Local gluing

Urs Frauenfelder
Universität Augsburg

Joa Weber*
UNICAMP

January 16, 2024

Abstract

In the local gluing one glues local neighborhoods around the critical point of the stable and unstable manifolds to gradient flow lines defined on a finite time interval $[-T, T]$ for large T. If the Riemannian metric around the critical point is locally Euclidean, the local gluing map can be written down explicitly. In the non-Euclidean case the construction of the local gluing map requires an intricate version of the implicit function theorem.

In this paper we explain a functional analytic approach how the local gluing map can be defined. For that we are working on infinite dimensional path spaces and also interpret stable and unstable manifolds as submanifolds of path spaces. The advantage of this approach is that similar functional analytical techniques can as well be generalized to infinite dimensional versions of Morse theory, for example Floer theory.

A crucial ingredient is the Newton-Picard map. We work out an abstract version of it which does not involve troublesome quadratic estimates.

Contents

*Email: urs.frauenfelder@math.uni-augsburg.de joa@unicamp.br

1 Introduction and main results

1.1 Local gluing map for the Euclidean metric

Consider a diagonal matrix with monotone decreasing diagonal entries

$$A = \operatorname{diag}(a_1, \dots, a_n), \qquad a_1 \geq \dots \geq a_{n-k} > 0 > a_{n-k+1} \geq \dots \geq a_n.$$

Consider the smooth function given by the euclidean inner product

$$f\colon \mathbb{R}^n \to \mathbb{R}, \quad z \mapsto \tfrac{1}{2} \langle z, Az \rangle . \tag{1.1}$$

This function is Morse and has a unique critical point at the origin of Morse index k. The gradient of f for the standard metric on $\mathbb{R}^n$ is $\nabla f(z) = Az$. Hence the downward gradient flow for time s is given by

$$\varphi_s^{-\nabla f}(z) = e^{-sA} z = \left(e^{-sa_1} z_1, \dots, e^{-sa_n} z_n \right) .$$

The stable and the unstable manifold of the origin are given by the sets

$$W^{\mathrm{s}} = \mathbb{R}^{n-k} \times \{0\}, \qquad W^{\mathrm{u}} = \{0\} \times \mathbb{R}^k .$$

Each point $z_0 = (x_0, 0) \in \mathbb{R}^{n-k} \times \{0\}$ determines an element $s \mapsto w_+(s) := e^{-sA} z_0$ in the function space $W^{1,2}([0,\infty), \mathbb{R}^n)$. Each point $z_0 = (0, y_0) \in \{0\} \times \mathbb{R}^k$ determines an element $s \mapsto w_-(s) := e^{-sA} z_0$ in the function space $W^{1,2}((-\infty,0], \mathbb{R}^n)$.

In our functional analytic approach to local gluing it is more convenient for us to think of the stable and the unstable manifold as function space subsets

$$\mathcal{W}^{\mathrm{s}} \subset W^{1,2}([0,\infty), \mathbb{R}^n), \qquad \mathcal{W}^{\mathrm{u}} \subset W^{1,2}((-\infty,0], \mathbb{R}^n).$$

A further advantage of this point of view is that many techniques discussed in this article can be generalized from $\mathbb{R}^n$ to the Hardy approach of gluing in the infinite dimensional case of Floer homology [Sim14].

With the interpretation of stable and unstable manifolds as function spaces we can easily recover the traditional interpretation as subsets of $\mathbb{R}^n$ using the evaluation maps

$$\mathrm{ev}_+ \colon \mathcal{W}^{\mathrm{s}} \to \mathbb{R}^{n-k} \times \{0\}, \quad w_+ \mapsto w_+(0)$$

and

$$\mathrm{ev}_- \colon \mathcal{W}^{\mathrm{u}} \to \{0\} \times \mathbb{R}^k, \quad w_- \mapsto w_-(0).$$

Given $T > 0$, let $\mathcal{M}_T \subset W^{1,2}([-T,T],\mathbb{R}^n)$ be the subset of all finite time gradient flow lines $w \colon [-T,T] \to \mathbb{R}^n$. Note that since in the euclidean case the gradient flow is linear and a gradient flow line is uniquely determined by its initial condition, the space $\mathcal{M}_T$ is an n-dimensional linear subspace of the infinite dimensional function space $W^{1,2}([-T,T],\mathbb{R}^n)$.

In the **euclidean case**, that is $\mathbb{R}^n$ endowed with the standard metric, there are natural linear isomorphisms

$$\Gamma_T \colon \mathcal{W}^{\mathrm{s}} \times \mathcal{W}^{\mathrm{u}} \to \mathcal{M}_T$$

called the **local gluing maps** and given at each time $s \in [-T,T]$ by

$$\Gamma_T(w_+,w_-)(s) = e^{-(s+T)A} w_+(0) + e^{(T-s)A} w_-(0). \tag{1.2}$$

Consider the evaluation map defined by

$$\mathrm{ev}_T \colon W^{1,2}([-T,T],\mathbb{R}^n) \to \mathbb{R}^n \times \mathbb{R}^n, \quad w \mapsto (w(-T), w(T)) .$$

The composition of the local gluing maps Γ_T with the evaluation map ev_T is a linear map, namely

$$\mathrm{ev}_T \circ \Gamma_T(w_+,w_-) = \left(w_+(0) + e^{2TA} w_-(0), w_-(0) + e^{-2TA} w_+(0) \right) .$$

Since w_- is in the unstable manifold and w_+ in the stable, both limits are zero

$$\lim_{T\to\infty} e^{2TA} w_-(0) = 0, \qquad \lim_{T\to\infty} e^{-2TA} w_+(0) = 0.$$

Therefore it holds that $\lim_{T\to\infty} \mathrm{ev}_T \circ \Gamma_T = \mathrm{ev}$ where

$$\mathrm{ev} = (\mathrm{ev}_+, \mathrm{ev}_-) \colon \mathcal{W}^{\mathrm{s}} \times \mathcal{W}^{\mathrm{u}} \to \mathbb{R}^n \times \mathbb{R}^n, \quad (w_+,w_-) \mapsto (w_+(0), w_-(0)) .$$

1.2 Local gluing map for a general Riemannian metric

Given a general Morse function f on a finite dimensional manifold, by the Morse Lemma one can always find locally around each critical point coordinates such that f has the form (1.1) after subtracting the critical value. In fact, it is even possible, after some additional scaling, to assume that all diagonal entries of the

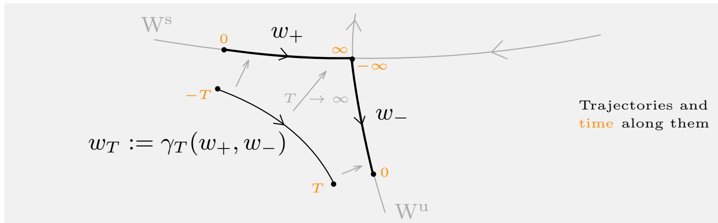

Figure 1: Convergence of local gluing $\mathrm{ev}_T \circ \gamma_T(w_+, w_-) \stackrel{T\to\infty}{\longrightarrow} \mathrm{ev}(w_+, w_-)$

matrix A are either 1 or -1. In infinite dimension this is usually not possible and therefore we don't use this fact.
Unfortunately, even in finite dimension, it is in general not possible to assume that in Morse coordinates the Riemannian metric is standard as well. Indeed curvature is an obstruction.

In this article we explain, based on a special version of Newton-Picard iteration, a functional analytic construction for **local gluing maps $\boldsymbol{\gamma_T}$** in the curved case. In sharp contrast to the Euclidean version Γ_T, the local gluing maps γ_T are in general not linear. However, still some of the major properties of the local gluing maps Γ_T in the flat case are preserved in the general case. More precisely, we have the following theorem.

Theorem A (Local gluing)**.** *There are open neighborhoods $\mathcal{U}_+$ and $\mathcal{U}_-$ of the origin in the stable and unstable manifold and gluing maps $\gamma_T\colon \mathcal{U}_+ \times \mathcal{U}_- \to \mathcal{M}_T$ for $T \ge T_0$, where $\mathcal{M}_T$ is the space of downward gradient flow lines on the finite time interval $[-T, T]$, which have the following properties.*

a) For every $T \ge T_0$ the gluing map γ_T is a diffeomorphism onto its image

b) In the limit $T \to \infty$ in the C^∞ topology the diagram

$$\begin{array}{ccccc} \mathcal{W}^{\mathrm{s}} \times \mathcal{W}^{\mathrm{u}} & \supset & \mathcal{U}_+ \times \mathcal{U}_- & \xrightarrow{\mathrm{ev}} & \mathbb{R}^n \times \mathbb{R}^n \\ & & {\scriptstyle\gamma_T} \searrow & & \nearrow {\scriptstyle\mathrm{ev}_T} \\ & & & \mathcal{M}_T & \end{array} \tag{1.3}$$

commutes, as illustrated by Figure 1, where ev and ev_T are the evaluation maps at the end points.

Remark 1.1. Our construction of local gluing maps γ_T has the following additional properties.

1. In the euclidean case it holds that $\gamma_T = \Gamma_T$.
2. In the general Riemannian case this still continues to hold for the differential of γ_T at the origin, in symbols $d\gamma_T(0_+, 0_-) = \Gamma_T$.

3. In particular, at the infinitesimal level, our construction is independent of any choices like the one of a cutoff function used to construct a pregluing map; see (2.11). The construction of the gluing map depends on the choice of a complement of the kernel $\mathbb{E}_T$ of the linearized gradient flow equation $D_T\colon \mathbb{W}_T \to \mathbb{V}_T$; see (4.39). There are different choices for such a complement. Possible choices are to take the complement orthogonal with respect to either the L^2 or the $W^{1,2}$ metric. We make a different choice, so that our complement $\mathbb{K}_T$ is not necessarily orthogonal, but instead has the property that the infinitesimal gluing map, see (3.26), does not depend on the choice of the cutoff function.

4. Furthermore, our construction uses a version of the Newton-Picard map which does not need quadratic estimates. We discuss properties of the Newton-Picard map and its derivatives in Appendix B.

   The results in Appendix B are quite general, so that they should also be applicable to the infinite dimensional version of the local gluing discussed in this article. Namely, the general Hardy approach to gluing, as discussed in the special case of Lagrangian Floer homology by Tatjana Simčević [Sim14].

We expect that the local gluing theorems will be useful for the construction of flow category theories [CJS95] by endowing, for Morse-Smale metrics, the moduli (solution) spaces of broken gradient flow lines with the structure of a manifold with boundary and corners [Qin18, Weh12].

1.3 Setup – path spaces and sections

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a smooth function such that the origin 0 is a Morse critical point of Morse index k. Suppose g is a Riemannian metric on $\mathbb{R}^n$ which is standard at 0, notation g_0. Let $\mathrm{Hess}_0 f$ be the Hessian bilinear form of f at 0. The Hessian linear operator $A\colon \mathbb{R}^n \to \mathbb{R}^n$ of f at 0 is defined with the help of the metric g_0 by the formula $\mathrm{Hess}_0 f(z_1, z_2) = g_0(z_1, A z_2)$ for every $z_1, z_2 \in \mathbb{R}^n$. After a linear change of coordinates we can assume that A is a diagonal matrix with monotone decreasing diagonal entries

$$A = \mathrm{diag}(a_1, \dots, a_n), \qquad a_1 \geq \dots \geq a_{n-k} > 0 > a_{n-k+1} \geq \dots \geq a_n. \tag{1.4}$$

Consider the g_0-orthogonal splitting

$$\mathbb{R}^n = \mathbb{R}^{n-k} \times \mathbb{R}^k \xrightarrow{p_-} \mathbb{R}^k, \quad (x, y) \mapsto y, \qquad p_+(x, y) := x. \tag{1.5}$$

Then the Hessian at 0 is positive definite on $E^+ = \mathbb{R}^{n-k} \times \{0\}$ and negative definite on $E^- = \{0\} \times \mathbb{R}^k$. The Hessian operator at 0 is of the form

$$A = \begin{pmatrix} A_+ & 0 \\ 0 & -A_- \end{pmatrix} \tag{1.6}$$

where $A_+ = \mathrm{diag}(a_1, \dots, a_{n-k})$ and $A_- = \mathrm{diag}(-a_{n-k+1}, \dots, -a_n)$ are positive definite diagonal matrices. The **spectral gap** $\sigma = \sigma(A) > 0$ is the smallest distance of an eigenvalue to the origin, in symbols

$$\sigma = \sigma(A) := \min_{1 \le \ell \le n} |a_\ell|. \tag{1.7}$$

Abbreviate $\mathbb{R}_+ = (0, \infty)$. For $T > 0$ consider the Sobolev spaces

$$\mathbb{W}_+ = W^{1,2}([0,\infty), \mathbb{R}^n),\ \mathbb{W}_- = W^{1,2}((-\infty,0], \mathbb{R}^n),\ \mathbb{W}_T = W^{1,2}([-T,T], \mathbb{R}^n),$$

$$\mathbb{V}_+ = L^2([0,\infty), \mathbb{R}^n), \quad \mathbb{V}_- = L^2((-\infty,0], \mathbb{R}^n), \quad \mathbb{V}_T = L^2([-T,T], \mathbb{R}^n).$$

Definition 1.2 (Constant maps to the critical point)**.** Let $0_+ \in \mathbb{W}_+$ and $0_- \in \mathbb{W}_-$, and $0_T \in \mathbb{W}_T$, be the constant maps to the critical point, in symbols

$$0_+ \colon [0,\infty) \to \mathbb{R}^n, \quad s \mapsto 0, \qquad 0_- \colon (\infty, 0] \to \mathbb{R}^n, \mapsto s \mapsto 0.$$

Let $0_T \in \mathbb{W}_T$ be the constant map $[-T,T] \to \mathbb{R}^n$, $s \mapsto 0$, to the critical point.

For $i \in \{+,-\} \cup \mathbb{R}_+$ consider the map defined by

$$\mathcal{F}_i \colon \mathbb{W}_i \to \mathbb{V}_i, \quad w \mapsto \partial_s w + \nabla f(w).$$

The zero sets of these maps are, respectively, the stable and the unstable manifold, and the set of gradient flow lines along the interval $[-T,T]$, in symbols

$$\mathcal{W}^{\mathrm{s}} := {\mathcal{F}_+}^{-1}(0_+) \subset \mathbb{W}_+, \qquad \mathcal{W}^{\mathrm{u}} := {\mathcal{F}_-}^{-1}(0_-) \subset \mathbb{W}_-,$$

and the **solution space**

$$\mathcal{M}_T := {\mathcal{F}_T}^{-1}(0_T) = \{w \colon [-T,T] \overset{W^{1,2}}{\to} \mathbb{R}^n \mid \partial_s w + \nabla f(w) = 0\} \subset \mathbb{W}_T.$$

The elements of the tangent spaces at the critical point

$$\xi \in \mathbb{E}^+ := T_{0_+}\mathcal{W}^{\mathrm{s}}, \qquad \eta \in \mathbb{E}^- := T_{0_-}\mathcal{W}^{\mathrm{u}}, \qquad \zeta \in \mathbb{E}_T := T_{0_T}\mathcal{M}_T,$$

are characterized by the linear autonomous ODEs (2.15) or, equivalently, by forward (backward) exponential decay (2.17) of $\xi = (\xi^+, 0)$ (of $\eta = (0, \eta^-)$).

Notation. The euclidean norm of $v \in \mathbb{R}^\ell$, $\ell \in \mathbb{N}$, is denoted by $|v|$.

1.4 Idea of proof

We construct the desired gluing map as a family of diffeomorphisms onto their images, one diffeomorphism for each $T \ge T_0$ given by composing two maps

$$\gamma_T \colon \mathcal{W}^{\mathrm{s}} \times \mathcal{W}^{\mathrm{u}} \supset \mathcal{U}_+ \times \mathcal{U}_- \xrightarrow{\wp_T} \mathbb{W}_T \xrightarrow{\mathcal{N}_T} \mathcal{M}_T. \tag{1.8}$$

Here $T_0 \ge 3$ is a constant and $\mathcal{U}_+ \subset \mathcal{W}^{\mathrm{s}}$ and $\mathcal{U}_- \subset \mathcal{W}^{\mathrm{u}}$ are open neighborhoods of 0_+ and 0_-, respectively, sufficiently small so that the image $\wp_T(\mathcal{U}_+ \times \mathcal{U}_-)$ of the pre-gluing map $\wp_T$ lies in the domain of the Newton-Picard map $\mathcal{N}_T$.

Newton-Picard. The Newton-Picard map on $X = \mathbb{W}_T$ associates to an approximate zero of a map, here $\mathcal{F}_T$, a true zero nearby. More precisely, after choosing a suitable initial point x_0, here 0_T, there are three ingredients needed:

1) an approximate zero x_1 of $\mathcal{F}_T$;
2) a uniformly bounded right inverse Q_T of $D_T := d\mathcal{F}_T(0_T)\colon \mathbb{W}_T \to \mathbb{V}_T$;
3) a slowly varying operator difference $d\mathcal{F}_T(\cdot) - D_T$ near the initial point.

The facts that $\mathcal{F}_T(0_T) = 0$ and that D_T is surjective suggest to choose as initial point $x_0 := 0_T$. 1) To provide an approximate zero of $\mathcal{F}_T$ will be the task of the pre-gluing map as described further below. 2) Right inverses of the linear operator $D_T\colon \mathbb{W}_T \to \mathbb{V}_T$ correspond to the topological complements of $\ker D_T$. A natural choice would be the orthogonal complement, but we shall choose another complement, notation $\mathbb{K}_T$, which represents the impossible paths for a downward gradient and makes the infinitesimal gluing map $\Gamma_T = d\gamma_T(0_+, 0_-)$ independent of the choice of cutoff function used to define the pre-gluing map. The corresponding right inverse Q_T indeed admits a uniform bound c. 3) The operator difference $d\mathcal{F}_T(\cdot) - D_T$ is usually controlled by calculating troublesome quadratic estimates. In Appendix B.1 we prove continuous differentiability of a version of the Newton-Picard map which does not require quadratic estimates.

Remark 1.3 (Higher smoothness of Newton-Picard map)**.** To obtain higher smoothness we use, roughly speaking, the fact that the supremum of the operator norm $\|d\mathcal{F}_T(\cdot) - D_T\|$ along smaller and smaller balls about the initial point x_0 admits bounds closer and closer to zero. Indeed there is a monotonically decreasing function $\delta\colon [2,\infty) \to (0,\infty)$, independent of T, such that along the $\delta(\mu)$-ball about x_0 the map $\|d\mathcal{F}_T(\cdot) - D_T\|$ is bounded by $1/\mu c$. See Corollary 4.5 for the case of $\mathcal{F}_T$ and Remark B.8 for the abstract theory.

For iteration arguments, such as to prove higher smoothness, tangent maps are much more suitable than differentials. Thus we prove in Appendix B.2 an estimate for the tangent map difference $T\mathcal{N} - \mathrm{Id}$ and we show that $T\mathcal{N}^F = \mathcal{N}^{TF}$, roughly speaking, where $\mathcal{N}^F$ is the Newton-Picard map for a map F.

Pre-gluing – approximate zero. Given a real $T \geq 3$, called *gluing parameter*, Floer's gluing construction associates to a pair $(w_+, w_-) \in \mathcal{W}^{\mathrm{s}} \times \mathcal{W}^{\mathrm{u}}$ of an (incoming, outgoing) flow trajectory the *pre-glued path* $w_T\colon [-T,T] \to \mathbb{R}^m$ defined as follows. One decomposes the time interval $[-T,T]$ into five subintervals. Along $[-T,-3]$ follow the backward shifted forward flow trajectory $w_+(T+\cdot)$, then along $[-3,-1]$ interpolate with the help of a cutoff function to the constant flow trajectory $s \mapsto 0$ sitting at the critical point at which w_T then rests along time $[-1,1]$. Next along time $[1,3]$ interpolate from the constant map $s \mapsto 0$ to the forward shifted backward flow trajectory $w_-(-T+\cdot)$ which then represents w_T along the final time interval $[3,T]$.

The behavior of the pre-glued path $w_T\colon [-T,T] \to \mathbb{R}^n$ along the five time intervals is detailed by formula (2.12) and illustrated by Figure 2. Observe that w_T takes on the boundary of its domain $[-T,T]$ values that do not depend on T, namely $w_+(0)$ and $w_-(0)$. Most importantly, the pre-glued path satisfies the gradient equation except, possibly, along the subinterval $[-3,-1]$ (and $[1,3]$) of

$[-T,T]$ along which it coincides, up to a cutoff function factor, with the forward flow trajectory w_+ along $[T-3,T-1]$. But $w_+|_{[T-3,T-1]}$ is very close to the critical point for large T. Consequently uniform exponential decay of $\partial_s w_+$ takes care of the L^2 norm of $\mathcal{F}_T(w_T)=\partial_s w_T+\nabla f(w_T)$ along $[-3,-1]$; same along $[1,3]$ where w_- appears. With this understood it follows that w_T is an approximate zero of $\mathcal{F}_T$ in the sense that

$$\|\mathcal{F}_T(w_T)\|_{\mathcal{V}_T}\le Ce^{-\varepsilon T} \tag{1.9}$$

whenever $T\ge 3$. The constant C serves all elements $w_\pm$ of any chosen pair of compact neighborhoods $K_\pm$ of $0_\pm$ in the stable/unstable manifold.

Gluing – smooth convergence. With Newton-Picard and pre-gluing in place the gluing map γ_T, given by composition (1.8), is well defined. Appendix A revisits the proof of the usual IFT explained in [MS04, App. A.3] to extract a quantitative version. It is applied in Section 5.1 to prove that γ_T is a diffeomorphism onto its image along a sufficiently small domain, uniformly in T.

In the limit $T\to\infty$ the diagram (1.3) commutes even after application of the m-fold tangent functor T^m. The proof uses techniques described by Remark 1.3 and is carried out in Section 5.2.

Outline of article

Section 2 "Pre-gluing map $\mathcal{P}_T$ and its restriction $\wp_T$" introduces for each parameter value $T\ge 3$ the pre-gluing map as the *linear* map $\mathcal{P}_T\colon \mathbb{W}_+\times\mathbb{W}_-\to\mathbb{W}_T$ defined by (2.11), equivalently by (2.12), and illustrated by Figure 2.

For $0_+\in\mathcal{W}^{\rm s}$ and $0_-\in\mathcal{W}^{\rm u}$ the pre-glued path is a true zero, more precisely $\mathcal{P}_T(0_+,0_-)=0_T\in\mathcal{F}_T^{-1}(0)\subset\mathbb{W}_T$. This motivates the expectation that pre-gluing pairs near $(0_+,0_-)\in\mathcal{W}^{\rm s}\times\mathcal{W}^{\rm u}$ should produce approximate zeroes. Thus we consider the **restriction** of the pre-gluing map $\mathcal{P}_T$, notation

$$\wp_T:=\mathcal{P}_T|_{\mathcal{W}^{\rm s}\times\mathcal{W}^{\rm u}}\colon \mathcal{W}^{\rm s}\times\mathcal{W}^{\rm u}\to\mathbb{W}_T. \tag{1.10}$$

This map is smooth by linearity of $\mathcal{P}_T$. Whereas the elements of the tangent spaces to $\mathcal{W}^{\rm s}$, $\mathcal{W}^{\rm u}$, and $\mathcal{M}_T$ at the origins 0_+, 0_-, and 0_T (notation $\mathbb{E}^\pm$ and $\mathbb{E}_T$) are the solutions of autonomous linear ODEs (2.15), at general points w_+, w_-, and w_T the characterizing linear ODE's (2.18) are non-autonomous.[1]

The linear identifications $\theta_{w_\pm}\colon T_{w_\pm}\mathcal{W}^{\rm s/u}\to\mathbb{E}^\pm$, defined via asymptotic limits, are used to prove Theorem 5.3 (gluing map γ_T is diffeomorphism onto its image).

Section 3 "Infinitesimal gluing" consists of two subsections. Subsection 3.1 introduces a complement $\mathbb{K}_T$ of the n-dimensional subspace $\mathbb{E}_T:=T_{0_T}\mathcal{M}_T$ of $\mathbb{W}_T$ and the corresponding projection $P_{\mathbb{E}_T,\mathbb{K}_T}$ onto $\mathbb{E}_T$ along $\mathbb{K}_T$, notation

$$\Pi_T:=P_{\mathbb{E}_T,\mathbb{K}_T}\colon \mathbb{W}_T=\mathbb{K}_T\oplus\mathbb{E}_T\to\mathbb{E}_T.$$

[1] whenever the Hessian operators $A_w(s)$ along a flow trajectory w depend on time s

Lemma 3.3 provides a formula for Π_T and asserts that the operator norm of $\Pi_T\colon \mathbb{W}_T \to \mathbb{W}_T$ is bounded by a constant $d = d(a_1, a_n)$, depending on the eigenvalues a_1 and a_n of the Hessian A in (1.4), but independent of $T \geq 1$. To prove this we establish the uniform-in-T Sobolev estimate $\|v\|_{L^\infty[-T,T]} \leq 2\|v\|_{W^{1,2}[-T,T]}$. Later on the estimate also enters the proof of Corollary 4.5 on existence of the monotone function $\delta(\mu)$ mentioned in Remark 1.3 on higher smoothness of the Newton-Picard map.
Subsection 3.2 introduces the infinitesimal gluing map, namely the linear map

$$\Gamma_T := \Pi_T \circ d\wp_T(0_+, 0_-)\colon \mathbb{E}^+ \times \mathbb{E}^- \to \mathbb{W}_T \to \mathbb{E}_T.$$

For Γ_T we obtain formula (3.29) which, firstly, by choice of $\mathbb{K}_T$, does not depend on the choice of cutoff function β in the pre-gluing map (2.11) and, secondly, reproduces the gluing map (1.2) in the Euclidean model case. Lemma 3.5 asserts that Γ_T is an isomorphism with inverse bounded by the constant $k := 1/(1 - e^{-12\sigma})$, independent of T, where σ is the spectral gap (1.7) of the Hessian A.

Section 4 "Newton-Picard map" consists of three subsections in which we verify the three ingredients 1) 2) 3) described earlier.
Subsection 4.1 shows 1) the pre-gluing provides an approximate zero $w_T := \wp_T(w_+, w_-)$ of $\mathcal{F}_T$ in the sense of (1.9). This hinges on Appendix C where we provide suitable exponential decay uniformly in T.
Subsection 4.2 shows that the linearization $D_T := d\mathcal{F}_T(0_T)\colon \mathbb{W}_T \to \mathbb{V}_T$ is surjective and 2) provides a bound $c = c(a_1, a_n)$ uniformly in T for the right inverse Q_T associated to the complement $\mathbb{K}_T$ of $\mathbb{E}_T$. Actually $\mathbb{E}_T = \ker D_T$.
Subsection 4.3 establishes 3) a bound on the difference $d\mathcal{F}_T(\cdot) - D_T$. Based on Proposition B.1 we define the Newton-Picard map $\mathcal{N}_T\colon \mathbb{W}_T \to \mathbb{W}_T$ along a neighborhood $U_0(\delta_4)$ of the initial point $x_0 := 0_T$. Then it is shown that for $T \geq 3$ pre-gluing map $\wp_T$ takes values in the domain of $\mathcal{N}_T$.

Section 5 "Gluing map" provides an open neighborhood $\mathcal{U}_+ \times \mathcal{U}_- \subset \mathcal{W}^{\rm s} \times \mathcal{W}^{\rm u}$ of the origin $(0_+, 0_-)$ which serves as domain for all gluing maps γ_T with gluing parameter $T \geq T_0$, see (5.50), and defined by pre-gluing $\wp_T$ followed by Newton-Picard zero detection $\mathcal{N}_T$, see (1.8). By Lemma 5.2 the linearized gluing map at the origin coincides with the infinitesimal gluing map Γ_T.
Subsection 5.1 "Diffeomorphism onto image" proves this property of the gluing maps along an open subset $\mathcal{O}_+ \times \mathcal{O}_- \subset \mathcal{U}_+ \times \mathcal{U}_-$, uniformly in T. This is an application of the quantitative inverse function Theorem A.1. Verification of (A.65) uses that $d\gamma_T|_{(0_+,0_-)}$ is the infinitesimal gluing map Γ_T (Lemma 5.2) and that Γ_T has an inverse bounded uniformly in T (Lemma 3.5). Verification of (A.66) uses Remark B.5 and B.8 on the linearized Newton-Picard map.
Subsection 5.2 "Evaluation maps and convergence in C^m" shows that in the limit $T \to \infty$ the diagram (1.3) commutes as illustrated by Figure 1.

The appendices provide abstract results which might be of general interest.
Appendix A is on the "Quantitative inverse function theorem".
Appendix B provides the "Newton-Picard map without quadratic estimates".
Appendix C "Exponential decay" proves such, uniformly in T, and for all time

derivatives. We use again the tangent map formalism for ease of induction. The proof is based on Lemma C.4 in which the exponential decay rate of η is inherited by ξ, as opposed to the original [RS01, Le. 3.1].

Acknowledgements. UF acknowledges support by DFG grant FR 2637/2-2.

2 Pre-gluing map $\mathcal{P}_T$ and its restriction $\wp_T$

Fix a cut-off function $\beta\colon \mathbb{R}\to[0,1]$, that is a smooth function such that $\beta(s)=0$ for $s\le -1$ and $\beta(s)=1$ for $s\ge 1$. For any real $T\ge 3$, the **gluing parameter**, the **pre-gluing map** is the <u>linear</u> bounded Hilbert space map defined by

$$\begin{aligned}\mathcal{P}_T\colon \mathbb{W}_+\times\mathbb{W}_-&\to \mathbb{W}_T\\ (w_+,w_-)&\mapsto \underbrace{(1-\beta(\cdot+2))\,w_+(T+\cdot)+\beta(\cdot-2)\,w_-(-T+\cdot)}_{=:w_T}.\end{aligned} \tag{2.11}$$

Lemma 2.1 (Uniform bound)**.** *There is a constant $b>0$, depending on the cut-off function β but not on T, such that $\|\mathcal{P}_T\|\le b$ for every $T\ge 3$.*

Proof. The shift map is an isometry in $W^{1,2}(\mathbb{R})$ and the cut-off function β is independent of T. □

The pre-gluing map has the two properties that, firstly, for times s on the boundary of $[-T,T]$ we have

$$\mathcal{P}_T(w_+,w_-)(-T)=w_+(0),\qquad \mathcal{P}_T(w_+,w_-)(T)=w_-(0),$$

and, secondly, during the time interval $[-1,1]$ the map rests in the critical point

$$\mathcal{P}_T(w_+,w_-)|_{[-1,1]}\equiv 0.$$

More precisely, for fixed $w_\pm$, the **pre-glued path $\boldsymbol{w_T}$** is of the form

$$\underbrace{\mathcal{P}_T(w_+,w_-)(s)}_{=:w_T(s)}=\begin{cases} w_+(T+s) & ,\ s\in[-T,-3]\\ (1-\beta(s+2))\,w_+(T+s) & ,\ s\in[-3,-1]\\ 0 & ,\ s\in[-1,1]\\ \beta(s-2)w_-(-T+s) & ,\ s\in[1,3]\\ w_-(-T+s) & ,\ s\in[3,T]\end{cases} \tag{2.12}$$

for $s\in[-T,T]$. The pre-glued path w_T for $w_\pm$ is illustrated by Figure 2.

Example 2.2 (Constant maps to the critical point)**.** It holds that

$$\mathcal{P}_T(0_+,0_-)=0_T. \tag{2.13}$$

Note that $0_+\in\mathcal{W}^{\mathrm{s}}$, $0_-\in\mathcal{W}^{\mathrm{u}}$, and $0_T\in\mathcal{M}_T$.

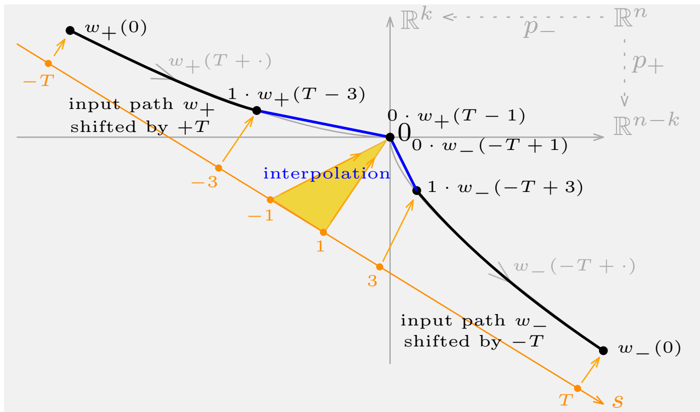

Figure 2: Pre-glued path $w_T(s) := \mathcal{P}_T(w_+, w_-)(s)$ for $s \in [-T, T]$

Restricted pre-gluing map

We denote the restriction of the pre-gluing map $\mathcal{P}_T$ to the stable and unstable manifolds $\mathcal{W}^{\rm s} \subset \mathbb{W}_+$ and $\mathcal{W}^{\rm u} \subset \mathbb{W}_-$ by

$$\wp_T := \mathcal{P}_T|_{\mathcal{W}^{\rm s}\times\mathcal{W}^{\rm u}} \colon \mathcal{W}^{\rm s} \times \mathcal{W}^{\rm u} \to \mathbb{W}_T \tag{2.14}$$

where $T \geq 3$. This map is smooth by linearity of $\mathcal{P}_T$.

Differential of $\wp_T$ at $(0_+, 0_-)$. Consider the tangent spaces to the trajectory spaces $\mathcal{W}^{\rm s}$, $\mathcal{W}^{\rm u}$, and $\mathcal{M}_T$, at the critical point, namely

$$\begin{aligned}
\mathbb{E}^+ &:= T_{0_+}\mathcal{W}^{\rm s} = \{\xi \in \mathbb{W}_+ \mid \partial_s\xi + A\xi = 0\} \subset C^\infty([0,\infty), \mathbb{R}^{n-k} \times \{0\})\\
\mathbb{E}^- &:= T_{0_-}\mathcal{W}^{\rm u} = \{\eta \in \mathbb{W}_- \mid \partial_s\eta + A\eta = 0\} \subset C^\infty((-\infty,0], \{0\} \times \mathbb{R}^{k})\\
\mathbb{E}_T &:= T_{0_T}\mathcal{M}_T = \{\zeta \in \mathbb{W}_T \mid \partial_s\zeta + A\zeta = 0\} \subset C^\infty([-T,T], \mathbb{R}^{n}).
\end{aligned} \tag{2.15}$$

By the theorem of Picard-Lindelöf the dimensions are given by

$$\dim \mathbb{E}^+ = n-k, \qquad \dim \mathbb{E}^- = k, \qquad \dim \mathbb{E}_T = n.$$

Then the linearization of $\wp_T$ at $(0_+, 0_-)$ is a map

$$d\wp_T(0_+, 0_-) \colon \mathbb{E}^+ \times \mathbb{E}^- \to \mathbb{W}_T.$$

For $\zeta \in \mathbb{E}^+, \mathbb{E}^-, \mathbb{E}_T$ we abbreviate

$$\zeta^+(s) := p_+(\zeta(s)), \quad \zeta^-(s) := p_-(\zeta(s)), \qquad \text{so } \zeta = (\zeta^+, \zeta^-). \tag{2.16}$$

Since A is diagonal, see (1.6), and by decay of the elements of $\mathbb{W}_\pm$, given maps $\xi \in C^\infty([0,\infty), \mathbb{R}^n)$ and $\eta \in C^\infty((-\infty,0], \mathbb{R}^n)$, there are the equivalences

$$\begin{aligned}
\xi \in \mathbb{E}^+ \quad &\Leftrightarrow \quad \xi = (\xi^+, 0) \quad \wedge \quad \xi^+(s) = e^{-sA_+}\xi^+(0)\ \forall s \geq 0,\\
\eta \in \mathbb{E}^- \quad &\Leftrightarrow \quad \eta = (0, \eta^-) \quad \wedge \quad \eta^-(s) = e^{+sA_-}\eta^-(0)\ \forall s \leq 0.
\end{aligned} \tag{2.17}$$

Differential of $\wp_T$ at $(w_+, w_-) \in \mathcal{W}^{\rm s} \times \mathcal{W}^{\rm u}$. The tangent spaces to the trajectory spaces $\mathcal{W}^{\rm s}$, $\mathcal{W}^{\rm u}$, and $\mathcal{M}_T$, at points w_+, w_-, and w_T, are

$$\begin{aligned} T_{w_+}\mathcal{W}^{\rm s} &= \{\xi \in \mathbb{W}_+ \mid \partial_s \xi + A_{w_+}\xi = 0\} \subset C^\infty([0,\infty),\mathbb{R}^n) \\ T_{w_-}\mathcal{W}^{\rm u} &= \{\eta \in \mathbb{W}_- \mid \partial_s \eta + A_{w_-}\eta = 0\} \subset C^\infty((-\infty,0],\mathbb{R}^n) \\ T_{w_T}\mathcal{M}_T &= \{\zeta \in \mathbb{W}_T \mid \partial_s \zeta + A_{w_T}\zeta = 0\} \subset C^\infty([-T,T],\mathbb{R}^n). \end{aligned} \tag{2.18}$$

Here, for $w \in \{w_+, w_-, w_T\}$, the family of **Hessian operators**

$$A_w = \{A_w(s)\}_s$$

is defined by the identities $\left(\mathrm{Hess}_{w(s)} f\right)(\cdot,\cdot) = g_{w(s)}\left(\cdot, A_w(s)\cdot\right)$, one identity for each s. There are canonical, continuous and linear, identifications[2]

$$\theta_{w_+} \colon \mathbb{W}_+ \supset T_{w_+}\mathcal{W}^{\rm s} \xrightarrow{\simeq} \mathbb{E}^+, \qquad \theta_{w_-} \colon \mathbb{W}- \supset T_{w_-}\mathcal{W}^{\rm u} \xrightarrow{\simeq} \mathbb{E}^-, \tag{2.19}$$

given by asymptotic limits where $\theta_{0_\pm} = \mathrm{Id}_{\mathbb{E}^\pm} =: \mathrm{Id}_\pm$ and the linear operators $\theta_{w_\pm}$ depend continuously on $w_\pm$; see [RS01, §3]. Since $\wp_T := \mathcal{P}_T|_{\mathcal{W}^{\rm s}\times\mathcal{W}^{\rm u}}$ is defined by restricting a linear map, the linearization is the linear map's restriction

$$d\wp_T(w_+, w_-) = d\left(\mathcal{P}_T|\right)(w_+, w_-) = \mathcal{P}_T| \colon T_{w_+}\mathcal{W}^{\rm s} \times T_{w_-}\mathcal{W}^{\rm u} \to \mathbb{W}_T. \tag{2.20}$$

Thus, given $(w_+, w_-) \in \mathcal{W}^{\rm s} \times \mathcal{W}^{\rm u}$, the map defined by

$$\begin{aligned} \Theta_T(w_+, w_-) \colon \mathbb{E}^+ \times \mathbb{E}^- &\to \mathbb{W}_T \\ (\xi, \eta) &\mapsto \mathcal{P}_T\left(\xi - \theta_{w_+}^{-1}\xi, \eta - \theta_{w_-}^{-1}\eta\right) \end{aligned} \tag{2.21}$$

is, by (2.20), equal to the difference

$$\begin{aligned} \Theta_T(w_+, w_-) &= \mathcal{P}_T\left((\mathrm{Id}_+, \mathrm{Id}_-) - (\theta_{w_+}^{-1}, \theta_{w_-}^{-1})\right) \\ &= d\wp_T|_{(0_+,0_-)} - d\wp_T|_{(w_+,w_-)} \circ (\theta_{w_+}^{-1}, \theta_{w_-}^{-1}). \end{aligned}$$

Lemma 2.3. *For any $\varepsilon > 0$ there are neighborhoods $\mathcal{O}_\varepsilon^+$ of 0_+ in $\mathcal{W}^{\rm s}$ and $\mathcal{O}_\varepsilon^-$ of 0_- in $\mathcal{W}^{\rm u}$ such that for every $T \geq 3$ and every $(w_+, w_-) \in \mathcal{O}_\varepsilon^+ \times \mathcal{O}_\varepsilon^-$ the operator norm of $\Theta_T(w_+, w_-) \colon \mathbb{E}^+ \times \mathbb{E}^- \to \mathbb{W}_T$ is less or equal ε.*

Proof. Lemma 2.1 and continuous dependence of $\theta_{w_\pm}$ on $w_\pm$ and $\theta_{0_\pm} = \mathrm{Id}_\pm$. □

[2] Observe that the elements

$$\xi \in \mathbb{E}^+ = T_{0_+}\mathcal{W}^{\rm s} \subset T_{0_+}\mathbb{W}_+ = \mathbb{W}_+, \qquad \theta_{w_+}\xi \in T_{w_+}\mathcal{W}^{\rm s} \subset T_{w_+}\mathbb{W}_+ = \mathbb{W}_+,$$

lie in the same ambient vector space $\mathbb{W}_+$, hence they can be added. Similarly $\eta \in \mathbb{E}^-$ and $\theta_{w_-}\eta$ both lie in $\mathbb{W}_-$.

3 Infinitesimal gluing

3.1 Projection associated to a particular complement

Complement $\mathbb{K}_T$ of n-dimensional linear solution space $\mathbb{E}_T$

For $T > 0$ we choose a Hilbert space complement $\mathbb{K}_T$ of $\mathbb{E}_T$ in $\mathbb{W}_T$ of the form

$$\mathbb{K}_T := \{\zeta = (\zeta^+, \zeta^-) \in \mathbb{W}_T \mid \zeta^+(-T) = 0,\ \zeta^-(T) = 0\}. \tag{3.22}$$

Note that the elements ζ of $\mathbb{K}_T$ start at points $\zeta(-T)$ in the negative definite space $\{0\} \times \mathbb{R}^k$ and end at points $\zeta(T)$ in the positive definite space $\mathbb{R}^{n-k} \times \{0\}$. Roughly speaking, the linear subspace $\mathbb{K}_T$ of $\mathbb{W}_T$ represents impossible paths for a downward gradient flow. The following lemma tells that $\operatorname{codim} \mathbb{K}_T = n$.

The complement $\mathbb{K}_T$ of $\mathbb{E}_T$ is not necessarily orthogonal. But it has the useful property that the infinitesimal gluing map Γ_T in (3.26) will not depend on the cutoff function β that was used to define the pre-gluing map $\mathcal{P}_T$ in (2.11).

Lemma 3.1 (Complement)**.** *a)* $\mathbb{K}_T \cap \mathbb{E}_T = \{0\}$ *and b)* $\mathbb{W}_T = \mathbb{K}_T + \mathbb{E}_T$.

Proof. a) Pick $\zeta = (\zeta^+, \zeta^-) \in \mathbb{E}_T$. Then $A(0, \zeta^-) = (0, -A_-\zeta^-)$ and $A(\zeta^+, 0) = (A_+\zeta^+, 0)$. Hence $\partial_s \zeta^- = A_-\zeta^-$ and $\partial_s \zeta^+ = -A_+\zeta^+$, and therefore

$$\zeta^-(s) = e^{(s-T)A_-}\zeta^-(T), \qquad \zeta^+(s) = e^{(-s-T)A_+}\zeta^+(-T),$$

for every $s \in [-T, T]$. Let $\zeta = (\zeta^+, \zeta^-) \in \mathbb{K}_T \cap \mathbb{E}_T$. Then $\zeta^+(-T) = 0$ and $\zeta^-(T) = 0$, so $\zeta^+ \equiv 0$ and $\zeta^- \equiv 0$ since each one solves a first order ODE. Hence $\zeta = (\zeta^+, \zeta^-) \equiv 0$. b) Let $\zeta \in \mathbb{W}_T$. Then the map defined by $Z(s) := \left(e^{(-s-T)A_+} p_+ \zeta(-T), e^{(s-T)A_-} p_- \zeta(T)\right)$ is element of $\mathbb{E}_T$ and the difference $\zeta - Z$ lies in $\mathbb{K}_T$ since $p_+(\zeta - Z)(-T) = 0$ and $p_-(\zeta - Z)(T) = 0$. □

Uniform Sobolev estimate

Lemma 3.2. *Let $T \geq 1$. Then any $v \colon [-T, T] \to \mathbb{R}$ of class $W^{1,2}$ satisfies*[3]

$$\|v\|_\infty \leq 2\|v\|_{1,2} \tag{3.23}$$

where the norms are over the domain $[-T, T]$.

Estimate (3.23) continues to hold for vector-valued maps $v \colon [-T, T] \to \mathbb{R}^\ell$ of class $W^{1,2}$ since

$$\|(v_1, \dots, v_\ell)\|_\infty = \max_{i=1,\dots,\ell} \|v_i\|_\infty \leq 2 \max_{i=1,\dots,\ell} \|v_i\|_{1,2} \leq 2\|v\|_{W^{1,2}([-T,T],\mathbb{R}^\ell)}.$$

Proof of Lemma 3.2. The proof has 4 steps.

Step 1. Let $T > 0$. Suppose $\|v\|_{1,2} \leq 1$ and at $s_0 \in [-T, T]$ we have $\kappa := |v(s_0)| > 0$. Then for $s \in [-T, T] \cap [s_0 - \kappa^2/4, s_0 + \kappa^2/4]$ it holds $|v(s)| \geq \kappa/2$.

[3] In [FW22a, (4.57)] we proved the case $T = \infty$ with constant 1, not 2.

Pointwise at s we have

$$
\begin{aligned}
|v(s)| &= \left| v(s_0) + \int_{s_0}^{s} v'(\sigma)\cdot 1\, d\sigma \right| \\
&\ge |v(s_0)| - \sqrt{\int_{s_0}^{s} \left(v'(\sigma)\right)^2\, d\sigma}\sqrt{\int_{s_0}^{s} 1\, d\sigma} \\
&\ge \kappa - \underbrace{\left(\int_{-T}^{T} \left(v'(\sigma)\right)^2\, d\sigma\right)}_{\le \|v\|_{1,2}^2 \le 1}^{1/2} \cdot \sqrt{|s-s_0|} \\
&\ge \kappa - \sqrt{|s-s_0|} \\
&\ge \kappa - \sqrt{\kappa^2/4} = \kappa/2.
\end{aligned}
$$

This proves Step 1.

Step 2. Under the assumption of Step 1 suppose, in addition, the inclusion $[s_0, s_0+1] \subset [-T,T] \cap [s_0 - \kappa^2/4, s_0 + \kappa^2/4]$, then $|v(s_0)| = 2$.

To prove Step 2 use Step 1 to obtain that

$$
1 \ge \|v\|_{1,2}^2 \ge \|v\|_2^2 \overset{\text{inclusion}}{\ge} \int_{s_0}^{s_0+1} \underbrace{v(\sigma)^2}_{\overset{\text{St.1}}{\ge} \kappa^2/4}\, d\sigma \ge \kappa^2/4.
$$

Therefore $2 \ge \kappa$. In view of the inclusion this implies $2 = \kappa := |v(s_0)|$.

Step 3. Under the assumption of Step 1 suppose, in addition, the inclusion $[s_0 - 1, s_0] \subset [-T,T] \cap [s_0 - \kappa^2/4, s_0 + \kappa^2/4]$, then $|v(s_0)| = 2$.

The same argument as in Step 2 proves Step 3.

Step 4. Let $T \ge 1$. If $\|v\|_{1,2} \le 1$, then $|v(s_0)| \le 2$ for every $s_0 \in [-T,T]$.

The assumption $T \ge 1$ guarantees $[s_0 - 1, s_0] \subset [-T,T]$ or $[s_0, s_0+1] \subset [-T,T]$. We argue by contradiction and assume that $\kappa := |v(s_0)| > 2$. Then $[s_0 - 1, s_0]$ or $[s_0, s_0+1]$ is contained in the intersection $[-T,T] \cap [s_0 - \kappa^2/4, s_0 + \kappa^2/4]$. Therefore by Step 2 or Step 3 we have $\kappa = 2$. Contradiction.

By homogeneity of the norm Step 4 implies Lemma 3.2. □

Projection Π_T onto $\mathbb{E}_T$ along $\mathbb{K}_T$

We denote the linear **projection** in the path space $\mathbb{W}_T := W^{1,2}([-T,T], \mathbb{R}^n)$ onto the n dimensional subspace $\mathbb{E}_T$ along the (not necessarily orthogonal) complement $\mathbb{K}_T$ by

$$
P_{\mathbb{E}_T, \mathbb{K}_T} \overset{(3.25)}{=} \Pi_T \colon \mathbb{W}_T = \mathbb{E}_T \oplus \mathbb{K}_T \to \mathbb{E}_T. \tag{3.24}
$$

Lemma 3.3. *The projection $P_{\mathbb{E}_T,\mathbb{K}_T}$ is given by the map $\Pi_T\colon \zeta\mapsto\zeta_{\mathbb{E}}$ where*

$$\zeta_{\mathbb{E}}(s) := \left(e^{-(s+T)A_+}\zeta^+(-T), e^{(s-T)A_-}\zeta^-(T)\right) \tag{3.25}$$

for $s\in[-T,T]$. There is a constant $d=d(a_1,a_n)$, depending on the eigenvalues a_1 and a_n of A in (1.4), but independent of $T\ge 1$, such that $\|\Pi_T\|\le d$.

Proof. Let $\zeta\in\mathbb{W}_T$. The map $\zeta\mapsto\zeta_{\mathbb{E}}$ is linear. Moreover $(\zeta_{\mathbb{E}})_{\mathbb{E}}(s)=\zeta_{\mathbb{E}}(s)$ since $\zeta^+_{\mathbb{E}}(-T)=\zeta^+(-T)$ and $\zeta^-_{\mathbb{E}}(T)=\zeta^-(T)$. This proves identity 1 in the following

$$\Pi_T\circ\Pi_T \overset{1}{=} \Pi_T,\qquad \operatorname{im}\Pi_T \overset{2}{=} \mathbb{E}_T,\qquad \ker\Pi_T\overset{3}{=}\mathbb{K}_T.$$

Identity 2: One readily checks that $\partial_s\zeta_{\mathbb{E}}=-A\zeta_{\mathbb{E}}$, therefore $\zeta_{\mathbb{E}}\in\mathbb{E}_T$. Hence $\operatorname{im}\Pi_T\subset\mathbb{E}_T$. Vice versa, given $\zeta\in\mathbb{E}_T$, the two components $\zeta^\pm:=p_\pm\zeta$ satisfy for $s\in[-T,T]$, and using (1.4), the ODE $\partial_s\zeta^+=-A_+\zeta^+$ with initial value $\zeta^+(-T)$ at $s=-T$ and the ODE $\partial_s\zeta^-=A_-\zeta^-$ with initial value $\zeta^-(T)$ at $s=T$. The solutions are given by $s\mapsto e^{-(s+T)A_+}\zeta^+(-T)$ and by $s\mapsto e^{(s-T)A_-}\zeta^-(T)$, respectively. Their direct sum is $\zeta_{\mathbb{E}}$, hence $\mathbb{E}_T\subset\operatorname{im}\Pi_T$.
Identity 3: Pointwise in s vanishing of the vector valued map $(\Pi_T\zeta)(s)=\left(e^{-(s+T)A_+}\zeta^+(-T),e^{(s-T)A_-}\zeta^-(T)\right)=(0,0)$ happens iff both components vanish, that is iff $\zeta\in\mathbb{K}_T$.

To find a bound for P, pick $\zeta\in\mathbb{W}_T$. Straightforward calculation shows that

$$\begin{aligned}
\|\zeta_{\mathbb{E}}\|^2_{\mathbb{W}_T} &= \|\zeta_{\mathbb{E}}\|^2_{\mathbb{V}_T}+\|\tfrac{d}{ds}\zeta_{\mathbb{E}}\|^2_{\mathbb{V}_T}\\
&\overset{(3.25)}{=} \int_{-T}^T\left|\left(e^{-(s+T)A_+}\zeta^+(-T),e^{(s-T)A_-}\zeta^-(T)\right)\right|^2 ds\\
&\quad+\int_{-T}^T\left|\left(-A_+e^{-(s+T)A_+}\zeta^+(-T),A_-e^{(s-T)A_-}\zeta^-(T)\right)\right|^2 ds\\
&\overset{(1.5)}{=}\int_0^{2T}\left|e^{-sA_+}\xi^+(-T)\right|^2+\left|A_+e^{-sA_+}\xi^+(-T)\right|^2 ds\\
&\quad+\int_{-2T}^0\left|e^{sA_-}\zeta^-(T)\right|^2+\left|A_-e^{sA_-}\zeta^-(T)\right|^2 ds\\
&\overset{(1.6)}{=}\sum_{i=1}^{n-k}\int_0^{2T}(1+a_i^2)e^{-2sa_i}\zeta_i(-T)^2\,ds\\
&\quad+\sum_{j=n-k+1}^n\int_{-2T}^0(1+a_j^2)e^{-2sa_j}\zeta_j(T)^2\,ds\\
&=\sum_{i=1}^{n-k}(1+a_i^2)\zeta_i(-T)^2\tfrac{1-e^{-4Ta_i}}{2a_i}+\sum_{j=n-k+1}^n(1+a_j^2)\zeta_j(T)^2\tfrac{1-e^{4Ta_j}}{-2a_j}\\
&\overset{(1.4)}{\le}\tfrac{1+a_1^2}{2\sigma}\sum_{i=1}^{n-k}\zeta_i(-T)^2+\tfrac{1+a_n^2}{2\sigma}\sum_{j=n-k+1}^n\zeta_j(T)^2\\
&\le\tfrac{d^2}{8}\left(|p_+\zeta(-T)|^2+|p_-\zeta(T)|^2\right)\qquad,\ \tfrac{d^2}{8}:=\tfrac{\max\{1+a_1^2,1+a_n^2\}}{2\sigma}
\end{aligned}$$

$$\leq \frac{d^2}{4}\|\zeta\|^2_{L^\infty([-T,T])}$$
$$\overset{(3.23)}{\leq} d^2\|\zeta\|^2_{\mathbb{W}_T}.$$

Here equality two is by definition (3.25) of $\zeta_{\mathbb{E}}$. Equality three is by the g_0-orthogonal splitting (1.5) which makes the mixed inner products zero. Equality four uses that $A = \operatorname{diag}(A_+, -A_-)$, by (1.6), and the $a_i > 0 > a_j$ are ordered by (1.4). Equality five is by integration. The first inequality uses the order (1.4) of the matrix entries a_ℓ and definition (1.7) of the spectral gap σ. The third inequality uses that the projections $p_\pm$ are orthogonal, hence of norm ≤ 1. The final inequality four is by the, uniform in T, Sobolev estimate (3.23). This concludes the proof of Lemma 3.3. □

3.2 Infinitesimal gluing map Γ_T

Definition 3.4. For $T \geq 3$ we call the linear map defined by the composition

$$\Gamma_T := \Pi_T \circ d\wp_T(0_+, 0_-) \colon \mathbb{E}^+ \times \mathbb{E}^- \to \mathbb{W}_T \to \mathbb{E}_T \tag{3.26}$$

the **infinitesimal gluing map**. It acts as shown in (3.29) and Figure 3.

To obtain a formula for Γ_T we proceed in three steps I–III. Fix elements $\xi = (\xi^+, 0) \in \mathbb{E}^+$ and $\eta = (0, \eta^-) \in \mathbb{E}^-$; see (2.17).

I. Time $s = -T$: By definition (2.14 of $\wp_T$ and (2.11) of the pre-gluing map $\mathcal{P}_T$ (using that $1 - \beta(-T+2) = 1$ and $\beta(-T-2) = 0$ for $T \geq 3$) we obtain

$$d\wp_T(0_+, 0_-)(\xi, \eta)(-T) \overset{(2.14)}{=} \left(\left.\tfrac{d}{d\tau}\right|_0 \mathcal{P}_T(\varepsilon\xi, \varepsilon\eta)\right)(-T)$$
$$\overset{(2.11)}{=} 1 \cdot \xi(T - T) + 0 \cdot \eta(-T - T) = (\xi^+(0), 0).$$

In view of the direct sum $\mathbb{W}_T = \mathbb{K}_T \oplus \mathbb{E}_T$ and since $\Gamma_T(\xi, \eta) := P_{\mathbb{E}_T, \mathbb{K}_T} \circ d\wp_T(0_+, 0_-)(\xi, \eta)$ is the projection to $\mathbb{E}_T$ there is an element ζ in the projection kernel $\mathbb{K}_T$ such that $d\wp_T(0_+, 0_-)(\xi, \eta) = \zeta + \Gamma_T(\xi, \eta)$. So we get identity 1 in

$$p_+\Big(\Gamma_T(\xi, \eta)(-T)\Big) \overset{1}{=} p_+\Big(d\wp_T(0_+, 0_-)(\xi, \eta)(-T) - \zeta(-T)\Big) \overset{2}{=} \xi^+(0). \tag{3.27}$$

Identity 2 holds as $p_+(\zeta(-T)) = 0$ by condition one in definition (3.22) of $\mathbb{K}_T$.

II. Time $s = T$: Similarly as in I. we obtain that

$$d\wp_T(0_+, 0_-)(\xi, \eta)(T) \overset{(2.11)}{=} \eta(0) = (0, \eta^-(0)).$$

Now use condition two in definition (3.22) of $\mathbb{K}_T$ to conclude that

$$p_-\left(\Gamma_T(\xi, \eta)(T)\right) = \eta^-(0). \tag{3.28}$$

III. Time $s \in [-T, T]$: Since $\Gamma_T(\xi, \eta)$ lies in $\mathbb{E}_T$ it satisfies the ODE given by $\partial_s \Gamma_T(\xi, \eta) + A\Gamma_T(\xi, \eta) = 0$ and so, by (3.27) and (3.28), we get the formula

$$\Gamma_T(\xi, \eta)(s) = \Big(e^{-(s+T)A_+}\xi^+(0), e^{(s-T)A_-}\eta^-(0)\Big). \tag{3.29}$$

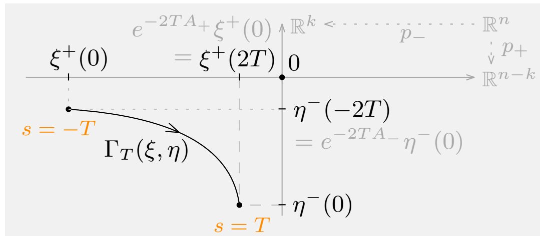

Figure 3: Infinitesimal gluing isomorphism $\Gamma_T \in \mathcal{L}(\mathbb{E}^+ \times \mathbb{E}^-, \mathbb{E}_T)$; cf. (3.29)

In particular, due to the choice of the complement $\mathbb{K}_T$ which just involves the ends $-T$ and T, the infinitesimal gluing map **does not depend** on the choice of the cutoff function β used to define the pre-gluing map (2.11).

Lemma 3.5 (Norm)**.** *Let $T \geq 3$. The linear map $\Gamma_T \colon \mathbb{E}^+ \times \mathbb{E}^- \to \mathbb{E}_T$ is an isomorphism of norm $\|\Gamma_T\| \leq 1$ and the norm of the inverse is bounded, uniformly in T, by $k := 1/(1 - e^{-12\sigma})$ where σ is the spectral gap (1.7) of A.*

Proof. We saw that $\dim \mathbb{E}^+ = n - k$, $\dim \mathbb{E}^- = k$, and $\dim \mathbb{E}_T = n$. Hence it suffices to show injectivity, i.e. that the kernel of Γ_T is trivial. Given $(\xi, \eta) \in \ker \Gamma_T$, then by (3.29) we have $\xi^+(0) = 0$ and $\eta^-(0) = 0$. So $\xi^+ \equiv 0$ and $\eta^- \equiv 0$, since ξ^+ and η^- are solutions of a linear first order ODE; see (2.17). Therefore $\xi = (\xi^+, 0) \equiv 0$ and $\eta = (0, \eta^-) \equiv 0$. This shows that Γ_T is an isomorphism whenever $T \geq 3$.

To see that Γ_T and ${\Gamma_T}^{-1}$ are bounded, uniformly in T, consider the identities

$$
\begin{aligned}
\|\Gamma_T(\xi,\eta)\|^2_{\mathbb{W}_T} &= \|\Gamma_T(\xi,\eta)\|^2_{\mathbb{V}_T} + \|\tfrac{d}{ds}\Gamma_T(\xi,\eta)\|^2_{\mathbb{V}_T} \\
&\overset{(3.29)}{=} \int_{-T}^{T} \left|\left(e^{-(s+T)A_+}\xi^+(0), e^{(s-T)A_-}\eta^-(0)\right)\right|^2 ds \\
&\quad + \int_{-T}^{T} \left|\left(-A_+ e^{(-s-T)A_+}\xi^+(0), A_- e^{(s-T)A_-}\eta^-(0)\right)\right|^2 ds \\
&\overset{(1.5)}{=} \int_0^{2T} \left|e^{-sA_+}\xi^+(0)\right|^2 + \left|A_+ e^{-sA_+}\xi^+(0)\right|^2 ds \\
&\quad + \int_{-2T}^{0} \left|e^{sA_-}\eta^-(0)\right|^2 + \left|A_- e^{sA_-}\eta^-(0)\right|^2 ds \\
&\overset{(1.6)}{=} \sum_{i=1}^{n-k} \int_0^{2T} (1 + a_i^2) e^{-2sa_i} \xi_i(0)^2\, ds \\
&\quad + \sum_{j=n-k+1}^{n} \int_{-2T}^{0} (1 + a_j^2) e^{-2sa_j} \eta_j(0)^2\, ds \\
&= \sum_{i=1}^{n-k} (1 + a_i^2)\xi_i(0)^2 \tfrac{1 - e^{-4Ta_i}}{2a_i} + \sum_{j=n-k+1}^{n} (1 + a_j^2)\eta_j(0)^2 \tfrac{1 - e^{4Ta_j}}{-2a_j}
\end{aligned}
$$

where equality two is by formula (3.29) for Γ_T. Equality three is by the g_0-orthogonal splitting (1.5). Equality four uses that $A = \mathrm{diag}(A_+, -A_-)$, by (1.6), and the $a_i > 0 > a_j$ are ordered by (1.4). Equality five is by integration.

The $\mathbb{W}_+$ norm of $\xi = (\xi^+, 0) \in \mathbb{E}^+$, see (2.17), is given by

$$
\begin{aligned}
\|\xi\|_{\mathbb{W}_+}^2 &= \|\xi\|_{\mathbb{V}_+}^2 + \|\tfrac{d}{ds}\xi\|_{\mathbb{V}_+}^2 \\
&\overset{(2.17)}{=} \int_0^\infty |e^{-sA_+}\xi^+(0)|^2\, ds + \int_0^\infty |A_+ e^{-sA_+}\xi^+(0)|^2\, ds \\
&\overset{(1.4)}{=} \sum_{i=1}^{n-k} \int_0^\infty (1+a_i^2) e^{-2sa_i} \xi_i(0)^2\, ds \\
&= \sum_{i=1}^{n-k} \tfrac{1+a_i^2}{2a_i} \xi_i(0)^2
\end{aligned}
\tag{3.30}
$$

and analogously for the $\mathbb{W}_-$ norm of $\eta = (0, \eta^-) \in \mathbb{E}^-$.

To see that ${\Gamma_T}^{-1}$ is bounded, uniformly in T, we estimate Γ_T from below

$$
\begin{aligned}
\|\Gamma_T(\xi,\eta)\|_{\mathbb{W}_T}^2 &= \sum_{i=1}^{n-k} (1+a_i^2)\xi_i(0)^2 \tfrac{1-e^{-4Ta_i}}{2a_i} + \sum_{j=n-k+1}^{n} (1+a_j^2)\eta_j(0)^2 \tfrac{1-e^{4Ta_j}}{-2a_j} \\
&\geq \left(1-e^{-12\sigma}\right) \left(\sum_{i=1}^{n-k} \tfrac{1+a_i^2}{2a_i}\xi_i(0)^2 + \sum_{i=n-k+1}^{n} \tfrac{1+a_j^2}{-2a_j}\eta_j(0)^2 \right) \\
&= \left(1-e^{-12\sigma}\right) \left(\|\xi\|_{\mathbb{W}_+}^2 + \|\eta\|_{\mathbb{W}_-}^2 \right).
\end{aligned}
$$

To obtain the inequality we use the assumption $T \geq 3$ and the spectral gap σ of A defined by (1.7) of A. The last equality is explained right above.

To see that Γ_T is bounded, uniformly in T, we estimate Γ_T from above

$$
\begin{aligned}
\|\Gamma_T(\xi,\eta)\|_{\mathbb{W}_T}^2 &= \sum_{i=1}^{n-k} (1+a_i^2)\xi_i(0)^2 \tfrac{1-e^{-4Ta_i}}{2a_i} + \sum_{j=n-k+1}^{n} (1+a_j^2)\eta_j(0)^2 \tfrac{1-e^{4Ta_j}}{-2a_j} \\
&\leq \sum_{i=1}^{n-k} \tfrac{1+a_i^2}{2a_i}\xi_i(0)^2 + \sum_{j=n-k+1}^{n} \tfrac{1+a_j^2}{-2a_j}\eta_j(0)^2 \\
&= \|\xi\|_{\mathbb{W}_+}^2 + \|\eta\|_{\mathbb{W}_-}^2
\end{aligned}
$$

where the inequality holds since $1 - e^{-4Ta_i} \leq 1$ and $1 - e^{4Ta_j} \leq 1$. □

4 Newton-Picard map

Given two elements $w_+ \in \mathcal{W}^{\mathrm{s}}$ and $w_- \in \mathcal{W}^{\mathrm{u}}$ near the critical point, we view the pre-glued path

$$
w_T := \wp_T(w_+, w_-) \in \mathbb{W}_T
$$

as an approximate flow trajectory, equivalently an approximate zero x_1 of the section $\mathcal{F}_T$, and then detect a nearby solution using the implicit function theorem with initial point $x_0 := 0_T$, see Appendix B. Thus we need that $\mathcal{F}_T(w_T)$ is suitably close to zero. We also need a uniformly in T bounded right inverse Q_T of the linearization $D_T := d\mathcal{F}_T(0_T)$. These are the next two subsections.

4.1 Approximate zero

Proposition 4.1 (Pre-glued path is approximate zero)**.** *Pick $\varepsilon \in (0, \sigma)$ where $\sigma = \sigma(A)$ is the spectral gap (1.7). Let $m \in \mathbb{N}_0$ and choose compact neighborhoods $K_+ = K_+(m)$ of $0_+ \in \mathcal{W}^{\mathrm{s}}$ in $T^m\mathcal{W}^{\mathrm{s}}$ and $K_- = K_-(m)$ of $0_- \in \mathcal{W}^{\mathrm{u}}$ in $T^m\mathcal{W}^{\mathrm{u}}$. Then the following is true. Given two elements*

$$W_+ \in K_+ \subset T^m\mathcal{W}^{\mathrm{s}}, \qquad W_- \in K_- \subset T^m\mathcal{W}^{\mathrm{u}},$$

pre-glue them with the m-fold tangent map of $\wp_T$, see (2.14), to get

$$W_T := T^m\wp_T(W_+, W_-).$$

Then there is a constant $C(K_+, K_-)$, independent of T, such that

$$\|T^m\mathcal{F}_T(W_T)\|_{T^m\mathbb{V}_T} \le C(K_+, K_-) \cdot e^{-\varepsilon T} \tag{4.31}$$

whenever $T \ge 3$.

Proof. Let $m \in \mathbb{N}_0$. An element $W_+ \in T^m\mathcal{W}^{\mathrm{s}}$ is a map satisfying the equation

$$W_+ \colon [0, \infty) \to T^m\mathbb{R}^n, \qquad \partial_s W_+ + T^m\nabla f(W_+) = 0. \tag{4.32}$$

An element $W_- \in T^m\mathcal{W}^{\mathrm{s}}$ is a map that satisfies the equation

$$W_- \colon (-\infty, 0] \to T^m\mathbb{R}^n, \qquad \partial_s W_- + T^m\nabla f(W_-) = 0. \tag{4.33}$$

Since 0 is a non-degenerate critical point of f the solutions W_+ and W_- decay with all their derivatives exponentially. In particular, by Theorem C.1 and compactness of $K_\pm$, there is a constant $c = c(K_+, K_-)$ with

$$|W_+(s)| + |\partial_s W_+(s)| \le ce^{-\varepsilon s}, \quad |W_-(s)| + |\partial_s W_-(s)| \le ce^{\varepsilon s}, \tag{4.34}$$

for every $s \ge 0$, respectively $s \le 0$. Since f has a critical point at the origin, we get $T^m\nabla f(0) = 0$ where $0 \in T^m\mathbb{R}^n$.[4] Hence there is a constant $\mu_m > 0$ with

$$|T^m\nabla f(W)| \le \mu_m |W| \tag{4.35}$$

whenever $|W| \le c$.

By linearity of the pre-gluing map $\mathcal{P}_T$, the $m-$fold tangent map is given by

$$W_T := T^m\mathcal{P}_T(W_+, W_-) \stackrel{(2.11)}{=} (1 - \beta(s+2))\, W_+(T+s) + \beta(s-2)\, W_-(-T+s)$$

[4] It holds $T^1\nabla f(0,0) = (\nabla f(0), D\nabla f(0)0) = (0,0)$, similarly $T^m\nabla f(0,\dots,0) = (0,\dots,0)$.

pointwise at $s \in [-T, T]$. Similarly, in analogy to (2.12), we have

$$W_T(s) = \begin{cases} W_+(T+s) & , s \in [-T, -3] \\ (1-\beta(s+2))\, W_+(T+s) & , s \in [-3, -1] \\ 0 & , s \in [-1, 1] \\ \beta(s-2) W_-(-T+s) & , s \in [1, 3] \\ W_-(-T+s) & , s \in [3, T] \end{cases} \tag{4.36}$$

for $s \in [-T, T]$. The tangent map of $\mathcal{F}_T(w) := \partial_s w + \nabla f(w)$ at W_T is given by

$$T^m \mathcal{F}_T(W_T) = \partial_s W_T + T^m \nabla f(W_T). \tag{4.37}$$

Now there are three cases.
1) For $s \in [-T, -3] \cup [-1, 1] \cup [3, T]$ the map $T^m \mathcal{F}_T(W_T)$ vanishes:

$$T^m \mathcal{F}_T(W_T)(s) \overset{(4.36)}{=} \partial_s W_+(T+s) + T^m \nabla f|_{W_T(T+s)} \overset{(4.32)}{=} 0, \quad \forall s \in [-T, -3].$$

$$T^m \mathcal{F}_T(W_T)(s) \overset{(4.36)}{=} \partial_s W_-(-T+s) + T^m \nabla f|_{W_-(-T+s)} \overset{(4.33)}{=} 0, \quad \forall s \in [3, T].$$

$$T^m \mathcal{F}_T(W_T)(s) = 0 \quad \text{since} \quad W_T(s) \overset{(4.36)}{=} 0, \quad \forall s \in [-1, 1].$$

2) For $s \in [-3, -1]$, using (4.37) for $T^m \mathcal{F}_T(W_T)$ and (4.36) for W_T, we obtain

$$\begin{aligned} T^m \mathcal{F}_T(W_T)(s) = & -\beta'(s+2) W_+(T+s) + (1-\beta(s+2)) \partial_s W_+(T+s) \\ & + T^m \nabla f|_{(1-\beta(s+2)) W_+(T+s)}. \end{aligned}$$

Now, by (4.35) and (4.32), we estimate the pointwise length by

$$\begin{aligned} & |T^m \mathcal{F}_T(W_T)(s)| \\ & \le \|\beta'\|_\infty |W_+(T+s)| + |\partial_s W_+(T+s)| + \left| T^m \nabla f|_{(1-\beta(s+2)) W_+(T+s)} \right| \\ & \le (\|\beta'\|_\infty + 1 + \mu_m)\, c \cdot e^{-\varepsilon(s+T)}. \end{aligned}$$

3) For $s \in [1, 3]$ we obtain analogously the formula

$$\begin{aligned} T^m \mathcal{F}_T(W_T)(s) = & -\beta'(s-2) W_-(-T+s) + \beta(s-2) \partial_s W_-(-T+s) \\ & + T^m \nabla f|_{\beta(s-2) W_-(-T+s)} \end{aligned}$$

and the estimate $|T^m \mathcal{F}_T(W_T)(s)| \le (\|\beta'\|_\infty + 1 + \mu_m)\, c e^{\varepsilon(-T+s)}$.

Thus for the L^2 norm we get, by integration, the estimate

$$\|T^m \mathcal{F}_T(W_T)\|^2_{T^m \mathbb{V}_T} \le \underbrace{2 \left(\|\beta'\|_\infty + 1 + \mu_m\right)^2 c^2 \frac{e^{6\varepsilon} - e^{2\varepsilon}}{\varepsilon}}_{=: C(K_+, K_-)^2} e^{-2\varepsilon T}$$

where $c = c(K_+, K_-)$ and $K_\pm$ depend on m. This proves Proposition 4.1. $\square$

4.2 Surjectivity and right inverse

Surjective linearization at 0_T. Let $T>0$. The linearization at general $w\in\mathbb{W}_T$ is given by

$$d\mathcal{F}_T(w)\colon \mathbb{W}_T\to\mathbb{V}_T,\quad \zeta\mapsto \partial_s\zeta+A_w\zeta,\qquad A_w(s):=\nabla\nabla f(w(s)),\tag{4.38}$$

where $\nabla\nabla f(w(s))$ is the Jacobian of the vector field $\nabla f\colon\mathbb{R}^n\to\mathbb{R}^n$ at $w(s)$. The linearization at the origin 0_T, namely the operator

$$D_T:=d\mathcal{F}_T(0_T)\colon \mathbb{W}_T=\mathbb{E}_T\oplus\mathbb{K}_T\to\mathbb{V}_T,\quad \zeta\mapsto\partial_s\zeta+A\zeta,\tag{4.39}$$

where $A_{0_T}=A=\operatorname{diag}(A_+,-A_-)$ is the diagonal block matrix (1.6), is surjective: Given $\eta\in\mathbb{V}_T$, then the element defined by

$$\zeta(s):=e^{-sA}\left(\zeta_0+\int_0^s e^{\sigma A}\eta(\sigma)\,d\sigma\right),\quad \zeta_0\in\mathbb{R}^n,\tag{4.40}$$

for $s\in[-T,T]$, lies in $\mathbb{W}_T$ and satisfies $d\mathcal{F}_T(0_T)\zeta=\eta$.[5] For $\mathcal{M}_T:=\mathcal{F}_T^{-1}(0)$ there is the natural inclusion $\mathbb{E}_T:=T_{0_T}\mathcal{M}_T\subset\ker d\mathcal{F}_T(0_T)$. On the other hand, both spaces are determined by the initial conditions which are given by $\mathbb{R}^n$, thus $\dim\mathbb{E}_T=n=\dim\ker D_T$. Therefore the two spaces coincide

$$\ker D_T=\mathbb{E}_T.\tag{4.41}$$

Since $\mathbb{K}_T$ is a complement of $\mathbb{E}_T$, the restriction

$$\begin{pmatrix}F_T^+ & 0\\ 0 & F_T^-\end{pmatrix}=F_T:=d\mathcal{F}_T(0_T)|_{\mathbb{K}_T}=\left.\begin{pmatrix}\frac{d}{ds}+A_+ & 0\\ 0 & \frac{d}{ds}-A_-\end{pmatrix}\right|_{\mathbb{K}_T}\colon\mathbb{K}_T\to\mathbb{V}_T\tag{4.42}$$

is injective, hence a continuous linear bijection. Hence, by the open mapping theorem, the inverse

$$\begin{pmatrix}Q_T^+ & 0\\ 0 & Q_T^-\end{pmatrix}=Q_T:={F_T}^{-1}=\begin{pmatrix}(F_T^+)^{-1} & 0\\ 0 & (F_T^-)^{-1}\end{pmatrix}\colon\mathbb{V}_T\to\mathbb{K}_T\tag{4.43}$$

is also continuous. Thus the map $F_T\colon\mathbb{K}_T\to\mathbb{V}_T$ is a Hilbert space isomorphism.

Right inverse Q_T of D_T.

Remark 4.2 (Q_T is a right inverse of D_T)**.** Given $\eta\in\mathbb{V}_T$, then $\zeta:=Q_T\eta\in\mathbb{K}_T$, hence $d\mathcal{F}_T(0_T)\zeta=F_T\zeta$. Therefore

$$d\mathcal{F}_T(0_T)\circ Q_T\eta=F_T\circ Q_T\eta=\eta.$$

Lemma 4.3. *There is a constant c, independent of $T>0$, such that $\|Q_T\|\le c$.*

[5] Since the interval $[-T,T]$ is finite, the Morse condition is not needed here.

Proof. In the proof we distinguish three cases.

I. A POSITIVE DEFINITE: In this case $k = 0$ and $p_- = 0$, see (1.5), in particular $a_1 \geq \dots \geq a_n > 0$. Thus $\mathbb{K}_T = \{\xi \in \mathbb{W}_T \mid \xi(-T) = 0\}$; see (3.22). Given $\eta \in \mathbb{V}_T$, let $\zeta := Q_T\eta$, equivalently $\eta = F_T\zeta$. Since $\zeta \in \mathbb{K}_T$, we know that $\zeta(-T) = 0$. By (4.40), where we changed the start of the integration from 0 to $-T$, we get the formula

$$\begin{aligned}\zeta(s) = e^{-sA}\left(\zeta(-T) + \int_{-T}^{s} e^{\sigma A}\eta(\sigma)\, d\sigma\right) &\overset{\zeta(-T)=0}{=} \int_{-T}^{s} e^{-(s-\sigma)A}\eta(\sigma)\, d\sigma \\ &= \int_{[-T,s]\cup(s,T]} \phi(s-\sigma)\eta(\sigma)\, d\sigma = (\phi * \eta)(s)\end{aligned}$$

whenever $s \in [-T, T]$ and where the function ϕ is defined by $\phi(r) := e^{-rA}$ for $r \geq 0$, and by 0 for $r < 0$. Hence, by Young's inequality, we have

$$\|\zeta\|_2 \leq \|\phi\|_1 \|\eta\|_2 \leq \frac{1}{a_n}\|\eta\|_2$$

where the L^2 and L^1 norms are over $[-T, T]$ and since

$$\begin{aligned}\|\phi\|_1 = \int_{-T}^{T} \|\phi(s)\|_{\mathcal{L}(\mathbb{R}^n)}\, ds &= \int_{0}^{T} \|e^{-sA}\|_{\mathcal{L}(\mathbb{R}^n)}\, ds \\ &= \int_0^T e^{-sa_n}\, ds = \tfrac{1-e^{-a_nT}}{a_n} \leq \tfrac{1}{a_n}.\end{aligned}$$

Note that $a_n > 0$ is the smallest eigenvalue of the positive definite operator A. Since $\partial_s\zeta = \eta - A\zeta$, and by the triangle inequality and $\|\zeta\|_2 \leq \frac{1}{a_n}\|\eta\|_2$, we get

$$\begin{aligned}\|Q_T\eta\|_{1,2}^2 = \|\zeta\|_{1,2}^2 &= \|\partial_s\zeta\|_2^2 + \|\zeta\|_2^2 \\ &= \|\eta - A\zeta\|_2^2 + \|\zeta\|_2^2 \\ &\leq (\|\eta\|_2 + \|A\zeta\|_2)^2 + \|\zeta\|_2^2 \\ &\leq \left(1 + \tfrac{a_1}{a_n}\right)^2 \|\eta\|_2^2 + \tfrac{1}{a_n^2}\|\eta\|_2^2.\end{aligned}$$

This proves Step 1 for $c^2 = \frac{(a_1+a_n)^2+1}{a_n^2}$.

II. A NEGATIVE DEFINITE: So $k = n$ and $p_- = \mathbb{1}$ and $\mathbb{K}_T = \{\xi \in \mathbb{W}_T \mid \xi(T) = 0\}$. Given $\eta \in \mathbb{V}_T$, let $\zeta := Q_T\eta$, equivalently $\eta = F_T\zeta$. Since $\zeta \in \mathbb{K}_T$, we know that $\zeta(T) = 0$. By (4.40), where we changed the start of the integration from 0 to T, we obtain the formula

$$\begin{aligned}\zeta(s) = e^{sA}\left(\zeta(T) + \int_{T}^{s} e^{-\sigma A}\eta(\sigma)\, d\sigma\right) &\overset{\zeta(T)=0}{=} -\int_{s}^{T} e^{(s-\sigma)A}\eta(\sigma)\, d\sigma \\ &= -\int_{[-T,s)\cup[s,T]} \phi(s-\sigma)\eta(\sigma)\, d\sigma = -(\phi * \eta)(s)\end{aligned}$$

whenever $s \in [-T,T]$ and where ϕ was defined in Step 1. Continue as in Step 1.

III. General case: Given $\eta = (\eta^+, \eta^-) \in \mathbb{V}_T$. Let $\zeta := Q_T\eta$, then

$$\zeta = (\zeta^+, \zeta^-), \qquad \zeta^+ = Q_T^+\eta^+, \quad \zeta^- = Q_T^-\eta^-.$$

From Step I and Step II there exists a constant $c > 0$ such that $\|\zeta^+\|_{1,2} \le c\|\eta^+\|_2$ and $\|\zeta^-\|_{1,2} \le c\|\eta^-\|_2$. Since the splitting $\mathbb{R}^{n-k} \times \mathbb{R}^k$ is orthogonal, we have

$$\|\zeta\|_{1,2}^2 \stackrel{\perp}{=} \|\zeta^+\|_{1,2}^2 + \|\zeta^-\|_{1,2}^2 \le c^2\|\eta^+\|_2^2 + c^2\|\eta^-\|_2^2 \stackrel{\perp}{=} c^2\|\eta\|_2^2.$$

This proves Step III and Lemma 4.3. □

4.3 Definition of $\mathcal{N}_T$

Let c be the right inverse bound from Lemma 4.3. In order to use later on Remark B.8 to satisfy hypothesis (B.82), as opposed to only (B.70), we define, for $\mu \ge 2$, a nested family of open neighborhoods of 0 in $\mathbb{R}^n$ as the pre-image of $[0, 1/\mu c)$ under the continuous map $\|d\nabla f(\cdot) - A\| \colon \mathbb{R}^n \to [0,\infty)$, in symbols

$$B^\mu := \|d\nabla f(\cdot) - A\|^{-1}\,[0, \tfrac{1}{\mu c}), \quad B^\mu \subset B^2.$$

For $T > 0$ define an open neighborhood $\mathcal{B}_T^\mu$ of 0_T in $\mathbb{W}_T$ by

$$\mathcal{B}_T^\mu := \{w \in \mathbb{W}_T \mid w(s) \in B^\mu \;\forall s \in [-T,T]\} \subset \mathcal{B}_T^2.$$

Lemma 4.4. *Let $T > 0$ and $\mu \ge 2$. If $w \in \mathcal{B}_T^\mu$, then $\|d\mathcal{F}_T(w) - D_T\| \le \frac{1}{\mu c}$.*

Proof. Given $w \in \mathcal{B}_T^\mu$, there is the estimate

$$\|(d\mathcal{F}_T(w) - D_T)\zeta\|_2 \stackrel{(4.38)}{=} \|(d\nabla f(w) - A)\zeta\|_2 \le \tfrac{1}{\mu c}\|\zeta\|_2 \le \tfrac{1}{\mu c}\|\zeta\|_{1,2}$$

for every $\zeta \in \mathbb{W}_T$. □

Corollary 4.5 (to Lemma 3.2)**.** *There is a monotone decreasing function $\delta\colon [2,\infty) \to (0,\infty)$, $\mu \mapsto \delta(\mu) =: \delta_\mu$, independent of T, such that for $\mu \in [2,\infty)$ the δ_μ-ball about 0_T in $\mathbb{W}_T$ is contained in $\mathcal{B}_T^\mu$, in symbols $B_{\delta_\mu}(0_T;\mathbb{W}_T) \subset \mathcal{B}_T^\mu$.*

Proof. By Lemma 3.2 for $w \in B_{\delta_\mu}(0_T;\mathbb{W}_T)$ we have $\|w\|_\infty \le 2\delta_\mu$. □

Definition 4.6 (Newton-Picard map)**.** Let $c > 0$ be the right inverse bound of Lemma 4.3 and let

$$\delta := \delta_4 > 0 \tag{4.44}$$

be the value in Corollary 4.5 for $\mu = 4$.[6] For $T \ge 1$ we can now, in view of Lemma 4.4 with $\mu = 4$ and the McDuff-Salamon Proposition B.1 with $x_0 = 0_T$,

[6] To define the Newton-Picard map via the McDuff-Salamon Proposition B.1 and obtain C^0-convergence (Theorem 5.6 with $m = 0$) for the gluing map it is sufficient to pick δ_μ for $\mu = 2$. However, to obtain C^1-convergence (Theorem 5.6 with $m = 1$) we need to choose $\mu = 4$ in order to satisfy assumption (B.82) ($\frac{1}{4c}$ as opposed to $\frac{1}{2c}$) in the tangent map Theorem B.10.

define a **Newton-Picard map**

$$\mathcal{N}_T \colon \mathbb{W}_T \supset \mathcal{B}^4_T \supset U_0(\delta) \overset{\text{(B.74)}}{:=} \Big(B_{\frac{\delta}{8}}(0_T; \mathbb{W}_T) \cap \big\{ \|\mathcal{F}_T\| < \tfrac{\delta}{4c} \big\} \Big) \to \mathbb{W}_T. \quad (4.45)$$

Here the inclusion $U_0(\delta) \subset \mathcal{B}^4_T$ holds by Corollary 4.5.

By (B.72) the Newton-Picard map $\mathcal{N}_T$ enjoys the following properties:

$$\mathcal{F}_T \circ \mathcal{N}_T(w) = 0, \qquad \mathcal{N}_T(w) - w \in \operatorname{im} Q_T, \qquad \mathcal{N}_T(w) \in B_\delta(0_T; \mathbb{W}_T),$$

and, moreover, one has the estimate

$$\|(\mathcal{N}_T - \mathrm{id})(w)\|_{\mathbb{W}_T} \le 2c \|\mathcal{F}_T(w)\|_{\mathbb{V}_T}. \quad (4.46)$$

Furthermore, by Corollary B.7, respectively identity (5.53)), we have

$$\mathcal{N}_T(0_T) = 0_T, \qquad d\mathcal{N}_T|_{0_T} \overset{(5.53)}{=} \Pi_T, \quad (4.47)$$

where the projection Π_T, see (3.24), is uniformly bounded in T, by Lemma 3.3.

Pre-gluing takes values in domain of Newton-Picard map

The next lemma and Proposition 4.1 show that, for $T \ge 3$ large enough, the pre-gluing map $\mathcal{P}_T$ takes values in the domain of the Newton-Picard map $\mathcal{N}_T$.

Lemma 4.7 (The neighborhoods $\mathcal{U}^\mu_\pm$.)**.** *Let $\delta \colon [2,\infty) \to (0,\infty)$, $\mu \mapsto \delta_\mu$, be the monotone decreasing function in Corollary 4.5. We abbreviate $\delta := \delta_4$. Then there exists a nested family of open and bounded neighborhoods $\mathcal{U}^\mu_+ \subset \mathcal{W}^{\mathrm{s}}$ of 0_+ and $\mathcal{U}^\mu_- \subset \mathcal{W}^{\mathrm{u}}$ of 0_- such that $\|\mathcal{P}_T(w_+, w_-)\|_{1,2} < \min\{\frac{\delta}{8}, \delta_\mu\}$ whenever $\mu \ge 2$, $w_+ \in \mathcal{U}^\mu_+$, $w_- \in \mathcal{U}^\mu_-$, and $T \ge 3$.*

While the estimate by $\frac{\delta}{8}$ serves in (4.45), the estimate by δ_μ for some $\mu \ge 5$ will be used in the proof of Theorem 5.3 further below.

Proof. For $\delta' > 0$ let $B_{\delta'}(0_+) \subset \mathbb{W}_+$ be the open radius δ' ball about 0_+, analogously for $B_{\delta'}(0_-)$. Pick $w_+ \in B_{\delta'}(0_+) \cap \mathcal{W}^{\mathrm{s}}$ and $w_- \in B_{\delta'}(0_-) \cap \mathcal{W}^{\mathrm{u}}$ and, for $T \ge 3$, abbreviate $w_T := \mathcal{P}_T(w_+, w_-) \colon [-T, T] \to \mathbb{R}^n$. Since by (2.12) at each time s at most one of $w_+(T+s)$ and $w_-(-T+s)$ comes with a nonzero factor, we obtain inequality one in the following estimate

$$\begin{aligned}
&\|w_T\|^2_{1,2} \\
&= \|w_T\|_2^2 + \|\partial_s w_T\|_2^2 \\
&\le \|w_+\|_2^2 + \|w_-\|_2^2 + 2\|\beta'\|^2_\infty \left(\|w_+\|_2^2 + \|w_-\|_2^2\right) + 2\left(\|\partial_s w_+\|_2^2 + \|\partial_s w_-\|_2^2\right) \\
&\le 2(1 + \|\beta'\|^2_\infty)\left(\|w_+\|^2_{1,2} + \|w_-\|^2_{1,2}\right) \\
&\le 4(1 + \|\beta'\|^2_\infty)(\delta')^2.
\end{aligned}$$

Choose $\delta' = \delta'(\mu) := \frac{1}{4}\min\{\frac{\delta}{8}, \delta_\mu\}\sqrt{1 + \|\beta'\|^2_\infty}$ and define the open δ'-neighborhoods in the stable, respectively unstable, manifolds as follows

$$\mathcal{U}^\mu_+ := (B_{\delta'}(0_+) \cap \mathcal{W}^{\mathrm{s}}) \subset \mathcal{U}_+ := \mathcal{U}^2_+, \quad \mathcal{U}^\mu_- := (B_{\delta'}(0_-) \cap \mathcal{W}^{\mathrm{u}}) \subset \mathcal{U}_- := \mathcal{U}^2_-.$$

Then the lemma holds by the previous displayed estimate. □

5 Gluing

Pick $\varepsilon \in (0,\sigma)$ where $\sigma = \sigma(A)$ is the spectral gap (1.7). Let $c > 0$ be the constant in the right inverse estimate, Lemma 4.3. Let $\delta = \delta_4 > 0$ be the constant in Corollary 4.5 and let

$$\mathcal{U}_{+/-} := \mathcal{U}^2_{+/-} \subset \mathcal{W}^{s/u}, \qquad K_\pm := \mathrm{cl}\,\mathcal{U}_\pm, \tag{5.48}$$

be the open sets in Lemma 4.7 and, respectively, the compact sets given by the closure of $\mathcal{U}_\pm$ in the (finite dimensional) stable/unstable manifold. Thus, by Proposition 4.1 for $m = 0$, we get a constant $C = C(K_+, K_-) > 0$. Pick $T_0 \geq 3$ such that

$$Ce^{-\varepsilon T_0} < \frac{\delta}{4c} \tag{5.49}$$

see (4.31). By Proposition 4.1 for $m = 0$ and Lemma 4.7 it holds that

$$\|\mathcal{F}_T \circ \wp_T(w_+, w_-)\| \overset{(4.31)}{<} \tfrac{\delta}{4c}, \qquad \|\wp_T(w_+, w_-)\| < \tfrac{\delta}{8}, \tag{5.50}$$

whenever $T \geq T_0$ and $w_\pm \in \mathcal{U}_\pm$ and were $\wp_T := \mathcal{P}_T|_{\mathcal{W}^s \times \mathcal{W}^u} \colon \mathcal{W}^s \times \mathcal{W}^u \to \mathbb{W}_T$, see (2.14), is the restriction of the (linear) pre-gluing map $\mathcal{P}_T \colon \mathbb{W}_+ \times \mathbb{W}_- \to \mathbb{W}_T$, see (2.11). In other words, the pre-gluing map $\wp_T$ maps $\mathcal{U}_+ \times \mathcal{U}_-$ into the domain of the Newton-Picard map $\mathcal{N}_T$, see (4.45), whenever $T \geq T_0$.

Definition 5.1 (Gluing map)**.** For $T \geq T_0$ the gluing map is the composition of smooth maps

$$\gamma_T := \mathcal{N}_T \circ \wp_T \colon \mathcal{U}_+ \times \mathcal{U}_- \longrightarrow U_0(\delta) \longrightarrow \mathbb{W}_T. \tag{5.51}$$

The linearized gluing map is the composition

$$d\gamma_T|_{(w_+,w_-)} = d\mathcal{N}_T|_{\wp_T(w_+,w_-)} \circ d\wp_T|_{(w_+,w_-)} \colon T_{w_+}\mathcal{W}^s \times T_{w_-}\mathcal{W}^u \to \mathbb{W}_T.$$

Lemma 5.2. *It holds $\gamma_T(0_+, 0_-) = 0_T$. Furthermore, the differential of the gluing map γ_T at $(0_+, 0_-)$ is the infinitesimal gluing map Γ_T, in symbols*

$$d\gamma_T|_{(0_+,0_-)} = \Pi_T \circ d\wp_T|_{(0_+,0_-)} \overset{(3.26)}{=:} \Gamma_T \colon \mathbb{E}^+ \times \mathbb{E}^- \to \mathbb{E}_T.$$

Proof. We get that $\gamma_T(0_+, 0_-) \overset{(2.13)}{=} \mathcal{N}_T(\wp_T(0_+, 0_-)) \overset{(4.47)}{=} \mathcal{N}_T(0_T) = 0_T$.

By definition of γ_T and the chain rule we get the first equality

$$\begin{aligned} d\gamma_T|_{(0_+,0_-)} &= d\mathcal{N}_T|_{\underbrace{\wp_T(0_+,0_-)}_{= 0_T \text{ by } (2.13)}} \circ d\wp_T|_{(0_+,0_-)} \\ &= (\mathrm{Id} - Q_T D_T) \circ d\wp_T|_{(0_+,0_-)} \\ &= \Pi_T \circ d\wp_T|_{(0_+,0_-)} \\ &=: \Gamma_T \end{aligned} \tag{5.52}$$

and the second equality holds since $d\mathcal{N}_T(0_T) = \mathrm{Id} - Q_T D_T$, by Corollary B.7 with $x_0 = 0_T$ and $P = Q_T D_T$.
Now $\Pi_T \colon \mathbb{W}_T \to \mathbb{W}_T$ is the projection onto $\mathbb{E}_T$ along $\mathbb{K}_T$, by definition (3.24), in symbols $\Pi_T = P_{\mathbb{E}_T,\mathbb{K}_T}$. Thus, to see that

$$d\mathcal{N}_T(0_T) = \mathrm{Id} - Q_T D_T = P_{\mathbb{E}_T,\mathbb{K}_T} =: \Pi_T, \tag{5.53}$$

it remains to show that the composition

$$Q_T D_T = P_{\mathbb{K}_T,\mathbb{E}_T}$$

is the projection onto $\mathbb{K}_T$ along $\mathbb{E}_T$. This follows since

$$Q_T D_T \overset{(4.43)}{=} F_T^{-1} d\mathcal{F}_T|_{0_T} = \left(d\mathcal{F}_T|_{0_T}|_{\mathbb{K}_T}\right)^{-1} d\mathcal{F}_T|_{0_T} \colon W_T \to \mathbb{V}_T \to \mathbb{K}_T$$

and $\mathbb{E}_T = \ker d\mathcal{F}_T|_{0_T} = \ker D_T$ and $Q_T = F_T^{-1} \colon \mathbb{V}_T \to \mathbb{K}_T$ where F_T is the restriction of D_T to $\mathbb{K}_T$, see (4.42). □

5.1 Diffeomorphism onto image

Theorem 5.3. *There are open neighborhoods $\mathcal{O}_+ \subset \mathcal{U}_+ \subset \mathcal{W}^{\mathrm{s}}$ of 0_+ and $\mathcal{O}_- \subset \mathcal{U}_- \subset \mathcal{W}^{\mathrm{u}}$ of 0_- such that for every $T \geq T_0$ the restricted gluing map*

$$\gamma_T \colon \mathcal{O}_+ \times \mathcal{O}_- \to \mathcal{M}_T, \quad (w_+, w_-) \mapsto \mathcal{N}_T \circ \wp_T(w_+, w_-)$$

is a diffeomorphism onto its image $\mathcal{O}_T$.

Note that the domain $\mathcal{O}_+ \times \mathcal{O}_-$ of γ_T does not depend on T.

Proof. Given $T_0 \geq 3$ as prior to (5.50), pick $T \geq T_0$. The theorem is a consequence of the quantitative inverse function Theorem A.1 (IFT), where F is given by a representative of γ_T in local coordinate charts; see (2.19). In order to apply the quantitative IFT two conditions, (A.65) and (A.66), are to be checked.

We verify (A.65): Recall that the inverse of the infinitesimal gluing map Γ_T is uniformly bounded by a constant $k = 1/(1-e^{-12\sigma}) > 1$, see Lemma 3.5. So, by Lemma 5.2, the inverse of $d\gamma_T|_{(0_+,0_-)} = dF|_0$ is bounded by k, uniformly in T.

We verify (A.66): To check this condition we choose

$$\boxed{\mathcal{O}_\pm := \mathcal{O}^\pm_{1/8kd} \cap \mathcal{U}_\pm^{4k+1},} \qquad \varepsilon = \tfrac{1}{8kd}, \quad \mu = 4k+1 \geq 5, \tag{5.54}$$

where d is the (T-independent) bound of Π_T from Lemma 3.3 and where the open origin neighborhoods $\mathcal{O}^\pm_\varepsilon$ and $\mathcal{U}^\mu_\pm$ in $\mathcal{W}^{\mathrm{s/u}}$ were defined in Lemma 2.3 and Lemma 4.7, respectively. Since $k \geq 1$, Lemma 4.7 tells that $\mathcal{U}_\pm^{4k+1} \subset \mathcal{U}_\pm^2 =: \mathcal{U}_\pm$, hence $\mathcal{O}_\pm \subset \mathcal{U}_\pm$. Pick $(w_+, w_-) \in \mathcal{O}_+ \times \mathcal{O}_- \subset \mathcal{U}_+ \times \mathcal{U}_-$.

Recalling (2.19) we shall investigate the operator norm of the difference

$$d\gamma_T|_{(w_+,w_-)} \circ (\theta_{w_+}^{-1}, \theta_{w_-}^{-1}) - d\gamma_T|_{(0_+,0_-)} \colon \mathbb{E}^+ \times \mathbb{E}^- \to \mathbb{W}_T.$$

Abbreviate $x_1 := \wp_T(w_+, w_-)$. By definition of γ_T and of $\Theta_T(w_+, w_-)$ we get

$$\begin{aligned} & d\gamma_T|_{(w_+,w_-)} \circ (\theta_{w_+}^{-1}, \theta_{w_-}^{-1}) - d\gamma_T|_{(0_+,0_-)} \\ & \stackrel{(5.51)}{=} d\mathcal{N}_T|_{x_1} \circ \mathcal{P}_T \circ (\theta_{w_+}^{-1}, \theta_{w_-}^{-1}) - d\mathcal{N}_T|_{0_T} \circ \mathcal{P}_T - d\mathcal{N}_T|_{x_1} \circ \mathcal{P}_T + d\mathcal{N}_T|_{x_1} \circ \mathcal{P}_T \\ & \stackrel{(2.21)}{=} -d\mathcal{N}_T|_{x_1} \circ \Theta_T(w_+, w_-) + (d\mathcal{N}_T|_{x_1} - d\mathcal{N}_T|_{0_T}) \circ \mathcal{P}_T. \end{aligned} \tag{5.55}$$

By Remark B.5 for $Q := Q_T$ and $P := Q_T D_T$ and since the projection $\mathrm{Id} - P = \Pi_T$, see (5.53), has a (T-independent) bound d by Lemma 3.3 we get that

$$\|d\mathcal{N}_T|_{x_1}\| \stackrel{(B.79)}{\le} \|(\mathrm{Id} + Q_T\, df(x_1) - P)^{-1}\| \cdot \|\mathrm{Id} - P\| \stackrel{(B.81)}{\le} 2d. \tag{5.56}$$

Since $(w_+, w_-) \in \mathcal{O}^+_{1/8kd} \times \mathcal{O}^-_{1/8kd}$, by Lemma 2.3, we have that

$$\|\Theta_T(w_+, w_-)\| \le \tfrac{1}{8kd}. \tag{5.57}$$

Furthermore, abbreviating $x_0 := 0_T$, then $d\mathcal{N}_T|_{0_T} = \mathrm{Id} - P$ by (B.79) with x_1 replaced by x_0 and using that $Q_T\, df(x_0) = Q_T D_T = P$. Thus we get that

$$(d\mathcal{N}_T|_{x_1} - d\mathcal{N}_T|_{0_T}) \circ \mathcal{P}_T \stackrel{(B.79)}{=} \Big(\big(\mathrm{Id} + Q_T\, df|_{x_1} - P\big)^{-1} - \mathrm{Id}\Big)\, (\mathrm{Id} - P) \circ \mathcal{P}_T. \tag{5.58}$$

Observe that

$$(\mathrm{Id} - P) \circ \mathcal{P}_T \stackrel{(2.20)}{=} (\mathrm{Id} - Q_T D_T) \circ d\wp_T|_{(0_+,0_-)} \stackrel{(5.52)}{=} \Gamma_T. \tag{5.59}$$

Since $(w_+, w_-) \in \mathcal{U}_+^{4k+1} \times \mathcal{U}_-^{4k+1}$, it follows from Lemma 4.7 that $\|x_1\|_{1,2} = \|\mathcal{P}_T(w_+, w_-)\|_{1,2} < \min\{\frac{\delta}{8}, \delta_{4k+1}\}$. Hence x_1 lies in $\mathcal{U}_T^{4k+1}$ by Corollary 4.5. Therefore, by Lemma 4.4, it follows that $\|d\mathcal{F}_T(x_1) - D_T\| \le \frac{1}{(4k+1)c}$.
In view of Remark B.8 with U_0 given by $\mathcal{U}_0$ we obtain

$$\|\big(\mathrm{Id} + Q_T\, df|_{x_1} - P\big)^{-1} - \mathrm{Id}\| \le \tfrac{1}{4k}. \tag{5.60}$$

By Lemma 3.5 we have $\|\Gamma_T\| \le 1$. Combining this fact with (5.58), (5.59), and (5.60) we conclude

$$\|(d\mathcal{N}_T|_{x_1} - d\mathcal{N}_T|_{0_T}) \circ \mathcal{P}_T\| \le \|\big(\mathrm{Id} + Q_T\, df|_{x_1} - P\big)^{-1} - \mathrm{Id}\| \cdot \|\Gamma_T\| \le \tfrac{1}{4k}. \tag{5.61}$$

By (5.55), (5.56), (5.57), and (5.61) we conclude

$$\|d\gamma_T|_{(w_+,w_-)} \circ (\theta_{w_+}^{-1}, \theta_{w_-}^{-1}) - d\gamma_T|_{(0_+,0_-)}\| \le \tfrac{2d}{8kd} + \tfrac{1}{4k} = \tfrac{1}{2k}.$$

This verifies (A.66). Corollary A.2 concludes the proof of Theorem 5.3. ☐

5.2 Evaluation maps and convergence in C^m

Definition 5.4. Consider the evaluation maps defined by

$$\mathrm{ev}\colon \mathbb{W}^+ \times \mathbb{W}^- \to \mathbb{R}^n \times \mathbb{R}^n, \quad (w_+, w_-) \mapsto (w_+(0), w_-(0))$$

and, for $T > 0$, by

$$\mathrm{ev}_T\colon \mathbb{W}_T \to \mathbb{R}^n \times \mathbb{R}^n, \quad w \mapsto (w(-T), w(T))\,.$$

Observe that both evaluation maps are linear. Furthermore, for $T \geq 3$ we have $\mathrm{ev}_T \circ \mathcal{P}_T(w_+, w_-) = (w_+(0), w_-(0) = \mathrm{ev}(w_+, w_-)$.[7] So there is the identity

$$\mathrm{ev}_T \circ \mathcal{P}_T = \mathrm{ev}\colon \mathbb{W}^+ \times \mathbb{W}^- \to \mathbb{R}^n \times \mathbb{R}^n \tag{5.62}$$

whenever $T \geq 3$. Therefore, for tangent maps, we get the identity

$$T^m\mathrm{ev}_T \circ T^m\mathcal{P}_T = T^m\mathrm{ev}\colon T^m\mathbb{W}^+ \times T^m\mathbb{W}^- \to T^m\mathbb{R}^n \times T^m\mathbb{R}^n \tag{5.63}$$

whenever $m \in \mathbb{N}$ and $T \geq 3$.

Lemma 5.5. $\|\mathrm{ev}_T\| \leq 2\sqrt{2}$.

Proof. $|\mathrm{ev}_T w|^2_{\mathbb{R}^n\times\mathbb{R}^n} = |w(-T)|^2 + |w(T)|^2 \leq 2\|w\|^2_\infty \leq 8\|w\|^2_{\mathbb{W}_T}$ by (3.23). □

To motivate Theorem 5.6 below we first check the infinitesimal version in case $m = 0$, see (1.3). The linearized evaluation maps are given by

$$\begin{aligned} d\mathrm{ev}|_{(0_+,0_-)} = \mathrm{ev}\colon \mathbb{E}^+ \times \mathbb{E}^- &\to \mathbb{R}^n \times \mathbb{R}^n \\ (\xi, \eta) &\mapsto (\xi(0), \eta(0)) \end{aligned}$$

and

$$\begin{aligned} d\mathrm{ev}_T|_{0_T} = \mathrm{ev}_T\colon \mathbb{E}_T &\to \mathbb{R}^n \times \mathbb{R}^n \\ \zeta &\mapsto (\zeta(-T), \zeta(T))\,. \end{aligned}$$

By Lemma 5.2 we get that

$$d\left(\mathrm{ev}_T \circ \gamma_T\right)|_{(0_+,0_-)} = \mathrm{ev}_T \circ d\gamma_T|_{(0_+,0_-)} = \mathrm{ev}_T \circ \Gamma_T\colon \mathbb{E}^+ \times \mathbb{E}^- \to \mathbb{E}_T.$$

[7] By definition (2.11) of w_T and the cut-off function β, we get the identities

$$\begin{aligned} w_T(-T) - w_+(0) &= (1 - \beta(-T+2))\, w_+(0) + \beta(-T-2)\, w_-(-2T) - w_+(0) \\ &= -\beta(-T+2) w_+(0) + \beta(-T-2)\, w_-(-2T) \\ &= 0 \quad \text{for } T \geq 3 \\ w_T(T) - w_-(0) &= (1 - \beta(T+2))\, w_+(2T) + \beta(T-2)\, w_-(0) - w_-(0) \\ &= (1 - \beta(T+2))\, w_+(2T) - (1 - \beta(T-2))\, w_-(0) \\ &= 0 \quad \text{for } T \geq 3. \end{aligned}$$

Thus, by definition of the evaluation maps, for $T \geq 3$ we get that

$$\mathrm{ev}_T(w_T) - \mathrm{ev}(w_+, w_-) = (w_T(-T) - w_+(0), w_T(T) - w_-(0)) = (0, 0).$$

Thus, for $(\xi,\eta) \in \mathbb{E}^+ \times \mathbb{E}^-$ and by (3.29), we obtain

$$\begin{aligned} d\left(\mathrm{ev}_T \circ \gamma_T\right)|_{(0_+,0_-)}(\xi,\eta) &= \left(\Gamma_T(\xi,\eta)(-T), \Gamma_T(\xi,\eta)(T)\right) \\ &= \big(\xi(0) + e^{-2TA_-}\eta(0), e^{-2TA_+}\xi(0) + \eta(0)\big) \\ &\overset{T\to\infty}{\to} (\xi(0),\eta(0)) = d\mathrm{ev}|_{(0_+,0_-)}(\xi,\eta). \end{aligned}$$

This confirms the infinitesimal version of Theorem 5.6 in case $m = 0$.

Theorem 5.6 (Local gluing – C^m)**.** *Let $m \in \mathbb{N}_0$. Consider the gluing map $\gamma_T : \mathcal{U}_+ \times \mathcal{U}_- \to \mathbb{W}_T$ from (5.51). In the limit $T \to \infty$ the tangent map diagram*

$$\begin{array}{ccc} T^m\mathcal{U}_+ \times T^m\mathcal{U}_- & \xrightarrow{\quad T^m\mathrm{ev} \quad} & T^m\mathbb{R}^n \times T^m\mathbb{R}^n \\ & {\scriptstyle T^m\gamma_T = T^m\mathcal{N}_T \circ T^m\mathcal{P}_T} \searrow \qquad \nearrow {\scriptstyle T^m\mathrm{ev}_T} & \\ & T^m\mathcal{M}_T & \end{array}$$

commutes. More precisely, it holds that

$$\lim_{T\to\infty} T^m\mathrm{ev}_T \circ T^m\gamma_T = T^m\mathrm{ev}$$

in $C^0(T^m\mathcal{U}_+ \times T^m\mathcal{U}_-, T^m\mathbb{R}^n \times T^m\mathbb{R}^n)$.

Proof. Let $W_\pm \in T^m\mathcal{U}_\pm$. For $W_T := T^m\mathcal{P}_T(W_+, W_-) \in T^m\mathbb{W}_T$ we obtain

$$\begin{aligned} \overbrace{(V_-, V_+)}^{T^m\mathbb{R}^n \times T^m\mathbb{R}^n} &:= \left(T^m\mathrm{ev}_T \circ T^m\gamma_T\right)(W_+, W_-) - T^m\mathrm{ev}(W_+, W_-) \\ &= T^m\mathrm{ev}_T \circ T^m\mathcal{N}_T(W_T) - T^m\mathrm{ev}(W_+, W_-) \\ &= T^m\mathrm{ev}_T\big(T^m\mathcal{N}_T(W_T) - W_T\big) + \underbrace{T^m\mathrm{ev}_T(W_T) - T^m\mathrm{ev}(W_+, W_-)}_{=0 \text{ by } (5.63)} \end{aligned}$$

where equality one is definition (5.51) of γ_T and equality two by adding zero.

In view of Lemma 5.7 and Lemma 5.8 below, by Theorem B.12 there exists a constant c, independent of T, such that

$$\begin{aligned} &\|(T^m\mathcal{N}_T - \mathrm{id})(W_T)\|_{T^m\mathbb{W}_T} \\ &\overset{(B.85)}{\leq} c\,\|T^m\mathcal{F}_T(W_T)\|_{T^m\mathbb{V}_T}\Big(1 + \|T^m\mathcal{F}_T(W_T)\|_{T^m\mathbb{V}_T}\Big) \\ &\overset{(4.31)}{\leq} cCe^{-\varepsilon T}\Big(1 + Ce^{-\varepsilon T}\Big) \end{aligned}$$

where the second inequality is by exponential decay (4.31) with constant $C = C(K_+, K_-)$ and $K_\pm$ depending on m. In particular, there exists a constant $T_m = T_m(K_+, K_-) > 0$ such that if $T > T_m$, then

$$1 + Ce^{-\varepsilon T} \leq 2.$$

Therefore, by the uniform–in–T Sobolev estimate (3.23) we get

$$\begin{aligned} \|(T^m\mathcal{N}_T - \mathrm{id})(W_T)\|_{L^\infty_{[-T,T]}} &\le 2\,\|(T^m\mathcal{N}_T - \mathrm{id})(W_T)\|_{T^m\mathbb{W}_T} \\ &\le 4cCe^{-\varepsilon T} \end{aligned} \tag{5.64}$$

for every $T > T_m$. Putting things together, using that $\mathrm{ev}_T \circ \mathcal{P}_T = \mathrm{ev}$ by (5.62), we obtain exponential decay

$$\begin{aligned} &|(T^m\mathrm{ev}_T \circ T^m\gamma_T)\,(W_+, W_-) - T^m\mathrm{ev}(W_+, W_-)|^2_{T^m\mathbb{R}^n \times T^m\mathbb{R}^n} \\ &= |V_-|^2_{T^m\mathbb{R}^n} + |V_+|^2_{T^m\mathbb{R}^n} \\ &\le 2(4cC)^2 e^{-2\varepsilon T} \end{aligned}$$

whenever $T > T_m$. But in finite dimensions pointwise convergence implies convergence of the operators.[8] By the uniformity of exponential decay in Proposition 4.1 we have uniform convergence in $C^0(T^m\mathcal{U}_+ \times T^m\mathcal{U}_-, T^m\mathbb{R}^n \times T^m\mathbb{R}^n)$. □

In the following, by iterated identification of the space with the zero section of its tangent space, we can interpret $0_T \in \mathbb{W}_T$ as an element of $T^m\mathbb{W}_T$. Since $\mathcal{N}_T(0_T) = 0_T$, see (4.47), we have $T^m\mathcal{N}_T(0_T) = 0_T$ and therefore

$$dT^m\mathcal{N}_T(0_T)\colon T_{0_T}T^m\mathbb{W}_T \to T_{0_T}T^m\mathbb{W}_T.$$

Since $\mathbb{W}_T$ itself is a vector space, we have a canonical isomorphism of $T_{0_T}T^m\mathbb{W}_T$ with $(\mathbb{W}_T)^{\times 2^m}$.

Lemma 5.7. *Given $m \in \mathbb{N}_0$ and $T \ge 1$, let d be the T-independent constant provided by Lemma 3.3, then $\|dT^m\mathcal{N}_T(0_T)\| \le d^{2^m}$. In particular, the norm is uniformly bounded independent of T.*

Proof. With respect to the splitting $T_{0_T}T^m\mathbb{W}_T = (\mathbb{W}_T)^{\times 2^m}$ we have the block decomposition $dT^m\mathcal{N}_T(0_T) = \mathrm{diag}\,(d\mathcal{N}_T(0_T), \dots, d\mathcal{N}_T(0_T))$. By formula (4.47) we have $d\mathcal{N}_T(0_T) = \Pi_T$. Hence the estimate follows from Lemma 3.3. □

Lemma 5.8. *Given $W_\pm \in T^m\mathcal{U}_\pm$, the norm of $W_T := T^m\mathcal{P}_T(W_+, W_-)$ is uniformly bounded in terms of the norms of W_+ and W_-, independent of T.*

Proof. By (4.36) the same estimate as in the proof of Lemma 4.7 for w_T yields

$$\|W_T\|^2_{T^m\mathbb{W}_T} \le 2(1 + \|\beta'\|^2_\infty)\left(\|W_+\|^2_{T^m\mathbb{W}_+} + \|W_-\|^2_{T^m\mathbb{W}_-}\right).$$

□

[8] In finite dimension, given a sequence of matrizes, then weak (and strong) convergence means that each matrix entry converges. In particular, the two notions of convergence are equivalent.

A Quantitative inverse function theorem

Let $F\colon X \to Y$ be a map between Banach spaces. Suppose that at a point $x \in X$ the derivative $dF|_x\colon X \to Y$ exists. If this bounded linear map is bijective then its inverse ${dF|_x}^{-1}$ is not only linear but, by the open mapping theorem, also bounded.

The quantitative version of the inverse function theorem (**IFT**) follows of the proof of the usual IFT explained in [MS04, App. A.3], although McDuff-Salamon never state explicitly the quantitative version. Therefore, for the reader's convenience, we state the quantitative version of the IFT and explain how it follows from the arguments in [MS04, App. A.3].

We denote by $B_r(x;X)$ the open ball of radius r centered at x in the Banach space X. We often abbreviate $B_r(x) := B_r(x;X)$ and $B_r := B_r(0;X)$.

Theorem A.1 (Quantitative inverse function theorem)**.** *Let $k, \delta > 0$ be constants. Let $F\colon X \to Y$ be a map between Banach spaces, continuously differentiable on the open radius-δ ball B_δ about $0 \in X$, such that $dF|_0$ is bijective and*

$$\|{dF|_0}^{-1}\| \le k \tag{A.65}$$

and

$$\|dF|_x - dF|_0\| \le \tfrac{1}{2k}, \quad \forall x \in B_\delta. \tag{A.66}$$

In this case the following is true. The restriction of F to B_δ is injective, the image $F(B_\delta)$ is open and contains the ball $B_{\delta/2k}$, the inverse $F^{-1}\colon F(B_\delta) \to B_\delta$ is of class C^1, and

$$d(F^{-1})|_y = (dF|_{F^{-1}(y)})^{-1} \tag{A.67}$$

for every $y \in F(B_\delta)$.

Corollary A.2. *If in Theorem A.1 in addition $F\colon B_\delta \to Y$ is of class C^ℓ for some $\ell \in \mathbb{N}$, then so is F^{-1}. In particular, in case $\ell = \infty$ the restriction $F|\colon B_\delta \to F(B_\delta)$ is a diffeomorphism onto its image.*

Proof. Induction, using the chain and Leibniz rules, together with (A.67). $\square$

The proof of Theorem A.1 is based on the following lemma.

Lemma A.3 ([MS04, Le. A.3.2])**.** *Let $\gamma < 1$ and R be positive real numbers. Let X be a Banach space, $x_0 \in X$, and $\psi\colon B_R(x_0) \to X$ be a continuously differentiable map such that*

$$\|\mathbb{1} - d\psi(x)\| \le \gamma$$

for every $x \in B_R(x_0)$. Then the following holds. The map ψ is injective and ψ maps $B_R(x_0)$ into an open set in X such that

$$B_{R(1-\gamma)}(\psi(x_0)) \subset \psi(B_R(x_0)) \subset B_{R(1+\gamma)}(x_0). \tag{A.68}$$

The inverse $\psi^{-1}\colon \psi(B_R(x_0)) \to B_R(x_0)$ is continuously differentiable and

$$d(\psi^{-1})|_y = (d\psi|_{\psi^{-1}(y)})^{-1}. \tag{A.69}$$

Proof of Theorem A.1. This is basically the proof of the usual IFT given in [MS04, App. A.3]. We assume without loss of generality $x_0 = 0$ and $F(0) = 0$. We consider the map $\psi\colon B_\delta \to X$ defined by

$$\psi(x) := D^{-1}F(x)$$

where $D := dF|_0$. For $x \in B_\delta$ we estimate

$$\|\mathbb{1} - d\psi|_x\| = \|D^{-1}(D - dF|_x)\| \leq \|D^{-1}\| \cdot \|D - dF|_x\| \leq k \cdot \tfrac{1}{2k} = \tfrac{1}{2}.$$

It follows from Lemma A.3 with $R = \delta$ and $\gamma = \frac{1}{2}$ that ψ has a continuously differentiable inverse on $B_\delta(0; X)$ and that $\psi(B_\delta(0; X))$ is an open set containing $B_{\delta/2}(0; X)$. Since $F = D \circ \psi$ and $\|dF|_0{}^{-1}\| \leq k$ we get, respectively, inclusion one and two

$$F(B_\delta(0; X)) = D \circ \psi(B_\delta(0; X)) \supset DB_{\delta/2}(0; X) \supset B_{\delta/2k}(0; Y).$$

The inverse of $F = D\psi$ is given by

$$F^{-1}(y) = \psi^{-1}(D^{-1}y).$$

The inverse F^{-1} is continuously differentiable, since ψ^{-1} is, and the formula $d(F^{-1})|_y = (dF|_{F^{-1}(y)})^{-1}$ follows by differentiating $F \circ F^{-1} = \mathrm{id}_Y$. □

B Newton-Picard without quadratic estimates

The Newton-Picard map is usually defined via the Newton-Picard iteration method. To show that Newton-Picard iteration is a contraction one needs to calculate troublesome quadratic estimates. Based on [MS04, App. A.3] we explain how the Newton-Picard map can as well be defined even if there are no quadratic estimates available. The Newton-Picard map $\mathcal{N}$ obtained in this way is still continuously differentiable. This fact is not mentioned in [MS04, App. A.3] and therefore we prove this fact in the present article; see Appendix B.1.

For induction arguments, e.g. the one in Section 5.2, tangent maps are much more suitable than differentials. Therefore we estimate, in Appendix B.2, the tangent map difference $T\mathcal{N} - \mathrm{Id}$.

Notation. Throughout Appendix B the letter f denotes a map between Banach spaces, not a Morse function as in the principal part of this article.

B.1 Newton-Picard map

The definition of the Newton-Picard map requires the following proposition from [MS04, App. A.3]. The proof can actually be interpreted in terms of the Newton-Picard iteration as is explained in [MS04, Rmk. A.3.5] in case $x_0 = x_1$.

Proposition B.1 ([MS04, Prop. A.3.4])**.** *Let X and Y be Banach spaces, $U \subset X$ be an open set, and $f\colon U \to Y$ be a continuously differentiable map. Let*

$x_0 \in U$ be a suitable initial point in the sense that $D := df(x_0)\colon X \to Y$ is surjective and has a (bounded linear) right inverse $Q\colon Y \to X$. Choose positive constants δ and c such that $\|Q\| \le c$, the open radius-δ ball about x_0 satisfies $B_\delta(x_0; X) \subset U$, and

$$\|x - x_0\| < \delta \qquad \Rightarrow \qquad \|df(x) - D\| \le \frac{1}{2c}. \tag{B.70}$$

Suppose that $x_1 \in X$ is an approximate zero of f near x_0 in the sense that

$$\|x_1 - x_0\| < \frac{\delta}{8}, \qquad \|f(x_1)\| < \frac{\delta}{4c}. \tag{B.71}$$

Then there exists a unique zero $x \in X$ near the initial point x_0 such that

$$f(x) = 0, \qquad x - x_1 \in \operatorname{im} Q, \qquad \|x - x_0\| < \delta. \tag{B.72}$$

Moreover, the distance between the detected zero x and the chosen approximate zero x_1 is controlled by $f(x_1)$, more precisely

$$\|x - x_1\| \le 2c\|f(x_1)\|. \tag{B.73}$$

Definition B.2. Based on the proposition we define the **Newton-Picard map** $\mathcal{N}$ as follows. Define an open subset of U by

$$\begin{aligned} U_0 = U_0(\delta) :&= \left(B_{\frac{\delta}{8}}(x_0; X) \cap \left\{\|f(\cdot)\|_Y < \tfrac{\delta}{4c}\right\}\right) \\ &\overset{(B.70)}{\subset} \left\{\|df(\cdot) - D\|_{\mathcal{L}(X,Y)} < \tfrac{1}{2c}\right\} \end{aligned} \tag{B.74}$$

and a map

$$\mathcal{N} = \mathcal{N}^f_{x_0,Q}\colon X \supset U \supset U_0 \to X, \quad x_1 \mapsto x \tag{B.75}$$

which maps a point x_1, thought of as an approximate zero of f, to the unique zero x in the δ-ball about x_0 whose difference $x - x_1$ lies in the image of the right inverse Q. Note that the domain U_0 of definition of the Newton-Picard map $\mathcal{N}$ depends on the choice of equivalent norms on X and Y.

Remark B.3. The uniqueness statement implies that $\mathcal{N}|_{f^{-1}(0)\cap U_0} = \mathrm{id}$.

Theorem B.4. *The Newton-Picard map $\mathcal{N}$ is continuously differentiable.*

Proof. We first recall how the zero x in Proposition B.1 is found from a given approximate zero $x_1 \in U_0$. One considers the map defined by

$$\psi_{x_1}\colon X \supset U \supset B_\delta(x_0) \to X, \quad x \mapsto x + Q\big(f(x) - D\,(x - x_1)\big). \tag{B.76}$$

The map ψ_{x_1} is continuously differentiable, because f is, and by (B.70) the differential at any $x \in B_\delta(x_0)$ satisfies

$$\|\mathrm{Id} - d\psi_{x_1}(x)\| \le c\|df(x) - D\| \le \frac{1}{2}. \tag{B.77}$$

Moreover, according to [MS04, Proof of Prop. A.3.4] the map $\psi_{x_1}\colon B_\delta(x_0)\to X$ is injective and the Newton-Picard map is given by $\mathcal{N}(x_1) := \psi_{x_1}^{-1}(x_1)$.[9]

To show that $\mathcal{N}$ is differentiable we consider the map

$$\Psi\colon B_\delta(x_0)\times U_0 \to X\times U_0,\quad (x,x_1)\mapsto (\psi_{x_1}(x),x_1)\,.$$

The differential of Ψ at a point $(x,x_1)\in B_\delta(x_0)\times U_0$ is the linear map

$$d\Psi|_{(x,x_1)} = \begin{pmatrix} d\psi_{x_1}|_x & (\partial_{x_1}\psi_{x_1})|_x \\ 0 & \mathrm{Id}\end{pmatrix} \overset{\text{(B.76)}}{=} \begin{pmatrix} d\psi_{x_1}|_x & P \\ 0 & \mathrm{Id}\end{pmatrix} : X\times X\to X\times X$$

where

$$P := QD\colon X\to X$$

is a projection.[10] Since the bound in (B.77) is <1, the linear map $d\psi_{x_1}(x)\colon X\to X$ is invertible, therefore so is $d\Psi|_{(x,x_1)}$ with inverse

$$(d\Psi|_{(x,x_1)})^{-1} = \begin{pmatrix} (d\psi_{x_1}|_x)^{-1} & -(d\psi_{x_1}|_x)^{-1}P \\ 0 & \mathrm{Id}\end{pmatrix}.$$

Therefore, by the inverse function theorem [MS04, Thm. A.3.1], the map Ψ is injective in a neighborhood of $(x,x_1)\in B_\delta(x_0)\times U_0$ with continuously differentiable inverse.
It follows that the Newton-Picard map $\mathcal{N}$ is differentiable, too, with differential $d\mathcal{N}(x_1)\colon X\to X$ given by the formula

$$d\mathcal{N}(x_1) = (d\psi_{x_1}|_{x_1})^{-1}(\mathrm{Id}-P) \overset{\text{(B.76)}}{=} \Big(\mathrm{Id}+Q\,df(x_1)-P\Big)^{-1}(\mathrm{Id}-P)\,. \tag{B.79}$$

The differential $d\mathcal{N}(x_1)$ depends continuously on x_1 since $df(x_1)$ does. □

Remark B.5 (The inverse in (B.79))**.** Let $A := Q\,(D-df|_{x_1}) = P - Q\,df|_{x_1}$, then $\|A\|\le\frac{1}{2}$ since $\|Q\|\le c$ and by (B.70). The inverse for $\mathrm{Id}-A$ is given by

$$\sum_{n=0}^\infty A^n = (\mathrm{Id}-A)^{-1} = \Big(\mathrm{Id}+Q\,df|_{x_1}-P\Big)^{-1} \tag{B.80}$$

where the sum, called **Neumann series**, converges in $\mathcal{L}(X)$ whenever A has operator norm less than 1; for details see e.g. [RS80, (VI.2) p. 191] with $\lambda=1$. Thus there is the estimate

$$\|(\mathrm{Id}+Q\,df|_{x_1}-P)^{-1}\| \le \sum_{n=0}^\infty \|A\|^n \le \sum_{n=0}^\infty \frac{1}{2^n} = \frac{1}{1-\frac{1}{2}} = 2 \tag{B.81}$$

[9] In [MS04, Proof of Prop. A.3.4] it is shown that $x_1\in B_{\delta/2}(\psi_{x_1}(x_0))\subset\psi_{x_1}(B_\delta(x_0))$.

[10] Since $DQ=\mathrm{Id}$, the map

$$P := QD,\quad P^2=P,\quad \operatorname{im}P=\operatorname{im}Q,\quad \ker P=\ker D, \tag{B.78}$$

is the projection $P=P_{\operatorname{im}Q,\ker D}$ onto the image of Q along the kernel of D. Indeed $P^2 = QDQD = QD = P$. It holds $\operatorname{im}P=\operatorname{im}Q$: '$\subset$' true by definition of P. '$\supset$' If $\xi=Q\eta$, then $P\xi = PQ\eta = QDQ\eta = Q\eta = \xi$. It holds $\ker P = \ker D$: '$\subset$' If $P\xi=0$, then $D\xi = DQD\xi = DP\xi = 0$. '$\supset$' true by definition of P.

and therefore

$$\|(\mathrm{Id} + Q\, df|_{x_1} - P)^{-1} - \mathrm{Id}\| = \left\|\sum_{n=1}^{\infty} A^n\right\| \le \sum_{n=1}^{\infty} \|A\|^n = 2 - \frac{1}{2^0} = 1.$$

Remark B.6 (Smoothness)**.** If f in Proposition B.1 is assumed to be not only continuously differentiable, but smooth, then it follows from (B.79) that the Newton-Picard map $\mathcal{N}$ is smooth as well.

Corollary B.7. *By (B.79), it holds $d\mathcal{N}(x_0) = \mathrm{Id} - P$. If additionally $f(x_0) = 0$ in Proposition B.1, then $\mathcal{N}(x_0) = x_0$ by uniqueness.*

Remark B.8. Given $\mu \ge 2$, the restriction of the Newton-Picard map $\mathcal{N} : X \supset U_0 \to X$ in (B.75) to the subset

$$U_0^\mu := U_0 \cap \|df(\cdot) - D\|^{-1}[0, \tfrac{1}{\mu c}), \qquad U_0^2 \overset{\text{(B.74)}}{=} U_0,$$

satisfies, just as above, an estimate of the form

$$\|\big(\mathrm{Id} + Q\, df|_{x_1} - P\big)^{-1} - \mathrm{Id}\| \le \sum_{n=1}^{\infty} \|A\|^n \le \sum_{n=1}^{\infty} \tfrac{1}{\mu^n} = \frac{1}{\mu} \cdot \frac{1}{1 - \frac{1}{\mu}} = \frac{1}{\mu - 1}$$

for every $x_1 \in U_0^\mu$ and where we used that

$$\|A\| = \|Q(D - df|_{x_1})\| \le \|Q\| \cdot \|D - df|_{x_1}\| \le \tfrac{1}{\mu}.$$

B.2 Tangent map

Hypothesis B.9. Consider the situation of Proposition B.1. In this section we assume, in addition, that the map $f \colon X \supset U \to Y$ is two times continuously differentiable. Recall that $x_0 \in U$ is a suitable initial point and δ and c are positive constants, the three of them related by assumption (B.70). Choose $\delta > 0$ smaller, if necessary, such that

$$\|x - x_0\| < \delta \qquad \Rightarrow \qquad \|df(x) - D\| \le \frac{1}{4c}. \tag{B.82}$$

Suppose that there is a constant $c_2 > 0$ such that $\|d^2 f(x)\| \le c_2$ for all $x \in B_\delta(x_0)$. Define

$$\hat{\delta} := \min\big\{\delta, \tfrac{1}{4cc_2}\big\} \subset (0,1).$$

Theorem B.10. *Under Hypothesis B.9 suppose that $x_1 \in U_0(\delta)$; see (B.74). Abbreviate $\mathcal{N} := \mathcal{N}^f_{x_0,Q}$. Then for each $\xi_1 \in X$ which is small in the sense that*

$$\|\xi_1\| < \frac{\delta}{8(1 + \|d\mathcal{N}|_{x_0}\|)}, \qquad \|df|_{x_1}\xi_1\|_Y < \frac{\delta}{4c}, \tag{B.83}$$

there is the estimate

$$\|d\mathcal{N}|_{x_1}\xi_1 - \xi_1\|_X \le 2c \max\left\{\tfrac{\delta}{\hat{\delta}}\|f(x_1)\|_Y, \|df|_{x_1}\xi_1\|_Y\right\}. \tag{B.84}$$

Corollary B.11. *For $(x_1, \xi_1) \in TU_0(\delta)$, see (B.74), there is the estimate*

$$\begin{aligned} &\|d\mathcal{N}|_{x_1}\xi_1 - \xi_1\|_X \\ &\le 2c \max\left\{\tfrac{9(1+\|d\mathcal{N}|_{x_0}\|)\|\xi_1\|}{\hat\delta}\|f(x_1)\|_Y, \tfrac{5c\,\|df|_{x_1}\xi_1\|}{\hat\delta}\|f(x_1)\|_Y, \|df|_{x_1}\xi_1\|_Y\right\}. \end{aligned}$$

Proof of Corollary B.11. Let $x_1 \in U_0(\delta)$; see (B.74). For $\xi_1 \in X$ we define

$$\lambda = \lambda(x_1, \xi_1) := \min\left\{\frac{\delta}{9(1+\|d\mathcal{N}|_{x_0}\|)\|\xi_1\|}, \frac{\delta}{5c\,\|df|_{x_1}\xi_1\|}\right\}.$$

Observe that

$$\frac{1}{\lambda} = \max\left\{\frac{9(1+\|d\mathcal{N}|_{x_0}\|)\|\xi_1\|}{\delta}, \frac{5c\,\|df|_{x_1}\xi_1\|}{\delta}\right\}.$$

Then $\lambda\xi_1$ meets condition (B.83), so there is the estimate

$$\begin{aligned} &\|d\mathcal{N}|_{x_1}\xi_1 - \xi_1\|_X \\ &= \frac{1}{\lambda}\|(d\mathcal{N}|_{x_1} - \mathrm{Id})\lambda\xi_1\|_X \\ &\le \frac{2c}{\lambda}\max\left\{\tfrac{\delta}{\hat\delta}\|f(x_1)\|_Y, \|df|_{x_1}\lambda\xi_1\|_Y\right\} \\ &\le 2c\max\left\{\tfrac{\delta}{\hat\delta}\tfrac{1}{\lambda}\|f(x_1)\|_Y, \|df|_{x_1}\xi_1\|_Y\right\} \\ &= 2c \max\left\{\tfrac{9(1+\|d\mathcal{N}|_{x_0}\|)\|\xi_1\|}{\hat\delta}\|f(x_1)\|_Y, \tfrac{5c\,\|df|_{x_1}\xi_1\|}{\hat\delta}\|f(x_1)\|_Y, \|df|_{x_1}\xi_1\|_Y\right\}. \end{aligned}$$

□

We can summarize the result of this section more compactly in tangent map notation by the following theorem.

Theorem B.12. *Under Hypothesis B.9, given $W_1 = (x_1, \xi_1) \in TU_0(\delta)$, see (B.74), there is a constant C depending on $\|\xi_1\|_X$ and $\|d\mathcal{N}|_{x_0}\|$ such that*

$$\|T\mathcal{N}(W_1) - W_1\|_{TX} \le C\|Tf(W_1)\|_{TY}\Big(1 + \|Tf(W_1)\|_{TY}\Big). \tag{B.85}$$

Proof of Theorem B.12. Write $W_1 = (x_1, \xi_1)$ and observe that

$$T\mathcal{N}(W_1) - W_1 = (\mathcal{N}(x_1) - x_1, d\mathcal{N}|_{x_1}\xi_1 - \xi_1).$$

To the first component apply estimate (B.73) in the McDuff-Salamon Proposition B.1 and to the second component apply Corollary B.11. □

Proof of Theorem B.10. We first consider the Newton-Picard map $\mathcal{N}^{Tf}$ for the tangent map Tf and then compare it with the tangent map $T\mathcal{N}^f$ of $\mathcal{N}^f$.

On $\hat X := TX = X \oplus X \ni (x, \xi)$ and $\hat Y := TY = Y \oplus Y \ni (y, \eta)$ We define norms for $(x, \xi) \in TX = X \oplus X$, respectively $(y, \eta) \in TY = Y \oplus Y$ by

$$\|(x,\xi)\| := \max\{\|x\|, \tfrac{\hat\delta}{\delta}\|\xi\|\}, \qquad \|(y,\eta)\| := \max\{\|y\|, \tfrac{\hat\delta}{\delta}\|\eta\|\}.$$

In the following we study the tangent map $\hat{f} := Tf$ which is defined by $Tf(x,\xi) = (f(x), df|_x\xi)$. The task at hand is to choose the corresponding quantities $\hat{x_0}$, $\hat{D}$, $\hat{Q}$, $\hat{c}$, $\hat{\delta}$, and $\hat{U_0}$ in order to apply Proposition B.1 to $\hat{f}$.
As initial point for $\hat{f}$ on $\hat{U} := TU = U \times X$ we pick $\hat{x_0} := (x_0, 0)$. The operator $\hat{D} := d\hat{f}|_{(x_0,0)} = D \oplus D$ is onto. A right inverse is given by the sum $\hat{Q} := Q \oplus Q$ with bound $\hat{c} = c$.[11] Observe that $B_\delta(\hat{x_0}; TX) \subset U \times X = \hat{U}$. Suppose that $x, \xi \in X$ satisfy the estimate

$$\|(x,\xi) - (x_0, 0)\| = \max\{\|x - x_0\|, \tfrac{\hat{\delta}}{\delta}\|\xi\|\} < \delta.$$

In particular, we have $x \in B_\delta(x_0)$, hence $\|d^2 f(x)\| \le c_2$ and $\|df|_x - D\| \le \frac{1}{4c}$, by (B.82). For elements $\tilde{x}$ and $\tilde{\xi}$ of X consider the operator difference

$$\begin{aligned}
&\left\| \left(d\hat{f}|_{(x,\xi)} - \hat{D} \right) (\tilde{x}, \tilde{\xi}) \right\|_{TY} \\
&= \left\| \begin{bmatrix} df|_x - D & 0 \\ d^2 f|_x(\xi, \cdot) & df|_x - D \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{\xi} \end{bmatrix} \right\|_{TY} \\
&= \max \left\{ \|(df|_x - D)\tilde{x}\|, \tfrac{\hat{\delta}}{\delta} \|(df|_x - D)\tilde{\xi} + d^2 f|_x(\underline{\xi}, \tilde{x})\| \right\} \\
&\overset{\text{(B.82)}}{\le} \max \left\{ \tfrac{1}{4c}\|\tilde{x}\|, \tfrac{\hat{\delta}}{\delta} \tfrac{1}{4c} \|\tilde{\xi}\| \right\} + \tfrac{\hat{\delta}}{\delta} \underline{\delta} \|d^2 f(x)\| \cdot \|\tilde{x}\| \\
&= \tfrac{1}{4c}\|(\tilde{x}, \tilde{\xi})\| + \hat{\delta}\|d^2 f(x)\| \cdot \|\tilde{x}\|.
\end{aligned}$$

Take the supremum over all $\|(\tilde{x}, \tilde{\xi})\| = 1$ to get the operator norm estimate

$$\|d\hat{f}|_{(x,\xi)} - \hat{D}\| \le \tfrac{1}{4c} + \hat{\delta}\|d^2 f(x)\| \le \tfrac{1}{4c} + \tfrac{1}{4cc_2} c_2 = \tfrac{1}{2c}.$$

Thus we have verified condition (B.70) in Proposition B.1 with $\hat{f}$ and $\hat{\delta}$ in place of f and δ. We next check condition (B.71) for $\hat{f}$ and $\hat{\delta}$ and any

$$\hat{x_1} := (x_1, \xi_1) \in X \oplus X, \quad \text{where } x_1 \in U_0(\delta) \text{ and } \xi_1 \text{ satisfies (B.83).}$$

By $x_1 \in U_0(\delta)$ it holds $\|f(x_1)\| < \frac{\delta}{4c}$ and $\|x_1 - x_0\| < \frac{\delta}{8}$. By (B.83) we get (B.71)

$$\|\hat{f}(\hat{x_1})\| = \max\{\|f(x_1)\|, \tfrac{\hat{\delta}}{\delta}\|df|_{x_1}\xi_1\|\} < \max\{\tfrac{\delta}{4c}, \tfrac{\hat{\delta}}{4c}\} = \frac{\delta}{4c} \tag{B.86}$$

and

$$\|\hat{x_1} - \hat{x_0}\| = \max\{\|x_1 - x_0\|, \tfrac{\hat{\delta}}{\delta}\|\xi_1\|\} < \max\{\tfrac{\delta}{8}, \tfrac{\hat{\delta}}{8}\} = \frac{\delta}{8}.$$

[11] Use the bound $\|Q\| \le c$ to obtain the inequality in what follows

$$\begin{aligned}
\|Q \oplus Q\|_{\mathcal{L}(TY,TX)} :&= \sup_{\|(\hat{y},\hat{\eta})\|_{TY}=1} \|(Q \oplus Q)(\hat{y}, \hat{\eta})\|_{TX} \\
&= \sup_{\|(\hat{y},\hat{\eta})\|_{TY}=1} \max\{\|Q\hat{y}\|, \tfrac{\hat{\delta}}{\delta}\|Q\hat{\eta}\|\} \\
&\le c \sup_{\|(\hat{y},\hat{\eta})\|_{TY}=1} \underbrace{\max\{\|\hat{y}\|, \tfrac{\hat{\delta}}{\delta}\|\hat{\eta}\|\}}_{\|(\hat{y},\hat{\eta})\|_{TY}} = c.
\end{aligned}$$

Then Proposition B.1 for $\hat{f}$ and δ yields a unique zero $\hat{x} = (x, \xi)$ of $\hat{f}$ such that

$$f(x) = 0,\ df|_x\xi = 0, \quad x - x_1, \xi - \xi_1 \in \operatorname{im} Q, \quad \max\{\|x - x_0\|, \tfrac{\hat{\delta}}{\delta}\|\xi\|\} < \delta. \tag{B.87}$$

In particular, since $\|x - x_0\| < \delta$ the element x is the same as the one uniquely determined by (B.72) and baptized $\mathcal{N}^f_{x_0,Q}(x_1)$ in (B.75). Moreover, Proposition B.1 for $\hat{f}$ and δ yields that

$$\max\{\|x - x_1\|, \tfrac{\hat{\delta}}{\delta}\|\xi - \xi_1\|\} \le 2c \max\{\|f(x_1)\|, \tfrac{\hat{\delta}}{\delta}\|df|_{x_1}\xi_1\|\}. \tag{B.88}$$

To conclude the proof of Theorem B.10 we need

Proposition B.13. *There is the identity $T\mathcal{N}^f_{x_0,Q} = \mathcal{N}^{Tf}_{(x_0,0),Q\oplus Q}$, that is*

$$x = \mathcal{N}(x_1), \qquad \xi = d\mathcal{N}|_{x_1}\xi_1. \tag{B.89}$$

where we abbreviated $\mathcal{N} = \mathcal{N}^f_{x_0,Q}$.

Proof of Proposition B.13. After (B.87) we already proved $x = \mathcal{N}(x_1)$. By uniqueness it suffices to verify the (three) properties in (B.87) for $d\mathcal{N}|_{x_1}\xi_1$ in place of ξ.

Property 1: $d\mathcal{N}|_{x_1}\xi_1 \in \ker df|_x$. Since $x = \mathcal{N}(x_1)$ and since x_1 lies in the open domain $U_0(\delta)$ of the Newton-Picard map $\mathcal{N}$ in (B.75), hence so does $x_1 + \tau\xi_1$ for any sufficiently small $\tau > 0$, we obtain that

$$df|_x \circ d\mathcal{N}|_{x_1}\xi_1 = \tfrac{d}{d\tau}\big|_0 f \circ \underbrace{\mathcal{N}(x_1 + \tau\xi_1)}_{\in f^{-1}(0)} = 0.$$

Property 2: $d\mathcal{N}|_{x_1}\xi_1 - \xi_1 \in \operatorname{im} Q$. Observe that

$$-\left(d\mathcal{N}|_{x_1}\xi_1 - \xi_1\right) \stackrel{\text{(B.79)}}{=} \Big(\mathrm{Id} + Q\,df|_{x_1} - QD\Big)^{-1} Q \circ df|_{x_1}\xi_1 = Q\eta.$$

That the last equality indeed holds for some $\eta \in Y$ is equivalent to

$$Q \circ df|_{x_1}\xi_1 = \Big(\mathrm{Id} + Q\,df|_{x_1} - QD\Big)Q\eta = Q\,df|_{x_1}Q\eta$$

for some $\eta \in Y$. Since Q is injective it remains to find an $\eta \in Y$ such that

$$df|_{x_1}\xi_1 = df|_{x_1}Q\eta.$$

But the operator $df|_{x_1}Q$ is invertible since it is of the form $\mathrm{Id} - B$ where $B := (D - df|_{x_1})Q$ has norm $\|B\| \le \frac{1}{4}$ due to $\|Q\| \le c$ and by (B.82); cf. Remark B.5.

Property 3: $\|d\mathcal{N}|_{x_1}\xi_1\| < \delta$. As $\mathrm{Id} - QD = d\mathcal{N}|_{x_0}$, see Corollary B.7, we get

$$\begin{aligned}
\|d\mathcal{N}|_{x_1}\xi_1\| &\stackrel{\text{(B.79)}}{=} \|(\mathrm{Id} + Q\,df|_{x_1} - QD)^{-1}(\mathrm{Id} - QD)\xi_1\| \\
&\le \|(\mathrm{Id} + Q\,df|_{x_1} - QD)^{-1}\| \cdot \|d\mathcal{N}|_{x_0}\| \cdot \|\xi_1\| \\
&\stackrel{\substack{\text{(B.81)}\\\text{(B.83)}}}{<} 2\tfrac{\delta}{8}.
\end{aligned}$$

This proves the identities (B.89) and Proposition B.13 ∎

We continue and conclude the proof of Theorem B.10. Since $\xi = d\mathcal{N}|_{x_1}\xi_1$, by (B.89), the estimate (B.88) multiplied by $\frac{\delta}{\bar{\delta}}$ leads to

$$\|(d\mathcal{N}|_{x_1} - \mathrm{Id})\xi_1\|_X \le 2c \max\left\{\tfrac{\delta}{\bar{\delta}}\|f(x_1)\|_Y, \|df|_{x_1}\xi_1\|_Y\right\}.$$

This concludes the proof of Theorem B.10. □

C Exponential decay

Theorem C.1 (Linear uniform exponential decay)**.** *Pick $\varepsilon \in (0, \sigma)$ where $\sigma = \sigma(A)$ is the spectral gap (1.7). Let $m \in \mathbb{N}_0$. Suppose W is a map of Sobolev class $W^{1,2}$ such that*

$$W\colon [0,\infty) \to T^m\mathbb{R}^n, \qquad \partial_s W + T^m\nabla f(W) = 0. \tag{C.90}$$

Then there is a positive constant $c(W)$, depending continuously on W, such that

$$|W(s)| + |\partial_s W(s)| \le c(W)\cdot e^{-\varepsilon s} \tag{C.91}$$

for every $s \ge 0$,

Preparation of proof

By $\mathcal{P}^*(\mathbb{N})$ we denote the collection of all finite non-empty subsets of $\mathbb{N}$. The evaluation map is defined by

$$e\colon \mathcal{P}^*(\mathbb{N}) \to \mathbb{N}, \quad D \mapsto \sum_{j\in D} 2^{j-1}$$

and its inverse is the digit map

$$\mathcal{D} := e^{-1}\colon \mathbb{N} \to \mathcal{P}^*(\mathbb{N}).$$

It can be described as follows. Write $k \in \mathbb{N}$ in binary representation and map it to the subset of $\mathbb{N}$ consisting of all positions of the binary representation of k at which you can find a 1, for example $9 = 1001 \mapsto \{1, 4\}$.

Given a finite non-empty subset $D \subset \mathbb{N}$, in symbols $D \in \mathcal{P}^*(\mathbb{N})$, we consider all partitions of D into $\ell \in \mathbb{N}$ non-empty subsets, namely

$$\mathrm{Part}_\ell(D) := \{\{A_1, \dots, A_\ell\} \subset \mathcal{P}(D) \mid \cup_{i=1}^\ell A_i = D, A_i \cap A_j \overset{i\ne j}{=} \emptyset, \forall i\colon A_i \ne \emptyset\}.$$

Given $\ell \in \mathbb{N}_0$, the ODE (C.90) for the map $W\colon [0,\infty) \to T^\ell\mathbb{R}^n$ is equivalent to a system of 2^ℓ ODEs for 2^ℓ maps $W_0, W_1, \dots, W_{2^\ell-1}\colon [0,\infty) \to \mathbb{R}^n$, namely

$$\partial_s W_0 + \nabla f(W_0) = 0 \tag{C.92}$$

and the $2^\ell - 1$ equations

$$0 = \partial_s W_k + \sum_{\ell\in\mathbb{N}} \Biggl(\sum_{\{A_1,\dots,A_\ell\}\in \mathrm{Part}_\ell \mathcal{D}(k)} D^\ell \nabla f|_{W_0}[W_{e(A_1)}, \dots, W_{e(A_\ell)}] \Biggr) \tag{C.93}$$

where $k = 1, \dots, 2^\ell - 1$.

Remark C.2 (Reformulation of (C.93))**.** Given $k \in \mathbb{N}$, let $\mathfrak{S}(k)$ be the digit sum of the binary representation of k, also referred to as the Hamming weight. Observe that $\mathfrak{S}(k)$ is the cardinality of $\mathcal{D}(k)$. Note that for $\ell > \mathfrak{S}(k)$ the partition set $\mathrm{Part}_\ell \mathcal{D}(k) = \emptyset$ is empty. Note also that $\mathrm{Part}_1 \mathcal{D}(k) = \{\{\mathcal{D}(k)\}\}$. Therefore we can write (C.93) equivalently as the finite sum

$$\begin{aligned} 0 = \;& \partial_s W_k + D\nabla f|_{W_0}[W_k] \\ & + \underbrace{\sum_{\ell=2}^{\mathfrak{S}(k)} \Biggl(\sum_{\{A_1,\dots,A_\ell\}\in \mathrm{Part}_\ell \mathcal{D}(k)} D^\ell \nabla f|_{W_0}[W_{e(A_1)},\dots,W_{e(A_\ell)}] \Biggr)}_{=:\eta}. \end{aligned} \tag{C.94}$$

In the special case where $k = 2^m$ we have $\mathfrak{S}(k) = 1$, hence (C.94) simplifies to

$$0 = \partial_s W_{2^m} + D\nabla f|_{W_0}[W_{2^m}]. \tag{C.95}$$

The following table illustrates (C.94) for $k = 0, \dots, 7$. It is written in binary notation, so the structure of the system becomes visible

$$\begin{aligned}
0)\quad & 0 = \partial_s W_0 + \nabla f(W_0) \\
1)\quad & 0 = \partial_s W_1 + D\nabla f|_{W_0} W_1 \\
10)\quad & 0 = \partial_s W_{10} + D\nabla f|_{W_0} W_{10} \\
11)\quad & 0 = \partial_s W_{11} + D^2\nabla f|_{W_0}[W_1, W_{10}] + D\nabla f|_{W_0} W_{11} \\
100)\quad & 0 = \partial_s W_{100} + D\nabla f|_{W_0} W_{100} \\
101)\quad & 0 = \partial_s W_{101} + D^2\nabla f|_{W_0}[W_1, W_{100}] + D\nabla f|_{W_0} W_{101} \\
110)\quad & 0 = \partial_s W_{110} + D^2\nabla f|_{W_0}[W_{10}, W_{100}] + D\nabla f|_{W_0} W_{110} \\
111)\quad & 0 = \partial_s W_{111} + D^3\nabla f|_{W_0}[W_1, W_{10}, W_{100}] + D^2\nabla f|_{W_0}[W_{10}, W_{101}] \\
& \quad + D^2\nabla f|_{W_0}[W_1, W_{110}] + D^2\nabla f|_{W_0}[W_{11}, W_{100}] + D\nabla f|_{W_0} W_{111}.
\end{aligned}$$

Lemma C.3. *Given $m \in \mathbb{N}_0$, consider maps $W_0, W_1, \dots, W_{2^m-1} \in W^{1,2}([0,\infty), \mathbb{R}^n)$ that satisfy the ODE system (C.92) and (C.93) for every $k = 1, \dots, 2^m-1$. Then the tuple $W := (W_0, \dots, W_{2^m-1}) \in W^{1,2}([0,\infty), \mathbb{R}^{n\cdot 2^m})$ lies in the m-fold tangent space $T^m\mathcal{W}^{\mathrm{s}}$ which means that*

$$\partial_s W + T^m \nabla f(W) = 0.$$

Proof. The proof is by induction on $m \in \mathbb{N}$.

Case $m = 0$. True by assumption.

Induction step $m \Rightarrow m+1$. There are three cases I-III. I. For $k \in \{1, \dots, 2^m - 1\}$ equation (C.93) holds directly by induction hypothesis. II. For $k \in \{2^m + 1, \dots, 2^{m+1} - 1\}$ we linearize (C.93) with respect to W_{k-2^m}. This yields

$$
\begin{aligned}
0 &= \partial_s W_k + \sum_{\ell\in\mathbb{N}} \sum_{\substack{\{A_1,\dots,A_\ell\}\\ \in \mathrm{Part}_\ell \mathcal{D}(k-2^m)}} \\
&\Bigg(\sum_{j=1}^{\ell} D^\ell \nabla f|_{W_0}[W_{e(A_1)},\dots,W_{e(A_{j-1})},W_{e(A_j)+2^m},W_{e(A_{j+1})},\dots,W_{e(A_\ell)}] \\
&\quad + D^{\ell+1}\nabla f|_{W_0}[W_{2^m},W_{e(A_1)},\dots,W_{e(A_\ell)}]\Bigg) \\
&= \partial_s W_k + \sum_{\ell\in\mathbb{N}} \Bigg(\sum_{\{A_1,\dots,A_\ell\}\in \mathrm{Part}_\ell \mathcal{D}(k)} D^\ell \nabla f|_{W_0}[W_{e(A_1)},\dots,W_{e(A_\ell)}]\Bigg)
\end{aligned} \tag{C.96}
$$

To see why the second equation in (C.96) holds note the identity of digit sets

$$\mathcal{D}(k) = \mathcal{D}(k-2^m) \cup \{m+1\}.$$

Moreover, consider the injections defined for $j = 1,\dots,\ell$ by

$$
\begin{aligned}
\iota_j \colon \mathrm{Part}_\ell(\mathcal{D}(k-2^m)) &\hookrightarrow \mathrm{Part}_\ell(\mathcal{D}(k)) = \mathrm{Part}_\ell(\mathcal{D}(k-2^m)\cup\{m+1\}) \\
\{A_1,\dots,A_\ell\} &\mapsto \{A_1,\dots,A_{j-1},A_j\cup\{m+1\},A_{j+1},\dots,A_\ell\}
\end{aligned}
$$

and the injection defined by

$$
\begin{aligned}
I \colon \mathrm{Part}_{\ell-1}(\mathcal{D}(k-2^m)) &\hookrightarrow \mathrm{Part}_\ell(\mathcal{D}(k)) \\
\{A_1,\dots,A_{\ell-1}\} &\mapsto \{\{m+1\},A_1,\dots,A_{\ell-1}\}.
\end{aligned}
$$

Using this notion we can write $\mathrm{Part}_\ell\mathcal{D}(k)$ as the union of pairwise disjoint subsets, namely

$$\mathrm{Part}_\ell\mathcal{D}(k) = \left(\bigcup_{j=1}^{\ell} \iota_j(\mathrm{Part}_\ell(\mathcal{D}(k-2^m)))\right) \cup I(\mathrm{Part}_{\ell-1}(\mathcal{D}(k-2^m))). \tag{C.97}$$

Now the second equation in (C.96) follows from (C.97).

III. It remains to consider the case $k = 2^m$. Linearizing (C.92) with respect to W_0 in direction W_{2^m} we obtain

$$0 = \partial_s W_{2^m} + D\nabla f|_{W_0} W_{2^m}$$

and this equation coincides with (C.95). This proves Lemma C.3. □

Proof of exponential decay

Proof of Theorem C.1 – Exponential decay. The proof is by induction on m.

Case $m = 0$. This follows for instance from the action-energy inequality; see e.g. [FW22b].

Induction step $m \Rightarrow m+1$. Suppose (C.91) is true for m. Then we want to show (C.91) for $m+1$. By induction hypothesis W_k and its derivative $\partial_s W_k$ decay exponentially for $k = 0, \dots 2^m - 1$. It remains to show that as well W_k and its derivative $\partial_s W_k$ decay exponentially for $k = 2^m, \dots, 2^{m+1} - 1$. This follows from Lemma C.4 below in view of (C.94) combined with the induction hypothesis. More precisely, we prove this by induction on k. In the notation A, ξ, η of Lemma C.4 we have $W_k = \xi$, $D\nabla f|_{W_0} = A$ and η is the sum indicated in (C.94).

Observe that if $\ell \geq 2$ and $\{A_1, \dots, A_\ell\} \in \mathrm{Part}_\ell(\mathcal{D}(k))$ then $e(A_j) < k$ for $j = 1, \dots, \ell$. Therefore by induction hypothesis $W_{e(A_j)}$ decays exponentially so that η decays exponentially. Now the exponential decay of W_k follows from Lemma C.4. □

Lemma C.4. *Consider a continuously differentiable family of quadratic matrizes $\mathbb{A} \colon [0, \infty) \to \mathbb{R}^{n \times n}$ and an invertible symmetric matrix $A \in \mathbb{R}^{n \times n}$ with*

$$\lim_{s \to \infty} \|\mathbb{A}(s) - A\| = 0 = \lim_{s \to \infty} \|\mathbb{A}'(s)\|, \qquad \mathbb{A}'(s) := \tfrac{d}{ds}\mathbb{A}(s).$$

Let $\sigma = \sigma(A) > 0$ be the spectral gap, see (1.7). Let $\xi, \eta \colon [0, \infty) \to \mathbb{R}^n$ be continuously differentiable maps such that ξ is of Sobolev class $W^{1,2}$ and

$$\xi'(s) + \mathbb{A}(s)\xi(s) = \eta(s) \tag{C.98}$$

for every $s \geq 0$. Suppose that there are constants $C > 0$ and $\varepsilon \in (0, \sigma)$ such that

$$|\eta(s)| + |\eta'(s)| \leq C e^{-\varepsilon s} \tag{C.99}$$

for every $s \geq 0$. Then there is a positive constant c, depending continuously on the $W^{1,2}$ norm of ξ and the constant C, such that

$$|\xi(s)| \leq c e^{-\varepsilon s}$$

for every $s \geq 0$.

Observe that the exponential decay rate of η is inherited by ξ, as opposed to [RS01, Le. 3.1].

Proof. We follow the proof of [RS01, Le. 3.1]. We shall employ the following facts and assumptions. The norms of a quadratic real matrix B and its transpose B^t are equal. By definition of the spectral gap $\sigma > 0$ it holds that

$$|Av| \geq \sigma |v|$$

for every $v \in \mathbb{R}^n$. Given $\delta > 0$ and $\varepsilon \in (0, \sigma)$, by assumption there is a large time $s_0 = s_0(\delta; \sigma, \varepsilon) > 0$ such that

$$\Big(\|\mathbb{A}'(s)\| + \big(\tfrac{13}{4} + 16 \tfrac{\sigma^2}{(\sigma^2 - \varepsilon^2)} \big) \|\mathbb{A}(s) - A\|^2 \Big) \leq \frac{\sigma^2 - \varepsilon^2}{4} \tag{C.100}$$

pointwise for $s \geq s_0$. The function defined for $s \geq 0$ by

$$\alpha(s) := \tfrac{1}{2}|\xi(s)|^2$$

has derivatives

$$\alpha' = \langle \xi, \xi' \rangle = \langle \xi, \eta - \mathbb{A}\xi \rangle$$

and

$$\alpha'' = \left\langle \xi', \eta - (\mathbb{A} + \mathbb{A}^t)\xi \right\rangle + \langle \xi, \eta' - \mathbb{A}'\xi \rangle .$$

Substitute ξ' according to (C.98), then add $-A + A$ various times, to obtain

$$\begin{aligned}
\alpha'' &= |\mathbb{A}\xi|^2 + |\eta|^2 - 2\langle \mathbb{A}\xi, \eta \rangle - \left\langle \eta, \mathbb{A}^t\xi \right\rangle + \langle \xi, \eta' - \mathbb{A}'\xi \rangle + \left\langle \mathbb{A}\xi, \mathbb{A}^t\xi \right\rangle \\
&= |(\mathbb{A} - A + A)\xi|^2 + |\eta|^2 - 2\langle (\mathbb{A} - A)\xi, \eta \rangle - 2\langle A\xi, \eta \rangle - \left\langle \eta, (\mathbb{A}^t - A)\xi \right\rangle \\
&\quad - \langle \eta, A\xi \rangle + \langle \xi, \eta' - \mathbb{A}'\xi \rangle + \left\langle (\mathbb{A} - A + A)\xi, (\mathbb{A}^t - A + A)\xi \right\rangle \\
&= |(\mathbb{A} - A)\xi|^2 + |A\xi|^2 + 2\langle (\mathbb{A} - A)\xi, A\xi \rangle \\
&\quad + |\eta|^2 - 2\langle (\mathbb{A} - A)\xi, \eta \rangle - 3\langle A\xi, \eta \rangle - \left\langle \eta, (\mathbb{A} - A)^t\xi \right\rangle + \langle \xi, \eta' \rangle - \langle \xi, \mathbb{A}'\xi \rangle \\
&\quad + \left\langle (\mathbb{A} - A)\xi, (\mathbb{A} - A)^t\xi \right\rangle + \left\langle A\xi, (\mathbb{A} - A)^t\xi \right\rangle + \langle (\mathbb{A} - A)\xi, A\xi \rangle + |A\xi|^2.
\end{aligned}$$

Observe that $|A\xi|^2$ appears twice and, in the following, we write this coefficient in the form $2 = \frac{\sigma^2+\varepsilon^2}{\sigma^2} + \frac{\sigma^2-\varepsilon^2}{\sigma^2}$. By Cauchy-Schwarz and Peter-Paul[12] we obtain

$$\begin{aligned}
\alpha'' &\geq \tfrac{\sigma^2+\varepsilon^2}{\sigma^2}|A\xi|^2 + \tfrac{\sigma^2-\varepsilon^2}{\sigma^2}|A\xi|^2 + |\eta|^2 - 3\|\mathbb{A} - A\| \cdot |\xi| \cdot |\eta| \underline{\ - 3|A\xi| \cdot |\eta|} \\
&\quad - |\xi| \cdot |\eta'| - \|\mathbb{A}'\| \cdot |\xi|^2 - \|\mathbb{A} - A\|^2 \cdot |\xi|^2 - 4|A\xi| \cdot \|\mathbb{A} - A\| \cdot |\xi| \\
&\geq (\sigma^2 + \varepsilon^2)|\xi|^2 + \tfrac{\sigma^2-\varepsilon^2}{2\sigma^2} \underbrace{|A\xi|^2}_{\geq \sigma^2|\xi|^2} + \underbrace{\left(\tfrac{\sigma^2-\varepsilon^2}{2\sigma^2} - \underline{\tfrac{\sigma^2-\varepsilon^2}{4\sigma^2}} - \tfrac{\sigma^2-\varepsilon^2}{4\sigma^2} \right)}_{=0} |A\xi|^2 - |\xi| \cdot |\eta'| \\
&\quad - \underbrace{\left(\|\mathbb{A}'\| + \left(1 + \tfrac{9}{4} + 4^2 \tfrac{\sigma^2}{(\sigma^2-\varepsilon^2)}\right) \|\mathbb{A} - A\|^2 \right)}_{\leq \frac{\sigma^2-\varepsilon^2}{4} \text{ by (C.100)}} |\xi|^2 + (1 - 1 \underline{- \tfrac{3^2\sigma^2}{(\sigma^2-\varepsilon^2)}})|\eta|^2 \\
&\geq (\sigma^2 + \varepsilon^2)|\xi|^2 + \tfrac{\sigma^2-\varepsilon^2}{2}(1 - \tfrac{1}{2} - \tfrac{1}{2})|\xi|^2 - \tfrac{3^2\sigma^2}{\sigma^2-\varepsilon^2}|\eta|^2 - \tfrac{1}{\sigma^2-\varepsilon^2}|\eta'|^2 \\
&\geq (2\delta)^2\alpha - c_0 e^{-2\varepsilon s}, \qquad 2\delta^2 := \sigma^2 + \varepsilon^2, \quad c_0 := \tfrac{9\sigma^2+1}{\sigma^2-\varepsilon^2}C^2,
\end{aligned}$$

pointwise for $s \geq s_0$. Inequality two and three is by $|A\xi| \geq \sigma|\xi|$, the final inequality by the η, η' decay assumption (C.99). Observe the estimate

$$2\delta = 2\sqrt{\tfrac{\sigma^2+\varepsilon^2}{2}} = 2\sqrt{\varepsilon^2 + \tfrac{\sigma^2-\varepsilon^2}{2}} > 2\varepsilon.$$

The function defined by

$$\beta(s) := \alpha(s) + \frac{c_0 e^{-2\varepsilon s}}{(2\varepsilon)^2 - (2\delta)^2}$$

[12] $ab \leq \frac{a^2+b^2}{2}$ whenever $a, b \geq 0$

satisfies

$$\begin{aligned}\beta''(s) &= \alpha''(s) + \tfrac{c_0 (2\varepsilon)^2 e^{-2\varepsilon s}}{(2\varepsilon)^2-(2\delta)^2}\\ &\ge (2\delta)^2\alpha + \tfrac{c_0 (2\varepsilon)^2 e^{-2\varepsilon s}}{(2\varepsilon)^2-(2\delta)^2} - c_0 e^{-2\varepsilon s}\,\tfrac{(2\varepsilon)^2-(2\delta)^2}{(2\varepsilon)^2-(2\delta)^2}\\ &= (2\delta)^2\beta(s)\end{aligned}$$

for $s \ge s_0$. This implies, exactly as in the proof of [RS01, Le. 3.1], the following. Firstly $\frac{d}{ds} e^{2\delta s}\beta(s) \le 0$ for $s \ge s_0$,[13] so secondly $e^{2\delta s_0}\beta(s_0) \ge e^{2\delta s}\beta(s)$, and therefore thirdly $\beta(s) \le e^{-2\delta(s-s_0)}\beta(s_0)$ decays even faster than $e^{-2\varepsilon s}$. Thus

$$\alpha(s) = \beta(s) - \tfrac{c_0 e^{-2\varepsilon s}}{(2\varepsilon)^2-(2\delta)^2} < \left(e^{-(2\delta-2\varepsilon)s} e^{2\delta s_0}\beta(s_0) + \tfrac{c_0}{(2\delta)^2-(2\varepsilon)^2}\right) e^{-2\varepsilon s}$$

and therefore

$$|\xi(s)| = \sqrt{2\alpha(s)} < \sqrt{2e^{-(2\delta-2\varepsilon)s} e^{2\delta s_0}\beta(s_0) + \tfrac{2c_0}{(2\delta)^2-(2\varepsilon)^2}}\; e^{-\varepsilon s}$$

for $s \ge s_0$. □

References

[CJS95] R. L. Cohen, J. D. S. Jones, and G. B. Segal. Floer's infinite-dimensional Morse theory and homotopy theory. In *The Floer memorial volume*, volume 133 of *Progr. Math.*, pages 297–325. Birkhäuser, Basel, 1995.

[FW22a] Urs Frauenfelder and Joa Weber. Lagrange multipliers and adiabatic limits I. *viXra e-prints science, freedom, dignity*, pages 1–60, 2022. viXra:2210.0061 v2.

[FW22b] Urs Frauenfelder and Joa Weber. Lagrange multipliers and adiabatic limits II. *viXra e-prints science, freedom, dignity*, pages 1–47, 2022. viXra:2210.0057.

[MS04] Dusa McDuff and Dietmar Salamon. *J-holomorphic curves and symplectic topology*, volume 52 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2004.

[Qin18] Lizhen Qin. On the associativity of gluing. *J. Topol. Anal.*, 10(3):585–604, 2018. arXiv:1107.5527.

[RS80] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I. Functional analysis*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, second edition, 1980. Functional analysis.

[13] Here boundedness of $|\xi(s)|$ enters which is true by the assumption $\xi \in W^{1,2}$.

[RS01] Joel W. Robbin and Dietmar A. Salamon. Asymptotic behaviour of holomorphic strips. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 18(5):573–612, 2001.

[Sim14] Tatjana Simčević. *A Hardy Space Approach to Lagrangian Floer Gluing*. PhD thesis, ETH Zürich, October 2014. https://doi.org/10.3929/ethz-a-010271531.

[Weh12] Katrin Wehrheim. Smooth structures on Morse trajectory spaces, featuring finite ends and associative gluing. In *Proceedings of the Freedman Fest*, volume 18 of *Geom. Topol. Monogr.*, pages 369–450. Geom. Topol. Publ., Coventry, 2012.