with respect to the norm topology.

ii) $A(q)$ is invertible for every $q \in U$, the map $q \mapsto A(q)$ is strongly continuous, *i.e. for every $v \in H^1$ the map $q \mapsto A(q)v$ is continuous.*

iii) The maps $q \mapsto \|A(q)\|$ and $q \mapsto \|A(q)^{-1}\|$ are locally uniformly bounded.

If G acts on G by isometries and γ_0 is a constant orbit with value a fixed point of G, there is an induced orthogonal G-operation on $T_{\gamma_0}X \cong H^1$, and the chart can be chosen G-equivariant (Remark 3.1.1), so that it maps γ_0 to $0 \in H^1$. $\tilde{H}_0$, $A(0)$ and $K(0)$ are G-equivariant.

Proof: We know from 3.2.6 that $D^2\tilde{S}_L$ (= the Hadamard derivative of the Fréchet derivative of $\tilde{S}_L$) is a symmetric bilinear form on the domain of our parametrisation. We should verify in which sense this defines a "Hessian" H_γ on $T_\gamma X$ at critical points.

Suppose we have a functional f on a C^2 Riemannian Hilbert manifold X and local parametrisations $\phi_i : U_i \to V_i \subset X$, $\tilde{f}_i := f \circ \phi_i$, $i = 1, 2$. Suppose $V := V_1 \cap V_2$ is non-empty and

$$\psi := \phi_2^{-1} \circ \phi_1|_{\phi_1^{-1}(V)} : \phi_1^{-1}(V) \to \phi_2^{-1}(V)$$

the change of charts. We have

$$\tilde{f}_1|_{\phi_1^{-1}(V)} = \tilde{f}_2 \circ \psi \ .$$

Differentiation at $y \in \phi_1^{-1}(D)$ yields

$$D\tilde{f}_1|_y(u) = D\tilde{f}_2|_{\psi(y)}(D\psi|_y(u)).$$

Suppose $\phi_1(y) = x$ is a critical point of f, and $\tilde{f}$ has a Hadamard differentiable Fréchet derivative. Hadamard differentiability satisfies a chain rule (s. [Yam74], (1.2.9)), hence

$$\begin{aligned} D^2\tilde{f}_1|_y(u,v) &= D^2\tilde{f}_2|_{\psi(y)}(D\psi|_y(u), D\psi|_y(v)) \\ &\quad + D\tilde{f}_2|_{\psi(y)}(D^2\psi|_y(u,v)) \\ &= D^2\tilde{f}_2|_{\psi(y)}(D\psi|_y(u), D\psi|_y(v)) \; . \end{aligned} \tag{3.35}$$

This expression defines a bilinear form $D^2 f$ on $T_xX \times T_xX$ as follows

$$D^2 f|_x(u,v) := D^2\tilde{f}_1|_y(D\phi_1|_y^{-1}(u), D\phi_1|_y^{-1}(v)) \; ,$$

which by (3.35) is independent of the parametrisation chosen. Its representation with respect to the Riemannian Hilbert structure on T_xX is the Hessian H_x of f at x.

With the notations from the proof of 3.2.6, we write the bilinear form as $D^2\tilde{S}_L|_q(u,v) = B_q(u,v) + C_q(u,v)$. B_q is representable by a compact operator, whereas C_q can be written as

$$\begin{aligned} C_q(u,v) = \int_0^T &[D_{vv}\hat{L}(t,q(t),\dot{q}(t))(\dot{u}(t),\dot{v}(t)) + \langle u(t),v(t)\rangle]dt \\ &- \int_0^t \langle u(t),v(t)\rangle dt \; . \end{aligned}$$

By (CON) with $c(q) := \min(\min_{t\in[0,T]} l_6(q(t)), 1)$

$$|C_q(u,u) + \int_0^t \langle u(t),u(t)\rangle dt| \geq c(q)\|u\|^2_{H^1_{T^*}} \; . \tag{3.36}$$

Hence $C_q(u,v)+\int_0^t\langle u(t),v(t)\rangle dt$ is represented by a positive self-addjoint operator $A(q)$ (with respect to the equivalent scalar product on $H^1_{T^*}$ induced from the Riemannian structure of X). The second summand $\int_0^t\langle u(t),v(t)\rangle dt$ is represented by a compact operator. On the whole we obtain that $D^2\tilde{S}_L$ is represented by an operator that is the sum of a positive operator and a compact operator $K(q)$, hence a Fredholm operator with finite dimensional kernel and finite dimensional negative eigenspace.

The positive self-adjoint operator $A(q)$ is invertible. If W is an ϵ-neighbourhood of q_0 in H^1, the set $\{q(t) \; | q \in W, t \in [0,T]\}$ is contained in a compact subset M' of M. Set

$$c_1 := \sup_{q\in W} \frac{1}{c(q)} \leq \max(\max_{x\in M'} \frac{1}{l_6(x)}, 1) \; .$$

Inequality (3.36) implies

$$\|A(q)^{-1}\| \leq c_1 \quad \forall q \in W \; .$$

On the other hand by (SQ”)$_{loc}$

$$\begin{aligned}|\langle u, A(q)v\rangle| &= |\int_0^T [D_{vv}\hat{L}(t,q(t),\dot{q}(t))(\dot{u}(t),\dot{v}(t)) + \langle u(t),v(t)\rangle]dt| \\ &\leq c_2\|u\|_{H^1}\|v\|_{H^1} \ ,\end{aligned}$$

where $c_2 = 1 + \max_{x\in M'} l_3(x)$. Thus

$$\|A(q)\| \leq c_2 \quad \forall q \in W \ .$$

Please note that this locally uniform boundedness would be evident for continuous $q \mapsto A(q)$.

The statements about the equivariant case are easily checked. $\square$

The special form of the Hessian allows us to ’replace’ the gradient of S_L in the neighbourhood of a critical point with nondegenerate Hessian by the gradient of a smooth function. We put this into an abstract lemma, where we state more properties than we actually need, as we found them useful to understand the situation.

Lemma 3.2.12 *Let U be an open subset of a Hilbert space H and $f : U \to \mathbb{R}$ a C^1-functional with Hadamard differentiable derivative. So for $q \in U$, the second Gâteaux derivative $D^2 f(q)$ defines a symmetric bilinear form on H, which is represented by a self-adjoint operator H_q.*

Now suppose that H_q allows a decomposition

$$H_q = A(q) + K(q) \ ,$$

with the following properties:

i) $K(q)$ is compact for every $q \in U$, and the map $q \mapsto K(q)$ is continuous with respect to the norm topology.

ii) $A(q)$ is invertible for every $q \in U$, the map $q \mapsto A(q)$ is strongly continuous*, i.e. for every $v \in H$ the map $q \mapsto A(q)v$ is continuous.*

iii) The maps $q \mapsto \|A(q)\|$ and $q \mapsto \|A(q)^{-1}\|$ are locally uniformly bounded.

In this case every critical point $q_0 \in U$ of f with invertible H_{q_0} is isolated, and there is a neighbourhood W of q_0 in U such that the smooth function

$$g : W \to \mathbb{R}, \ q \mapsto g(q) = \frac{1}{2}\langle (q - q_0), H_{q_0}(q - q_0)\rangle$$

is a Lyapunov function for the negative gradient flow of f with respect to the equivalent inner product

$$(v, w) := \langle v, A(q_0)w\rangle$$

and f is a Lyapunov function for the negative gradient flow of g with respect to $(\cdot,\cdot)$.

More precisely

$$\begin{aligned}
&Df(q)(\nabla_{(,)}g(q)) = Dg(q)(\nabla_{(,)}f(q))\\
&= (\nabla_{(,)}f(q), \nabla_{(,)}g(q)) = \langle A(q_0)^{-1}\nabla f(q), A(q_0)^{-1}\nabla g(q)\rangle\\
&\geq c\|q\|^2 \;.
\end{aligned}$$

for some constant $c > 0$.

$\nabla_{(,)}g$ is a pseudo-gradient field for f.

Suppose G acts on H by isometries, $q_0 = 0$, and $A(0)$ amd $K(0)$ are G-equivariant. Then g is G-invariant and $\nabla_{(,)}g$ is G-equivariant.

Proof: As

$$\langle A(0)\nabla_{(,)}g(q), v\rangle = (\nabla_{(,)}g(q), v) = Dg(q)(v) = \langle\nabla g(q), v\rangle \;,$$

we conclude

$$\nabla_{(,)}g = A(0)^{-1}\nabla g(q) = A(0)^{-1}(A(q) + K(q)) \;.$$

We can assume $q_0 = 0$ without restriction of generality. By the locally uniform boundedness of $\|A(q)^{-1}\|$ and $\|A(q)^{-1}\|$ and the positiveness of these operators we can assume

$$\langle v, A(q)v\rangle \geq c\|v\|^2$$

and

$$\|A(q)v\| \leq C\|v\|$$

on some convex neighbourhood W of 0.

By assumption $A(0) + K(0)$ is bounded and invertible, hence $1 + A(0)^{-1}K(0)$ is bounded and invertible. There are $\lambda > 0$, $\Lambda > 0$ such that

$$\|[1 + A(0)^{-1}K(0)]v\| \geq \lambda\|v\| \;.$$

and

$$\|[1 + A(0)^{-1}K(0)]v\| \leq \Lambda\|v\| \;.$$

Now

$$\begin{aligned}
&[Df(q) - Df(0)](A(0)^{-1}(A(0) + K(0))q) \\
&= \int_0^1 D^2 f(sq)(q, A(0)^{-1}(A(0) + K(0))q)ds \\
&= \int_0^1 \langle [A(sq) + K(sq)]q, A(0)^{-1}[A(0) + K(0)]q\rangle ds \\
&= \int_0^1 \langle A(sq)[1 + A(sq)^{-1}K(sq)]q, [1 + A(0)^{-1}K(0)]q\rangle ds \\
&= \int_0^1 \langle [1 + A(sq)^{-1}K(sq)]q, A(sq)[1 + A(0)^{-1}K(0)]q\rangle ds \\
&= \int_0^1 \langle [1 + A(0)^{-1}K(0)]q, A(sq)[1 + A(0)^{-1}K(0)]q\rangle ds + R(q) \\
&\geq c\|[1 + A(0)^{-1}K(0)]q\|^2 + R(q) \\
&\geq c\lambda^2\|q\|^2 + R(q)\ .
\end{aligned}$$

We will now prove that the remainder term $R(q)$ is $o(\|q\|^2)$ for $q \to 0$.

$$\begin{aligned}
|R(q)| &= \left| \int_0^1 \langle [A(sq)^{-1}K(sq) - A(0)^{-1}K(0)]q, A(sq)[1 + A(0)^{-1}K(0)]q\rangle ds \right| \\
&\leq \sup_{s\in[0,1]} \|A(sq)^{-1}K(sq) - A(0)^{-1}K(0)\| C\Lambda \|q\|^2\ .
\end{aligned}$$

Now we need the important observation that the map

$$q \mapsto A(q)^{-1}K(q)$$

is continuous with respect to the norm operator topology.

(As we have not found a reference we sketch the simple argument: Suppose, (q_n) converges to 0 and (v_n) is a bounded sequence in H, then (v_n) contains a weakly convergent subsequence $v_n \rightharpoonup v$. The operator $K(0)$ maps this sequence into a strongly convergent sequence. Now

$$\begin{aligned}
&\|[A(q_n)^{-1}K(q_n) - A(0)^{-1}K(0)]v_n\| \\
&\leq \|A(q_n)^{-1}(K(q_n) - K(0))v_n\| + \|A(q_n)^{-1}K(0)(v_n - v)\| \\
&\qquad + \|(A(q_n)^{-1} - A(0)^{-1})K(0)v\| + \|A(0)^{-1}K(0)(v - v_n)\|\ .
\end{aligned}$$

For all but the third summand the convergence is immediate. For the third summand we need that $A(q_n)^{-1}$ converges strongly to $A(0)^{-1}$. But this follows easily from the boundedness of $A(q_n)^{-1}$ and the strong convergence of $A(q_n)$ or by a more abstract argument from the strong continuity of continuous functional calculus for normal operators, s. e.g. [Tak02], Lemma II 4.6.)

Thus for every $\epsilon > 0$ there is a $\delta > 0$ such that $\sup_{s\in[0,1]} \|A(sq)^{-1}K(sq) - A(0)^{-1}K(0)\| < \epsilon$ for $\|q\| < \delta$. By chosing ϵ so small that

$$\epsilon C\Lambda < \frac{1}{2}c\lambda^2$$

we get

$$Df(q)(A(0)^{-1}(A(0)+K(0))q) \geq \frac{1}{2}c\lambda^2\|q\|^2 \geq B\|Df(q)\|^2 ,$$

where we use the local Lipschitz continuity of $Df(q)$, which is a consequence of $Df(0)$ being Hadamard differentiable with our local bounds on the derivatives. As

$$|Df(q)(A(0)^{-1}(A(0)+K(0))q)| \leq \|Df(q)\| \; \Lambda \; \|q\| ,$$

we conclude

$$\|Df(q)\| \geq \frac{c\lambda^2}{2\Lambda}\|q\| \geq \frac{1}{A}\|A(0)^{-1}(A(0)+K(0))q\| .$$

We have thus proved that $A(0)^{-1}(A(0)+K(0))q$ is a pseudo-gradient field for f in $U_\delta(0)$. □

The following proposition generalises Proposition 2.2.24 to our less regular situation.

Proposition 3.2.13 *Under the assumptions of Proposition 3.2.11, for G-invariant L ($G \in \{\mathbb{Z}_2, \mathbb{Z}_p, S^1\}$) we have at any fixed (and hence critical) point γ of S_L*

$$k_G(\gamma) \geq \mu(\gamma) ,$$

where $\mu(\gamma)$ is the dimension of the negative eigenspace of the Hessian of S_L at γ.

Proof: By Proposition 3.2.11 the Hadamard derivative of the Fréchet derivative of S_L at γ defines a symmetric bilinear form on $T_\gamma \times T_\gamma$, represented by the self-adjoint Hessian $H : T_\gamma \to T_\gamma$. Let $i : W \to X$ be the smooth embedding of a $\mu(\gamma)$-dimensional submanifold of X, such that $i(p) = \gamma$ and $i_* T_p W = E^-$. Then by Proposition 3.2.6, $S_L \circ i$ is C^2 and $D^2(S_L \circ i)(p)$ is negative definite. The Morse lemma provides equivariant embeddings $(D^k, S^{k-1}) \to (U^c, U^{c-\epsilon})$ for a representation disk k and a neighbourhood of γ in $i(W)$, hence $k_G(\gamma) \geq \mu(\gamma)$. □

In order to do critical point theory we need a compactness condition.

Proposition 3.2.14 *If L satisfies (TP), (SQ') and (CON) and M is compact, the functional S_L satisfies the Palais-Smale condition.*

Proof: This is the standard proof as e. g. in [Ben86]. However, we prefer to write it down, as our assumptions on L are not precisely the same.

The property (CON) implies

$$\begin{aligned}
L(t,(q,v)) &= L(t,(q,0)) + \int_0^1 D_V L(t,(q,sv))(v)ds \\
&= L(t,(q,0)) \\
&\qquad + \int_0^1 \left(D_V L(t,(q,0))(v) + \int_0^1 D_{VV} L(t,(q,tsv))(sv,v)dt \right) ds \\
&\geq -|L(t,(q,0))| - \|D_V L(t,(q,0))\| \|v\|_q + \frac{1}{2} l_6(q) \|v\|_q^2 \\
&\geq l_7(q)\|v\|_q^2 - l_8(q) \quad .
\end{aligned}$$

with $l_6, l_7, l_8, : U \to \mathbb{R}^+$ continuous.

Hence

$$\|v\|_q^2 \leq \frac{l_8(q)}{l_7(q)} + \frac{1}{l_7(q)} L(t,(q,v))$$

and, as M is compact, we get for $(q,v) \in TM$

$$\langle v, v \rangle_q \leq c_1 + c_2 L(t,(q,v)) \tag{3.37}$$

for some $c_1, c_2 > 0$.

Suppose (γ_n) is a sequence in X with

$$\|DS_L(\gamma_n)\| \to 0 \text{ and} S_L(\gamma_n) \leq c \, ,$$

where $\|DS_L(\gamma_n)\|$ is the dual norm on $T^*_{\gamma_n} X$.

Then by (3.37)

$$\|\dot{\gamma}_n\|_{L^2([0,T],\mathbb{R}^N)} \leq Tc_1 + Tc_2 S_L(\gamma_n) \leq Tc_1 + Tc_2 c \, .$$

As $M \subset \mathbb{R}^N$ is compact, the sequence $(\|\gamma_n\|_{L^2([0,T],\mathbb{R}^N)})_n$ is bounded.

We conclude, that $(\gamma_n)_n$ is bounded in $H^1_T([0,T],\mathbb{R}^N)$ and has a subsequence $(\gamma_n)_n$ that converges uniformly and H^1-weakly to some $\gamma \in H^1_T([0,T],\mathbb{R}^N)$. As X is closed in $H^1_T(S^1,\mathbb{R}^N)$ with respect to uniform convergence, we have $\gamma \in X$.

Now by the following Lemma 3.2.16 the sequence converges strongly to γ, which therefore is a critical point of S_L. □

Remark 3.2.15 On noncompact manifolds we can still prove the Palais-Smale condition, if further conditions on L are imposed.

If, for example, $L(t,(q,v)) = \frac{1}{2}\langle v,v\rangle_q - V(q)$ is a classical Lagrangian with $-V$ coercive on M, it follows that $S_L(q) \leq c$ implies a $H^1_T(S^1,\mathbb{R}^N)$-bound on q, hence (PS). $\diamond$

Lemma 3.2.16 *Assume (TP),(SQ') and (CON). Then any bounded and uniformly convergent sequence $\gamma_n \in X$ with*

$$\|DS_L(\gamma_n)\| \to 0$$

is strongly convergent.

Proof: After possibly dropping the first elements, we can assume that all γ_n and $\gamma_0 := \gamma$ are contained in one of our exponential charts of X. More precisely there is an $\epsilon > 0$ and a smooth family of diffeomorphisms $\phi_t : U_{2\epsilon}(0) \to M$ such that

$$q(t) \mapsto \phi_t(q(t))$$

defines a map Φ from $H^1_{T^*}(S^1,\mathbb{R}^n) \cap C^0_{T^*}(S^1, U_{2\epsilon}(0))$ to an open subset V of X, such that $\gamma_n(t) \in U_\epsilon(0)$ for all $n \in \mathbb{N}$. If $\delta > 0$ is sufficiently small (i.e. smaller than some δ_0) the open δ-neighbourhood N_δ of 0 in $H^1_{T^*}(S^1,\mathbb{R}^n)$ is contained in $C^0_{T^*}(S^1,\overline{U_\epsilon(0)})$. From now on we abbreviate $L^2 := L^2(S^1,\mathbb{R}^n)$, $H^1 := H^1_{T^*}(S^1,\mathbb{R}^n)$, and we suppose $\delta < \delta_0$.

By Proposition 3.1.2 the intrinsical Hilbert-Riemannian structure is equivalent to the extrinsical. The pull-back of the extrinsical Riemannian structure to N_δ is uniformly equivalent to the Hilbert-structure on $H^1_{T^*}(S^1,\mathbb{R}^n)$, if δ is sufficiently small.

This may be considered *evident*, but presumably for this very reason, we could not find a proper reference and provide a proof:

Proof of the claim: Please note that $\Phi(q)^*TM$ is an n-dimensional vector bundle over $[0,T]/(0 \sim T) \cong S^1$ with the inherited inner product on the fibres. If F is one of $\{C^0, H^1, L^2\}$ we denote the space of F-sections by $F(\Phi(q)^*TM)$. The tangent space $T_{\Phi(q)}X$ is parametrised by

$$H^1_{T^*}(S^1,\mathbb{R}^n) \to T_{\Phi(q)}X, \quad v \mapsto D\Phi|_q(v)\ ,$$

where

$$D\Phi|_q(v)(t) = D\phi_t|_{q(t)}(v(t))\ .$$

The derivative of $D\Phi|_q(v)(t)$ with respect to t is given by

$$D\phi_t|_{q(t)}(\dot v(t)) + (\frac{d}{dt}D\phi_t)|_{q(t)}(v(t)) + D^2\phi_t|_{q(t)}(\dot q(t), v(t))\ .$$

We can rewrite this as
$$D\Phi|_q(v) = A_q(v)$$
and
$$\frac{d}{dt} D\Phi|_q(v) = A_q(\dot{v}) + B_q(v) \ .$$
where $A_q : L^2 \to L^2(\Phi(q)^*TM)$ and A_q^{-1} are uniformly bounded isomorphisms for $q \in N_\delta$ and $B_q : H^1 \to L^2(\Phi(q)^*TM))$ is uniformly bounded for $q \in N$, more precisely
$$\begin{aligned}
\|A_q(v)\|_{L^2(\Phi(q)^*TM)} &\leq c\|v\|_{L^2} \\
\|B_q(v)\|_{L^2(\Phi(q)^*TM))} &\leq d_1|v\|_{L^2} + d_2\delta\|v\|_{H^1} \\
\|A_q(v)\|_{L^2(\Phi(q)^*TM)} &\geq e\|v\|_{L^2}
\end{aligned}$$
for all $q \in N$ and all $v \in H^1$. We deduce
$$\begin{aligned}
\|D\Phi|_q(v)\|^2_{H^1(\Phi(q)^*TM)} &= \|A_q(v)\|^2_{L^2(\Phi(q)^*TM)} + \|A_q(\dot{v}) + B_q(v)\|^2_{L^2(\Phi(q)^*TM)} \\
&\leq c^2\|v\|^2_{L^2} + (c\|\dot{v}\|_{L^2} + d\|v\|_{H^1})^2 \quad \text{for} \quad d := d_1 + d_2\delta \\
&\leq c^2\|v\|^2_{L^2} + 2c^2\|\dot{v}\|^2_{L^2} + 2d^2\|v\|^2_{H^1} \\
&\leq f^2\|v\|^2_{H^1}
\end{aligned}$$
for a convenient constant $f > 0$. And
$$\begin{aligned}
&\|D\Phi|_q(v)\|^2_{H^1(\Phi(q)^*TM)} \\
&= \|A_q(v)\|^2_{L^2(\Phi(q)^*TM)} + \|A_q(\dot{v}) + B_q(v)\|^2_{L^2(\Phi(q)^*TM)} \\
&\geq \|A_q(v)\|^2_{L^2(\Phi(q)^*TM)} + (\|A_q(\dot{v})\|_{L^2(\Phi(q)^*TM)} - \|B_q(v)\|_{L^2(\Phi(q)^*TM)})^2 \\
&\geq \|A_q(v)\|^2_{L^2(\Phi(q)^*TM)} + \|A_q(\dot{v})\|^2_{L^2(\Phi(q)^*TM)} - 2\|A_q(\dot{v})\|_{L^2(\Phi(q)^*TM)}\|B_q(v)\|_{L^2(\Phi(q)^*TM)} \\
&\qquad + \|B_q(v)\|^2_{L^2(\Phi(q)^*TM)} \\
&\geq \|A_q(v)\|^2_{L^2(\Phi(q)^*TM)} + \|A_q(\dot{v})\|^2_{L^2(\Phi(q)^*TM)} - \lambda^2\|A_q(\dot{v})\|^2_{L^2(\Phi(q)^*TM)} \\
&\qquad - \frac{1}{\lambda^2}\|B_q(v)\|^2_{L^2(\Phi(q)^*TM)} + \|B_q(v)\|^2_{L^2(\Phi(q)^*TM)} \\
&= \|A_q(v)\|^2_{L^2(\Phi(q)^*TM)} + (1-\lambda^2)\|A_q(\dot{v})\|^2_{L^2(\Phi(q)^*TM)} - (\frac{1}{\lambda^2} - 1)\|B_q(v)\|^2_{L^2(\Phi(q)^*TM)} \\
&\geq e^2\|v\|^2_{L^2} + (1-\lambda^2)e^2\|\dot{v}\|^2_{L^2} - (\frac{1}{\lambda^2} - 1)(d_1\|v\|_{L^2} + d_2\delta\|v\|_{H^1})^2 \\
&\geq e^2\|v\|^2_{L^2} + (1-\lambda^2)e^2\|\dot{v}\|^2_{L^2} - (\frac{1}{\lambda^2} - 1)(2d_1^2\|v\|^2_{L^2} + 2d_2^2\delta^2\|v\|^2_{H^1}) \\
&= ((1-\lambda^2)e^2 - (\frac{1}{\lambda^2} - 1)2d_2^2\delta^2)\|v\|^2_{H^1} + (\lambda^2 e^2 - (\frac{1}{\lambda^2} - 1)2d_1^2)\|v\|^2_{L^2}
\end{aligned}$$
for some $\lambda \in (0,1)$. We can chose λ large enough to ensure that
$$\lambda^2 e^2 - (\frac{1}{\lambda^2} - 1)2d_1^2 > 0$$

and then chose δ small enough to satisfy

$$(1-\lambda^2)e^2 - (\frac{1}{\lambda^2}-1)2d_2^2\delta^2 > 0 ,$$

and thus have proved the claim. □

The rest of the proof goes as usual. Suppose $\gamma_n = \Phi(q_n)$ and $\gamma = \Phi(q)$ with $q_n \in N_\delta$. The sequence $(q_n)_{n\in\mathbb{N}}$ converges uniformly to q, and $\|q_n\|_{H^1} \le \delta$. To prove the strong convergence of $(q_n)_{n\in\mathbb{N}}$ in H^1 it is sufficient to prove that $(\dot{q}_n)_{n\in\mathbb{N}}$ converges strongly in L^2.

Any subsequence of $(q_n)_{n\in\mathbb{N}}$ contains a weakly convergent subsequence, also denoted by $(q_n)_{n\in\mathbb{N}}$. It follows that $(\dot{q}_n)$ converges weakly to $\dot{q}$ in L^2. We are now going to prove that $(q_n)_{n\in\mathbb{N}}$ converges strongly to q. (It follows by the 'subsequence principle', that the original sequence (q_n) converges strongly.)

We note that $D\tilde{S}_L(q_n)(q_n - q)$ converges to zero, as $q_n - q$ is bounded in H^1.

$$\begin{aligned} & D\tilde{S}_L(q_n)(q_n-q) \\ = & \int_0^T \Big(D_q\hat{L}(t,q_n(t),\dot{q}_n(t))(q_n(t)-q(t)) + D_v\hat{L}(t,q_n(t),\dot{q}_n(t))(\dot{q}_n(t)-\dot{q}(t))\Big)dt \ . \end{aligned}$$

The integral over the first summand converges to zero, as $D_q\hat{L}(t,q_n(t),\dot{q}_n(t))$ is bounded in $L^1([0,T],(\mathbb{R}^n)^*)$ and $q_n \to q$ uniformly. Hence the integral over the second summand converges to zero, too.

$$\begin{aligned} & \int_0^T D_v\hat{L}(t,q_n(t),\dot{q}_n(t))(\dot{q}_n(t)-\dot{q}(t))dt \\ = & \int_0^T D_v\hat{L}(t,q_n(t),\dot{q}(t))(\dot{q}_n(t)-\dot{q}(t))dt \\ & + \int_0^T\int_0^1 D_{vv}\hat{L}(t,q_n(t),\dot{q}(t)+s(\dot{q}_n(t)-\dot{q}(t)))(\dot{q}_n(t)-\dot{q}(t),\dot{q}_n(t)-\dot{q}(t)) \ ds \ dt \end{aligned} \tag{3.38}$$

We observe that $D_v\hat{L}(t,q_n(t),\dot{q}(t))$ converges in L^2 to $D_v\hat{L}(t,q(t),\dot{q}(t))$:
All $q_n(t)$ are contained in a compact neighbourhood K of $q([0,T])$, hence by (SQ')$_{loc}$ the sequence of functions

$$t \mapsto \|D_v\hat{L}(t,q_n(t),\dot{q}(t))\|$$

is dominated by an L^2-function. As $D_v\hat{L}(t,q_n(t),\dot{q}(t))$ converges almost everywhere to $D_v\hat{L}(t,q(t),\dot{q}(t))$, the assertion follows by Lebesgue's theorem on dominated convergence.

As $D_v\hat{L}(t, q_n(t), \dot{q}(t))$ converges in L^2 and $\dot{q}_n(t) - \dot{q}(t)$ converges weakly to zero in L^2, the pairing

$$\int_0^T D_v\hat{L}(t, q_n(t), \dot{q}(t))(\dot{q}_n(t) - \dot{q}(t))dt$$

converges to zero.

Hence the double integral (3.38) converges to zero. We estimate by means of $(\mathrm{CON})_{loc}$

$$\begin{aligned}&\int_0^T \int_0^1 D_{vv}\hat{L}(t, q_n(t), \dot{q}(t) + s(\dot{q}_n(t) - \dot{q}(t)))(\dot{q}_n(t) - \dot{q}(t), \dot{q}_n(t) - \dot{q}(t))\ ds\ dt\\ &\geq c\int_0^T \int_0^1 \|\dot{q}_n(t) - \dot{q}(t))\|^2 ds\ dt\\ &\geq c\|\dot{q}_n - \dot{q}\|^2_{L^2}\ ,\end{aligned}$$

where c is a lower bound for $l_6(q)$ on K.

We conclude the L^2 convergence of $\dot{q}_n$ to $\dot{q}$. □

3.3. Multiplicity results for symmetric Lagrangian systems

Remark 3.3.1 We consider a Lagrangian system on TM satisfying (SQ'') and (CON). Although S_L is usually not C^2 under these conditions, a Hessian is defined at critical points p, and by Proposition 3.2.11 the dimension of the negative eigenspace $\mu(p) := E^-$ is finite. If the Hessian at a critical orbit γ is non-degenerate, we conclude by Lemma 3.2.12 and Corollary 2.2.21 that γ is a non-degenerate critical point in the sense of Definition 2.2.11 with Morse index $i(\gamma) = \mu(\gamma)$.

This Morse index can be understood and calculated in different ways. (CON) allows to obtain an equivalent Hamiltonian system on T^*M by means of Legendre transformation. The Maslov index and the relative Morse index (in the sense of Abbondandolo, who generalises the idea of Conley and Zehnder) of a solution of the Hamiltonian system are equal to the Morse index of the corresponding solution of the Lagrangian system (s. [Abb03]).

Let X_0 be the component of $X = H^1_T(S^1, M)$ consisting of contractible loops. There is a projection $\pi : X_0 \to M$, $\gamma \mapsto \gamma(0)$ (which is a Hurewicz fibration) and a section $\sigma : M \to X_0$ which maps each $x \in M$ to the constant loop with value

x. A smooth G-action, induces a smooth G-action on X_0, for which σ and π are G-equivariant. Suppose M has only isolated fixed points, then all fixed points of X_0 are constant loops contained in $\sigma(M)$. In order to make the statements more readable we write $\overline{p} := \sigma(p)$

Thus we can apply Theorem 2.2.28 or Theorem 2.2.31. $\diamond$

Theorem 3.3.2 *Suppose (G, R, d) is one of the triples $(\mathbb{Z}_2, \mathbb{Z}_2, 0)$, $(\mathbb{Z}_p, \mathbb{Z}_p, 1)$, $(S^1, \mathbb{Q}, 1)$ and h^* is Čech cohomology with coefficients R.*

Suppose M is a compact n-dimensional differentiable G-manifold and $L : \mathbb{R} \times TM \to \mathbb{R}$ a G-invariant T-periodic Lagrange function that satisfies (SQ") and (CON).

Now let $p \in M$ be a fixed point of the G-operation and

$$F := \{x \in M^G \mid S_L(\overline{x}) < S_L(\overline{p})\}$$

the set of fixed points with action below $S_L(\overline{p})$.

Suppose for all $x \in F$ the constant solution $\sigma(x)$ is a non-degenerate critical point of S_L. This is the case, if the Hessian of S_L at the constant solution is non-degenerate.

Then, if

$$k_G(\overline{p}) > \sigma(p, F) \ ,$$

we have the following result:

i) $G \in \{\mathbb{Z}_2, S^1\}$**:** *There are at least*

$$(k_G(\overline{p}) - \sigma(p, F))\frac{1}{d+1} \geq (i(\overline{p}) - \sigma(p, F))\frac{1}{d+1}$$

non-fixed G-orbits of T-periodical solutions of the Lagrangian system with action below $S_L(\sigma(p))$.

ii) $G = \mathbb{Z}_p$: *There are at least*

$$(k_G(\overline{p}) - \sigma_T(p, F))\frac{1}{d+1} \geq (i(\overline{p}) - \sigma_T(p, F))\frac{1}{d+1}$$

non-fixed G-orbits of T-periodical solutions of the Lagrangian system with action below $S_L(\sigma(p))$.

Proof: After Remark 3.3.1, nothing remains to prove. $\square$

Remark 3.3.3 1. In order to use the theorem we have to estimate $k_G(\overline{p})$. The rough estimate (Proposition 3.2.13)

$$k_G(\overline{p}) \geq \mu(\overline{p})$$

can be improved in some cases. For example, if we have a classical Lagrangian

$$L(t,(q,v)) = \frac{1}{2}\|v\|_q^2 - V(t,q) \; .$$

and

$$V(t,q) = V_2(t,q) + o(d(p,q)^2) \; ,$$

the following condition (V_4) from [BW97b]

$$V(t,q) > V_2(t,q) \quad \text{for} \quad 0 < d(q,p) < \epsilon$$

(for a sufficiently small $\epsilon > 0$) guarantees

$$k_G(\overline{p}) \geq \mu(\overline{p}) + \nu(\overline{p}) \; ,$$

where $\nu(\overline{p})$ is the nullity, i.e. the dimension of the Kernel of the Hessian at $\overline{p}$.

2. In some cases the non-degeneracy assumption can be easily verified. If there is a point q such that

$$V(t,x) \leq V(t,q)$$

for all $x \in M$ and all $t \in \mathbb{R}$, the corresponding T-periodical solution $\overline{q}$ is an absolute minimum of S_L and hence a non-degenerate critical point in the sense of Definition 2.2.11 with Morse index 0. $\diamond$

Remark 3.3.4 This statement improves Theorem 3.13. of Bartsch and Wang [BW97b] in several respects. It works not only on the torus, but for Lagrangian systems on the tangent bundle of any compact manifold. Furthermore, we consider general Lagrangians, that satisfy the growth condition (SQ") and the convexity condition (CON), whereas [BW97b] considers classical Lagrangians

$$L(t,(q,v)) = \frac{1}{2}\|v\|_q^2 - V(t,q) \; .$$

In the case of the torus, it allows to predict more critical points. The improvement of our result with respect to Bartsch and Wang is measured by the difference

$$n - \sigma(p.F) \; .$$

In *general situations* it may be difficult to get better estimates than $\sigma(p, F) \leq n$, yet for G-manifolds that are (TNHZ) this is an algebraic invariant, which we can estimate from above (s. Example 2.1.34) or even calculate. The n-torus with 2^n fixed points is (TNHZ), it is one of the cases dealt with in Example 2.1.34.

We mention again a few special cases, where our result improves the result of Bartsch and Wang.

If $|F| \leq 2^n - 2$ (that is, there is at least one fixed point with action above $S_L(\overline{p})$)

$$\sigma(p, F) \leq n - 1 \ .$$

And the estimate

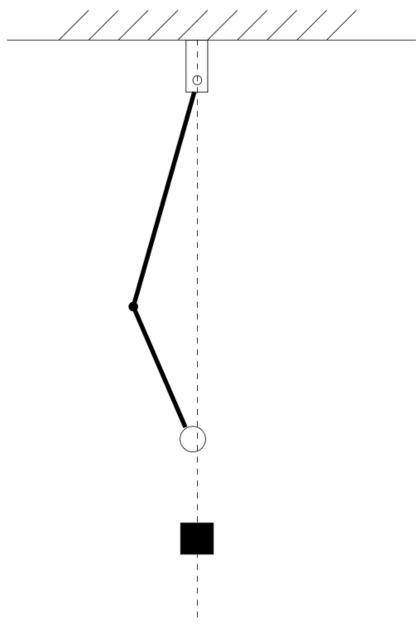
$$\sigma(p, F) \leq \left[\frac{|F|}{2}\right] + 1 \ ,$$

yields an improvement whenever $|F| \leq 2N - 2$.

If we happen to know that p is contained in a k-dimensional G-invariant submanifold of T^n that does not intersect F, we have (Proposition 2.1.40)

$$\sigma(p, F) \leq n - k \ .$$

Example 3.3.5 Just to give an illustration of the result consider a double pendulum.



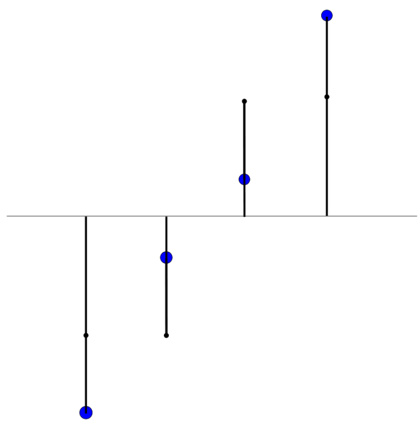
We make the idealising assumptions that the mass of the rods is negligeable compared to the mass m of the pivot and the mass M of the bob, that gravitation is pointing downwards, and that there is no friction. The configuration space of the double pendulum is T^2. The pendulum is symmetric with respect to reflection in the dashed line. The movement of the pendulum is described by an autonomous system. However, we can add a T-periodic symmetric magnetic field via an electromagnet in the little black box underneath the pendulum, assuming that only the bob is paramagnetic.

Let us analyse the autonomous case first. If $r_1 > r_2$ are the lengths of the upper and lower rod, respectively, and α_1, α_2 are the angles of the rods with the vertical line, the Lagrangian function is

$$\begin{aligned} &L((\alpha_1, \alpha_2), (\omega_1, \omega_2)) \\ =& \frac{1}{2} M \left((r_1 \omega_1 \sin \alpha_1 + r_2 \omega_2 \sin \alpha_2)^2 + (r_1 \omega_1 \cos \alpha_1 + r_2 \omega_2 \cos \alpha_2)^2 \right) \\ &+ \frac{1}{2} m r_1^2 \omega_1^2 + g((m + M) r_1 \cos \alpha_1 + M r_2 \cos \alpha_2) \ , \end{aligned}$$

where $g > 0$ is a constant (the graviation). (For $m = 0$ the condition (CON) would be violated, so we actually *need* a certain lack of idealisation.)

The double pendulum has four stationary points (down-down, down-up, up-down, up-up).



The equilibrium point down-down has the highest action, the equilibrium point up-up the lowest. Let us now consider the point up-down $p = (\pi, 0)$. As there is only one fixed point with lower action, we have

$$\sigma(p, F) = 1 \ .$$

The Hessian of S_L at $\overline{p}$ is

$$\begin{aligned} & D^2 S_L(\overline{p})(v,v) \\ = & \int_0^T \left(\left\langle v, g \begin{pmatrix} (M+m)r_1 & 0 \\ 0 & -Mr_2 \end{pmatrix} v \right\rangle \right. \\ & \left. + \left\langle \dot{v}, \begin{pmatrix} (M+m)r_1^2 & -Mr_1r_2 \\ -Mr_1r_2 & Mr_2^2 \end{pmatrix} \dot{v} \right\rangle \right) dt \end{aligned}$$

We set

$$A := \begin{pmatrix} (M+m)r_1 & 0 \\ 0 & -Mr_2 \end{pmatrix} , \qquad B := \begin{pmatrix} (M+m)r_1^2 & -Mr_1r_2 \\ -Mr_1r_2 & Mr_2^2 \end{pmatrix} .$$

The bilinear form $\langle \cdot, A \cdot \rangle$ is represented by a matrix H with respect to the scalar product $\langle \cdot, B \cdot \rangle$. H is non-degenerate and indefinite. Let $-\alpha$ be the negative eigenvalue of H. Then a Fourier decomposition allows to calculate the Morse index of $\overline{p}$ as in [MW89], Proposition 9.1.

$$\mu(\overline{p}) = 1 + 2\#\{j \in \mathbb{N}^* | \ \frac{4\pi^2 j^2}{T^2} < \alpha\} .$$

The Morse index grows roughly linearly in T.

In a similar way we argue, that for all but a discrete set of values of T all fixed points are non-degenerate. We can chose a large T, so that all fixed points are non-degenerate orbits and

$$\mu(\overline{p}) > 1 .$$

By our theorem there are at least $\mu(\overline{p}) - 1$ pairs of (S^1-orbits) of contractible T-periodical orbits with action below $S_L(\overline{p})$. The T-average of the potential

$$V(\alpha_1, \alpha_2) := -g((m+M)r_1 \cos\alpha_1 + Mr_2 \cos\alpha_2)$$

over such a solution must be greater than the potential at p

$$V(\pi, 0) = g((m+M)r_1 - Mr_2) ,$$

hence these orbits presumably look strange, as the longer rod stays more or less upright.

These results will remain true, when a symmetric T-periodic perturbation is "switched on", but it will be more difficult to calculate the Morse index of the trivial solution $\overline{p}$.

By the way, the result implies that there are infinitely many prime periodic solutions with periods kT, $k \in \mathbb{N}$ and actions below the action of p as a kT-periodic curve. Here the improvement of our result with respect to Bartsch and Wang is irrelevant, because the Morse index of p as a kT-periodic curve tends to infinity for $k \to \infty$, anyway. (The assumption that there are only a finite number of prime curves leads to a contradiction.) $\diamond$

A. A lemma about smooth retractions

The following Lemma seems to be well known, but lacking a standard reference, we provide a proof.

Lemma A.1 Suppose we have a C^k-retraction R of some open subset U of a Banach-space B to $X \subset B$. Then X is a C^k-submanifold of U. For any $x \in X$ the derivative $L_x := DR(x)$ is a linear retraction onto T_xX. Every $x \in X$ has a neighbourhood V in X such that the mapping

$$y \mapsto L_x(y - x)$$

provides a chart onto some open subset of $\mathrm{Im}\ (L_x) = T_xX$. $\diamond$

Proof: As L_x is an idempotent continuous linear operator, it defines a splitting of $B = \mathrm{Im}\ L_x \oplus \mathrm{Ker}\ L_x$ of B as the direct sum of two closed subspaces.

Let $P_x := \mathrm{id}_B - L_x$ be the projection of B onto $\mathrm{Ker}\ (L_x)$ along $\mathrm{Im}\ (L_x)$ according to this splitting.

Now consider the map

$$\Phi_x : B \to B,\ y \mapsto P_x(y - R(y)) + L_x(y - x)\ .$$

It is continuously differentiable with

$$D\Phi_x(x)(v) = P_x(v - L_x(v)) + L_x(v) = P_x(v) + L_x(v) = v\ .$$

As the set of isomorphisms is open in the set of bounded operators on B, there is a neighbourhood V of x such that $D\Phi_x(y)$ is an isomorphism for all $y \in V$. Hence, $\Psi_x := P_x \circ \Phi_x|_V$ is a submersion from V to $\mathrm{Ker}\ L_x$, $\tilde{X}_V := \Psi_x^{-1}(0)$ is a C^k-submanifold of $V \subset B$ and L_x defines a chart $\tilde{X}_V \to \mathrm{Im}\ L_x$. The tangent space of $\tilde{X}_V$ at x is given by

$$\mathrm{Ker}\ D\Psi_x(x) = \mathrm{Ker}\ P_x = \mathrm{Im}\ L_x\ .$$

We will prove below, that this chart provides a local chart for X, more precisely, there is a (possibly smaller) open neighbourhood $W \subset V$ of x such that $X \cap W = \tilde{X}_V \cap W$:

It is obvious, that $X \cap V \subset \tilde{X}_V$, as for $y \in X \cap V$ we have $R(y) = y$ and thus $\Phi_x(y) = 0$, i.e. $y \in \tilde{X}_V$. On the other hand, for $y \in \tilde{X}_V$,

$$P_x(y - R(y)) = 0 \ .$$

From the following statement we deduce that for y in some neighbourhood $W \subset V$, we have indeed $y \in X$.

Claim: There is a neighbourhood $W \subset V$ of x, such -that

$$\|P_x(y - R(y))\| \geq \frac{1}{3}\|y - R(y)\| \ .$$

Proof: $R(y)$ is continuously differentiable, so there is a neighbourhood W of x such that $\|DR(z) - DR(y)\| \leq \frac{1}{3}$ for all $y, z \in W$. In the Taylor expansion of degree 1

$$\begin{aligned} R(z) &= R(y) + \int_0^1 DR(t(z-y)+y)(z-y)dt \\ &= R(y) + DR(y)(z-y) + \int_0^1 (DR(t(z-y)+y) - DR(y))(z-y)dt \ , \end{aligned}$$

the norm of the remainder term

$$\tilde{R}(y,z) := \int_0^1 (DR(t(z-y)+y) - DR(y))(z-y)dt$$

can be estimated from above

$$\|\tilde{R}(y,z)\| \leq \frac{1}{3}\|z - y\|$$

for all $y, z \in W$.

From

$$\begin{aligned} R(y) &= R(R(y)) + L_{R(y)}(y - R(y)) + \tilde{R}(R(y), y) \\ &= R(y) + L_x(y - R(y)) + (L_{R(y)} - L_x)(y - R(y)) + \tilde{R}(R(y), y) \end{aligned}$$

we get

$$\begin{aligned} \|P_x(y - R(y))\| &= \|y - R(y) + (L_{R(y)} - L_x)(y - R(y)) + \tilde{R}(R(y), y)\| \\ &\geq \|y - R(y)\| - \frac{1}{3}\|y - R(y)\| - \frac{1}{3}\|y - R(y)\| \\ &= \frac{1}{3}\|y - R(y)\| \ . \end{aligned}$$

□

With the proof of the above claim the lemma is proved. □

B. A covariant proof of Proposition 3.2.9

We recall Proposition 3.2.9:

If L satisfies (TP), (SQ') and (REG), the critical points of S_L are precisely the T-periodic C^2 solutions of

$$\nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor} - \frac{D}{dt}\nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert} = 0\ . \tag{B.1}$$

In local coordinates as above they correspond to the smooth solutions of

$$D_q\hat L(t,\gamma(t),\dot\gamma(t)) - \frac{d}{dt}D_v\hat L(t,\gamma(t),\dot\gamma(t)) = 0\ . \tag{B.2}$$

If L is C^k the solutions are C^k.

Proof: Suppose γ is a critical point of S_L. This means

$$\begin{aligned}0 = DS_L(\gamma)(u) = \int_0^T &\langle \nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert}, \frac{D}{dt}u(t)\rangle \\ &+ \langle \nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor}, u(t)\rangle dt\ . \end{aligned}\tag{B.3}$$

$$\tag{B.4}$$

We need to verify that integration by parts works in a Riemannian setting. The product rule is satisfied almost everywhere for (absolutely continuous representatives) of functions $u, v \in T_\gamma X \subset H^1_T(S^1,\mathbb{R}^N)$:

$$\frac{d}{dt}\langle u(t), v(t)\rangle = \langle \frac{D}{dt}u(t), v(t)\rangle + \langle u(t), \frac{D}{dt}v(t)\rangle\ .$$

And therefore

$$0 = \langle u(T), v(T)\rangle - \langle u(0), v(0)\rangle = \int_0^T (\langle \frac{D}{dt}u(t), v(t)\rangle + \langle u(t), \frac{D}{dt}v(t)\rangle)dt\ . \tag{B.5}$$

We need, moreover, a notion of *covariant primitive* $U(t)$ [1] of a vector field u along γ, which is supposed to be L^p, $1 \leq p \leq 2$.

The covariant primitive $U := I(u, U_0)$ with initial value $U(0) = U_0 \in T_{\gamma(0)}M$ is the unique vector field along γ that satisfies

$$\frac{D}{dt}U(t) = u(t) \text{ and } U(0) = U_0 \ .$$

For $U_0 = 0_{\gamma(0)}$ we chose the notation $U(t) = \int_0^t u(s)ds$ by analogy with the euclidean case. We have to justify the existence of this primitive in local coordinates. It is sufficient to solve the problem for a curve $\gamma \in X$ with values in one coordinate chart. If this is not the case we can subdivide the curve into pieces for which it is. The local problem reads as

$$\begin{aligned}
&\frac{D}{dt}(\sum_i U^i(t)\frac{\partial}{\partial x_i}) \\
&= \sum_i \dot{U}^i \frac{\partial}{\partial x_i} + \sum_{ijk} \Gamma^i_{jk} U^j \dot{\gamma}^k \frac{\partial}{\partial x_i} \\
&= \sum_i u^i \frac{\partial}{\partial x_k} \ .
\end{aligned}$$

This defines a first order linear inhomogenous differential equation for $(U_i)_i$ with right hand side $(u_i)_i \in L^p([0,T], \mathbb{R}^n)$. The coefficients $(\sum_k \Gamma^i_{jk}\dot{\gamma}^k)_{ij}$ of the zero order terms are in $L^2([0,T], \mathbb{R}^{n\times n})$. This equation has a unique absolutely continuous almost everywhere solution $(U_i)_i$ with initial value $0_{\mathbb{R}^n}$ (s. e.g. Theorem 2.2. of [Zet97]). By a boostrapping argument we see that $(U_i)_i \in W^{1,p}([0,T], \mathbb{R}^n)$. If $\gamma \in C^k$ and $u_i \in C^{k-1}$ we have $U_i \in C^k$.

Any two covariant primitives of u differ by a parallel vector field. The covariant primitive $I(0, U_0)$ of the zero vector field with initial value U_0 is almost everywhere differentiable with covariant derivative 0, i.e. it is the parallel vector field achieved by parallel transport of U_0 along γ, which generalises the parallel transport along a C^1-curve to the H^1-case. We define the n-dimensional space $P_\gamma := \{I(0, U_0) | U_0 \in T_{\gamma(9)}X\}$ of *parallel vector fields* along γ, which is a subspace of the Hilbert space H of L^2-sections of $\gamma^*(TM)$, *but not* a subspace of $T_\gamma X$, which only contains periodical vector fields. The map $\Gamma_1 : T_{\gamma(0}M \to T_{\gamma(T)}M$, $u_0 \mapsto I(0, U_0)(T)$ is linear. It is bijective and in fact an isometry, as for any $u_0 \in T_{\gamma(0)}M$

$$\frac{d}{dt}\langle I(0, U_0)(t), I(0, U_0)(t)\rangle = 2\langle \frac{D}{dt} I(0, U_0)(t), I(0, U_0)(t)\rangle = 0 \ ,$$

[1] The notion of covariant primitive, though quite natural, does not appear often in the literature, therefore we supply the necessary details.

hence $\|I(0,U_0)(T)\| = \|U_0\|$.

Likewise we consider the double primitive $I(I(0,U_0),0)$ and obtain

$$\begin{aligned}&\frac{d}{dt}\langle I(I(0,U_0),0)(t), I(0,U_0)(t)\rangle\\ &= \langle \frac{D}{dt} I(I(0,U_0),0)(t), I(0,U_0)(t)\rangle\\ &= \langle I(0,U_0)(t), I(0,U_0)(t)\rangle = \|U_0\|^2\end{aligned}$$

and hence

$$\langle I(I(0,U_0),0)(t), I(0,U_0)(t)\rangle = t\|U_0\|^2 \ .$$

In particular the map $\Gamma_2 : T_{\gamma(0)}M \to T_{\gamma(1)}M, U_0 \mapsto I(I(0,U_0),0)(T)$ is a linear isomorphism.

Please note that the parallel vector fields $I(0,U_0)$ correspond to the constant vector fields in the euclidean setting, and the vector fields $I((0,U_0),0)$ correspond to the vector fields linear in t.

From the above we obtain the rule of integration by parts for $u \in L^2([0,T],\mathbb{R}^n)$ and $v \in H^1([0,T],\mathbb{R}^n)$

$$\int_0^T \langle u(t), v(t)\rangle dt = \langle \int_0^T u(t)dt, v(T)\rangle - \int_0^T \langle \int_0^t u(s)ds, \frac{D}{dt}v(t)\rangle dt \ . \qquad \text{(B.6)}$$

Hence

$$\begin{aligned}&\int_0^T \langle \nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor}, u(t)\rangle dt\\ =&\langle \int_0^T \nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor} dt, u(T)\rangle - \int_0^T \langle \int_0^t \nabla_H L(s,(\gamma(s),\dot\gamma(s)))_{hor} ds, \frac{D}{dt}u(t)\rangle \ .\end{aligned}$$

For any $u \in T_\gamma X \cap C^\infty([0,T],M)$ with $u(T) = 0_{\gamma(T)}$ we conclude

$$\int_0^T \langle \nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert} - \int_0^t \nabla_H L(s,(\gamma(s),\dot\gamma(s)))_{hor} ds, \frac{D}{dt}u(t)\rangle dt = 0 \ .$$

Now an arbitrary $v \in T_\gamma X \cap C^\infty([0,T],M)$ has a H^1-primitive $V := I(v,0)$, which is not periodic in general. However the vector field

$$\tilde V := V - I(I(0,\Gamma_2^{-1}(V(T))),0)$$

is T-periodic with $\tilde V(T) = 0_{\gamma(T)}$ and

$$\frac{D}{dt}\tilde V(t) = v(t) - I(0,\Gamma_2^{-1}(V(T)))(t)$$

hence

$$\int_0^T \langle \nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert} \\ - \int_0^t \nabla_H L(s,(\gamma(s),\dot\gamma(s)))_{hor} ds \ , \ v(t) - I(0,\Gamma_2^{-1}(V(T)))(t)\rangle dt = 0 \ .$$

Now observe that $T_\gamma X$ is dense in the space H of L^2-sections of $\gamma^*(TM)$. Therefore the above relation implies that

$$\nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert} - \int_0^t \nabla_H L(s,(\gamma(s),\dot\gamma(s)))_{hor} ds$$

is an element of the space P_γ of parallel vector fields, hence it is weakly differentiable and an absolutely continuous representative of $\nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert}$ is differentiable almost everywhere with a. e.

$$\frac{D}{dt}\nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert} - \nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor} = 0 \ . \tag{B.7}$$

As $(t \mapsto \nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor}) \in L^1$, we conclude that $t \mapsto \nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert}$ is in $W^{1,1}$. By assumption (REG) the map Δ^{-1} is linearly bounded on the fibres and

$$\dot\gamma \in W^{1,1}([0,T],\mathbb{R}^N) \subset C^0([0,T],\mathbb{R}^N) \ .$$

Furthermore $\nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert}$ is T-periodic as by (B.3),(B.5) and (B.7)

$$\begin{aligned} 0 &= \int_0^T (\langle \nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert}, \frac{D}{dt}u(t)\rangle + \langle \nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor}, u(t)\rangle)dt \\ &= \int_0^T (\langle \nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert}, \frac{D}{dt}u(t)\rangle + \langle \frac{D}{dt}\nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert}, u(t)\rangle)dt \\ &= \langle \nabla_V L(T,(\gamma(T),\dot\gamma(T)))_{vert}, u(T)\rangle - \langle \nabla_V L(T,(\gamma(0),\dot\gamma(0)))_{vert}, u(0)\rangle \end{aligned}$$

for any $u \in T_\gamma X$. We conclude by assumption (REG) that $\dot\gamma$ is T-periodic.

As $\nabla_H L(t,(\gamma(t),\dot\gamma(t)))_{hor}$ is continuous, it follows by (B.7) that $\nabla_V L(t,(\gamma(t),\dot\gamma(t)))_{vert}$ is C^1. As L is C^2 we conclude that Δ is C^1, hence $\dot\gamma$ is C^1 and γ is C^2. In order to see that γ defines a T-periodic C^2-function we have to check $\frac{D}{dt}\dot\gamma(0) = \frac{D}{dt}\dot\gamma(T)$.

Observe that Δ and Δ^{-1} preserve the fibres of TM, hence their derivatives map vertical vector fields on TM in vertical vector fields. Now the covariant derivative $\frac{D}{dt}\dot\gamma(t)$ is the vertical part of the derivative $\frac{d}{dt}\dot\gamma(t) \in T_{(\gamma(t),\dot\gamma(t))}TM$, hence with

$z(t) = (\gamma(t), \dot\gamma(t)) \in T_{\gamma(t)}M$

$$\begin{aligned}
\frac{D}{dt}\dot\gamma &= \frac{D}{dt}[\pi_2\Delta^{-1}(t, (\gamma(t), \nabla_V L(t, z(t))_{vert}))] \\
&= \left\{\frac{d}{dt}[\pi_2\Delta^{-1}(t, (\gamma(t), \nabla_V L(t, (z(t))_{vert}))]\right\}_{vert} \\
&= \left\{\pi_2 D\Delta|^{-1}_{(t,z(t))}(1, \frac{d}{dt}(\gamma(t), \nabla_V L(t, z(t))_{vert}))\right\}_{vert} \\
&= \left\{\pi_2 D\Delta|^{-1}_{(t,z(t))}(1, vl_{z(t)}(\frac{D}{dt}(\nabla_V L(t, z(t))_{vert}))\right\}_{vert} \\
&= \left\{\pi_2 D\Delta|^{-1}_{(t,z(t))}(1, vl_{z(t)}(\nabla_H L(t, z(t))_{hor})\right\}_{vert} .
\end{aligned} \tag{B.8}$$

As the right hand side is T-periodic, the left hand side must be T-periodic as well.

Now from (B.8) we easily obtain that γ is a T-periodic C^k-function whenever L is C^k and thus $D\Delta$ is of class C^{k-2}. □

Remark B.1 Please note that (B.8) is the explicit second order system that corresponds to solutions of our variational problem. For a classical Lagrangian $L(t, (q, v)) = \frac{1}{2}\|v\|_q^2 - V(q)$ the diffeomorphism Δ is the identity map, and we have the classical equation

$$\frac{D}{dt}\dot\gamma(t) = -\nabla_{\gamma(t)}(V) . \tag{B.9}$$

Bibliography

[AB82] Michael Atiyah and Raoul Bott. On the Yang-Mill equations over Riemann surfaces. *Philos. Trans. Roy. Soc. London Ser. A*, 308:523–614, 1982.

[Abb03] Alberto Abbondandolo. On the Morse index of Lagrangian systems. *Nonlinear Anal.*, 53(3-4):551–566, 2003.

[AF07] Alberto Abbondandolo and Alessio Figalli. High action orbits for Tonelli Lagrangians and superlinear Hamiltonians on compact configuration spaces. *J. Differential Equations*, 234(2):626–653, 2007.

[AM78] Ralph Abraham and Jerrold E. Marsden. *Foundations of mechanics. 2nd ed., rev., enl., and reset. With the assistance of Tudor Ratiu and Richard Cushman.* Benjamin/Cummings Publishing Company, Reading, MA., 1978.

[AP93] Christopher Allday and Volker Puppe. *Cohomological methods in transformation groups.* Cambridge University Press, Cambridge, 1993.

[AS06] Alberto Abbondandolo and Matthias Schwarz. On the Floer homology of cotangent bundles. *Comm. Pure Appl. Math.*, 59(2):254–316, 2006.

[AS09] Alberto Abbondandolo and Matthias Schwarz. A smooth pseudo-gradient for the Lagrangian action functional. *Adv. Nonlinear Stud.*, 9(4):597–623, 2009.

[Bar93] Thomas Bartsch. *Topological Methods for Variational Problems with Symmetries*, volume 1560 of *Lecture Notes in Math.* Springer, Berlin, 1993.

[Ben86] Vieri Benci. Periodic solutions of Lagrangian systems on a compact manifold. *J. Differential Equations*, 63:135–161, 1986.

[Bre67] Glen E. Bredon. *Sheaf theory.* McGraw-Hill Book Company, New York-St. Louis-San Francisco-Toronto-London-Sydney, 1967.

[Bre72] Glen E. Bredon. *Introduction to Compact Transformation Groups.* Academic Press, London, 1972.

[BT82] Raoul Bott and Loring W. Tu. *Differential forms in algebraic topology.* Graduate Texts in Mathematics, 82, Springer, New York, 1982.

[BW97a] Thomas Bartsch and Zhi-Qiang Wang. Even Hamiltonian systems on tori and cotangent spaces of tori. *Nonlinear Anal., Theory Methods Appl.*, 30(7):4123–4134, 1997.

[BW97b] Thomas Bartsch and Zhi-Qiang Wang. Periodic solutions of even Hamiltonian systems on the torus. *Math. Z.*, 224(1):65–76, 1997.

[BW97c] Thomas Bartsch and Zhi-Qiang Wang. Periodic solutions of spatially periodic, even Hamiltonian systems. *J. Differential Equations*, 135(1):103–128, 1997.

[CG96] Kung-Ching Chang and Nassif Ghoussoub. The Conley index and the critical groups via an extension of Gromoll-Meyer theory. *Topol. Methods Nonlinear Anal.*, 7(1):77–93, 1996.

[Cha93] Kung-Ching Chang. *Infinite Dimensional Morse Theory and Multiple Solution Problems*, volume 6 of *Progr. Nonlin. Diff Eq. Appl.* Birkhäuser, Boston, 1993.

[CM00] Mónica Clapp and Wacław Marzantowicz. Essential equivariant maps and Borsuk-Ulam theorems. *J. Lond. Math. Soc., II. Ser.*, 61(3):950–960, 2000.

[CP91] Mónica Clapp and Dieter Puppe. Critical point theory with symmetries. *J. Reine Angew. Math.*, 418:1–29, 1991.

[CZ83] Charles C. Conley and Eduard Zehnder. The Birkhoff-Lewis fixed point theorem and a conjecture of V. Arnold. *Invent. Math.*, 73:33–49, 1983.

[dC92] Manfredo Perdigão do Carmo. *Riemannian geometry.* Birkhäuser, Boston, MA, 1992.

[Die60] Jean Dieudonné. *Foundations of modern analysis.* Pure and Applied Mathematics, Vol. X. Academic Press, New York, 1960.

[DK01] James F. Davis and Paul Kirk. *Lecture notes in algebraic topology.* AMS Graduate Studies in Mathematics. 35. Providence, RI , 2001.

[Dol63] Albrecht Dold. Partitions of unity in the theory of fibrations. Proc. Int. Congr. Math. 1962, 459-461 (1963)., 1963.

[Dug73] James Dugundji. *Topology.* Allyn and Bacon, Boston, 1973.

[Fad92] Sufian Fadell, Edward Husseini. Category Weight and Steenrod Operations. *Boletín de la Sociedad Matemática Mexicana*, 37, 1992.

[Fel92] Patricio L. Felmer. Periodic solutions of spatially periodic Hamiltonian systems. *J. Differential Equations*, 98:143–168, 1992.

[FHT01] Yves Félix, Stephen Halperin, and Jean-Claude Thomas. *Rational homotopy theory*, volume 205 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2001.

[Ful05] William Fulton. *Equivariant Cohomology in Algebraic Geometry.* Lecture Notes, Web version by D. Anderson, 2005.

[Hat02] Allen Hatcher. *Algebraic topology.* Cambridge University Press, Cambridge, 2002.

[HP57] Einar Hille and Ralph S. Phillips. *Functional analysis and semi-groups.* American Mathematical Society Colloquium Publications, vol. 31. American Mathematical Society, Providence, R. I., 1957. rev. ed.

[Hub61] Peter J. Huber. Homotopical cohomology and Čech cohomology. *Math. Ann.*, 144:73–76, 1961.

[Hus93] Dale H. Husemoller. *Fibre bundles. 3rd ed.* Springer, Berlin, 1993.

[HW48] Witold Hurewicz and Henry Wallman. *Dimension theory.* Princeton University Press, Princeton, 1948.

[Ill83] Sören Illman. The equivariant triangulation theorem for actions of compact Lie groups. *Math. Ann.*, 262:487–501, 1983.

[Kli78] Wilhelm Klingenberg. *Lectures on Closed Geodesics.* Grundlehren. Springer, Berlin Heidelberg New York, 1978.

[LN83] Rudolf Lidl and Harald Niederreiter. *Finite fields*, volume 20 of *Encyclopedia of Mathematics and its Applications.* Addison-Wesley, Reading, MA, 1983.

[Lu09] Guangcun Lu. The Conley conjecture for Hamiltonian systems on the cotangent bundle and its analogue for Lagrangian systems. *J. Funct. Anal.*, 256(9):2967–3034, 2009.

[Mor63] Kiiti Morita. On the product of paracompact spaces. *Proc. Japan Acad.*, 39:559–563, 1963.

[Mos57] George D. Mostow. Equivariant embeddings in Euclidian space. *Ann. of Math. (2)*, 65:432–446, 1957.

[MS80] John Douglas Moore and Roger Schlafly. On equivariant isometric embeddings. *Math. Z.*, 173:119–133, 1980.

[MW89] Jean Mawhin and Michel Willem. *Critical Point Theory and Hamiltonian Systems*, volume 74 of *Appl. Math. Sci.* Springer, New York Berlin Heidelberg London Paris Tokyo, 1989.

[PT88] Richard S. Palais and Chuu-Lian Terng. *Critical point theory and submanifold geometry*, volume 1353 of *Lecture Notes in Mathematics*. Springer, Berlin, 1988.

[Qui71] Daniel Quillen. The spectrum of an equivariant cohomology ring. I. II. *Ann. Math.*, 94, 1971.

[Rab88] Paul H. Rabinowitz. On a class of functionals invariant under a $\mathbb{Z}^n$-action. *Trans. Amer. Math. Soc.*, 310:303–311, 1988.

[Rud98] Yuli B. Rudyak. Category weight: new ideas concerning Lusternik-Schnirelmann category. In *Homotopy and geometry (Warsaw, 1997)*, volume 45 of *Banach Center Publ.*, pages 47–61. Polish Acad. Sci., Warsaw, 1998.

[Rud99] Yuli B. Rudyak. On category weight and its applications. *Topology*, 38(1):37–55, 1999.

[Spa66] Edwin H. Spanier. *Algebraic Topology.* McGraw Hill, New York, 1966.

[Str90] Michael Struwe. *Variational Methods and Their Applications to Non-Linear Partial Differential Equations and Hamiltonian Systems.* Springer, Berlin, 1990.

[Str02] Jeffrey Strom. Generalizations of category weight. In *Lusternik-Schnirelmann category and related topics (South Hadley, MA, 2001)*, volume 316 of *Contemp. Math.*, pages 193–203. Amer. Math. Soc., Providence, RI, 2002.

[Tak02] Masamichi Takesaki. *Theory of operator algebras. I*, volume 124 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, Berlin, 2002. Reprint of the first (1979) edition, Operator Algebras and Non-commutative Geometry, 5.

[tD87] Tammo T. Dieck. *Transformation Groups.* Walter de Gruyter, Berlin-New York, 1987.

[Whi78] George W. Whitehead. *Elements of Homotopy Theory.* Springer, New York, 1978.

[Yam74] Sadayuki Yamamuro. *Differential calculus in topological linear spaces.* Lecture Notes in Mathematics, Vol. 374. Springer, Berlin, 1974.

[Zet97] Anton Zettl. Sturm-Liouville problems. In *Spectral theory and computational methods of Sturm-Liouville problems (Knoxville, TN, 1996)*, volume 191 of *Lecture Notes in Pure and Appl. Math.*, pages 1–104. Dekker, New York, 1997.