

Morse theory for Lagrange multipliers and adiabatic limits

Stephen Schecter^{a,*}, Guangbo Xu^{b,1}

^a *Department of Mathematics
North Carolina State University
Box 8205
Raleigh, NC 27695 USA
+1-919-515-6533*

^b *Department of Mathematics
Princeton University
Fine Hall, Washington Road
Princeton, NJ 08544 USA
+1-312-646-9515*

Abstract

Given two Morse functions f, μ on a compact manifold M , we study the Morse homology for the Lagrange multiplier function on $M \times \mathbb{R}$, which sends (x, η) to $f(x) + \eta\mu(x)$. Take a product metric on $M \times \mathbb{R}$, and rescale its \mathbb{R} -component by a factor λ^2 . We show that generically, for large λ , the Morse-Smale-Witten chain complex is isomorphic to the one for f and the metric restricted to $\mu^{-1}(0)$, with grading shifted by one. On the other hand, in the limit $\lambda \rightarrow 0$, we obtain another chain complex, which is geometrically quite different but has the same homology as the singular homology of $\mu^{-1}(0)$. The isomorphism between the chain complexes is provided by the homotopy obtained by varying λ . Our proofs use both the implicit function theorem on Banach manifolds and geometric singular perturbation theory.

Keywords: Morse homology, geometric singular perturbation theory, exchange lemma, adiabatic limit

Contents

1	Introduction	2
2	Morse homology and Morse-Smale-Witten complex of (\mathcal{F}, g_λ)	7

*Corresponding author

Email addresses: schecter@ncsu.edu (Stephen Schecter), guangbox@math.uci.edu (Guangbo Xu)

¹Current address: Department of Mathematics, University of California, Irvine, Irvine, CA 92697 USA.

3	Adiabatic limit $\lambda \rightarrow \infty$	11
4	Fast-slow system associated with the $\lambda \rightarrow 0$ limit	20
5	Adiabatic limit $\lambda \rightarrow 0$	27
6	Acknowledgements	39
	Appendix A Exchange lemmas	39
	Appendix B Choosing the center manifold	43

1. Introduction

Let M be a compact manifold. Suppose f and μ are Morse functions on M , and 0 is a regular value of μ . Then we consider the Lagrange multiplier function

$$\begin{aligned} \mathcal{F} : M \times \mathbb{R} &\rightarrow \mathbb{R}, \\ (x, \eta) &\mapsto f(x) + \eta\mu(x). \end{aligned}$$

The critical point set of \mathcal{F} is

$$\text{Crit}(\mathcal{F}) = \{(x, \eta) \mid \mu(x) = 0, df(x) + \eta d\mu(x) = 0\}, \quad (1.1)$$

and there is a bijection

$$\begin{aligned} \text{Crit}(\mathcal{F}) &\simeq \text{Crit}(f|_{\mu^{-1}(0)}), \\ (x, \eta) &\mapsto x. \end{aligned}$$

This is a topic which is taught in college calculus.

A deeper story is the Morse theory of \mathcal{F} . Take a Riemannian metric g on M and the standard Euclidean metric e on \mathbb{R} . Denote by g_1 the product metric $g \oplus e$ on $M \times \mathbb{R}$. Then the gradient vector field of \mathcal{F} with respect to g_1 is

$$\nabla \mathcal{F}(x, \eta) = (\nabla f + \eta \nabla \mu, \mu(x)). \quad (1.2)$$

The differential equation for the negative gradient flow of \mathcal{F} is

$$x' = -(\nabla f(x) + \eta \nabla \mu(x)), \quad (1.3)$$

$$\eta' = -\mu(x). \quad (1.4)$$

Let $p_{\pm} = (x_{\pm}, \eta_{\pm}) \in \text{Crit}(\mathcal{F})$, and let $\mathcal{M}(p_-, p_+)$ be the space of orbits of the negative gradient flow that approach p_{\pm} as $t \rightarrow \pm\infty$ respectively. $\mathcal{M}(p_-, p_+)$ is called a *moduli space* of orbits. For generic choice of data (f, μ, g) , this moduli space will be a smooth manifold with dimension

$$\dim \mathcal{M}(p_-, p_+) = \text{index}(p_-) - \text{index}(p_+) - 1. \quad (1.5)$$

When the dimension is zero, we can count the number of elements of the corresponding moduli space and use the counting to define the Morse-Smale-Witten complex. Its homology is called the Morse homology of the pair $(\mathcal{F}, g \oplus e)$.

We shall show that if we rescale the metric on the \mathbb{R} -part, obtaining $g_\lambda := g \oplus \lambda^{-2}e$, $\lambda \in \mathbb{R}^+$, then the chain complex is defined for λ in a generic set Λ^{reg} , and for all such λ we have isomorphic homology groups. The system (1.3)–(1.4) is replaced by

$$x' = -(\nabla f(x) + \eta \nabla \mu(x)), \quad (1.6)$$

$$\eta' = -\lambda^2 \mu(x). \quad (1.7)$$

We then consider the limits of the complex for (\mathcal{F}, g_λ) as λ approaches ∞ and zero. Orbits in the two limits will be completely different geometric objects, but the counting of them gives the same homology groups.

The main results of this paper can be summarized as the following theorems. For precise meaning and necessary assumptions of these theorems, the reader is referred to Theorem 5, Theorem 25 and Theorem 23.

Theorem. *For $\lambda \in \Lambda^{reg}$ sufficiently large, the Morse-Smale-Witten complex \mathcal{C}^λ of the pair (\mathcal{F}, g_λ) is canonically isomorphic to the Morse-Smale-Witten complex of the pair $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$ with grading shifted by one.*

Theorem. *When $\lambda \rightarrow 0$, orbits of solutions to (1.6)–(1.7) converges to certain singular orbits. The counting of isolated singular orbits defines a chain complex \mathcal{C}^0 , which is isomorphic to \mathcal{C}^λ for $\lambda \in \Lambda^{reg}$ sufficiently small.*

As a corollary, we have

Corollary. *The homology of \mathcal{C}^0 is isomorphic to the singular homology of $\mu^{-1}(0)$ with grading shifted by one.*

Our results give a new perspective on classical Morse theory. If we replace μ by $\mu - c$ for any regular value c of μ , then the homology of the hypersurface $\mu^{-1}(c)$ can be computed as the homology of the chain complex \mathcal{C}^0 defined in this paper. The fast flow determined by (1.6)–(1.7) with $\lambda = 0$ is independent of c . Only the slow flow changes as c changes. This allows tracking of the change in homology of $\mu^{-1}(c)$ when c crosses a critical value. However, we do not pursue this point in the present paper.

Throughout this paper, we shall always assume that M , f , μ and g are C^∞ , or *smooth*. A property will be called *generic* if it is true on a countable intersection of open dense sets in a Fréchet space of C^∞ functions.

All chain complexes and homology groups will have \mathbb{Z}_2 coefficients. It should be straightforward to extend our results to integer coefficients. In order to do so, one has to orient the manifold M and various moduli spaces. The isolated orbits will then obtain orientations, and each will be counted with coefficient 1 or -1 in various sums. One then has to keep track of the signs in the adiabatic limit processes. We use \mathbb{Z}_2 coefficients to avoid these complications.

We shall use the notation $\mathbb{R}_+ = (0, \infty)$, $\mathbb{Z}_+ = \{n \in \mathbb{Z} \mid n \geq 0\}$, and we use $\text{index}(p, f)$ to denote the Morse index for p as a critical point of f .

1.1. Outline

In Section 2, we review basic facts about the Morse-Smale-Witten complex, and we construct this complex \mathcal{C}^λ for the function \mathcal{F} and the metric g_λ . Throughout the paper we work with generic data (f, μ, g) .

Let p_- , p_+ be two critical points of \mathcal{F} , and consider orbits of (1.6)–(1.7) from p_- to p_+ for different λ . The energy of such an orbit is $\mathcal{F}(p_-) - \mathcal{F}(p_+)$, which is λ -independent.

For large λ , because of the finiteness of energy, one can show that all orbits from p_- to p_+ will be confined *a priori* in a small neighborhood of $\mu^{-1}(0) \times \mathbb{R}$. In fact, those orbits will be close to the orbits of the negative gradient flow of $f|_{\mu^{-1}(0)}$. This gives the correspondence between the Morse-Smale-Witten chain complex for (\mathcal{F}, g_λ) for large λ , and the complex for $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$. In Section 3 we give two proofs of these facts, one using Fredholm theory and one using geometric singular perturbation theory [1].

On the other hand, in the limit $\lambda \rightarrow 0$, one obtains singular orbits of a fast-slow system. There is one slow variable, η . We will show that for $\text{index}(p_-, \mathcal{F}) - \text{index}(p_+, \mathcal{F}) = 1$, if $\lambda > 0$ is small enough, then near each singular orbit from p_- to p_+ , there exists a unique orbit of (1.6)–(1.7) from p_- to p_+ . This gives a correspondence between the Morse-Smale-Witten chain complex of (\mathcal{F}, g_λ) for λ small, and a complex \mathcal{C}^0 defined by counting singular orbits. Details are in Sections 4 and 5.

Our method of studying the $\lambda \rightarrow 0$ limit is similar to that used in [2]. However, in [2], the slow manifolds are normally hyperbolic, so one can use the classical exchange lemma [1] to relate singular orbits for $\lambda = 0$ to true orbits for small $\lambda > 0$. In the current paper, on the other hand, one cannot avoid the existence of points on the slow manifold where normal hyperbolicity is lost. To overcome this difficulty one needs a small extension of the more general exchange lemma proved in [3]. A review of exchange lemmas including this extension is provided in Appendix A.

1.2. Motivation

The motivation for our work comes from quantum physics and symplectic geometry. This subsection is independent of the rest of the paper.

To study quantum physics theories, mathematicians usually consider moduli spaces instead of path integrals. A moduli space is the space of solutions to some nonlinear differential equation. For example, in Morse theory, which was interpreted by Witten [4] as supersymmetric quantum mechanics, we consider the moduli space of gradient flow lines, which are solutions to the gradient flow equation. Moduli spaces are also used in symplectic geometry. For example, to define the Gromov-Witten invariant of a symplectic manifold (X, ω) , we consider the moduli space of pseudoholomorphic curves introduced in [5], which are solutions to a nonlinear Cauchy-Riemann equation on a Riemann surface Σ . These two examples are both “conformally invariant,” that is, the differential equations are independent of the size of the domain, which are the real line \mathbb{R} in the first example and the Riemann surface Σ in the second example.

We can integrate certain cohomology classes over the moduli spaces. In quantum field theory these integrals are actually correlation functions, which are the most important computable quantities. There are algebraic structures on the set of such integrals. In Morse theory, the integrals can be interpreted as counting the isolated gradient flow lines; they define a chain complex, whose homology is the singular homology of the manifold. In Gromov-Witten theory, the counting of pseudoholomorphic curves gives rise to the so-called quantum cohomology of the symplectic manifold (X, ω) .

Our two examples are related by the theory of Floer homology, which was introduced by Floer in 1980s. In Gromov-Witten theory, if the Riemann surface Σ is the cylinder $\mathbb{R} \times S^1$, then we can use a Hamiltonian function $H : X \rightarrow \mathbb{R}$ to perturb the Cauchy-Riemann equation on Σ . More precisely, if (t, s) is the standard coordinate on the cylinder and J is an almost complex structure on (X, ω) , then Floer's equation is the following PDE for maps $u : \mathbb{R} \times S^1 \rightarrow X$:

$$\frac{\partial u}{\partial t} + J \left(\frac{\partial u}{\partial s} \right) + \nabla H(u) = 0. \quad (1.8)$$

This can be viewed as the negative gradient flow equation of a certain function on the loop space of X . Moreover, if we shrink loops to points, then solutions should be s -independent, so (1.8) reduces to an equation for maps $u : \mathbb{R} \rightarrow X$:

$$\frac{du}{dt} + \nabla H(u) = 0. \quad (1.9)$$

This is the equation considered in the Morse theory. In this sense, in symplectic geometry, Morse theory is often considered to be a finite-dimensional model for the theory of pseudoholomorphic curves.

There are quantum physics theories that are not conformally invariant. Such theories depend on one or more scale parameters. An example is the gauged σ -model. The corresponding PDE is the symplectic vortex equation, introduced by Mundet in his thesis [6] and by Cieliebak-Gaio-Salamon [7]. More precisely, suppose we have a symplectic manifold $(\tilde{X}, \tilde{\omega})$ and a Hamiltonian G -action, where G is a compact Lie group, with moment map $\mu : \tilde{X} \rightarrow \mathfrak{g}$. (This is the motivation for using μ to denote one of the two functions at the beginning of this paper.) Then the symplectic vortex equation is an elliptic system on a Riemann surface Σ associated with the triple $(\tilde{X}, \tilde{\omega}, \mu)$. This system depends on a scale parameter $\lambda > 0$, which determines the size of Σ .

One can use a G -invariant Hamiltonian function $\tilde{H} : \tilde{X} \rightarrow \mathbb{R}$ to perturb the symplectic vortex equation on the cylinder $\mathbb{R} \times S^1$ and study the corresponding Floer homology theory. This was proposed in [7] and has been studied by the second named author [8]. In this case, the equation is for maps $(u, \Psi) : \mathbb{R} \times S^1 \rightarrow \tilde{X} \times \mathfrak{g}$, and reads

$$\frac{\partial u}{\partial t} + J \left(\frac{\partial u}{\partial s} \right) + \nabla \left(\tilde{H} + \langle \mu, \Psi \rangle \right) = 0, \quad (1.10)$$

$$\frac{\partial \Psi}{\partial t} + \lambda^2 \mu(u) = 0. \quad (1.11)$$

Here $\langle \cdot, \cdot \rangle$ is the inner product on the Lie algebra \mathfrak{g} , and $\lambda > 0$ is the scale parameter.

As shown in [9], the adiabatic limit of the symplectic vortex equation as $\lambda \rightarrow \infty$ is nearly the same as the nonlinear Cauchy-Riemann equation that arises in the Gromov-Witten theory of the symplectic quotient $X := \mu^{-1}(0)/G$. In the present case, the $\lambda \rightarrow \infty$ limit of (1.10), (1.11) is related to (1.8) on the symplectic quotient X .

A natural question arises: what can be said about the opposite adiabatic limit $\lambda \rightarrow 0$? This limit could lead to a different perspective on the quantum cohomology of the symplectic quotient. Moreover, the $\lambda \rightarrow 0$ limit is related to the following argument of Witten in [10]. The complex Grassmannian $Gr(k, n)$ (the space of k -planes in \mathbb{C}^n) can be viewed as the symplectic quotient of the Euclidean space \mathbb{C}^{nk} with respect to a Hamiltonian $U(k)$ -action. Thus as above, we could use the symplectic vortex equation in \mathbb{C}^{nk} and its $\lambda \rightarrow \infty$ adiabatic limit to study the quantum cohomology of $Gr(k, n)$. Witten argued nonrigorously, using path integrals, that the opposite adiabatic limit leads to the Verlinde algebra, discovered by E. Verlinde in [11], which is therefore related to the quantum cohomology of $Gr(k, n)$. This argument explains earlier work by Gepner [12], Vafa [13], and Intriligator [14].

A mathematical proof of the isomorphism between the quantum cohomology and the Verlinde algebra was given by Agnihotri in his thesis [15] by directly calculating the two objects. However, geometers would prefer a proof closer to Witten's argument, using the symplectic vortex equation with varying λ and its adiabatic limits. Such a theory should apply not just to Grassmannians but to general symplectic quotients. The difficulty is that, unlike the $\lambda \rightarrow \infty$ limit, the $\lambda \rightarrow 0$ limit of the symplectic vortex equation is little understood. Only a few isolated results have been obtained, such as [16].

Because of this lack of understanding, it is natural to symplectic geometers to consider a Morse theory model. More precisely, if a solution (u, Ψ) to (1.10)-(1.11) is independent of s , then it satisfies the ODE

$$u'(t) + \nabla \left(\tilde{H} + \langle \mu, \Psi \rangle \right) = 0, \quad (1.12)$$

$$\Psi'(t) + \lambda^2 \mu(x) = 0. \quad (1.13)$$

We can further simplify the equation by removing the group action and replacing the moment map $\mu : \tilde{X} \rightarrow \mathfrak{g}$ by a smooth function $\mu : \tilde{X} \rightarrow \mathbb{R}$.² Then (1.12)-(1.13) becomes our equation (1.3)-(1.4), and the symplectic quotient becomes the hypersurface $\mu^{-1}(0)$.

It may be possible to generalize our result to other situations. For example, we can consider a vector-valued function $\mu : X \rightarrow \mathbb{R}^k$ instead of a scalar-valued one. A similar approach might also apply to Picard-Lefschetz theory, the holomorphic analogue of Morse theory. The case of equation (1.10)-(1.11) can

²The second named author would like to thank Urs Frauenfelder for suggesting removal of the group action.

perhaps be attacked in a similar way, presumably using an infinite-dimensional version of the exchange lemma and more-involved functional analytic techniques.

2. Morse homology and Morse-Smale-Witten complex of (\mathcal{F}, g_λ)

2.1. Morse homology

For the topological aspects of Morse theory, there is the classical book [17]. Here we will adopt the viewpoint of Witten [4], which is by now standard.

Let M be a smooth manifold without boundary. A smooth function $f : M \rightarrow \mathbb{R}$ is a *Morse function* if its differential df is a transverse section of the cotangent bundle T^*M . Equivalently, at each critical point p of f , the second derivative $D^2f(p)$ is nondegenerate. By the Morse lemma, near each $p \in \text{Crit}(f)$, there exists a local coordinate chart (x_1, \dots, x_n) such that

$$f(x) = f(p) - \sum_{i=1}^{k_p} x_i^2 + \sum_{i=k_p+1}^n x_i^2.$$

The integer k_p is called the *Morse index* of the critical point p . We write

$$\text{index}(p, f) = k_p \in \mathbb{Z}_+,$$

and we denote the set of critical points of f of index k by $\text{Crit}_k(f)$.

For any complete Riemannian metric g on M , the gradient vector field ∇f is the dual of the 1-form df . We denote by $\phi_t : M \rightarrow M$ the flow of $-\nabla f$. For each $p \in \text{Crit}(f)$, the unstable and stable manifolds are defined by

$$W^u(p) = \left\{ x \in M \mid \lim_{t \rightarrow -\infty} \phi_t(x) = p \right\}, \quad W^s(p) = \left\{ x \in M \mid \lim_{t \rightarrow +\infty} \phi_t(x) = p \right\}$$

and

$$\dim W^u(p) = \text{index}(p, f), \quad \dim W^s(p) = n - \text{index}(p, f).$$

The pair (f, g) is called *Morse-Smale* if for any two critical points p, q , $W^u(p)$ and $W^s(q)$ intersect transversally in M . This condition holds for generic (f, g) [18].

Associated to a Morse-Smale pair (f, g) is its *Morse-Smale-Witten chain complex* $\mathcal{C}(f, g)$ of \mathbb{Z}_2 -modules. The module of k -chains is generated by the critical points of index k . The boundary of a critical point of index k is a linear combination of critical points of index $k - 1$, where the coefficient of a critical point q is the number of orbits of the negative gradient flow from p to q . We write

$$\mathcal{C}(f, g) = (C_*, \partial), \quad C_k = \bigoplus_{p \in \text{Crit}_k(f)} \mathbb{Z}_2 \langle p \rangle, \quad (2.1)$$

where $\partial : C_k \rightarrow C_{k-1}$ is given by

$$\partial p = \sum_{q \in \text{Crit}_{k-1}(f)} n_{p,q} \cdot q, \quad n_{p,q} = \#[(W^u(p) \cap W^s(q))/\mathbb{R}] \in \mathbb{Z}_2. \quad (2.2)$$

One can show that $\#[(W^u(p) \cap W^s(q))/\mathbb{R}]$ is finite and $\partial \circ \partial = 0$ (under the Palais-Smale condition on f , see [18]), hence $\mathcal{C}(f, g)$ is a chain complex. Its homology is called the *Morse homology* of (f, g) with \mathbb{Z}_2 -coefficients, denoted $H(f, g; \mathbb{Z}_2)$. If M is compact, this homology is independent of the choice of Morse-Smale pair (f, g) and is isomorphic to the singular homology of M .

2.2. Morse homology in the analytic setting

The book [18] give a comprehensive treatment of the analytical perspective on finite-dimensional Morse theory. Here we give a brief review.

Let p and q be two critical points of f , and consider the nonlinear ODE

$$x' = -\nabla f(x), \quad (2.3)$$

with the boundary conditions

$$\lim_{t \rightarrow -\infty} x(t) = p, \quad \lim_{t \rightarrow +\infty} x(t) = q. \quad (2.4)$$

The space of solutions to this boundary value problem can be identified with $W^u(p) \cap W^s(q)$ by identifying the solution $x(t)$ with the point $x(0)$.

Indeed, the space of solutions to such a boundary value problem can be viewed as a finite-dimensional submanifold of an infinite-dimensional Banach manifold. We consider the Banach manifold \mathcal{B} of those $W_{loc}^{1,2}$ maps $x : \mathbb{R} \rightarrow M$ for which there exists $T > 0$ such that for $t \in [T, +\infty)$ (respectively $t \in (-\infty, -T]$), $d(x(t), q)$ (respectively $d(x(t), p)$) is less than the injectivity radius of (M, g) and $\exp_q^{-1}(x(t)) \in W^{1,2}([T, +\infty), T_q M)$ (respectively $\exp_p^{-1}(x(t)) \in W^{1,2}((-\infty, -T], T_p M)$). There is a Banach space bundle $\mathcal{E} \rightarrow \mathcal{B}$ whose fibre over $x : \mathbb{R} \rightarrow M$ is $L^2(\mathbb{R}, x^* TM)$. There is a smooth section

$$\begin{aligned} \mathcal{S} : \mathcal{B} &\rightarrow \mathcal{E} \\ x &\mapsto x' + \nabla f(x) \end{aligned}$$

whose zero locus is exactly $\widetilde{\mathcal{M}}(p, q) \simeq W^u(p) \cap W^s(q)$, the moduli space of solutions to (2.3)–(2.4). \mathcal{S} is a Fredholm section, which means that the linearization of \mathcal{S} at each $x \in \mathcal{S}^{-1}(0)$, given by

$$\begin{aligned} D\mathcal{S}_x : T_x \mathcal{B} &\rightarrow \mathcal{E}_x \\ V &\mapsto \nabla_t V + \nabla_V(\nabla f) \end{aligned}$$

is a Fredholm operator. The Fredholm index is $\text{index}(p) - \text{index}(q)$.

$\widetilde{\mathcal{M}}(p, q)$ is a smooth manifold if the linearization of \mathcal{S} is surjective along $\mathcal{S}^{-1}(0)$. In this case, we say that $\widetilde{\mathcal{M}}(p, q) = \mathcal{S}^{-1}(0)$ is *transverse*. This condition is equivalent to transversality of $W^u(p)$ and $W^s(q)$.

$\widetilde{\mathcal{M}}(p, q)$ has a free \mathbb{R} -action by time translation, so we can define $\mathcal{M}(p, q) = \widetilde{\mathcal{M}}(p, q)/\mathbb{R}$, the moduli space of orbits from p to q . A useful assumption is the Palais-Smale condition on f (which is automatic when M is compact): any sequence $x_i \in M$ for which $f(x_i)$ is uniformly bounded and $|\nabla f(x_i)| \rightarrow 0$ has a convergent subsequence. Under the transversality and Palais-Smale assumptions, if $\text{index}(q, f) = \text{index}(p, f) - 1$, then $\mathcal{M}(p, q)$ is finite. Moreover, if $\text{index}(q, f) = \text{index}(p, f) - 2$, then $\mathcal{M}(p, q)$ can be compactified to become a smooth one-dimensional manifold with boundary; the boundary points correspond to the broken orbits from p to q that pass through another critical point r with $\text{index}(r, f) = \text{index}(p, f) - 1$. These two facts imply that the boundary operator (2.2) is well-defined and $\partial \circ \partial = 0$.

A solution of (2.3)–(2.4) has *energy*

$$E(x) = \int_{\mathbb{R}} \|x'(t)\|^2 dt = - \int_{\mathbb{R}} \langle \nabla f(x(t)), x'(t) \rangle dt = f(p) - f(q), \quad (2.5)$$

which only depends on the values of f at p and q . The energy is essential to many useful estimates.

2.3. The Morse-Smale-Witten complex of (\mathcal{F}, g_λ)

We shall assume

(A1) M is a compact manifold with metric g .

We shall work with generic triples (f, μ, g) defined on M . For the sake of precision, we shall list our generic assumptions, and denote them (Ak) , $k \geq 2$. In each of our results, we shall assume without comment that all assumptions (Ak) made up to that point hold. The assumptions (Ak) are all independent

We assume:

(A2) f and μ are Morse functions on M .

(A3) 0 is a regular value of μ and $f|_{\mu^{-1}(0)}$ is Morse.

We shall regard the metric g and the Morse function μ having 0 as a regular value as arbitrary and fixed. Then all our assumptions (Ak) , $k \geq 2$, are true for generic f . Alternatively, we could fix the Morse function f ; the assumptions are then true for generic (μ, g) .

Lemma 1. [19, Proposition A.2] *The Lagrange multiplier \mathcal{F} is a Morse function on $M \times \mathbb{R}$ and for any $p = (x_p, \eta_p) \in \text{Crit}(\mathcal{F})$, $\text{index}(p, \mathcal{F}) = \text{index}(x_p, f|_{\mu^{-1}(0)}) + 1$.*

Now we consider the unstable and stable manifolds of critical points of \mathcal{F} . Since we are dealing with the family of metrics g_λ , we want transversality of unstable and stable manifolds not just for a single g_λ , but as much as possible for the family. Let ϕ_t^λ denote the flow of (1.6)–(1.7) for the given value of λ , let

$p_-, p_+ \in \text{Crit}(\mathcal{F})$, and let I be an interval in λ -space. Define the sets

$$\begin{aligned} W^u(p_-, \lambda) &= \left\{ p \in M \times \mathbb{R} \mid \lim_{t \rightarrow -\infty} \phi_t^\lambda(p) = p_- \right\}, \\ W^s(p_+, \lambda) &= \left\{ p \in (M \times \mathbb{R}) \mid \lim_{t \rightarrow +\infty} \phi_t^\lambda(p) = p_+ \right\}, \\ W^u(p_-, I) &= \{(p, \lambda) \in (M \times \mathbb{R}) \times I \mid p \in W^u(p_-, \lambda)\}, \\ W^s(p_+, I) &= \{(p, \lambda) \in (M \times \mathbb{R}) \times I \mid p \in W^s(p_+, \lambda)\}; \end{aligned}$$

$$\widetilde{\mathcal{M}}^I(p_-, p_+) = W^u(p_-, I) \cap W^s(p_+, I), \quad \mathcal{M}^I(p_-, p_+) = \widetilde{\mathcal{M}}^I(p_-, p_+)/\mathbb{R}.$$

We assume:

(A4) For all $p_-, p_+ \in \text{Crit}(\mathcal{F})$, $W^u(p_-, \mathbb{R}_+)$ and $W^s(p_+, \mathbb{R}_+)$ are transverse.

Assumption (A4) implies that for all but discretely many $\lambda \in \mathbb{R}_+$, $W^u(p_-, \lambda)$ and $W^s(p_+, \lambda)$ are transverse. We denote by $\Lambda^{reg} \subset \mathbb{R}_+$ the subset of λ 's for which $W^u(p_-, \lambda)$ and $W^s(p_+, \lambda)$ are transverse for all $p_-, p_+ \in \text{Crit}(\mathcal{F})$.

Lemma 2. *For each $\lambda \in \mathbb{R}_+$, the function \mathcal{F} satisfies the Palais-Smale condition with respect to the metric g_λ , i.e., for any sequence $\tilde{x}_i \in M \times \mathbb{R}$ such that $\mathcal{F}(\tilde{x}_i)$ is bounded and $|\nabla \mathcal{F}(\tilde{x}_i)|_{g_\lambda} \rightarrow 0$, there exists a convergent subsequence of \tilde{x}_i .*

Proof. Suppose $\tilde{x}_i = (x_i, \eta_i)$. Since M is compact, we may assume that $\eta_i \rightarrow +\infty$ or $-\infty$. The condition that $\mathcal{F}(\tilde{x}_i)$ is bounded implies that $\mu(x_i) \rightarrow 0$; but since 0 is a regular value of μ , $\nabla \mathcal{F}(\tilde{x}_i) = \nabla f(x_i) + \eta_i \nabla \mu(x_i)$ cannot have arbitrary small norm. \square

Now, for any $\lambda \in \Lambda^{reg}$, we can define the associated Morse-Smale-Witten complex of (\mathcal{F}, g_λ) , which is denoted by $\mathcal{C}^\lambda = \mathcal{C}(\mathcal{F}, g_\lambda)$. Its homology is denoted by H^λ . All \mathcal{C}^λ share the same generators and gradings, but the boundary operator ∂^λ may change when λ crosses a value in $\mathbb{R}_+ \setminus \Lambda^{reg}$.

In general, the Morse homology of a pair (f, g) is not independent of (f, g) if the underlying manifold is noncompact; see for example [20]. Despite the fact that $M \times \mathbb{R}$ is noncompact, we will show that H^λ is independent of λ for $\lambda \in \Lambda^{reg}$.

Lemma 3. *For any $L > 0$, there exists $K_L > 0$ such that for any $p_\pm \in \text{Crit}(\mathcal{F})$, $W^u(p_-, (0, L]) \cap W^s(p_+, (0, L]) \subset (M \times [-K_L, K_L]) \times [0, L]$.*

Proof. If this statement is false, then there exists $p_-, p_+ \in \text{Crit}(\mathcal{F})$, a sequence λ_i converging to $\lambda_\infty \in [0, L]$ and $(x_i, \eta_i) \in W^u(p_-, \lambda_i) \cap W^s(p_+, \lambda_i)$ such that $\lim_{i \rightarrow \infty} |\eta_i| = +\infty$. Let $\tilde{p}_i(t) = (\tilde{x}_i(t), \tilde{\eta}_i(t))$ be the solution of (1.6)–(1.7) for $\lambda = \lambda_i$ with $\tilde{p}_i(0) = (x_i, \eta_i)$. Then since the λ_i are bounded, we see that for any $R > 0$, $|\tilde{\eta}_i(t)| \rightarrow \infty$ uniformly on $-R \leq t \leq R$ as $i \rightarrow \infty$.

Now

$$\mathcal{F}(p_-) \geq \mathcal{F}(\tilde{p}_i(t)) = f(\tilde{x}_i(t)) + \tilde{\eta}_i(t)\mu(\tilde{x}_i(t)) \geq \mathcal{F}(p_+) \quad (2.6)$$

which implies that $\mu(\tilde{x}_i(t)) \rightarrow 0$ uniformly on $-R \leq t \leq R$. Hence, since 0 is a regular value of μ , $\|\nabla\mu(\tilde{x}_i(t))\|$ on $-R \leq t \leq R$ is bounded away from 0 for large i . Therefore, by the definition (2.5) of the energy of the solution \tilde{p}_i , we see

$$\begin{aligned} \mathcal{F}(p_+) - \mathcal{F}(p_-) &\geq \int_{-R}^R \|\tilde{p}'_i(t)\|^2 dt \geq \int_{-R}^R \|\tilde{x}'_i(t)\|^2 dt \\ &= \int_{-R}^R \|\nabla f(x_i(t)) + \eta_i(t)\nabla\mu(x_i(t))\|^2 dt \rightarrow \infty, \end{aligned}$$

which is impossible. \square

Proposition 4. *For any $\lambda_1, \lambda_2 \in \Lambda^{reg}$, there is a canonical isomorphism $\Phi_{\lambda_1, \lambda_2} : H^{\lambda_1} \rightarrow H^{\lambda_2}$ such that for $\lambda_1, \lambda_2, \lambda_3 \in \Lambda^{reg}$, $\Phi_{\lambda_2, \lambda_3} \circ \Phi_{\lambda_1, \lambda_2} = \Phi_{\lambda_1, \lambda_3}$.*

Proof. For any $\lambda_1, \lambda_2 \in \Lambda^{reg}$, $\lambda_1 < \lambda_2$, we can compare the two complexes \mathcal{C}^{λ_1} and \mathcal{C}^{λ_2} as in the case of compact manifolds, thanks to the compactness provided by the previous lemma. More precisely, we can either use the continuation principle [18], or bifurcation analysis as in [21], to show that the two chain complexes have isomorphic homology. Note that as we vary λ from λ_1 to λ_2 , the critical point set is fixed so there is no “birth-death” of critical points, but there may be a loss of transversality between unstable and stable manifolds at discrete values of λ . \square

3. Adiabatic limit $\lambda \rightarrow \infty$

We assume

(A5) $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$ is Morse-Smale.

Theorem 5. *For $\lambda \in \Lambda^{reg}$ sufficiently large, the complex \mathcal{C}^λ is canonically isomorphic to the Morse-Smale-Witten complex of the pair $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$ with grading shifted by one.*

Corollary 6. *For any $\lambda \in \Lambda^{reg}$, there is a canonical isomorphism*

$$\Phi_{\lambda, +\infty} : H_*^\lambda \rightarrow H_{*-1}(\mu^{-1}(0), \mathbb{Z}_2).$$

We give two proofs of Theorem 5, one using the infinite-dimensional implicit function theorem and one using geometric singular perturbation theory. The first is more likely to generalize. The second gives more geometric intuition. There are other ways to relate the two chain complexes. For example, in the appendix of [19] Frauenfelder had a different approach by deforming the function f in the normal direction of $\mu^{-1}(0)$.

3.1. Proof of Theorem 5 using the implicit function theorem

We first give a sketch of this proof. For fixed $p_{\pm} \in \text{Crit}\mathcal{F}$, we prove in Proposition 7 that, for λ sufficiently large, any orbit in $\mathcal{M}^\lambda(p_-, p_+)$ will be close to some orbit of the negative gradient flow of $f|_{\mu^{-1}(0)}$ from p_- to p_+ . Such orbits form a moduli space $\mathcal{N}^\infty(p_-, p_+)$. Next we show that, for large λ , there exists a homeomorphism $\Phi^\lambda : \mathcal{N}^\infty(p_-, p_+) \rightarrow \mathcal{M}^\lambda(p_-, p_+)$, which is constructed by using the infinite dimensional implicit function theorem. Proposition 7 is used to prove the surjectivity Φ^λ . Since the Morse homology is defined by counting orbits connecting two critical points with adjacent Morse indices, the homeomorphism (Theorem 10) means that the counting of $\mathcal{M}^\lambda(p_-, p_+)$ and that of $\mathcal{N}^\infty(p_-, p_+)$ are the same. Noting that the indices of p_{\pm} are dropped by 1 when regarded as critical points of $f|_{\mu^{-1}(0)}$, Theorem 5 follows immediately.

3.1.1. Convergence to orbits in the level set

For any $\tilde{p} = (\tilde{x}, \tilde{\eta}) \in \widetilde{\mathcal{M}}^\lambda(p_-, p_+)$, its energy, calculated by the metric g_λ on $M \times \mathbb{R}$, is

$$E = \mathcal{F}(p_-) - \mathcal{F}(p_+) = \|\tilde{x}'\|_{L^2}^2 + \lambda^2 \|\mu(\tilde{x})\|_{L^2}^2. \quad (3.1)$$

Here the L^2 -norms are still defined using the fixed metric g_1 . So

$$\|\mu(\tilde{x})\|_{L^2} \leq \frac{E^{\frac{1}{2}}}{\lambda}. \quad (3.2)$$

On the other hand, one has

$$\left\| \frac{d}{dt} \mu(\tilde{x}(t)) \right\|_{L^2} \leq \|d\mu\|_{L^\infty} \|\tilde{x}'(t)\|_{L^2}. \quad (3.3)$$

So by Sobolev embedding $W^{1,2} \rightarrow C^0$, for any $\epsilon > 0$, there exists $\Lambda_\epsilon > 0$ such that for $\lambda > \Lambda_\epsilon$ and $\tilde{x} \in \widetilde{\mathcal{M}}^\lambda(p_-, p_+)$, we have $\tilde{x}(t) \in U_\epsilon = \mu^{-1}((-\epsilon, \epsilon))$.

Now, consider the Banach manifold $\mathcal{B} = \mathcal{B}^{1,2}$ of $W_{loc}^{1,2}$ -maps $\tilde{p} = (\tilde{x}, \tilde{\eta})$ from $\mathbb{R} \rightarrow U_\epsilon \times \mathbb{R}$ such that \tilde{p} is asymptotic to $p_{\pm} = (x_{\pm}, \eta_{\pm}) \in \text{Crit}(\mathcal{F})$ at $\pm\infty$ in the following sense: there exists $R > 0$ and $\tilde{W}_- \in W^{1,2}((-\infty, -R], T_{p_-}M \oplus \mathbb{R})$ and $\tilde{W}_+ \in W^{1,2}([R, +\infty), T_{p_+}M \oplus \mathbb{R})$ such that $\tilde{p}|_{(-\infty, -R]} = \exp_{p_-} \tilde{W}_-$, $\tilde{p}|_{[R, +\infty)} = \exp_{p_+} \tilde{W}_+$. Consider $\mathcal{E} \rightarrow \mathcal{B}$ the Banach space bundle whose fibre over \tilde{p} is $\mathcal{E}_{\tilde{p}} = L^2(\tilde{x}^*TM \oplus \mathbb{R})$. For $\lambda > \Lambda_\epsilon$, consider the Fredholm section $\tilde{S}^\lambda : \mathcal{B} \rightarrow \mathcal{E}$ given by

$$\tilde{S}^\lambda(\tilde{p}) = \tilde{S}^\lambda(\tilde{x}, \tilde{\eta}) = \left(\frac{d\tilde{x}}{dt} + \nabla f + \tilde{\eta} \nabla \mu, \frac{d\tilde{\eta}}{dt} + \lambda^2 \mu(\tilde{x}) \right). \quad (3.4)$$

Then $\widetilde{\mathcal{M}}^\lambda(p_-, p_+) = \left(\tilde{S}^\lambda \right)^{-1}(0)$, where 0 is the zero section of $\mathcal{E} \rightarrow \mathcal{B}$. The linearization of \tilde{S}^λ at $\tilde{p} = (\tilde{x}, \tilde{\eta}) \in \mathcal{B}$ is given by

$$\begin{aligned} \tilde{\mathcal{D}}_{\tilde{p}} : \quad & \begin{pmatrix} T_{\tilde{p}}\mathcal{B} \\ V \\ H \end{pmatrix} \rightarrow \begin{pmatrix} \mathcal{E}_{\tilde{p}} \\ \nabla_t V + \nabla_V(\nabla f + \eta \nabla \mu) + H \nabla \mu \\ \frac{dH}{dt} + \lambda^2 d\mu(V) \end{pmatrix} \end{aligned} \quad (3.5)$$

On U_ϵ , there is the line bundle $L \subset TM$ generated by $\nabla\mu$, and denote by L^\perp the orthogonal complement with respect to the Riemannian metric g . Then, for any $\tilde{p} \in \mathcal{B}$, we can decompose the domain and target space of $\tilde{D}_{\tilde{p}}$ as

$$T_{\tilde{p}}\mathcal{B} \simeq W_L(\tilde{p}) \oplus W_T(\tilde{p}), \quad W_L(\tilde{p}) = W^{1,2}(\tilde{x}^*L \oplus \mathbb{R}), \quad W_T(\tilde{p}) = W^{1,2}(\tilde{x}^*L^\perp); \quad (3.6)$$

$$\mathcal{E}_{\tilde{p}} \simeq \mathcal{E}_L(\tilde{p}) \oplus \mathcal{E}_T(\tilde{p}), \quad \mathcal{E}_L(\tilde{p}) = L^2(\tilde{x}^*L \oplus \mathbb{R}), \quad \mathcal{E}_T(\tilde{p}) = L^2(\tilde{x}^*L^\perp). \quad (3.7)$$

We rescale the norms on $W_L(\tilde{p})$ and $\mathcal{E}_L(\tilde{p})$ as follows. We identify $(h_1\nabla\mu, h_2) \in W_L(\tilde{p})$ with $(h_1, h_2) \in W^{1,2} \oplus W^{1,2}$ and define

$$\|(h_1, h_2)\|_{W_\lambda} = \lambda \|h_1\|_{L^2} + \|h_1'\|_{L^2} + \|h_2\|_{L^2} + \lambda^{-1} \|h_2'\|_{L^2}; \quad (3.8)$$

and for $(h_1\nabla\mu, h_2) \in \mathcal{E}_L(\tilde{p})$, identify it with $(h_1, h_2) \in L^2 \oplus L^2$ and define

$$\|(h_1, h_2)\|_{L_\lambda} = \|h_1\|_{L^2} + \lambda^{-1} \|h_2\|_{L^2}. \quad (3.9)$$

We leave the norms on their complements unchanged, and use W_λ and L_λ to denote the norms on $T_{\tilde{p}}\mathcal{B}$ and $\mathcal{E}_{\tilde{p}}$ respectively.

Now we describe the limit objects. There exists a smooth function $\zeta : U_\epsilon \rightarrow \mathbb{R}$ defined by the condition

$$\langle \nabla\mu(x), \nabla f(x) + \zeta(x)\nabla\mu(x) \rangle = 0. \quad (3.10)$$

Then $\nabla f + \zeta\nabla\mu$ is a smooth vector field whose restriction to $\mu^{-1}(0)$ is the gradient of $\bar{f} = f|_{\mu^{-1}(0)}$ with respect to the restriction of the Riemannian metric. We denote by $\tilde{\mathcal{N}}^\infty(p_-, p_+)$ the space of solutions of the negative gradient flow of \bar{f} , whose elements are denoted by $y : \mathbb{R} \rightarrow \mu^{-1}(0)$; and by $\mathcal{N}^\infty(p_-, p_+)$ the quotient space by identifying reparametrizations, and by $\overline{\mathcal{N}}^\infty(p_-, p_+)$ the compactified moduli space by adding broken orbits.

Proposition 7. *Suppose $\lambda_\nu \rightarrow \infty$ and $\tilde{p}_\nu = (\tilde{x}_\nu, \tilde{\eta}_\nu) \in \tilde{\mathcal{M}}^{\lambda_\nu}(p_-, p_+)$. Then there is a subsequence, still indexed by ν , and a broken orbit $\mathcal{Y} = ([y_i])_{i=1}^n \in \overline{\mathcal{N}}^\infty(p_-, p_+)$ (where $[y_i]$ is the orbit of y_i) and a constant $c_0 > 0$ such that*

1. $\|(\mu(\tilde{x}_\nu), \tilde{\eta}_\nu - \zeta(\tilde{x}_\nu))\|_{W_{\lambda_\nu}} \leq c_0 \lambda_\nu^{-1}$;
2. *There exists $t_{1,\nu}, t_{2,\nu}, \dots, t_{n,\nu} \in \mathbb{R}$ such that $\tilde{x}_\nu(t_{i,\nu} + \cdot)$ converges to y_i in C_{loc}^1 -topology;*

Proof. Apply $d\mu$ to the equation $\tilde{x}_\nu'(t) + \nabla f(\tilde{x}_\nu(t)) + \tilde{\eta}_\nu(t)\nabla\mu(\tilde{x}_\nu(t)) = 0$, we obtain

$$\begin{aligned} 0 &= \frac{d}{dt} \mu(\tilde{x}_\nu(t)) + \langle \nabla\mu(\tilde{x}_\nu(t)), \nabla f(\tilde{x}_\nu(t)) + \tilde{\eta}_\nu(t)\nabla\mu(\tilde{x}_\nu(t)) \rangle \\ &= \frac{d}{dt} \mu(\tilde{x}_\nu(t)) + \lambda_\nu |\nabla\mu|^2 \left(\frac{1}{\lambda_\nu} (\tilde{\eta}_\nu(t) - \zeta(\tilde{x}_\nu(t))) \right). \end{aligned} \quad (3.11)$$

Also we have

$$\frac{d}{dt} \left(\frac{1}{\lambda_\nu} (\tilde{\eta}_\nu - \zeta(\tilde{x}_\nu)) \right) + \lambda \mu(\tilde{x}_\nu) = -\frac{1}{\lambda_\nu} \frac{d}{dt} \zeta(\tilde{x}_\nu). \quad (3.12)$$

Consider the linear operator

$$\begin{aligned} D : W^{1,2}(\mathbb{R}, \mathbb{R}^2) &\rightarrow L^2(\mathbb{R}, \mathbb{R}^2) \\ (f_1, f_2) &\mapsto \left(\frac{d}{dt} f_1 + \lambda |\nabla \mu|^2 f_2, \frac{d}{dt} f_2 + \lambda f_1 \right). \end{aligned} \quad (3.13)$$

If regarded as an unbounded operator from L^2 to L^2 , then it is bounded from below by $c\lambda$ for some constant c . Then (3.11) and (3.12) imply that

$$\|\mu(\tilde{x}_\nu)\|_{L^2} + \lambda_\nu^{-1} \|\tilde{\eta}_\nu - \zeta(\tilde{x}_\nu)\|_{L^2} \leq \frac{c}{\lambda_\nu^2} \left\| \frac{d}{dt} \zeta(\tilde{x}_\nu) \right\|_{L^2} \leq c\lambda_\nu^{-2}. \quad (3.14)$$

Also, D has a uniformly bounded right inverse, so

$$\|\mu(\tilde{x}_\nu)\|_{W^{1,2}} + \lambda_\nu^{-1} \|\tilde{\eta}_\nu - \zeta(\tilde{x}_\nu)\|_{W^{1,2}} \leq c\lambda_\nu^{-1}. \quad (3.15)$$

This implies the first claim of this proposition. Moreover, the Sobolev embedding $W^{1,2} \rightarrow C^0$ implies in particular that $\tilde{\eta}_\nu$ is uniformly bounded since $\zeta(\tilde{x}_\nu)$ is. Then by the differential equation on \tilde{x}_ν , we have that $|\tilde{x}'_\nu|$ is uniformly bounded.

Now we identify $U_\epsilon \simeq \mu^{-1}(0) \times (-\epsilon, \epsilon)$ such that the projection to the second component is equal to μ . Then we can write

$$\tilde{x}_\nu(t) = (\bar{x}_\nu(t), \mu(\tilde{x}_\nu(t))). \quad (3.16)$$

Projecting to the first factor and using the fact that $|\tilde{x}'_\nu|$ is uniformly bounded, we see that there exists $K > 0$ independent of ν , such that

$$|\bar{x}'_\nu(t) + \nabla \bar{f}(\bar{x}_\nu(t))| \leq K |\mu(\tilde{x}_\nu(t))|. \quad (3.17)$$

This implies that a subsequence of \bar{x}_ν converges to a broken orbit for the induced function \bar{f} in C_{loc}^0 -topology. More precisely, there exists $t_{1,\nu}, \dots, t_{k,\nu} \in \mathbb{R}$ such that $\bar{x}_\nu(t_{i,\nu} + \cdot)$ converges to an orbit y_i in $\mu^{-1}(0)$ in C_{loc}^0 -topology.

Then, by the Sobolev embedding $W^{1,2} \rightarrow C^0$ and the C_{loc}^0 convergence of $\tilde{x}_\nu(t_{i,\nu} + \cdot)$ to y_i , we see that $\tilde{\eta}_\nu(t_{i,\nu} + \cdot)$ converges to $\zeta(y_i)$ in C_{loc}^0 . This implies that $\tilde{x}_\nu(t_{i,\nu} + \cdot)$ converges to y_i in C_{loc}^1 . \square

3.1.2. Applying implicit function theorem and the isomorphism of chain complexes

Now we want to prove that any $y \in \tilde{\mathcal{N}}^\infty(p_-, p_+)$ can be approximated by $\tilde{p} \in \tilde{\mathcal{M}}^\lambda(p_-, p_+)$, so that in particular, when $\text{index}(p_-, \mathcal{F}) - \text{index}(p_+, \mathcal{F}) = 1$, there is a canonical one-to-one correspondence between $\mathcal{N}^\infty(p_-, p_+)$ and $\mathcal{M}^\lambda(p_-, p_+)$.

Since we have assumed that the restriction of (f, g) to $\mu^{-1}(0)$ is Morse-Smale, the trajectory y is transverse. Namely, the linearized operator

$$\begin{aligned} \overline{D}_y : \quad W^{1,2}(y^*T\mu^{-1}(0)) &\rightarrow L^2(y^*T\mu^{-1}(0)) \\ V &\mapsto \nabla_{y'}V + \nabla_V(\nabla f + \zeta\nabla\mu) \end{aligned} \quad (3.18)$$

is surjective and has a bounded right inverse \overline{Q}_y . In particular, we can choose \overline{Q}_y such that

$$\text{Im } \overline{Q}_y = \{V \in W^{1,2}(y^*T\mu^{-1}(0)) \mid g(V(0), W(0)) = 0 \ \forall W \in \ker \overline{D}_y\}. \quad (3.19)$$

For any large λ , we take our approximate solution just to be

$$\tilde{y}(t) = (y(t), \zeta(y(t))). \quad (3.20)$$

Note that $y(t)$ converges to p_{\pm} exponentially as $t \rightarrow \pm\infty$, so $\tilde{y} \in \mathcal{B}$. Then denote the linearization of \tilde{S}^{λ} at \tilde{y} by

$$\begin{aligned} \tilde{D}_{\tilde{y}} : \quad W^{1,2}(y^*TM \oplus \mathbb{R}) &\rightarrow L^2(y^*TM \oplus \mathbb{R}) \\ \begin{pmatrix} V \\ h \end{pmatrix} &\mapsto \begin{pmatrix} \nabla_{y'}V + \nabla_V(\nabla f + \zeta\nabla\mu) + h\nabla\mu \\ h' + \lambda^2 d\mu(V) \end{pmatrix}. \end{aligned} \quad (3.21)$$

Now we construct a right inverse to $\tilde{D}_{\tilde{y}}$ out of \overline{Q}_y . Note that, for any $\overline{V} \in W_T(\tilde{y})$

$$\tilde{D}_{\tilde{y}}(\overline{V}) = (\nabla_t \overline{V} + \nabla_{\overline{V}}(\nabla f + \zeta\nabla\mu), \lambda^2 d\mu(\overline{V})) = (\overline{D}_y(\overline{V}), 0) \in \mathcal{E}_T(\tilde{y}). \quad (3.22)$$

Hence with respect to the decomposition (3.6) and (3.7), the linearized operator can be written as

$$\tilde{D}_{\tilde{y}} = \begin{pmatrix} \overline{D}_y & A_y \\ 0 & D'_{\tilde{y}} \end{pmatrix}. \quad (3.23)$$

Here A_y is given by

$$A_y(h_1 \nabla \mu, h_2) = h_1 W \quad (3.24)$$

where W is a smooth tangent vector field on $\mu^{-1}(0)$. Hence there exists $c_1 > 0$ such that

$$\|A_y(h_1 \nabla \mu, h_2)\|_{L_{\lambda}} \leq c_1 \|h_1\|_{L^2} \leq c_1 \lambda^{-1} \|h_1\|_{W_{\lambda}}. \quad (3.25)$$

Now we look at the operator $D'_{\tilde{y}}$ in (3.23). After trivialize the bundle $\mathbb{R}\{\nabla\mu\} \oplus \mathbb{R}$ isometrically to \mathbb{R}^2 , we see the operator $D'_{\tilde{y}}$ is transformed into

$$\mathfrak{D}'_{\tilde{y}} = \frac{d}{dt} + \begin{pmatrix} a(t) & \lambda b(t) \\ \lambda b(t) & 0 \end{pmatrix} := \frac{d}{dt} + B(t). \quad (3.26)$$

It is easy to see that, for λ large, $B(t)$ has 1 positive eigenvalue and one negative eigenvalue, both of which are bounded away from zero by $c\lambda$, where c is a constant independent of λ . Hence we have a (unique) bounded right inverse

$$Q'_y : \mathcal{E}_L(\tilde{y}) \rightarrow W_L(\tilde{y}), \quad \|Q'_y\| \leq c_2 \quad (3.27)$$

for some $c_2 > 0$. Then define

$$Q_{\tilde{y}} = \begin{pmatrix} \overline{Q}_y & 0 \\ 0 & Q'_y \end{pmatrix} \quad (3.28)$$

which serves as the approximate right inverse.

Lemma 8. *There exists $\Lambda_0 > 0$, such that for all $\lambda > \Lambda_0$,*

$$\left\| \text{Id} - \tilde{D}_{\tilde{y}} Q_{\tilde{y}} \right\|_{L_\lambda} < \frac{1}{2} \quad (3.29)$$

with respect to the operator norm of the space L_λ .

Proof. Note that $\text{Id} - \tilde{D}_{\tilde{y}} Q_{\tilde{y}} = A_y \circ Q'_y$ and

$$\left\| A_{\tilde{y}} Q'_y(h_1 \nabla \mu, h_2) \right\|_{L_\lambda} \leq \frac{c_1}{\lambda} \left\| Q'_y(h_1 \nabla \mu, h_2) \right\|_{W_\lambda} \leq \frac{c_1 c_2}{\lambda} \left\| (h_1 \nabla \mu, h_2) \right\|_{L_\lambda}. \quad (3.30)$$

□

Hence $\tilde{D}_{\tilde{y}} Q_{\tilde{y}}$ is invertible and a right inverse can be constructed as

$$\tilde{Q}_{\tilde{y}} = Q_{\tilde{y}} \left(\tilde{D}_{\tilde{y}} Q_{\tilde{y}} \right)^{-1} : L_\lambda \rightarrow W_\lambda. \quad (3.31)$$

It is uniformly bounded by some constant which is independent of λ .

Now since we want to apply the implicit function theorem (cf. [22, Appendix A]), we need to trivialize the Banach manifold \mathcal{B} and the Banach space bundle \mathcal{E} locally near \tilde{y} . We identify a neighborhood of \tilde{y} in \mathcal{B} with a small ball (with respect to the W_λ -norm) centered at the origin of $T_{\tilde{y}}\mathcal{B}$ as follows: using the identification $U_\epsilon \simeq \mu^{-1}(0) \times (-\epsilon, \epsilon)$, for any $\tilde{W} = (W, a \nabla \mu, h) \in T_{\tilde{y}}\mathcal{B}$, we identify it with the map $\rho(\tilde{W}) = (\overline{\text{exp}}_{\tilde{y}} W, a, \zeta(y) + h)$ into $\mu^{-1}(0) \times (-\epsilon, \epsilon) \times \mathbb{R}$ where $\overline{\text{exp}}$ is the exponential map inside $\mu^{-1}(0)$. We trivialize \mathcal{E} over such a neighborhood around \tilde{y} by parallel transport along the radial geodesics, which is denoted by $\Phi_{\tilde{W}} : \mathcal{E}_{\rho(\tilde{W})} \rightarrow \mathcal{E}_{\tilde{y}}$. Then for $\tilde{W} \in T_{\tilde{y}}\mathcal{B}$ with $\left\| \tilde{W} \right\|_{W_\lambda}$ small, denote by $\tilde{D}_{\tilde{W}} : T_{\tilde{y}}\mathcal{B} \rightarrow \mathcal{E}_{\tilde{y}}$ the linearization of $\Phi_{\tilde{W}} \circ \tilde{S}^\lambda(\rho(\tilde{W}))$ at \tilde{W} . We have

Lemma 9. *There exists $\epsilon_3, c_3 > 0$, independent of λ , such that for all $\tilde{W} \in T_{\tilde{y}}\mathcal{B}$ with $\left\| \tilde{W} \right\|_{W_\lambda} < \epsilon_3$, one has*

$$\left\| \tilde{D}_{\tilde{W}} - \tilde{D}_{\tilde{y}} \right\| \leq c_3 \left\| \tilde{W} \right\|_{W_\lambda}. \quad (3.32)$$

Proof. First consider the case when the \mathbb{R} -component of \widetilde{W} is zero. We denote by \widetilde{S}_1 the first component of \widetilde{S}^λ , which is independent of λ . Then we see

$$\left\| D\widetilde{S}_1(\widetilde{W}) - D\widetilde{S}_1(\widetilde{y}) \right\| \leq c \left\| \widetilde{W} \right\|_{L^\infty} \quad (3.33)$$

which is standard. Now because the \mathbb{R} -component of \widetilde{W} is zero, we have $\left\| \widetilde{W} \right\|_{L^\infty} \leq c \left\| \widetilde{W} \right\|_{W_\lambda}$ by our definition of the W_λ -norm (3.8).

The second component of \widetilde{S}^λ is denoted by \widetilde{S}_2^λ . Then we see in this case, for any $\widetilde{V} = (V, h) \in T_{\widetilde{y}}\mathcal{B}$

$$D\widetilde{S}_2^\lambda(\widetilde{W})(\widetilde{V}) = \lambda^2 h. \quad (3.34)$$

Hence the variation of the derivative of \widetilde{S}_2^λ is always zero.

Now we consider the case $\widetilde{W} = (0, h_0)$ with $h_0 \in W^{1,2}(\mathbb{R})$. Then for any $\widetilde{W}_0 \in T_{\widetilde{y}}\mathcal{B}$ small, by our definition of our norms (3.8) and (3.9),

$$\begin{aligned} \left\| (\widetilde{D}_{\widetilde{W}+\widetilde{W}_0} - \widetilde{D}_{\widetilde{W}_0})(V, h) \right\|_{L_\lambda} &= \|h_0 \nabla_V \nabla \mu\|_{L_\lambda} = \|h_0 \nabla_V \nabla \mu\|_{L^2} \\ &\leq c \left\| \widetilde{W} \right\|_{W_\lambda} \|V\|_{L^2} \leq c \left\| \widetilde{W} \right\|_{W_\lambda} \|(V, h)\|_{W_\lambda} \end{aligned} \quad (3.35)$$

for some $c > 0$.

Combining (3.33), (3.34) and (3.35) we obtain (3.32). \square

On the other hand, we see

$$\left\| \widetilde{S}^\lambda(\widetilde{y}) \right\|_{L_\lambda} = \lambda^{-1} \left\| \frac{d}{dt} \zeta(y) \right\|_{L^2} \leq c_4 \lambda^{-1} \quad (3.36)$$

for some constant $c_4 > 0$. By the implicit function theorem, there exists $\epsilon_4 > 0$ such that for sufficiently large λ , there exists a unique $\widetilde{W}_\lambda(y) \in T_{\widetilde{y}}\mathcal{B}$ satisfying

$$\left\| \widetilde{W}_\lambda(y) \right\|_{W_\lambda} \leq \epsilon_4, \quad \widetilde{W}_\lambda(y) \in \text{Im } Q_{\widetilde{y}}, \quad \widetilde{S}^\lambda(\exp_{\widetilde{y}} \widetilde{W}_\lambda(y)) = 0. \quad (3.37)$$

Moreover, there exists $c_5 > 0$ such that

$$\left\| \widetilde{W}_\lambda(y) \right\|_{W_\lambda} \leq \frac{c_5}{\lambda}. \quad (3.38)$$

We hence define the *gluing map* (for large Λ_0) to be

$$\begin{aligned} \Phi : \quad [\Lambda_0, +\infty) \times \widetilde{\mathcal{N}}^\infty(p_-, p_+) &\rightarrow \cup_{\lambda \geq \Lambda_0} \widetilde{\mathcal{M}}^\lambda(p_-, p_+) \\ (\lambda, y) &\mapsto \exp_{\widetilde{y}} \widetilde{W}_\lambda(y). \end{aligned} \quad (3.39)$$

Theorem 10. *For each $\lambda \in [\Lambda_0, +\infty) \cap \Lambda^{reg}$, the restriction of Φ to $\{\lambda\} \times \widetilde{\mathcal{N}}^\infty(p_-, p_+)$ is a homeomorphism onto $\widetilde{\mathcal{M}}^\lambda(p_-, p_+)$.*

Proof. To complete the proof of this theorem, it remains to prove the local surjectivity of the gluing map Φ . For simplicity, we only prove this for the case in which the Morse indices of p_- and p_+ differ by one, which is enough for our purpose. If, in this case, the gluing map is not locally surjective, then there exists a sequence $\lambda_i \rightarrow \infty$ and $\tilde{p}_i \in \widetilde{\mathcal{M}}^{\lambda_i}(p_-, p_+)$ which converges to some $\tilde{y} \in \widetilde{\mathcal{N}}^\infty(p_-, p_+)$ as described by Proposition 7 (with $n = 1$ and $t_{1,\nu} = 0$) but don't lie in the range of Φ . Consider the hyperplanes $N(y(s)) = (y'(s))^\perp \subset T_{y(s)}M$ for $s \in \mathbb{R}$. For each i , there exists $s_i \in \mathbb{R}$ such that $s_i \rightarrow 0$ and

$$\tilde{p}_i(0) \in \left\{ (\exp_{y(s_i)} V, \eta) \mid \eta \in \mathbb{R}, V \in N(y(s_i)), \|V\| \leq \epsilon \right\}. \quad (3.40)$$

Then if we write $\tilde{p}_i(t) = \exp_{\tilde{y}(s_i+t)} \widetilde{W}_i(t)$, then $\widetilde{W}_i \in \text{Im } Q_{\tilde{y}(s_i+\cdot)}$ and $\|\widetilde{W}_i\|_{W_{\lambda_i}} \leq c_0 \lambda_i^{-1}$ by Proposition 7. By the uniqueness part of the implicit function theorem, this implies that $\tilde{p}_i = \Phi(\lambda_i, \tilde{y}(s_i + \cdot))$, which contradicts our assumption. \square

Now Theorem 5 is an immediate consequence.

3.2. Proof of Theorem 5 using geometric singular perturbation theory

In the system (1.6)–(1.7), in which the time variable is t , we make the rescalings

$$\eta = \frac{\rho}{\epsilon}, \quad \lambda = \frac{1}{\epsilon}, \quad t = \epsilon\tau.$$

The system becomes

$$\frac{dx}{d\tau} = -(\epsilon \nabla f(x) + \rho \nabla \mu(x)), \quad (3.41)$$

$$\frac{d\rho}{d\tau} = -\mu(x). \quad (3.42)$$

Studying the limit $\lambda \rightarrow \infty$ in (1.6)–(1.7) is equivalent to studying the limit $\epsilon \rightarrow 0$ in (3.41)–(3.42). Setting $\epsilon = 0$ in (3.41)–(3.42), we obtain

$$\frac{dx}{d\tau} = -\rho \nabla \mu(x), \quad (3.43)$$

$$\frac{d\rho}{d\tau} = -\mu(x). \quad (3.44)$$

Let N denote the submanifold of M defined by $\mu = 0$. Then the set of equilibria of (3.43)–(3.44) is the set $\mu = \rho = 0$, i.e., $N_0 = N \times \{0\} \subset M \times \mathbb{R}$. N_0 is a compact codimension-two submanifold of $M \times \mathbb{R}$.

Choosing coordinates on M and linearizing (3.43)–(3.44) at an equilibrium $(x, 0)$, we obtain

$$\begin{pmatrix} \dot{v} \\ \dot{\rho} \end{pmatrix} = \begin{pmatrix} 0 & -\nabla \mu(x) \\ -d\mu(x) & 0 \end{pmatrix} \begin{pmatrix} v \\ \rho \end{pmatrix}. \quad (3.45)$$

Since 0 is a regular value of μ by (A3), $d\mu(x)$ and $\nabla \mu(x)$ are nonzero row and column vectors respectively. Therefore the matrix in (3.45) has rank 2, so it has

at least $n - 2$ zero eigenvalues. (Of course this is a consequence of the fact that we linearized (3.43)–(3.44) at a point on a manifold of equilibria of dimension $n - 2$.) The other eigenvalues are $\pm \|\nabla \mu(x)\|$, where $\|v\| = \|v\|_{g(x)}$; this is shown by the following calculations, which use the fact

$$d\mu(x)v = \langle \nabla \mu(x), v \rangle_{g(x)}.$$

$$\begin{aligned} \begin{pmatrix} 0 & -\nabla \mu(x) \\ -d\mu(x) & 0 \end{pmatrix} \begin{pmatrix} \nabla \mu(x) \\ -\|\nabla \mu(x)\| \end{pmatrix} &= \begin{pmatrix} \|\nabla \mu(x)\| \nabla \mu(x) \\ -\|\nabla \mu(x)\|^2 \end{pmatrix}, \\ \begin{pmatrix} 0 & -\nabla \mu(x) \\ -d\mu(x) & 0 \end{pmatrix} \begin{pmatrix} \nabla \mu(x) \\ \|\nabla \mu(x)\| \end{pmatrix} &= \begin{pmatrix} -\|\nabla \mu(x)\| \nabla \mu(x) \\ -\|\nabla \mu(x)\|^2 \end{pmatrix}. \end{aligned}$$

Therefore $N \times \{0\}$ is a compact normally hyperbolic manifold of equilibria for (3.41)–(3.42) with $\epsilon = 0$ [1]. It follows that for small $\epsilon > 0$, (3.41)–(3.42) has a normally hyperbolic invariant manifold N_ϵ near N_0 .

Locally we may assume the coordinates on M are chosen so that $\mu = x_n$. Let $y = (x_1, \dots, x_{n-1})$, so $x = (y, x_n)$. Then locally N_ϵ is parameterized by y and is given by

$$x_n = x_n(y, \epsilon) = x_n(x_1, \dots, x_{n-1}, \epsilon), \quad \rho = \rho(y, \epsilon) = \rho(x_1, \dots, x_{n-1}, \epsilon),$$

with $x_n(y, 0) = \rho(y, 0) = 0$. After division by ϵ , the system (3.41)–(3.42) restricted to N_ϵ is given by

$$\dot{y} = -\nabla_y f(y, 0) + \mathcal{O}(\epsilon),$$

where $\nabla_y f(y, x_n)$ denotes the first $n - 1$ components of $\nabla f(y, x_n)$.

It follows that the system (3.41)–(3.42) restricted to N_ϵ is a perturbation of the negative gradient flow of $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$. Since $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$ is Morse-Smale by (A5), its negative gradient flow is structurally stable. Therefore, for small $\epsilon > 0$, the flow of (3.41)–(3.42) restricted to N_ϵ is topologically equivalent to the negative gradient flow of $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$.

An equilibrium x of the negative gradient flow of $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$ corresponds, for each $\lambda > 0$, to an equilibrium (x, η) of (1.6)–(1.7), which turn corresponds to the equilibrium $(x, \epsilon\eta)$ of (3.41)–(3.42); the latter lies in N_ϵ , which contains all complete orbits nearby. It has one higher index than that of x for the negative gradient flow of $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$. (The reason is that one of the transverse eigenvalues computed above is positive.) If the only connections between these equilibria are those in N_ϵ , the resulting Morse-Smale-Witten chain complex is the same as that of $(f|_{\mu^{-1}(0)}, g|_{\mu^{-1}(0)})$ with degree shifted by one.

To rule out the existence of other connections, note that for $\epsilon > 0$, the energy $E_\epsilon(x, \rho) = \epsilon f(x) + \rho \mu(x)$ decreases along solutions of (3.41)–(3.42). For small $\epsilon > 0$, the energy difference between two equilibria $(x_-, \epsilon\eta_-)$ and $(x_+, \epsilon\eta_+)$ of (3.41)–(3.42) is of order ϵ .

On the other hand, by the normal hyperbolicity of N_0 , any sufficiently small neighborhood V of N_0 has the property that for (3.41)–(3.42) with ϵ small, a

solution of (3.41)–(3.42) that starts in $V \setminus N_\epsilon$ must leave V in forward or backward time. Therefore a solution of (3.41)–(3.42) that connects two equilibria but does not lie in N_ϵ must at some time leave V through its boundary. If we can show that it must do so at a point where E_ϵ is of order one, we have a contradiction.

For a small $\alpha > 0$, we can take $V = \{(x, \rho) : |\mu(x)| < \alpha, |\rho| < \alpha, |\rho\mu(x)| < \frac{\alpha^2}{4}\}$; see Figure 1. At points on the boundary where $|\rho\mu(x)| = \frac{\alpha^2}{4}$, $|E_\epsilon|$ is close to $\frac{\alpha^2}{4}$ for small ϵ . Thus if we can show that a solution that connects two equilibria must leave V through such a point, we are done.

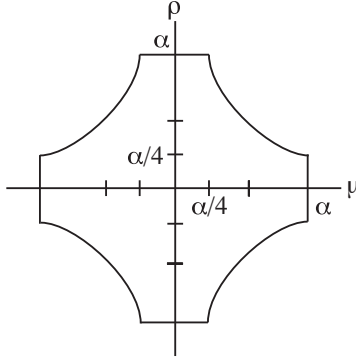


Figure 1: The set V .

This is not true, but by making a small ϵ -dependent alteration in V , we can make it true: we replace the portions of ∂V on which $\mu = \pm\alpha$ or $\rho = \pm\alpha$ by nearby invariant surfaces, so that solutions cannot cross them. More precisely, we replace the portion of the boundary on which $\mu = \pm\alpha$ by a union of integral curves of (3.41)–(3.42) that start on the codimension-two surfaces $\mu = \pm\alpha$, $\rho = 0$, and we replace the portion of the boundary on which $\rho = \pm\alpha$ by a union of integral curves of (3.41)–(3.42) that start on the codimension-two surfaces $\mu = 0$, $\rho = \pm\alpha$. Details are left to the reader.

4. Fast-slow system associated with the $\lambda \rightarrow 0$ limit

Now we consider the limit of (1.6)–(1.7) as $\lambda \rightarrow 0$. In this limit, (1.6)–(1.7) is a fast-slow system [1]. There is one slow variable, η . In this section, we will identify the slow manifold, and study the fast flow and slow equation, under appropriate generic assumptions. We will also relate various Morse indices.

4.1. The slow manifold

The *slow manifold* for (1.6)–(1.7) is the set of equilibria for $\lambda = 0$:

$$\mathcal{C}_{\mathcal{F}} = \{(x, \eta) \in M \times \mathbb{R} \mid \nabla f(x) + \eta \nabla \mu(x) = 0\}.$$

We assume:

(A6) $\text{Crit}(f) \cap \text{Crit}(\mu)$ is empty.

(A7) $\nabla f + \eta \nabla \mu$ is a transverse section of $\pi^* TM$, where $\pi : M \times \mathbb{R} \rightarrow M$ is the projection.

(A7) implies that $\mathcal{C}_{\mathcal{F}}$ is a 1-dimensional smooth submanifold of $M \times \mathbb{R}$. Indeed, it is noncompact and its end at infinity is asymptotic to $(\text{Crit}\mu \times [R, +\infty)) \cup (\text{Crit}\mu \times (-\infty, -R])$ for R large. This can be seen as follows. If there is a sequence $(x_i, \eta_i) \in \mathcal{C}_{\mathcal{F}}$ with $\lim_{i \rightarrow \infty} |\eta_i| = \infty$, then we see that $\nabla \mu(x_i) \rightarrow 0$ and a subsequence of x_i converging to some $y \in \text{Crit}\mu$. Conversely, for any $y \in \text{Crit}\mu$, let $S_\epsilon(y) \subset M$ denote the sphere of radius ϵ around y . Since by (A2) y is a nondegenerate critical point of μ , we see that for small $\epsilon > 0$, the map

$$S_\epsilon(y) \rightarrow S(T_y M), \quad x \mapsto \frac{\nabla \mu(x)}{\|\nabla \mu(x)\|} \quad (4.1)$$

has degree ± 1 . On the other hand, by (A6) we can use local coordinates near y in which the vector field ∇f is constant. Hence there exist a unique pair of points $(x_\epsilon^\pm, \eta_\epsilon^\pm)$, with $x_\epsilon^\pm \in S_\epsilon(y)$, such that

$$(x_\epsilon^\pm, \eta_\epsilon^\pm) \in \mathcal{C}_{\mathcal{F}}. \quad (4.2)$$

We can order the two points so that $\lim_{\epsilon \rightarrow 0} \eta_\epsilon^\pm = \pm \infty$.

For fixed η , denote the function $f_\eta = f + \eta \mu : M \rightarrow \mathbb{R}$. Then consider the function

$$\begin{aligned} d_{\mathcal{C}} : \quad \mathcal{C}_{\mathcal{F}} &\rightarrow \mathbb{R} \\ p = (x, \eta) &\mapsto \det \text{Hess} f_\eta(x). \end{aligned} \quad (4.3)$$

We assume

(A8) 0 is a regular value of $d_{\mathcal{C}}$.

(A9) For any $p = (x, \eta) \in d_{\mathcal{C}}^{-1}(0)$, $\mu(x) \neq 0$. Therefore $d_{\mathcal{C}}^{-1}(0) \cap \text{Crit}(\mathcal{F}) = \emptyset$.

We denote $\mathcal{C}_{\mathcal{F}}^{\text{sing}} = d_{\mathcal{C}}^{-1}(0)$. From (A8), as $(x, \eta) \in \mathcal{C}_{\mathcal{F}}$ crosses a point in $\mathcal{C}_{\mathcal{F}}^{\text{sing}}$, one and only one of the eigenvalues of $\text{Hess} f_\eta(x)$ changes its sign.

By the implicit function theorem, $\mathcal{C}_{\mathcal{F}}$ can be smoothly parameterized by η near any point of $\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{\text{sing}}$ as $x = x(\eta)$. On the other hand, near $\mathcal{C}_{\mathcal{F}}^{\text{sing}}$ we have the following result.

Proposition 11. *Let $p = (x_p, \eta_p) \in \mathcal{C}_{\mathcal{F}}^{\text{sing}}$. We can choose local coordinates (x_1, \dots, x_n) on M such that x_p corresponds to $(0, \dots, 0)$, and near p , $\mathcal{C}_{\mathcal{F}}$ is parameterized by x_n . Moreover, for some $c \neq 0$,*

$$\eta = \eta_p + cx_n^2 + O(x_n^3). \quad (4.4)$$

Thus $\eta|_{\mathcal{C}_{\mathcal{F}}}$ has a nondegenerate critical point at p .

This is a standard result, but we shall give a proof in Subsection 4.3.

4.2. The fast flow

The fast flow $\Phi_t : M \times \mathbb{R} \rightarrow M \times \mathbb{R}$ of (1.6)–(1.7) is the flow on $M \times \mathbb{R}$ determined by (1.6)–(1.7) for $\lambda = 0$. The set of equilibria of the fast flow is just $\mathcal{C}_{\mathcal{F}}$, and along the fast flow, η is constant.

If $p = (x_p, \eta_p) \in \mathcal{C}_{\mathcal{F}}$, then x_p is a critical point of the function $f_{\eta_p} : M \rightarrow \mathbb{R}$. A *fast solution* from $p_- \in \mathcal{C}_{\mathcal{F}}$ to $p_+ \in \mathcal{C}_{\mathcal{F}}$ is a solution $\tilde{x}(t)$ of (1.6)–(1.7) for $\lambda = 0$ such that $\lim_{t \rightarrow \pm\infty} \tilde{x}(t) = p_{\pm}$. A *fast orbit* from p_- to p_+ is an equivalence class of fast solutions from p_- to p_+ modulo time translation. A fast solution or orbit is *trivial* if the orbit consists of a single point.

For $p = (x_p, \eta_p) \in \mathcal{C}_{\mathcal{F}}$, x_p is a nondegenerate critical point if and only if $p \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$. Note that the Morse index of a nondegenerate critical point x of f_{η} can be defined as the number of negative eigenvalues of the matrix $\text{Hess} f_{\eta}(x)$. We define the Morse index of a degenerate critical point the same way. We shall use the already-introduced notation $\text{index}(x_p, f_{\eta_p})$ to denote the Morse index in this sense of x_p as a critical point of f_{η_p} .

$\text{index}(x_p, f_{\eta_p})$ is locally constant on $\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$. Near $p \in \mathcal{C}_{\mathcal{F}}^{sing}$, the values of $\text{index}(x_p, f_{\eta_p})$ on the two branches differ by one, and $\text{index}(x_p, f_{\eta_p})$ is the lower of these two numbers.

4.2.1. Transversality assumptions

On $\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$, the fast flow is normally hyperbolic. For any subset β of $\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$ define

$$W^u(\beta) = \left\{ (x, \eta) \in M \times \mathbb{R} \mid \lim_{t \rightarrow -\infty} \Phi_t(x, \eta) \in \beta \right\} \quad (4.5)$$

and

$$W^s(\beta) = \left\{ (x, \eta) \in M \times \mathbb{R} \mid \lim_{t \rightarrow +\infty} \Phi_t(x, \eta) \in \beta \right\}. \quad (4.6)$$

Note that in this section we always have $\lambda = 0$, so we will not use the notation of Subsection 2.3 to specify a value of λ . If β is connected, define $\text{index}(\beta)$ to be $\text{index}(x_p, f_{\eta_p})$ for any $p = (x_p, \eta_p)$ in β .

We assume

- (A10) $W^u(\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing})$ and $W^s(\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing})$ intersect transversally in $M \times \mathbb{R}$.
- (A11) For each $p = (x_p, \eta_p) \in \text{Crit}(\mathcal{F})$, the pair (f_{η_p}, g) is a Morse-Smale pair on M .

Assumption (A10) can be thought of as a weak version of the Morse-Bott-Smale transversality condition (see [23]) for a Morse-Bott function and a metric. (A10) does not imply (A11). However, (A11) implies that if $p = (x_p, \eta_p) \in \text{Crit}(\mathcal{F})$, then for any pair of critical points $y_-, y_+ \in M$ of the function f_{η_p} , the unstable manifold of y_- and the stable manifold of y_+ intersect transversally in M . This implies that the transversality in (A10) holds for η near η_p .

There are two special types of nontrivial fast orbits that play important roles in our construction. We introduce them in the remainder of this subsection.

4.2.2. Handle-slides and cusp orbits

By (A10) we see, if $\beta_1, \beta_2 \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$ are connected, then

$$\dim (W^u(\beta_1) \cap W^s(\beta_2)) = \text{index}(\beta_1) - \text{index}(\beta_2) + 1. \quad (4.7)$$

In particular, when $\text{index}(\beta_1) = \text{index}(\beta_2)$, the dimension is one, so the intersection consists of discrete nontrivial fast orbits. We call such an orbit a *handle-slide*. If a handle-slide is contained in $M \times \{\eta\}$, we say that a handle-slide happens at η .

We assume:

- (A12) The composition $\pi : \text{Crit}(\mathcal{F}) \cup \mathcal{C}_{\mathcal{F}}^{sing} \hookrightarrow M \times \mathbb{R} \rightarrow \mathbb{R}$ is injective. Moreover, handle-slides happen at distinct values of η that are not in the image of π .

Now we look near a point $p = (x_p, \eta_p) \in \mathcal{C}_{\mathcal{F}}^{sing}$. All the eigenvalues of $\text{Hess}f_{\eta_p}(x_p)$ have absolute values great than a constant $a > 0$, except for the one that changes its sign at p . Then we define $W_a^u(p)$ to be the space of maps $u : (-\infty, 0] \rightarrow M$ satisfying

1. $(u(t), \eta_p) = \Phi_t(u(0), \eta_p)$ for $t \leq 0$;
2. For sufficiently large T , for $t \in (-\infty, -T]$,

$$u(t) = \exp_{x_p} V(t), \quad V \in W_a^{1,2}((-\infty, -T], T_{x_p}M).$$

Here for any open subset $\Omega \subset \mathbb{R}$, the space $W_a^{1,2}(\Omega)$ is the space of functions f with $e^{a|t|}f \in W^{1,2}(\Omega)$.

Similarly we can define $W_a^s(p)$. By the map $u \mapsto u(0)$, $W_a^u(p)$ and $W_a^s(p)$ are naturally identified with smooth submanifolds of M , which are called the a -exponential unstable and stable manifolds of p , having dimensions

$$\dim W_a^u(p) = \text{index}(x_p, f_{\eta_p}), \quad \dim W_a^s(p) = n - \text{index}(x_p, f_{\eta_p}) - 1. \quad (4.8)$$

Let $\mathring{W}^u(p) = W^u(p) \setminus W_a^u(p)$ and $\mathring{W}^s(p) = W^s(p) \setminus W_a^s(p)$. These manifolds are unions of orbits that approach p polynomially rather than exponentially. They are submanifolds of $M \times \{\eta_p\}$ of dimension $\dim W_a^u(p) + 1$ and $\dim W_a^s(p) + 1$. We assume

- (A13) For each $p = (x_p, \eta_p) \in \mathcal{C}_{\mathcal{F}}^{sing}$ and $q = (x_q, \eta_q) \in \mathcal{C}_{\mathcal{F}}$ with $\eta_p = \eta_q = \eta$, $W_a^u(p)$ and $W^s(q)$ (respectively $W_a^s(p)$ and $W^u(q)$) intersect transversely in $M \times \{\eta\}$, $\mathring{W}^u(p)$ and $W^s(q)$ (respectively $\mathring{W}^s(p)$ and $W^u(q)$) intersect transversally in $M \times \{\eta\}$.

The reader may refer to [24] and [21] for this transversality in the case of Lagrangian intersections.

(A13) implies that for $p = (x_p, \eta_p) \in \mathcal{C}_{\mathcal{F}}^{sing}$ and $q = (x_q, \eta_q) \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$, $W_a^u(p) \cap W^s(q) = \emptyset$ if $\text{index}(x_q, f_{\eta_q}) \geq \text{index}(x_p, f_{\eta_p}) - 1$; and $W^u(q) \cap W_a^s(p) = \emptyset$ if $\text{index}(x_p, f_{\eta_p}) \geq \text{index}(x_q, f_{\eta_q}) - 2$. For example, $W^u(q) \cap W_a^s(p) = \emptyset$ if

$$\begin{aligned} 0 &\geq \dim W^u(q) + \dim W_a^s(p) - (n+1) = \text{index}(x_q, f_{\eta_q}) + (n - \text{index}(x_p, f_{\eta_p}) - 1) \\ &\quad - (n+1) = \text{index}(x_q, f_{\eta_q}) - \text{index}(x_p, f_{\eta_p}) - 2. \end{aligned} \quad (4.9)$$

When the inequalities are equalities, $W^u(p) \cap W^s(q) = \mathring{W}^u(p) \cap W^s(q)$ and $W^u(q) \cap W^s(p) = W^u(q) \cap \mathring{W}^s(p)$ consist of isolated orbits corresponding to solutions that approach p like $\frac{1}{t}$ rather than exponentially. We call such orbits *cuspidal orbits*; parametrized ones are *cuspidal solutions*. An argument similar to one in [21] shows that between p and q there are only finitely many cuspidal orbits.

4.2.3. Short fast orbits

By Proposition 11, near $p \in \mathcal{C}_{\mathcal{F}}^{sing}$, one can parametrize $\mathcal{C}_{\mathcal{F}}$ by $\gamma_p : (-\epsilon, \epsilon) \rightarrow \mathcal{C}_{\mathcal{F}}$ such that

$$\gamma_p(0) = p, \quad \eta(t) = \eta_p \pm s^2. \quad (4.10)$$

We choose the orientation of γ_p such that $\text{index}(\gamma_p((-\epsilon, 0))) = \text{index}(\gamma_p((0, \epsilon))) + 1$. For ϵ small enough there exists a unique such parametrization γ and we call γ the *canonical parametrization* near p . Then, for ϵ small enough, there is a unique orbit of the flow of $-\nabla f_{\eta(s)}$ from $\gamma(-s)$ to $\gamma(s)$ and they are all contained in a small neighborhood of x_p .

Lemma 12. *There exists $\epsilon_0 > 0$ such that for all $p \in \mathcal{C}_{\mathcal{F}}^{sing}$ and for $s \in (0, \epsilon_0)$ there exists a unique orbit of the fast flow from $\gamma_p(-s)$ to $\gamma_p(s)$, where $\gamma_p : (-\epsilon_0, \epsilon_0) \rightarrow \mathcal{C}_{\mathcal{F}}$ is the canonical parametrization of $\mathcal{C}_{\mathcal{F}}$ near p .*

If $0 < s \leq \epsilon < \epsilon_0$, then we call a fast orbit from $\gamma_p(-s)$ to $\gamma_p(s)$ an ϵ -short orbit.

Lemma 13. *For any ϵ small enough, there exists $\delta = \delta(\epsilon) > 0$ such that, for all $\eta \in \mathbb{R}$, any orbit of the flow of $-\nabla f_{\eta}$ which is not an ϵ -short orbit has energy no less than δ .*

Proof. Denote $\mathcal{C}_{\mathcal{F}}^{sing, \epsilon} = \bigcup_{p \in \mathcal{C}_{\mathcal{F}}^{sing}} \gamma_p((-\epsilon, \epsilon))$. Because we have assumed that the map $\mathcal{C}_{\mathcal{F}}^{sing} \rightarrow M \times \mathbb{R} \rightarrow \mathbb{R}$ is injective, any fast orbit both of whose beginning and end lie in $\mathcal{C}_{\mathcal{F}}^{sing, \epsilon}$ must be an ϵ -short orbit. So suppose the lemma doesn't hold, then there exists $\epsilon > 0$ and a sequence of nontrivial fast orbits $(y_k, \eta_k) \subset M \times \{\eta_k\}$ such that

$$\lim_{k \rightarrow \infty} f_{\eta_k}(y_k(-\infty)) - f_{\eta_k}(y_k(+\infty)) = 0 \quad (4.11)$$

and they are not ϵ -short ones. Hence without loss of generality, we may assume that for all k , $y_k(-\infty) \notin \mathcal{C}_{\mathcal{F}}^{sing, \epsilon}$.

1. $\lim_{k \rightarrow \infty} \eta_k = \pm\infty$, which implies that $y_k(\pm\infty)$ converges to points in $\text{Crit}(\mu)$. However, if we rescale the orbit by $z_k(t) = y_k(\eta_k^{-1}t)$, then z_k converges to a (possibly broken) nontrivial orbit of the flow of $-\nabla \mu$, which contradicts with (4.11).
2. $\lim_{k \rightarrow \infty} \eta_k = \eta_{\infty} \in \mathbb{R}$. Then a subsequence of y_k converges to a broken orbit of $f_{\eta_{\infty}}$, which must be a constant one. Hence

$$\lim_{k \rightarrow \infty} (y_k(\pm\infty)) := y_{\infty} \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing, \epsilon}. \quad (4.12)$$

So for k large enough, $y_k(\pm\infty)$ lie in the same connected component of $\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing, \frac{5}{2}}$. This is impossible because on each such component, distinct points have distinct values of η .

□

4.3. The slow equation

On $\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$, which is parameterized by η , the slow equation is given by restricting (1.7) to $\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$ and dividing by λ^2 :

$$\eta' = -\mu(x(\eta)). \quad (4.13)$$

Thus η' changes sign only when $\mu = 0$. We have

$$\text{Crit}(\mathcal{F}) = \{(x, \eta) \in \mathcal{C}_{\mathcal{F}} \mid \mu(x) = 0\}.$$

Thus η' changes sign only at points of $\text{Crit}(\mathcal{F})$. By (A9), η' does not change sign at points of $\mathcal{C}_{\mathcal{F}}^{sing}$.

Proposition 14. *Equilibria of (4.13) are hyperbolic.*

Proof. Let $p = (x_p, \eta_p)$ be an equilibrium of (4.13). Then $\mu(x_p) = 0$, so by (A3), $d\mu(x_p) \neq 0$. Hence we can choose local coordinates (x_1, \dots, x_n) near x_p such that $\mu(x_1, \dots, x_n) = x_n$. Near p , $\mathcal{C}_{\mathcal{F}}$ can be defined without reference to the metric g by the equations

$$\frac{\partial f}{\partial x_i} + \eta \delta_{nj} = 0, \quad j = 1, \dots, n; \quad \delta_{ij} = \text{Kronecker delta}. \quad (4.14)$$

Let $e_n = (0 \ \dots \ 0 \ 1)^T$, the n th standard unit basis vector in \mathbb{R}^n . Then the derivative of the system (4.14) is the $n \times (n+1)$ matrix

$$(\text{Hess}f(x) \ e_n) = \begin{pmatrix} A & b & 0 \\ b^T & c & 1 \end{pmatrix}, \quad (4.15)$$

where A is an $(n-1) \times (n-1)$ matrix, and $(b^T, c) = \left(\frac{\partial^2 f}{\partial x_1 \partial x_n}, \dots, \frac{\partial^2 f}{\partial x_{n-1} \partial x_n}, \frac{\partial^2 f}{\partial x_n^2} \right)$.

Since x_p is a critical point of $f|_{\mu^{-1}(0)}$, by (A5), A is invertible. The matrix (4.15) has full rank; a nonzero vector in its kernel is tangent at p to $\mathcal{C}_{\mathcal{F}}$. Since $p \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$, such a vector has nonzero η -component, and one can check that it has nonzero x_n -component. Therefore near p , when $\mathcal{C}_{\mathcal{F}}$ is parameterized by η , we have $x(\eta) = (y(\eta), x_n(\eta))$ with

$$x_n(\eta) = a(\eta - \eta_p) + \mathcal{O}(\eta - \eta_p)^2, \quad a \neq 0.$$

But then (4.13) becomes

$$\eta' = a(\eta - \eta_p) + \mathcal{O}(\eta - \eta_p)^2,$$

which proves the result. □

Next we discuss the slow equation near points of $\mathcal{C}_{\mathcal{F}}^{sing}$. Recall that if p_0 is an equilibrium of $\dot{p} = h(p)$, h is C^s , $Dh(p_0)$ has m eigenvalues (counting multiplicity) with real part 0, and E is the corresponding m -dimensional invariant subspace of $\dot{v} = Dh(p_0)v$, then there is a C^s invariant manifold through p_0 and tangent there to E , called the *center manifold* of $\dot{p} = h(p)$ at p_0 .

Proposition 15. *Let $p = (x_p, \eta_p) \in \mathcal{C}_{\mathcal{F}}^{sing}$. We can choose local coordinates (x_1, \dots, x_n) on M near x_p , with x_p corresponding to 0, in which*

$$D\nabla f(0) = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}, \quad A \text{ an invertible symmetric } (n-1) \times (n-1) \text{ matrix.} \quad (4.16)$$

The center manifold of (1.6)–(1.7) at p is parameterized by (x_n, η) . There are numbers $c \neq 0$ and $d \neq 0$ such that the system (1.6)–(1.7), restricted to the center manifold, is

$$x'_n = c(\eta - \eta_p) + dx_n^2 + \dots, \quad (4.17)$$

$$\eta' = -\lambda^2(\mu(x_p) + \dots), \quad (4.18)$$

where \dots indicates terms of higher order (and, in (4.17), other second-order terms). Therefore near p , $\mathcal{C}_{\mathcal{F}}$ is parameterized by x_n , and $\eta|_{\mathcal{C}_{\mathcal{F}}}$ has a nondegenerate critical point at p .

Proof. (A7) implies that we can choose local coordinates (x_1, \dots, x_n) on M near x_p , with x_p corresponding to 0, in which (4.16) holds. In these coordinates, $\mu = \mu(x_p) + a^T x + \dots$ with $a \in \mathbb{R}^n$. Write $x = (y, x_n)$ with $y \in \mathbb{R}^{n-1}$, and write $a = (b, c)$ with $b \in \mathbb{R}^{n-1}$ and $c \in \mathbb{R}$. Near p the system (1.6)–(1.7) becomes

$$y' = Ay + b(\eta - \eta_p) + \dots, \quad (4.19)$$

$$x'_n = c(\eta - \eta_p) + dx_n^2 + \dots, \quad (4.20)$$

$$\eta' = -\lambda^2(\mu(x_p) + \dots), \quad (4.21)$$

where \dots indicates higher-order terms (and, in (4.20), other second-order terms). (A7) implies that $c \neq 0$, and (A8) implies that $d \neq 0$. The linearization of (4.19)–(4.21) at $(y, x_n, \eta) = (0, 0, \eta_p)$ is

$$\begin{pmatrix} y' \\ x'_n \\ \eta' \end{pmatrix} \begin{pmatrix} A & 0 & b \\ 0 & 0 & c \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} y \\ x_n \\ \eta - \eta_p \end{pmatrix}.$$

The matrix has a two-dimensional generalized eigenspace for the eigenvalue 0 that is tangent to the center manifold of (4.19)–(4.21) at p . The center manifold is parameterized by (x_n, η) , and to lowest order it is given by

$$y = -A^{-1}b(\eta - \eta_p)$$

The restriction of (4.19)–(4.21) to the center manifold is therefore given by (4.17)–(4.18). \square

Proposition 11 is of course just a partial restatement of Proposition 15.

4.4. Three indices

In this subsection we describe the relation between three different Morse indices:

1. For any $p = (x_p, \eta_p) \in \mathcal{C}_{\mathcal{F}}$, we have $\text{index}(x_p, f_{\eta_p})$, used in Subsection 4.2.
2. For $p = (x_p, \eta_p) \in \text{Crit}(\mathcal{F})$, we can consider $\text{index}(p, \mathcal{F})$.
3. If $p = (x_p, \eta_p) \in \text{Crit}(\mathcal{F})$, then $x_p \in \text{Crit}(f|_{\mu^{-1}(0)})$, and we can consider $\text{index}(x_p, f|_{\mu^{-1}(0)})$.

We have already seen in Lemma 1 that $\text{index}(p, \mathcal{F}) = \text{index}(x_p, f|_{\mu^{-1}(0)}) + 1$ for $p = (x_p, \eta_p) \in \text{Crit}(\mathcal{F})$. The best way to see the relation of the different indices is to look at the corresponding unstable manifolds.

Proposition 16. *Let $p = (x_p, \eta_p) \in \text{Crit}(\mathcal{F})$.*

1. *If p is a repeller of the slow equation, then $\text{index}(p, \mathcal{F}) = \text{index}(x_p, f_{\eta_p}) + 1$.*
2. *If p is an attractor of the slow equation, then $\text{index}(p, \mathcal{F}) = \text{index}(x_p, f_{\eta_p})$.*

Proof. If p is a repeller of the slow equation, then for small $\lambda > 0$, the equilibrium p of (1.6)–(1.7) has one more positive eigenvalue than the equilibrium x_p of f_{η_p} . Since the index of a critical point of a Morse function h is the number of positive eigenvalues of the corresponding equilibrium of ∇h for any metric, (1) follows. (2) is similar. \square

5. Adiabatic limit $\lambda \rightarrow 0$

5.1. Slow solutions, fast solutions, and fast-slow solutions

Let $p_-, p_+ \in \mathcal{C}_{\mathcal{F}}$.

Recall that a nontrivial fast solution from p_- to p_+ is a nonconstant solution $\tilde{x}(t)$, $-\infty < t < \infty$, of (1.6)–(1.7) for $\lambda = 0$, such that $\lim_{t \rightarrow \pm\infty} \tilde{x}(t) = p_{\pm}$.

Let $(x(\eta), \eta)$ parameterize the closure of a component of $\mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{\text{sing}}$. Let $I(\alpha_-, \alpha_+) \subset \mathbb{R}$ be the closed interval from α_- to α_+ , with $\alpha_- \leq \alpha_+$; we allow $\alpha_- = -\infty$, $\alpha_+ = +\infty$, and $\alpha_- = \alpha_+ \in \mathbb{R}$. Let $\eta(t)$, t in the interior of $I(\alpha_-, \alpha_+)$, be a solution of $\eta' = -\mu(x(\eta))$. If α_- (respectively α_+) is finite, we extend $\eta(t)$ continuously to α_- (respectively α_+). Let $p(t) = (x(\eta(t)), \eta(t))$, $t \in I(\alpha_-, \alpha_+)$, and let $p_{\pm} \in \mathcal{C}_{\mathcal{F}}$. Then $p(t)$ is a *slow solution* from p_- to p_+ (for short, a *slow solution*) provided:

1. if α_- (respectively α_+) is finite, then $p(\alpha_-) = p_-$ (respectively $p(\alpha_+) = p_+$);
2. if $\alpha_- = -\infty$ (respectively $\alpha_+ = +\infty$), then $\lim_{t \rightarrow -\infty} p(t) = p_-$ (respectively $\lim_{t \rightarrow +\infty} p(t) = p_+$).

A slow solution or its orbit is *trivial* if the orbit is a single point in $\mathcal{C}_{\mathcal{F}}$.

A *fast-slow solution* from p_- to p_+ of (1.6)–(1.7) (for short, a *fast-slow solution*) is a sequence

$$\mathcal{X} = (p_0, \sigma_1, p_1, \sigma_2, p_2, \dots, p_{n-1}, \sigma_n, p_n) \quad (5.1)$$

such that:

1. $p_0 = p_-, p_n = p_+$.
2. Each $p_i \in \mathcal{C}_{\mathcal{F}}$.
3. Each σ_i is a nontrivial fast solution or a slow solution from p_{i-1} to p_i . Trivial slow solutions are allowed, but σ_1 and σ_n are not allowed to be trivial slow solutions.
4. Either (i) σ_i is slow for i even and fast for i odd, or (ii) σ_i is slow for i odd and fast for i even.

Trivial slow solutions are allowed to deal with the possibility that a fast solution to p is followed by a fast solution from p .

An orbit (fast orbit, slow orbit, or fast-slow orbit) is an equivalence class of solutions obtained by forgetting the parametrization (but remembering the orientation in which \mathcal{F} decreases). In particular, a fast-slow orbit is denoted by $[\mathcal{X}]$ for \mathcal{X} as in (5.1).

Remark 17. A fast-slow orbit defined above is similar to a “flow line with cascades” in Morse-Bott theory (see [25]), where the fast solutions correspond to “cascades”. The difference is that the slow manifold $\mathcal{C}_{\mathcal{F}}$ is not everywhere normally hyperbolic, and we can have $p_i \in \mathcal{C}_{\mathcal{F}}^{sing}$.

In the remainder of this subsection, we will discuss some properties of fast-slow orbits from $p_- \in \mathcal{C}_{\mathcal{F}}$ to $p_+ \in \mathcal{C}_{\mathcal{F}}$. We are particularly interested in the case $p_{\pm} \in \text{Crit}(\mathcal{F})$ and $\text{index}(p_-, \mathcal{F}) - \text{index}(p_+, \mathcal{F}) = 1$.

The following proposition is obvious.

Proposition 18. *Consider a nontrivial slow solution $p(t)$, $t \in I(\alpha_-, \alpha_+)$, from p_- to p_+ .*

1. *If $p_- \in \mathcal{C}_{\mathcal{F}}^{sing}$, then either (i) for all $t \in (\alpha_-, \alpha_+)$, $\text{index}(x_{p(t)}, f_{\eta_{p(t)}}) = \text{index}(x_-, f_{\eta_-})$, or (ii) for all $t \in (\alpha_-, \alpha_+)$, $\text{index}(x_{p(t)}, f_{\eta_{p(t)}}) = \text{index}(x_-, f_{\eta_-}) + 1$.*
2. *If $p_+ \in \mathcal{C}_{\mathcal{F}}^{sing}$, then either (i) for all $t \in (\alpha_-, \alpha_+)$, $\text{index}(x_{p(t)}, f_{\eta_{p(t)}}) = \text{index}(x_+, f_{\eta_+})$, or (ii) for all $t \in (\alpha_-, \alpha_+)$, $\text{index}(x_{p(t)}, f_{\eta_{p(t)}}) = \text{index}(x_+, f_{\eta_+}) + 1$.*

A slow solution $p(t)$ for $t \in I(\alpha_-, \alpha_+)$ is *regular* if it is nontrivial and satisfies:

1. If $p_- \in \mathcal{C}_{\mathcal{F}}^{sing}$, then for all $t \in (\alpha_-, \alpha_+)$, $\text{index}(x_{p(t)}, f_{\eta_{p(t)}}) = \text{index}(x_-, f_{\eta_-}) + 1$.
2. If $p_+ \in \mathcal{C}_{\mathcal{F}}^{sing}$, then for all $t \in (\alpha_-, \alpha_+)$, $\text{index}(x_{p(t)}, f_{\eta_{p(t)}}) = \text{index}(x_+, f_{\eta_+})$.

A slow orbit is regular if it is the orbit of a regular slow solution.

See Figure 2.

Proposition 19. *Let γ be either a slow or fast solution from $p_- = (x_-, \eta_-)$ to $p_+ = (x_+, \eta_+)$.*

1. *If γ is a regular slow solution, then:*
 - (a) *If $p_- \in \mathcal{C}_{\mathcal{F}}^{sing}$, then $\text{index}(x_+, f_{\eta_+}) = \text{index}(x_-, f_{\eta_-}) + 1$.*
 - (b) *Otherwise $\text{index}(x_+, f_{\eta_+}) = \text{index}(x_-, f_{\eta_-})$*

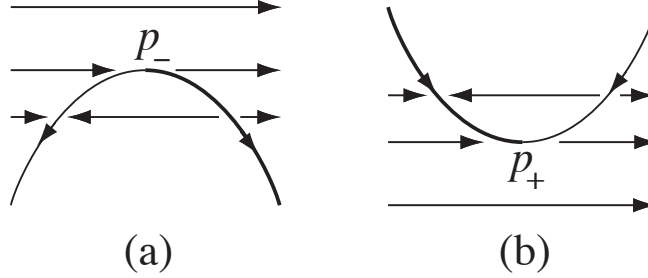


Figure 2: (a) Direction of slow flow is downward. The thick curve is a regular slow orbit that starts at p_- . Points on it have index one greater than p_- . (b) Direction of slow flow is downward. The thick curve is a regular slow orbit that ends at p_+ . Points on it have the same index as p_+ .

2. If γ is a nontrivial fast solution (so $\eta_+ = \eta_-$), then:
 - (a) At most one of p_{\pm} is in $\text{Crit}(\mathcal{F}) \cup \mathcal{C}_{\mathcal{F}}^{\text{sing}}$.
 - (b) $\text{index}(x_+, f_{\eta_+}) \leq \text{index}(x_-, f_{\eta_-})$.
 - (c) If $p_+ \in \text{Crit}(\mathcal{F}) \cup \mathcal{C}_{\mathcal{F}}^{\text{sing}}$ or $p_- \in \text{Crit}(\mathcal{F})$, then $\text{index}(x_+, f_{\eta_+}) \leq \text{index}(x_-, f_{\eta_-}) - 1$.

Proof. The result for slow solutions is immediate from their definition.

For fast solutions, (2)(a) follows from assumption (A12)

Let $p(t)$ be a fast solution, let $\text{index}(x_-, f_{\eta_-}) = k$, and $\text{index}(x_+, f_{\eta_+}) = l$. (Of course $\eta_+ = \eta_-$.)

If $p_+ \in \text{Crit}(\mathcal{F}) \cup \mathcal{C}_{\mathcal{F}}^{\text{sing}}$, then (A11) or (A13) implies that $W^u(p_-)$ and $W^s(p_+)$ intersect transversally within $M \times \{\eta_-\}$. We have $\dim W^u(p_-) = k$, and $\dim W^s(p_+) = n - l$, so the intersection is nonempty provided $k + (n - l) \geq n + 1$, i.e., $l \leq k - 1$.

If $p_- \in \text{Crit}(\mathcal{F})$, the same argument applies. However, if $p_- \in \mathcal{C}_{\mathcal{F}}^{\text{sing}}$, then $\dim W^u(p_-) = k + 1$. We again have $\dim W^s(p_+) = n - l$, so the intersection is nonempty provided $(k + 1) + (n - l) \geq n + 1$, i.e., $l \leq k$.

If neither p_- nor p_+ is in $\text{Crit}(\mathcal{F}) \cup \mathcal{C}_{\mathcal{F}}^{\text{sing}}$, then we do not know that $W^u(p_-)$ and $W^s(p_+)$ intersect transversally within $M \times \{\eta_-\}$. However, by (A10), $(k + 1) + (n - l + 1) \geq n + 2$, so $l \leq k$. \square

Proposition 20. Consider a fast-slow solution

$$\mathcal{X} = (p_0, \sigma_1, p_1, \sigma_2, p_2, \dots, p_{n-1}, \sigma_n, p_n)$$

from $p_- = (x_-, \eta_-)$ to $p_+ = (x_+, \eta_+)$, with neither in $\mathcal{C}_{\mathcal{F}}^{\text{sing}}$ and $\text{index}(x_+, f_{\eta_+}) \geq \text{index}(x_-, f_{\eta_-}) = k$. Then

1. For all $i \geq 1$, $\text{index}(x_{p_i}, f_{p_i}) = k$, except for the case where $p_i \in \mathcal{C}_{\mathcal{F}}^{\text{sing}}$ and σ_i is a fast solution, in which case $\text{index}(x_{p_i}, f_{p_i}) = k - 1$. In particular, $\text{index}(x_+, f_{\eta_+}) = k$.
2. All slow solutions appearing in \mathcal{X} are regular. In particular, they are nontrivial.

3. If $p_i \in \text{Crit}(\mathcal{F})$, then $i = 0$ or $i = n$. In the first case, $p_0 = p_- \in \text{Crit}^-(\mathcal{F})$ and σ_1 is slow; in the second case, $p_n = p_+ \in \text{Crit}^+(\mathcal{F})$ and σ_n is slow.
4. The fast solutions are handle-slides or cusp solutions.

Proof. Along a fast-slow solution, $\text{index}(x_{p_i}, f_{p_i}) \leq \text{index}(x_{p_{i-1}}, f_{p_{i-1}})$, except along a regular slow solution for which $p_{i-1} \in \mathcal{C}_{\mathcal{F}}^{\text{sing}}$ and $p_i \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{\text{sing}}$; in this case $\text{index}(x_{p_i}, f_{p_i}) = \text{index}(x_{p_{i-1}}, f_{p_{i-1}}) + 1$. However, the index drops along the fast solution from p_{i-2} to p_{i-1} . Thus (1) holds; also, nontrivial slow solutions must be regular and (3) must hold, otherwise we would have $\text{index}(x_+, f_{\eta_+}) < \text{index}(x_-, f_{\eta_-})$. (3) rules out fast solutions that start or end in $\text{Crit}(\mathcal{F})$; any remaining fast solutions are handle-slides or cusp solutions, so (4) holds. Since handle-slides and cusp solutions occur at different values of η , slow solutions must be nontrivial, which completes the proof of (2). \square

Now we describe fast-slow solutions between two critical points whose indices differ by 1. Recall from Subsection 2.1 that $\text{Crit}_k(\mathcal{F})$ denotes the set of index k critical points of \mathcal{F} . Let $\text{Crit}_k^+(\mathcal{F})$ (respectively $\text{Crit}_k^-(\mathcal{F})$) denote the set of equilibria in $\text{Crit}_k(\mathcal{F})$ that are stable (respectively unstable) equilibria of the slow equation. Let $\text{Crit}^\pm(\mathcal{F}) = \cup_{k \geq 0} \text{Crit}_k^\pm(\mathcal{F})$.

Proposition 21. *Let $p_\pm = (x_\pm, \eta_\pm) \in \text{Crit}(\mathcal{F})$ with $p_- \in \text{Crit}_k(\mathcal{F})$, $p_+ \in \text{Crit}_{k-1}(\mathcal{F})$. A fast-slow solution $\mathcal{X} = (p_0, \sigma_1, p_1, \sigma_2, p_2, \dots, p_{n-1}, \sigma_n, p_n)$ from p_- to p_+ has the following properties.*

1. If σ_i is a fast solution from $p_{i-1} = (x_{p_{i-1}}, \eta_{p_{i-1}})$ to $p_i = (x_{p_i}, \eta_{p_i})$ (of course $\eta_{p_i} = \eta_{p_{i-1}}$), then $\text{index}(x_{p_i}, f_{\eta_{p_i}}) = \text{index}(x_{p_{i-1}}, f_{\eta_{p_{i-1}}})$, unless:
 - (a) $p_i \in \mathcal{C}_{\mathcal{F}}^{\text{sing}}$, in which case $\text{index}(x_{p_i}, f_{\eta_{p_i}}) = \text{index}(x_{p_{i-1}}, f_{\eta_{p_{i-1}}}) - 1$.
 - (b) $i = 1$, $p_0 = p_- \in \text{Crit}_k^+(\mathcal{F})$, in which case σ_1 is fast (so $\eta_{p_1} = \eta_p$) and $\text{index}(x_{p_1}, f_{\eta_{p_1}}) = \text{index}(x_-, f_{\eta_-}) - 1$.
 - (c) $i = n$, $p_n = p_+ \in \text{Crit}_{k-1}^-(\mathcal{F})$, in which case σ_n is fast (so $\eta_q = \eta_{p_{n-1}}$) and $\text{index}(x_+, f_{\eta_+}) = \text{index}(x_{p_{n-1}}, f_{\eta_{p_{n-1}}}) - 1$.
2. All slow solutions appearing in \mathcal{X} are regular. In particular, they are nontrivial.
3. For $i = 1, \dots, n-1$, $p_i \notin \text{Crit}(\mathcal{F})$.
4. The fast solutions in \mathcal{X} are handle-slides or cusp solutions, except for σ_1 when $p_- \in \text{Crit}_k^+(\mathcal{F})$ and σ_n when $p_+ \in \text{Crit}_{k-1}^-(\mathcal{F})$.

In addition:

- (I) Suppose $p_- \in \text{Crit}_k^+(\mathcal{F})$ and $p_+ \in \text{Crit}_{k-1}^+(\mathcal{F})$. Then the odd σ_i are fast, and n is even.
- (II) Suppose $p_- \in \text{Crit}_k^+(\mathcal{F})$ and $p_+ \in \text{Crit}_{k-1}^-(\mathcal{F})$. Then the odd σ_i are fast, and n is odd.
- (III) Suppose $p_- \in \text{Crit}_k^-(\mathcal{F})$ and $p_+ \in \text{Crit}_{k-1}^+(\mathcal{F})$. Then the odd σ_i are slow, and n is odd.
- (IV) Suppose $p_- \in \text{Crit}_k^-(\mathcal{F})$ and $p_+ \in \text{Crit}_{k-1}^-(\mathcal{F})$. Then the odd σ_i are slow, and n is even.

Proof. We prove (1)–(4) separately for the cases (I) and (II). The proofs for cases (III) and (IV) are similar so we omit them. We remark that in one of the cases, case (III), we can have $n = 1$, which means \mathcal{X} may contain a single slow solution and no fast solutions.

Proof in case (I). In this case, $\text{index}(x_-, f_{\eta_-}) = k$ and $\text{index}(x_+, f_{\eta_+}) = k - 1$. Since p_- is an attractor for the slow equation, the first solution σ_1 must be fast, so the odd σ_i are fast. By Proposition 19 (2), $p_1 \notin \text{Crit}(\mathcal{F}) \cup \mathcal{C}_{\mathcal{F}}^{\text{sing}}$ and $\text{index}(x_{p_1}, f_{\eta_{p_1}}) \leq k - 1$. Consider the portion of the fast-slow solution from p_1 to q . Then Proposition 20 implies that $\text{index}(x_{p_1}, f_{\eta_{p_1}}) = k - 1$, conclusion (2) of Proposition 21 holds, and $p_i \notin \text{Crit}^{\pm}(\mathcal{F})$ for $i = 2, \dots, n - 1$ so (3) holds. Proposition 20 also implies that all fast solutions except the first are handle-slides or cusp solutions, so (4) holds.

Since p_+ is an attractor for the slow equation, *a priori* the last solution could be slow or fast. However, if the last solution were fast, from Proposition 19 (2) we would have $\text{index}(x_+, f_{\eta_+}) < \text{index}(x_{p_{n-1}}, f_{\eta_{p_{n-1}}}) = k - 1$, contradiction. Therefore the last solution is slow, so n is even. Then (1) and (4) follow from Proposition 20.

Proof in case (II). As with case (I), $\text{index}(x_-, f_{\eta_-}) = k$; σ_1 must be fast, so the odd σ_i are fast; $p_1 \notin \text{Crit}(\mathcal{F}) \cup \mathcal{C}_{\mathcal{F}}^{\text{sing}}$; and $\text{index}(x_{p_1}, f_{\eta_{p_1}}) \leq k - 1$. Since p_+ is a repeller for the slow equation, $\text{index}(x_+, f_{\eta_+}) = k - 2$, and the last solution must be fast. Then $p_{n-1} \notin \text{Crit}(\mathcal{F}) \cup \mathcal{C}_{\mathcal{F}}^{\text{sing}}$, and by Proposition 19 (2), $\text{index}(x_{p_{n-1}}, f_{\eta_{p_{n-1}}}) \geq k - 1$. As in the previous argument, for the portion of the fast-slow solution from p_1 to p_{n-1} , apply Proposition 20; we obtain (1)–(4). Since the first and last solutions are fast, n is odd. \square

5.2. The chain complex

Now let $p, q \in \text{Crit}(\mathcal{F})$. We denote by $\tilde{\mathcal{N}}^0(p, q)$ the space of all fast-slow solutions from p to q and $\mathcal{N}^0(p, q)$ the space of all fast-slow orbits from p to q .

Proposition 22. *If $p, q \in \text{Crit}(\mathcal{F})$ and $\text{index}(q, \mathcal{F}) = \text{index}(p, \mathcal{F}) - 1$, then $\#\mathcal{N}^0(p, q) < +\infty$.*

Proof. By Proposition 21, we only need to show that there are finitely many handle-slides or cusp orbits contained in $\mathcal{F}^{-1}([\mathcal{F}(q), \mathcal{F}(p)])$. But this follows from our description of $\mathcal{C}_{\mathcal{F}}$ at the beginning of Section 4 and the transversality assumption (A10), (A11), (A13). \square

Now we define a chain complex $(\mathcal{C}^0, \partial^0)$ of \mathbb{Z}_2 -modules. It has the same generators and gradings as \mathcal{C}^{λ} . Its boundary operator $\partial^0 : \mathcal{C}_k^0 \rightarrow \mathcal{C}_{k-1}^0$ is defined by

$$\langle \partial^0 p, q \rangle = \#\mathcal{N}^0(p, q) \bmod 2. \quad (5.2)$$

The main theorem of this section and of the paper is:

Theorem 23. *$(\mathcal{C}^0, \partial^0)$ is a chain complex, i.e., $\partial^0 \circ \partial^0 = 0$. Moreover, the homology $H_k(\mathcal{C}^0, \partial^0)$ is canonically isomorphic to $H_{k-1}(\mu^{-1}(0), \mathbb{Z}_2)$.*

The proof of this theorem is carried out in the following way. Let $p, q \in \text{Crit}(\mathcal{F})$. First we prove that for small $\lambda > 0$, any orbit in $\mathcal{M}^\lambda(p, q)$ is close to a fast-slow orbit in $\mathcal{N}^0(p, q)$ (the compactness theorem, Theorem 25). Then we restrict to the case $\text{index}(p, \mathcal{F}) - \text{index}(q, \mathcal{F}) = 1$ and show that for each small $\lambda > 0$, each orbit in $\mathcal{N}^0(p, q)$ is close to exactly one orbit in $\mathcal{M}^\lambda(p, q)$ (the gluing theorem, Theorem 30). The resulting one-to-one correspondence implies that $(\mathcal{C}^0, \partial^0)$ is isomorphic to $(\mathcal{C}^\lambda, \partial^\lambda)$ for small $\lambda \in \Lambda^{reg}$. Then by Corollary 6, we obtain the isomorphism.

5.3. The compactness theorem

Definition 24. Let $p, q \in \text{Crit}(\mathcal{F})$. Let $\lambda_\nu \in \mathbb{R}^+$ be a sequence such that $\lim_{\nu \rightarrow \infty} \lambda_\nu = 0$. We say that a sequence of (parametrized) $\tilde{p}_\nu = (\tilde{x}_\nu, \tilde{\eta}_\nu) \in \widetilde{\mathcal{M}}^{\lambda_\nu}(p, q)$ converges to a parametrized fast-slow solution

$$\mathcal{X} = (p_0, \sigma_1, p_1, \sigma_2, \dots, p_{n-1}, \sigma_n, p_n) \in \tilde{\mathcal{N}}^0(p, q),$$

if the following hold.

1. For each fast solution σ_i contained in \mathcal{X} , there exists $t_{i,\nu} \in \mathbb{R}$ such that the sequence of maps $\tilde{p}_\nu(\cdot + t_{i,\nu})$ converges to σ_i in the C_{loc}^∞ -topology.
2. For each slow solution σ_j contained in \mathcal{X} , there exists a sequence of intervals $I_{j,\nu} \subset \mathbb{R}$ that satisfies the following two conditions.
 - (a) For every $t_{i,\nu}$ in the first condition, we have

$$\lim_{\nu \rightarrow \infty} d(I_{j,\nu}, t_{i,\nu}) = \infty.$$

- (b) $\lim_{\nu \rightarrow \infty} \tilde{p}_\nu(I_{j,\nu}) = \sigma_j$ in the Hausdorff topology.

It is easy to see that the limit is unique in $\mathcal{N}^0(p, q)$ and only depends on the sequence of orbits $[\tilde{p}_\nu]$, not the representatives. Moreover, the image of \tilde{p}_ν converges to the image of \mathcal{X} in the Hausdorff topology.

In the remainder of this section we prove the following theorem.

Theorem 25. Suppose $p, q \in \text{Crit}(\mathcal{F})$ and $\lambda_\nu \rightarrow 0^+$ be a sequence of real numbers. Suppose $\tilde{p}_\nu = (\tilde{x}_\nu, \tilde{\eta}_\nu) \in \widetilde{\mathcal{M}}^{\lambda_\nu}(p, q)$. Then there exists a subsequence (still indexed by ν), and a fast-slow solution $\mathcal{X} = (p_0, \sigma_1, p_1, \sigma_2, \dots, p_{n-1}, \sigma_n, p_n) \in \tilde{\mathcal{N}}^0(p, q)$, such that \tilde{p}_ν converges to \mathcal{X} in the sense of Definition 24.

A key point in proving this theorem is to use the energy control to bound the number of pieces of fast orbits appearing in the limit. This is similar to the proof of the Gromov compactness theorem for J -holomorphic curves (see, for example, [22]) where there is a lower bound of energy for any nontrivial bubble. After finding all the fast orbits in the limit, each adjacent pair of them is connected by a slow orbit. Before the first fast orbit and after the last fast orbit, one may or may not need to add nontrivial slow ones, depending on the types of p and q (i.e., repeller or attractor of the slow flow).

From now on the sequence \tilde{p}_ν is given and we are free to take subsequences as many times as we want.

5.3.1. The limit set

For any subsequence ν' of ν and sequence of intervals $I_{\nu'} \subset \mathbb{R}$, set

$$L_\infty(\nu', I_{\nu'}) = \bigcap_{k \geq 1} \overline{\bigcup_{\nu' \geq k} \tilde{p}_{\nu'}(I_{\nu'})} \subset M \times \mathbb{R}. \quad (5.3)$$

Lemma 26. *The following are true.*

- (I) $(x, \eta) \in L_\infty(\nu', I_{\nu'})$ if and only if there exists a subsequence ν'' of ν' and a sequence of numbers $t_{\nu''} \in I_{\nu''}$, such that $\lim_{\nu'' \rightarrow \infty} \tilde{p}_{\nu''}(t_{\nu''}) = (x, \eta)$.
- (II) $L_\infty(\nu', I_{\nu'})$ is compact.
- (III) If there exist $s_{\nu'} \in I_{\nu'}$ such that $\lim_{\nu' \rightarrow \infty} \tilde{p}_{\nu'}(s_{\nu'})$ exists, then $L_\infty(\nu', I_{\nu'})$ is connected.

Proof. The first statement and the closedness are by definition, and compactness follows from Lemma 3. It remains to show the connectedness. If $L_\infty(\nu', I_{\nu'})$ is not connected, then there are two nonempty closed subset $L_1, L_2 \subset M \times \mathbb{R}$ such that $L_\infty(\nu', I_{\nu'}) = L_1 \cup L_2$, and L_1, L_2 has a nonzero distance. Suppose $\lim_{\nu' \rightarrow \infty} \tilde{p}_{\nu'}(s_{\nu'}) \in L_1$. Then there exists a subsequence (still indexed by ν') and $w_{\nu'} \in I_{\nu'}$ such that $d(\tilde{p}_{\nu'}(w_{\nu'}), L_2) \rightarrow 0$. Then there exists $w'_{\nu'} \in [s_{\nu'}, w_{\nu'}]$ such that $d(\tilde{p}_{\nu'}(w'_{\nu'}), L_1) = d(\tilde{p}_{\nu'}(w'_{\nu'}), L_2) > 0$ for the same subsequence. Then there is a subsequence of $\tilde{p}_{\nu'}(w'_{\nu'})$ that converges to a point not in L_1 and L_2 , which contradicts $L_\infty(\nu', I_{\nu'}) = L_1 \cup L_2$. \square

The following lemma is obvious.

Lemma 27. *Suppose we have two sequences of numbers, $s_\nu < t_\nu$ with $\lim_{\nu \rightarrow \infty} \tilde{p}_\nu(s_\nu) = a$, $\lim_{\nu \rightarrow \infty} \tilde{p}_\nu(t_\nu) = b$. Then $\mathcal{F}(a) \geq \mathcal{F}(b)$, and, for any $z \in L_\infty(\nu, [s_\nu, t_\nu])$, $\mathcal{F}(z) \in [\mathcal{F}(a), \mathcal{F}(b)]$.*

5.3.2. Identify all fast orbits

By the compactness of $L_\infty(\nu, \mathbb{R})$ and the fact that $\lambda_\nu \rightarrow 0$, for any sequence $t_\nu \in \mathbb{R}$, there exists a subsequence such that on any finite interval, $\tilde{\eta}_\nu(\cdot + t_\nu)$ converges to a constant function η_1 . Then, by the usual argument of Morse theory, there is a subsequence of $\tilde{p}_\nu(\cdot + t_\nu)$ converging to a fast solution $\tilde{y} := (y_1, \eta_1) : \mathbb{R} \rightarrow M \times \mathbb{R}$ with $\text{Im } \tilde{y}_1 \subset M \times \{\eta_1\}$, in C_{loc}^∞ -topology. So $L_\infty(\nu, \mathbb{R})$ is the union of fast orbits (which could contain constant orbits). We first prove that for a suitable subsequence it only contains finitely many nonconstant ones.

Proposition 28. *There exists a subsequence (still indexed by ν without loss of generality), and $t_{1,\nu}, \dots, t_{n,\nu} \in \mathbb{R}$ (n could be zero), satisfying*

- (I) $t_{1,\nu} < t_{2,\nu} < \dots < t_{n,\nu}$, $\lim_{\nu \rightarrow \infty} |t_{i,\nu} - t_{j,\nu}| = \infty$ for all $i \neq j$;
- (II) $\tilde{p}_\nu(\cdot + t_{i,\nu})$ converges to a (nonconstant) fast solution $\tilde{y}_i := (y_i, \eta_i) : \mathbb{R} \rightarrow M \times \mathbb{R}$ in the C_{loc}^∞ -topology;
- (III) for any sequence $s_\nu \in \mathbb{R}$ such that $\lim_{\nu \rightarrow \infty} |s_\nu - t_{i,\nu}| = \infty$ for all i , any convergent subsequence of $\tilde{p}_\nu(s_\nu)$ has limit in $\mathcal{C}_\mathcal{F}$.

Proof. If $L_\infty(\nu, \mathbb{R}) \subset \mathcal{C}_\mathcal{F}$, then nothing has to be proved. Suppose it is not the case. Then we take $\delta, \epsilon > 0$ as in Lemma 13 and small enough, and $r > 0$ small enough, such that the following are true.

- All ϵ -short orbits are contained in $\cup_{w \in \mathcal{C}_\mathcal{F}^{sing}} B_r(w)$.
- All nontrivial fast orbits which are not ϵ -short have energy at least δ .
- For each $w \in \mathcal{C}_\mathcal{F}^{sing}$, $\mathcal{C}_\mathcal{F} \cap B_{2r}(w)$ can be parametrized using the canonical parametrization as in (4.10), so that \mathcal{F} restricted to both components of $\mathcal{C}_\mathcal{F} \cap B_{2r}(w) \setminus \{w\}$ is monotonic.

Let $N = \lfloor \frac{E}{\delta} \rfloor$, where $E = \mathcal{F}(p_-) - \mathcal{F}(p_+)$ is the total energy. It is easy to prove by induction that there exists $k \leq N$ (possibly zero), a subsequence ν' of the original sequence, a sequence of numbers $\{t_{i,\nu'}\}_{i=1,\dots,k}$, and fast solutions $\tilde{y}_i := (y_i, \eta_i)$ satisfying

1. $\lim_{\nu' \rightarrow \infty} t_{i+1,\nu'} - t_{i,\nu'} = +\infty$, for $i = 1, \dots, k-1$;
2. $\lim_{\nu' \rightarrow \infty} \tilde{p}_{\nu'}(t_{i,\nu'} + \cdot) = (y_i, \eta_i)$ in C_{loc}^∞ -topology;
3. each \tilde{y}_i has energy $\geq \delta$;
4. any other nonconstant fast orbit contained in $L_\infty(\nu', \mathbb{R})$ is ϵ -short.

We replace ν' by ν for simplicity. Let $N' = \#\mathcal{C}_\mathcal{F}^{sing}$. We can continue the induction to find all ϵ -short orbits, and we claim that the induction stops at finite time and we can find at most $N'(N+1)$ ϵ -short orbits. Suppose not, then there exists a $w \in \mathcal{C}_\mathcal{F}^{sing}$ and a subsequence, still indexed by ν , and sequence of numbers

$$s_{1,\nu} < s_{2,\nu} < \dots < s_{N+2,\nu} \quad (5.4)$$

such that for each j , $\tilde{p}_\nu(s_{j,\nu} + \cdot)$ converges in C_{loc}^∞ to an ϵ -short solution $\tilde{z}_j := (z_j, \xi_j)$ whose orbit is contained in $B_r(w)$. Moreover, there exists $j \in \{1, \dots, N+1\}$ such that $L_\infty(\nu, [s_{j,\nu}, s_{j+1,\nu}])$ contains no orbits that are not ϵ -short. Then we take $[t_{j,\nu}, t_{j+1,\nu}] \subset [s_{j,\nu}, s_{j+1,\nu}]$ such that

$$\begin{aligned} \lim_{\nu \rightarrow \infty} \tilde{p}_\nu(t_{j,\nu}) &= \lim_{t \rightarrow +\infty} \tilde{z}_j(t) := a_+ \in \mathcal{C}_\mathcal{F}, \\ \lim_{\nu \rightarrow \infty} \tilde{p}_\nu(t_{j+1,\nu}) &= \lim_{t \rightarrow -\infty} \tilde{z}_{j+1}(t) := b_- \in \mathcal{C}_\mathcal{F} \end{aligned}$$

and denote $\lim_{t \rightarrow -\infty} \tilde{z}_j(t) = a_-$, $\lim_{t \rightarrow +\infty} \tilde{z}_{j+1}(t) = b_+$. Suppose $L_\infty(\nu, [t_{j,\nu}, t_{j+1,\nu}])$ contains some other ϵ -short orbit contained in $B_r(w)$. It can contain only finitely many, because those orbits must start from the arc between a_- and b_- and ends at the arc between a_+ and b_+ ; their energy therefore has a nonzero lower bound. Hence, by taking a subsequence if necessary, and choosing subintervals of $[t_{j,\nu}, t_{j+1,\nu}]$, we may assume that $T = L_\infty(\nu, [t_{j,\nu}, t_{j+1,\nu}])$ contains no other ϵ -short orbit that is also contained in $B_r(w)$. This implies that

$$T \subset \mathcal{C}_\mathcal{F} \cup \bigcup_{v \in \mathcal{C}_\mathcal{F}^{sing}, v \neq w} B_r(v).$$

Indeed, T is contained in a connected component of the latter set.

Now by the local behavior of $\mathcal{F}|_{\mathcal{C}_{\mathcal{F}}}$ and the monotonicity requirement in Lemma 27, we see that $w \notin T$. And for the same reason, one component of $\mathcal{C}_{\mathcal{F}} \cap (B_{2r}(w) \setminus B_r(w))$ is disjoint from T . This contradicts the connectedness of T . \square

5.3.3. Identify all slow orbits

One can not go from p to q only through fast orbits because $\eta_p \neq \eta_q$ according to our assumption (A12). Hence in the limit slow orbits appear. The identification of them is easy intuitively, because $\mathcal{C}_{\mathcal{F}}$ is 1-dimensional and our limit object should be connected.

Proposition 29. *Assume the sequence \tilde{p}_ν and $\{t_{i,\nu}\}_{i=1,\dots,n}$ satisfy the conditions of the last proposition. Let $t_{0,\nu} = -\infty$ and $t_{n+1,\nu} = +\infty$. Then there exists a subsequence (still indexed by ν), and slow orbits $\gamma_i \subset \mathcal{C}_{\mathcal{F}}$, $i = 0, \dots, n$, satisfying the following.*

- (I) *For each $i = 1, \dots, n$, γ_i is from $\tilde{y}_i(+\infty)$ to $\tilde{y}_{i+1}(-\infty)$.*
- (II) *If $p \in \text{Crit}^+(\mathcal{F})$, then γ_0 is the single point p ; if $q \in \text{Crit}^-(\mathcal{F})$, then γ_n is the single point q .*
- (III) *For each $i = 1, \dots, n-1$, there exist a sequence of intervals $(a_{i,\nu}, b_{i,\nu}) \subset (t_{i-1,\nu}, t_{i,\nu})$ such that $\lim_{\nu \rightarrow \infty} |a_{i,\nu} - t_{i-1,\nu}| = \lim_{\nu \rightarrow \infty} |b_{i,\nu} - a_{i,\nu}| = \lim_{\nu \rightarrow \infty} |t_{i,\nu} - b_{i,\nu}| = \infty$ and $\tilde{p}_\nu([a_{i,\nu}, b_{i,\nu}])$ converges to γ_i in Hausdorff topology.*

Proof. By the previous proposition, we can find sequences of intervals $I_{i,\nu} = [a_{i,\nu}, b_{i,\nu}] \subset \mathbb{R}$, $i = 1, \dots, n-1$ such that for each ν , $I_{i,\nu} \cap I_{j,\nu} = \emptyset$, for $i \neq j$, $b_{i,\nu} - a_{i,\nu} \rightarrow \infty$, and

$$\lim_{\nu \rightarrow \infty} \tilde{p}_\nu(a_{i,\nu}) = \tilde{y}_i(+\infty), \quad \lim_{\nu \rightarrow \infty} \tilde{p}_\nu(b_{i,\nu}) = \tilde{y}_{i+1}(-\infty).$$

Take $I_{0,\nu} = (-\infty, a_{0,\nu}]$, $I_{n,\nu} = [b_{n,\nu}, +\infty)$ such that

$$\lim_{\nu \rightarrow \infty} \tilde{p}_\nu(a_{0,\nu}) = \tilde{y}_1(-\infty), \quad \lim_{\nu \rightarrow \infty} \tilde{p}_\nu(b_{n,\nu}) = \tilde{y}_n(+\infty).$$

1. For each $i \in \{1, \dots, n\}$, set $y_{i,\pm} = \tilde{y}_i(\pm\infty)$, and set $y_{0,+} = p$, $y_{n+1,-} = q$. For $i = 0, \dots, n$, let $\Gamma_i \subset \mathcal{C}_{\mathcal{F}}$ be the union of all slow orbits from $y_{i,+}$ to $y_{i+1,-}$. Since $\mathcal{C}_{\mathcal{F}}$ is a 1-dimensional manifold, Γ_i is either a single point, or is one nontrivial slow orbit, or is the union of two nontrivial slow orbits. We would like to prove that $L_\infty(\nu, I_{i,\nu}) \subset \Gamma_i$. If Γ_i is the union of two slow orbits, then Γ_i is a connected component of $\mathcal{C}_{\mathcal{F}}$. By the connectedness of $L_\infty(\nu, I_{i,\nu})$ the claim is true. Suppose Γ_i is simply-connected and $L_\infty(\nu, I_{i,\nu})$ is also simply-connected. Suppose in this case the claim is not true, then there exists a subsequence (still indexed by ν) and $t_\nu \in I_{i,\nu}$ such that $\lim_{\nu \rightarrow \infty} \tilde{p}_\nu(t_\nu) = z \in L_\infty(\nu, I_{i,\nu}) \setminus \Gamma_i$ and

$$\mathcal{F}(y_{i,+}) > \mathcal{F}(z) > \mathcal{F}(y_{i+1,-}). \quad (5.5)$$

But since $L_\infty(\nu, [t_\nu, b_{i,\nu}])$ and $L_\infty(\nu, [a_{i,\nu}, t_\nu])$ are connected, either the former contains $y_{i,+}$ or the latter contains $y_{i+1,-}$, either of which contradicts Lemma 27.

We claim that it is impossible that Γ_i is homeomorphic to a closed interval and $L_\infty(\nu, I_{i,\nu})$ is homeomorphic to a circle $C \subset \mathcal{C}_\mathcal{F}$. Suppose not, then Lemma 27 implies that $y_{i,+}$ and $y_{i+1,-}$ must be the absolute maximum and minimum of the function \mathcal{F} restricted to this component C . The critical points of $\mathcal{F}|_C$ can be put in a cyclic order and, adjacent to $y_{i,+}$, the two local minimum are $y_{i+1,-}$ and some $z \in C \setminus \Gamma_i$. After taking a subsequence, we can find $t_\nu \in I_{i,\nu}$ such that $\lim_{\nu \rightarrow \infty} \tilde{p}_\nu(t_\nu) = z$. Consider the interval $J_{i,\nu} = [t_\nu, b_{i,\nu}]$. Then $L_\infty(\nu, J_{i,\nu})$ is homeomorphic to a closed interval and doesn't contain $y_{i,+}$. But this is impossible because then it must contain another local maximum z' between z and $y_{i+1,-}$ to keep the connectedness of $L_\infty(\nu, J_{i,\nu})$, while $\mathcal{F}(z') \notin [\mathcal{F}(z), \mathcal{F}(y_{i+1,0})]$, which contradicts with Lemma 27.

2. We can identify one slow orbit contained in Γ_i . More precisely, we would like to show that, there exists $\gamma_i \subset \Gamma_i$ a slow orbit from $y_{i,+}$ to $y_{i+1,-}$ and a subsequence ν' such that $L_\infty(\nu', I_{i,\nu'}) = \gamma_i$. Assume it is impossible, then $\Gamma_i = \gamma_i^1 \cup \gamma_i^2$, the union of two different (nonconstant) slow orbits, and we can assume that there exists a subsequence ν' such that for *any* further subsequence $\{\nu''\} \subset \{\nu'\}$, $\gamma_i^1 \subset L_\infty(\nu'', I_{i,\nu''})$. Now suppose there exists $\tilde{z} \in L_\infty(\nu', I_{i,\nu'}) \cap (\Gamma_i \setminus \gamma_i^1)$. Then there exists a subsequence ν'' , and $t_{\nu''}^-, t_{\nu''}, t_{\nu''}^+ \in I_{i,\nu''}$, $t_{\nu''}^- < t_{\nu''} < t_{\nu''}^+$, such that

$$\lim_{\nu'' \rightarrow \infty} \tilde{p}_{\nu''}(t_{\nu''}) = \tilde{z}, \quad \lim_{\nu'' \rightarrow \infty} \tilde{p}_{\nu''}(t_{\nu''}^\pm) = \tilde{z}^\pm \in \text{Int}(\gamma_i^1)$$

and $\mathcal{F}(y_{i,+}) > \mathcal{F}(\tilde{z}^-) > \mathcal{F}(\tilde{z}) > \mathcal{F}(\tilde{z}^+) > \mathcal{F}(y_{i+1,-})$.

Since

$$(\Gamma_i \setminus \gamma_i^1) \cap \mathcal{F}^{-1}([\mathcal{F}(\tilde{z}^-), \mathcal{F}(\tilde{z}^+)])$$

has a nonzero distance from γ_1 , we see that $L_\infty(\nu'', [t_{\nu''}^-, t_{\nu''}^+])$ is disconnected, which contradicts Lemma 26. Hence we have $L_\infty(\nu', I_{i,\nu'}) = \gamma_i^1$.

3. In summary, there exists a slow orbit γ_i from $y_{i,+}$ to $y_{i+1,-}$, and a subsequence (still indexed by ν), such that $L_\infty(\nu, I_{i,\nu}) = \gamma_i$. By the connectedness of $L_\infty(\nu, I_{i,\nu})$, this implies that $\tilde{p}_\nu(I_{i,\nu})$ converges to γ_i in Hausdorff topology. Conclusion (II) of the proposition follows from the fact that \mathcal{F} is decreasing along γ_0 and γ_n , and p is attracting and q is repelling in this case.

□

This completes the proof of the compactness theorem.

5.4. Gluing

In this subsection we prove the following theorem.

Theorem 30. Suppose $p \in \text{Crit}_k(\mathcal{F})$, $q \in \text{Crit}_{k-1}(\mathcal{F})$. Then, there exists $\epsilon_0 > 0$, such that for all $\lambda \in (0, \epsilon_0]$, there exists a bijection

$$\Phi^\lambda : \mathcal{N}^0(p, q) \rightarrow \mathcal{M}^\lambda(p, q) \quad (5.6)$$

such that for any $\mathcal{X} \in \tilde{\mathcal{N}}^0(p, q)$, there exist representatives $\tilde{p}^\lambda \in \tilde{\mathcal{M}}^\lambda(p, q)$ of $\Phi^\lambda([\mathcal{X}])$ such that as $\lambda \rightarrow 0$, $\{\tilde{p}^\lambda\}$ converges to \mathcal{X} in the sense of Definition 24.

Proof. Let $p \in \text{Crit}_k(\mathcal{F})$, $q \in \text{Crit}_{k-1}(\mathcal{F})$. Theorem 25 implies that for $\lambda > 0$ small, any orbit in $\mathcal{M}^\lambda(p, q)$ is close to a fast-slow orbit in $\mathcal{N}^0(p, q)$. We must show that for any fast-slow orbit from p to q , if $\lambda > 0$ is small, then there exists a unique orbit from p to q that lies near the given fast-slow orbit.

The proof consists of two parts. In the first part, we construct the gluing map Φ^λ , by using the exchange lemma (Lemma 31) and general exchange lemma (Lemma 32). In the second part, we show that Φ^λ is a bijection.

We remark that the exchange lemma and general exchange lemma are for C^r differential equations, and some regularity is lost in the derivative of the constructed solution with respect to ϵ . In our situation the differential equations are of class C^∞ , so of course each solution is of class C^∞ .

Now we start to construct the gluing map. We will frequently use Proposition 21 without citing it, and we will use the notation $W^u(p, \lambda)$, etc., from Subsection 2.3.

Let $\mathcal{X} = (p_0, \sigma_1, p_1, \dots, p_{n-1}, \sigma_n, p_n)$ be a fast-slow solution from p to q . For small $\lambda > 0$, we will follow $W^u(p, \lambda)$ around the fast-slow solution until it meets $W^s(q, \lambda)$ transversally. Any orbit in $\mathcal{M}^\lambda(p, q)$ that is near the given fast-slow solution must be in the portion of $W^u(p, \lambda)$ that we follow; uniqueness is a consequence of the transversality at the end of the proof (Step 8).

We have $p \in \text{Crit}_k^+(\mathcal{F}) \cup \text{Crit}_k^-(\mathcal{F})$ and $q \in \text{Crit}_{k-1}^+(\mathcal{F}) \cup \text{Crit}_{k-1}^-(\mathcal{F})$.

Step 1. Suppose $p = (x_p, \eta_p) \in \text{Crit}_k^+(\mathcal{F})$. Then $\text{index}(x_p, f_{\eta_p}) = k$, and σ_1 is a fast solution from p to $p_1 = (x_{p_1}, \eta_{p_1}) \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{\text{sing}}$, with $\text{index}(x_{p_1}, f_{\eta_{p_1}}) = k-1$. By (A13), $W^u(p, 0)$ meets $W^s(p_1, 0)$ transversally within $M \times \{\eta_p\}$ along σ_1 . Then $W^u(p, 0)$ meets $W^s(\sigma_2, 0)$ transversally in $M \times \mathbb{R}$ along σ_1 . (More precisely, σ_2 should be replaced by a compact portion of the complete orbit corresponding to σ_2 ; we shall use this sort of abuse of notation throughout this section.) The dimension of the intersection is $\dim W^u(p, 0) + \dim W^s(\sigma_2, 0) - (n+1) = k + (n - (k-1) + 1) - (n+1) = 1$.

Thus σ_1 is isolated in the intersection. The next orbit σ_2 is slow, and we have (i) $p_2 \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{\text{sing}}$, or (ii) $p_2 \in \mathcal{C}_{\mathcal{F}}^{\text{sing}}$. In both cases, $\text{index}(x_{p_2}, f_{\eta_{p_2}}) = k-1$.

Step 2. After Step 1, in case (i), the exchange lemma [1] implies that for $\lambda > 0$ small, $W^u(p, \lambda)$, followed along the flow, becomes close to $W^u(\sigma_2, 0)$ near p_2 . By *close* we mean close in the C^s -topology for some large s , which decreases in the course of the proof; see Appendix A. Notice that the dimension of $W^u(\sigma_2, 0)$ is $(k-1) + 1 = k$ as it should be.

The next orbit σ_3 is fast, and we have (i) $p_3 \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{\text{sing}}$, or (ii) $p_3 \in \mathcal{C}_{\mathcal{F}}^{\text{sing}}$. In case (i), $\text{index}(x_{p_3}, f_{\eta_{p_3}}) = k-1$; in case (ii), $\text{index}(x_{p_3}, f_{\eta_{p_3}}) = k-2$.

Step 3. After Step 1, in case (ii), let N_0 denote the center manifold of (1.6)–(1.7) for $\lambda = 0$ at p_2 , which has dimension 2. The choice of center manifold is not unique; we may choose it to include the start of σ_3 (see Appendix B). By the exchange lemma, for $\lambda > 0$ small, $W^u(p, \lambda)$, followed along the flow, becomes close to $W^u(\sigma_2, 0)$ where σ_2 enters N_0 . Of course $W^u(\sigma_2, 0)$ (dimension k) is transverse to $W^s(N_0, 0)$ (dimension $(n - (k - 1) - 1) + 2 = n - k + 2$); $n - (k - 1) - 1$ is the number of negative eigenvalues at p_2 . The dimension of the intersection is 1. For small $\lambda > 0$, N_0 perturbs to a normally hyperbolic invariant manifold N_λ , and $W^u(p, \lambda)$ is transverse to $W^s(N_\lambda, \lambda)$.

The solution in $W^u(p, \lambda) \cap W^s(N_\lambda, \lambda)$ approaches a solution $p_2^\lambda(t)$ in N_λ that is initially near $\sigma_2 \cap N_0$. The system restricted to N_λ , $\lambda \geq 0$, has been analyzed in [26]; see Appendix A. The result is that $p_2^\lambda(t)$ leaves N_λ close to $\sigma_3 \cap N_0$. Then the general exchange lemma implies that $W^u(p, \lambda)$ becomes close to the restriction of $W^u(N_0, 0)$ to $\sigma_3 \cap N_0$, i.e., $W^u(p_2, 0)$, as $W^u(p, \lambda)$ exits a neighborhood of N_λ . $W^u(p_2, 0)$ has dimension $(k - 1) + 1 = k$ as it should. The orbit σ_3 is fast, $p_3 \in \mathcal{C}_\mathcal{F} \setminus \mathcal{C}_\mathcal{F}^{sing}$, and $\text{index}(x_{p_3}, f_{\eta_{p_3}}) = k - 1$.

Step 4. After Step 2, in case (i), $W^u(p, \lambda)$ is close to $W^u(\sigma_2, 0)$ near p_2 , and by (A10), $W^u(\sigma_2, 0)$ is transverse to $W^s(\sigma_4, 0)$ along σ_3 . Continuation from here is like continuation after Step 1, described in Steps 2–3.

Step 5. After Step 2, in case (ii), we note that by (A13), $W^u(p_2, 0)$ (dimension $k - 1$) is transverse to $W^s(p_3, 0)$ (dimension $n - (k - 2)$) within $M \times \{\eta_{p_3}\}$ along σ_3 . Let N_0 denote the center manifold of (1.6)–(1.7) for $\lambda = 0$ at p_3 , which has dimension 2; we choose it to include the end of σ_3 (see Appendix B). Then $W^u(\sigma_2, 0)$ (dimension k) is transverse to $W^s(N_0, 0)$ (dimension $(n - (k - 2) - 1) + 2 = n - k + 3$) along σ_3 . The intersection is 2-dimensional and consists of solutions that track an open set of solutions in N_0 around σ_3 .

Step 6. After Step 3, we note that by (A13), $W^u(p_2, 0)$ (dimension k) meets $W^s(p_3, 0)$ (dimension $n - (k - 1)$) transversally within $M \times \{\eta_{p_3}\}$ along σ_3 . As in Step 1, $W^u(p_2, 0)$ meets $W^s(\sigma_4, 0)$ transversally in $M \times \mathbb{R}$ along σ_3 , and $W^u(p, \lambda)$ is close to $W^u(p_2, 0)$ near σ_3 . The next orbit σ_4 is a slow solution from p_3 to (i) $p_4 \in \mathcal{C}_\mathcal{F} \setminus \mathcal{C}_\mathcal{F}^{sing}$, or (ii) $p_4 \in \mathcal{C}_\mathcal{F}^{sing}$. In both cases, $\text{index}(x_{p_4}, f_{\eta_{p_4}}) = k - 1$. Continuation from here is like continuation after Step 1, described in Steps 2–3.

Step 7. After Step 5, for small $\lambda > 0$, the tracked orbits include an open set U_λ in N_λ , the perturbation of N_0 , that lies above an open set U_0 in N_0 that contains $\sigma_4 \cap N_0$ in its interior. The general exchange lemma implies that $W^u(p, \lambda)$ is close to the restriction of $W^u(N_0, 0)$ to U_0 , i.e., to $W^u(\sigma_4 \cap N_0)$, as it exits a neighborhood of N_λ . Continuation from here is like continuation after Step 2, described in Steps 4–5.

Step 8. Continuation proceeds using analogs of the steps previously described until $W^u(p)$ arrives near p_{n-1} .

If $q \in \text{Crit}_k^+(\mathcal{F})$, then σ_n is slow, and $\text{index}(x_{p_{n-1}}, f_{\eta_{p_{n-1}}}) = \text{index}(x_q, f_{\eta_q}) = k - 1$. $W^u(p, \lambda)$ arrives near p_n along the slow orbit σ_n , close to $W^u(\sigma_n, 0)$, which has dimension k . $W^u(\sigma_n, 0)$ is transverse to $W^s(\sigma_n, 0)$, with dimension $n - (k - 1) + 1 = n - k + 2$. Therefore $W^u(p, \lambda)$ is transverse to $W^s(q, \lambda)$, which is close to $W^s(\sigma_n, 0)$. The intersection has dimension 1 and gives the solution

from p to q .

If $q \in \text{Crit}_k^-(\mathcal{F})$, then σ_n is a fast orbit from $p_{n-1} \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$ to $q \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$; $\text{index}(x_{p_{n-1}}, f_{\eta_{p_{n-1}}}) = k-1$ and $\text{index}(x_q, f_{\eta_q}) = k-2$. By (A11), $W^u(p_{n-1}, 0)$ meets $W^s(q, 0)$ transversally in $M \times \{\eta_q\}$ along σ_n . Therefore $W^u(\sigma_n, 0)$ meets $W^s(q, 0)$ transversally in $M \times \mathbb{R}$ along σ_n . Since $W^u(p, \lambda)$ is close to $W^u(\sigma_n, 0)$, $W^u(p, \lambda)$ meets $W^s(q, \lambda)$ transversally in $M \times \mathbb{R}$ near σ_n . The intersection has dimension one and gives the solution from p to q .

Step 9. Suppose $p = (x_p, \eta_p) \in \text{Crit}_k^-(\mathcal{F})$. Then $\text{index}(x_p, f_{\eta_p}) = k-1$, and σ_1 is a slow solution from p to $p_1 = (x_{p_1}, \eta_{p_1})$. Either (i) $p_1 \in \mathcal{C}_{\mathcal{F}} \setminus \mathcal{C}_{\mathcal{F}}^{sing}$, or (ii) $p_1 \in \mathcal{C}_{\mathcal{F}}^{sing}$. In either case, $\text{index}(x_{p_1}, f_{\eta_{p_1}}) = k-1$.

In case (i), $W^u(p, \lambda)$ is close to $W^u(\sigma_1, 0)$. Step 3 above describes how to continue from p_1 . Case (ii) is left to the reader.

Now we show that the gluing map just constructed is a bijection. We first note that there exist small $\delta > 0$ and $\lambda_0 > 0$ such that for all $\lambda \in (0, \lambda_0)$, the orbit we constructed is the unique one in $\mathcal{M}^\lambda(p, q)$ such that its Hausdorff distance from the fast-slow orbit $[\mathcal{X}]$ is less than δ . This is a general fact about exchange lemma constructions; see for example [27], p. 1021.

Thus the gluing map is a bijection provided it is surjective. We claim that there exists $\lambda_0 > 0$ such that the gluing map is surjective for all $\lambda \in (0, \lambda_0)$. If not, then there exist a sequence $\lambda_i \rightarrow 0$ and a sequence of orbits $[\tilde{x}_i] \in \mathcal{M}^{\lambda_i}(p, q)$ such that for each i , $[\tilde{x}_i]$ is not in the image of Φ^{λ_i} . Then, by the compactness theorem (Theorem 25), without loss of generality, $[\tilde{x}_i]$ converges to a fast-slow orbit $[\mathcal{Y}]$ in the sense of Definition 24. In particular, this sequence converges in the Hausdorff topology. Then for large i , $\Phi^{\lambda_i}([\mathcal{Y}])$, which is the unique orbit in $\mathcal{M}^{\lambda_i}(p, q)$ in a small Hausdorff neighborhood of $[\mathcal{Y}]$, must be $[\tilde{x}_i]$. This contradiction proves the claim. \square

6. Acknowledgements

The first author's work was partially supported by NSF under award DMS-1211707. The second author would like first to thank his advisor Professor Gang Tian, for his help and encouragement, guidance and corrections. He also would like to thank Urs Frauenfelder for many helpful discussions, to thank Robert Lipshitz and Dietmar Salamon for their interest in this work, to thank Hongbin Sun for answering topology questions. At the early stage of this project, the second author was funded by Professor Helmut Hofer for the Extramural Summer Research Project of Princeton University in Summer 2011, and he would like to express his gratitude for Professor Hofer's generosity.

Appendix A. Exchange lemmas

Let Z be a manifold, let $\dot{p} = f(p, \epsilon) = f_\epsilon(p)$ be a differential equation on Z with parameter ϵ (i.e., each f_ϵ is a section of TZ), let ϕ_t^ϵ be the flow, and let N_0 be a compact submanifold of Z that is invariant under the flow of $\dot{p} = f_0(p)$. N_0 is called *normally hyperbolic* if there is a splitting of $TZ|_{N_0}$,

$TZ|N_0 = S \oplus U \oplus TN_0$, such that under $D\phi_t^0$, all vectors in S shrink at a faster exponential rate than any vector in TN_0 , and under $D\phi_{-t}$, all vectors in U shrink at a faster exponential rate than any vector in TN_0 . (There are less restrictive definitions, but this one suffices for our purposes.)

Normally hyperbolic invariant manifolds have stable and unstable manifolds with flow-preserved fibrations, and the whole structure persists under perturbation. This structure is most easily described in local coordinates.

Let us assume that $\dim N_0 = m$, $\dim Z = n+m$, and fibers of S (respectively U) have dimension k (respectively l), with $k+l = n$. Near a point of N_0 one can choose coordinates $p = \Phi(x, y, z, \epsilon)$, $(x, y, z, \epsilon) \in \Omega_1 \times \Omega_2 \times (-\epsilon_0, \epsilon_0)$, Ω_1 open subset of $\mathbb{R}^k \times \mathbb{R}^l$, Ω_2 an open subset of \mathbb{R}^m , such that, for small ϵ , $\dot{p} = f(p, \epsilon)$ becomes

$$\dot{x} = A(x, y, z, \epsilon)x, \quad (\text{A.1})$$

$$\dot{y} = B(x, y, z, \epsilon)y, \quad (\text{A.2})$$

$$\dot{z} = h(z, \epsilon) + x^T C(x, y, z, \epsilon)y; \quad (\text{A.3})$$

the matrices A , B , and C are $k \times k$, $l \times l$, and $k \times l$ respectively.

We list some facts and terminology.

1. If $\dot{p} = f(p, \epsilon)$ is C^{r+3} , the coordinate change can be chosen so that the new system is C^{r+1} .
2. For each ϵ , the subspaces $y = 0$, $x = 0$, and their intersection are invariant. For fixed ϵ , the set $\{(x, y, z) \mid x = 0 \text{ and } y = 0\}$ (dimension m) corresponds to part of a normally hyperbolic invariant manifold N_ϵ ; the set $y = 0$ (dimension $m+k$) corresponds to part of the stable manifold of N_ϵ , $W^s(N_\epsilon)$; and the set $x = 0$ (dimension $m+l$) corresponds to part of the unstable manifold of N_ϵ , $W^u(N_\epsilon)$.
3. If $(x(t), 0, z(t))$ is a solution in $W^s(N_\epsilon)$, then $(0, 0, z(t))$ is a solution in N_ϵ ; and if $(0, y(t), z(t))$ is a solution in $W^u(N_\epsilon)$, then $(0, 0, z(t))$ is again a solution in N_ϵ . Each solution in $W^s(N_\epsilon)$ (respectively $W^u(N_\epsilon)$) approaches exponentially a solution in N_ϵ as time increases (respectively decreases).
4. Given a point $p = (0, 0, z_0)$ in N_ϵ , the *stable* (respectively *unstable*) *fiber* of p is the set of all points $(x, 0, z_0)$ (dimension k) (respectively $(0, y, z_0)$ (dimension l)). For each t , the time- t map of the flow takes fibers to fibers; in this sense the fibration is flow-invariant. Solutions that start in the stable (respectively) unstable fiber of p approach the solution that starts at p exponentially as t increases (respectively decreases).
5. Given an $P_\epsilon \subset N_\epsilon$, we shall refer to the union of the stable (respectively unstable) fibers of points in P_ϵ as $W^s(N_\epsilon)$ (respectively $W^u(N_\epsilon)$) *restricted to* P_ϵ . If P_ϵ is invariant, $W^s(N_\epsilon)$ restricted to P_ϵ , for example, may be smaller than $W^s(P_\epsilon)$, since the latter may include solutions in N_ϵ that approach P_ϵ at a slower exponential rate, together with points in their stable fibers.

Suppose N_0 is a compact manifold with boundary of equilibria of dimension m , and each equilibrium in N_0 has k eigenvalues with negative real part and l

eigenvalues with positive real part. N_0 does not fit the definition of a normally hyperbolic invariant manifold that we have given, but a coordinate system as above still exists, except that Ω_2 may have boundary and N_ϵ may be only locally invariant (i.e., solutions may leave through the boundary). We shall abuse terminology and refer to the N_ϵ as normally hyperbolic invariant manifolds in this situation.

Suppose we wish to follow an $(l + 1)$ -dimensional manifold of solutions M_ϵ as it passes near a normally hyperbolic invariant manifold N_ϵ in a manifold Z , where N_0 is a normally hyperbolic invariant manifold of equilibria. We choose coordinates so that the system is (A.1)–(A.3), with $h(z, \epsilon) = \epsilon \tilde{h}(z, \epsilon)$. We assume that near the point $(x^*, 0, z^*)$, the M_ϵ , $\epsilon \geq 0$, fit together to form a C^{r+1} manifold with boundary in $Z \times \mathbb{R}$; and we assume that M_0 meets $W^s(N_0)$ transversally in the solution through $(x^*, 0, z^*)$, which approaches the equilibrium $(0, 0, z^*)$ as $t \rightarrow \infty$. The system restricted to N_ϵ has the form $\dot{z} = \epsilon \tilde{h}(z, \epsilon)$; we assume $\tilde{h}(z^*, 0) \neq 0$. Let ψ_t denote the flow of $\dot{z} = \tilde{h}(z, 0)$, and let $z^\dagger = \psi_T(z^*)$ for some $T > 0$. Let $I = \{(\psi_t(z^*) : T - \delta < t < T + \delta)\}$, a small interval around z^\dagger in the orbit of z^* for $\dot{z} = \tilde{h}(z, 0)$. Let $J = \{0\} \times \{0\} \times I$. Let V be a small open set around $(0, 0, z^\dagger)$ in $W^u(N_0)$ restricted to J , which has dimension $l + 1$. Then we have:

Theorem 31 (Exchange Lemma). *For small $\epsilon > 0$, parts of M_ϵ fit together with V to form a C^r manifold in $Z \times \mathbb{R}$.*

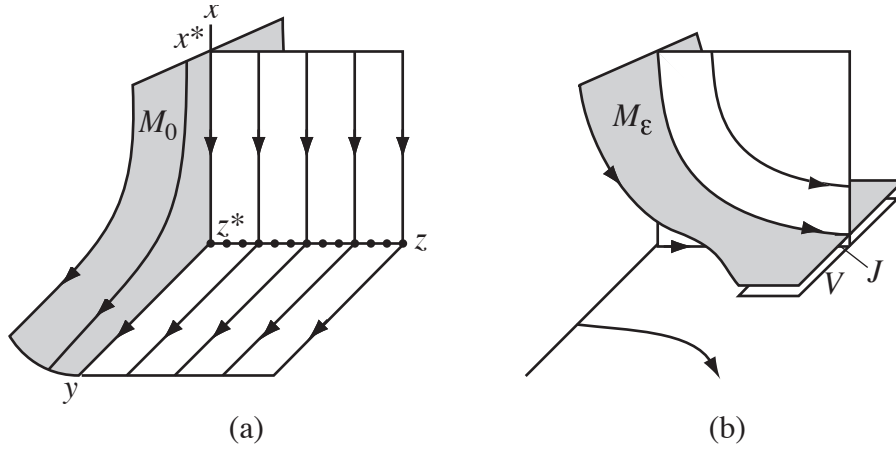


Figure A.3: The exchange lemma with $k = l = m = 1$: (a) $\epsilon = 0$, (b) $\epsilon > 0$.

See Figure A.3. Notice that for small $\epsilon > 0$, M_ϵ meets $W^s(N_\epsilon)$ transversally in a solution that tracks a solution in N_ϵ that starts at a point $(0, 0, z(\epsilon))$, with $z(\epsilon)$ near z^* . For $\epsilon > 0$, the orbit of $\dot{z} = \epsilon \tilde{h}(z_\epsilon)$ through z_ϵ equals the orbit of $\dot{z} = \tilde{h}(z, \epsilon)$ through $z(\epsilon)$. This orbit is close to the orbit of $\dot{z} = \tilde{h}(z, 0)$ through z^* . Thus we find that portions of the orbits tracked in N_ϵ for $\epsilon > 0$ limit on

a curve J in N_0 that is not an orbit of the system (A.1)–(A.3) on N_0 (these orbits are just points). Nevertheless the M_ϵ for $\epsilon > 0$ become close to $W^u(N_0)$ restricted to J .

Now suppose we wish to track an $(l + s + 1)$ -dimensional manifold of solutions M_ϵ , with $0 \leq s \leq m - 1$, as it passes near a normally hyperbolic invariant manifold N_ϵ in a manifold Z , where N_0 is a normally hyperbolic invariant manifold but does not consist of equilibria. (This occurs, for example, Step 3 of the proof of Theorem 30.) We again choose coordinates so that the system is (A.1)–(A.3). We choose a cross-section \tilde{M}_ϵ to the flow within M_ϵ , of dimension $l + s$ (so that M_ϵ is the union of orbits that start in \tilde{M}_ϵ), and we assume that \tilde{M}_0 meets $W^s(N_0)$ transversally at $(x^*, 0, z^*)$. For small $\epsilon \geq 0$, \tilde{M}_ϵ meets $W^s(N_\epsilon)$ in a manifold Q_ϵ of dimension s ; we assume that Q_ϵ projects regularly along the stable fibration to a submanifold P_ϵ of N_ϵ of dimension s . We also assume that for small $\epsilon \geq 0$, the vector field $(0, 0, h(x, \epsilon))$ is not tangent to P_ϵ . (Of course these assumptions follow from the corresponding ones at $\epsilon = 0$.) Finally, we assume that for $\epsilon > 0$, following P_ϵ along the flow for time $\mathcal{O}(\frac{1}{\epsilon})$ produces a submanifold P_ϵ^* of dimension $s + 1$ of N_ϵ , and the manifolds P_ϵ^* fit together with a submanifold P_0^* of N_0 , of dimension $s + 1$, to form a C^{r+1} manifold in $Z \times \mathbb{R}$. This assumption typically requires that some solution of the system restricted to N_0 that starts in P_0 approaches an equilibrium. We emphasize that, analogous to the usual exchange lemma, P_0^* is not the result of following P_0 along the flow for $\epsilon = 0$. Let V be a small open neighborhood of P_0^* in $W^u(N_0)$ restricted to P_0^* , which has dimension $l + s + 1$.

Theorem 32 (General Exchange Lemma). *For small $\epsilon > 0$, parts of M_ϵ fit together with V to form a C^r manifold in $Z \times \mathbb{R}$.*

Some technical hypotheses that are not relevant to the present paper have been omitted. We actually need a small generalization of Theorems 31 and 32, whose proofs are essentially the same.

We will use C^r -topology to measure submanifolds. If $S \subset X$ is a smooth submanifold, we say another C^r -submanifold S' (of the same dimension) is C^r -close to S , if with some smooth identification of the normal bundle of S with a tubular neighborhood of S , S' can be identified with a C^r -small section of the normal bundle.

In Theorem 31, replace the assumption that near the point $(x^*, 0, z^*)$, the M_ϵ , $\epsilon \geq 0$, fit together to form a C^{r+1} manifold with boundary, with the assumption that $M_\epsilon \rightarrow M_0$ in the C^{r+1} -topology. The conclusion becomes that parts of M_ϵ converge to V in the C^r -topology.

In Theorem 32, make the replacement in the assumptions just mentioned, and replace the assumption that the P_ϵ^* , $\epsilon \geq 0$, fit together to form a C^{r+1} manifold with the assumption that $P_\epsilon^* \rightarrow P_0^*$ in the C^{r+1} -topology. The conclusion again becomes that parts of M_ϵ converge to V in the C^r -topology.

These generalizations are required for the following reason. Consider the

fast-slow system

$$\dot{z}_1 = -z_2 + z_1^2, \quad (\text{A.4})$$

$$\dot{z}_2 = -\epsilon, \quad (\text{A.5})$$

which is related to (4.17)–(4.18). See Figure B.4. For $\epsilon = 0$, any compact portion N_0 of the curve of equilibria $z_1 = -z_2^{\frac{1}{2}}$ is normally hyperbolic (in fact attracting). It perturbs to the normally hyperbolic manifold N_ϵ , on which system reduces to $\dot{y} = -\epsilon$. For small $\epsilon > 0$, a solution in or close to N_ϵ arrives in the region $\delta < z_1 < 2\delta$ ($\delta > 0$) along a curve given by $z_2 = \rho(z_1, \epsilon)$, $\delta < z_1 < 2\delta$; $\rho = \mathcal{O}(\epsilon^{\frac{2}{3}})$. As $\epsilon \rightarrow 0$, this curve, which in examples with greater dimension may be P_ϵ^* , approaches $z_2 = 0$, which in examples may be P_0^* , in the C^s -topology for any s , but does not fit together with $z_2 = 0$ to form a manifold with a high degree of differentiability in $z_1 z_2 \epsilon$ -space.

Appendix B. Choosing the center manifold

Let us consider for concreteness a system in the form (A.1)–(A.3), with N_0 two-dimensional, and the equation on N_ϵ given by (A.4)–(A.5):

$$\dot{x} = A(x, y, z_1, z_2, \epsilon)x, \quad (\text{B.1})$$

$$\dot{y} = B(x, y, z_1, z_2, \epsilon)y, \quad (\text{B.2})$$

$$\dot{z}_1 = -z_2 + z_1^2 + x^T c_1(x, y, z_1, z_2, \epsilon)y, \quad (\text{B.3})$$

$$\dot{z}_2 = -\epsilon + x^T c_2(x, y, z_1, z_2, \epsilon)y, \quad (\text{B.4})$$

$(x, y) \in \mathbb{R}^k \times \mathbb{R}^l$. The system arises by center manifold reduction at the origin, so $A(0, 0, z_1, z_2, 0)$ has eigenvalues with negative real part, and $B(0, 0, z_1, z_2, 0)$ has eigenvalues with positive real part. We wish to follow a manifold M_ϵ of dimension $l+1$ as it passes N_ϵ (i.e., $z_1 z_2$ -space). M_0 meets $W^s(N_0)$ transversally at a point $(x^*, 0, z_1^*, 0)$ with $z_1^* < 0$; hence the intersection includes the semiorbit γ that starts at $(x^*, 0, z_1^*, 0)$, which approaches the origin as $t \rightarrow \infty$. We wish to replace the center manifold N_0 for $\epsilon = 0$ by one that contains γ . To do this, replace all semiorbits in Figure B.4 that start at points $(0, 0, z_1^*, z_2)$, $|z_2| < (z_1^*)^2$, with the semiorbits that start at $(x^*, 0, z_1^*, z_2)$. (The semiorbits that start at $(0, 0, z_1^*, z_2)$ and at $(x^*, 0, z_1^*, z_2)$ both approach $(0, 0, -z_2^{\frac{1}{2}}, z_2)$ as $t \rightarrow \infty$, and the second arrives tangent to $z_1 z_2$ -space.) The result will be a new center manifold, in a smaller neighborhood of the origin, that contains γ . The differentiability class of this manifold will decrease as we move away from the origin. Now N_0 perturbs to new center manifold N_ϵ for $\epsilon > 0$.

References

- [1] C. Jones, Geometric singular perturbation theory, in: R. Johnson (Ed.), Dynamical systems, Vol. 1609 of Lecture Notes in Mathematics, Springer, 1995.

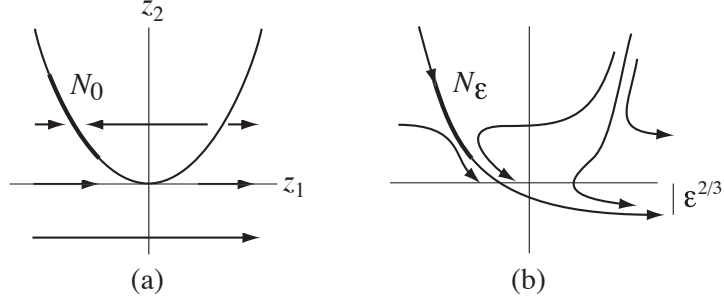


Figure B.4: Flow of (A.4)–(A.5): (a) $\epsilon = 0$, (b) $\epsilon > 0$.

- [2] A. Banyaga, D. Hurtubise, Cascades and perturbed Morse-Bott functions, arXiv: 1110.4609 (2011).
- [3] S. Schecter, Exchange lemmas. II. General exchange lemma, *Journal of Differential Equations* 245 (2) (2008) 411–441.
- [4] E. Witten, Supersymmetry and Morse theory, *Journal of Differential Geometry* 17 (4) (1982) 661–692.
- [5] M. Gromov, Pseudoholomorphic curves in symplectic manifolds, *Inventiones Mathematicae* 82 (2) (1985) 307–347.
- [6] I. Mundet i Riera, Teoría de Yang-Mills-Higgs para fibraciones simpécticas, Ph.D. thesis, Universidad Autónoma de Madrid (1999).
- [7] K. Cieliebak, A. Gaio, D. Salamon, J -holomorphic curves, moment maps, and invariants of Hamiltonian group actions, *International Mathematics Research Notices* 16 (2000) 831–882.
- [8] G. Xu, Gauged Floer homology for Hamiltonian isotopies I: definition of the Floer homology groups, arXiv: 1312.6923 (2013).
- [9] A. Gaio, D. Salamon, Gromov-Witten invariants of symplectic quotients and adiabatic limits, *Journal of Symplectic Geometry* 3 (1) (2005) 55–159.
- [10] E. Witten, The Verlinde algebra and the cohomology of the Grassmannian, in: *Geometry, Topology and Physics*, International Press, Cambridge, MA, 357–422.
- [11] E. Verlinde, Fusion rules and modular transformations in 2d conformal field theory, *Nuclear Physics. B* 300 (3) (1988) 360–376.
- [12] D. Gepner, Fusion rings and geometry, *Communications in Mathematical Physics* 141 (2) (1991) 381–411.

- [13] C. Vafa, Topological mirrors and quantum rings, in: S.-T. Yau (Ed.), *Essays On Mirror Manifolds*, International Press, 1992.
- [14] K. Intriligator, Fusion residues, *Modern Physics Letter A* 6 (1991) 3543–3556.
- [15] S. Agnihotri, Quantum cohomology and the Verlinde algebra, Ph.D. thesis, Oxford University (1995).
- [16] E. González, C. Woodward, Gauged Gromov-Witten invariants for small spheres, to appear in *Mathematische Zeitschrift* (2009).
- [17] J. Milnor, Morse theory, no. 51 in *Annals of Mathematics Studies*, Princeton University Press, 1963.
- [18] M. Schwarz, Morse homology, *Progress in Mathematics*, Birkhauser, 1993.
- [19] U. Frauenfelder, Vortices on the cylinder, *International Mathematics Research Notices* 2006 (2006) 1–34.
- [20] J. Kang, Invariance property of Morse homology on noncompact manifolds, *arXiv:math/1012.5571* (2010).
- [21] A. Floer, Morse theory for Lagrangian intersections, *Journal of Differential Geometry* 28 (1988) 513–547.
- [22] D. McDuff, D. Salamon, *J*-holomorphic curves and symplectic topology, Vol. 52 of *Colloquium Publications*, American Mathematical Society, 2004.
- [23] A. Banyaga, D. Hurtubise, Morse-Bott homology, *Transactions of the American Mathematical Society* 362 (8) (2010) 3997–4043.
- [24] A. Floer, The unregularized gradient flow of the symplectic action, *Communications on Pure and Applied Mathematics* 41 (6) (1988) 775–813.
- [25] U. Frauenfelder, The Arnold-Givental conjecture and moment Floer homology, *International Mathematics Research Notices* 2004 (42) (2004) 2179–2269.
- [26] M. Krupa, P. Szmolyan, Extending geometric singular perturbation theory to nonhyperbolic points-fold and canard points in two dimensions, *SIAM Journal of Mathematical Analysis* 33 (2) (2001) 286–314.
- [27] C. Jones, S.-K. Tin, Generalized exchange lemmas and orbits heteroclinic to invariant manifolds, *Discrete Contin. Dyn. Syst. Ser. S* 2 (4) (2009) 967–1023.