EFFECTIVE DYNAMICS OF NONLINEAR EVOLUTIONS

JINGXUAN ZHANG

1. Introduction

Consider an abstract evolution equation

$$\partial_t u = -J^{-1}E'(u). \tag{1}$$

Here $u(t) \in X$ is a path in some closed set X in a real Hilbert space H. $E : H \to \mathbb{R} \cup \{\infty\}$ is a functional, which is finite and smooth on X. $E'(u) \in H$ is the gradient (w.r.t. the inner product of H) of E at u. The operator $J : T_u^* X \to T_u X$ is either the identity or a symplectic operator, which respectively turns (1) into either a first-order energy dissipative dynamics, or a Hamiltonian system. In the first case we take $H = H^s_{\text{loc}}$, consisting of functions from $\mathbb{R}^n \to \mathbb{R}^m$ whose restriction on any domain $\Omega \subset \mathbb{R}^n$ belongs to the Sobolev space $H^s(\Omega, \mathbb{R}^m)$ of order $s > 0$. In the latter case, we take $H = H^s_{\text{loc}} \times H^r_{\text{loc}}$, whence J can be written as the standard block matrix. We are interested in the following problem:

> Given a low energy intial configuration $u_0 \in H$ for (1), which is stable and spatially concentrated in some subset of its domain, can one reduce the system (1) to an evolution of the concentration set?

Such low energy, stable, and localized configurations are called solitons. Depending on the specific context, one can make different definitions for these three qualitative descriptions. The reduced evolution we seek is the *effective dynamics* of the full evolution (1). The point is that so long as a path $u(t) \in X$ solving (1) remains solitonic, (1) does not contain much information away from the concentration sets of $u(t)$. In this study we propose an abstract scheme to derive effective dynamics for (1).

Essentially, our method is a nonlinear perturbation theory. This is known in the physics literature as *adiabatic approximation*. The idea dates back at least to Manton's moduli space approximation scheme for monopole dynamics [Man82]. Rigorous results along this line include [Stu94a, Stu94b, OS98a, OS98b, GS06, DS09, Tin10, CFS18], with diverse applications to superfluidity, superconductivity, particle physics, and geometric flow. The point in common of these applications is that solitonic equilibria of (1) arise naturally due to focusing nonlinearity, or some kind of fixed content constraint.

Our argument goes roughly as follows: First, we construct a parametrized family of configurations $u_\sigma \in X, \sigma \in \Sigma$, where Σ is the space of all possible concentration sets. The space Σ is finite-dimensional if the concentration sets are finitely many points. In general Σ is a space of geometric objects. Each v_σ is an approximate minimizer of the energy E, so that $E'(v_\sigma)$ is uniformly close to the zero function, and the Hessian (i.e. the linearized operator of $E'(u)$ at u) is in some sense positive. We call this family v_σ the *approximate solitons*. Next, we find an evolution equation $\partial_t \sigma = F(\sigma)$ for a path σ with the following property: If $u(t)$ solves (1) with an intitial configuration close to an approximate soliton v_{σ_0}, then $u(t)$ stays close to the path of approximate solitons $v_{\sigma(t)}$, with $\sigma(t)$ generated by the equation for σ from some σ_1 close to σ_0. This equation for σ is the effective dynamics for (1). The function $F(\sigma)$ can be calculated explicitly, provided one has precise information of the parametrization $\sigma \mapsto u_\sigma$ as well as its Fréchet derivative. This argument also shows that the evolution for σ is actually independent of the parametrization. In this step we use the properties of u_σ as approximate energy minimizer. Lastly, we show the converse holds, in the sense that for a path $v_{\sigma(t)}$ as above, we can find a uniformly small correction $w(t)$ s.th. $v(t) + w(t)$ solves (1).

We refer the readers to an excellent review on specific applications of the adiabatic approximation to classical field theory [Stu07]. We believe however the present abstract formulation is new. Indeed, we list three conditions (6)-(8) under which addiabatic approximation is valid. One of our main points is that these conditions do not require a large space of equilibria of (1). We can do with just approximate ones. This has a particular advantage when one is interested in multi-soliton dynamics, for which approximation through equilibria can only be done with satisficatory validity within a certain range of parameters that comes with the problem, which is close to a critical value variously known as the *Bogomolnyi* or *self-dual* regime. The problem at this critical parameter tends to assume an unusually large space of equilibria, which can then serve as the space of approximate solitons for the problem with near-critical parameter. Heuristically, this is because the space of equilibria changes continuously as the parameter in the problem varies. The aforementioned review [Stu07] surveys some key problems that fall into this category.

The validity of approximation is determined by how close the approximate solitons are to the exact ones. We isolate a constraint (18), which shows how the conditions (6)-(8) affect the time interval on which the adiabatic approximation is valid. In particular, (18) shows that as the approximation space tends to the space of equilibria, the validity of approximation improves. This then agrees with the known results [Stu94a, Stu94b, DS09] in near Bogolmonyi regime.

This is largely motivated by a seminal work of Gustafson and Sigal [GS06], on the motion of magnetic Ginzburg-Landau vortices under parabolic and hyperbolic dynamics. Similar methods are used in [OS98a, OS98b, Per04] to study multi-solitons in nonlinear Schrödinger equations. In these papers, the authors construct some explicit low energy *approximate* equilibira, using only the fast localizing property of the single exact solitons. This is readily availble even apart from the Bogomolnyi regime. which extends earlier result of [Stu94a] beyond the Bogomolnyi case. The caveat is that one must then require some spatial separation condition, which ensures each manufactured approximate multi-soliton indeed behaves like a single soliton near each of its localization point. This undermines applications to scattering theory, an issue we shall return to at the end of Section 2.

We shall also mention another context when the problem of effective dynamics arise. Namely, the effective dynamics of the concentration sets can be determined by the geometry of the latter. This is the now classical geometric theory of phase transitions [Mod87, ESS92, PR03], which connects the flow of a real order parameter under the Allen-Cahn equation to the mean curvature flow of its nodal set. This connection has important consequences to both sides of statistical physics and geometric analysis. On the mathematical side, we mention the resolution of De Giorgi's conjecture [dPKW11] among others. See an excellent review [Jer14] on the analogous codimension-two problems, which remain largely open. The author is undertaking to apply the theory we develop here to this end.

2. Effective dynamics

Let Σ be a manifold representing all possible concentration sets of interest. For instance, if one is interested in the nucleation at n points in $\mathbb{R}^d$, then one can take $\Sigma = \mathbb{R}^{nd}$. If one is intereted in the motion of a closed filament in $\mathbb{R}^d$, then one can take $\Sigma = \mathrm{Emb}(S^1, \mathbb{R}^d)$, an so on.

Suppose we have a C^2 map $f : \Sigma \to X$ whose Fréchet derivative $df(\sigma) : T_\sigma\Sigma \to T_{f(\sigma)}X$ is injective for every σ. *NB*: the tangent space $T_{f(\sigma)}X$ can be locally trivialized as a Hilbert space H_1, but H_1 *does not* necessarily coincide with either X or H. For instance, if $X = 1 + H^s \subset H^s\mathrm{loc}$, then $H_1 = H^s$. If $X = \{u \in H : J(u) = 0\}$ for some constraint functional $J : H \to \mathbb{R} \cup \{\infty\}$, then then $H_1 \cong \{v \in H : \langle J'(u), v\rangle = 0\} = \ker dJ(u)$. This is a nontrivial technicality. See Sections 3-4 for examples when this problem arises.

The map f is an immersion, so $M := f(\Sigma)$ forms a submanifold of X, with tangent space at $v_\sigma := f(\sigma)$ given by $T_{v_\sigma} = df(\sigma)(T_\sigma\Sigma)$. This is a subspace of $T_{v_\sigma}X$. We shall call M the space of *approximate solitons*. Here and below, by solitons we shall mean qualitatively such configurations that remain coherent under the flow (1), and quantitatively satisfy (6)-(8). An important characteristic of nonlinear dynamics is the existence of solitons due to the focusing effect of the nonlinearity.

We are interested in the initial value problem

$$\begin{cases} \partial_t u = -E'(u), \\ u|_{t=0} = u_0. \end{cases} \tag{2}$$

Suppose $u_0 \in X_0 \subset X$ where X_0 is a sufficiently small tubular neighbourhood of M. This means that u_0 is close to some $v_{\sigma_0} = f(\sigma_0)$. Let $u(t)$ be a path solving (2). The goal is to reduce (2) under suitable assumptions to an evolution of σ

$$\begin{cases} \partial_t \sigma = F(\sigma), \\ \sigma|_{t=0} = \sigma_1 \end{cases} \tag{3}$$

Here F is independent of v, and σ_1 is close to σ_0. Moreover, $\|u(t) - v_{\sigma(t)}\|_H$ remains uniformly small up to a large but possibly finite time, with $\sigma(t)$ solving (3).

There are three key steps in this reduction. First, we construct a *coordinate* map $S : X_0 \to \Sigma$, s.th. the approximate soliton $v_{S(u)} = f(S(u)) \in M$ serves as an optimal approximation of a given configuration $u \in X_0$. Next, revertin]g the parametrization of the approximation path $v_{\sigma(t)}$, with $\sigma(t) = S(u(t))$, we get an equation for σ as in (3), *so long as* $u(t)$ *remains in* X_0. Lastly, we show under suitable conditions on the approximate solitons, namely condition (6)-(8), the path $u(t)$ actually remains in X_0 for a long time.

The coordinate map. Here we determine a subset of M on which the subsequent orthogonal decomposition is possible. The point is to keep the operators involved to be uniformly bounded on this set.

Lemma 2.1. *Suppose the Fréchet derivative* $\mathrm{res}\, df : \Sigma_0 \to L(Y, H_1)$ *is uniformly bounded on a submanifold* $\Sigma_0 \subset \Sigma$. *Then the (orthogonal or symplectic) projections* $Q_\sigma : H_1 \to T_{f(\sigma)}M$ *are uniformly bounded in* $\sigma \in \Sigma_0$.

Proof. This essentially follows the definition of Q_σ. Define $g_\sigma : Y \to H_1$ by $g_\sigma\eta = df(\sigma)\eta$ where abusing notation we identify the coordinate of $\eta \in T_\sigma\Sigma$ with its coordinate in Y. This map is uniformly bounded in σ as df does. Let $g_\sigma^* : H_1^* \to Y^*$ be the adjoint of g_σ. The map $P_\sigma := g_\sigma^* J^{-1} g_\sigma : Y \to Y^*$ is invertible because df is injective.

The projections from H_1 onto the tagent space H_2 are given by $Q_\sigma := g_\sigma P_\sigma^{-1} g_\sigma^* J^{-1}$. From this formula one observes that Q_σ is uniformly bounded in σ as each of its factor does. ☐

Remark. Here and below, to simplify notation, whenever a basis is understood, we shall identify the coordinate of an element in a tangent space with the element itself.

For later use, we record that if J is symplectic, then P_σ induces a simplectic form on Y by $\langle \tau, P_\sigma \rho\rangle$. The normal space of M at u_σ is related to Q_σ by the relation $\ker Q_\sigma = (JH_2)^\perp \subset H_1$.

Lemma 2.2 (Existence of coordinate)**.** *There exists an open set $X_0 \subset X$ and a C^1 map $S : X_0 \to \Sigma$ s.th.*

(4) $$Q_{S(u)}(u - f(S(u))) = 0.$$

Moreover, if σ' satisfies $Q_{\sigma'}(u - f(\sigma') = 0$, then $\sigma' = S(u)$.

Moreover, if the three maps f, $\operatorname{res} df : \Sigma_0 \times L(Y, H_1)$ and $\operatorname{res} d^2 f : \Sigma_0 \times L(Y, L(Y, H_1)) \cong L_2(Y, H_1)$ are all uniformly bounded in σ, then one can make $S(X_0) \subset \Sigma_0$.

Proof. Let X_0 be a tubular neighbourhood of $f(\Sigma_0)$, whose size is to be determined. Consider the map

$$h : X_0 \times \Sigma_0 \to Y$$

given by $h(u, \sigma) \mapsto g_\sigma^*(u - f(\sigma))$. This map is well-defined for every pair $(u, \sigma) \in X_0 \times \Sigma_0$ s.th. $\|u - f(\sigma)\|_H \le \beta$ for any fixed $\beta > 0$ independent of σ. The point is that that the difference $u - f(\sigma) \in H$, but needs to to be a subset of H_1.

By construction, $d_\sigma h = -P_\sigma$ is invertible (see Lemma 2.1). The equation

$$h(u, \sigma) = 0$$

obviously has the solution $(f(\sigma'), \sigma')$ for every $\sigma' \in \Sigma_0$. Thus by the Implicit Function Theorem, for every σ' there is a neighbourhood $U_f(\sigma') \subset X_0$ around $v_{\sigma'} = f(\sigma')$ and a map $S : U_f(\sigma') \to \Sigma$ that solves the last equation.

Shrink each $U_f(\sigma')$ so that $S(U_f(\sigma')) \subset \Sigma_0$. Shrink X_0 to be the union of all these $U_f(\sigma')$. Then we get a map

$$h(u, S(u)) = 0 \quad (u \in X_0,\ S(X_0) \subset \Sigma_0.)$$

The orthogonality condition (4) follows from the last equation (see the definition of Q_σ in Lemma 2.1). Uniqueness follows by the Implicit Function Theorem.

A priori, as σ' vary in Σ_0, the size of $U_{f(\sigma')}$ can depend on σ'. Yet so long as the maps

$$h(\cdot, \sigma') = g_{\sigma'}^*(\cdot - f(\sigma')),$$
$$d_\sigma h_{\sigma'}(\cdot, \sigma') = -P_{\sigma'},$$
$$d_\sigma^2 h_{\sigma'}(\cdot, \sigma') = g_{\sigma'}^* d^2 f(\sigma')$$

are all uniformly bounded in $\sigma' \in \Sigma$, the size of $U_{f(\sigma)}$ can be made uniform in σ' as well [AP95, Section 2]. These maps are indeed uniformly bounded by the assumptions on f and its Fréchet derivatives. Thus it is possible to shrink X_0 to be a small ϵ-tubular neighbourhood of $f(\Sigma_0)$. ☐

Remark. For $u \in X_0$, the element $S(u) \in \Sigma_0$ serves as the coordinate of u in Σ. The orthogonality condition (4) ensures that the approximate soliton $v_{S(u)} \in M$ is the optimal approximation of u. The map $f \circ S$ can be viewed as a nonlinear projection from X_0 into M. Geometrically, the coordinate $\sigma := S(u)$ is characterized by the condition

$$(u - v_\sigma) \perp JT_{v_\sigma} M.$$

This implies that if u, u' differs by an element in $(JT_{v_\sigma} M)^\perp$, or an element in M, then $S(u) = S(u')$.

In what follows we shall call an immmersion f with the uniformly boundedness property as in Lemma 2.2 *admissible*. If f is only admissible on $\Sigma_0 \subsetneq$, then we shrink Σ to Σ_0.

The orthogonal decomposition. By Lemmas 1-2, any path $u(t) \in X_0$ admits a decomposition

(5) $$u = v + w, \quad v(t) := f(\sigma(t)), \quad \sigma(t) := S(u(t)), \quad Q_{\sigma(t)} w(t) \equiv 0,$$

We now state our main result in this section. *NB*. For simplicity, in what follows we make an implicit assumption that the linearized operator $L_\sigma = E''(f(\sigma))$ is self-adjoint. This is realistic for most applications, see the examples in the following sections.

Theorem 2.3 (Main)**.** *Given $0 < \epsilon \ll 1$, let $u_0 \in X_0$ be an element s.th. $\|u_0 - v_0\|_H \le \epsilon$ for some $v_0 \in M$. Let $u(t) \in X$, $t \le T$ be a solution to (2).*

Suppose f is an admissible immersion of Σ. Suppose there are $\epsilon_1, \epsilon_2 \ge 0, \delta > 0$ s.th. $\delta \gg \epsilon_1, \epsilon_2$, and for every $(\sigma, v, \xi, w) \in \Sigma \times M \times \operatorname{ran} Q_\sigma \times \ker Q_\sigma$, we have

(6) $$\|E'(v)\|_H \le \epsilon_1 \quad \textit{(approximate critical point)},$$

(7) $$\|L_\sigma \xi\|_H \le \epsilon_2 \|\xi\|_H \quad \textit{(approximate zero-mode)},$$

(8) $$|\langle L_\sigma w, w\rangle| \ge \delta \|w\|_H^2 \quad \textit{(coercivity)}.$$

Then $u(t) \in X_0$ for $t \le T_1 = O(\delta/\epsilon_2^2)$, and there exist $c, C, C' > 0$ independent of $t \le T_1$ s.th. for the decomposition $u(t) = v(t) + w(t)$ as in (5),

$$\partial_t \sigma = -P_\sigma^{-1} d_\sigma E(f(\sigma)) + O(\epsilon), \quad \|w(t)\|_H \le C'(\epsilon + \epsilon_1) \quad (t \le T_1). \tag{9}$$

Conversely, suppose $\partial_t \sigma = -P_\sigma^{-1} d_\sigma E(f(\sigma))$ and $v(t) = f(\sigma(t)$. Then there exists $u(t) = v(t) + w(t) \in X_0$, $t \le T_2 = O(1/\epsilon_1)$ s.th. $u(t)$ solves (1), and $\|w(t)\| \le C\epsilon_1$.

Remark. Let $\mathcal{E}(\sigma) := E(f(\sigma))$. Define $\mathcal{J} : T_\sigma^*\Sigma \to T_\sigma\Sigma$ by its action on the coordinate, $Y \ni \rho \mapsto P_\sigma \rho$. If J in (1) is a symplectic operator on the tangent bundle TX, then $\mathcal{J}$ a symplectic operator on $T\sigma$. This follows from the construction of P_σ in Lemma 2.1.

We call the evolution

$$\begin{cases} \partial_t \sigma = -\mathcal{J}^{-1}\mathcal{E}'(\sigma), \\ \sigma|_{t=0} = S(u_0), \end{cases} \tag{10}$$

the *effective dynamics* associated to (2). It follows from the discussion above that this evolution is of the same type (i.e. gradient flow or Hamiltonian system) as (1).

The existence of F depends on the injectivity of $df(\sigma)$. The time interval on which (10) is valid depends on the conditions (6)-(8). These conditions specify how the manifold M should interplay with the energy E: (6),(8) suggest that M should lie near the energy bottom of E, while (7) says that the tangent space $T_{f(\sigma))}M$ should consists of approximately the zero-modes of the linearized operator $E''(f(\sigma))$. Lemma 2.4 below show how (6)-(8) provide the remainder estimate on $\|w(t)\|_H$, depending on the type of evolution (1).

Proof of Theorem 2.3. To simplify notations, in what follows we do not display the dependence on t.

Plugging (5) into (2), we have

$$\partial_t v + \partial_t w = -J^{-1}E'(v + w). \tag{11}$$

Expanding r.h.s. of (11) at v, and then applying $Q_\sigma = Q_{\sigma(t)}$ to both sides of (11), we have

$$\partial_t v + Q_\sigma J^{-1}E'(v) = -Q_\sigma J^{-1}L_\sigma w - Q_\sigma J^{-1}N_\sigma(w) - Q_\sigma \partial_t w. \tag{12}$$

Here the linear operator $L_\sigma := E''(v_\sigma)$, and the nonlinearity N_σ is deterined by this expression.

L.h.s. of (12) can be written as

$$\partial_t v + Q_\sigma J^{-1}E'(v) = g_\sigma(\partial_t \sigma + J^{-1}\mathcal{E}'(\sigma)). \tag{13}$$

Here we use the identities $\partial_t v = g_\sigma \partial_t \sigma$, and $g_\sigma^* E'(v) = d_\sigma E(f(\sigma)) = \mathcal{E}'(\sigma)$, where the last term is the Y-gradient of the functional $E \circ f : Y \to \mathbb{R}$. These identities follow directly from definition and the chain rule. Since g_σ^* is injective, it has bounded inverse on its range. It follows that $\|\partial_t \sigma + \mathcal{J}^{-1}\mathcal{E}'(\sigma)\|_Y$ is bounded by the H-norm of the r.h.s (12).

We now bound the r.h.s. of (12) uniformly in time. The first two terms are of the order $O(\|w\|_H)$, since Q_σ is uniformly bounded and E is smooth on X. The third term is bounded as $\|Q\partial_t w\|_H \le \|\partial_t \sigma\|\|w\|$. In Lemma 2.4 below we show $\|\partial_t \sigma\|_Y$ is uniformly bounded by $\epsilon_1\epsilon_2$, so in fact $\|Q\partial_t w\|_H \le C\|w\|$ in dependent of $\partial_t \sigma$.

Now each term in the r.h.s. of (12) is bounded by $\|w(t)\|_H$. To conclude (9), we show $\|w(0)\|_H \le C\|w_0\|_H$ where $w_0 := u_0 - v_0$. (This is trivial if v_0 is X-closest to u_0, which we *do not* assume.) The rest follows from (15).

Write $v_0 = f(\sigma_0)$ with $\sigma_0 \in \Sigma$. Using the uniform boundedness of S and its local inverse (which exists by the IVT, with the regularity of S as in Lemma 2.2), we compute

$$\|w_0\|_H \ge C\|S(w_0)\|_Y = C\|S(v(0) - v_0)\|_Y \ge C'\|v(0) - v_0\|_Y.$$

Here we use that $S(w_0) = S(v(0) - v_0)$ as w_0 and $v(0) - v_0$ differs by an vector perpenticular to $T_{v(0)}M$, and $S(v(0)) = S(w_0)$. Combining the preceding estimate in the identity,

$$w(0) = v_0 - v(0) + w_0$$

we have $\|w(0)\|_H \le C\|w_0\|$.

Conversely, we want to solve for w from (11), with a given path $v(t) \in M$ satisfying $\partial_t v + Q_\sigma J^{-1}E'(v) = 0$. This gives an evolution for w as

$$\partial_t w = -J^{-1}L_\sigma w - J^{-1}N_\sigma(w) - \bar{Q_\sigma}E'(v).$$

Here $\bar{Q_\sigma} = 1 - Q_\sigma$. The last term is of the order $O(\epsilon_1)$. We can cast this into a fixed point problem with self-adjoint L_σ, which is also invertible on $\ker Q_\sigma \ni w$ by (8), and smooth (in particular Lipschitz) nonlinearity N, using the map

$$w \mapsto e^{-J^{-1}L_\sigma t}w_0 + \int_0^t e^{-J^{-1}L_\sigma s} - J^{-1}N_\sigma(w(s)) + O(\epsilon_1)\, ds. \tag{14}$$

This map is a contraction up to a time T s.th. $CT\epsilon_1 < 1$, where C depends on $-J^{-1}L_\sigma$ and $-J^{-1}N_\sigma$. □

Lemma 2.4. *Suppose there are $\epsilon_1, \epsilon_2 \geq 0, \delta > 0$, with $\delta \gg \epsilon_1, \epsilon_2$, s.th. for every $(\sigma, v, \xi, w) \in \Sigma \times M \times \operatorname{ran} Q_\sigma \times \ker Q_\sigma$, we have*

$$\begin{cases} \|E'(v)\|_H \leq \epsilon_1 \quad \textit{(approximate critical point)}, \\ \|L_\sigma \xi\|_H \leq \epsilon_2 \|\xi\|_H \quad \textit{(approximate zero-mode)}, \\ |\langle L_\sigma w, w\rangle|,\ \|Lw\|_H^2 \geq \delta \|w\|_H^2 \quad \textit{(coercivity)}, \end{cases}$$

Then for a path $w(t) \in \ker Q_{\sigma(t)}$ as in (5), there is $T_1 = O(\delta/\epsilon_2^2)$ s.th.

$$\|w(t)\|_H \leq C(e^{-\frac{\delta}{2\beta}t}\beta\|w(0)\|_H + \delta\epsilon_1) \quad (t \leq T_1) \tag{15}$$

for some $\beta, C > 0$ independent of t.

Proof. We consider two separate cases for gradient and Hamiltonian evolution. 1. Gradient flow. In this case $J = 1$. We derive a differential inequality for the quantity $\frac{1}{2}\langle L_\sigma w, w\rangle$, which contributes most to the energy. Then the desired estimate follows from the coercivity of L_σ.

For paths $v(t), w(t)$ as in (5), compute

$$\begin{aligned} \frac{d}{dt}\frac{1}{2}\langle L_\sigma w, w\rangle &= \langle \partial_t w, L_\sigma w\rangle + \frac{1}{2}\langle [\partial_t, L_\sigma] w, w\rangle \\ &\overset{(11)}{\leq} -\langle E'(v) + L_\sigma w + N_\sigma(w), L_\sigma w\rangle - \langle \partial_t v, L_\sigma w\rangle + C_1\|\partial_t \sigma\|_Y \|w\|_H^2 \\ &\leq -\langle E'(v) + L_\sigma w + N_\sigma(w), L_\sigma w\rangle + \gamma\|\partial_t\sigma\|_Y\|w\|_H + C_1\|\partial_t\sigma\|_Y\|w\|_H^2, \end{aligned} \tag{16}$$

where $\gamma > 0$ is independent of t. For the first inequality, use the commutator estimate $\|[\partial_t, L_\sigma]\|_{H\to H} \leq C_1\|\partial_t\sigma\|_Y$.

For the last inequality in (16), use the self-adjointness, the approximate zero-mode property of v, and boundness of L_σ on $\ker Q_\sigma$, to get

$$\langle \partial_t v, L_\sigma w\rangle = \langle L_\sigma \partial_t v, w\rangle \leq \epsilon_2 \|g_\sigma\|_{Y\to X}\|\partial_t\sigma\|_Y\|w\|_H.$$

By this, one can choose $\gamma = C\epsilon_2$ where C is a uniform upper bound of $\|g_\sigma\|_{Y\to X}$.

Next, we show $\|L_\sigma w\|_H^2 \geq \frac{\delta}{2}\|w\|_H^2$. This is a generic argument, so we drop σ dependence here. The point is that L is coercive and therefore positive, so $L^{1/2}$ is well-defined. Write $L^{1/2}w = w_1 + w_2$ with $w_1 \in \operatorname{ran} Q$ and $w_2 \in \ker Q$. Compute

$$\langle Lw, Lw\rangle = \langle LL^{1/2}w, L^{1/2}w\rangle = \langle L(w_1 + w_2), (w_1 + w_2)\rangle \geq (\delta + O(\epsilon_2))\|w\|_H,$$

where we use (8) to bound $\langle Lw_2, w_2\rangle$ from below, and the smallness property (7) of $L|_{\operatorname{ran} Q}$ to bound the rest terms. For $\epsilon_2 \ll \delta$ this gives the desired result.

Using (6), the coercivity of L_σ, the remainder estimate $\|N_\sigma(w)\|_H \leq C_2\|w\|_H^2$, and (16), we have

$$\left(\frac{d}{dt} + \frac{\delta}{2\beta}\right)\frac{1}{2}\langle L_\sigma w, w\rangle \leq \beta\epsilon_1\|w\|_H + \|w\|_H(\|\partial_t\sigma\|_Y(\gamma + C_1\|w\|_H) + C_2\beta\|w\|^2 - \frac{\delta}{4}), \tag{17}$$

where $\beta < \infty$ is a uniform upperbound of the operator norm of L_σ.

Let T_1 be the maximal time s.th.

$$\|\partial_t\sigma\|_Y(\gamma + C_1\|w\|_H) + C_2\beta\|w\|^2 \leq \frac{\delta}{4} \tag{18}$$

holds. We show $T_1 = O(\delta/\epsilon_2^2)$. First, we bound $\|\partial_t v\|_Y$. Differentiating the square norm of the velocity using (10), we have

$$\begin{aligned} \frac{d}{dt}\frac{1}{2}\|\partial_t\sigma\|_Y^2 &\leq C\frac{d}{dt}\frac{1}{2}\|E'(v)\|_H^2 \\ &= C\langle L_\sigma\partial_t v, E'(v)\rangle \\ &\leq C\epsilon_1\epsilon_2\|\partial_t v\|_Y. \end{aligned}$$

Integrating this, we get

$$M(t) \leq C\epsilon_1\epsilon_2 t + C'\epsilon_1 \quad (M(t) := \sup_{t'\leq t}\|\partial_t\sigma(t')\|_Y).$$

This shows that (18) is valid on a time interval $t \leq T_1$ with $T_1 = O(\delta/\epsilon_2^2)$, if $\epsilon_2 > 0$. By (15), if $\epsilon_2 = 0$ then (18) is valid for all time, since the raminder $\|w\|_H = C\epsilon_1 + O(e^{-\frac{\delta}{2\beta}t})$.

Next, we show $\|w(t)\|_H$ has a exponentially decaying envelope, and therefore (18) holds up to T_1. Shrinking the set of admissible initial configurations X_0 to a smaller tubular neighbourhood of M, one can make $\|w(0)\|_H \leq C\|w_0\| \ll \delta$. Then by continuity of the flow, $\|w(t)\|_H \ll \delta$ up to some $T_2 > 0$. For $t \leq \min(T_1, T_2)$, condition (18) holds, so (17) implies

$$\frac{d}{dt}(e^{\frac{\delta}{2\beta}t}\frac{1}{2}\langle L_\sigma w, w\rangle) \leq e^{\frac{\delta}{2\beta}t}\beta\epsilon_1\|w\|_H. \tag{19}$$

Integrating (19) in time, we have

$$\frac{1}{2}\langle L_\sigma w, w\rangle \le e^{-\frac{\delta}{2\beta}t}\langle L_{\sigma(0)}w(0), w(0)\rangle + \frac{\delta\epsilon_1}{2}M(t), \tag{20}$$

where we set $M(t) := \sup_{t'\le t}\|w(t')\|_H$. Using the coercivity condition in (20), we get

$$M(t)^2 \le \frac{2}{\delta}(e^{-\frac{\delta}{2\beta}t}\beta\|w(0)\|_H^2 + \beta\epsilon_1 M(t)) \implies M(t) \le \frac{2}{\delta}(e^{-\frac{\delta}{2\beta}t}\beta\|w(0)\|_H + \beta\epsilon_1).$$

This implies $\|w(T_2)\| \ll \delta$, and therefore (18) still holds at T_2. Thus we can iteratre this process all the way up to $t = T_1$.

2. Hamiltonian system. In this case J is a symplectic operator, and we would like to derive a approximate conservation law. Consider the expansion

$$E(v+w) = E(v) + \langle E'(v), w\rangle + \frac{1}{2}\langle L_\sigma w, w\rangle + O(\|w\|_H^3).$$

Using (6) and (8), we find

$$\|w\|^2 \le E(v+w) - E(v) + \epsilon_1\|w\| + C\|w\|_H^3.$$

□

In the physics literature, the procedure described above is known as the *adiabatic approximation*. The constraint (18) is the adiabatic condition: it restricts the validity of approximation to configurations with slowly moving parameters.

The conditions (6) strengthens the estimate (15) as $\epsilon_1 \to 0$, viz. as $v \in M$ tends to exact critical point of E. Indeed, if M are critical points of E, then differentiating the stationary equation $E'(f(\sigma)) = 0$ w.r.t. σ we see that $L_\sigma df(\sigma) = 0$, and consequently $T_{f(\sigma)}M = df(\sigma)(T_\sigma\Sigma)$ lies in the kernel of L_σ. Thus (7) are satisfied with $\epsilon_2 = 0$. In turn, as $\delta \to \infty$, the exponential decay rate increases through (15).

We now show that conditions (6)-(8) are robust, in the sense that if they hold at one point, then they hold in a neighbourhood around that point. This shows that if one has at least one candidate for the approximate soliton, then automatically one has a manifold of such.

Proposition 2.5. *Suppose $v_0 = f(\sigma_0)$ for some $\sigma_0 \in \Sigma$, where $f : \Sigma \to X$ is a smooth immersion of some manifold Σ. Suppose 0 is an isolated eigenvalue of the linearized operator $L_0 = E''(v_0)$. Suppose (6)-(8) hold at v_0. Then (6)-(8) hold for for every $(\sigma, v, \xi, w) \in \Sigma_0 \times M_0 \times \operatorname{ran} Q_\sigma \times \ker Q_\sigma$ where Σ_0, M_0 are resp. neighbourhoods around σ_0, v_0.*

Proof. These estimates follow by contuinuity argument. First, since E' and f vary smoothly, for $\|\sigma - \sigma_0\| \ll \epsilon_1$, we have $\|E'(f(\sigma)) - E'(f(\sigma_0))\|_H \le \epsilon_1/2$, which gives (6).

The tangent space to M varies smoothly near σ_0, in the sense that

$$\|\bar{Q_\sigma} - \bar{Q_0}\|_{H\to H} = \|Q_\sigma - Q_0\|_{H\to H} \le C\|\sigma\|_Y. \tag{21}$$

This equation holds because Q_σ is C^1 and therefore Lipshitz in σ, as f is C^2. (See the definition of Q_σ in Lemma 2.1.) Similarly, L_σ also varies smoothly with $\|L_\sigma - L_0\|_{H\to H} \le C\|\sigma\|_Y$. So for $\|\sigma - \sigma_0\| \ll \epsilon_2$ we have (7) for every tangent vector in $T_{f(\sigma)}M$.

Lastly, the coercivity condition (8) follows because

$$\bar{Q_\sigma} L_\sigma \bar{Q_\sigma} = \bar{Q_0} L_0 \bar{Q_0} + O(\|\sigma\|_Y) \implies \langle L_\sigma w, w\rangle \ge (\delta - O(\|\sigma\|_Y))\|w\|^2 \quad (w \in (\ker Q_\sigma)^\perp). \tag{22}$$

□

Next, we shall exploit the fact that M needs not to consist of only critical points of E. This leaves considerable freedom in constructing configurations that resembles a multi-soliton. As long as these manufactured configurations satisfy (6)-(8), we can derive an effective motion law (10) for these non-stationary solutions to (2), which then captures the interaction of multiple single solitons.

For instance, consider the following construction in [Per04]. Denote $\psi_{\lambda,z}$ the NLS gound state, which is the unique smoooth, radial, positive solution to the nonlinear eigenvalue problem $-\Delta u + \lambda u + g(u^2)u = 0$. (Here g is suitable nonlinearity.) This ground state decays exponentially fast away from $z \in \mathbb{R}^d$. Then simply adding several gound states together produces a multi-soliton $f(z,\lambda) := \sum_{i=1}^n \psi_{\lambda_j, z_j}$. When the point solitons are weakly interacting, i.e. if $\min_{i\ne j}|z_i - z_j| \gg 1$, then heuristically $f(z,\lambda)$ is itself close to a gound state of $E = \int \frac{1}{2}|\nabla u|^2 + G(|u|^2)$ (where $G' = g$). Now taking into account the symmetry of E (rotations with dimension $d(d-1)/2$ and modulation $u \mapsto e^{i\theta}u$ with one dimension), we get an immersion $F : \Sigma \subset \mathbb{R}^{d^2(d+1)/2+2d} \to H^1(\mathbb{R}^d, \mathbb{C})$. With a suitable submanifold Σ we can make $M := F(\Sigma)$ compatible with the energy E, in the sense that (6)-(8) hold. Then one can produce an effective motion law for the interacting solitons $f(z,\lambda)$ under (1), in this case a nonlinear heat equation, in terms of z, as in (10). The effective motion is valid up to a exponentially decaying error, as in (15). This makes rigorous an essential feature of solitons, namely this asymptotic linear aspect that resembles the principle of superposition for linear equations.

Finally, let us mention the limitation of our approximation scheme. Suppose now the physical system consists of n identical particles. Then two points in $\mathbb{R}^{nd}$ related by a permutation in S_n should parametrize the same configuration. This way Σ becomes a submanifold of $\mathbb{R}^n/S_n$. So far we have not discuss the possibility of Σ being a manifold with boundary of less dimension. This problem comes to the front when one consider the scattering of n identical particle. With the trivial coordinate inherit from $\mathbb{R}^{nd}$, the manifold Σ is singular along the diagonal $D := \{(z, \dots, z)\} \subset \mathbb{R}^{nd}$. However, scattering happens precisely when the trajectory crosses D. One can use the coordinate given by the n elementary symmetric polynomials $c_1, \dots, c_n$ on Σ so as to make the latter a smooth manifold. But then though the map $(c_1, \dots, c_n) \mapsto f(z(c_i))$ can be well-defined, it is not C^1 in general, because the derivatives $\partial z_i / \partial_{c_j}$ blows up at the diagonal D. Moreover, as z approaches D, with the kind of constrution we described above (i.e. with naive superposition), it is not likely that the approximate soliton remains to have low energy as $z \to D$, since the separation $\min_{i \neq j} |z_i - z_j| \to 0$.

References

[AP95] Antonio Ambrosetti and Giovanni Prodi, *A primer of nonlinear analysis*, Cambridge Studies in Advanced Mathematics, vol. 34, Cambridge University Press, Cambridge, 1995. Corrected reprint of the 1993 original. MR1336591

[CFS18] Ilias Chenn, G. Fournodavlos, and I. M. Sigal, *The effective dynamics of the volume preserving mean curvature flow*, J. Stat. Phys. **172** (2018), no. 2, 458–476. MR3824945

[dPKW11] Manuel del Pino, Michał Kowalczyk, and Juncheng Wei, *On De Giorgi's conjecture in dimension* $N \geq 9$, Ann. of Math. (2) **174** (2011), no. 3, 1485–1569. MR2846486

[DS09] Sophia Demoulini and David Stuart, *Adiabatic limit and the slow motion of vortices in a Chern-Simons-Schrödinger system*, Comm. Math. Phys. **290** (2009), no. 2, 597–632. MR2525632

[ESS92] L. C. Evans, H. M. Soner, and P. E. Souganidis, *Phase transitions and generalized motion by mean curvature*, Comm. Pure Appl. Math. **45** (1992), no. 9, 1097–1123. MR1177477

[GS06] S. Gustafson and I. M. Sigal, *Effective dynamics of magnetic vortices*, Adv. Math. **199** (2006), no. 2, 448–498. MR2189216

[Jer14] Robert L. Jerrard, *Quantized vortex filaments in complex scalar fields*, Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. III, 2014, pp. 789–810. MR3729052

[Man82] N. S. Manton, *A remark on the scattering of BPS monopoles*, Phys. Lett. B **110** (1982), no. 1, 54–56. MR647883

[Mod87] Luciano Modica, *The gradient theory of phase transitions and the minimal interface criterion*, Arch. Rational Mech. Anal. **98** (1987), no. 2, 123–142. MR866718

[OS98a] Yu. N. Ovchinnikov and I. M. Sigal, *The Ginzburg-Landau equation. III. Vortex dynamics*, Nonlinearity **11** (1998), no. 5, 1277–1294. MR1644389

[OS98b] ______, *Long-time behaviour of Ginzburg-Landau vortices*, Nonlinearity **11** (1998), no. 5, 1295–1309. MR1644393

[Per04] Galina Perelman, *Asymptotic stability of multi-soliton solutions for nonlinear Schrödinger equations*, Comm. Partial Differential Equations **29** (2004), no. 7-8, 1051–1095. MR2097576

[PR03] Frank Pacard and Manuel Ritoré, *From constant mean curvature hypersurfaces to the gradient theory of phase transitions*, J. Differential Geom. **64** (2003), no. 3, 359–423. MR2032110

[Stu07] David M. A. Stuart, *Analysis of the adiabatic limit for solitons in classical field theory*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **463** (2007), no. 2087, 2753–2781. MR2360179

[Stu94a] D. Stuart, *Dynamics of abelian Higgs vortices in the near Bogomolny regime*, Comm. Math. Phys. **159** (1994), no. 1, 51–91. MR1257242

[Stu94b] ______, *The geodesic approximation for the Yang-Mills-Higgs equations*, Comm. Math. Phys. **166** (1994), no. 1, 149–190. MR1309545

[Tin10] F. Ting, *Effective dynamics of multi-vortices in an external potential for the Ginzburg-Landau gradient flow*, Nonlinearity **23** (2010), no. 1, 179–210. MR2576380