Commun. Math. Phys. 193, 337 – 371 (1998)

Communications in
**Mathematical
Physics**

Homotopy Classes for Stable Connections between Hamiltonian Saddle-Focus Equilibria

W. D. Kalies*, **J. Kwapisz***,**, **R.C.A.M. VanderVorst***

Center for Dynamical Systems and Nonlinear Studies, Georgia Institute of Technology, Atlanta, GA 30332, USA

Received: 10 December 1996 / Accepted: 5 September 1997

Abstract: For a class of Hamiltonian systems in $\mathbb{R}^4$ the set of homoclinic and heteroclinic orbits which connect saddle-focus equilibria is studied using a variational approach. The oscillatory properties of a saddle-focus equilibrium and the variational nature of the problem give rise to connections in many homotopy classes of the configuration plane punctured at the saddle-foci. This variational approach does not require *any* assumptions on the intersections of stable and unstable manifolds, such as transversality. Moreover, these connections are shown to be local minimizers of an associated action functional. This result has applications to spatial pattern formation in a class of fourth-order bistable evolution equations.

1. Introduction

Hamiltonian systems obtained from second-order Lagrangian densities of the form $L = L(u, u', u'')$ are used in many physical models including problems in nonlinear optics, nonlinear elasticity, and mechanics. The Euler-Lagrange equations associated with such densities are fourth-order differential equations, and their solutions are critical points of the Lagrangian functional $J[u] = \int_{\mathbb{R}} L(u, u', u'')\, dt$. In these systems stationary solutions $\widehat{u}$, which satisfy $\partial L/\partial u\,(\widehat{u}, 0, 0) = 0$, can have four hyperbolic complex eigenvalues – two with positive real part and two with negative real part. This type of equilibrium solution is called a **saddle-focus**, and it is well-known that systems with homoclinic or heteroclinic connections between such points can exhibit complicated (chaotic) behavior, cf. [12, 20, 26, 33]. The purpose of this paper is to investigate the structure of the homoclinic and heteroclinic orbits connecting saddle-focus equilibria in a class of fourth-order equations.

The specific Lagrangians that we study are given by

* Partially supported by grants ARO DAAH-0493G0199 and NIST G-06-605.

** Partially supported by Polish KBN Grant 2 P301 01307 Iteracje i Fraktale, II.

$$J[u] = \int_{\mathbb{R}} \left[\frac{\gamma}{2}|u''|^2 + \frac{\beta}{2}|u'|^2 + F(u) \right] dt \quad \text{with } \gamma, \beta > 0. \tag{1.1}$$

The nonlinearity F is assumed to be a double-well potential with two nondegenerate global minima at ± 1 of which the prototypical example is $F(u) = (u^2 - 1)^2/4$. The number of global minima is not crucial, and our analysis extends to potentials with arbitrarily many wells. The Euler-Lagrange equation for (1.1) is

$$\gamma u'''' - \beta u'' + F'(u) = 0 \quad \text{with } \gamma, \beta > 0. \tag{1.2}$$

This equation has been proposed as a generalization of the second-order stationary Allen-Cahn or Fisher-Kolmogorov equation ($\gamma = 0$) and arises in the study of phase transitions in the neighborhood of Lifshitz points [18, 19, 24, 50], see Sect. 8. This extended Fisher-Kolmogorov equation is valid near parameter values where the Ginzburg-Landau formulation becomes degenerate [40]. When $\gamma > \beta^2/4F''(\pm 1)$, the equilibrium points $u = \pm 1$ are saddle-foci, and we are interested in the heteroclinic and homoclinic orbits connecting these two points in the four-dimensional flow generated by (1.2). Note that this ODE generates only a local flow on $\mathbb{R}^4$, and there are solutions which blow-up in finite time.

Before describing the history of this problem, we would like to state three characteristics of our results which differ from much of the previous work. First, the assumptions on the nonlinearity are very mild. In particular we do not require symmetry or analyticity of F, nor do we place any transversality or nondegeneracy conditions on the intersections of the stable and unstable manifolds of ± 1. Second, we produce multitransition solutions of (1.2) with any number of transitions, all of which are *local minimizers* of the action functional (1.1). Finally, these multitransition solutions do not all lie in some small neighborhood of the principal loop in the phase space. In particular the distance between transitions is not required to be large. Our results imply that the dynamics of equations of the form (1.2) with a double-well potential and saddle-focus equilibria are always chaotic and hence never completely integrable.

The methods used in this paper also seem to be applicable to mechanical systems with two degrees of freedom. As for fourth order problems, this would require a nonnegative Lagrangian density and saddle-foci which are global minima. These examples may be the subject of future work.

Finding multitransition and multibump solutions for Hamiltonian systems has become an active field of study in recent years. In this context we mention the work of Buffoni, Coti-Zelati, Ekeland, Rabinowitz, and Séré, [4, 10, 16, 17, 21, 38, 39, 43, 44]. The initial work is due to Séré [43] who finds infinitely many two-bump homoclinic orbits for a general class of nonautonomous, periodically-forced Hamiltonian systems with a subsequent generalization to multibump homoclinics [44]. Coti-Zelati and Rabinowitz [17, 39] consider the problem of finding multibump homoclinic connections for mechanical systems with Lagrangians of the form $L(t, q, q') = \frac{1}{2}|q'|^2 - V(t, q)$, where $q : \mathbb{R} \to \mathbb{R}^n$ and V is a periodically-forced potential. Nondegeneracy conditions are imposed on the primary homoclinic connections in order to construct multibump solutions. These variational results are analogous to those obtained from the study of Poincaré maps of time-periodic systems via Melnikov theory which detects transverse intersections, cf. [23, 31].

Using such techniques from dynamical systems theory, Devaney [20] has shown that *autonomous* Hamiltonian systems in $\mathbb{R}^4$ display horseshoe-like dynamics in a neighborhood of a *transverse* homoclinic or heteroclinic loop connecting saddle-foci. This principal loop is the four-dimensional equivalent of a Shil'nikov orbit [45, 46]. In particular this implies that a countable family of multibump or multitransition solutions exists near the primary loop in the phase space. In general the existence of a primary loop composed of transverse intersections of stable and unstable manifolds is difficult to verify in fourth-order problems. In this paper we prove that for a significant class of autonomous Hamiltonian systems no transversality condition is required to obtain multitransition solutions of (1.2). Multibump or multiple-pulse solutions and their stability have been investigated using dynamical systems techniques in a wider variety of contexts with varying hypotheses on the primary solutions. For references to this literature, the reader is referred to Alexander, Gardner, and Jones [3], Lin [30], Nishura [32], and Sandstede [41, 42].

Multibump homoclinic connections near a Shil'nikov orbit in conservative, autonomous systems in $\mathbb{R}^4$ have been studied variationally by Buffoni and Séré [12]. Their results require an intersection condition on the stable and unstable manifolds which is weaker than transversality and is simpler to check for certain examples. For the system (1.2) this condition has been verified *for the specific potential* $F = (u^2 - 1)^2/4$ and is used in [26] to construct multitransition solutions of

$$\gamma u'''' - \beta u'' + u^3 - u = 0, \tag{1.3}$$

which is often referred to as the (stationary) extended Fisher-Kolmogorov equation. This approach yields solutions which are close to a primary heteroclinic loop, composed of single-transition solutions, by gluing together well-separated copies of these primary solutions. However, checking the intersection condition of Buffoni and Séré [12] can be involved, and it has not been verified for the general problem (1.2).

The EFK-Eq. (1.3) has been extensively studied by Peletier and Troy [33, 34, 35, 36, 37] who show that heteroclinic connections between ± 1 exist for all $\beta, \gamma > 0$. These primary heteroclinic connections have exactly one monotone transition and minimize the action J in a suitable class of functions [37]. Using topological shooting methods they explore the set of bounded solutions of (1.3) when $\gamma > \beta^2/8$ (the saddle-focus case). They prove the existence of a countable family of heteroclinic connections which are qualitatively different from those found in this paper and in [26], as well as various types of periodic and chaotic solutions. For $\gamma < \beta^2/8$ the points ± 1 are saddles (two negative and two positive real eigenvalues), and the primary heteroclinic connection is monotone and unique (up to translations) within the class of odd monotone functions. A similar approach in [28], using a global Poincaré section, extends this result to show that the primary heteroclinic is unique without restriction (see also [48]). The approach taken in this paper will be completely different. Also the method of Gardner and Jones [22] proves that for small values of γ/β^2 the heteroclinics are the result of transverse intersections of the stable and unstable manifolds. Transversality for all $\gamma/\beta^2 \leq \frac{1}{8}$ was proved by Vanden Berg [49]. This is still an open question for large values of γ/β^2.

A related equation

$$u'''' + Pu'' + u - u^2 = 0,$$

which arises in nonlinear elasticity and the theory of shallow water waves, has been extensively studied by Amick, Buffoni, Champneys, and Toland [5, 10, 11, 13, 14] who also develop a shooting method suitable for fourth-order problems of this type. Note that the Lagrangian density is given by $L(u, u', u'') = \frac{1}{2}|u''|^2 - \frac{P}{2}|u'|^2 + \frac{1}{2}u^2 - \frac{1}{3}u^3$ and is not bounded from below as is (1.1). The primary homoclinic connection occurs as a mountain pass critical point, and our methods are not directly applicable. However we believe that many of the same ideas are important for both classes of problems. The parameter P plays the same role as the ratio γ/β^2 in Eq. (1.2), and for $-2 < P < 2$ the stationary point $u = 0$ is a saddle-focus, which leads to a complicated set of multibump homoclinic connections [10, 12, 15]. This equation with u^2 replaced by u^3 is also used in certain optical models [2].

Since the points $u = \pm 1$ are hyperbolic equilibria of (1.2), homoclinic connections are contained in the affine Sobolev spaces $\pm 1 + H^2(\mathbb{R})$. Heteroclinic connections lie in the spaces $\pm\widehat{\chi} + H^2(\mathbb{R})$, where $\widehat{\chi} \in C^\infty(\mathbb{R})$ is a fixed function such that $\widehat{\chi}(t) = -1$ for $t \le -1$ and $\widehat{\chi}(t) = 1$ for $t \ge 1$. In the sequel, $\chi + H^2(\mathbb{R})$ will denote the affine space in which χ is chosen appropriately to be $1, -1, \widehat{\chi}$, or $-\widehat{\chi}$. Functions in all of these spaces have infinite tails which are bi-asymptotic to ± 1 as $t \to \pm\infty$. Disregarding these tails, the solution can make a finite number of transitions between the values -1 and +1 with oscillations around ± 1 between transitions, see Fig. 1.1. Our approach to finding multitransition solutions is to define open subclasses of the above spaces in which functions make a specific number of transitions and oscillations. We then minimize the functional J in these classes. When the minimum is attained in the interior of a class, a local minimizer of J is found which is a smooth solution to the Euler-Lagrange Eq. (1.2) with the corresponding properties. The precise definitions of the subclasses are given in Sect. 2, but we present here a brief, informal description in terms of homotopy classes of curves in the plane.

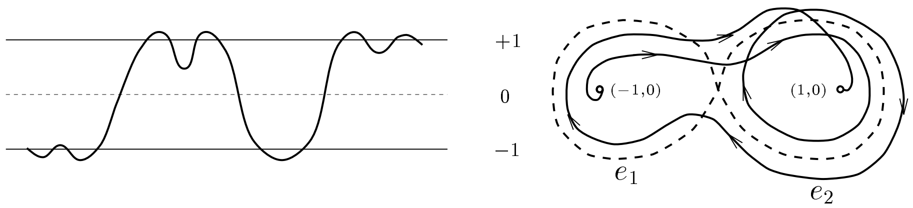

Fig. 1.1. A typical heteroclinic orbit with homotopy type $e_1 e_2^2$

Viewed in the configuration plane (u, v) where $v = u'$, a heteroclinic or homoclinic orbit is a curve connecting the points $(\pm 1, 0)$, and the transitions and oscillations record the homotopy type relative to these points. All homotopy types arising in this way can be represented by a free semigroup generated by two (clockwise) oriented loops, e_1 and e_2 around $(\pm 1, 0)$, see Fig. 1.1. Note that the winding in the tails is disregarded in this representation, and the orientation of the loops is due to the relation $v = u'$. Thus for every heteroclinic and homoclinic orbit u there exists a representative word of the form

$$e_{i_m}^{\theta_m} \cdot e_{i_{m-1}}^{\theta_{m-1}} \cdot \ldots \cdot e_{i_2}^{\theta_2} \cdot e_{i_1}^{\theta_1},$$

where $\theta(u) = (\theta_1, \ldots, \theta_m) \in \mathbb{N}^m$ and $i_{k+1} - 1 = i_k \bmod 2$. In the sequel it will be more convenient to consider $\mathbf{g}(u) = 2\theta(u)$ rather than the winding vector θ. The vector $\mathbf{g} \in 2\mathbb{N}^m$ specifies the number of crossings of ± 1 which u makes between transitions, and $\mathbf{g} = \mathbf{0}$ for functions with only one transition. For each vector $\mathbf{g} \in G = 2\mathbb{N}^m \cup \{\mathbf{0}\}$ there are two distinct homotopy classes which correspond to words beginning with e_1 and words beginning with e_2. These two classes contain functions for which $\lim_{t\to-\infty} u(t) = +1$ and $\lim_{t\to-\infty} u(t) = -1$ respectively. Let $M^\pm(\mathbf{g})$ denote these homotopy classes of functions with winding vector $\mathbf{g}/2$, and define the numbers

$$\mathcal{J}^\pm(\mathbf{g}) = \inf_{M^\pm(\mathbf{g})} J[u].$$

The above infima are well-defined since J is bounded from below, although it is not clear whether the infima are attained in each of the homotopy classes. Our goal is to prove that $\mathcal{J}^{\pm}(\mathbf{g})$ is attained in many classes $M^{\pm}(\mathbf{g})$.

Before stating our main result, we introduce an order relation $\prec$ on G defined by the following rule:

$$(g_1, \ldots, g_m) \prec (g_1, \cdots, g_{k-1}, g_k^-, 2, g_k^+, g_{k+1}, \ldots, g_m)$$
$$\text{for any } k \in \{1, \ldots, m\} \text{ and } g_k^{\pm} \in 2\mathbb{N} \text{ such that } g_k^- + g_k^+ = g_k,$$

which is extended on G by transitivity. In terms of words representing homotopy types,

$$e_1^2 \prec e_1 \cdot e_2 \cdot e_1 \quad \text{and} \quad e_2^2 \prec e_2 \cdot e_1 \cdot e_2.$$

This order relation determines the classes in which we can find minimizers as stated in the following theorem.

Theorem 1.1. *Suppose that $F \in C^2(\mathbb{R})$ has exactly two nondegenerate global minima at $u = \pm 1$, and F grows superquadratically as $u \to \pm\infty$. If $\beta, \gamma > 0$ are chosen such that ± 1 are saddle-focus equilibria of* (1.2)*, then for any $\mathbf{g} \in G$ there exist $\mathbf{h}^{\pm} \succeq \mathbf{g}$ such that $\mathcal{J}^{\pm}(\mathbf{h}^{\pm})$ are attained by functions $\widehat{u}^{\pm} \in C^4(\mathbb{R}) \cap M^{\pm}(\mathbf{h})$.*

Remark. Since there are only finitely many $\mathbf{h} \in G$ with $\mathbf{h} \succ \mathbf{g}$, the following alternative holds for either + or $-$ throughout: either $\mathcal{J}^{\pm}(\mathbf{g}) < \mathcal{J}^{\pm}(\mathbf{h})$ for all $\mathbf{h} \succ \mathbf{g}$ and $\mathcal{J}^{\pm}(\mathbf{g})$ is attained in $M^{\pm}(\mathbf{g})$, or there are finitely many vectors $\mathbf{h}_i^{\pm} \succ \mathbf{g}$, $i = 1, .., n$, such that $\mathcal{J}^{\pm}(\mathbf{h}_1^{\pm}) = \ldots = \mathcal{J}^{\pm}(\mathbf{h}_n^{\pm}) \leq \mathcal{J}^{\pm}(\mathbf{g})$, and $\mathcal{J}^{\pm}(\mathbf{h}_i^{\pm})$ is attained in $M^{\pm}(\mathbf{h}_i^{\pm})$ for each $i \leq n$.

The minima obtained in Theorem 1.1 are local minima of J in the appropriate function spaces $\chi + H^2(\mathbb{R})$. The theorem does not imply that the infimum is attained in every homotopy class, but it can be shown that there are certain classes in which a local minimizer must exist. In particular, if the winding numbers g_i are all small enough or large enough, then the infimum $\mathcal{J}^{\pm}(\mathbf{g})$ is attained in $M^{\pm}(\mathbf{g})$. In the latter case, the resulting local minimizers are multitransition solutions whose transitions are close to the single transition minimizers in $\mathcal{J}^{\pm}(\mathbf{0})$ and are separated by large distances. These solutions are analogous to those found in typical multibump constructions, cf. [26].

Theorem 1.2. *Let F, β, γ be as in Theorem* 1.1*. If $\mathbf{g} = \mathbf{0}$ or $\mathbf{g} \in 2\mathbb{N}^m$ with $g_i = 2$ for all $i \leq m$, then $\mathcal{J}^{\pm}(\mathbf{g})$ are attained by minimizers in $M^{\pm}(\mathbf{g})$. Futhermore, there exists an $N > 0$ such that, if $\mathbf{g} \in 2\mathbb{N}^m$ and $g_i > N$ for all $i \leq m$, then $\mathcal{J}^{\pm}(\mathbf{g})$ are attained by minimizers in $M^{\pm}(\mathbf{g})$.*

In Theorems 1.1 and 1.2 the hypotheses on F are fairly mild. If additional symmetry for F is assumed, then local minima exist in every homotopy class.

Theorem 1.3. *Let F, β, γ be as in Theorem* 1.1*, and assume in addition that $F(u) = F(-u)$ for all $u \in \mathbb{R}$. Then for any $\mathbf{g} \in G$ the infima $\mathcal{J}^-(\mathbf{g}) = \mathcal{J}^+(\mathbf{g})$ are attained by minimizers in the associated homotopy classes.* [1]

[1] Theorem 1.3 was originally proved with the restriction that $g_i = 2$ or $g_i \geq 4$ for all i. It was demonstrated to us by J.B. VandenBerg [47] and the referee that this restriction is unnecessary, see Sect. 7.

The proofs of these theorems are based on a constrained minimization principle in which J is minimized on open sets $M^{\pm}(\mathbf{g})$ in $\chi + H^2(\mathbb{R})$. The main difficulty is to show that there exist minimizing sequences which are bounded with respect to the appropriate norm and whose weak limits are contained in the interior of the class $M^{\pm}(\mathbf{g})$. The oscillatory nature of solutions which lie in a neighborhood of a saddle-focus equilibrium is crucial to the control of minimizing sequences and is described in Sect. 4. In Sect. 3 we develop tools for removing spurious oscillations from minimizing sequences. The results of these two sections complement each other, and combined with fairly simple a priori estimates, they comprise the essential ingredients of the proofs of Theorems 1.1, 1.2, and 1.3 in Sects. 5, 6, and 7 respectively.

Minimizing sequences in a class $M^{\pm}(\mathbf{g})$ can lose complexity in the limit, i.e. approach the boundary of $M^{\pm}(\mathbf{g})$, in two ways: crossings of ± 1 can coalesce, or the distance between crossings can grow to infinity. In both cases minimizing sequences can be adjusted by replacing pieces of the functions in the sequence with pieces of orbits near a saddle-focus equilibrium. The oscillatory properties of such orbits ensure that the limits of these specially-constructed minimizing sequences remain in the interior of the class $M^{\pm}(\mathbf{g})$. Intuitively this is the main idea of this paper, but the implementation requires some technical adjustments, see Sect. 5. We begin in Sect. 2 with a precise description of the functional analytic framework for these minimization problems, and in Sect. 8 we briefly describe other problems in which these techniques might be useful.

Throughout this paper C will denote an arbitrary constant which may change from line to line. Generally C will depend on the parameters γ, β, and the nonlinearity F. Any other important dependence will be explicitly specified.

2. Preliminaries

The Hamiltonian of (1.2) is given by

$$\begin{aligned} H(u,u',u'',u''') &= -\gamma u''' u' + \frac{\gamma}{2}|u''|^2 + \frac{\beta}{2}|u'|^2 - F(u) \\ &= p_1 q_2 + \frac{1}{2\gamma}p_2^2 - \frac{\beta}{2}q_2^2 - F(q_1) \end{aligned}$$

in the symplectic coordinates $(q,p) = (q_1,q_2,p_1,p_2)$ defined by $(q_1,q_2) = (u,u')$ and $(p_1,p_2) = (\beta u' - \gamma u''', \gamma u'')$. The canonical Lagrangian then has the form $I[q,p] = \int \{\langle q', p\rangle - H(q,p)\}$. Since this Lagrangian is strongly indefinite, it is more convenient to study homoclinic and heteroclinic orbits of (1.2) as critical points of the action functional

$$J[u] = \int_{\mathbb{R}} \left[\frac{\gamma}{2}|u''|^2 + \frac{\beta}{2}|u'|^2 + F(u)\right] dt$$

with $\gamma, \beta > 0$, which is obtained from $I[q,p]$ by substituting the above definiton of (q,p).

Recall that the function $F \in C^2(\mathbb{R})$ is a nondegenerate double-well potential which grows superquadratically as $|u| \to \infty$. The specific hypothesis is

(H1) $F(\pm 1) = F'(\pm 1) = 0$, $F''(\pm 1) > 0$, and $F(u) > 0$ for $u \neq \pm 1$. Moreover there are constants c_1 and c_2 such that $F(u) \geq -c_1 + c_2 u^2$.

This implies the following property which will be used in the sequel:

(H2) for every $\alpha \in (-1,1)$ there exists $\eta(\alpha) > 0$ such that

$$F(u) \geq \begin{cases} \eta(\alpha)^2(u-1)^2 & \text{for } u \in (\alpha,\infty), \\ \eta(\alpha)^2(u+1)^2 & \text{for } u \in (-\infty,\alpha). \end{cases}$$

As described in the introduction, we will consider classes of functions in the affine spaces $\chi + H^2(\mathbb{R})$. We will restrict attention to functions for which $\lim_{t\to-\infty} u(t) = -1$, in which case χ is either $\chi_{-1} \equiv -1$ or $\chi_1 = \widehat{\chi}$, a fixed smooth function with $\widehat{\chi}(t) = -1$ for $t \leq -1$ and $\widehat{\chi}(t) = 1$ for $t \geq 1$. The other cases are completely analogous, and thus we will drop the superscripts $\pm$ from the notation for the classes $M(\mathbf{g})$. For $m \geq 1$ and $\mathbf{g} \in 2\mathbb{N}^m$ we define the subclass $M(\mathbf{g})$ of $\chi_{(-1)^m} + H^2(\mathbb{R})$ as follows.

Definition 2.1. *A function u is in $M(\mathbf{g})$ if there are nonempty sets $\{A_i\}_{i=0}^{m+1}$ such that*

i) $u^{-1}(\pm 1) = \bigcup_{i=0}^{m+1} A_i$,
ii) $\# A_i = g_i$ *for* $i = 1,\ldots,m$,
iii) $\max A_i < \min A_{i+1}$ *for* $i = 0,\ldots,m$,
iv) $u(A_i) = (-1)^{i+1}$, *and*
v) $\{\max A_0\} \cup \left(\bigcup_{i=1}^m A_i\right) \cup \{\min A_{m+1}\}$ *consists of transverse* **crossings** *of* ± 1.

Under these conditions $M(\mathbf{g})$ is an open subset of $\chi_{(-1)^m} + H^2(\mathbb{R})$. For $m = 0$ define $M(\mathbf{0})$ as above with two sets A_0 and A_1 each with at least one transverse crossing. For convenience we will suppress the dependence of χ on m and use the notation $|\mathbf{g}| = m$ if $\mathbf{g} \in 2\mathbb{N}^m$ and $|\mathbf{0}| = 0$.

For $\mathbf{g} \in G = \cup_{m=1}^{\infty} 2\mathbb{N}^m \cup \{\mathbf{0}\}$ functions in $M(\mathbf{g})$ make $|\mathbf{g}| + 1$ **transitions** between ± 1, and these occur on the intervals from $\max A_i$ to $\min A_{i+1}$ for $i = 0,\ldots,m$. The numbers g_i count the crossings of either -1 or $+1$ between consecutive transitions, and these crossings are transverse as well as the crossings at the beginning of the first transition and at the end of the last transition. Functions in $\chi + H^2(\mathbb{R})$ can make infinitely many crossings of ± 1 as $t \to \pm\infty$, but we do not make any assumptions about their transversality. For $u \in M(\mathbf{g})$ we will call the interval from the beginning of the first transition to the end of the last transition, i.e. from $\max A_0$ to $\min A_{m+1}$, the **core interval** of u, see Fig. 2.1.

These classes have been defined so that between any two crossings in A_i functions in $M(\mathbf{g})$ stay strictly above/below $(-1)^i$. Hence a function which crosses +1, then has a tangency at -1, and subsequently crosses +1 again is on the boundary between two distinct classes. Initially we will allow minimizing sequences to move from one class to another in this manner. To formalize this idea we define the following partial ordering $\prec$ on the set $G = \cup_{m=1}^{\infty} 2\mathbb{N}^m \cup \{\mathbf{0}\}$.

Let $\mathbf{g} = (g_1,\ldots,g_m) \in G$. For any $k \in \{1,\ldots,m\}$ and any $g_k^{\pm} \in 2\mathbb{N}$ such that $g_k^- + g_k^+ = g_k$, we declare

$$(g_1,\ldots,g_m) \prec (g_1,\cdots,g_{k-1},g_k^-,2,g_k^+,g_{k+1},\ldots,g_m).$$

Also if $\mathbf{g}$, $\mathbf{h}$, $\mathbf{k} \in G$ with $\mathbf{g} \prec \mathbf{h}$ and $\mathbf{h} \prec \mathbf{k}$, then $\mathbf{g} \prec \mathbf{k}$. These relations define a partial ordering on the set G.

For any $\mathbf{g} \in G$ define $\mathcal{M}(\mathbf{g}) = \cup_{\mathbf{h}\succeq\mathbf{g}} M(\mathbf{h})$. The functions in $\mathcal{M}(\mathbf{g})$ are at least as topologically complex as those in $M(\mathbf{g})$, i.e. they have at least $|\mathbf{g}| + 1$ transitions and at least $\sum_i g_i$ crossings of ± 1 between the first and last transitions. Let $\mathcal{J}(\mathbf{g}) = \inf_{\mathcal{M}(\mathbf{g})} J$. We will prove a more detailed version of Theorem 1.1 which states that given any $\mathbf{g} \in G$ there is a local minimizer of J which is at least as topologically complex as functions in $M(\mathbf{g})$ in the sense of the above ordering, see Fig. 2.1.

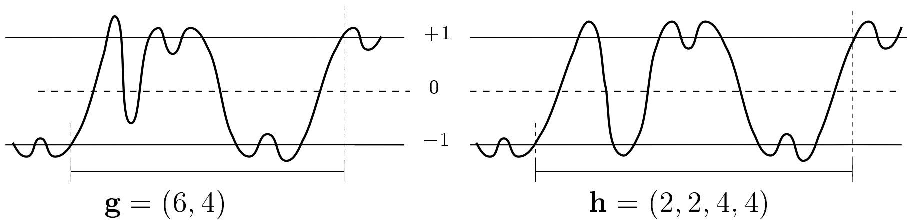

Fig. 2.1. $(6, 4) \prec (2, 2, 4, 4)$. The intervals pictured are the core intervals

Theorem 2.2. *Suppose F satisfies the hypothesis* (H1) *and $\gamma, \beta > 0$ are chosen such that ± 1 are saddle-foci. Then for any $\mathbf{g} \in G$ there exists $\widehat{u} \in \mathcal{M}(\mathbf{g})$ which is a local minimizer of J in $\chi + H^2(\mathbb{R})$ with the following properties:*

i) *$\widehat{u}$ has strictly monotone transitions,*
ii) *$\widehat{u}$ has only one local extremum between consecutive crossings, and*
iii) *the tails of $\widehat{u}$ have a countable infinity of crossings of ± 1, all of which are transverse, and between consecutive crossings $\widehat{u}$ has one local extremum. In each tail, there are two sequences $(\widehat{u}(t_n^{\max}))$ and $(\widehat{u}(t_n^{\min}))$ consisting of all local maxima and minima. These sequences are strictly monotone, and $\widehat{u}(t_n^{\max}) \searrow \pm 1$ and $\widehat{u}(t_n^{\min}) \nearrow \pm 1$ as $n \to \infty$.*

Moreover, either $\mathcal{J}(\mathbf{g}) < \mathcal{J}(\mathbf{h})$ for all $\mathbf{h} \succ \mathbf{g}$ and $\widehat{u} \in M(\mathbf{g})$, or there are finitely many $\mathbf{h}_i \succ \mathbf{g},\ i = 1, \ldots, n$, such that $\mathcal{J}(\mathbf{h}_1) = \ldots = \mathcal{J}(\mathbf{h}_n) \leq \mathcal{J}(\mathbf{g})$, and there exist local minimizers $\widehat{u}_i$ in each $M(\mathbf{h}_i),\ i = 1, \ldots, n$.

Remark. This theorem establishes the existence of locally minimizing heteroclinic and homoclinic solutions emanating from -1. Obviously the same result holds for solutions starting at +1.

For certain vectors $\mathbf{g} \in G$ the theorem immediately implies that there is a local minimizer in the class $M(\mathbf{g})$.

Corollary 2.3. *There are local minimizers of J in any class $M(\mathbf{g})$ for which $g_i = 2$ for all $i \leq |\mathbf{g}|$. In particular there exist minimizers in the classes $M(\mathbf{0})$ and $M((2))$ which correspond to a single transition heteroclinic orbit and a single pulse homoclinic orbit with no oscillations between two transitions.*

Remark. The minimizers in the class $M(\mathbf{0})$ are global minimizers in $\chi_1 + H^2(\mathbb{R})$. Global minimizers in these affine spaces can be found without assuming ± 1 are saddle-foci. In the saddle-focus case these global minimizers must a priori be in $M(\mathbf{0})$, i.e. have oscillations in the tails, but this is not necessary for other types of equilibria.

3. Clipping

Let $u \in C^1[a, b]$. Suppose there is a subinterval $I = [\alpha, \beta]$ of $[a, b]$ such that $u(\alpha) = u(\beta)$ and $u'(\alpha) = u'(\beta)$. Then we can clip out the interval I from $[a, b]$ by collapsing it to a point (cf. Fig. 3.1) to obtain a function $u^* \in C^1[a, b - |I|]$ which is formally defined by

$$u^*|_{[a,\alpha]} \equiv u|_{[a,\alpha]} \quad \text{and} \quad u^*|_{[\alpha, b-|I|]} \equiv u|_{[\beta,b]}.$$

Here $|I|$ denotes the length of the interval I. More generally, a function u^* will be called a **clip** of u if it is obtained by clipping out a finite number of intervals from u. Note that the values of the functions u and u^* along with their derivatives coincide at the corresponding endpoints of their domains. Clipping is a well-defined operation on $H^2(\mathbb{R})$ functions. Since the integrand of J is nonnegative, it also has the fundamental property that $J[u^*] \leq J[u]$ for any clip u^* of u. The next three lemmas will be tools for clipping functions, and these ideas are best understood by examining intersections of the corresponding curves in the configuration plane (u, u'), cf. Fig. 1.1.

Lemma 3.1. *Let $a_1 < b_1 \leq a_2 < b_2$, and $I_j = [a_j, b_j]$, $j = 1$ or 2. Suppose a function $u \in C^1(I_1) \cap C^1(I_2)$ is increasing on both I_1 and I_2 with $u(I_1) \cap u(I_2) \neq \emptyset$ and satifies one of the following two properties:*

i) $u(a_1) = u(a_2),\ u(b_1) = u(b_2)$,
and $(u'(a_1) - u'(a_2)) \cdot (u'(b_1) - u'(b_2)) \leq 0$, *or*
ii) $u'(a_1) = u'(a_2) = u'(b_1) = u'(b_2) = 0$
and $(u(a_1) - u(a_2)) \cdot (u(b_1) - u(b_2)) \geq 0$.

Then there exist $c_j \in I_j$ such that $u(c_1) = u(c_2)$ and $u'(c_1) = u'(c_2)$. Hence the interval (c_1, c_2) can be clipped out of u to produce an increasing function. The same result holds for functions which decrease on I_1 and I_2, and if u is strictly monotone over these intervals, then the clip of u is also strictly monotone.

Proof. First consider the case in which the hypothesis (i) is satisfied. Assume $u(a_1) < u(b_1)$, and let $I = [u(a_1), u(b_1)]$, as the other case is similar. Since u is C^1 and monotone on the intervals I_1 and I_2, the function

$$\varphi(s) = u'(u|_{I_1}^{-1}(s)) - u'(u|_{I_2}^{-1}(s))$$

is well-defined and continuous for $s \in I$. By hypothesis

$$\varphi(u(a_1)) \cdot \varphi(u(b_1)) = (u'(a_1) - u'(a_2)) \cdot (u'(b_1) - u'(b_2)) \leq 0.$$

Therefore $\varphi(u(a_1))$ and $\varphi(u(b_1))$ have opposite signs, and $\varphi(s_0) = 0$ for some $s_0 \in I$. Let $c_j \in u|_{I_j}^{-1}(s_0)$, $j = 1$ or 2. Then $u(c_1) = u(c_2)$ and $u'(c_1) = u'(c_2)$ by construction. The clip of u inherits the monotonicity properties of u on the intervals $[a_1, c_1]$ and $[c_2, b_2]$.

Now suppose u satisfies hypothesis (ii) and u is increasing on I_1 and I_2. Then there are two cases, and we consider $u(a_1) \leq u(a_2)$ and $u(b_1) \leq u(b_2)$, as the other case is similar. Since $u(I_1) \cap u(I_2) \neq \emptyset$, there are points $\widehat{a}_1 \in I_1$ and $\widehat{b}_2 \in I_2$ such that $u(\widehat{a}_1) = u(a_2)$ and $u(\widehat{b}_2) = u(b_1)$. Then u satisfies hypothesis (i) on the intervals $[\widehat{a}_1, b_1]$ and $[a_2, \widehat{b}_2]$. $\square$

Typically, we will apply this lemma to functions defined on all of $\mathbb{R}$. In this case, the clipping operation is localized and removes a finite interval so that the resulting function is again defined on all of $\mathbb{R}$, see Fig. 3.1.

The next two lemmas apply to Morse functions, i.e. C^2 functions whose critical points are all nondegenerate. Morse functions have finitely many critical points on a compact interval, all of which are local maxima or minima. We will denote the closed convex hull of a set A by conv (A).

Lemma 3.2. *Let $u \in C^2[a, b]$ be a Morse function with $u(a) \neq u(b)$. Suppose $u([a, b]) \subset$ conv $(\{u(a), u(b)\})$. Then u can be clipped to a strictly monotone function.*

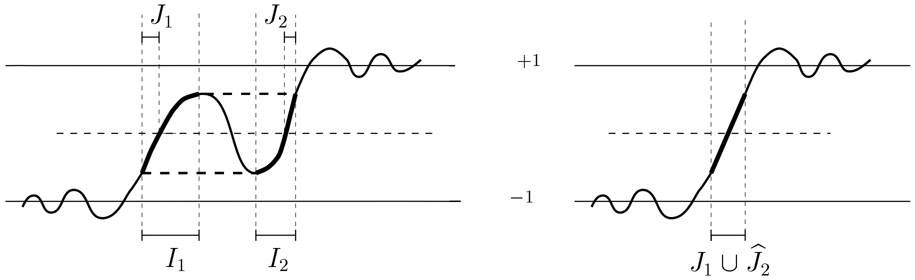

Fig. 3.1. A typical example of the clipping operation. As in Lemma 3.1, the intervals $J_1 = [a_1, c_1]$ and $J_2 = [c_2, b_2]$ are concatenated to $[a_1, b_2 - c_2 + c_1] = [a_1, c_1] \cup [c_1, b_2 - c_2 + c_1] = J_1 \cup \widehat{J_2}$

Proof. If u has no critical points, then u is monotone, and there is nothing to prove. Otherwise we will show that u can be clipped to remove at least two critical points.

Let $b_1 > a$ be the first critical point of u. Consider the case $u(a) < u(b)$ as the other case is similar. Since $u(t) \geq u(a)$ and u is Morse, u is strictly increasing on $[a, b_1]$, and u has a local maximum at b_1. Let $b_2 = \sup\{t : u(t) = u(b_1)$ and u is increasing at $t\}$ and $a_2 = \sup\{t < b_2 : u'(t) = 0\}$, the location of the local minimum to the left of b_2. Finally, since u is increasing on (a, b_1), let a_1 be the unique point in $[a, b_1)$ with $u(a_1) = u(a_2)$. Note that u is strictly increasing on $[a_2, b_2]$ and on $[a_1, b_1]$. Applying Lemma 3.1 (hypothesis (i)) to u with subintervals $I_j = [a_j, b_j],\ j = 1, 2$, we obtain points $c_j \in I_j$ for $j = 1, 2$ at which u can be clipped to a function u^*. Since b_1 and a_2 are in the interval which is clipped out, u^* has at least two critical points less than u. An even number of critical points are removed.

Since there are only finitely many critical points, this process can be repeated to obtain a strictly monotone function. □

Lemma 3.3. *Let $u \in C^2[a, b]$ be Morse with $u(a) = u(b),\ u'(a) \cdot u'(b) < 0$, and $u([a, b]) \subset [u(a), \infty)$ or $(-\infty, u(b)]$. Then there exists a clip of u which has exactly one critical point in (a, b).*

Proof. Assume $u'(b) < 0 < u'(a)$, the other case is similar. Let $M = \max u(t)$ and $S = \{t : u(t) = M\}$. Let $c_1 = \min S$ and $c_2 = \max S$. Then $u(c_1) = u(c_2) = M$ and $u'(c_1) = u'(c_2) = 0$. Apply Lemma 3.2 to the intervals $[a, c_1]$ and $[c_2, b]$ to obtain two clips, u_1^* which is increasing from $u(a)$ to M and u_2^* which is decreasing from M to $u(b)$. These two clips can be glued together at c_1 and c_2 to get a clip of u with one critical point. □

Definition 3.4. *A function $u \in M(\mathbf{g})$ is* **normalized** *if u is monotone on each transition, and there are points $-\infty \leq a < b \leq \infty$ such that*

i) *all crossings of ± 1 are transverse in (a, b),*

ii) *u contains exactly one local extremum between each pair of consecutive crossings of either $+1$ or -1 in (a, b),*

iii) *a and b are accumulation points of crossings of ± 1, and*

iv) *u contains no intervals of critical points except* $(-\infty, a]$ *and* $[b, \infty)$ *on which u is identically* ± 1.

Several comments about this definition are in order. First, the basic property of normalized functions, which will be used extensively throughout the sequel, is that all local extrema are isolated except possibly on infinite intervals at the ends of the tails where the function is identically ± 1. Moreover, each maximal monotonicity interval either contains exactly one crossing of ± 1 or exactly one crossing of both -1 and $+1$ if it is a transition. These monotonicity intervals will be used to identify places where a function can be clipped using Lemma 3.1 with hypothesis (ii). Note that the crossings are transverse in the core interval by the definition of the classes $M(\mathbf{g})$, but for normalized functions all crossings are transverse except possibly at the ends of the tails. So the core interval is contained in (a, b). Finally, the normalized functions are prevalent in the classes $M(\mathbf{g})$ as the following key lemma indicates.

Lemma 3.5. *Let $u \in M(\mathbf{g})$. For every $\epsilon > 0$ there exists a normalized $u^* \in M(\mathbf{g})$ such that $J[u^*] \leq J[u] + \epsilon$.*

Proof. The first step in the proof is to perturb u so that it possesses infinitely many transverse crossings. Choose $\mu > 0$ and a fixed cutoff function $\omega \in C^\infty(\mathbb{R})$ with $\operatorname{supp} \omega = [-1, 1]$ and $\omega(0) = 1$ and $\omega'(0) = 0$. Since $u \in \chi + H^2(\mathbb{R})$, there is a bi-infinite sequence $\ldots t_{-1} < t_0 < t_1 \ldots$, which can be chosen such that $|t_{i+1} - t_i| > 4$ and

$$|(u(t_i), u'(t_i)) - (\chi(t_i), \chi'(t_i))| < \mu \cdot 2^{-|i|-1}.$$

Clearly the sequence (t_i) can also be chosen so that each t_i is at least distance four from the core interval and $\chi(t_i) = \pm 1$ for all $i \in \mathbb{Z}$ by the definition of χ. The function

$$v(t) = u(t) + \sum_{i=-\infty}^{\infty} (\mu \cdot 2^{-|i|}(t - t_i) - u(t_i) + \chi(t_i))\omega(t - t_i)$$

has transverse crossings with positive derivative at the points $t_i \to \pm\infty$ as $i \to \pm\infty$, and $\|v - u\|_{H^2} \leq C\mu\|\omega\|_{H^2}$. Thus for μ sufficiently small $J[v] \leq J[u] + \epsilon/2$. Let $(z_i)_{i \in \mathbb{Z}}$ denote the sequence of all the crossings t_i created above along with all the crossings in the core interval of v. On the interior of each interval $[z_i, z_{i+1}]$ the function v can be perturbed to a Morse function in the usual way without altering the crossings z_i. This perturbation can be made arbitrarily small so that the resulting function w satisfies $\|w - u\|_{H^2} \leq C\mu$. For sufficiently small μ, we have $J[w] \leq J[u] + \epsilon$. Futhermore, the nontransverse crossings must be local extrema of w.

Now let $(z_i)_{i \in \mathbb{Z}}$ be the sequence of all transverse crossings of w. On each interval $[z_i, z_{i+1}]$ either $w(z_i) = w(z_{i+1}) = \pm 1$ or $w(z_i) = -w(z_{i+1}) = \pm 1$ in which case w makes a transition on this interval. By the choice of (z_i) the hypotheses of Lemmas 3.3 and 3.2 hold, and we can apply them to the two cases repectively to "normalize" w on $[z_i, z_{i+1}]$. More precisely, for each $i \in \mathbb{Z}$ we obtain an interval I_i and $w^* : I_i \to \mathbb{R}$ which is a clip of $w|_{[z_i, z_{i+1}]}$. Note that the values of w^* and its derivative match at the right endpoint of I_i and the left endpoint of I_{i+1}. Therefore, by concatenating the intervals I_i to I, we construct a C^1 and piecewise C^2 function $w^* : I \to \mathbb{R}$ with $|I| = \sum |I_i|$. Since clipping reduces the action, $J[w^*] \leq J[w] \leq J[u] + \epsilon$. Now w^* has all the properties of a normalized function in Definition 3.4 except that I need not be all of $\mathbb{R}$ (note that possibly infinitely many intervals were clipped from w). However

$$\lim_{t\to\inf I}(w^*(t),(w^*)'(t)) = \lim_{t\to\sup I}(w^*(t),(w^*)'(t)) = (\pm 1, 0).$$

Indeed any sequence $t_n \to \inf I$ or $\sup I$ can be associated with a sequence $\tau_n \to \pm\infty$ such that $(w^*(t_n),(w^*)'(t_n)) = (w(\tau_n), w'(\tau_n))$ which tends to $(\pm 1, 0)$ since $w \in \chi + H^2(\mathbb{R})$. In this way

$$u^*(t) = \begin{cases} w^*(t) & \text{for } t \in I, \\ \pm 1 & \text{for } t \notin I, \end{cases}$$

is a normalized function in $\chi + H^2(\mathbb{R})$ as in Definition 3.4 with $J[u^*] \leq J[u] + \epsilon$. □

The previous lemma allows us to consider minimizing sequences which are normalized functions. The clipping lemmas also imply that adding more crossings in the core interval can only increase the action J.

Lemma 3.6. *Suppose* $\mathbf{g}, \mathbf{h} \in 2\mathbb{N}^m$ *with* $g_i \leq h_i$ *for all* $i \leq m$. *Then* $\mathcal{J}(\mathbf{g}) \leq \mathcal{J}(\mathbf{h})$.

Proof. We will consider the case where $g_i = h_i$ for all $i \geq 2$ and $g_1 = h_1 - 2$. The case $i \neq 1$ is similar, and the general case follows by induction. For any $u \in M(\mathbf{h})$ which is normalized we construct $v \in M(\mathbf{g})$ with strictly lower action.

Let $\tau_1 = \max A_0$ and $\tau_2 = \min A_2$, and consider u restricted to $[\tau_1, \tau_2]$. Let s_1 be such that u has a local minimum at s_1 and no other local minimum of $u|_{(\tau_1,\tau_2)}$ is larger than $u(s_1)$. Let $s_2 < s_1 < s_3$ be such that u is strictly monotone on $[s_2, s_1]$ and $[s_1, s_3]$, and u has a local maximum at s_2 and s_3. Assume $u(s_2) \leq u(s_3)$, as the other case is similar. Define $a = \sup\{t < s_1 : u(t) = u(s_1)\}$ and $b = \inf\{t > s_1 : u(t) = u(s_2)\}$. By construction $u(a) = u(s_1) < 1 < u(s_2) = u(b)$. Also $u'(a) \geq 0,\ u'(b) \geq 0$, and $u'(s_1) = u'(s_2) = 0$. Since u is normalized, the hypothesis (i) of Lemma 3.1 holds on the intervals $I_1 = [a, s_2]$ and $I_2 = [s_1, b]$. Therefore we can clip u over I_1 and I_2 to a monotone function v. Exactly two crossings are removed because u has one crossing in each interval $I_1 = [a, s_2],\ [s_2, s_1]$, and $I_2 = [s_1, b]$.

Let $\epsilon > 0$ and $u \in M(\mathbf{h})$ be such that $J[u] \leq \mathcal{J}(\mathbf{h}) + \epsilon/2$. By Lemma 3.5 there is a normalized $u^* \in M(\mathbf{h})$ sufficiently close to u such that $|J[u] - J[u^*]| < \epsilon/2$. By the above argument we can find $v^* \in M(\mathbf{g})$ such that $J[v^*] < J[u^*] < \mathcal{J}(\mathbf{h}) + \epsilon$. Therefore $\mathcal{J}(\mathbf{g}) \leq \mathcal{J}(\mathbf{h})$. □

4. Saddle-Focus Equilibria

In this section we analyze the minimizers of J on a finite interval $[0, T]$ which satisfy the following boundary value problem,

$$\begin{cases} \gamma v'''' - \beta v'' + G'(v) = 0, \\ (v(0), v'(0)) = x \quad \text{and} \quad (v(T), v'(T)) = y, \end{cases} \tag{4.1}$$

where $x, y \in \mathbb{R}^2$. For notational convenience we consider a potential $G \in C^2(\mathbb{R})$ for which $G(0) = 0$ is a nondegenerate global minimum. Since the wells at ± 1 in the original potential F can be translated to the origin, the analysis of this section will apply to both cases, i.e. $G(v) = F(v \pm 1)$. Hence we want to minimize

$$J = \int_0^T \left[\frac{\gamma}{2}|v''|^2 + \frac{\beta}{2}|v'|^2 + G(v)\right] dt$$

over the space $X_T = \{v \in H^2[0,T] : (v(0), v'(0)) = x, (v(T), v'(T)) = y\}$, and we will be interested in the properties of the minimizer with small boundary data, $\|x\|, \|y\| \ll 1$. In particular we will show that if the boundary data are sufficiently small, then the minimizer in X_T is unique and small along with all its derivatives up to third-order, and the minimizer oscillates around the origin.

Theorem 4.1. *There exists $\delta_0(G) > 0$ such that if $\|x\|, \|y\| \leq \delta \leq \delta_0$ and $T \geq 1$, then there exists a unique global minimizer $\widehat{v}$ of J in X_T satisfying the boundary value problem* (4.1). *Furthermore, $\|\widehat{v}\|_{W^{3,\infty}} \leq C\delta$ and $J[\widehat{v}] \leq C\delta^2$, where C is independent of $T \geq 1$.*

Proof. We will separate the proof into several steps.

Since G has a nondegenerate global minimum at the origin, there are constants $\delta_1 > 0$ and $\eta > 0$ such that $G(v) \geq \eta^2 v^2$ for $|v| \leq \delta_1$.

Step 1. There exists $C_1(G, \delta_1) > 0$ such that, if $\|x\|, \|y\| \leq \delta \leq \delta_1$, then $\inf_{X_T} J \leq C_1\delta^2$.

Choose any functions $\varphi_0, \varphi_1 \in C^\infty[0,1]$ such that $\operatorname{supp}(\varphi_j) \subset [0, 1/2]$ with $\varphi_0(0) = 1$, $\varphi_0'(0) = 0$, $\varphi_1(0) = 0$, and $\varphi_1'(0) = 1$, and define $\psi_j(t) = (-1)^j \varphi_j(T - t)$. Consider the function $\phi \in X_T$ defined by $\phi = x_0\varphi_0 + x_1\varphi_1 + y_0\psi_0 + y_1\psi_1$. Note that there is a constant $\xi > 0$ such that $G(v) \leq \xi v^2$, and hence $\inf_{X_T} J \leq J[\phi] \leq C_1\,\delta^2$.

Step 2. There exists $C_2(\eta) > 0$ such that for every $v \in X_T$ with $\|x\|, \|y\| \leq \delta \leq \delta_1/2$ we have $J[v] \geq C_2 \min\{\|v\|_\infty^2, \delta_1^2\}$.

First suppose $\|v\|_\infty \leq \delta_1$. If $|v(t)| \geq \|v\|_\infty/2$ for all $t \in [0,T]$, then

$$J[v] \geq \int_0^T G(v)\,dt \geq \int_0^T \eta^2 v^2\,dt \geq \frac{1}{4}\eta^2\|v\|_\infty^2 \geq C\|v\|_\infty^2.$$

Otherwise there are points t_0 and $t_1 \in [0,T]$ such that $|v(t_0)| = \|v\|_\infty/2$ and $|v(t_1)| = \|v\|_\infty$. Then

$$\begin{aligned} J[v] \geq C(\beta)\int_{t_0}^{t_1} |v'|\sqrt{G(v)}\,dt &\geq C\left|\int_{t_0}^{t_1} \eta v v'\,dt\right| \\ &= C|v(t_1)^2 - v(t_0)^2| \\ &= C(\|v\|_\infty^2 - \frac{1}{4}\|v\|_\infty^2) \geq C\|v\|_\infty^2. \end{aligned} \tag{4.2}$$

Now, if $\|v\|_\infty > \delta_1$, then there are points t_0 and t_1 such that $|v(t_0)| = \delta_1/2$ and $|v(t_1)| = \delta_1$ because the boundary conditions are smaller than $\delta_1/2$. Hence $J[v] \geq C\delta_1^2$ by (4.2).

Step 3. There exists a $\delta_0 < \delta_1/2$ and $C(\delta_0) > 0$ such that, if $\|x\|, \|y\| \leq \delta < \delta_0$ and $v \in X_T$ with $J[v] \leq 2\inf_{X_T} J$, then $\|v\|_\infty, \|v\|_{H^2} \leq C\delta$.

If $\|v\|_\infty \geq \delta_1$, then $J[v] \geq C_2\delta_1^2$ by Step 2. From Step 1, $J[v] \leq 2C_1\delta_0^2$. Thus δ_0 can be chosen small enough so that $\|v\|_\infty < \delta_1$. Again by Steps 1 and 2, $C_2\|v\|_\infty^2 \leq J[v] \leq C_1\delta^2$, which implies $\|v\|_\infty \leq C\delta$. Since $G(v) \geq \eta^2 v^2$, we have that $C(\gamma, \beta, \eta)\|v\|_{H^2}^2 \leq J[v] \leq C_1\delta^2$.

Step 4. For δ_0 sufficiently small J has a unique minimizer $\widehat{v} \in X_T$ such that $\|\widehat{v}\|_{W^{3,\infty}} \leq C\delta$, where C is independent of $T \geq 1$.

Using the a priori estimates in Step 3 and the weakly lower semicontinuity of J on X_T, a minimizer $\widehat{v} \in X_T$ of J can be found by the standard theory, and $\widehat{v}$ is a solution to (4.1).

From the differential equation $\|\widehat{v}''''\|_{L^2} \leq C(\|\widehat{v}\|_{L^2} + \|\widehat{v}''\|_{L^2}) \leq C\delta$. A straightforward interpolation inequality yields

$$\|v^{(k)}\|_{L^2[0,T]} \leq C\left\{\|v^{(k-1)}\|_{L^2[0,T]} + \|v^{(k+1)}\|_{L^2[0,T]}\right\}$$

with a constant C independent of $T \geq 1$.

Remark. Using Fourier transforms, the above inequality is easily established for functions defined on all of $\mathbb{R}$. To obtain the desired estimate for functions on the finite interval $[0,T]$, extension operators are used, and a description of their properties can be found in [9]. The independence of C on T follows from combining properties of extension operators and the estimate on $\mathbb{R}$. A complete proof of a more general result is contained in the appendix to Kalies, VanderVorst, and Wanner [27]. Similar estimates are also performed in [29].

Combining these estimates we obtain $\|\widehat{v}\|_{H^4} \leq C\delta$ which gives the bound in $W^{3,\infty}$, cf. [29]. From the assumptions on G near the origin, δ_0 can further be chosen small enough so that the standard estimates on the difference of two solutions yields the uniqueness of the minimizer.

This completes the proof of Theorem 4.1. □

In the construction of convergent minimizing sequences of J we will need to know that the minimizer $\widehat{v}$ of J found in the previous theorem has many oscillations.

Theorem 4.2. *Suppose $\gamma > \beta^2/4G''(0)$ so that the origin is a saddle-focus equilibrium in the four-dimensional flow. Then there exist $\delta_0(G) > 0$ and $\tau_0(G) > 0$ such that if $\|x\|, \|y\| \leq \delta_0$, the unique global minimizer $\widehat{v}$ of J in X_T satisfying* (4.1) *changes sign in any subinterval of length τ_0 in $[0,T]$ for $T \geq 1$.*

Proof. First we consider solutions to the linear differential equation

$$\gamma w'''' - \beta w'' + G''(0)w = 0. \tag{4.3}$$

Since the origin is a saddle-focus, it has complex eigenvalues $\pm\lambda \pm \mu i$. By rescaling time we can assume without loss of generality that $\mu = 1$ and $\lambda > 0$. Therefore all solutions to (4.3) have the form

$$w(t) = Ae^{-\lambda t}\sin(t + \varphi) + Be^{\lambda t}\sin(t + \psi)$$

for some A, B, φ, and ψ.

Step 1. There exists $\tau_0 > 0$ depending only on λ such that for every A, B, φ, and ψ there are points $\tau_\pm \in [0,\tau_0]$ such that

$$\pm\, w(\tau_\pm) \geq \frac{1}{\tau_0}\|w\|_{L^\infty[0,\tau_\pm]}. \tag{4.4}$$

We prove only the existence of $\tau = \tau_+$, as the other case is similar. The calculation is separated into two cases. First suppose

$$|B|e^{2\pi\lambda} \le \frac{1}{2}|A|e^{-2\pi\lambda}.$$

Choose $\tau \in [0, 2\pi]$ such that $\sin(\tau + \varphi) = \operatorname{sgn} A$. Then we can estimate

$$w(\tau) \ge |A|e^{-2\pi\lambda} - |B|e^{2\pi\lambda} \ge \frac{1}{2}|A|e^{-2\pi\lambda}, \text{ and}$$
$$\|w\|_{L^\infty[0,\tau]} \le |A| + |B|e^{2\pi\lambda} \le |A| + \frac{1}{2}|A|e^{-2\pi\lambda} \le 2|A|.$$

Otherwise

$$|B|e^{2\pi\lambda} \ge \frac{1}{2}|A|e^{-2\pi\lambda}.$$

Choose $\tau \in [2\pi + \lambda^{-1}\ln 4, 4\pi + \lambda^{-1}\ln 4]$ such that $\sin(\tau + \psi) = \operatorname{sgn} B$. For this choice of τ we have

$$\frac{1}{2}|B|e^{\lambda\tau} \ge 2|B|e^{2\pi\lambda} \ge |A|e^{-2\pi\lambda} \ge |A|e^{-\lambda\tau}.$$

Thus we can estimate

$$w(\tau) \ge |B|e^{\lambda\tau} - |A|e^{-\lambda\tau} \ge \frac{1}{2}|B|e^{\lambda\tau} \ge \frac{1}{2}|B|e^{2\pi\lambda}, \text{ and}$$
$$\|w\|_{L^\infty[0,\tau]} \le |A| + |B|e^{\lambda\tau} \le 2|B|e^{4\pi\lambda} + |B|e^{\lambda\tau} \le |B|\left[2e^{4\pi\lambda} + e^{4\pi+\lambda^{-1}\ln 4}\right].$$

If τ_0 is chosen larger than

$$\max\{4\pi + \frac{\ln 4}{\lambda}, 4e^{2\pi\lambda} + 2e^{4\pi-2\pi\lambda+\lambda^{-1}\ln 4}\} > 1,$$

then for every w there is a $\tau_+ \in [0, \tau_0]$ such that (4.4) holds.

Step 2. There exists $\delta_1 > 0$ such that if v is the solution to the nonlinear differential equation

$$\gamma v'''' - \beta v'' + G'(v) = 0$$

with initial conditions $v_0 = (v(0), v'(0), v''(0), v'''(0))$ and $\|v_0\| < \delta_1$, then v changes sign in $[0, \tau_0]$.

First note that if w is the solution to the linear Eq. (4.3) with the same initial conditions, and δ_1 is small enough so that $\|v\|_{L^\infty[0,\tau_0]}, \|w\|_{L^\infty[0,\tau_0]} \le 1$, then there is a constant $C = C(\tau_0)$ such that

$$\|v - w\|_{L^\infty[0,t]} \le C(\tau_0)\|\rho(v)\|_{L^\infty[0,t]} \cdot \|v\|_{L^\infty[0,t]} \quad \text{for all } t \in [0, \tau_0], \tag{4.5}$$

where $\rho(v) = (G'(v) - G''(0)v)/v$. This estimate is obtained from the variation of constants formula

$$\mathbf{v}(t) = \mathbf{w}(t) + \int_0^t e^{L(t-s)}\mathbf{N}(v(s))\,ds,$$

where $\mathbf{v} = (v, v', v'', v''')$, $\mathbf{w} = (w, w', w'', w''')$, $\mathbf{N}(\mathbf{v}) = (0, 0, 0, G'(v) - G''(0)v)$, and L is the 4×4 matrix of the linear vector field obtained by writing (4.3) as a first-order system. The estimate (4.5) follows from the fact that $|G'(v) - G''(0)v| = o(|v|)$

as $|v| \to 0$. Now choose $\kappa < 1$ such that $0 < C\kappa/(1 - C\kappa) \le 1/2\tau_0$ and δ_1 such that $\|\rho(v)\|_{L^\infty[0,t]}, \|v\|_{L^\infty[0,\tau_0]}, \|w\|_{L^\infty[0,\tau_0]} \le \kappa$.

We now estimate as follows,

$$\|v\|_{L^\infty[0,t]} \le \|w\|_{L^\infty[0,t]} + \|v - w\|_{L^\infty[0,t]} \le \|w\|_{L^\infty[0,t]} + C\|\rho(v)\|_{L^\infty[0,t]} \cdot \|v\|_{L^\infty[0,t]},$$

and hence

$$(1 - C\kappa)\|v\|_{L^\infty[0,t]} \le \|w\|_{L^\infty[0,t]}.$$

This implies that

$$\|v-w\|_{L^\infty[0,t]} \le C\|\rho(v)\|_{L^\infty[0,t]} \cdot \|v\|_{L^\infty[0,t]} \le \frac{C\kappa}{1 - C\kappa}\|w\|_{L^\infty[0,t]} \le \frac{1}{2\tau_0}\|w\|_{L^\infty[0,t]}.$$

Now take $t = \tau = \tau_+$ as in Step 1. Then

$$v(\tau) \ge w(\tau) - \|v - w\|_{L^\infty[0,\tau]} \ge \frac{1}{\tau_0}\|w\|_{L^\infty[0,\tau]} - \frac{1}{2\tau_0}\|w\|_{L^\infty[0,\tau]} > 0.$$

So $v(\tau_+) > 0$ and similarly $v(\tau_-) < 0$.

Finally let $T \ge 1$ and $\widehat{v}$ be the minimizer from Theorem 4.1 on the interval $[0, T]$, and choose δ_0 sufficiently small such that $\|\widehat{v}\|_{W^{3,\infty}} < \delta_1$. Note that in the above analysis we rescaled time, and hence redefine the constant τ_0 to be τ_0/μ. Then either $T < \tau_0$ and the theorem is vacuously satisfied, or $T \ge \tau_0$. In the latter case, Step 2 above implies that $\widehat{v}$ changes sign on every subinterval of length τ_0 in $[0, T]$. This completes the proof of Theorem 4.2. $\square$

5. Minimization

In this section we minimize J in the classes $\mathcal{M}(\mathbf{g})$ defined in Sect. 2 and prove Theorem 2.2 and Corollary 2.3. The main idea in this minimization problem is to use the clipping lemmas and local theory of the previous sections to construct minimizing sequences which have a weak limit in the class $\mathcal{M}(\mathbf{g})$. The limiting function is then a local minimizer of J in $\chi + H^2(\mathbb{R})$.

First we obtain estimates for functions which stay away from a neighborhood of ± 1.

Lemma 5.1. *Let $u \in H^2[a, b]$ and $\delta > 0$. Then there exists a constant $C(\beta, \delta)$ such that*

$$J[u] \ge C|u(b) - u(a)| \quad \text{and} \quad J[u] \ge C(b - a) \tag{5.1}$$

whenever $|u \pm 1| > \delta$ on $[a, b]$. Moreover, J is uniformly continuous on the sublevel set $J^c = \{u \in H^2[a, b] : J[u] \le c\}$.

Proof. We estimate

$$J[u] \ge \int_a^b \frac{\beta}{2}(u')^2 + F(u)\, dt \ge C\left(\frac{(u(b) - u(a))^2}{b - a} + b - a\right) \tag{5.2}$$

using the Schwartz inequality and that $F(u) \ge C(\delta)$ for $|u \pm 1| > \delta$. The first estimate in (5.1) follows from the arithmetic-geometric mean (Young's) inequality and the second estimate is clear.

Let $u \in J^c$. Since $J[u] \le c$ and $F(u) \ge -C_1 + C_2 u^2$, we have that $\|u\|_{H^2} \le C$ which implies a uniform bound on u in L^∞ for all $u \in J^c$.

Therefore, using the local Lipschitz continuity of F, we have

$$\begin{aligned} |J[u+\varphi] - J[u]| &\le C \int_a^b |u''\varphi''| + (\varphi'')^2 + |u'\varphi'| + (\varphi')^2 + |F(u+\varphi) - F(u)|\, dt \\ &\le C(\|\varphi\|_{H^2}, c, b-a) \cdot \|\varphi\|_{H^2} \end{aligned}$$

for any $u \in J^c$ and $\varphi \in H^2[a,b]$ which establishes uniform continuity. $\square$

Corollary 5.2. *There is a constant $\kappa(\beta, F)$ so that, whenever $u \in \chi + H^2(\mathbb{R})$ makes a transition on an interval I, we have $J[u|_I] \ge \kappa(\beta, F)$. Therefore $J[u] \ge \kappa(\beta, F)\cdot(|\mathbf{g}|+1)$ for all $u \in M(\mathbf{g})$.*

Proof. In each transition there must be an interval $[a,b] \subset I$ such that $u([a,b]) = [1/4, 3/4]$. The previous lemma implies that $J[u|_I] \ge C(\beta,\delta)(3/4 - 1/4) = \kappa(\beta, F)$. Since $u \in M(\mathbf{g})$ has $|\mathbf{g}| + 1$ transitions, the result follows. $\square$

Remark. The above estimates require $\beta > 0$, but if $\beta = 0$, similar but more delicate estimates can be obtained.

Now consider the minimization problem

$$\inf \left\{ J[u] : u \in \mathcal{M}(\mathbf{g}) = \bigcup_{\mathbf{h} \succeq \mathbf{g}} M(\mathbf{h}) \right\} = \min \left\{ \mathcal{J}(\mathbf{h}) : \mathbf{h} \succeq \mathbf{g} \right\}. \tag{5.3}$$

The last equality holds because $\{\mathbf{h} : \mathbf{h} \succeq \mathbf{g}\}$ is a finite set, and also note that there is a maximal element of the form $\mathbf{g}_{\max} = (2, 2, \dots, 2)$ with $|\mathbf{g}_{\max}| = \sum g_i - |\mathbf{g}|$. Within this set it is convenient to consider only those vectors for which $\mathcal{J}(\mathbf{h}) = \inf_{\mathcal{M}(\mathbf{g})} J$ and which are maximal with respect to $\prec$, and we define

$$E(\mathbf{g}) = \left\{ \mathbf{h} \in G : \mathbf{h} \succeq \mathbf{g},\ \mathcal{J}(\mathbf{h}) = \inf_{\mathcal{M}(\mathbf{g})} J, \text{ and } \mathcal{J}(\mathbf{h}) < \mathcal{J}(\mathbf{k}) \text{ for all } \mathbf{k} \succ \mathbf{h} \right\}.$$

Since no two elements in $E(\mathbf{g})$ can be related by $\prec$, either $E(\mathbf{g}) = \{\mathbf{g}\}$ or $E(\mathbf{g}) = \{\mathbf{h}_1, \dots, \mathbf{h}_n : \mathbf{h}_i \succ \mathbf{g}\}$. In the first case $\mathcal{J}(\mathbf{g}) < \mathcal{J}(\mathbf{h})$ for all $\mathbf{h} \succ \mathbf{g}$, and otherwise $\mathcal{J}(\mathbf{h}_1) = \ldots = \mathcal{J}(\mathbf{h}_m) \le \mathcal{J}(\mathbf{g})$. Also let $I_i = \operatorname{conv}(A_i)$ for $i = 1, \dots, m,\ \ I_0 = (-\infty, \max A_0)$, and $I_{m+1} = (\min A_{m+1}, \infty)$, where the sets A_i are those used in the definition of the classes in Sect. 2, i.e. the intervals I_i for $i = 1, \dots, m$ are simply the intervals between transitions.

We will minimize J in a fixed class $M(\mathbf{h})$. This minimization can only be accomplished in those classes which have the following property:

Uniform Separation Property. There exists $\epsilon_0 > 0$ and $0 < \delta(\mathbf{h}) < 1/4$ such that for any normalized $u \in M(\mathbf{h})$ with $J[u] \le \mathcal{J}(\mathbf{h}) + \epsilon_0$ we have $|u(t) - (-1)^i| > \delta(\mathbf{h})$ for all $t \in I_i,\ i = 0, \dots, |\mathbf{h}| + 1$.

This property asserts that minimizing sequences in the class $M(\mathbf{h})$ cannot gain complexity by forming new transitions, i.e. between two crossings of ± 1, functions with small enough action in $M(\mathbf{h})$ are uniformly bounded away from ∓ 1. The next lemma states that any class $M(\mathbf{h})$ with $\mathbf{h} \in E(\mathbf{g})$ satisfies this property. In Sect. 7 we will prove that if F is even, then all classes have this property.

Lemma 5.3. *Let* $\mathbf{g} \in G$. *Then there exists an* $\kappa > 0$ *such that for every* $\mathbf{h} \in E(\mathbf{g})$ *we have* $\mathcal{J}(\mathbf{k}) \geq \mathcal{J}(\mathbf{h}) + \kappa$ *for all* $\mathbf{k} \succ \mathbf{h}$. *Moreover, the class* $M(\mathbf{h})$ *satisfies the uniform separation property.*

Proof. Since $E(\mathbf{g})$ and $\{\mathbf{k} : \mathbf{k} \succ \mathbf{h}\}$ are both finite sets, there is $\kappa > 0$ such that $\mathcal{J}(\mathbf{k}) \geq \mathcal{J}(\mathbf{h}) + \kappa$ for all $\mathbf{k} \succ \mathbf{h}$ and all $\mathbf{h} \in E(\mathbf{g})$.

Before proving the uniform separation property, we define a family of functions $\psi_n(t) \in H^2(\mathbb{R})$ by

$$\psi_n(t) = \begin{cases} (t^2-1)^2 \cos(n\pi t) & \text{for } t \in [-1,1], \\ 0 & \text{for } t \notin [-1,1]. \end{cases} \tag{5.4}$$

The support and range of ψ_n are contained in $[-1, 1]$, and by a suitable scaling, the H^2 norm, the domain and range of ψ_n can all be made arbitrarily small. We will use various scalings of these functions in this proof and subsequent proofs to perturb functions in $\chi + H^2(\mathbb{R})$. In most cases there will be many different ways to accomplish such perturbations, but we will use this family of functions for specificity.

Now let $\epsilon_0 = \kappa/2$. We will show that the uniform separation property holds on the core interval with this ϵ_0 and some small $\delta > 0$. Suppose to the contrary that for any $\delta < 1/4$ there exists a normalized $u \in M(\mathbf{h})$ with $J[u] \leq \mathcal{J}(\mathbf{h}) + \epsilon_0$ and a point $t_0 \in I_i$ with $1 \leq i \leq m$ such that $|u(t_0) - (-1)^i| < \delta$. Let $v = u + 2\delta\psi_0((t-t_0)/\delta^{1/2})(-1)^i$. Then $v = u$ outside the interval of size $\delta^{1/2}$ around t_0. To specify δ, let $I = [a, b]$ be the largest interval containing t_0 for which $(-1)^i u \geq 0$, and choose $t_1 \in I$ such that $|u(t_1)| = 1/2$. The size of the interval I is uniformly bounded from below because

$$C\frac{(1/4-0)}{b-a} \leq C\frac{|u^2(t_1) - u^2(a)|}{t_1 - a} \leq J[u] \leq \mathcal{J}(\mathbf{h}) + \epsilon_0,$$

using the estimate (5.2). Now choose δ small enough so that the perturbation is restricted to the interval I. Therefore the function v has exactly two more transitions than u, since $|v(t_0) - u(t_0)| = 2\delta$. Furthermore $\|v - u\|_{H^2} < C\delta^{1/4}$. For δ sufficiently small (independently of u) we have $J[v] < J[u] + \epsilon_0/2$, because J is uniformly continuous with respect to perturbations with support on the finite interval $[t_0 - 1/2, t_0 + 1/2]$ by Lemma 5.1. Since u is normalized in $M(\mathbf{h})$, the point t_0 can be chosen to be the unique local extremum between two crossings. In this case the two new crossings created in v are transverse because of the form of the perturbation. Thus $v \in M(\mathbf{k})$ for some $\mathbf{k} \succ \mathbf{h}$ and $J[v] < J[u] + \epsilon_0/2$. Hence

$$\mathcal{J}(\mathbf{k}) \leq J[v] \leq J[u] + \frac{\epsilon_0}{2} \leq \mathcal{J}(\mathbf{h}) + \frac{3\epsilon_0}{2} \leq \mathcal{J}(\mathbf{h}) + \frac{3\kappa}{4},$$

which contradicts the choice of ϵ_0, since $\mathcal{J}(\mathbf{k}) \geq \mathcal{J}(\mathbf{h}) + \kappa$. This establishes the uniform separation property in the core interval.

Finally we will show that the uniform separation property holds in the tails. We begin with the following claim.

Claim. There exists $\kappa > 0$ such that if $\mathbf{k} = (2, 2, \mathbf{h})$ or $(\mathbf{h}, 2, 2)$ for any $\mathbf{h} \in G$, then $\mathcal{J}(\mathbf{k}) \geq \mathcal{J}(\mathbf{h}) + \kappa$.

Let u be any normalized function in $M(\mathbf{k})$, where $\mathbf{k} = (2, 2, \mathbf{h})$ as the other case is similar. Then the core interval of u begins with an interval I containing some number of transitions (at least three) with no extra oscillations around ± 1 between them, i.e. there are (at least) three maximal monotonicity intervals each contributing a transition. Let I be the union of the maximal monotonicity intervals of all these transitions. Then u does not have transitions on either side of I without first oscillating around ± 1. Let $t_{\max}$ be the location of the maximum value of u on I, and let $s_1 < t_{\max} < s_2$ be the locations of the adjacent local minima. Let $u(a_1)$ and $u(a_2)$ be the adjacent local maxima located to the left of s_1 and to the right of s_2 respectively. If $u(s_1) = u(s_2)$, then the interval $[s_1, s_2]$ can be clipped out of u removing two transitions. If $u(s_1) < u(s_2)$, then the intervals $[a_1, s_1]$ and $[t_{\max}, s_2]$ satisfy the hypothesis (ii) of Lemma 3.1. Note that one of the intervals $[a_1, s_1]$ and $[s_1, t_{\max}]$ need not be a transition, but in any case two transitions can be clipped from u. Note that s_1 must be at the bottom of a transition so that $u(s_1) < -1$. If $u(s_1) > u(s_2)$, then $u(s_2) < -1$, and hence $a_2 \in I$ so that $u(a_2) < u(t_{\max})$. Therefore u can be clipped using the intervals $[s_1, t_{\max}]$ and $[s_2, a_2]$. This clipping yields a $u^* \in M(\mathbf{h})$ such that $J[u^*]+\kappa \leq J[u]$ for some $\kappa > 0$ independent of u by Corollary 5.2. Therefore $\mathcal{J}(\mathbf{h}) + \kappa \leq J[u]$ for all $u \in M(\mathbf{k})$ which implies that $J(\mathbf{k}) \geq J(\mathbf{h}) + \kappa$. This completes the proof of the claim.

Now assume the uniform separation property fails in the left tail of u. Let t_0 be the largest point such that $u(t_0)$ is the global maximum of u on I_0. Arguing as in the proof of Lemma 3.6 each of the oscillations around -1 can be removed one-by-one so that there are exactly two crossings of -1 between t_0 and $\max A_0$. As above assume that $|u(t_0) - 1| < \delta < 1/4$. For δ sufficiently small depending only on $\max\{\mathcal{J}(\mathbf{k}) : \mathbf{k} = (2, 2, \mathbf{h})$ or $\mathbf{k} = (\mathbf{h}, 2, 2)$ for some $\mathbf{h} \in E(\mathbf{g})\}$, we can perturb u near t_0 exactly as above to obtain $v \in M(\mathbf{k})$ where $\mathbf{k} = (2, 2, \mathbf{h})$ with $J[v] \leq \mathcal{J}(\mathbf{h}) + 3\kappa/4$. This contradicts the claim. Hence the uniform separation property holds in $M(\mathbf{h})$ with ϵ_0 chosen above and some $\delta(h) < 1/4$. $\square$

The above results show that to minimize in the extended class $\mathcal{M}(\mathbf{g})$ one can minimize in a fixed class $M(\mathbf{h})$ for some $\mathbf{h} \in E(\mathbf{g})$, and that minimizing sequences in this class never gain complexity. New transitions cannot form in the limit of a minimizing sequence by Lemma 5.3, and the number of crossings between transitions is fixed by the class $M(\mathbf{h})$. The remainder of this section is primarily devoted to constructing minimizing sequences which do not lose complexity in the limit and which are bounded in $\chi + H^2$ to exploit the weakly lower semicontinuity of J. As we will show, this amounts to controlling how much time functions spend near ± 1 in their core interval, and hence we now need to examine what happens when functions are close to ± 1.

For the remainder of this section the constants $\mathbf{h} \in G,\ \epsilon_0 > 0,\ \delta > 0$, *and* $\tau_0 = \max\{\tau_0(F(u-1)), \tau_0(F(u+1))\} > 0$ *will be fixed so that the uniform separation property holds in* $M(\mathbf{h})$ *with* ϵ_0, δ *and so that the results of the local theory in Sect.* 4 *apply near both wells. In particular we choose* $\delta < \min\{\delta_0(F(u+1)), \delta_0(F(u-1)), \delta(\mathbf{h}), 1/4\}$.

Let $u \in M(\mathbf{h})$, and consider the interval $\mathcal{C}(u) = (c_1, c_2)$, where $c_1 = \inf\{t : |u(t) - 1| < \delta\}$ and $c_2 = \sup\{t : |u(t) + (-1)^m| < \delta\}$. We will call $\mathcal{C}(u)$ the δ**-core interval** of u. Observe that if u has the properties stated in the uniform separation condition, then its δ-core interval is contained in its core interval. Now let $S(u) = \{t : |u(t) \pm 1| < \delta\}$ be the set of all points for which u is in the strips of width δ around ± 1, and define $B[u] = |\mathcal{C}(u) \cap S(u)|$, where $|\cdot|$ denotes the Lebesgue measure. To control minimizing sequences for J in $M(\mathbf{h})$ we first analyze the minimization problem,

$$B_\epsilon = \inf\{B[u] : u \in M(\mathbf{h}),\ u \text{ is normalized, and } J[u] \leq \mathcal{J}(\mathbf{h}) + \epsilon\}, \tag{5.5}$$

for $\epsilon < \epsilon_0$. The purpose of this auxiliary minimization is contained in the following lemma.

Lemma 5.4. *For any normalized $u \in M(\mathbf{h})$ with $J[u] \leq \mathcal{J}[\mathbf{h}] + \epsilon_0$,*

$$\|u - \chi\|_{H^2} \leq C(\mathbf{h})\,(1 + J[u] + B[u])\,,$$

and $|\mathcal{C}(u)| \leq C(\mathbf{h}) \cdot J[u] + B[u]$.

Before proving this lemma we introduce some notation.

> *Due to the translation invariance of both J and B, we will always translate normalized functions in $M(\mathbf{h})$ so that $c_1 = 0$, and the δ-core interval is $\mathcal{C}(u) = (0, c_2)$.*

Note that $\|u'\|^2_{H^1} \leq CJ[u]$, and hence we must control $\|u - \chi\|_{L^2}$ to prove Lemma 5.4. Define $H_{-1} = \chi_{-1} \equiv -1$ and H_1 to be the Heavyside function, i.e. $H_1(t) = -1$ for $t < 0$ and $H_1(t) = 1$ for $t > 0$. Then $(H_{\pm 1} - \chi_{\pm 1}) \in L^2$, and it suffices to estimate $\|u - H_{(-1)^m}\|_{L^2}$, where $m = |\mathbf{h}|$. Since $|\mathbf{h}|$ determines $H_{\pm 1}$ for the entire class $M(\mathbf{h})$, we will drop the subscript on H as well as on χ.

Proof of Lemma 5.4. Using the uniform separation property, we can estimate

$$\begin{aligned}\|u - H\|^2_{L^2} &\leq \int_{-\infty}^{0} |u - (-1)|^2\,dt + \int_{\mathcal{C}(u)} |u - (-1)^m|^2\,dt + \int_{c_2}^{\infty} |u - (-1)^m|^2\,dt \\ &\leq \frac{1}{\eta^2}\int_{-\infty}^{0} F(u)\,dt + \int_{\mathcal{C}(u)} |u - (-1)^m|^2\,dt + \frac{1}{\eta^2}\int_{c_2}^{\infty} F(u)\,dt.\end{aligned}$$

where η satisfies property (H2) in Sect. 2 for both $\alpha = \delta - 1$ and $\alpha = (-1)^m(1-\delta)$. Since δ is fixed, η can also be chosen small enough so that $F(u) \geq \eta^2(u \pm 1)^2$ for $|u \pm 1| > \delta$ to obtain

$$\begin{aligned}\int_{\mathcal{C}(u)} |u - (-1)^m|^2\,dt &= \int_{\mathcal{C}(u)\cap S(u)} |u - (-1)^m|^2\,dt + \int_{\mathcal{C}(u)\setminus S(u)} |u - (-1)^m|^2\,dt \\ &\leq (2+\delta)^2 B[u] + \frac{1}{\eta^2}\int_{\mathcal{C}(u)\setminus S(u)} F(u)\,dt.\end{aligned}$$

Therefore combining these two estimates yields

$$\|u - H\|^2_{L^2} \leq \frac{1}{\eta^2}\int_{\mathbb{R}} F[u]\,dt + (2+\delta)^2 B[u] \leq C(J[u] + B[u]),$$

and the estimate of $\|u - \chi\|_{H^2}$ follows. The bound on the length of the δ-core interval follows immediately from the second inequality in Lemma 5.1. □

Theorem 5.5. *There exists a constant $K = K(\mathbf{h}, \tau_0)$ such that for every $\epsilon < \epsilon_0$ there is a normalized $u \in M(\mathbf{h})$ with $J[u] \leq \mathcal{J}(\mathbf{h}) + 2\epsilon$ and $B[u] \leq K$.*

Proof. Let $m = |\mathbf{h}|$. Choose a minimizing sequence $u_n \in M(\mathbf{h})$ for B_ϵ in the minimization problem (5.5) with a fixed $\epsilon < \epsilon_0$. Then $\|u_n - \chi\|_{H^2}$ is bounded by Lemma 5.4. Hence there exists $\widehat{u} \in \chi + H^2(\mathbb{R})$ such that u_n approaches $\widehat{u}$ weakly in $\chi + H^2(\mathbb{R})$, denoted $u_n \rightharpoonup \widehat{u}$, which implies in particular that $u_n \to \widehat{u}$ uniformly in C^1 on any compact interval. Since $J[u_n]$ and $B[u_n]$ are uniformly bounded, the points c_1, c_2 in the definition of the δ-core interval are finite for $\widehat{u}$, so $\mathcal{C}(\widehat{u})$ is a compact interval. Since the set $S(u)$ is defined to be the points where u is in the *open* strips around ± 1, it follows from Fatou's lemma that

$$|\mathcal{C}(\widehat{u}) \cap S(\widehat{u})| \le \liminf_{n\to\infty} |\mathcal{C}(u_n) \cap S(u_n)|,$$

and hence $B[\widehat{u}] \le \liminf_{n\to\infty} B[u_n]$. So $\widehat{u}$ satisfies the estimate in (5.5), and $B[\widehat{u}] \le B_\epsilon$.

Let

$$K(\mathbf{h}, \tau_0) = 2\left\{\tau_0\left(\max_{1\le i\le m} h_i + 2\right) + 2\right\}\sum_{i=1}^m h_i.$$

Note that the number of intervals of maximal monotonicity in the δ-core interval of any normalized function in $M(\mathbf{h})$ is exactly $\sum_i h_i - m + 1$. Since the sequence u_n is normalized and converges strongly in C^1_{loc}, the number of maximal monotonicity intervals in the δ-core of $\widehat{u}$ is bounded by $\sum_i h_i$. Likewise, $\widehat{u}$ has finitely many local extrema except for possibly intervals of local extrema. However, all intervals of critical points can be clipped to a point, which reduces both J and B. Therefore we can assume that $\widehat{u}$ has finitely many local extrema in its δ-core interval. The number of transitions is preserved in the limit at $m + 1$.

Now there are two possibilities: either $B[\widehat{u}] > K$ or $B[\widehat{u}] \le K$. In the first case, we will construct a normalized function $\widehat{v} \in M(\mathbf{h})$ such that

$$J[\widehat{v}] < J[\widehat{u}] \le \mathcal{J}(\mathbf{h}) + \epsilon, \quad \text{and} \quad B[\widehat{v}] < B[\widehat{u}] \le B_\epsilon,$$

which contradicts the definition of B_ϵ in (5.5) and therefore excludes this case. In the second case, we will construct a normalized function $\widehat{v} \in M(\mathbf{h})$ such that

$$J[\widehat{v}] < J[\widehat{u}] + \epsilon \le \mathcal{J}(\mathbf{h}) + 2\epsilon, \quad \text{and} \quad B[\widehat{v}] < B[\widehat{u}] \le K,$$

which means that $\widehat{u}$ can be pushed back into the class $M(\mathbf{h})$ at the expense of slightly increasing the action only. Then $\widehat{v}$ satisfies the requirements of the theorem. The construction of $\widehat{v}$ in both cases is very similar, and we now implement it in detail for the first case.

Step 1. Modify $\widehat{u}$ on an interval where J is not optimal, see Fig. 5.1. First, note that the number of components of $\mathcal{C}(\widehat{u}) \cap S(\widehat{u})$ is at most twice the number of maximal monotonicity intervals in $\mathcal{C}(\widehat{u})$ which is therefore at most $2\sum_i h_i$. Hence there must be a component I with $|I|$ larger than $K/2\sum_i h_i = \tau_0(\max h_i + 2) + 2$. This implies that I contains a subinterval $[a_1, a_2]$ with small boundary data $|\widehat{u}(a_i) \pm 1| < \delta,\ |\widehat{u}'(a_i)| < \delta$, and yet large length $|a_2 - a_1| > \tau_0(\max h_i + 2)$. Note that $\widehat{u}$ has at most $\max h_i$ crossings of ± 1 in $[a_1, a_2]$, and by Theorem 4.2 the global minimizer of J over the finite interval $[a_1, a_2]$ with these small boundary conditions has at least $\max h_i + 2$ crossings. Hence $\widehat{u}$ does not minimize J on the interval $[a_1, a_2]$. Therefore we can replace $\widehat{u}$ by the minimizer on $[a_1, a_2]$ to construct a new function $\widehat{v} \in \chi + H^2(\mathbb{R})$ with $J[\widehat{v}] < J[\widehat{u}] \le \mathcal{J}(\mathbf{h}) + \epsilon$. Also $B[\widehat{v}] \le B[\widehat{u}]$ since $[a_1, a_2] \subset \mathcal{C}(\widehat{u}) \cap S(\widehat{u})$.

Step 2. Modify $\widehat{v}$ so that $B[\widehat{v}] < B[\widehat{u}]$. Since I is a component of $\mathcal{C}(\widehat{u}) \cap S(\widehat{u})$, we know that $|\widehat{v} \pm 1| = \delta$ at the endpoints of I. Also $\widehat{v}$ is strictly monotone at the endpoints of I, since intervals of critical points have been clipped out. One can easily perturb $\widehat{v}$ near the endpoint of I so that $B[\widehat{v}] < B[\widehat{u}]$. Specifically, suppose $b = \inf I$ and $\widehat{v}$ is increasing at b, but possibly $\widehat{v}'(b) = 0$. Choose $b_1 < b < b_2$ such that $\widehat{v}'(b_i) > 0$ and $\widehat{v}$ is increasing on $[b_1, b_2]$. Define

$$\widehat{\psi}_\nu(t) = \begin{cases} \psi_0(\frac{t-b_1}{\nu}) & \text{for } t \leq b_1, \\ 1 & \text{for } b_1 \leq t \leq b_2, \\ \psi_0(\frac{t-b_2}{\nu}) & \text{for } t \geq b_2, \end{cases}$$

where ψ_0 is defined by formula (5.4). The support of $\widehat{\psi}_\nu$ is $[-\nu+b_1, b_2+\nu]$, and note that b_1, b_2 can be chosen arbitrarily close to b. Let $\widehat{w} = \widehat{v} \pm \nu^3 \widehat{\psi}_\nu(t)$. For $\nu > 0$ sufficiently small, this perturbation shrinks the component I of $\mathcal{C}(\widehat{v}) \cap S(\widehat{v})$ to the right of b so that $B[\widehat{w}] < B[\widehat{v}] \leq B[\widehat{u}]$. Also $J[\widehat{w}] < \mathcal{J}(\mathbf{h}) + \epsilon$ since J is uniformly continuous over the support of $\widehat{\psi}_\nu$ by Lemma 5.1, and $\widehat{w}$ has exactly the same number of local extrema as $\widehat{v}$ by the choice of b_i. Set $\widehat{v} = \widehat{w}$, and $\rho = B[\widehat{u}] - B[\widehat{v}] > 0$.

Step 3. Eliminate tangencies in $\widehat{v}$ at ± 1. The function $\widehat{v}$ is possibly not in any class $M(\mathbf{h})$ as defined in Sect. 2 because crossings of ± 1 could have coalesced into tangencies at ± 1 in the limit $u_n \rightharpoonup \widehat{u}$. No crossings are added due to the postulated uniform separation property. However, as will be described below, near each place where such a tangency occurs $\widehat{v}$ can be perturbed to add as many tiny oscillations as are necessary to put $\widehat{v}$ back into a class $M(\mathbf{k})$ with $k_i \geq h_i$ for $i \leq m$.

Suppose $\widehat{v}$ has N tangencies at ± 1 in its core interval. Each point of tangency, t_0, can be widened to an interval of length $\rho/2N$ on which $\widehat{v}$ is identically ± 1. Then $B[\widehat{v}] = B[\widehat{u}] - \rho/2$, but J is unchanged. For $0 < \nu < \rho/4N$ sufficiently small and $n = (\max h_i + 2)/2$, the scaled function $\nu^3 \psi_n((t - t_0)/\nu)$, defined by formula (5.4), can be added to $\widehat{v}$ on this interval to create at least $\max h_i + 2$ crossings of ± 1 in a neighborhood of t_0. This procedure can be done at each point of tangency. For ν small enough, we still have $J[\widehat{v}] < \mathcal{J}(\mathbf{h}) + \epsilon$ and $B[\widehat{v}] = B[\widehat{u}] - \rho/2$.

Step 4 Adjust crossings in the tails. Still the function $\widehat{v}$ is possibly not in any class as defined in Sect. 2 because all of the crossings in the tails may have been lost at $\pm\infty$ in the limit $u_n \to \widehat{u}$. The definition of the classes $M(\mathbf{h})$ require at least one crossing in each of the tails. However, since $\widehat{v} \in \chi + H^2(\mathbb{R})$, $\|(\widehat{v}(t), \widehat{v}'(t)) - (\pm 1, 0)\| \to 0$ as $|t| \to \infty$. We perturb $\widehat{v}$ by adding functions of the form $\pm \nu^3 \psi_0((t \pm t_0)/\nu)$ to each tail. For ν sufficiently small, there is a point t_0 sufficiently large such that two transverse crossings are added to each tail in $\widehat{v}$, and $J[\widehat{v}] < \mathcal{J}(\mathbf{h}) + \epsilon$ and $B[\widehat{v}] = B[\widehat{u}] - \rho/2$.

Step 5 Clip $\widehat{v}$ so that $\widehat{v}$ is normalized in $M(\mathbf{h})$, see Fig. 5.1. Finally we can perturb $\widehat{v}$ to make it a Morse function on its δ-core interval. Outside of I, the function $\widehat{v}$ has finitely many local extrema, and hence, for a small enough perturbation, B increases only slightly. Since I is a component of $\mathcal{C}(\widehat{u}) \cap S(\widehat{u})$, no perturbation in I can increase B beyond $B[\widehat{u}]$. Thus, for a small enough perturbation, $B[\widehat{v}] < B[\widehat{u}]$, and also $J[\widehat{v}] < \mathcal{J}(\mathbf{h}) + \epsilon$. Now $\widehat{v}$ can be normalized by clipping which possibly reduces but does not increase either B or J. Thus $\widehat{v}$ is a normalized function in $M(\mathbf{k})$ with $k_i \geq h_i$ for $i = 1 \leq m$. As in Lemma 3.6, crossings can be removed by clipping to put $\widehat{v}$ into $M(\mathbf{h})$. Again note that clipping can only reduce both B and J.

Therefore we have constructed a normalized $\widehat{v} \in M(\mathbf{h})$ with $J[\widehat{v}] < \mathcal{J}(\mathbf{h}) + \epsilon$ and $B[\widehat{v}] < B[\widehat{u}]$ which completes the first case described above. In the second case, we skip Step 1 and begin with $\widehat{v} = \widehat{u}$ and I any component of $\mathcal{C}(\widehat{u}) \cap S(\widehat{u})$ in Step 2. Since the action J has not been decreased (as in Step 1 of the previous case), Step 2 possibly results in a small increase in J and yields a function $\widehat{v}$ with

$$J[\widehat{v}] < \mathcal{J}(\mathbf{h}) + 2\epsilon \quad \text{and} \quad B[\widehat{v}] < B[\widehat{u}].$$

Then Steps 3–5 further modify $\widehat{v}$ to a normalized function in $M(\mathbf{h})$ conserving the above inequalities which are required by the theorem.

This completes the proof of the theorem. □

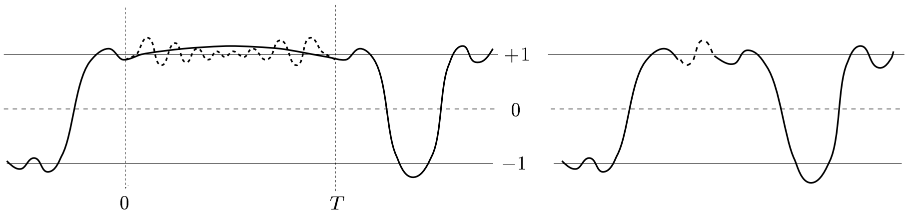

Fig. 5.1. The dashed curve represents the minimizer of J on the interval $[0, T]$ from Theorem 4.1 which oscillates by Theorem 4.2. After pasting in the dashed curve, the function is clipped to restore the correct number of crossings. See Steps 1 and 5 in the proof of Theorem 5.6. A similar technique is used to prove Theorem 2.2

Now we are in a position to prove Theorem 2.2 as stated in Sect. 2.

Proof of Theorem 2.2. Given $\mathbf{g} \in G$, consider any $\mathbf{h} \in E(\mathbf{g})$. By Lemma 5.3, the class $M(\mathbf{h})$ has the uniform separation property. Hence ϵ_0 and K are chosen so that Theorem 5.5 holds, and minimizing sequences for J in the minimization problem (5.3) can be chosen in the fixed class $M(\mathbf{h})$. Therefore we finish the proof with the following general statement.

Claim 5.6. If the uniform separation property holds in $M(\mathbf{h})$, then there exists $\widehat{u} \in M(\mathbf{h})$ which is a local minimizer of J and satisfies all of the properties listed in Theorem 2.2.

Theorem 5.5 implies that a minimizing sequence of normalized functions $u_n \in M(\mathbf{h})$ exists with the property that $B[u_n]$ is uniformly bounded. Hence $\|u_n - \chi\|_{H^2}$ is uniformly bounded by Lemma 5.4, and there exists $\widehat{u} \in \chi + H^2(\mathbb{R})$ such that $u_n \rightharpoonup \widehat{u}$. Since J is sequentially weakly lower semicontinuous, $J[\widehat{u}] \leq \mathcal{J}(\mathbf{h})$. We will now show that $\widehat{u}$ is in the class $M(\mathbf{h})$ and satisfies the desired properties.

Since $B[u_n]$ is uniformly bounded, the δ-core interval is uniformly bounded by Lemma 5.4, and hence $\widehat{u}$ has $|\mathbf{h}| + 1$ transitions. Recall that we have factored out translations by pinning down the left endpoint of the δ-core interval at the origin. First we would like to conclude that $\widehat{u} \in M(\mathbf{h})$ which would require that $\widehat{u}$ have the correct number $\sum_i h_i + 2$ of transverse crossings in its core interval. Note that $\widehat{u}$ cannot gain crossings due to the uniform separation property. Since $u_n \in M(\mathbf{h})$ are normalized and converge in C^1_{loc} to $\widehat{u}$, the only way crossings can be lost is if some number of them coalesce into a tangency at ± 1 or the crossings at the endpoints of the core interval approach $\pm\infty$. If neither of these situations occurs, then the intervals of maximal monotonicity in the core interval are preserved in $\widehat{u}$, and hence $\widehat{u} \in M(\mathbf{h})$.

First suppose that one of the endpoints of the core interval approaches $\pm\infty$ as $n \to \infty$. Then $\widehat{u}$ has no crossings in one or both of its tails, but $\lim_{t\to\pm\infty}(\widehat{u}(t), \widehat{u}'(t)) = (\pm 1, 0)$

since $\widehat{u} \in \chi + H^2(\mathbb{R})$. Therefore there is an interval I in the tail with length $|I| > \tau_0$ and $\|(\widehat{u}, \widehat{u}') - (\pm 1, 0)\| < \delta$ at the endpoints of I. As in Step 1 of the proof of Theorem 5.5, we can replace $\widehat{u}$ by the local minimizer of J on I with the appropriate boundary conditions (Theorem 4.1). By Theorem 4.2 this yields a function $\widehat{w}$ with at least one crossing in both tails and $J[\widehat{w}] < J[\widehat{u}]$.

Now suppose $\widehat{w}$ has a certain number of tangencies at ± 1. Each point of tangency p can be stretched to an arbitrarily long interval K_p by gluing in a long segment on which the function is identically ± 1. The resulting function has the same action as $\widehat{w}$ and will still be denoted by $\widehat{w}$. On a slightly larger interval L_p containing K_p the function $\widehat{w}$ has boundary data close but not equal to $(\pm 1, 0)$. Again as in Step 1 of the proof of Theorem 5.5, we can replace $\widehat{w}$ on L_p with the minimizer of the boundary value problem to obtain $\widehat{v}$ with reduced action $J[\widehat{v}] < J[\widehat{u}]$, see Fig. 5.1. By Theorem 4.2, we can assume each interval L_p contains at least max h_i crossings by choosing $|K_p| \geq \tau_0 \max h_i$. A small perturbation will make $\widehat{v}$ a Morse function with transverse crossings in the core interval. So $\widehat{v} \in M(\mathbf{k})$ for some $\mathbf{k} \in G$ such that $|\mathbf{k}| = |\mathbf{h}| = m$ and $k_i \geq h_i$ for all $i \leq m$. However $\mathcal{J}(\mathbf{k}) \leq J[\widehat{v}] < J[\widehat{u}] \leq \mathcal{J}(\mathbf{h})$ which contradicts Lemma 3.6. This proves that $\widehat{u} \in M(\mathbf{h})$, and hence $J[\widehat{u}] \geq \mathcal{J}(\mathbf{h})$. Therefore $\widehat{u}$ is a minimizer of J over $M(\mathbf{h})$.

Now that the existence of a $\widehat{u} \in M(\mathbf{h})$ which minimizes over the class $\mathcal{M}(\mathbf{g})$ has been established, we can characterize the behavior of $\widehat{u}$ in the tails. First, all crossings of ± 1 must be transverse by the above argument. Therefore the number of crossings is at most countable. As $t \to \pm\infty$, $\widehat{u}$ must have infinitely many crossings – otherwise there would be an arbitrarily long interval on which $\widehat{u}$ could be replaced by the local solution from Theorems 4.1 and 4.2 which reduces J and contradicts the fact that $\widehat{u}$ is a minimizer. Since $\widehat{u}$ is the weak H^2 limit (strong C^1_{loc} limit) of normalized functions $u_n \in M(\mathbf{h})$, between each pair of consecutive crossings there is exactly one local extremum, and the transitions are monotone. Recall that intervals of critical points are removed.

Finally, without loss of generality assume $\widehat{u} \to +1$ as $t \to \infty$ and consider the right tail, the other cases are similar. Let t_1 be the location of the first maximum to the right of the core interval, and suppose the absolute maximum in the right tail is located at $t_2 > t_1$. Let s_1 and s_2 be the locations of the local minima immediately preceding t_1 and t_2 respectively. Then $s_1 < t_1 < s_2 < t_2$ and Lemma 3.1 can be applied to the intervals $[s_1, t_1]$ and $[s_2, t_2]$ to clip $\widehat{u}$ which is a contradiction. A similar argument shows that the first minimum to the right of the core interval is the absolute minimum in the right tail. Repeating an analogous argument shows that the maxima are strictly decreasing and the minima are strictly increasing as $t \to \infty$.

This completes the proof of the claim and Theorem 2.2. $\square$

Proof of Corollary 2.3. The corollary follows immediately from the definition of the partial order $\prec$ since there is no $\mathbf{h} \in G$ such that $\mathbf{h} \succ \mathbf{g}$ when $g_i = 2$ for all $i \geq 1$. $\square$

6. Well-Separated Transitions

Now we would like to investigate in more detail the relationship between the infima over the various classes $M(\mathbf{g})$, and prove Theorem 1.2. In particular, we are interested in finding conditions either on the class or on the nonlinearity F which imply that a local minimizer over $\mathcal{M}(\mathbf{g})$ exists in the class $M(\mathbf{g})$ without increasing the complexity to some $M(\mathbf{h})$ with $\mathbf{h} \succ \mathbf{g}$. As will be proven in the next section, this can be done for all $\mathbf{g} \in G$ if F is even, because the symmetry provides an extra tool for clipping. Also this automatically holds for all $\mathbf{g}$ with $g_i = 2$ for all $i = 1, \ldots, |\mathbf{g}|$, since there is no $\mathbf{h} \in G$ with $\mathbf{h} \succ \mathbf{g}$ (Corollary 2.3).

In this section we provide additional arguments which yield results similar to those in more conventional multibump constructions, cf. [26]. Intuitively the basic idea is that the interaction between transitions should be weak if they are all separated by large distances. Thus, if the numbers of oscillations between all the transitions are large enough, i.e. $g_i \gg 1$ for all $i \geq 1$, one would expect that a local minimum should be attained in the class $M(\mathbf{g})$. We begin with the following lemma.

Lemma 6.1. *Suppose that $g_i > 2$ for all $i = 1, ..., m$. There exists a universal constant C_0 such that if $u \in M(\mathbf{g})$ with $J[u] \leq \mathcal{J}(\mathbf{g}) + 1$, then $J[u|_{I_i}] \leq C_0$.*

Proof. Fix $i \in \{1, ..., m\}$, and for specificity assume that it is odd. Let $u(a)$ be the last local maximum before $I_i = \operatorname{conv}(A_i)$ and $u(b)$ be the first local maximum after I_i. Since $g_{i\pm1} > 2$, we have $-1 < u(a), u(b) < 1$. Note that in the case $i = 1$, it is possible that there is no local maximum in the left tail with $-1 < u(a)$. However, in this case a small perturbation of u can be made to create such a point with $J(u) \leq \mathcal{J}(\mathbf{g}) + 1 + \epsilon$, which does not substantially effect the following argument. A similar situation can occur for $i = m + 1$.

Let $\xi_u(t)$ be defined by

$$\xi_u(t) = \begin{cases} (u(a)+1)\psi_0(t+3) + 2\psi_0(t+1) - 1 & \text{for } t \in [-3,-1] \\ 1 & \text{for } t \in [-1,1] \\ (u(b)+1)\psi_0(t-3) + 2\psi_0(t-1) - 1 & \text{for } t \in [1,3], \end{cases} \tag{6.1}$$

where ψ_0 is defined in (5.4). Clearly $J[\xi_u] < 8J[\psi_0]$. Note that $\xi_u(-3) = u(a),\ \xi_u(3) = u(b)$, and $\xi_u'(\pm3) = 0$. Also ξ_u has a "W" shape which can easily be adjusted via an arbitrarily small C^∞ perturbation to obtain $\widehat{\xi_u}$ with two transversal crossings of -1 followed by g_i transversal crossings of 1 and another two transversal crossings of -1. This perturbation can be accomplished exactly as in Step 3 in the proof of Theorem 5.5.

Now, replace u over the interval $[a, b]$ by $\widehat{\xi_u}$ (after adjusting the domain) to obtain a function $v \in M(\mathbf{g})$. Using $\mathcal{J}(\mathbf{g}) \leq J[v]$ and $J[u] \leq \mathcal{J}(\mathbf{g}) + 1$, we can estimate

$$-1 \leq J[v] - J[u] = J[\widehat{\xi_u}] - J[u|_{[a,b]}].$$

Hence

$$J[u|_{I_i}] \leq J[u|_{[a,b]}] \leq J[\widehat{\xi_u}] + 1 \leq 8J[\psi_0] + 1 = C_0,$$

which completes the proof. □

Lemma 6.2. *There exists $N > 0$ such that whenever $\mathbf{g} \in G$ with $g_i > N$ for all $i \leq |\mathbf{g}|$, the uniform separation property holds in $M(\mathbf{g})$.*

Proof. Let κ be the lower bound on the action of any transition from Corollary 5.2, and fix $\rho \leq \min\{\delta_0(F(u+1)), \delta_0(F(u-1)), \sqrt{\kappa/3C}, 1/4\}$, where δ_0 and C are as in Theorem 4.1.

Claim. There exists $N \geq 4$ such that for any $\mathbf{g} \in G$ satisfying the above conditions the following property holds: given any normalized $u \in M(\mathbf{g})$ with $J[u] \leq \mathcal{J}(\mathbf{g}) + 1$ and any interval $I_i,\ i \leq m$, which necessarily contains N consecutive crossings of $(-1)^{i+1}$, then $|u(t_0) - (-1)^{i+1}| < \rho$ for some $t_0 \in I_i$ at which u has a local extremum.

From Lemma 6.1, we have $J[u|_{I_i}] \leq C_0$. The set A of locations of local extrema of u in I_i has cardinality at least $N-1$. From Lemma 5.1 we obtain

$$J[u|_{I_i}] \geq \#(A \cap \{t \in I_i : |u(t) \mp 1| > \rho\})\, C\rho.$$

Combining these two estimates, $A \cap \{t \in I_i : |u(t) \mp 1| \leq \rho\}$ is nonempty for N sufficiently large. Choosing t_0 in this set proves the claim.

Assume the uniform separation property fails in $M(\mathbf{g})$ for $\epsilon_0 = \kappa/2$ and all $\delta > 0$. Then there exists a normalized $u \in M(\mathbf{g})$ with $J[u] \leq \mathcal{J}(\mathbf{g}) + \kappa/2$ and $s_0 \in I_i$ such that $|u(s_0) - (-1)^i| \leq \delta$ for some $i \leq m$. For specificity, we can assume $i = 1$, and without loss of generality we can take $u(s_0)$ to be the smallest local minimum on the interval between the zeroes of the first two transitions. Since $g_1 > N$, the interval I_1 contains a point t_0 satisfying the properties of the claim. First we will clip out the large oscillation near s_0, which lowers J by at least κ, and then we will insert crossings at t_0, which increases J by at most $C\rho^2$. Thus, for any sufficiently small $\delta > 0$, we will construct $u^* \in M(\mathbf{g})$ such that

$$\mathcal{J}(\mathbf{g}) \leq J[u^*] \leq \mathcal{J}(\mathbf{g}) - \frac{1}{12}\kappa,$$

which will yield a contradiction.

The insertion of crossings at t_0 is straightforward. First cut the function u at t_0. Then insert the minimizer of J on $[t_0, t_0 + T]$ with boundary conditions $(u(t_0), u'(t_0))$ at both ends. For T sufficiently large any number of crossings are added to u by Theorem 4.2 with an increase in the action of less than $C\rho^2$ by Theorem 4.1.

Clipping out the large oscillation near s_0 requires a more technical argument which is similar to the proof of Lemma 3.6. Let $s_1 < s_0 < s_2$ be the locations of the neighboring local maxima immediately adjacent to s_0. First assume that neither s_1 or s_2 is the endpoint of a transition, i.e. if a and b are the left/right endpoints of the maximal interval of monotonicity ending/beginning at s_1 and s_2 respectively, then $u(a), u(b) \geq u(s_0) > -1$. In this case the oscillation can be removed by clipping over the intervals $[a, s_1]$ and $[s_0, s_2]$ or $[s_1, s_0]$ and $[s_2, b]$ if $u(s_1) > u(s_2)$ or $u(s_1) < u(s_2)$ respectively. There is the trivial clipping if $u(s_1) = u(s_2)$.

A case where the above argument does not work is when $u(s_2) > u(s_1)$ and $[s_2, b]$ is a transition, i.e. $u(b) < -1$. The other case involving a transition is similar. Immediately to the right of b is either an interval with at least N crossings or the tail of u which has infinitely many transverse crossings because u is normalized. Let c and d be the locations of the first local maximum to the right of b and the adjacent local minimum to its right. Since there cannot be another transition immediately to the right of b, we know that $u(c) < 1$ and $u(d) < -1$. To clip we need $u(s_0) < u(c)$. If $u(s_0) > u(c)$, then since $|u(s_0) + 1| < \delta$, we can perturb u in a neighborhood of s_0 so that $-1 < u(s_0) < u(c)$. For δ small enough, this can be accomplished with at most $\kappa/12$ increase in the action.

Now the large oscillation can be removed by clipping over the intervals $[s_1, s_0]$ and $[c, d]$ which yields

$$\mathcal{J}(\mathbf{g}) \leq J[u^*] \leq J[u] - \kappa + C\rho^2 + \frac{1}{12}\kappa \leq \mathcal{J}(\mathbf{g}) - \frac{1}{12}\kappa.$$

This estimate provides the desired contradiction. □

Lemma 6.2 states that the uniform separation property holds in $M(\mathbf{g})$, and hence Claim 5.6 in the proof of Theorem 2.2 immediately yields the following theorem.

Theorem 6.3. *Let* $\mathbf{g} \in G$*. If* g_i *is sufficiently large for all* $i \leq |\mathbf{g}|$*, then there exists a solution* $\widehat{u} \in M(\mathbf{g})$ *of* (1.2) *which is a local minimizer of* J *restricted to* $M(\mathbf{g})$*.*

7. Symmetric Case

In Sect. 5 we found a local minimizer $\widehat{u}$ of J in every class $\mathcal{M}(\mathbf{g})$ which is a union of classes $M(\mathbf{h})$ with $\mathbf{h} \succeq \mathbf{g}$. A priori there is no guarantee that $\widehat{u}$ lies in the primary subset $M(\mathbf{g})$ unless $g_i = 2$ for all $i \leq |\mathbf{g}|$ (Corollary 2.3), or g_i are all large (Theorem 6.3). The purpose of this section is to prove that if F is even, then a local minimizer exists in every class $M(\mathbf{g})$. For example the result will apply to the EFK Eq. (1.3) in which the potential $F(u) = (u^2 - 1)^2/4$.

Theorem 7.1. *Suppose that* F *satisfies the hypothesis* (H1) *and* ± 1 *are saddle-foci. If* F *is* even*, then for each* $\mathbf{g} = (g_1, \ldots, g_m) \in G$ *there is a local minimizer* $\widehat{u} \in M(\mathbf{g})$ *which satisfies all the properties listed in Theorem* 2.2*.*

Remark. The original proof of this theorem contained the restriction that $g_i > 2$ for all i. An argument by J. VandenBerg [47], given as part of the proof of Lemma 7.4 below, allows this restriction to be removed.

The existence of $\widehat{u}$ will follow from Claim 5.6 in the proof of Theorem 2.2, after we have established that the uniform separation property holds in $M(\mathbf{g})$.

Proposition 7.2. *If* F *is even, then the uniform separation property holds in* $M(\mathbf{g})$ *for any* $\mathbf{g} = (g_1, \ldots, g_m) \in G$*.*

In this section we will assume that F is even, and we will need to consider all classes $M^{\pm}(\mathbf{g})$ as defined in the introduction, where $\lim_{t\to-\infty} u(t) = \pm 1$ for $u \in M^{\pm}(\mathbf{g})$. In the preceding sections we used only the classes $M(\mathbf{g}) = M^{-}(\mathbf{g})$, but the classes $M^{+}(\mathbf{g})$ are simply their reflections, $M^{+}(\mathbf{g}) = \{-u : u \in M(\mathbf{g})\}$. Since F is even, the infima in these classes are the same, $\mathcal{J}(\mathbf{g}) = \mathcal{J}^{-}(\mathbf{g}) = \mathcal{J}^{+}(\mathbf{g})$. Proposition 7.2 relies on two lemmas that allow for comparison of the infima $\mathcal{J}(\mathbf{g})$ for various $\mathbf{g}$ in a way similar to Lemma 3.6.

Lemma 7.3. *For any* $\mathbf{g} = (g_1, ..., g_m) \in G$*,*

$$\mathcal{J}(\mathbf{g}) \leq \mathcal{J}((g_1, ..., g_{i-1})) + \mathcal{J}((g_{i+1}, ..., g_m)) \textit{ for } 1 < i < m,$$
$$\mathcal{J}(\mathbf{g}) \leq \mathcal{J}(\mathbf{0}) + \mathcal{J}((g_2, \ldots, g_m)), \textit{ and } \mathcal{J}(\mathbf{g}) \leq \mathcal{J}((g_1, \ldots, g_{m-1})) + \mathcal{J}(\mathbf{0}).$$

Remark. As will be evident in the proof, estimates similar to those in Lemma 7.3 using $\mathcal{J}^{\pm}$ hold without requiring symmetry. However, the next lemma does assume F to be even. An example of a non-even F and a class $M(\mathbf{g})$ without a local minimizer is not known.

Lemma 7.4. *There is a universal constant* $\kappa = \kappa(\beta, F) > 0$ *such that*

i) $\mathcal{J}(\mathbf{0}) < \mathcal{J}((g_1)) - \kappa$ *for any* $g_1 \in 2\mathbb{N}$*,*
ii) $\mathcal{J}((g_2, ..., g_m)) < \mathcal{J}(\mathbf{g}) - \kappa$ *for* $\mathbf{g} = (g_1, ..., g_m) \in G$*,*
iii) $\mathcal{J}((g_1, ..., g_{m-1})) < \mathcal{J}(\mathbf{g}) - \kappa$ *for* $\mathbf{g} = (g_1, ..., g_m) \in G$*.*

First we will assume these lemmas to prove Proposition 7.2 which implies Theorem 7.1.

Proof of Proposition 7.2. Set $\epsilon_0 = \kappa$, where κ is as in Lemma 7.4 and Corollary 5.2. Take an arbitrary $u \in M(\mathbf{g})$ such that $J[u] \leq \mathcal{J}(\mathbf{g}) + \epsilon_0/2$, and assume, contrary to the assertion of the proposition, that $|u(t_0) - (-1)^i| < \delta$ for some $t_0 \in I_i,\ i \in \{0, 1, \ldots, m, m+1\}$, and small δ to be specified later. We will consider the case of i odd, as the other case is completely analogous. Let $g_i' := \#(A_i \cap (-\infty, t_0])$ and $g_i'' := \#(A_i \cap [t_0, \infty))$. Note that if $i = m+1$, then g_i'' is infinite. From Lemma 7.4, for $1 < i < m$ we have

$$\mathcal{J}((g_1, ..., g_{i-1})) + \mathcal{J}((g_{i+1}, ..., g_m)) \leq \mathcal{J}((g_1, ..., g_{i-1}, g_i')) + \mathcal{J}((g_i'', g_{i+1}, ..., g_m)) - 2\epsilon_0,$$

which has a natural counterpart for $i = m$

$$\mathcal{J}((g_1, \ldots, g_{m-1})) + \mathcal{J}(\mathbf{0}) \leq \mathcal{J}((g_1, \ldots, g_{m-1}, g_m')) + \mathcal{J}((g_m'')) - 2\epsilon_0,$$

and for $i = m+1$ we have

$$\mathcal{J}(\mathbf{g}) \leq \mathcal{J}((g_1, \ldots, g_m, g_{m+1}')) - \epsilon_0.$$

This will immediately contradict Lemma 7.3 if we prove the following inequality

$$\mathcal{J}((g_1, ..., g_{i-1}, g_i')) + \mathcal{J}((g_i'', g_{i+1}, ..., g_m)) - \epsilon_0 \leq \mathcal{J}(\mathbf{g}) \tag{7.1}$$

for $1 \leq i \leq m$ and

$$\mathcal{J}((g_1, \ldots, g_m, g_{m+1}')) - \epsilon_0 < \mathcal{J}(\mathbf{g}) \tag{7.2}$$

for $i = m+1$. The same simple geometric construction leads to both these inequalities, and we regret that our notation forced these rather cumbersome formulations.

First we will perturb u to a function v which has a tangency with $(-1)^i$. This perturbation will cause a small increase in J, which is uniform in u and dependent on δ, and is accomplished as follows. Let I be the largest interval around t_0 such that $(-1)^i u|_I \geq 0$. Exactly as in the proof of Lemma 5.4, $|I|$ is uniformly bounded from below because u can not grow too sharply without increasing J. Therefore, there is a uniform $\delta > 0$ for which the perturbations $v := u + \tau\psi_0((t - t_0)/\sqrt{\delta})$, $\tau \in [0, \delta]$, are localized to I, and hence all the crossings in u are preserved.

Choose $\tau \in [0, \delta]$ so that v is tangent to $(-1)^i$ at some $t_* \in I$. For δ sufficiently small $J[v] < \mathcal{J}(\mathbf{g}) + \epsilon_0$. Again, as in the proof of Lemma 5.4, uniformity in u comes from Lemma 5.1.

Cutting u at t_* and extending the two pieces by a constant ± 1 we obtain C^1 functions u_- and u_+ defined by

$$u_-(t) = \begin{cases} u(t) & \text{for } t \leq t_* \\ (-1)^{i-1} & \text{for } t \geq t_* \end{cases} \quad \text{and} \quad u_+(t) = \begin{cases} (-1)^{i-1} & \text{for } t \leq t_* \\ u(t) & \text{for } t \geq t_*. \end{cases}$$

Note that $J[u_+] + J[u_-] = J[v]$, hence $J[u_+] + J[u_-] < \mathcal{J}(\mathbf{g}) + \epsilon_0$. The two functions account for all the crossings in u. Therefore, if not for tangency that u_- and u_+ have in their tails, we would have $u_- \in M((g_1, ..., g_i'))$ and $u_+ \in M((g_i'', g_{i+1}, ..., g_m))$. Then the last inequality immediately yields the desired inequalities (7.1) and (7.2). The tangencies can be easily removed with arbitrarily small perturbations that leave the inequality intact. □

Now we proceed with the proofs of Lemmas 7.3 and 7.4.

Proof of Lemma 7.3. Fix $\epsilon > 0$ arbitrarily. We will argue for $0 < i < m$. Let $u_- \in M((g_1, ..., g_{i-1}))$ and $u_+ \in M((g_{i+1}, ..., g_m))$ be normalized such that

$$J[u_-] + J[u_+] \leq \mathcal{J}((g_1, ..., g_{i-1})) + \mathcal{J}((g_{i+1}, ..., g_m)) + \epsilon/2.$$

We will paste u_- and u_+ together to get $u \in M(\mathbf{g})$. For $s_1, s_2 > 0$ consider the function $v(t)$ defined on $\mathbb{R} \setminus [-1, 1]$ by

$$v(t) = \begin{cases} u_-(t + s_1) & \text{for } t < -1, \\ u_+(t - s_2) & \text{for } t > 1. \end{cases}$$

Since normalized functions have infinitely many transverse crossings which accumulate in the tails at points $-\infty \leq a < b \leq \infty$, we can choose s_1 and s_2 so that $v(\pm 1)$ is close to $(-1)^{i+1}$, and $v'(\pm 1)$ is very small. In particular v can be extended across $[-1, 1]$ (using the local solutions of Sect. 4 for example) to construct a C^1-function w with $J[w|_{[-1,1]}] < \epsilon/2$. Then

$$J[w] < \mathcal{J}((g_1, ..., g_{i-1})) + \mathcal{J}((g_{i+1}, ..., g_m)) + \epsilon,$$

and the construction can be accomplished so that w has transverse crossings of $(-1)^{i+1}$, i.e. $w \in M((g_1, ..., l, ..., g_m))$ for some $l > 0$. A number of crossings l is inherited by w from the tails of u_- and u_+. Thus l can be arbitrarily large since normalized functions have infinitely many oscillations in their tails. Assume then that $l \geq g_i$. The above estimate together with Lemma 3.6 yield

$$\mathcal{J}(\mathbf{g}) \leq J((g_1, ..., l, ..., g_m)) \leq \mathcal{J}((g_1, ..., g_{i-1})) + \mathcal{J}((g_{i+1}, ..., g_m)) + \epsilon.$$

Since ϵ was arbitrary, this completes the proof of the first inequality; the other two are proved analogously. □

Proof of Lemma 7.4. (i) From Lemma 3.6 we can assume that $g_1 = 2$. Let $u \in M((2))$ be normalized and translated so that $u(0)$ is the global maximum. Also assume that $J[u|_{(-\infty,0]}] \leq J[u|_{[0,+\infty)}]$, the other case is similar.

Let $u(a)$ be the local minimum preceding $u(0)$. Define v by $v(t) = u(t)$ for $t < 0$, and $v(t) = -u(-t)$ for $t > 0$. At $t = 0$ there is a jump with the one-sided limits $v(0^-) = u(0)$ and $v(0^+) = -u(0)$. Clearly, $(v(0^+) - v(a))(v(-a) - v(0^-)) = (u(0) + u(a))^2 \geq 0$, so $v|_{[a,0]}$ and $v|_{[0,-a]}$ are **latched**, i.e. the hypothesis (ii) of Lemma 3.1 is satisfied. Clipping v yields a function w which is in $M(\mathbf{0})$. Indeed, clipping merges the two transitions of v; one is included in w and the other is removed. In particular $J[w] \leq J[u] - \kappa$, where κ is the uniform (a priori) lower bound on the action of any transition given by Corollary 5.2 (Lemma 5.1). Since u is arbitrary, $\mathcal{J}((g_1)) \geq J[w] + \kappa \geq \mathcal{J}(\mathbf{0})$.

Statements (ii) and (iii) are completely analogous by considering the map $t \mapsto -t$, and therefore we will only prove (iii).

(iii) As before it is enough to consider the case when $g_m = 2$. Take an arbitrary normalized $u \in M(\mathbf{g})$. For specificity let us assume that m is even so that $u(+\infty) = 1$. Let $u(d)$ be the first local maximum in the tail, $u(b)$ be the last local maximum before the tail, and $u(c)$ be the unique local minimum in $[b, d]$, see Fig. 7.1. In this way $u|_{[b,c]}$ and $u|_{[c,d]}$ are the two last transitions. There are three possibilities.

Case 1a: $u(b) > -u(c)$ with $g_{m-1} > 2$. Let $u(a)$ be the local minimum preceding $u(b)$. Since $g_{m-1} > 2$, then $u(a) > -1$, and so $(-u(a) - u(b)) < 0$. Cutting u at $t = b$ and flipping the tail part leads to $v(t) := u(t)$ for $t \leq b$ and $v(t) := -u(t)$ for $t \geq b$. The

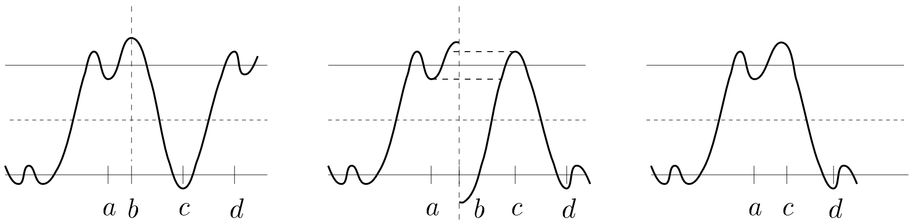

Fig. 7.1. An example of flipping and clipping in the proof of Lemma 7.4 (iii), Case 1

above inequality and the assumption guarantee that $v|_{[a,b]}$ and $v|_{[b,c]}$ are latched. Thus we can clip v, and the resulting function w, having lost one transition and none of the g_{m-1} crossings, belongs to $M((g_1, ..., g_{m-1}))$, see Fig. 7.1. Since a full transition was clipped out, $J[w] \leq J[u] - \kappa$, where κ is as in (i). Since u is arbitrary, we conclude that $\mathcal{J}((g_1, ..., g_{m-1})) \leq \mathcal{J}(\mathbf{g}) - \kappa$.

Case 1b [J. VandenBerg]: $u(b) > -u(c)$ with $g_{m-1} = 2$. If $u(b) \geq -u(a)$, the proof is similar to Case 1a, but notice that $a \in I_{m-2}$ and $u(a) < -1$.

If $u(b) < -u(a)$ and $g_{m-2} > 2$, then let $u(e)$ be the local maximum immediately preceding $u(a)$. Cut and flip u at $t = a$ (as described above) to obtain $v(t)$ for which $v|_{[e,a]}$ and $v|_{[a,b]}$ are latched. Clipping v then yields the desired function $w \in M((g_1, ..., g_{m-2}, 2)) = M((g_1, ..., g_{m-1}))$.

If $u(b) < -u(a)$ and $g_{m-2} = 2$, we define new points $\bar{a}, \bar{b}, \bar{c}$, and $\bar{d}$ as follows. Let $\bar{d} = b$ and $\bar{c} = a$, and let $u(\bar{b})$ be the local maximum preceding $u(\bar{c})$ and $u(\bar{a})$ be the local minimum preceding $u(\bar{b})$. Now $u(\bar{d}) = u(b) < -u(a) = -u(\bar{c})$. So either $-u(\bar{c}) > u(\bar{b})$, in which case we can use Case 3 below to obtain the desired result, or $-u(\bar{c}) < u(\bar{b})$. In the latter case, if $g_{m-3} > 2$, then applying Case 1a with the newly defined points yields the result, and if $g_{m-3} = 2$, we repeat the argument in Case 1b. This process terminates after finitely many steps.

Case 2: $u(d) > -u(c)$. Let $u(e)$ be the local minimum succeeding $u(d)$. Since $u(e) > -1$, we have $-u(e) - u(d) < 0$. As before, cut and flip u at d to get $v(t) = u(t)$ for $t \leq d$ and $v(t) = -u(t)$ for $t \geq d$. The above inequality and the assumption imply that $v|_{[c,d]}$ and $v|_{[d,e]}$ are latched, so clip v to w which has one less transition. Actually $w \in M((g_1, ..., g_{m-1}))$, and again $J[w] \leq J[u] - \kappa$. Consequently, $\mathcal{J}((g_1, ..., g_{m-1})) \leq \mathcal{J}(\mathbf{g}) - \kappa$.

Case 3: $-u(c) \geq \max\{u(b), u(d)\}$. Cut and flip the tail at c to obtain $v(t) = u(t)$ for $t \leq c$ and $v(t) = -u(t)$ for $t \geq c$. The functions $v|_{[b,c]}$ and $v|_{[c,d]}$ are latched; clip out one transition and finish the argument exactly as in the previous case. □

8. Concluding Remarks

In this section we will discuss some related observations and possible directions for future work which are worthy of mention. However, we will be brief and omit details.

8.1. Stable states for fourth-order evolution equations. In relation to the study of phase transitions in the neighborhood of Lifshitz points the following partial differential equation has been proposed in [18, 19, 50]:

$$u_t = -\gamma u_{xxxx} + \beta u_{xx} - F'(u), \tag{8.1}$$

where F is a double-well potential. In the case that the spatial domain is all of $\mathbb{R}$, the local minima of J are weakly stable states of (8.1) in the following sense. Let $\widehat{u}$ be a local minimum of J, and define $\widehat{u}_s$ to be the curve $\widehat{u}(\cdot - s)$, $s \in \mathbb{R}$. From [12, 26] it follows (at least in the analytic case) that these curves are isolated in $\chi + H^2(\mathbb{R})$. Now let $u_0 \in B_\epsilon(\widehat{u}_s)$ for ϵ sufficiently small. Then, since $\widehat{u}$ is an isolated local minimum of J, the flow $u(t,\cdot)$ generated by (8.1) remains in some tubular neighborhood $B_{\epsilon_0}(\widehat{u}_s)$, with $\epsilon_0 \geq \epsilon$. Furthermore, for arbitrary analytic F, the kernel of $J''(\widehat{u})$ satisfies $1 \leq \dim N(J''(\widehat{u})) \leq 2$. When $\dim N(J''(\widehat{u})) = 1$, which corresponds to the transverse intersection of stable and unstable manifolds, it follows from [6] that $u(t, x; u_0)$ approaches an element $\widehat{u}(\cdot - s_0)$ on the curve $\widehat{u}_s$ for some s_0. Otherwise it is not clear whether the ω-limit set of the solution $u(t, x; u_0)$ is a singleton.

In the case of a finite domain, the local minimizers found in this paper provide stable equilibrium solutions with complex spatial patterns to initial boundary value problems of the form

$$\begin{cases} u_t = -\gamma\epsilon^4 u_{xxxx} + \beta\epsilon^2 u_{xx} - F'(u) & \text{for } x \in (0,1), \\ (u(t,0), u'(t,0)) = A \quad \text{and} \quad (u(t,1), u'(t,1)) = B, & \\ u(0,x) = u_0(x). & \end{cases} \tag{8.2}$$

Indeed, let $\widehat{v} \in M(\mathbf{g})$ be a local minimizer of J, and let $[0, L]$ be the core interval of $\widehat{v}$. Truncating $\widehat{v}$ to this interval and rescaling by $\epsilon = 1/L$ yields a function $v : [0,1] \to \mathbb{R}$ which is a local minimizer of

$$J_\epsilon[u] = \int_0^1 \left[\frac{\gamma\epsilon^4}{2} u_{xx}^2 + \frac{\beta\epsilon^2}{2} u_x^2 + F(u) \right] dx$$

over the space $X = \{u \in H^2[0,1] : (u(0), u'(0)) = A \text{ and } (u(1), u'(1)) = B\}$ where A and B are the values of $(\widehat{v}, \widehat{v}')$ at the points $0, L$. It is easily verified that J_ϵ is a Lyapunov function for the flow generated on X by the initial boundary value problem, and hence v is an asymptotically stable equilibrium. The function v has the complexity specified by $\mathbf{g}$. Note that the values of A, B, and ϵ are determined by the particular minimizer on $\mathbb{R}$, and $\epsilon \to 0$ as $\sum g_i \to \infty$.

This is reminiscent of a result due to Afraimovich, Babin and Chow [1] who study second-order parabolic systems of the form

$$q_t = q_{xx} - \nabla V(q), \quad q : [0,T] \to \mathbb{R}^2,$$

derived from potentials V with two infinite spike-like singularities at $z_1, z_2 \in \mathbb{R}^2$. This makes the configuration space nontrivial homotopically (unlike our case with wells at $(\pm 1, 0)$). They prove that homotopy classes of initial data are invariant under this gradient flow with periodic boundary conditions, which yields stable equilibria in every homotopy class of $\mathbb{R}^2 \setminus \{z_1, z_2\}$. This invariance of classes will not occur for the fourth-order PDE (8.2), but we do obtain spatially complex stable equilibria.

The dynamics near the attractor is governed by the motion of transition layers in solutions to (8.2), see Kalies, VanderVorst, and Wanner [27]. Sandstede [41] has developed a general theory for studying the dynamics of higher-order parabolic PDE's in one space dimension. If the single pulse heteroclinics in $M(\mathbf{0})$ are transverse, then his results apply to (8.2) and describe the dynamics near solutions with well-separated transitions such as those in Theorem 1.2 (or 6.3).

The influence of saddle-focus equilibria in stationary or traveling-wave equations for a variety of fourth-order PDE's has been discussed by many authors, cf. [7, 8] and the references therein, see also Remark 7 below.

8.2. Multiple-well potentials. In the previous sections we discussed only equal depth double-well potentials $F(u)$. However the theory developed there easily extends to the case when F has arbitrarily many wells all of equal depth and all global minima. As in the case of two wells, topological classes of functions can be defined with limits at $t = \pm\infty$ in one of the wells. Naturally all of these equilibrium solutions are required to be of saddle-focus type. The necessary clipping procedures are identical to those in this paper, and the analysis proceeds in a completely analogous way. Again by comparing infima one can conclude that there are local minimizers in many of these topological classes. Potentials F with infinitely many wells can also be considered, cf. [26].

8.3. Critical levels for large winding vectors. For vectors $\mathbf{g} \in G$ for which the components g_i are large, a result more detailed than Theorem 6.3 can be obtained. In particular precise estimates on the critical levels of the minimizers can be found. There exist positive constants k and $N = N(|\mathbf{g}|)$ such that if $\min g_i \geq N$, then

$$|\mathcal{J}(\mathbf{g}) - n\mathcal{J}^-(\mathbf{0}) - m\mathcal{J}^+(\mathbf{0})| = O(e^{-k\ \min g_i}), \quad n + m = |\mathbf{g}|.$$

As $\min g_i \to \infty$, the critical levels cluster around sums of the global minima $\mathcal{J}^-(\mathbf{0})$ and $\mathcal{J}^+(\mathbf{0})$. Moreover the minimizers are $W^{1,\infty}$-close to $\sum_{i\leq|\mathbf{g}|+1} \widehat{u}_i(\cdot - t_i)$, where u_i are minimizers of $\mathcal{J}^-(\mathbf{0})$ for i odd and $\mathcal{J}^+(\mathbf{0})$ for i even. The points t_i satisfy $t_1 < t_2 < \ldots < t_{|\mathbf{g}|+1}$ and $|t_i - t_{i-1}| \to \infty$ as $\min g_i \to \infty$. The proof is based on the following observation. Recall $I_i = \operatorname{conv} A_i$ are the intervals between transitions. Let $u \in M(\mathbf{g})$ be a minimizer, then $|I_i| \to \infty$ as $g_i \to \infty$ for all i, which follows from the proof of Theorem 2.8 in [26]. The exponential bound in the above estimate follows from the fact that ± 1 are hyperbolic equilibria. A detailed proof of this estimate from below is completely analogous to arguments found in [27]. Note that the solutions constructed in [26], which lie in a $W^{1,\infty}$-neighborhood of $\sum_i [u_-(\cdot - s_{2i+1}) + u_+(\cdot - s_{2i})]$, $s \to \infty$ also have the property that the critical value goes to $n\mathcal{J}^- + m\mathcal{J}^+$. Also the uniform a priori bound on all bounded solutions established in the appendix of [28] allows the constant N in the above estimate to be chosen independently of $|\mathbf{g}|$.

8.4. Non-integrability and postive entropy. The multitransition solutions found in this paper can be used to estimate the topological entropy of the local flow generated by (1.2). Such a computation is performed in [26] for the EFK Eq. (1.3), where $F(u) = \frac{1}{4}(u^2 - 1)^2$. A similar argument can be made for a general double-well potential using the solutions in Theorem 1.2 provided one can find estimates on the distances between transitions. These estimates can be done, but we will not furnish them here. Since the topological entropy is positive, the flow is chaotic. Consequently Eq. (1.2), with any C^2 double-well potential F, is not a completely integrable Hamiltonian system when $u = \pm 1$ are saddle-focus equilibria points. In fact, the above results could possibly be extended to produce chaotic solutions with infinitely many transitions directly by a minimization procedure (see [25]).

8.5. Connections of higher index. The homoclinic and heteroclinic solutions discussed here are local minimizers of J. Therefore one would expect to also find critical points of higher index such as mountain pass critical points. If the minimizers were known to be nondegenerate, i.e. the translation eigenvalue is simple (which is difficult to verify), then many of the critical points of arbitrarily high index can be derived from work by Sandstede [41]. Recent numerical results concerning such solutions can be found in VandenBerg [48]. A possible variational approach can be illustrated by the following example. The shooting methods of Peletier and Troy [36] yield heteroclinics which oscillate around 0 on an interval $[-T, T]$ with amplitude less than 1. Such functions are in the class $M(\mathbf{0})$ but are not the local minimizers from Theorem 2.2, because their transitions are not monotone. This does not preclude them from being local minimizers of J, but they cannot be global minimizers in $\chi + H^2(\mathbb{R})$, since they can be clipped. However we conjecture that there are solutions of this type which are mountain pass critical points between local minima. Let $u_0 \in M(\mathbf{0})$ and $u_1 \in M((2, 2))$ and consider the connection $u_s = \{(1-s)u_0 + su_1\}_{s\in[0,1]}$. Then for some $s^* \in [0, 1]$, the function u_{s^*} has this property. It is possible that the techniques used in this paper can be extended to find a minimax of mountain pass type in the class of odd functions between the solutions u_0 and u_1 in this way.

8.6. Different types of Lagrangians. The Lagrangians that we consider in this paper are bounded from below by zero which allows for minimization. There are related problems for which such a formulation is not immediately applicable. However, the clipping and pasting techniques developed here could be useful in the application of other variational methods to the study of fourth-order equations with saddle-foci.

In this context we mention two equations. One class of problems comes from nonlinear optics [2], where the following equation arises: $u'''' - \alpha u'' - u^3 + u = 0$. In this case the Lagrangian density is not bounded from below, and one can find a homoclinic connection to the origin as a mountain pass, and the origin is a saddle-focus for $-2\sqrt{2} < \alpha < 2\sqrt{2}$. A similar situation occurs in (8.1) when the wells of F do not have equal depth. One form of the stationary Swift-Hohenberg or Mizel equation is $u'''' + \alpha u'' + u^3 - u = 0$, $\alpha > 0$ (see e.g. [29]), and the points $u = \pm 1$ are saddle-foci when $0 < \alpha < 2\sqrt{2}$. Again the action functional is not bounded below which causes problems in minimization. The techniques of this paper might be extendable to find higher types of minimax critical points in these types of equations.

Acknowledgement. The authors thank their colleagues at the Center for Dynamical Systems and Nonlinear Studies for their support and encouragement. In particular, we also thank J.B. VandenBerg and the referee for the improvement of Theorem 7.1 and the referee for his careful reading of the manuscript.

References

1. Afraimovich, V., Babin, A. and Chow, S.-N.: Spatial chaotic structure of attractors of reaction-diffusion systems. Trans. Am. Math. Soc. **348**, 5031–5063 (1996)
2. Akhmediev, N.N., Buryak, A.V., and Karlsson, M.: Radiationless optical solutions with oscillating tails. Optics Comm. **110**, 540–544 (1994)
3. Alexander, J.C., Gardner, R.A., and Jones, C.K.R.T.: Existence and stability of asymptotically oscillatory double pulses. J. Reine Angew. Math. **446**, 49–79 (1994)
4. Ambrosetti, A. and Coti-Zelati, V.: Multiplicité des orbites homoclines pour des systèm conservatifs. C. R. Acad. Sci. Paris **314** II, 601–604 (1992)
5. Amick, C.J. and Toland, J.F.: Homoclinic orbits in the dynamic phase-space analogy of an elastic strut. Eur. J. Appl. Math. **3**, 97–114 (1992)
6. Bates, P. and Jones, C.K.R.T.: Invariant manifolds for semilinear partial differential equations. Dynamics Reported **2**, 1–38 (1989)

7. Belyakov, L.Y. and Shil'nikov, L.P.: On the complex stationary nearly solitary waves. In: *Self-Organisation, Autowaves and Structures Far from Equilibrium*. Berlin–Heidelberg–New York: Springer-Verlag, 1984, pp. 106–110
8. Belyakov, L.Y. and Shil'nikov, L.P.: Homoclinic curves and complex waves. Selecta Math. Sov. **9**, 214–228 (1990)
9. Brezis, H.: *Analyse Functionnelle, Theorie et Applications*. Masson: Paris, 1988
10. Buffoni, B.: Multiplicity of homoclinic orbits in fourth-order conservative systems. In: *Variational and Local Methods in the Study of Hamiltonian Systems*, River Edge, NJ: World Sci Publishing, 1995, pp. 129–136
11. Buffoni, B., Champneys, A.R. and Toland, J.F.: Bifurcation and coalescence of a plethora of homoclinic orbits. J. Dyn. Diff. Eq. **8**, 221–279 (1996)
12. Buffoni, B. and Séré, E.: A global condition for quasi-random behavior in a class of conservative systems. Comm. Pure Appl. Math. **49**, 285–305 (1996)
13. Buffoni, B. and Toland, J.F.: Global existence of homoclinic and periodic orbits for a class of autonomous Hamiltonian systems. J. Diff. Eq. **118**, 104–120 (1995)
14. Champneys, A.R.: Subsidiary homoclinic orbits to a saddle-focus for reversible Hamiltonian systems. Intl. J. Bifurcation and Chaos **4**, 1447–1482 (1994)
15. Champneys, A.R. and Toland, J.F.: Bifurcation of a plethora of multi-modal homoclinic orbits for autonomous Hamiltonian systems. Nonlinearity **6**, 665–772 (1993)
16. Coti-Zelati, V., Ekeland, I. and Séré, E.: A variational approach to homoclinic orbits in Hamiltonian systems. Math. Ann. **288**, 133–160 (1990)
17. Coti-Zelati, V. and Rabinowitz, P.H.: Homoclinic orbits for second order Hamiltonian systems possessing superquadratic potentials. J. AMS **4**, 693–727 (1991)
18. Coullet, P., Elphick, C., and Repaux, D.: The nature of spatial chaos. Phys. Rev. Lett. **58**, 431–434 (1987)
19. Dee, G.T. and van Saarloos, W.: Bistable systems with propagating front leading to pattern formation. Phys. Rev. Lett. **60**, 2641–2644 (1988)
20. Devaney, R.L.: Homoclinic orbits in Hamiltonian systems. J. Diff. Eq. **21**, 431–438 (1976)
21. Felmer, P.: Heteroclinic orbits for spatially periodic Hamiltonian systems. Ann. Inst. H. Poincaré – Anal. Nonl. **8**, 477–497 (1991)
22. Gardner, R.A. and Jones, C.K.R.T.: Traveling waves of a perturbed diffusion equation arising in a phase field model. Indiana Univ. Math. Journ. **38**, 1197–1222 (1989)
23. Guckenheimer, J. and Holmes, P.: *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Fields*. Number 42 in Appl. Math. Sciences. Berlin–Heidelberg–New York: Springer-Verlag, 1982
24. Hornreich, R.M., Luban, M. and Shtrikman, S.: Critical behavior at the onset of k-space instability at the λ line. Phys. Rev. Lett. **35**, 1678 (1975)
25. Kalies, W.D., Kwapiesz, J., VandenBerg, J.B., VanderVorst, R.C.M.: Homotopy classes for stable periodic and random patterns in fourth order Hamiltonian systems. In preparation
26. Kalies, W.D. and VanderVorst, R.C.A.M.: Multitransition homoclinic and heteroclinic solutions of the extended Fisher-Kolmogorov equation. J. Diff. Eq. **131**, 209–228 (1996)
27. Kalies, W.D., VanderVorst, R.C.A.M. and Wanner, T.: Slow motion in higher-order systems and Γ-convergence in one space dimension. CDSNS Report Series 97–259 (1997)
28. Kwapisz, J.: Uniqueness of the stationary wave for the extended Fisher–Kolmogorov equation. Preprint, 1997
29. Leizarowitz, A. and Mizel, V.: One dimensional infinite-horizon variational problems arising in continuum mechanics. Arch. Rational Mech. Anal. **106**, 161–194 (1989)
30. Lin, X.-B.: Using Melnikov's method to solve Silnikov's problems. Proc. Royal Soc. Edin. **116A**, 295–325 (1990)
31. Melnikov, V.K.: On the stability of the center for time periodic perturbations. Trans. Moscow Math. Soc. **12**, 1–57 (1963)
32. Nishiura, Y.: Coexistence of infinitely many stable solutions to reaction- diffusion systems in the singular limit. Dynamics Reported **3**, 25–103 (1994)
33. Peletier, L.A. and Troy, W.C.: Spatial patterns described by the extended Fisher-Kolmogorov equatio: Kinks. Differential Integral Equations **8**, 1279–1304 (1995)
34. Peletier, L.A. and Troy, W.C.: Spatial patterns described by the extended Fisher-Kolmogorov equation: Periodic solutions. Preprint 1995
35. Peletier, L.A. and Troy, W.C.: Chaotic spatial patterns described by the extended Fisher-Kolmogorov equation. J. Diff. Eq. **129**, 458–508 (1996)

36. Peletier, L.A. and Troy, W.C.: A topological shooting method and the existence of kinks of the extended Fisher-Kolmogorov equation. Topol. Methods Nonlinear Anal. **6**, 331–355 (1997)
37. Peletier, L.A., Troy, W.C. and VanderVorst, R.C.A.M.: Stationary solutions of a fourth-order nonlinear diffusion equation. Differentsialnye Uravneniya **31**, 327–337 (1995)
38. Rabinowitz, P.H.: Periodic and heteroclinic solutions for a periodic Hamiltonian system. Ann. Inst. H. Poincaré – Anal. Nonl. **6**, 331–346 (1989)
39. Rabinowitz, P.H.: Homoclinic and heteroclinic orbits for a class of Hamiltonian systems. Calc. Var. **1**, 1–36 (1993)
40. Rottschäfer, V. and Doelman, A.: On the transition from the Ginzburg-Landau equation to the extended Fisher-Kolmogorov equation. Preprint (1997)
41. Sandstede, B.: Interaction of pulses in dissipative systems. In preparation (1997)
42. Sandstede, B. Stability of multiple-pulse solutions. To appear in Trans. Am. Math. Soc. 1997
43. Séré, E. Existence of infinitely many homoclinic orbits in Hamiltonian systems. Math. Ann. **209**, 27–42 (1992)
44. Séré, E.: Looking for the Bernoulli shift. Ann. Inst. H. Poincaré – Anal. Nonl. **10**, 561–590 (1993)
45. Shil'nikov, L.P.: The existence of a denumerable set of periodic motions in four-dimensional space in an extended neighborhood of a saddle-focus. Soviet Math. Dokl. **8**, 54–58 (1967)
46. Turaev, D.V. and Shil'nikov, L.P. On Hamiltonian systems with homoclinic saddle curves. Soviet Math. Dokl. **39**, 165–168 (1989)
47. VandenBerg, J.B.: Private communication 1997
48. VandenBerg, J.B.: Branches of heteroclinic solutions in fourth-order bi-stable systems. In preparation (1997)
49. VandenBerg, J.B.: Uniqueness for solutions of the extended Fisher–Kolmogorov equation. To appear in Comptes Rendus
50. Zimmerman, W.: Propagating fronts near a Lifshitz point. Phys. Rev. Lett. **66**, 1546 (1991)

Communicated by J. L. Lebowitz