



**HAL**  
open science

# Orbital stability via the energy-momentum method: the case of higher dimensional symmetry groups

Stephan de Bievre, Simona Rota Nodari

## ► To cite this version:

Stephan de Bievre, Simona Rota Nodari. Orbital stability via the energy-momentum method: the case of higher dimensional symmetry groups. 2016. hal-01312534v2

**HAL Id: hal-01312534**

**<https://hal.science/hal-01312534v2>**

Preprint submitted on 5 Feb 2018 (v2), last revised 11 Jul 2018 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# ORBITAL STABILITY VIA THE ENERGY-MOMENTUM METHOD: THE CASE OF HIGHER DIMENSIONAL SYMMETRY GROUPS

STEPHAN DE BIÈVRE<sup>1,2</sup> AND SIMONA ROTA NODARI<sup>3</sup>

ABSTRACT. We consider the orbital stability of relative equilibria of Hamiltonian dynamical systems on Banach spaces, in the presence of a multi-dimensional invariance group for the dynamics. We prove a persistence result for such relative equilibria, present a generalization of the Vakhitov-Kolokolov slope condition to this higher dimensional setting, and show how it allows to prove the local coercivity of the Lyapunov function, which in turn implies orbital stability. The method is applied to study the orbital stability of relative equilibria of nonlinear Schrödinger and Manakov equations. We provide a comparison of our approach to the one by Grillakis-Shatah-Strauss.

## 1. INTRODUCTION

The *Vakhitov-Kolokolov slope condition* [17, 34, 37] is an often used ingredient in the proof of orbital stability of relative equilibria via the energy-momentum method for Hamiltonian systems with a one-dimensional symmetry group. For example, it has been applied in the proofs of stability of stationary or traveling waves of a variety of nonlinear partial differential equations (see [3, 5, 14, 15, 16, 19, 34, 38] and references therein). It is our goal here to present a natural generalization of this condition to the case where the Hamiltonian system admits a higher dimensional invariance group and to show how to obtain orbital stability from it.

The overall strategy underlying the energy-momentum method is well understood. Simply stated, it is a generalization of the standard Lyapunov method for proving the stability of fixed points to Hamiltonian systems having a Lie symmetry group  $G$ . Indeed, relative equilibria can be seen as fixed points “modulo symmetry”: they are fixed points of the dynamics induced on the space obtained by quotienting the phase space by the action of an appropriate subgroup of the invariance group. For finite dimensional systems, the theory goes back to the nineteenth century. It is concisely explained in [1, 22], in the modern language of Hamiltonian systems with symmetry, through the use of the properties of the momentum map, notably. This theory first of all gives a simple geometric characterization of all relative equilibria. It also naturally provides a candidate Lyapunov function as well as subgroup of  $G$  with respect to which the relative equilibria can be hoped to be relatively stable. More recent developments in the finite dimensional setting can be found in [23, 24, 29, 28, 21, 27, 30, 33].

When the Hamiltonian system is infinite dimensional, such as is the case for nonlinear Hamiltonian PDE’s, the general philosophy of the energy-momentum method remains the same, but many technical complications arise, as expected. In [17] and [34], the theory is worked out in a Hilbert space setting, and when the symmetry group  $G$  is a *one-dimensional* Lie group. More recently, in [10], a version of the energy-momentum method has been presented for Hamiltonian dynamical systems on a Banach space  $E$  having as invariance group a Lie group  $G$  of *arbitrary finite dimension*. What is shown there is that the proof of orbital stability can be reduced to a “local coercivity estimate” on an appropriately constructed Lyapunov function  $\mathcal{L}$ . It is shown in [10] that, in the above infinite dimensional setting, the construction of the latter follows naturally from the Hamiltonian structure and basic properties of the momentum map, in complete analogy with the finite dimensional situation. In specific models, it then remains to show the appropriate local coercivity estimate on  $\mathcal{L}$  which amounts to a lower bound on its Hessian restricted to an appropriate subspace of  $E$  (see (2.14)).

When the invariance group  $G$  of the system is one-dimensional, one way to obtain such an estimate is via the aforementioned *Vakhitov-Kolokolov slope condition*. Our main result here is a generalization of this condition to situations with a higher dimensional invariance group  $G$ , and the proof that it implies the desired coercivity of the Lyapunov function (Theorem 3.1 and Theorem 6.3). Using this property, orbital stability can then for example be obtained using the techniques described in [10].

The rest of this paper is organized as follows. The elements of [10] needed here are summarized in Section 2. In Section 3, we state our main result (Theorem 3.1) in the simplest setting, when the phase space of the system is a Hilbert space. Some preliminary lemmas are proven in Section 4 and the proof of Theorem 3.1 is given in Section 5. In Section 6, we generalize our result to the Banach space setting, see Theorem 6.3. Section 7 deals with the important question of “persistence” of relative equilibria. In Section 8, we use our approach to study several applications of our results to the stability study for relative equilibria of the nonlinear Schrödinger and Manakov equations. In Section 9, finally, a detailed analysis of the differences between our work here and the approach of [18] is provided. In the latter paper, the energy-momentum method had previously been adapted to the case of higher symmetry groups, and a generalization of the *Vakhitov-Kolokolov slope condition* proposed. We have in particular amended, completed and generalized some of the results and arguments of this seminal work.

**Acknowledgments.** This work was supported in part by the Labex CEMPI (ANR-11-LABX-0007-01) and by FEDER (PIA-LABEX-CEMPI 42527). The authors are grateful to Dr. M. Conforti, Prof. F. Genoud, Prof. S. Keraani, Prof. S. Mehdi, Prof S. Trillo and Prof. G. Tuynman for helpful discussions on the subject matter of this paper. They also thank an anonymous referee for instructive comments and a careful reading of the manuscript.

## 2. THE ENERGY-MOMENTUM METHOD

To make this paper self-contained and to fix our notation, we summarize in this section the energy-momentum method as described in [10]. We refer there for more details and for examples of the structures introduced here.

**2.1. Hamiltonian systems with symmetry.** Let  $E$  be a Banach space,  $\mathcal{D}$  a domain in  $E$  (*i.e.* a dense subset of  $E$ ) and  $\mathcal{J}$  a symplector, that is an injective continuous linear map  $\mathcal{J} : E \rightarrow E^*$  such that  $(\mathcal{J}u)(v) = -(\mathcal{J}v)(u)$ . We will refer to  $(E, \mathcal{D}, \mathcal{J})$  as a *symplectic Banach triple*. Next, let  $H : E \rightarrow \mathbb{R}$  be differentiable on  $\mathcal{D} \subset E$ . In other words,  $H$  is globally defined on  $E$ , and differentiable at each point  $u \in \mathcal{D}$ . We say that the function  $H$  has a  $\mathcal{J}$ -compatible derivative if, for all  $u \in \mathcal{D}$ ,  $D_u H \in \mathcal{R}_{\mathcal{J}}$ , where  $\mathcal{R}_{\mathcal{J}}$  is the range of  $\mathcal{J}$ . In that case we write  $H \in \text{Dif}(\mathcal{D}, \mathcal{J})$ .

We define a *Hamiltonian flow* for  $H \in \text{Dif}(\mathcal{D}, \mathcal{J})$  as a separately continuous map  $\Phi^H : \mathbb{R} \times E \rightarrow E$  with the following properties:

- (i) For all  $t, s \in \mathbb{R}$ ,  $\Phi_{t+s}^H = \Phi_t^H \circ \Phi_s^H$ ,  $\Phi_0^H = \text{Id}$ ;
- (ii) For all  $t \in \mathbb{R}$ ,  $\Phi_t^H(\mathcal{D}) = \mathcal{D}$ ;
- (iii) For all  $u \in \mathcal{D}$ , the curve  $t \in \mathbb{R} \rightarrow u(t) := \Phi_t^H(u) \in \mathcal{D} \subset E$  is differentiable and is the unique solution of

$$\mathcal{J}\dot{u}(t) = D_{u(t)}H, \quad u(0) = u. \quad (2.1)$$

Note that here and below,  $D_u H \in E^*$  is our notation for the derivative of  $H$  at  $u$ . We refer to (2.1) as the Hamiltonian differential equation associated to  $H$  and to its solutions as Hamiltonian flow lines.

Next, let  $G$  be a Lie group,  $\mathfrak{g}$  the Lie algebra of  $G$  and  $\Phi : (g, x) \in G \times E \rightarrow \Phi_g(x) \in E$ , an action of  $G$  on  $E$ . In what follows we will suppose all Lie groups are connected. We will say  $\Phi$  is a *globally Hamiltonian action* if the following conditions are satisfied:

- (i) For all  $g \in G$ ,  $\Phi_g \in C^1(E, E)$  is symplectic.
- (ii) For all  $g \in G$ ,  $\Phi_g(\mathcal{D}) = \mathcal{D}$ .

- (iii) For all  $\xi \in \mathfrak{g}$ , there exists  $F_\xi \in C^1(E, \mathbb{R}) \cap \text{Dif}(\mathcal{D}, \mathcal{J})$  such that  $\Phi_{\exp(t\xi)} = \Phi_t^{F_\xi}$  is a Hamiltonian flow as defined above and the map  $\xi \rightarrow F_\xi$  is linear.

Here and in what follows, we say  $\Psi \in C^0(E, E) \cap C^1(\mathcal{D}, E)$  is a symplectic transformation if

$$\forall u \in \mathcal{D}, \forall v, w \in E, (\mathcal{J}D_u\Psi(v))(D_u\Psi(w)) = (\mathcal{J}v)(w). \quad (2.2)$$

Note that, in the above definition of globally Hamiltonian action,  $\Psi = \Phi_g \in C^1(E, E)$ . For further reference, we introduce, for all  $u \in \mathcal{D}$  and for all  $\xi \in \mathfrak{g}$ ,

$$X_\xi(u) = \frac{d}{dt}\Phi_{\exp(\xi t)}(u)|_{t=0}. \quad (2.3)$$

It follows from the preceding definitions that

$$X_\xi(u) = \mathcal{J}^{-1}D_u F_\xi. \quad (2.4)$$

We will always suppose  $G$  is a matrix group, in fact, a subgroup of  $\text{GL}(\mathbb{R}^N)$ . We can then think of the Lie algebra  $\mathfrak{g}$  as a sub-algebra of the  $N \times N$  matrices  $\mathcal{M}(N, \mathbb{R})$  and define the adjoint action of  $G$  on  $\mathfrak{g}$  via

$$\text{Ad}_g \xi = g\xi g^{-1},$$

where in the right hand side we have a product of matrices. We will write  $m = \dim \mathfrak{g} = \dim \mathfrak{g}^*$ , where  $\mathfrak{g}^*$  designates the vector space dual of the Lie algebra  $\mathfrak{g}$ . For details, we refer to Section A.2 of [10], [1] or [22]. Note that, for each  $u \in E$  fixed, one can think of  $\xi \in \mathfrak{g} \rightarrow F_\xi(u) \in \mathbb{R}$  as an element of  $\mathfrak{g}^*$ . Hence, if we identify (as we always will)  $\mathfrak{g}$  and  $\mathfrak{g}^*$  with  $\mathbb{R}^m$  and view  $F$  as a map  $F : E \rightarrow \mathbb{R}^m \simeq \mathfrak{g}^*$ , we can write

$$F_\xi = \xi \cdot F,$$

where  $\cdot$  refers to the canonical inner product on  $\mathbb{R}^m$ . The map  $F$  is called the momentum map of the symplectic group action and, in what follows, we will suppose that  $F$  is  $\text{Ad}^*$ -equivariant which means that for all  $g \in G$ , for all  $\xi \in \mathfrak{g}$

$$F_\xi \circ \Phi_g = F_{\text{Ad}_{g^{-1}}\xi},$$

or equivalently,  $F \circ \Phi_g = \text{Ad}_g^* F$ . Here  $\text{Ad}^*$  is the co-adjoint action of  $G$  on  $\mathfrak{g}^*$ .

Now, for all  $\mu \in \mathfrak{g}^*$ , we define the isotropy group or stabilizer of  $\mu$  as

$$G_\mu = \{g \in G \mid \text{Ad}_g^* \mu = \mu\};$$

$\mathfrak{g}_\mu$  is the Lie algebra of  $G_\mu$ , and  $\mathfrak{g}_\mu^*$  its dual. Finally, for all  $\mu \in \mathfrak{g}^* \simeq \mathbb{R}^m$ , let

$$\Sigma_\mu = \{u \in E \mid F(u) = \mu\}.$$

We will say  $\mu$  is a *regular value* of  $F$  if  $\Sigma_\mu \neq \emptyset$  and if, for all  $u \in \Sigma_\mu$ ,  $D_u F$  is surjective (maximal rank). Then  $\Sigma_\mu$  is a codimension  $m$  sub-manifold of  $E$  and its tangent space at  $u \in \Sigma_\mu$  is

$$T_u \Sigma_\mu = \text{Ker } D_u F. \quad (2.5)$$

Finally, since the momentum map is  $\text{Ad}^*$ -equivariant, it is easy to see  $G_\mu = G_{\Sigma_\mu}$ , where  $G_{\Sigma_\mu}$  is the subgroup of  $G$  leaving  $\Sigma_\mu$  invariant.

Below,  $G$  will be an invariance group of  $H$ , in the sense that  $H \circ \Phi_g = H$ , for all  $g \in G$ . This implies  $G$  is an invariance group for the dynamics generated by  $H$ , meaning that for all  $g \in G, t \in \mathbb{R}$ ,  $\Phi_g \circ \Phi_t^H = \Phi_t^H \circ \Phi_g$  (See Theorem 2.2 (i) below). Noether's Theorem then implies that the components  $F_i$  of the moment map are constants of the motion (See Theorem 2.2 (ii)) and hence that, for any  $\mu \in \mathbb{R}^m \simeq \mathfrak{g}^*$ , the level set  $\Sigma_\mu$  is invariant under the dynamics  $\Phi_t^H$ . We refer to Sections 8 and 9.4 for examples; see also [10].

**2.2. Relative equilibria and orbital stability.** We now recall the definition of a relative equilibrium. Let  $G$  be an invariance group for the dynamics  $\Phi_t^H$ , as above, and let  $\tilde{G}$  be a subgroup of  $G$ . Let  $u \in E$  and let  $\mathcal{O}_u^{\tilde{G}} = \Phi_{\tilde{G}}(u)$  be the  $\tilde{G}$ -orbit of  $u$ . We say  $u$  is a relative  $\tilde{G}$ -equilibrium of the dynamics if, for all  $t \in \mathbb{R}$ ,  $\Phi_t^H(u) \in \mathcal{O}_u^{\tilde{G}}$ . In other words, if the dynamical trajectory through  $u$  lies in the group orbit  $\mathcal{O}_u^{\tilde{G}}$ .

The goal is to investigate under which circumstances these relative equilibria are orbitally stable. Recall that a relative  $\tilde{G}$ -equilibrium  $u \in E$  is orbitally stable if

$$\forall \varepsilon > 0, \exists \delta > 0, \forall v \in E, \left( d(v, u) \leq \delta \Rightarrow \forall t \in \mathbb{R}, d(v(t), \mathcal{O}_u^{\tilde{G}}) \leq \varepsilon \right),$$

with  $v(t)$  the solution of the Hamiltonian equation of motion with initial condition  $v(0) = v$ . Here  $d(\cdot, \cdot)$  is the distance function induced by the norm on  $E$ . Note that the definitions of relative equilibrium and of orbital stability are increasingly restrictive as the subgroup  $\tilde{G}$  is taken smaller. Sharper statements are therefore obtained by choosing smaller subgroups  $\tilde{G}$ .

It turns out that, if  $H$  is  $G$  invariant and the action of  $G$  is  $\text{Ad}^*$ -equivariant, then  $u$  is a  $G$ -relative equilibrium if and only if  $u$  is a  $G_\mu$ -relative equilibrium, where  $\mu = F(u)$  (See Theorem 7 in [10]). This observation, familiar from the finite dimensional theory (See for instance [1, 22]), explains why it is natural to try to prove  $G_\mu$ -orbital stability. This is the approach we shall adopt here. It differs from the one in [18], where orbital stability is studied with respect to an a priori different subgroup, as we will explain in detail in Section 9. We will also show there that in many situations of interest, the two subgroups actually coincide.

We will write

$$\mathcal{O}_u = \Phi_{G_\mu}(u), \tag{2.6}$$

where  $\mu = F(u)$ . And, for all  $u \in \mathcal{D}$ ,

$$T_u \mathcal{O}_u = \{X_\xi(u) \mid \xi \in \mathfrak{g}_\mu\} \subset E. \tag{2.7}$$

For later reference, we recall the following definitions.

**Definition 2.1.** We say  $F$  is regular at  $u \in E$  if  $D_u F$  is of maximal rank. We say  $\mu$  is a regular value of  $F$ , if for all  $u \in \Sigma_\mu$ ,  $D_u F$  is of maximal rank. We will refer to relative equilibria  $u$  for which  $D_u F$  is of maximal rank, as *regular relative equilibria*.

To understand what follows, it is helpful to keep in mind that in practice, the action of the invariance group  $G$  is well known explicitly, and typically linear and isometric. The dynamical flow  $\Phi_t^H$ , on the other hand, is a complex object one tries to better understand using the invariance properties of  $H$ .

We now collect some results from [10] which give a characterization of the relative equilibria of Hamiltonian systems with symmetry and which also yield the candidate Lyapunov function that can be used to study their stability.

**Theorem 2.2.** *Let  $(E, \mathcal{D}, \mathcal{J})$  be a symplectic Banach triple. Let  $H \in C^1(E, \mathbb{R}) \cap \text{Dif}(\mathcal{D}, \mathcal{J})$  and suppose  $H$  has a Hamiltonian flow  $\Phi_t^H$ . Let furthermore  $G$  be a Lie group, and  $\Phi$  a globally Hamiltonian action on  $E$  with  $\text{Ad}^*$ -equivariant momentum map  $F$ . Suppose that,*

$$\forall g \in G, \quad H \circ \Phi_g = H. \tag{2.8}$$

- (i) *Then  $G$  is an invariance group for  $\Phi_t^H$ .*
- (ii) *For all  $t \in \mathbb{R}$ ,  $F \circ \Phi_t^H = F$ .*
- (iii)  *$u$  is a relative  $G$ -equilibrium if and only if  $u$  is a relative  $G_\mu$ -equilibrium.*
- (iv) *Let  $u \in \mathcal{D} \subset E$ . If there exists  $\xi \in \mathfrak{g}$  so that*

$$D_u H - \xi \cdot D_u F = 0, \tag{2.9}$$

*then  $u$  is a relative  $G_\mu$ -equilibrium. Let  $\mu = F(u) \in \mathbb{R}^m \simeq \mathfrak{g}^*$ ; if  $\mu$  is a regular value of  $F$ , then  $u$  is a critical point of  $H_\mu$  on  $\Sigma_\mu$ , where  $H_\mu = H|_{\Sigma_\mu}$ .*

Equation (2.9) is referred to as the stationary equation in the PDE literature. The theorem states that its solutions determine relative  $G$ - and hence relative  $G_\mu$ -equilibria.

We now turn to the stability analysis of those relative equilibria. Suppose we are given  $\xi \in \mathfrak{g}$  and  $u_\xi$ , solution of (2.9). We first note that the fact that  $u_\xi$  is a critical point of the restriction  $H_{\mu_\xi}$  of the Hamiltonian  $H$  to  $\Sigma_{\mu_\xi}$  ( $\mu_\xi = F(u_\xi)$ ) is an immediate consequence of the observation that  $u_\xi$  is a critical point of the Lagrange function

$$\mathcal{L}_\xi = H - \xi \cdot F : E \rightarrow \mathbb{R}. \quad (2.10)$$

The goal is now to prove that these relative equilibria are orbitally stable. As pointed out in [10], the basic idea underlying the energy-momentum method is that, modulo technical problems, a relative equilibrium is expected to be stable if it is not only a critical point but actually a local minimum of  $H_{\mu_\xi}$ . To establish such a result, it is natural to use the second variation of the Lagrange multiplier theory and to establish that the Hessian of  $\mathcal{L}_\xi$  is positive definite when restricted to  $T_{u_\xi} \Sigma_{\mu_\xi} \cap T_{u_\xi} \mathcal{O}_{u_\xi}^\perp$ . The precise statement is given in Proposition 2.3 below.

Let  $\langle \cdot, \cdot \rangle$  be a scalar product on  $E$ , which is continuous in the sense that

$$\forall v, w \in E, \quad |\langle v, w \rangle| \leq \|v\|_E \|w\|_E, \quad (2.11)$$

where  $\|\cdot\|_E$  is our notation for the Banach norm on  $E$ . Note that  $E$  is not necessarily a Hilbert space for this inner product. In addition, even if  $(E, \|\cdot\|_E)$  is in fact a Hilbert space with the norm  $\|\cdot\|_E$  coming from an inner product  $\langle \cdot, \cdot \rangle_E$ , the second inner product  $\langle \cdot, \cdot \rangle$  is not necessarily equal to  $\langle \cdot, \cdot \rangle_E$ .

Let  $\|\cdot\|$  be the norm associated to the scalar product  $\langle \cdot, \cdot \rangle$  and define  $\hat{E}$  to be the closure of  $E$  with respect to the  $\|\cdot\|$ -norm, that is

$$\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}, \quad \hat{E} = \overline{E}^{\|\cdot\|}. \quad (2.12)$$

Note that  $\hat{E}$  is a Hilbert space and  $E \subset \hat{E}$ . As an example, one can think of  $E = H^1(\mathbb{R}^n)$  and  $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{L^2(\mathbb{R}^n)}$  so that  $\hat{E} = L^2(\mathbb{R}^n)$  in that case. This is the typical situation for the nonlinear Schrödinger equation; we refer to [10] and Section 8 for details and further examples.

For further reference, we collect our main structural assumptions in the following hypotheses:

**Hypothesis A.** Let  $(E, \mathcal{J}, \mathcal{D}, \langle \cdot, \cdot \rangle, H, G, \Phi, F)$  and suppose:

- (i)  $(E, \mathcal{J}, \mathcal{D})$  is a symplectic Banach triple and  $\langle \cdot, \cdot \rangle$  a continuous scalar product on  $(E, \|\cdot\|_E)$  as in (2.11).
- (ii)  $H \in C^2(E, \mathbb{R}) \cap \text{Dif}(\mathcal{D}, \mathcal{J})$
- (iii)  $G$  is a Lie group, and  $\Phi$  a globally Hamiltonian  $G$ -action on  $E$  with  $\text{Ad}^*$ -equivariant momentum map  $F \in C^2(E, \mathbb{R}^m)$ .
- (iv)  $H \circ \Phi_g = H$  for all  $g \in G$ .

**Hypothesis B.**  $\Phi_g$  is linear and preserves both the structure  $\langle \cdot, \cdot \rangle$  and the norm  $\|\cdot\|_E$  for all  $g \in G$ .

One then has:

**Proposition 2.3.** *Suppose Hypotheses A and B hold. Let  $\xi \in \mathfrak{g}$  and suppose  $u_\xi \in \mathcal{D}$  satisfies (2.9), i.e.  $D_{u_\xi} \mathcal{L}_\xi = 0$ , with  $\mathcal{L}_\xi = H - \xi \cdot F$ . Let  $\mu_\xi = F(u_\xi) \in \mathbb{R}^m \simeq \mathfrak{g}^*$  and suppose  $\mu_\xi$  is a regular value of  $F$ . Suppose in addition that*

- (i)  $g \in G_{\mu_\xi} \rightarrow \Phi_g(u_\xi) \in E$  is  $C^2$ .
- (ii)  $\forall j = 1, \dots, m,$

$$\exists \nabla F_j(u_\xi) \in E \text{ such that } D_{u_\xi} F_j(w) = \langle \nabla F_j(u_\xi), w \rangle \quad \forall w \in E; \quad (2.13)$$

- (iii) There exists  $C > 0$  so that

$$\forall w \in E, \quad D_{u_\xi}^2 \mathcal{L}_\xi(w, w) \leq C \|w\|_E^2;$$

(iv) There exists  $c > 0$  so that

$$\forall w \in T_{u_\xi} \Sigma_{\mu_\xi} \cap (T_{u_\xi} \mathcal{O}_{u_{\mu_\xi}})^\perp, \quad D_{u_\xi}^2 \mathcal{L}_\xi(w, w) \geq c \|w\|_E^2, \quad (2.14)$$

where

$$(T_{u_\xi} \mathcal{O}_{u_\xi})^\perp = \{z \in E \mid \langle z, y \rangle = 0, \forall y \in T_{u_\xi} \mathcal{O}_{u_\xi}\}. \quad (2.15)$$

Then there exist  $\epsilon > 0, c > 0$  so that

$$\forall u \in \mathcal{O}_{u_\xi}, \forall u' \in \Sigma_{\mu_\xi}, \quad d(u, u') \leq \epsilon \Rightarrow H(u') - H(u) \geq cd^2(u', \mathcal{O}_{u_{\mu_\xi}}). \quad (2.16)$$

This result constitutes a generalization of Proposition 5 in [10]. In fact, if  $G_\mu$  is commutative, the latter result applies immediately. If not, the mild regularity condition (i) of the proposition suffices to obtain the result. We will give the proof of Proposition 2.3 in the next subsection. The basic message of this result is the following. If  $G$  is an invariance group for  $H$  that has a globally Hamiltonian action on  $E$  and if  $u_\xi$  satisfies the stationary equation  $D_{u_\xi} \mathcal{L}_\xi = 0$  for some  $\xi \in \mathfrak{g}$ , then, modulo the technical conditions of the proposition, the coercive estimate (2.14) on the Hessian of  $\mathcal{L}_\xi$  implies that the restriction of the Hamiltonian  $H$  to the constraint surface  $\Sigma_{\mu_\xi}$  attains a local minimum on the  $G_{\mu_\xi}$ -orbit  $\mathcal{O}_{u_\xi}$ . As explained in Section 8 of [10], modulo some further technical conditions, (2.16) allows one to show that  $u_\xi$  is  $G_{\mu_\xi}$ -orbitally stable. (See in particular Theorem 10 and Theorem 11 in [10]). We will give details in the examples of Section 8 below.

The difficulty in proving (2.14) comes from the fact that, in general, the bilinear symmetric form  $D_{u_\xi}^2 \mathcal{L}_\xi$  is not positive on  $E$ , but has instead a non-trivial negative cone

$$\mathcal{C}_- = \{v \in E \mid D_{u_\xi}^2 \mathcal{L}_\xi(v, v) < 0\}.$$

The estimate (2.14) implies that  $T_{u_\xi} \Sigma_{\mu_\xi}$  does not intersect  $\mathcal{C}_-$ . To prove this, we shall show that there exists a maximally negative subspace of  $E$  for  $D_{u_\xi}^2 \mathcal{L}_\xi$  which is  $D_{u_\xi}^2 \mathcal{L}_\xi$ -orthogonal to  $T_{u_\xi} \Sigma_{\mu_\xi}$ .

The main goal of this paper is to give a condition (see Theorem 3.1 (iv) and Theorem 6.3 (iv)), which is a generalization to the *Vakhitov-Kokolov slope condition*, that implies the coercivity estimate (2.14). This condition is in general easier to verify than the coercivity estimate itself and allows one to prove the orbital stability of relative equilibria of general Hamiltonian system. As an example of this method we study in Section 8 several applications of our results to the stability analysis of relative equilibria of nonlinear Schrödinger and Manakov equations.

**2.3. Proof of Proposition 2.3.** The general strategy of the proof is identical to the one of Proposition 5 in [10]. First, we need some simple preliminary results.

**Lemma 2.4.** *Suppose the hypotheses of Proposition 2.3 hold. Then  $\forall u \in \mathcal{O}_{u_\xi}$ , for all  $g \in G_{\mu_\xi}$ , we have*

$$\begin{aligned} (a) \quad & \Phi_g(T_u \mathcal{O}_{u_\xi}) = T_v \mathcal{O}_{u_\xi}, \quad \Phi_g\left((T_u \mathcal{O}_{u_\xi})^\perp\right) = (T_v \mathcal{O}_{u_\xi})^\perp, \\ (b) \quad & \Phi_g(T_u \Sigma_{\mu_\xi}) = T_v \Sigma_{\mu_\xi}, \quad \Phi_g\left((T_u \Sigma_{\mu_\xi})^\perp\right) = (T_v \Sigma_{\mu_\xi})^\perp, \end{aligned}$$

where  $v = \Phi_g(u)$ . In addition, defining  $V_u = \text{span}\{\nabla_j F(u) \mid j = 1 \dots m\}$ , we have

$$V_u^\perp = T_u \Sigma_{\mu_\xi} \quad \text{and} \quad E = T_u \Sigma_{\mu_\xi} \oplus V_u.$$

*Proof.* (a), respectively (b), follows from the observation that  $\Phi_g(\mathcal{O}_{u_\xi}) = \mathcal{O}_{u_\xi}$ , respectively  $\Phi_g(\Sigma_\mu) = \Sigma_\mu$ , and the fact that  $\Phi_g$  preserves the inner product  $\langle \cdot, \cdot \rangle$ . That  $V_u^\perp = T_u \Sigma_\mu$  follows from the definitions and the second statement is easily verified.  $\square$

*Proof. (of Proposition 2.3)* We start with some preliminaries. First of all, we prove that there exists  $\delta > 0$  so that, for all  $u' \in \Sigma_{\mu_\xi}$  for which  $d(u', \mathcal{O}_{u_\xi}) < \delta$ , there exists  $g \in G_{\mu_\xi}$  so that  $\Phi_g(u') - u_\xi \in (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp$ .

Let  $\delta > 0$  to be fixed later. We start by remarking that, since  $d(u', \mathcal{O}_{u_\xi}) < \delta$ , there exists  $v' \in \mathcal{O}_{u_\xi}$  such that  $\|u' - v'\|_E < \delta$ . Moreover, since  $v' \in \mathcal{O}_{u_\xi}$ , there exists  $g' \in G_{\mu_\xi}$  such that

$u_\xi = \Phi_{g'}(v')$ . It follows that  $\|\Phi_{g'}(u') - u_\xi\|_E < \delta$ . Next, let  $\eta_1, \dots, \eta_m$  be a basis of  $\mathfrak{g}_{\mu_\xi}$  and  $X_{\eta_i}(u_\xi)$  defined as in (2.3). We define

$$\begin{aligned} \mathcal{F} : E \times G_{\mu_\xi} &\rightarrow \mathbb{R}^m \\ (w, g) &\mapsto \begin{pmatrix} \langle \Phi_g(w) - u_\xi, X_{\eta_1}(u_\xi) \rangle \\ \dots \\ \langle \Phi_g(w) - u_\xi, X_{\eta_m}(u_\xi) \rangle \end{pmatrix}. \end{aligned} \quad (2.17)$$

Using Hypothesis B, we can write:

$$\langle \Phi_g(w) - u_\xi, X_{\eta_i}(u_\xi) \rangle = \langle w - \Phi_{g^{-1}}(u_\xi), \Phi_{g^{-1}}(X_{\eta_i}(u_\xi)) \rangle,$$

for all  $i = 1, \dots, m$ , it follows from hypothesis (i) of Proposition 2.3 and the fact that  $u_\xi \in \mathcal{D}$ , that  $\mathcal{F}$  is of class  $\mathcal{C}^1$ .

Next, remark that  $\mathcal{F}(u_\xi, e) = 0$ .

We now compute the partial derivative of  $\mathcal{F}$  along  $G_{\mu_\xi}$  at the point  $(u_\xi, e)$  denoted by

$$\partial_g \mathcal{F}(u_\xi, e) : \mathfrak{g}_{\mu_\xi} \rightarrow \mathbb{R}^m.$$

Note that  $\partial_g \mathcal{F}(u_\xi, e)$  is a  $m \times m$  matrix and the  $i, j$ -coefficient is obtained by writing  $g = \exp(t\eta_j)$  and computing

$$\begin{aligned} \frac{d}{dt} \mathcal{F}(u_\xi, \exp(t\eta_j))|_{t=0} &= \frac{d}{dt} \langle u_\xi, \Phi_{\exp(-t\eta_j)}(X_{\eta_i}(u_\xi)) \rangle|_{t=0} \\ &= - \langle u_\xi, X_{\eta_j}(X_{\eta_i}(u_\xi)) \rangle = \langle X_{\eta_j}(u_\xi), X_{\eta_i}(u_\xi) \rangle. \end{aligned}$$

Here we used Hypothesis B again. Since  $\mu_\xi$  is a regular value of  $F$ , it follows that the  $X_{\eta_i}$  are linearly independent. Hence,  $\partial_g \mathcal{F}(u_\xi, e)$  is invertible and we can apply the implicit function theorem to  $\mathcal{F}$ . As consequence, there exist  $\mathcal{V}_{u_\xi}$  a neighbourhood of  $u_\xi$  in  $E$ ,  $\mathcal{V}_e$  a neighbourhood of  $e$  in  $G_{\mu_\xi}$  and a function  $\Lambda : \mathcal{V}_{u_\xi} \rightarrow \mathcal{V}_e$  such that if  $v \in \mathcal{V}_{u_\xi}$  then there exists a unique  $g_v = \Lambda(v) \in \mathcal{V}_e \subset G_{\mu_\xi}$  such that  $\mathcal{F}(v, g_v) = 0$  which means  $\Phi_{g_v}(v) - u_\xi \in (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp$ .

Hence, by taking  $\delta$  sufficiently small, we have  $\Phi_{g'}(u') \in \mathcal{V}_{u_\xi}$  and we can conclude that there exists  $\tilde{g} = \Lambda(\Phi_{g'}(u'))$  such that  $\Phi_{\tilde{g}}(\Phi_{g'}(u')) - u_\xi \in (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp$ . Finally, by defining  $g = \tilde{g}g'$ , we obtain  $\Phi_g(u') - u_\xi \in (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp$ .

Let  $u'' = \Phi_g(u')$ . Then  $u'' - u_\xi \in (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp$  and, thanks to Lemma 2.4, we can write

$$u'' - u_\xi = (u'' - u_\xi)_1 + (u'' - u_\xi)_2,$$

where  $(u'' - u_\xi)_1 \in T_{u_\xi} \Sigma_\mu \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp$  and  $(u'' - u_\xi)_2 \in V_{u_\xi}$ . Since  $u'', u_\xi \in \Sigma_\mu$ , one has

$$0 = F(u'') - F(u_\xi) = D_{u_\xi} F((u'' - u_\xi)_2) + O(\|u'' - u_\xi\|_E^2).$$

As  $D_{u_\xi} F$  is of maximal rank, it has no kernel in  $V_{u_\xi}$ , and we conclude there exists  $c_0$  so that

$$\|(u'' - u_\xi)_2\|_E \leq O(\|u'' - u_\xi\|_E^2).$$

Hence,

$$\|(u'' - u_\xi)_1\|_E \geq C\|u'' - u_\xi\|_E$$

since  $\|u'' - u_\xi\|_E$  is small by construction.



We can then conclude the proof as follows. Let  $\epsilon > 0$  be small enough so that the previous inequalities hold. Then compute

$$\begin{aligned}
\mathcal{L}_\xi(u') - \mathcal{L}_\xi(u_\xi) &= \mathcal{L}_\xi(u'') - \mathcal{L}_\xi(u_\xi) \\
&= D_{u_\xi} \mathcal{L}_\xi(u'' - u_\xi) + \frac{1}{2} D_{u_\xi}^2 \mathcal{L}_\xi(u'' - u_\xi, u'' - u_\xi) \\
&\quad + o(\|u'' - u_\xi\|_E^2) \\
&= \frac{1}{2} D_{u_\xi}^2 \mathcal{L}_\xi((u'' - u_\xi)_1, (u'' - u_\xi)_1) + O(\|u'' - u_\xi\|_E^3) \\
&\quad + o(\|u'' - u_\xi\|_E^2) \\
&= \frac{1}{2} D_{u_\xi}^2 \mathcal{L}_\xi((u'' - u_\xi)_1, (u'' - u_\xi)_1) + o(\|u'' - u_\xi\|_E^2) \\
&\geq \frac{c}{2} \|(u'' - u_\xi)_1\|_E^2 + o(\|u'' - u_\xi\|_E^2) \\
&\geq \tilde{c} \|u'' - u_\xi\|_E^2 \geq \tilde{c} d^2(u'', \mathcal{O}_{u_\xi}) = \tilde{c} d^2(u', \mathcal{O}_{u_\xi}).
\end{aligned}$$

Note that, in the first equality above, we used the observation that, for all  $g \in G_{\mu_\xi}$ , for all  $u' \in \Sigma_{\mu_\xi}$ , one has

$$\mathcal{L}_\xi(\Phi_g(u')) = \mathcal{L}_\xi(u').$$

This follows from the  $G$ -invariance of  $H$  and from the fact that

$$\xi \cdot F(\Phi_g(u')) = \xi \cdot \mu_\xi = \xi \cdot F(u')$$

since both  $\Phi_g(u')$  and  $u'$  belong to  $\Sigma_{\mu_\xi}$ .  $\square$

### 3. MAIN RESULT: THE HILBERT SPACE SETTING

In this section, we state our main result (Theorem 3.1) in the setting where  $E$  is a Hilbert space, with inner product  $\langle \cdot, \cdot \rangle_E$ , and  $\|\cdot\|_E = \sqrt{\langle \cdot, \cdot \rangle_E}$ . In particular, the inner product  $\langle \cdot, \cdot \rangle$  in (2.11) and in Hypothesis A is, in this section, equal to  $\langle \cdot, \cdot \rangle_E$ . The Hilbert space structure will be used mainly to control the Hessian of  $\mathcal{L}_\xi$  through the spectral analysis of the associated bounded self-adjoint operator  $\nabla^2 \mathcal{L}_\xi$  (see below). This makes for a simpler statement and proof than in the more general setting of Section 6, and allows for an easier understanding of the philosophy of the result. We point out that the result we obtain in Theorem 3.1 may be of interest also in finite dimensional problems ( $\dim E < +\infty$ ). Indeed, the usual orbital stability results in the literature on finite dimensional Hamiltonian dynamical systems reduce their proof to the coercivity estimate (2.14) on the Hessian of  $\mathcal{L}_\xi$ , which is of dimension  $\dim E$ . We reduce the problem here to a control on the Hessian of the function  $W$  (see (3.3)), which is of dimension  $m = \dim G$ , typically much lower than  $\dim E$ , especially when the latter is high-dimensional.

We start with some preliminaries. We will make use of the following hypothesis:

**Hypothesis C.** There exists an open set  $\Omega \subset \mathfrak{g} \simeq \mathbb{R}^m$  and a map  $\tilde{u} \in C^1(\Omega \subset \mathfrak{g}, E)$

$$\tilde{u} : \xi \in \Omega \subset \mathfrak{g} \rightarrow u_\xi \in \mathcal{D} \subset E \tag{3.1}$$

satisfying, for all  $\xi \in \Omega$ ,

$$D_{u_\xi} H - \xi \cdot D_{u_\xi} F = 0. \tag{3.2}$$

As recalled in section 2, if  $u_\xi$  is a solution to (3.2), then  $u_\xi$  is  $G_{\mu_\xi}$ -relative equilibrium with  $\mu = \mu_\xi = F(u_\xi)$ .

So our starting point is equation (3.2), which in PDE applications is often an elliptic partial differential equation and we suppose we have an  $m$ -parameter family of solutions, indexed by  $\xi$ . One of the major difficulties to apply the theory is of course to find such families of solutions. In many cases, one has one single such solution for  $\xi_* \in \mathfrak{g}$  and one needs to ensure there exists a neighbourhood  $\Omega$  of  $\xi_*$  for which such solutions exist. We will come back to this property

of ‘‘persistence of relative equilibria’’ in Section 7 and present results ensuring Hypothesis C is satisfied. For now, we will suppose this is the case. Next, consider the Lyapunov function  $\mathcal{L}_\xi$  defined by (2.10) and remark that each  $u_\xi$  solution to (3.2) is a critical point of  $\mathcal{L}_\xi$ . Moreover, define for all  $\xi \in \Omega \subset \mathfrak{g}$ , the map

$$W : \xi \in \Omega \subset \mathfrak{g} \rightarrow \mathcal{L}_\xi(u_\xi) \in \mathbb{R}. \quad (3.3)$$

Note that

$$W(\xi) = H(u_\xi) - \xi \cdot \hat{F}(\xi),$$

where

$$\hat{F} : \xi \in \Omega \subset \mathfrak{g} \rightarrow F(u_\xi) \in \mathfrak{g}^* \simeq \mathbb{R}^m. \quad (3.4)$$

For each  $\xi \in \Omega$ , the Hessian  $D_\xi^2 W$  of  $W$  is a bilinear form on  $\mathbb{R}^m$ . Hence, we can consider the following decomposition

$$\mathbb{R}^m = W_- \oplus W_0 \oplus W_+,$$

where  $W_0$  is the kernel of  $D_\xi^2 W$  and where  $D_\xi^2 W$  is positive (negative) definite on  $W_+$  ( $W_-$ ). Let  $d_0(D_\xi^2 W), p(D_\xi^2 W), n(D_\xi^2 W)$  be the dimensions of these spaces. Note that the decomposition is not unique, but the respective dimensions are. In other words,  $W_\pm$  are maximal positive/negative definite spaces for  $D_\xi^2 W$ . Also, in order not to burden the notation, we have not made the  $\xi$ -dependence of the spaces  $W_0, W_\pm$  explicit. Recall that, given a symmetric bilinear form  $B$  on a Banach space  $E$ , a subspace  $\mathcal{X}$  of  $E$  is said to be a positive (negative) definite subspace for  $B$  on  $E$  if  $B|_{\mathcal{X} \times \mathcal{X}}$  is positive (negative) definite. A subspace is maximally positive (negative) definite if it is positive (negative) definite and it is not contained in any other positive (negative) definite subspace.

Similarly, the Hessian  $D_u^2 \mathcal{L}_\xi$  of  $\mathcal{L}_\xi$  is a symmetric bilinear form on  $E$ . For each  $u \in E$ , we define as usual the Morse index  $n(D_u^2 \mathcal{L}_\xi)$  of  $u$  for  $\mathcal{L}_\xi$  as the dimension of a maximally negative definite subspace for  $D_u^2 \mathcal{L}_\xi$  in  $E$ .

Finally, when  $E$  is a Hilbert space, we can define for each  $u \in E$  a bounded self-adjoint operator  $\nabla^2 \mathcal{L}_\xi(u)$  by

$$\langle v, \nabla^2 \mathcal{L}_\xi(u)w \rangle_E = D_u^2 \mathcal{L}_\xi(v, w). \quad (3.5)$$

As a consequence, we can consider the spectral decomposition of  $E$  for  $\nabla^2 \mathcal{L}_\xi(u_\xi)$

$$E = E_- \oplus E_0 \oplus E_+ \quad (3.6)$$

with  $E_0 = \text{Ker} \nabla^2 \mathcal{L}_\xi(u_\xi) = \text{Ker} D_{u_\xi}^2 \mathcal{L}_\xi$ , and  $E_\pm$  the positive and negative spectral subspaces of  $\nabla^2 \mathcal{L}_\xi(u)$ . Clearly  $E_\pm$  are maximally positive/negative subspaces for  $\nabla_{u_\xi}^2 \mathcal{L}_\xi$  so that  $n(D_{u_\xi}^2 \mathcal{L}_\xi) = \dim E_-$ . We can now state our main result.

**Theorem 3.1.** *Suppose  $(E, \langle \cdot, \cdot \rangle_E)$  is a Hilbert space and that Hypotheses A and C hold. Let  $\xi \in \Omega$  and suppose*

- (i)  $D_\xi^2 W$  is non-degenerate,
- (ii)  $\text{Ker} D_{u_\xi}^2 \mathcal{L}_\xi = T_{u_\xi} \mathcal{O}_{u_\xi}$ ,
- (iii)  $\inf(\sigma(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap (0, +\infty)) > 0$ ,
- (iv)  $p(D_\xi^2 W) = n(D_{u_\xi}^2 \mathcal{L}_\xi)$ .

Then there exists  $\delta > 0$  such that

$$\forall v \in T_{u_\xi} \Sigma_{\mu_\xi} \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp, \quad D_u^2 \mathcal{L}_\xi(v, v) \geq \delta \|v\|_E^2. \quad (3.7)$$

We will say a relative equilibrium is *non-degenerate* when  $D_\xi^2 W$  is non-degenerate. Since (3.7) is the same as (2.14), one can then use Proposition 2.3 together with Theorems 10 and 11 of [10] to show that  $u_\xi$  is orbitally stable. It is the fourth condition of the above theorem that generalizes the Vakhitov-Kolokolov slope condition, as we now explain. Suppose the group  $G$  is 1-dimensional, so that  $m = 1$ . Then  $W$  is a scalar function of  $\xi \in \mathbb{R} \simeq \mathfrak{g}$ . We will see below (See (4.13)) that

$$W''(\xi) = -\hat{F}'(\xi).$$

Hence the proof of orbital stability for  $u_\xi$  reduces to verifying that the spectral conditions on  $\nabla_\xi^2 \mathcal{L}_\xi$  are satisfied and notably that  $\dim E_- = 1$ , and that

$$\hat{F}'(\xi) < 0. \quad (3.8)$$

This is the Vakhitov-Kolokov slope condition. In applications to the Schrödinger equation, where  $F(u) = \frac{1}{2}\langle u, u \rangle$ , it says that the norm of  $u_\xi$  decreases as a function of  $\xi$ . In the case  $m = 1$ , the above result is proven in [17] (Corollary 3.3.1) and in [34] (Proposition 5.2).

The setup of the Hamiltonian dynamics with higher dimensional symmetry on a Hilbert space we used in this section is similar to the one proposed in [18] where the decomposition (3.6) of the bounded self-adjoint operator  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  as well as condition (iii) of Theorem 3.1 are also used to obtain a coercivity result of the type (3.7). Nevertheless, in [18] a different constraint surface and orbit are used and some of the arguments provided are incomplete: for a complete comparison between Theorem 3.1 and the coercivity results of [18], we refer to Section 9.

We finally note that, when  $E$  is infinite dimensional, and the equation under study a PDE, the more general formulation of Section 6 is often considerably more pertinent than the simpler Hilbert space formulation proposed here. Indeed, the operator  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  introduced in Theorem 3.1 is not a partial differential operator (it is bounded) making the analysis of its spectrum generally less convenient than for the operator  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  in Theorem 6.5, which in applications is a self-adjoint partial differential operator on a suitable auxiliary Hilbert space. We will come back to this point when treating examples in Section 8.

#### 4. USEFUL LEMMAS

The following lemmas collect some basic properties of the objects introduced above, that are essential in the further analysis of the Hessian of the Lyapunov function. We define, for  $\xi \in \Omega$ ,

$$\mathcal{U}_\xi = \{\eta \cdot \nabla_\xi u_\xi \in E \mid \eta \in \mathbb{R}^m\}, \quad (4.1)$$

where we used the notation

$$\eta \cdot \nabla_\xi u_\xi := D_\xi \tilde{u}(\eta). \quad (4.2)$$

**Lemma 4.1.** *Let  $E$  be a Banach space,  $\Omega$  an open set in  $\mathfrak{g}$ . Let  $\tilde{u} \in C^1(\Omega \subset \mathfrak{g}, E)$ . Let  $\xi \in \Omega$  and consider the following statements:*

- (1)  $\hat{F}$  is a local diffeomorphism;
- (2)  $D_\xi \tilde{u}$  is injective.
- (3)  $\mathcal{U}_\xi \cap \text{Ker} D_{u_\xi} F = \{0\}$ .
- (4) There is a neighbourhood of  $u_\xi$  where the moment map  $F$  is regular (i.e.  $D_{u_\xi} F$  has maximal rank).
- (5)

$$\mathcal{U}_\xi \oplus \text{Ker} D_{u_\xi} F = E; \quad (4.3)$$

Then (1)  $\Leftrightarrow$  ((2) and (3))  $\Leftrightarrow$  ((4) and (5)).

Note that the lemma does not use the fact that the  $u_\xi$  are solutions to the stationary equation:  $\tilde{u}$  takes values in  $E$ , without further condition. The lemma therefore strings together some useful facts on compositions of maps.

It is easy to see that, whenever  $u_\xi$  is a solution to (3.2) for every  $\xi \in \Omega$ , the map  $\tilde{u}$  is injective provided the  $u_\xi$  are regular relative equilibria. Indeed, if  $u_{\xi_1} = u_{\xi_2}$  are both solutions of (3.2), then

$$(\xi_1 - \xi_2) \cdot D_{u_{\xi_1}} F = 0.$$

Hence, if the  $u_\xi$  are regular relative equilibria (see Definition 2.1), one has  $\xi_1 = \xi_2$ . It is natural in that context to assume it is in fact an immersion, meaning that its derivative is injective, as in condition (2) of Lemma 4.1. One can then think of  $\tilde{u}(\Omega)$  as an  $m$ -parameter surface in  $E$ . In applications, this additional condition often arises naturally from the construction of  $\tilde{u}$ , as seen in

Section 7. It is also a consequence of assumption (i) in Theorem 3.1, as a result of Lemma 4.2 (1) together with Lemma 4.1 (1).

*Proof.* First note that, for all  $\eta_1, \eta_2 \in \mathfrak{g}$ ,

$$D_\xi(\eta_2 \cdot \hat{F})(\eta_1) = D_{u_\xi}(\eta_2 \cdot F)(\eta_1 \cdot \nabla_\xi u_\xi). \quad (4.4)$$

**(1)  $\Rightarrow$  ((2) and (3))** Let  $\eta_1 \in \mathfrak{g}$  and suppose  $\eta_1 \cdot \nabla_\xi u_\xi = 0$ . It follows from (4.4) that  $D_\xi \hat{F}(\eta_1) = 0$ . But since  $\hat{F}$  is a local diffeomorphism at  $\xi$ , this implies  $\eta_1 = 0$ . Hence  $D_\xi \tilde{u}$  is injective, which shows (2). To show (3), let  $\eta_1 \in \mathfrak{g}$  and suppose  $\eta_1 \cdot \nabla_\xi u_\xi \in \text{Ker} D_{u_\xi} F$ . Then, by definition,  $D_{u_\xi}(\eta_2 \cdot F)(\eta_1 \cdot \nabla_\xi u_\xi) = 0$  for all  $\eta_2 \in \mathfrak{g}$ . It follows from (4.4) that  $\eta_1 \in \text{Ker} D_\xi \hat{F}$  so that, by (1),  $\eta_1 = 0$ . This proves (3).

**((2) and (3))  $\Rightarrow$  (1)** Let  $\eta_1 \in \text{Ker} D_\xi \hat{F}$ . Then according to the above equality,  $\eta_1 \cdot \nabla_\xi u_\xi \in \text{Ker} D_{u_\xi} F$ . So, by (3),  $\eta_1 \cdot \nabla_\xi u_\xi = 0$  and by (2),  $\eta_1 = 0$ . This proves  $D_\xi \hat{F}$  is injective, hence surjective, which proves (1).

**((2) and (3))  $\Rightarrow$  ((4) and (5))** According to (3), the map

$$D_{u_\xi} F : \mathcal{U}_\xi \rightarrow \mathfrak{g}^* \simeq \mathbb{R}^m$$

is injective. But since by (2),  $D_\xi \tilde{u}$  is injective, the dimension of  $\mathcal{U}_\xi$  is  $m$ . Hence this map is a bijection. The rank of  $D_{u_\xi} F$  is therefore maximal. By continuity of  $D_u F$  in  $u$ , this remains true in a neighbourhood of  $u_\xi$ , which proves (4). It follows from (4) that locally,  $\Sigma_{\mu_\xi}$  is a co-dimension  $m$  submanifold of  $E$ . Since, by definition,  $T_{u_\xi} \Sigma_{\mu_\xi} = \text{Ker} D_{u_\xi} F$ , we know from (3) that  $\mathcal{U}_\xi \cap T_{u_\xi} \Sigma_{\mu_\xi} = \{0\}$ . Since, by (2), the dimension of  $\mathcal{U}_\xi$  is  $m$ , (5) follows.

**((4) and (5))  $\Rightarrow$  ((2) and (3)).** This is obvious and concludes the proof of the lemma.  $\square$

We introduce

$$G_\xi = \{g \in G \mid \text{Ad}_g \xi = \xi\}, \quad (4.5)$$

which is the subgroup of  $G$  for which  $\xi$  is a fixed point under the adjoint action. We will write  $\mathfrak{g}_\xi$  for its Lie-algebra. We furthermore need ( $u_\xi \in \mathcal{D}$ )

$$Z_\xi = \{X_\eta(u_\xi) \mid \eta \in \mathfrak{g}_\xi\} \subset E \quad (4.6)$$

where  $X_\eta(u_\xi)$  is defined in (2.3).

**Lemma 4.2.** *Suppose Hypotheses A and C hold. Let  $\xi \in \Omega$ . Then, one has:*

(1) For all  $\eta \in \mathbb{R}^m$ ,

$$\eta \in W_0 = \text{Ker} D_\xi^2 W \Leftrightarrow \eta \in \text{Ker} D_\xi \hat{F} \Leftrightarrow \eta \cdot \nabla_\xi u_\xi \in \text{Ker} D_{u_\xi} F. \quad (4.7)$$

In particular,  $D_\xi^2 W$  is non-degenerate if and only if  $\hat{F}$  is a local diffeomorphism at  $\xi$ .

(2) For all  $\eta_1, \eta_2 \in \mathbb{R}^m$ ,

$$D_{u_\xi}^2 \mathcal{L}_\xi(\eta_1 \cdot \nabla_\xi u_\xi, \eta_2 \cdot \nabla_\xi u_\xi) = -D_\xi^2 W(\eta_1, \eta_2). \quad (4.8)$$

(3) For all  $v \in \text{Ker} D_{u_\xi} F$ , for all  $\eta \in \mathfrak{g}$ ,

$$D_{u_\xi}^2 \mathcal{L}(\eta \cdot \nabla_\xi u_\xi, v) = 0. \quad (4.9)$$

(4)  $T_{u_\xi} \mathcal{O}_{u_\xi}$  is a subspace of the kernel of  $(D_{u_\xi}^2 \mathcal{L}_\xi \mid \text{Ker} D_{u_\xi} F)$ , which is the restriction of  $D_{u_\xi}^2 \mathcal{L}_\xi$  to  $\text{Ker} D_{u_\xi} F \times \text{Ker} D_{u_\xi} F$ .

(5)  $Z_\xi \subset \text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) \subset \text{Ker}(D_{u_\xi} F)$ .

(6)  $\mathfrak{g}_\xi \subset \mathfrak{g}_{\mu_\xi}$ .

Note that, combining (4.7) with Lemma 4.1, we can conclude that the directions  $\eta \cdot \nabla_\xi u_\xi$  form a complementary subspace to  $\text{Ker} D_{u_\xi} F$  when  $D_\xi^2 W$  is non-degenerate. Also,  $u_\xi$  is a regular relative equilibrium, and the subspace  $\mathcal{U}_\xi$  is complementary to the tangent space  $T_{u_\xi} \Sigma_{\mu_\xi}$ .

Expression (4.8) is of interest since it identifies part of the Hessian of the Lyapunov function  $\mathcal{L}_\xi$  in terms of the Hessian of  $W$ . More precisely, it is useful to determine a subspace of negative

directions of  $D_{u_\xi}^2 \mathcal{L}_\xi$ . Indeed, if  $n_+ = p(D_\xi^2 W)$  and if  $\{\eta_1, \dots, \eta_{n_+}\}$  is a family of linearly independent elements of  $\mathbb{R}^m$  such that  $\text{span}\{\eta_1, \dots, \eta_{n_+}\}$  is a positive definite subspace for  $D_\xi^2 W$ , then  $\text{span}\{\eta_1 \cdot \nabla_\xi u_\xi, \dots, \eta_{n_+} \cdot \nabla_\xi u_\xi\}$  is a negative definite subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$  (see (4.8)). Thus, the dimension of a maximally negative definite subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$  in  $E$  is at least  $p(D_\xi^2 W)$ :

$$n(D_{u_\xi}^2 \mathcal{L}_\xi) \geq p(D_\xi^2 W). \quad (4.10)$$

Expression (4.9) turns out to be crucial in what follows: it expresses the fact that  $\mathcal{U}_\xi = \{\eta \cdot \nabla_\xi u_\xi \mid \eta \in \mathbb{R}^m\}$  is  $D_{u_\xi}^2 \mathcal{L}_\xi$ -orthogonal to  $\text{Ker } D_{u_\xi} F$ .

*Proof.* First of all, note that, since  $u_{\xi'}$  is a solution to the stationary equation (3.2) for all  $\xi' \in \Omega$ , for all  $\eta \in \mathbb{R}^m$

$$D_\xi W(\eta) = -\eta \cdot F(u_\xi). \quad (4.11)$$

Then a straightforward calculation gives, for all  $\eta_1, \eta_2 \in \mathbb{R}^m$ ,

$$D_\xi^2 W(\eta_1, \eta_2) = -D_{u_\xi} \eta_1 \cdot F(\eta_2 \cdot \nabla_\xi u_\xi). \quad (4.12)$$

In other words,

$$D_\xi^2 W = -D_\xi \hat{F}. \quad (4.13)$$

Note that, as  $\hat{F}$  is a map from  $\mathbb{R}^m \simeq \mathfrak{g}$  to  $\mathbb{R}^m \simeq \mathfrak{g}^*$ ,  $D_\xi \hat{F}$  is linear from  $\mathbb{R}^m = \mathfrak{g}$  to  $\mathbb{R}^m = \mathfrak{g}^*$ . It therefore naturally defines a bilinear map on  $\mathbb{R}^m \simeq \mathfrak{g}$ . In our notation here, we identify  $\mathfrak{g}$  with  $\mathfrak{g}^*$  using an Euclidean structure, but even without the latter, the above is natural.

The first statement of (4.7) is now obvious and for the second, note that  $\eta \cdot \nabla_\xi u_\xi \in \text{Ker } D_{u_\xi} F$  if and only if, for all  $\eta' \in \mathbb{R}^m$ ,  $D_{u_\xi} \eta' \cdot F(\eta \cdot \nabla_\xi u_\xi) = 0$ , which yields the conclusion, thanks to (4.4) and (4.12).

To obtain (4.8), it is sufficient to take the derivative of the stationary equation (3.2) with respect to  $\xi \in \mathfrak{g}$  and to use (4.12). More precisely, by taking this derivative with respect to  $\xi$  in the direction  $\eta$ , we obtain for all  $\eta \in \mathfrak{g}$ ,

$$D_{u_\xi}^2 \mathcal{L}_\xi(\eta \cdot \nabla_\xi u_\xi) = D_{u_\xi} \eta \cdot F \in E^*. \quad (4.14)$$

Hence, using (4.12),

$$D_{u_\xi}^2 \mathcal{L}_\xi(\eta_1 \cdot \nabla_\xi u_\xi, \eta_2 \cdot \nabla_\xi u_\xi) = D_{u_\xi} \eta_1 \cdot F(\eta_2 \cdot \nabla_\xi u_\xi) = -D_\xi^2 W(\eta_1, \eta_2).$$

Next, (4.9) follows directly from (4.14). Indeed, for all  $v \in \text{Ker } D_{u_\xi} F$  and for all  $\eta \in \mathfrak{g}$ ,  $D_{u_\xi}^2 \mathcal{L}_\xi(\eta \cdot \nabla_\xi u_\xi, v) = D_{u_\xi} \eta \cdot F(v) = 0$ .

From  $F = (F \circ \Phi_{g^{-1}}) \circ \Phi_g$  and  $H = H \circ \Phi_g$  one finds, for all  $u \in E, g \in G$ ,

$$D_u F = (D_{\Phi_g(u)}(F \circ \Phi_{g^{-1}})) D_u \Phi_g, \quad D_u H = D_{\Phi_g(u)} H D_u \Phi_g.$$

Hence, by (3.2),

$$\begin{aligned} D_{\Phi_g(u_\xi)} H &= \xi \cdot D_{u_\xi} F (D_{u_\xi} \Phi_g)^{-1} = D_{\Phi_g(u_\xi)} (\xi \cdot F \circ \Phi_{g^{-1}}) \\ &= D_{\Phi_g(u_\xi)} (Ad_g \xi \cdot F) \end{aligned}$$

and therefore

$$D_{\Phi_g(u_\xi)} (H - \xi \cdot F) = D_{\Phi_g(u_\xi)} ((Ad_g \xi - \xi) \cdot F).$$

Now let  $\eta \in \mathfrak{g}$ , consider  $g = \exp(t\eta)$  and take the derivative at  $t = 0$  in the previous relation. One finds, for all  $v \in E$ ,

$$D_{u_\xi}^2 (H - \xi \cdot F)(X_\eta(u_\xi), v) = D_{u_\xi} [\eta, \xi] \cdot F(v). \quad (4.15)$$

Taking  $v \in \text{Ker } D_{u_\xi} F$ , the right hand side above vanishes for any  $\eta \in \mathfrak{g}$ , and one finds (4) follows.

To prove (5), note that, taking  $\eta \in \mathfrak{g}_\xi$  so that  $[\eta, \xi] = 0$ , we see that  $Z_\xi \subset \text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi$ . Finally, let  $v \in E$ , then (4.14) yields, for all  $\eta \in \mathfrak{g}$

$$D_{u_\xi}^2 \mathcal{L}_\xi(\eta \cdot \nabla_\xi u_\xi, v) = \eta \cdot D_{u_\xi} F(v).$$

Hence, if  $v \in \text{Ker } (D_{u_\xi}^2 \mathcal{L}_\xi)$ , it follows that  $v \in \text{Ker } (D_{u_\xi} F)$ .

To prove (6) note first that (5) implies that, if  $\eta \in \mathfrak{g}_\xi$ , we have

$$0 = D_{u_\xi} F(X_\eta(u_\xi)) = \frac{d}{dt} F \circ \Phi_{\exp(\eta t)}(u_\xi)|_{t=0} = \frac{d}{dt} \text{Ad}_{\exp(\eta t)}^* F(u_\xi)|_{t=0} = \text{ad}_\eta^* \mu_\xi.$$

It follows that  $\eta \in \mathfrak{g}_{\mu_\xi}$ . Hence  $\mathfrak{g}_\xi \subset \mathfrak{g}_{\mu_\xi}$ .  $\square$

Finally, we state some properties of symmetric bilinear forms and their associated quadratic forms in the form of a short lemma. In what follows, if  $B$  is a bilinear form on some vectorspace  $E$ , and  $\mathcal{Y}$  is a subspace of  $E$ , then we write  $(B|_{\mathcal{Y}})$  for the restriction of  $B$  to  $\mathcal{Y} \times \mathcal{Y}$ .

**Lemma 4.3.** *Let  $E$  be a vector space and  $B$  a symmetric bilinear form on  $E$ .*

- (i) *Let  $\mathcal{X}_-$  be a maximally negative definite subspace for  $B$  in  $E$ . Suppose  $\mathcal{Y} \subset E$  is a subspace of  $E$  with the property that  $\mathcal{X}_- \cap \mathcal{Y} = \{0\}$  and such that  $B(\mathcal{X}_-, \mathcal{Y}) = 0$ . Then  $\mathcal{Y}$  is a positive subspace for  $B$ .*
- (ii) *Let  $\mathcal{Y}_1, \mathcal{Y}_2$  be two subspaces of  $\mathcal{Y}$ , such that  $B(\mathcal{Y}_1, \mathcal{Y}_2) = 0$ . Then*

$$\text{Ker}(B|_{\mathcal{Y}_1 + \mathcal{Y}_2}) = \text{Ker}(B|_{\mathcal{Y}_1}) + \text{Ker}(B|_{\mathcal{Y}_2}). \quad (4.16)$$

- (iii) *Let  $\mathcal{Y}$  be a positive subspace for  $B$ . Suppose  $u \in \mathcal{Y}$  satisfies  $B(u, u) = 0$ . Then  $u \in \text{Ker}(B|_{\mathcal{Y}})$ .*

We say  $\mathcal{Y}$  is a positive subspace for  $B$  if for all  $y \in \mathcal{Y}$ ,  $B(y, y) \geq 0$ . Note that the  $B$ -orthogonality of the subspaces is crucial in parts (i) and (ii).

*Proof.* (i) Suppose the statement is false, then there exists  $y \in \mathcal{Y}$ , so that  $B(y, y) < 0$ . Clearly,  $y \neq 0$  and hence, by assumption,  $y \notin \mathcal{X}_-$ . Now consider  $\mathcal{Z} = \text{span}\{y, \mathcal{X}_-\}$ . Let  $0 \neq z \in \mathcal{Z}$ . Then, there exist  $\lambda \in \mathbb{R}$  and  $z_- \in \mathcal{X}_-$ , not both zero, so that  $z = \lambda y + z_-$ . It follows from the  $B$  orthogonality of  $\mathcal{X}_-$  and  $\mathcal{Y}$  that

$$B(z, z) = \lambda^2 B(y, y) + B(z_-, z_-) < 0.$$

Hence  $B$  is negative definite on  $\mathcal{Z}$ . Since  $\dim \mathcal{X}_- \subsetneq \mathcal{Z}$  this is a contradiction.

(ii) Immediate.

(iii) One has, for all  $v \in \mathcal{Y}$ , and for all  $\lambda \in \mathbb{R}$ ,

$$0 \leq B(\lambda u + v, \lambda u + v) = 2\lambda B(u, v) + B(v, v).$$

If  $B(u, v) \neq 0$ , this is a contradiction.  $\square$

## 5. PROOF OF THEOREM 3.1

Let  $n_- = \dim E_-$  and  $\{\eta_1, \dots, \eta_{n_-}\} \subset \mathbb{R}^m$  a family of linearly independent elements of  $\mathbb{R}^m$  such that  $D_\xi^2 W(\eta, \eta) > 0$  for all  $\eta \in \text{span}\{\eta_1, \dots, \eta_{n_-}\}$ . As a consequence of (4.8),  $\mathcal{X}_- := \text{span}\{\eta_1 \cdot \nabla_\xi u_\xi, \dots, \eta_{n_-} \cdot \nabla_\xi u_\xi\}$  is a negative definite subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$  in  $E$ . Moreover, since  $\dim \mathcal{X}_- = n_-$ ,  $\mathcal{X}_-$  is a maximally negative definite subspace.

Next, since  $D_\xi^2 W$  is non-degenerate by hypothesis, Lemma 4.2 (1) implies that  $\hat{F}$  is a local diffeomorphism. Hence, thanks to Lemma 4.1,  $\mathcal{X}_- \cap T_{u_\xi} \Sigma_{\mu_\xi} = \{0\}$ . Moreover, thanks to (4.9),  $\mathcal{X}_-$  and  $T_{u_\xi} \Sigma_{\mu_\xi}$  are  $D_{u_\xi}^2 \mathcal{L}_\xi$ -orthogonal. As a consequence, we can apply Lemma 4.3 (i) and conclude that  $T_{u_\xi} \Sigma_{\mu_\xi}$  is a positive subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$ .

Furthermore, since by hypothesis (ii) of the theorem,  $E_0 = \text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi = T_{u_\xi} \mathcal{O}_{u_\xi}$ , it follows from Lemma 4.3 (ii)-(iii) that

$$\mathcal{Y} := T_{u_\xi} \Sigma_{\mu_\xi} \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp \quad (5.1)$$

is a positive definite subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$ , meaning that

$$D_{u_\xi}^2 \mathcal{L}_\xi(v, v) > 0, \quad \forall v \in \mathcal{Y} = T_{u_\xi} \Sigma_{\mu_\xi} \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp, v \neq 0. \quad (5.2)$$

Here, and in the rest of the proof, the  $\perp$  denotes orthogonality with respect to the inner product  $\langle \cdot, \cdot \rangle_E$ .

To obtain the desired coercive bound, we now use the spectral information on  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  provided by the hypotheses of the theorem. Note first that, since  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  is self-adjoint,  $E_0^\perp = (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp \subset E$  is an invariant subspace for  $\nabla^2 \mathcal{L}_\xi(u_\xi)$ : it is the spectral subspace of  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  corresponding to  $\mathbb{R} \setminus \{0\}$ .

Let  $\nabla F_j(u_\xi) \in E$  be defined by  $D_{u_\xi} F_j(v) = \langle \nabla F_j(u_\xi), v \rangle_E$  for  $j = 1, \dots, m$ . Then, since  $\mu_\xi$  is a regular value,

$$(T_{u_\xi} \Sigma_{\mu_\xi})^\perp = \text{span} \{ \nabla F_j(u_\xi) \}_{j=1, \dots, m}.$$

and  $\dim (T_{u_\xi} \Sigma_{\mu_\xi})^\perp = m$ . Moreover, since  $T_{u_\xi} \mathcal{O}_{u_\xi} \subset T_{u_\xi} \Sigma_{\mu_\xi}$ , one obtains the following orthogonal decomposition of  $E_0^\perp$ :

$$E_0^\perp = (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp = \mathcal{Y} \oplus (T_{u_\xi} \Sigma_{\mu_\xi})^\perp.$$

Note that  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  does not leave this decomposition invariant: we are interested in controlling it on  $\mathcal{Y}$ .

For that purpose, let  $P$  be the projection on  $\mathcal{Y} \subset E_0^\perp$  and consider the following decomposition of the restriction of the operator  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  to  $E_0^\perp$ ,

$$\begin{aligned} \nabla^2 \mathcal{L}_\xi(u_\xi)|_{E_0^\perp} &= P \nabla^2 \mathcal{L}_\xi(u_\xi) P + P \nabla^2 \mathcal{L}_\xi(u_\xi) (I_{E_0^\perp} - P) \\ &\quad + (I_{E_0^\perp} - P) \nabla^2 \mathcal{L}_\xi(u_\xi) P + (I_{E_0^\perp} - P) \nabla^2 \mathcal{L}_\xi(u_\xi) (I_{E_0^\perp} - P). \end{aligned}$$

Since  $(T_{u_\xi} \Sigma_{\mu_\xi})^\perp$  is finite dimensional, the projector  $I_{E_0^\perp} - P$  is finite rank. Hence  $\nabla^2 \mathcal{L}_\xi(u_\xi)|_{E_0^\perp} = P \nabla^2 \mathcal{L}_\xi(u_\xi) P + K$  with  $K$  a compact operator and it follows that  $\sigma_{\text{ess}}(P \nabla^2 \mathcal{L}_\xi(u_\xi) P) = \sigma_{\text{ess}}(\nabla^2 \mathcal{L}_\xi(u_\xi)|_{E_0^\perp})$ . Here  $\sigma_{\text{ess}}(\cdot)$  designates the essential spectrum. In particular,  $0 \notin \sigma_{\text{ess}}(P \nabla^2 \mathcal{L}_\xi(u_\xi) P)$  by hypotheses (iii) and (iv) of Theorem 3.1.

Now, note that  $\text{Ker} P \nabla^2 \mathcal{L}_\xi(u_\xi) P = (T_{u_\xi} \Sigma_{\mu_\xi})^\perp \subset E_0^\perp$ . Indeed, let  $u \in E_0^\perp$  be such that  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P u = 0$ . Then

$$D_{u_\xi}^2 \mathcal{L}(u_\xi)(Pu, Pu) = \langle u, P \nabla^2 \mathcal{L}_\xi(u_\xi) P u \rangle_E = 0.$$

Since  $D_{u_\xi}^2 \mathcal{L}(u_\xi)$  is strictly positive on  $\mathcal{Y}$  (see (5.2)), it follows that  $Pu = 0$ , which means that  $u \in (T_{u_\xi} \Sigma_{\mu_\xi})^\perp$ .

We now consider  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P$  as an operator on  $\mathcal{Y}$ . We have just shown that 0 is not an eigenvalue of  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P$ , nor in its essential spectrum. It is therefore not in its spectrum. In addition, we showed  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P$  is a strictly positive operator on  $\mathcal{Y}$  (see (5.2)). It therefore has a spectral gap:

$$\delta = \inf_{v \in \mathcal{Y} \setminus \{0\}} \frac{\langle P \nabla^2 \mathcal{L}_\xi(u_\xi) P v, v \rangle_E}{\langle v, v \rangle_E} > 0.$$

Finally, for all  $v \in \mathcal{Y}$ ,  $v \neq 0$ , one finds

$$D_{u_\xi}^2 \mathcal{L}_\xi(v, v) = \langle \nabla^2 \mathcal{L}_\xi(u_\xi) v, v \rangle_E = \langle P \nabla^2 \mathcal{L}_\xi(u_\xi) P v, v \rangle_E \geq \delta \|v\|_E^2$$

which is the desired estimate.

## 6. MAIN RESULT: A MORE GENERAL SETTING

In this section, we extend Theorem 3.1 to a more general setting that we now describe. In order to state our main result, we first need to associate to  $D_{u_\xi}^2 \mathcal{L}_\xi$  a (typically unbounded) self-adjoint operator on  $\hat{E}$ .

**Lemma 6.1.** *Suppose Hypothesis A holds. Let  $\xi \in \mathfrak{g}$  and  $\mathcal{L}_\xi$  as in (2.10) and let  $u \in E$ . If there exist  $\varepsilon, C > 0$  such that*

$$D_u^2 \mathcal{L}_\xi(v, v) \geq \varepsilon \|v\|_E^2 - C \|v\|^2 \quad (6.1)$$

for all  $v \in E$ , then there is a self-adjoint operator  $\nabla^2 \mathcal{L}_\xi(u) : \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u)) \subset \hat{E} \rightarrow \hat{E}$  defined by

$$\begin{aligned} \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u)) &= \{z \in E \mid \exists w \in \hat{E} \text{ such that } D_u^2 \mathcal{L}_\xi(z, v) = \langle w, v \rangle \text{ for all } v \in E\}, \\ \nabla^2 \mathcal{L}_\xi(u)z &= w \text{ for all } z \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u)). \end{aligned} \quad (6.2)$$

Moreover,  $\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$  is a form core for  $D_{u_\xi}^2 \mathcal{L}_\xi$ .

*Remark 6.2.* Note that

- (1)  $E$  is the form domain of the operator  $\nabla^2 \mathcal{L}_\xi(u)$ ,
- (2) Since  $\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u))$  is a form core for  $D_u^2 \mathcal{L}_\xi$  and condition (6.1) holds,  $\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u))$  is dense in  $E$ .

The existence and the uniqueness of the operator  $\nabla^2 \mathcal{L}_\xi(u)$  is a consequence of the First Representation theorem in Kato [20, Theorem 2.1 and 2.6 in Chapter VI]. Condition (6.1) ensures that the hypotheses of the First Representation theorem are satisfied (see [34, Lemma 3.3]). See also [32, 36] for more details.

We can now state our main result. We define  $p(D_\xi^2 W)$ ,  $n(D_\xi^2 W)$ ,  $W$ , and  $n(D_{u_\xi}^2 \mathcal{L}_\xi)$  as in Section 3.

**Theorem 6.3.** *Suppose Hypotheses A, B, C hold. Let  $\xi \in \Omega$  and suppose*

$$D_{u_\xi}^2 \mathcal{L}_\xi(v, v) \geq \varepsilon \|v\|_E^2 - C \|v\|^2, \forall v \in E. \quad (6.3)$$

Suppose in addition that

- (i)  $D_\xi^2 W$  is non-degenerate,
- (ii)  $\text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi = T_{u_\xi} \mathcal{O}_{u_\xi}$ ,
- (iii)  $\inf(\sigma(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap (0, +\infty)) > 0$ ,
- (iv)  $p(D_\xi^2 W) = n(D_{u_\xi}^2 \mathcal{L}_\xi)$ ,
- (v) for all  $u \in \mathcal{O}_{u_\xi}$  and for all  $j = 1, \dots, m$ , there exists  $\nabla F_j(u) \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \subset E$  such that

$$D_u F_j(w) = \langle \nabla F_j(u), w \rangle, \quad \forall w \in E. \quad (6.4)$$

Then there exists  $\delta > 0$  such that

$$\forall v \in T_{u_\xi} \Sigma_{\mu_\xi} \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp, \quad D_{u_\xi}^2 \mathcal{L}_\xi(v, v) \geq \delta \|v\|_E^2 \quad (6.5)$$

with  $(T_{u_\xi} \mathcal{O}_{u_\xi})^\perp = \{v \in \hat{E} \mid \langle v, w \rangle = 0 \ \forall w \in T_{u_\xi} \mathcal{O}_{u_\xi}\}$ .

Note that here, and in the rest of this section, the orthogonality is with respect to the inner product  $\langle \cdot, \cdot \rangle$ .

Let us point out that the hypotheses on the bilinear form  $D_{u_\xi}^2 \mathcal{L}_\xi$  in Theorem 6.3 can be re-expressed in terms of spectral hypotheses on the (unbounded) self-adjoint operator  $\nabla^2 \mathcal{L}_\xi(u_\xi)$ , as shown in the following lemma. This is important in applications, since it allows one to use the tools of spectral analysis for partial differential operators to establish those conditions.

**Lemma 6.4.** *Under the hypotheses of Lemma 6.1, with  $u = u_\xi$ ,  $\text{Ker } \nabla^2 \mathcal{L}_\xi(u_\xi) = \text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi$ .*

*If, in addition  $\dim \text{Ker } \nabla^2 \mathcal{L}_\xi(u_\xi) < +\infty$ , the negative spectral subspace of  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  is finite dimensional, and hypothesis (iii) of Theorem 6.3 is satisfied, then the dimension of the negative spectral subspace of  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  in  $\hat{E}$  is equal to the Morse index  $n(D_{u_\xi}^2 \mathcal{L}_\xi)$  of  $u_\xi \in E$  for  $\mathcal{L}_\xi$ .*

This lemma constitutes a slight generalization of Lemma 5.4 in [34] and its proof follows along the same lines. We give it for completeness.



*Proof.* By definition  $\text{Ker } \nabla^2 \mathcal{L}_\xi(u_\xi) = \{v \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \mid \nabla^2 \mathcal{L}_\xi(u_\xi)v = 0\}$  and a straightforward calculation leads to  $\text{Ker } \nabla^2 \mathcal{L}_\xi(u_\xi) = \text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ . Moreover, using the definition (6.2), it is easy to see that  $\text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi \subset \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ . As a consequence,  $\text{Ker } \nabla^2 \mathcal{L}_\xi(u_\xi) = \text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi$ .

Now, we know that

$$\dim \text{Ker } \nabla^2 \mathcal{L}_\xi(u_\xi) = n_0 < +\infty, \quad \inf(\sigma(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap (0, +\infty)) > 0$$

and we denote by  $0 \leq n_- < +\infty$  the dimension of the negative spectral subspace of  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  in  $\hat{E}$ . It follows that there exists  $\Gamma > 0$  such that  $\sigma(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap (0, \Gamma] = \emptyset$ . Let  $P_{(-\infty, 0]} : \hat{E} \rightarrow \hat{E}$  denote the orthogonal projection onto the finite dimensional span of all the eigenvectors of  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  corresponding to the eigenvalues in  $(-\infty, 0]$ , and let  $Q = I - P_{(-\infty, 0]}$ . We have that  $P_{(-\infty, 0]}(\hat{E}) \subset \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$  and that  $Qz \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$  if and only if  $z \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ . Also  $\dim P_{(-\infty, 0]}(\hat{E}) = n_0 + n_-$  and  $\langle \nabla^2 \mathcal{L}_\xi(u_\xi)Qz, Qz \rangle \geq \Gamma \|Qz\|^2$  for all  $z \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ . Thus, for any  $z \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$

$$\begin{aligned} D_{u_\xi}^2 \mathcal{L}_\xi(z, z) &= \langle \nabla^2 \mathcal{L}_\xi(u_\xi)z, z \rangle = \langle Q\nabla^2 \mathcal{L}_\xi(u_\xi)z, z \rangle + \langle P_{(-\infty, 0]}\nabla^2 \mathcal{L}_\xi(u_\xi)z, z \rangle \\ &= \langle \nabla^2 \mathcal{L}_\xi(u_\xi)Qz, Qz \rangle + \langle P_{(-\infty, 0]}\nabla^2 \mathcal{L}_\xi(u_\xi)z, z \rangle \\ &\geq \Gamma \|Qz\|^2 + \langle P_{(-\infty, 0]}\nabla^2 \mathcal{L}_\xi(u_\xi)z, z \rangle \\ &\geq \Gamma \|z\|^2 - \Gamma \|P_{(-\infty, 0]}z\|^2 + \langle \nabla^2 \mathcal{L}_\xi(u_\xi)P_{(-\infty, 0]}z, z \rangle \\ &\geq \Gamma \|z\|^2 + \langle (\nabla^2 \mathcal{L}_\xi(u_\xi) - I\Gamma)P_{(-\infty, 0]}z, z \rangle. \end{aligned}$$

As a consequence, using (6.3) and the fact that  $\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$  is dense in  $E$ , we obtain

$$\left(1 + \frac{\Gamma}{C}\right) D_{u_\xi}^2 \mathcal{L}_\xi(z, z) - \langle (\nabla^2 \mathcal{L}_\xi(u_\xi) - I\Gamma)P_{(-\infty, 0]}z, z \rangle \geq \frac{\varepsilon\Gamma}{C} \|z\|_E^2$$

for all  $z \in E$ , which implies

$$D_{u_\xi}^2 \mathcal{L}_\xi(z, z) - \frac{C}{\Gamma + C} \langle (\nabla^2 \mathcal{L}_\xi(u_\xi) - I\Gamma)P_{(-\infty, 0]}z, z \rangle \geq \frac{\varepsilon\Gamma}{\Gamma + C} \|z\|_E^2$$

for all  $z \in E$ . Moreover, for all  $z \in Q(E)$ ,

$$D_{u_\xi}^2 \mathcal{L}_\xi(z, z) \geq \frac{\varepsilon\Gamma}{\Gamma + C} \|z\|_E^2 \quad (6.6)$$

since  $P_{(-\infty, 0]}z = 0$ . But  $Q(E) \subset E$  since  $P_{(-\infty, 0]}(E) \subset P_{(-\infty, 0]}(\hat{E}) \subset \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \subset E$ . So we have shown that  $Q(E)$  is a positive subspace of  $E$ , for  $D_{u_\xi}^2 \mathcal{L}_\xi$ . Now consider the direct sum decomposition of  $E$  given by

$$E = Q(E) \oplus P_{(-\infty, 0)}(E) \oplus \text{Ker}(D^2 \mathcal{L}_\xi).$$

Here  $P_{(-\infty, 0)}$  is the projector onto the  $n_-$ -dimensional space spanned by the eigenvectors of  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  with strictly negative eigenvalue. Clearly,  $P_{(-\infty, 0)}(E)$  is a negative definite space for  $D_{u_\xi}^2 \mathcal{L}_\xi$ . We now show it is maximal. For that purpose, suppose  $z_* \in E$ ,  $z_* \notin P_{(-\infty, 0)}(E)$  and suppose  $\text{span}\{z_*, P_{(-\infty, 0)}(E)\}$  is a negative definite subspace of  $E$  for  $D_{u_\xi}^2 \mathcal{L}_\xi$  of dimension  $n_- + 1$ . We can suppose, without loss of generality, that  $z_* \in \text{Ker}(D^2 \mathcal{L}_\xi) \oplus Q(E)$ . Writing  $z_* = z_{*,0} + z_{*,+}$ , we see

$$D_{u_\xi}^2 \mathcal{L}_\xi(z_*, z_*) = D_{u_\xi}^2 \mathcal{L}_\xi(z_{*,+}, z_{*,+}) \geq 0,$$

where we used (6.6). This contradicts the fact that  $\text{span}\{z_*, P_{(-\infty, 0)}(E)\}$  is a negative definite space for  $D_{u_\xi}^2 \mathcal{L}_\xi$  and shows that  $P_{(-\infty, 0)}(E)$  is a maximally negative definite subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$ . Thus  $n(D_{u_\xi}^2 \mathcal{L}_\xi) = n_-$ .  $\square$

For the proof of Theorem 6.3, we will need the following two lemmas.

**Lemma 6.5.** *Let  $(\mathcal{H}, \|\cdot\|)$  be a Hilbert space and  $\mathcal{M} \subset \mathcal{K} \subset \mathcal{H}$  with  $\mathcal{M}$  a closed subspace of  $\mathcal{H}$  and  $\mathcal{K}$  a dense subspace of  $\mathcal{H}$  ( $\overline{\mathcal{K}}^{\|\cdot\|} = \mathcal{H}$ ). Then*

$$\overline{\mathcal{M}^\perp \cap \mathcal{K}}^{\|\cdot\|} = \mathcal{M}^\perp \quad (6.7)$$

*Proof.* Let  $u \in \mathcal{M}^\perp$ . There exists a sequence  $k_n \in \mathcal{K}$  so that  $k_n \rightarrow u$ . Since  $\mathcal{M}$  is closed, we can write  $k_n = w_n + v_n$ , with  $w_n \in \mathcal{M}, v_n \in \mathcal{M}^\perp$ . Moreover, since  $k_n \in \mathcal{K}$  and  $w_n \in \mathcal{M} \subset \mathcal{K}$ ,  $v_n \in \mathcal{M}^\perp \cap \mathcal{K}$ . Clearly both sequences  $w_n$  and  $v_n$  converge, respectively to  $w \in \mathcal{M}, v \in \mathcal{M}^\perp$ . Since  $u = w + v \in \mathcal{M}^\perp$ , we find  $w = 0$  and  $v = u$ . Hence  $v_n \in \mathcal{K} \cap \mathcal{M}^\perp$  converges to  $u$ .  $\square$

We introduce, as before  $E_0 = \text{Ker } \nabla_{u_\xi}^2 \mathcal{L}_\xi$ . We know from Lemma 6.4 that  $\text{Ker } \nabla_{u_\xi}^2 \mathcal{L}_\xi = \text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi$ . Hypothesis (ii) of the theorem then implies

$$E_0 = T_{u_\xi} \mathcal{O}_{u_\xi} \subset \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)). \quad (6.8)$$

We define furthermore

$$V_{u_\xi} = \text{span}\{\nabla F_j(u_\xi), j = 1, \dots, m\}. \quad (6.9)$$

Note that, by hypothesis (v) of Theorem 6.3,  $V_{u_\xi} \subset \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \subset E$ .

**Lemma 6.6.** *Under the hypotheses of Theorem 6.3, we have*

- (i)  $E_0^\perp \cap \overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|} = \overline{E_0^\perp \cap T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}$ .
- (ii)  $\overline{T_{u_\xi} \Sigma_{\mu_\xi} \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))}^{\|\cdot\|} = \overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}$ .
- (iii) Define

$$\mathcal{Y} = T_{u_\xi} \Sigma_{\mu_\xi} \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap E_0^\perp. \quad (6.10)$$

Then

$$\hat{\mathcal{Y}} := \overline{\mathcal{Y}}^{\|\cdot\|} = E_0^\perp \cap V_{u_\xi}^\perp, \quad (6.11)$$

where  $V_{u_\xi}$  is defined in (6.9). Hence

$$E_0^\perp = T_{u_\xi} \mathcal{O}_{u_\xi}^\perp = \hat{\mathcal{Y}} \oplus_\perp V_{u_\xi}^\perp, \quad \hat{E} = E_0 \oplus_\perp \hat{\mathcal{Y}} \oplus_\perp V_{u_\xi}^\perp. \quad (6.12)$$

(iv) Let  $P$  be the orthogonal projector onto  $\hat{\mathcal{Y}}$ . Let  $u \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap E_0^\perp$ . Then  $Pu \in \mathcal{Y}$ .

We use the notation  $\oplus_\perp$  to indicate a direct sum that is orthogonal for the inner product  $\langle \cdot, \cdot \rangle$ .

*Proof.* (i) Note that  $E_0 \subset T_{u_\xi} \Sigma_{\mu_\xi} \subset \overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}$ . We now apply Lemma 6.5 with  $\mathcal{M} = E_0, \mathcal{K} = T_{u_\xi} \Sigma_{\mu_\xi}, \mathcal{H} = \overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}$  to obtain

$$E_0^{\perp\sigma} = \overline{E_0^{\perp\sigma} \cap T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}.$$

Here we wrote  $E_0^{\perp\sigma}$  for the orthogonal complement to  $E_0$  in  $\overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}$ , i.e.

$$E_0^{\perp\sigma} = E_0^\perp \cap \overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}. \quad (6.13)$$

The last two equations imply the result.

(ii) Since  $V_{u_\xi}$  is a closed finite dimensional subspace of  $\hat{E}$ , we have  $\hat{E} = V_{u_\xi}^\perp \oplus_\perp V_{u_\xi}$  with

$$V_{u_\xi}^\perp = \{w \in \hat{E} \mid \langle v, w \rangle = 0 \ \forall v \in V_{u_\xi}\}.$$

Since by hypothesis (v) of Theorem 6.3, for all  $w \in E$ ,  $D_{u_\xi} F_j(w) = \langle \nabla F_j(u_\xi), w \rangle$  for  $j = 1, \dots, m$ , we see that

$$T_{u_\xi} \Sigma_{\mu_\xi} = V_{u_\xi}^\perp \cap E \quad (6.14)$$

and hence

$$E = T_{u_\xi} \Sigma_{\mu_\xi} \oplus_\perp V_{u_\xi}. \quad (6.15)$$

Using Lemma 6.5, and the fact that  $V_{u_\xi} \subset E \subset \hat{E}$ , with  $E$  dense in  $\hat{E}$ , (6.14) implies  $\overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|} = \overline{V_{u_\xi}^\perp \cap E}^{\|\cdot\|} = V_{u_\xi}^\perp$  meaning that

$$\hat{E} = \overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|} \oplus_\perp V_{u_\xi}. \quad (6.16)$$

From (6.15) one concludes

$$\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) = (\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap T_{u_\xi} \Sigma_{\mu_\xi}) \oplus_\perp V_{u_\xi},$$

and hence

$$\hat{E} = \overline{\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|} \oplus_\perp V_{u_\xi}$$

Comparing this to (6.16), one concludes

$$\overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|} = \overline{\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}. \quad (6.17)$$

In other words,  $\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap T_{u_\xi} \Sigma_{\mu_\xi}$  is dense in  $\overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}$ . This proves (ii).

(iii) Note that  $E_0 \subset \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap T_{u_\xi} \Sigma_{\mu_\xi} \subset \overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}$ . Then we can, in view of part (ii) of the Lemma, apply Lemma 6.5 with  $\mathcal{M} = E_0$ ,  $\mathcal{K} = \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap T_{u_\xi} \Sigma_{\mu_\xi}$ , and  $\mathcal{H} = \overline{T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|}$  to obtain:

$$E_0^{\perp\sigma} = \overline{E_0^{\perp\sigma} \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|} = \overline{E_0^{\perp\sigma} \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap T_{u_\xi} \Sigma_{\mu_\xi}}^{\|\cdot\|} = \hat{\mathcal{Y}},$$

where  $E_0^{\perp\sigma}$  is defined in (6.13). Since

$$E_0^\perp = E_0^{\perp\sigma} \oplus_\perp V_{u_\xi},$$

part (iii) follows.

(iv) Let  $u \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap E_0^\perp$ . Since  $E_0^\perp = \hat{\mathcal{Y}} \oplus_\perp V_{u_\xi}$ ,  $u = Pu + v$  with  $Pu \in \hat{\mathcal{Y}}$  and  $v \in V_{u_\xi}$ . Moreover, since  $u \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$  and  $V_{u_\xi} \subset \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ , it follows  $Pu \in \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap \hat{\mathcal{Y}}$ . To conclude, we observe that

$$\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap \hat{\mathcal{Y}} = \mathcal{Y}.$$

Indeed, using (6.11) and (6.14),

$$\begin{aligned} \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap \hat{\mathcal{Y}} &= \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap E_0^\perp \cap V_{u_\xi}^\perp = \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap E_0^\perp \cap V_{u_\xi}^\perp \cap E \\ &= \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap E_0^\perp \cap T_{u_\xi} \Sigma_{\mu_\xi} = \mathcal{Y}. \end{aligned}$$

Finally,  $Pu \in \mathcal{Y}$ . □

We can now proceed with the proof of Theorem 6.3.

*Proof of Theorem 6.3.* Let  $n_- = n(D_{u_\xi}^2 \mathcal{L}_\xi)$ . As in the proof of Theorem 3.1,  $\mathcal{X}_- = \text{span} \{ \eta_1 \cdot \nabla_\xi u_\xi, \dots, \eta_{n_-} \cdot \nabla_\xi u_\xi \}$  is a maximally negative definite subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$ ,  $\mathcal{X}_- \cap T_{u_\xi} \Sigma_{\mu_\xi} = \{0\}$ , and  $T_{u_\xi} \Sigma_{\mu_\xi}$  is a positive subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$ . Furthermore, note that by hypothesis  $\text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi = T_{u_\xi} \mathcal{O}_{u_\xi} \subset T_{u_\xi} \Sigma_{\mu_\xi}$ . Hence, by Lemma 4.3 (ii),  $T_{u_\xi} \Sigma_{\mu_\xi} \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp$  is a positive definite subspace for  $D_{u_\xi}^2 \mathcal{L}_\xi$ , meaning

$$D_{u_\xi}^2 \mathcal{L}_\xi(v, v) > 0, \quad \forall v \in T_{u_\xi} \Sigma_{\mu_\xi} \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp, v \neq 0. \quad (6.18)$$

We note for further reference that, by Lemma 6.4,  $\text{Ker } D_{u_\xi}^2 \mathcal{L}_\xi = \text{Ker } \nabla^2 \mathcal{L}_\xi(u_\xi)$ , so that  $T_{u_\xi} \mathcal{O}_{u_\xi} \subset \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ .

Recall that, since  $E_0$  is the kernel of  $\nabla^2 \mathcal{L}_\xi(u_\xi)$ ,  $E_0^\perp$  is the spectral space associated to  $\mathbb{R}^*$ , and hence invariant under  $\nabla^2 \mathcal{L}_\xi(u_\xi)$ . Now, let  $P$  be the projection on  $\hat{\mathcal{Y}}$  and consider the following decomposition of the operator  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  on  $E_0^\perp = \hat{\mathcal{Y}} \oplus_\perp V_{u_\xi}$ ,

$$\begin{aligned} \nabla^2 \mathcal{L}_\xi(u_\xi)|_{E_0^\perp} &= P \nabla^2 \mathcal{L}_\xi(u_\xi) P + P \nabla^2 \mathcal{L}_\xi(u_\xi) (I_{E_0^\perp} - P) \\ &\quad + (I_{E_0^\perp} - P) \nabla^2 \mathcal{L}_\xi(u_\xi) P + (I_{E_0^\perp} - P) \nabla^2 \mathcal{L}_\xi(u_\xi) (I_{E_0^\perp} - P). \end{aligned}$$

We claim that  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P$  is a self-adjoint operator on  $E_0^\perp$  with domain  $E_0^\perp \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ . Indeed, since  $V_{u_\xi}$  is finite dimensional, we can easily show that  $P \nabla^2 \mathcal{L}_\xi(u_\xi) (I_{E_0^\perp} - P) + (I_{E_0^\perp} - P) \nabla^2 \mathcal{L}_\xi(u_\xi) P$  and  $(I_{E_0^\perp} - P) \nabla^2 \mathcal{L}_\xi(u_\xi) (I_{E_0^\perp} - P)$  are bounded self-adjoint operators on  $E_0^\perp$ . Hence  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P$  is the sum of a self-adjoint operator with domain  $E_0^\perp \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$  and a bounded operator on  $E_0^\perp$  and, by the Kato-Rellich theorem, it is self-adjoint on  $E_0^\perp$  with domain  $E_0^\perp \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ . In particular,  $\nabla^2 \mathcal{L}_\xi(u_\xi) = P \nabla^2 \mathcal{L}_\xi(u_\xi) P + K$  with  $K$  a finite rank operator and  $\sigma_{\text{ess}}(P \nabla^2 \mathcal{L}_\xi(u_\xi) P) = \sigma_{\text{ess}}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ . As a consequence,  $0 \notin \sigma_{\text{ess}}(P \nabla^2 \mathcal{L}_\xi(u_\xi) P)$  by hypotheses (iii) and (iv) of Theorem 6.3.

Now note that  $\ker P \nabla^2 \mathcal{L}_\xi(u_\xi) P = V_{u_\xi} \subset E_0^\perp \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$ . Indeed, let  $u \in E_0^\perp \cap \mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$  be such that  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P u = 0$ . Then

$$D_{u_\xi}^2 \mathcal{L}_\xi(Pu, Pu) = \langle u, P \nabla^2 \mathcal{L}_\xi(u_\xi) P u \rangle = 0$$

with  $Pu \in \mathcal{Y}$  by Lemma 6.6. Since  $D_{u_\xi}^2 \mathcal{L}_\xi$  is strictly positive on  $\mathcal{Y}$  (see (6.18)), it follows that  $Pu = 0$ , which means that  $u \in V_{u_\xi}$ .

We now consider  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P$  on  $\mathcal{Y}$ . We have just shown that

$$0 \notin \sigma(P \nabla^2 \mathcal{L}_\xi(u_\xi) P) = \sigma_{\text{ess}}(P \nabla^2 \mathcal{L}_\xi(u_\xi) P) \cup \sigma_{\text{d}}(P \nabla^2 \mathcal{L}_\xi(u_\xi) P)$$

and that  $P \nabla^2 \mathcal{L}_\xi(u_\xi) P$  is strictly positive on  $\mathcal{Y}$  (see (6.18)). It therefore has a spectral gap:

$$\tilde{\delta} = \inf_{v \in \mathcal{Y} \setminus \{0\}} \frac{\langle P \nabla^2 \mathcal{L}_\xi(u_\xi) P v, v \rangle}{\langle v, v \rangle} > 0.$$

Next, using the inequality (6.1), we obtain for all  $v \in \mathcal{Y}$

$$\left(1 + \frac{\tilde{\delta}}{C}\right) D_{u_\xi}^2 \mathcal{L}_\xi(v, v) \geq \frac{\epsilon \tilde{\delta}}{C} \|v\|_E^2$$

which implies

$$D_{u_\xi}^2 \mathcal{L}_\xi(v, v) \geq \frac{\epsilon \tilde{\delta}}{\tilde{\delta} + C} \|v\|_E^2.$$

Finally, the density of  $\mathcal{D}(\nabla^2 \mathcal{L}_\xi(u_\xi))$  in  $E$  for  $\|\cdot\|_E$  yields (6.5).  $\square$

## 7. PERSISTENCE OF RELATIVE EQUILIBRIA

In this section we come back to the question of persistence of relative equilibria, which is the question of the existence of a family of relative equilibria as in (3.1)-(3.2). Three situations occur. In some cases, such a family can be explicitly exhibited. In others, its existence can be proven by ad hoc methods adapted to the specific situation at hand. We will give examples of both these cases in the following section. Finally, under suitable conditions, general structural theorems asserting its existence can be proven. We give below a theorem guaranteeing the existence of a family of relative equilibria as in (3.1)-(3.2) in the infinite dimensional framework under study here, under a natural condition on the point  $\mu_* = F(u_*)$  in  $\mathfrak{g}^*$ , which is for example always satisfied when the symmetry group  $G$  is commutative and which is satisfied on an open dense subset of  $\mathfrak{g}^*$  in all cases. We will comment on the relation with the situation for finite dimensional systems at the end of this section. Applications will be given in the following section.

We will make the following hypothesis throughout this section:

**Hypothesis D.**  $E$  is a Hilbert space with inner product  $\langle \cdot, \cdot \rangle_E$  and  $\| \cdot \|_E = \sqrt{\langle \cdot, \cdot \rangle_E}$ .

With this hypothesis, one can view  $\langle \cdot, \cdot \rangle_E$  as a closed form on  $\hat{E}$  (defined in (2.12)), with form domain  $E$ . It follows (Theorem VIII.15 in [32]) that there exists a unique unbounded positive operator  $T^2$  on  $\hat{E}$ , with domain  $\mathcal{D}(T^2)$ , so that, for all  $u, v \in \mathcal{D}(T^2)$ ,

$$\langle u, v \rangle_E = \langle u, T^2 v \rangle,$$

and so that, in addition  $E = \mathcal{D}(T)$  and, for all  $u, v \in \mathcal{D}(T)$

$$\langle u, v \rangle_E = \langle Tu, Tv \rangle.$$

Here  $T$  is the positive square root of  $T^2$ . Note that

$$\langle u, T^2 u \rangle \geq \langle u, u \rangle,$$

so that 0 is in the resolvent set of  $T^2$  and hence  $T^{-2}$  is a bounded operator on  $\hat{E}$ .

Next, we introduce in the usual manner the scale of spaces  $E_\lambda = \overline{\mathcal{D}(T^\lambda)}^{\|\cdot\|_\lambda}$ , where  $\|u\|_\lambda = \|T^\lambda u\|$  and  $\lambda \in \mathbb{R}$ . In particular, we have  $E = E_1$  and  $\hat{E} = E_0$ .

Our persistence result then reads as follows:

**Theorem 7.1.** *Let Hypotheses A, B and D hold and suppose there exists  $\xi_* \in \mathfrak{g}$  and  $u_* \in \mathcal{D} \cap E_2$  so that  $D_{u_*} \mathcal{L}_{\xi_*} = 0$ . Suppose in addition:*

(a) *There exist  $\varepsilon, C > 0$  such that, for all  $v \in E$ ,*

$$D_{u_*}^2 \mathcal{L}_{\xi_*}(v, v) \geq \varepsilon \|v\|_E^2 - C \|v\|^2. \quad (7.1)$$

(b) *For all  $u \in E_2$  and for all  $\xi \in \mathfrak{g}$  there exists  $\nabla \mathcal{L}_\xi(u) \in \hat{E}$  such that*

$$D_u \mathcal{L}_\xi(v) = \langle \nabla \mathcal{L}_\xi(u), v \rangle, \quad \forall v \in E. \quad (7.2)$$

(c) *The function  $(\xi, v) \in \mathfrak{g} \times E_2 \rightarrow \nabla \mathcal{L}_\xi(v) \in \hat{E}$  belongs to  $C^1(\mathfrak{g} \times E_2; \hat{E})$ .*

(d) *For all  $v \in E_2$ ,  $g \in G \rightarrow \Phi_g(v) \in E$  is  $C^1$ .*

(e) *The function  $F$  is regular at  $u_*$ .*

(f) *For all  $\mu$  in a neighbourhood of  $\mu_* = F(u_*)$ ,  $\dim \mathfrak{g}_\mu = \dim \mathfrak{g}_{\mu_*}$ .*

If in addition,

- (i)  $\mathcal{D}(\nabla^2 \mathcal{L}_{\xi_*}(u_*)) = \mathcal{D}(T^2)$
- (ii)  $\text{Ker } D_{u_*}^2 \mathcal{L}_{\xi_*} = T_{u_*} \mathcal{O}_{u_*}$ ,
- (iii)  $\inf(\sigma(\nabla^2 \mathcal{L}_{\xi_*}(u_*)) \cap (0, +\infty)) > 0$ ,
- (iv)  $n(D_{u_*}^2 \mathcal{L}_{\xi_*}) < +\infty$ ,

Then there exists a neighbourhood  $\Omega$  of  $\xi_*$  in  $\mathfrak{g}$  and a  $C^1$  map  $\xi \in \Omega \rightarrow u_\xi \in E$  with  $u_{\xi_*} = u_*$  so that for all  $\xi$ , (3.2) holds. The map  $\xi \rightarrow u_\xi$  is an injective immersion.

The conditions that are central here are (ii)-(iii)-(iv): they are to be compared to the identically numbered conditions of Theorem 6.3. The other conditions, notably (a)-(e), are technical and usually readily verified in applications. They are virtually automatic in finite dimensional problems. Condition (f) is of purely group-theoretic nature. It is known to hold on an open dense set of any Lie algebra. In fact, on such a set, the orbits all have the same maximal dimension and the Lie algebra  $\mathfrak{g}$  of the isotropy group is commutative [11].

*Proof.* Let  $\hat{U}$  be the subspace of  $\hat{E}$  defined by

$$\hat{U} = \{v \in \hat{E} \mid \langle v, u \rangle = 0, \forall u \in T_{u_*} \mathcal{O}_{u_*}\} = T_{u_*} \mathcal{O}_{u_*}^\perp \quad (7.3)$$

and consider the map

$$\begin{aligned} \mathcal{F} : \mathfrak{g} \times (E_2 \cap \hat{U}) &\rightarrow \hat{U} \\ (\xi, w) &\mapsto Q \nabla \mathcal{L}_\xi(u_* + w) \end{aligned}$$

where  $Q$  is the orthogonal projector onto  $\hat{U} \subset \hat{E}$ . Note that  $(E_2 \cap \hat{U}, \|\cdot\|_2)$  and  $(\hat{U}, \|\cdot\|)$  are Banach spaces.

By hypothesis (c),  $\mathcal{F}$  is  $\mathcal{C}^1$ . Moreover, it is clear that  $\mathcal{F}(\xi_*, 0) = 0$ . Hence, to apply the implicit function theorem, we remark that the derivative of  $\mathcal{F}$  along  $E_2 \cap \hat{U}$  at the point  $(\xi_*, 0)$

$$\partial_w \mathcal{F}(\xi_*, 0) = Q \nabla^2 \mathcal{L}_{\xi_*}(u_*)$$

is an isomorphism from  $(E_2 \cap \hat{U}, \|\cdot\|_2)$  to  $(\hat{U}, \|\cdot\|)$ . Indeed, as a result of the hypotheses (ii), (iii) and (iv),  $Q \nabla^2 \mathcal{L}_{\xi_*}(u_*) Q$  is a self-adjoint operator with bounded inverse.

Therefore, there exist  $\mathcal{V}_{\xi_*}$  a neighbourhood of  $\xi_*$  in  $\mathfrak{g}$ ,  $\mathcal{V}_{u_*}$  a neighbourhood of  $u_*$  in  $E_2 \cap \hat{U}$ , and a function  $\Lambda : \mathcal{V}_{\xi_*} \rightarrow \mathcal{V}_{u_*}$  such that  $\mathcal{F}(\xi, \Lambda(\xi)) = 0$ . In particular, setting  $u_\xi = u_* + \Lambda(\xi) \in E_2$ , we have

$$Q \nabla \mathcal{L}_\xi(u_\xi) = 0.$$

This implies that for all  $v \in U = \{v \in E \mid \langle v, u \rangle = 0, \forall u \in T_{u_*} \mathcal{O}_{u_*}\} \subset \hat{U}$ ,

$$0 = \langle v, Q \nabla \mathcal{L}_\xi(u_\xi) \rangle = D_{u_\xi} \mathcal{L}_\xi(v).$$

On the other hand, we know that for all  $g \in G_{\mu_\xi}$

$$\mathcal{L}_\xi(\Phi_g(u_\xi)) = \mathcal{L}_\xi(u_\xi).$$

Hence, by writing  $g = \exp(t\eta)$  with  $\eta \in \mathfrak{g}_{\mu_\xi}$  and taking the derivative with respect to  $t$  (which is possible because of hypothesis (d)), it follows that

$$D_{u_\xi} \mathcal{L}_\xi(X_\eta(u_\xi)) = 0$$

which means that  $D_{u_\xi} \mathcal{L}_\xi(v) = 0$  for all  $v \in T_{u_\xi} \mathcal{O}_{u_\xi}$ .

To conclude it is sufficient to prove that  $E = U \oplus T_{u_\xi} \mathcal{O}_{u_\xi}$  for all  $\xi$  in a neighbourhood of  $\xi_*$ .

First of all, using the fact that  $\hat{E} = T_{u_*} \mathcal{O}_{u_*} \oplus_\perp T_{u_*} \mathcal{O}_{u_*}^\perp = T_{u_*} \mathcal{O}_{u_*} \oplus_\perp \hat{U}$ , we prove  $\hat{E} = \hat{U} \oplus T_{u_\xi} \mathcal{O}_{u_\xi}$ . Since  $\mu_* = F(u_*)$  is a regular point in  $\mathfrak{g}^*$ , one can choose, for every  $\mu$  in a neighbourhood  $\tilde{\mathcal{M}}$  of  $\mu_*$ , a basis  $\ell_i(\mu) \in \mathfrak{g}_\mu$ , for  $i = 1 \dots m'$ , smoothly in  $\mu$ . One can then construct

$$\Xi : (\eta, \mu) \in \mathbb{R}^{m'} \times \tilde{\mathcal{M}} \rightarrow \Xi(\eta, \mu) = \sum_{i=1}^{m'} \eta_i \ell_i(\mu) \in \mathfrak{g}_\mu \subset \mathfrak{g}.$$

Let  $e_1, \dots, e_{m'}$  the canonical basis of  $\mathbb{R}^{m'}$ . For all  $\xi \in \mathfrak{g}$  such that  $u_\xi$  is sufficiently close to  $u_*$ , we can define, using (2.3),

$$X_i(u_\xi) = \frac{d}{dt} \Phi_{\exp(\Xi(te_i, F))}(u_\xi)|_{t=0} = X_{\ell_i(F(u_\xi))}(u_\xi) \quad (7.4)$$

Note that  $X_i(u_\xi)$  are linearly independent and  $\text{span}\{X_i(u_\xi)\} = T_{u_\xi} \mathcal{O}_{u_\xi}$ . Then,  $\dim T_{u_\xi} \mathcal{O}_{u_\xi} = m' = \dim T_{u_*} \mathcal{O}_{u_*}$ .

Next, writing  $P = \mathbb{I} - Q$  the orthogonal projector onto  $T_{u_*} \mathcal{O}_{u_*}$ , we prove  $P : T_{u_\xi} \mathcal{O}_{u_\xi} \rightarrow T_{u_*} \mathcal{O}_{u_*}$  is a bijection. Noting that the matrix of  $P : T_{u_\xi} \mathcal{O}_{u_\xi} \rightarrow T_{u_*} \mathcal{O}_{u_*}$ , given by

$$\langle X_i(u_*), P X_i(u_\xi) \rangle$$

is invertible for  $\xi = \xi_*$ , this remains true by continuity for  $\xi$  close to  $\xi_*$ . As a conclusion,  $\hat{E} = \hat{U} \oplus T_{u_\xi} \mathcal{O}_{u_\xi}$ .

Finally, it is clear that  $E = U \oplus T_{u_\xi} \mathcal{O}_{u_\xi}$ . Since  $\hat{E} = \hat{U} \oplus T_{u_\xi} \mathcal{O}_{u_\xi}$ , we have, for each  $u \in E$ ,

$$u = \ell + k,$$

with  $k \in T_{u_\xi} \mathcal{O}_{u_\xi} \subset E, \ell \in \hat{U}$ . Hence  $\ell \in E \cap \hat{U} = U$ , which concludes the argument. Note that, here as before, the orthogonality is with respect to the inner product  $\langle \cdot, \cdot \rangle$ . □

For systems with a finite dimensional phase space, the persistence problem is addressed in [28, 23, 21]. The theorems provided there use various conditions on the group  $G$  and on its action  $\Phi$ , on the nature of the isotropy groups  $G_\xi$  and  $G_\mu$ , and finally on the Hessian of  $H$  and/or  $\mathcal{L}_\xi$  restricted to a suitable space. Note however that in applications to PDE the symplectic structure is always weak rather than strong. As a result, the finite dimensional arguments do not readily transpose to the infinite dimensional setting. Indeed, various topological complications manifest themselves essentially as domain questions for unbounded operators and lack of differentiability of the dynamical vector field  $X_H = \mathcal{J}^{-1}DH$  and of the vector fields generating the symmetries ( $X_\xi$  in (2.3)), as we have seen. On the other hand, the infinite dimensional setting offers a simplification over the usual finite dimensional one because the phase space  $E$ , rather than being a manifold, is a vector space, and because the action of the invariance group is usually linear, facts that we have very much exploited in the above proof and elsewhere in this paper.

## 8. EXAMPLES

**8.1. Stability of solitons for the nonlinear Schrödinger equation.** We consider the focusing nonlinear Schrödinger equation with a power nonlinearity given by

$$\begin{cases} i\partial_t u(t, x) + \Delta u(t, x) + |u(t, x)|^{p-1}u(t, x) = 0 & \text{in } \mathbb{R}^d \\ u(0, x) = u(x) \end{cases} \quad (8.1)$$

with  $u(t, x) \in \mathbb{C}$ ,  $1 < p < 1 + \frac{4}{d}$  and  $d = 1, 2, 3$ . This choice of parameters guarantees the global existence of solution to (8.1) in  $H^1(\mathbb{R}^d)$  (see [7]).

Equation (8.1) is the Hamiltonian differential equation associated to the Hamiltonian

$$H(u) = \frac{1}{2} \int_{\mathbb{R}^d} |\nabla u(x)|^2 dx - \frac{1}{p+1} \int_{\mathbb{R}^d} |u(x)|^{p+1} dx. \quad (8.2)$$

Next, let  $G = \mathbb{R} \times \mathbb{R}^d$  and define its action on  $E = H^1(\mathbb{R}^d, \mathbb{C})$  via

$$\forall u \in H^1(\mathbb{R}^d), \quad (\Phi_{\gamma_1, \gamma_2}(u))(x) = e^{-i\gamma_1} u(x - \gamma_2). \quad (8.3)$$

Clearly,  $H \circ \Phi_g = H$  and the group  $G$  is an invariance group for the dynamics and the quantities

$$F_1(u) = \frac{1}{2} \int_{\mathbb{R}^d} |u(x)|^2 dx \quad (8.4)$$

$$F_{1+j}(u) = \frac{1}{2} \int_{\mathbb{R}^d} u^* \left( \frac{1}{i} \partial_{x_j} \right) u dx \quad (8.5)$$

for  $j = 1, \dots, d$ , are the corresponding constants of the motion (see [7]).

The family of solitary waves

$$u_{\omega, c}(x) = e^{i\frac{c}{2} \cdot x} u_\omega(x) \quad (8.6)$$

with  $c \in \mathbb{R}^d$  and  $u_\omega$  the unique positive solution (see [35] for more details) to

$$\Delta u_\omega + |u_\omega|^{p-1} u_\omega = -\omega u_\omega \quad (8.7)$$

with  $\omega \in \mathbb{R}$ ,  $\omega < 0$ , are  $G$ -relative equilibria of (8.1). Indeed, if we define  $\mathcal{L}_\xi$  by

$$\mathcal{L}_\xi(u) = H(u) - \xi_1 F_1(u) - \sum_{j=1}^d \xi_{j+1} F_j(u), \quad (8.8)$$

we can easily verify that  $u_{c, \omega}$  is a solution to the stationary equation  $D_u \mathcal{L}_\xi = 0$  with  $\xi = (\omega - \frac{|c|^2}{4}, c)$ .

In other words, for each  $\xi \in \Omega = \left\{ (\xi_1, \hat{\xi}) \in \mathbb{R} \times \mathbb{R}^d, \xi_1 + \frac{|\hat{\xi}|^2}{4} < 0 \right\}$ ,

$$u_\xi(x) = e^{i\frac{\hat{\xi}}{2} \cdot x} u_{\xi_1 + \frac{|\hat{\xi}|^2}{4}}(x)$$

is a  $G$ -relative equilibrium of (8.1) with  $\mu_\xi = F(u_\xi)$ . Note that, since  $G$  is commutatif,  $G_{\mu_\xi} = G$ . Here, we use the notation  $\hat{\xi} = (\xi_2, \dots, \xi_{d+1})$ . Note that, if  $d = 1$  and  $p = 3$ , the unique positive solution of (8.7) is explicit:

$$u_\omega(x) = \sqrt{-2\omega} \operatorname{sech}(\sqrt{-\omega}x). \quad (8.9)$$

The  $G$ -orbit of the initial condition  $u_\xi(x)$  is given by

$$\mathcal{O}_{u_\xi} = \{e^{-i\gamma_1} u_\xi(x - \hat{\gamma}), (\gamma_1, \hat{\gamma}) \in G_{\mu_\xi}\}. \quad (8.10)$$

Our goal is to investigate the orbital stability of these relative equilibria and in particular to obtain the coercivity of  $\mathcal{L}_\xi$  by means of Theorem 6.3. This, together with Proposition 2.3 and the results of [10], leads to the orbital stability. Hypotheses A, B and C are easily seen to be satisfied, with  $\mathcal{D} = H^3(\mathbb{R}^d)$ . Note in particular that, since  $p > 1$ ,  $H \in C^2(E)$ . Also, we use for  $\langle \cdot, \cdot \rangle$  in Hypothesis B the usual  $L^2$ -scalar product, so that  $\hat{E} = L^2(\mathbb{R}^d, \mathbb{C})$  (viewed as a real Hilbert space). To check the further hypotheses of Theorem 6.3, we start by computing  $D_{u_\xi}^2 \mathcal{L}_\xi(v, w)$ . A straightforward calculation gives

$$\begin{aligned} D_{u_\xi}^2 \mathcal{L}_\xi(v, w) = \operatorname{Re} & \left[ \int_{\mathbb{R}^d} (-\Delta w) v^* - \int_{\mathbb{R}^d} |u_\xi|^{p-1} w v^* - \frac{p-1}{2} \int_{\mathbb{R}^d} |u_\xi|^{p-3} (w^* u_\xi + w u_\xi^*) (u_\xi v^*) \right. \\ & \left. - \xi_1 \int_{\mathbb{R}^d} w v^* - \sum_{j=1}^d \xi_{j+1} \int_{\mathbb{R}^d} \left( \frac{1}{i} \partial_{x_j} w \right) v^* \right]. \end{aligned}$$

Hence, writing  $v(x) = e^{i\frac{\xi}{2} \cdot x} \tilde{v}(x)$  and  $w(x) = e^{i\frac{\xi}{2} \cdot x} \tilde{w}(x)$ , we obtain

$$\begin{aligned} D_{u_\xi}^2 \mathcal{L}_\xi(v, w) = \operatorname{Re} & \left[ \int_{\mathbb{R}^d} (-\Delta \tilde{w}) \tilde{v}^* - \int_{\mathbb{R}^d} u_\omega^{p-1} \tilde{w} \tilde{v}^* - \frac{p-1}{2} \int_{\mathbb{R}^d} u_\omega^{p-1} (\tilde{w}^* + \tilde{w}) \tilde{v}^* - \omega \int_{\mathbb{R}^d} \tilde{w} \tilde{v}^* \right] \\ = \langle L \tilde{w}, \tilde{v} \rangle & = \langle \nabla^2 \mathcal{L}_\xi(u_\xi) w, v \rangle \end{aligned}$$

with

$$L \tilde{w} = \begin{pmatrix} -\Delta - p u_\omega^{p-1} - \omega & 0 \\ 0 & -\Delta - u_\omega^{p-1} - \omega \end{pmatrix} \begin{pmatrix} \operatorname{Re}(\tilde{w}) \\ \operatorname{Im}(\tilde{w}) \end{pmatrix}. \quad (8.11)$$

It then follows that the operator  $\nabla^2 \mathcal{L}_\xi(u_\xi)$  introduced in Lemma 6.1 is given by

$$\nabla^2 \mathcal{L}_\xi(u_\xi) = U^* L U \quad \text{with} \quad U = \begin{pmatrix} \cos\left(\frac{\xi}{2} \cdot x\right) & \sin\left(\frac{\xi}{2} \cdot x\right) \\ -\sin\left(\frac{\xi}{2} \cdot x\right) & \cos\left(\frac{\xi}{2} \cdot x\right) \end{pmatrix}.$$

Clearly, the estimate (6.3) is satisfied. Let  $L_+$  and  $L_-$  be defined by

$$L_+ = -\Delta - p u_\omega^{p-1} - \omega, \quad L_- = -\Delta - u_\omega^{p-1} - \omega.$$

Since  $u_\omega$  is the unique positive solution to (8.7), using a decomposition in spherical harmonics and proceeding as in [39, Lemma 4.1], one proves that  $\operatorname{Ker}(L_+) = \operatorname{span}\{\partial_{x_1} u_\omega, \dots, \partial_{x_n} u_\omega\}$  and  $\operatorname{Ker}(L_-) = \operatorname{span}\{u_\omega\}$ . Moreover, since  $u_\omega$  is strictly positive, one concludes that 0 is the first eigenvalue of  $L_-$ . Similarly, one proves  $L_+$  has exactly one negative eigenvalue.

As a consequence,  $\operatorname{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) = T_{u_\xi} \mathcal{O}_{u_\xi}$  and  $n(D_{u_\xi}^2 \mathcal{L}_\xi) = 1$ .

Next, we have to show that  $1 = n(D_{u_\xi}^2 \mathcal{L}_\xi) = p(D_\xi^2 W)$ . Since  $p(D_\xi^2 W) \leq n(D_{u_\xi}^2 \mathcal{L}_\xi)$ , we already know that  $D_\xi^2 W$ , which is a  $(d+1) \times (d+1)$  matrix, has at least  $d$  negative eigenvalue  $\lambda_1, \dots, \lambda_d$ . Let  $\lambda_0$  the remaining eigenvalue, then

$$(-1)^d \operatorname{sign}(\lambda_0) = \operatorname{sign}(\lambda_0 \lambda_1 \cdots \lambda_d) = \operatorname{sign}(\det(D_\xi^2 W)).$$



A straightforward calculation gives

$$\begin{aligned} W(\xi) &= H(u_\xi) - \xi_1 F_1(u_\xi) - \sum_{j=1}^d \xi_{j+1} F_{j+1}(u_\xi) \\ &= \frac{1}{2} \int_{\mathbb{R}^d} |\nabla u_\omega(x)|^2 dx - \frac{1}{p+1} \int_{\mathbb{R}^d} |u_\omega(x)|^{p+1} dx - \frac{\omega}{2} \int_{\mathbb{R}^d} |u_\omega(x)|^2 dx. \end{aligned} \quad (8.12)$$

Therefore,  $W(\xi)$  depends only on the single parameter  $\omega$  which is itself a function of  $\xi$ . As a consequence, for each  $k = 1, \dots, d+1$ ,

$$\frac{\partial W}{\partial \xi_k} = \frac{\partial W}{\partial \omega} \frac{\partial \omega}{\partial \xi_k} = \left( -\frac{1}{2} \int_{\mathbb{R}^d} |u_\omega(x)|^2 dx \right) \frac{\partial \omega}{\partial \xi_k}$$

and, writing  $f(\omega) = \left( -\frac{1}{2} \int_{\mathbb{R}^d} |u_\omega(x)|^2 dx \right)$ ,

$$\frac{\partial^2 W}{\partial \xi_\ell \partial \xi_k} = \frac{\partial f}{\partial \omega} \frac{\partial \omega}{\partial \xi_\ell} \frac{\partial \omega}{\partial \xi_k} + f(\omega) \frac{\partial^2 \omega}{\partial \xi_\ell \partial \xi_k}$$

for any  $\ell = 1, \dots, d+1$ . Recalling  $\omega(\xi) = \xi_1 + \frac{|\xi|^2}{4}$ , this gives

$$\begin{cases} \frac{\partial^2 W}{\partial \xi_1^2} = \frac{\partial f}{\partial \omega} \\ \frac{\partial^2 W}{\partial \xi_1 \partial \xi_k} = \frac{\xi_k}{2} \frac{\partial f}{\partial \omega} & \text{for } k = 2, \dots, d+1 \\ \frac{\partial^2 W}{\partial \xi_\ell \partial \xi_k} = \frac{\xi_\ell \xi_k}{2} \frac{\partial f}{\partial \omega} + \frac{1}{2} \delta_{k\ell} f(\omega) & \text{for } \ell, k = 2, \dots, d+1 \end{cases}$$

Hence,

$$\det(D_\xi^2 W) = \left( \frac{1}{2} f(\omega) \right)^d \frac{\partial f}{\partial \omega} \text{ and } \text{sign}(\det(D_\xi^2 W)) = (-1)^d \text{sign} \left( \frac{\partial f}{\partial \omega} \right).$$

As a consequence,  $\text{sign} \left( \frac{\partial f}{\partial \omega} \right) = \text{sign}(\lambda_0)$ . This implies that  $p(D_\xi^2 W) = 1$  if and only if  $\frac{\partial f}{\partial \omega} > 0$ . Using the definition of  $f(\omega)$ , we can conclude that  $p(D_\xi^2 W) = 1$  if and only if

$$\frac{\partial}{\partial \omega} \int_{\mathbb{R}^d} |u_\omega|^2 dx < 0. \quad (8.13)$$

Condition (8.13) can be rewritten as

$$\frac{\partial}{\partial \omega} \int_{\mathbb{R}^d} |u_\omega|^2 dx = 2 \int_{\mathbb{R}^d} u_\omega \frac{\partial u_\omega}{\partial \omega} dx = 2 \int_{\mathbb{R}^d} u_\omega L_+^{-1} u_\omega dx < 0$$

with  $L_+$  defined above. Hence, let  $S = x \cdot \nabla + \frac{2}{p-1}$ . A straightforward calculation gives  $L_+ S u_\omega = 2\omega u_\omega$ . As a consequence,

$$2 \int_{\mathbb{R}^d} u_\omega L_+^{-1} u_\omega dx = \frac{1}{\omega} \int_{\mathbb{R}^d} u_\omega S u_\omega dx = \frac{1}{\omega} \int_{\mathbb{R}^d} u_\omega \left( x \cdot \nabla + \frac{2}{p-1} \right) u_\omega dx = \frac{1}{\omega} \left( -\frac{d}{2} + \frac{2}{p-1} \right) \int_{\mathbb{R}^d} |u_\omega|^2 dx$$

which is strictly negative if and only if  $p < 1 + \frac{4}{d}$ .

As a consequence, if  $1 < p < 1 + \frac{4}{d}$ , Theorem 6.3 applies and gives the local coercivity of  $D_{u_\xi}^2 \mathcal{L}_\xi$ . We then have:

**Theorem 8.1.** *Let  $d = 1, 2, 3$  and  $1 < p < 1 + \frac{4}{d}$ . The solitary wave  $u_{\omega,c}$ , defined as in (8.6) is an orbitally stable relative equilibrium.*

When  $d = 1$  and  $3 \leq p < 5$ , this follows from Theorem 6.3 together with Proposition 2.3 and the results of [10]. When  $d = 1, 2, 3$  and  $1 < p < 3$ , the nonlinearity is not sufficiently smooth to ensure the “propagation of the regularity” for initial conditions in  $\mathcal{D} = H^3(\mathbb{R}^d)$ , as required in [10] (see [7]). Hence, the results of [10] cannot be directly applied in this case. Nevertheless, to prove the orbital stability once one has the coercivity of  $\mathcal{L}_\xi$ , we can use Theorem 10 of [10] the proof of which can be easily adapted in the case of the Schrödinger equation with a power nonlinearity.

*Remark 8.2.*

- (1) As announced at the end of Section 3,  $\nabla^2 \mathcal{L}_\xi$  is an unbounded partial differential operator and we are in the setting of Theorem 6.3, not of Theorem 3.1, nor of Theorem 9.1 below, which comes from [18].
- (2) A proof of the orbital stability of the soliton of the focusing NLSE for  $1 < p < 1 + \frac{4}{d}$ ,  $d = 1, 2, 3$  was given originally using concentration-compactness arguments in [8] and with a variational method in [40]. Finally, in [18], some of the spectral arguments we used to control  $\nabla^2 \mathcal{L}_\xi$  are provided, but a complete proof of orbital stability is lacking for reasons further explained in the next section.

**8.2. Stability of solitons for a system of coupled nonlinear Schrödinger equations.** We consider the system of two coupled nonlinear Schrödinger equations given by

$$\begin{cases} i\partial_t u_1(t, x) + \Delta u_1(t, x) + (\alpha|u_1(t, x)|^2 + \delta|u_2(t, x)|^2)u_1(t, x) = 0 \\ i\partial_t u_2(t, x) + \Delta u_2(t, x) + (\delta|u_1(t, x)|^2 + \gamma|u_2(t, x)|^2)u_2(t, x) = 0 \\ u(0, x) = u(x) \end{cases} \quad (8.14)$$

with  $u(t, x) = \begin{pmatrix} u_1(t, x) \\ u_2(t, x) \end{pmatrix} : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{C}^2$  and  $d = 1, 2, 3$ . Here,  $\alpha, \gamma \in \mathbb{R}_+$  and  $\delta \in \mathbb{R}_+^*$  are parameters of the model.

Coupled NLSEs have been used to model nonlinear wave propagation in a variety of physical systems. In nonlinear optics, they describe light propagation in birefringent fibers [2]. In the study of ocean waves, they have been proposed as a model for the generation of rogue waves in crossing sea states : these are two-component wave systems with different directions of propagation (See [26] and references therein). They also appear in the study of two-component Bose-Einstein condensates [4, 31]. A central topic in each of these situations is the stability or instability of solitons and plane wave solutions of those equations. We consider solitons in this subsection, and plane waves in the next one.

In dimension  $d = 1, 2, 3$ , the Cauchy problem (8.14) is locally well posed in  $H^1(\mathbb{R}^d, \mathbb{C}^2)$  [6]. Moreover, it has been proved in [12] that, in dimension  $d = 1$ , (8.14) is globally well posed in  $H^1(\mathbb{R}^d, \mathbb{C}^2)$ .

Equation (8.14) is the Hamiltonian differential equation associated to the function  $H$  defined by

$$\begin{aligned} H(u) &= \frac{1}{2} \int_{\mathbb{R}^d} (|\nabla u_1(x)|^2 + |\nabla u_2(x)|^2) \, dx \\ &\quad - \frac{1}{4} \int_{\mathbb{R}^d} (\alpha|u_1(x)|^4 + 2\delta|u_1(x)|^2|u_2(x)|^2 + \gamma|u_2(x)|^4) \, dx. \end{aligned} \quad (8.15)$$

Let  $G = \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d$  and define its action on  $E = H^1(\mathbb{R}^d, \mathbb{C}^2)$  via

$$\forall u \in H^1(\mathbb{R}^d, \mathbb{C}^2), \quad (\Phi_{\gamma_1, \gamma_2, \tau}(u))(x) = \begin{pmatrix} e^{-i\gamma_1} u_1(x - \tau) \\ e^{-i\gamma_2} u_2(x - \tau) \end{pmatrix}. \quad (8.16)$$

The group  $G$  is an invariance group for the dynamics and the quantities

$$F_1(u) = \frac{1}{2} \int_{\mathbb{R}^d} |u_1(x)|^2 dx \quad (8.17)$$

$$F_2(u) = \frac{1}{2} \int_{\mathbb{R}^d} |u_2(x)|^2 dx \quad (8.18)$$

$$F_{2+j}(u) = \frac{1}{2} \int_{\mathbb{R}^d} u_1^* \cdot \left( \frac{1}{i} \partial_{x_j} \right) u_1 dx + \frac{1}{2} \int_{\mathbb{R}^d} u_2^* \cdot \left( \frac{1}{i} \partial_{x_j} \right) u_2 dx \quad (8.19)$$

for  $j = 1, \dots, d$ , are the corresponding constants of the motion. The family of solitary waves

$$u_{\omega_1, \omega_2, c}(x) = e^{i \frac{c}{2} \cdot x} \Phi_{\omega_1, \omega_2} \quad (8.20)$$

with  $c \in \mathbb{R}^d$  and  $\Phi_{\omega_1, \omega_2} = \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}$  a solution to

$$\begin{cases} -\Delta \Phi_1 - \omega_1 \Phi_1 = (\alpha |\Phi_1|^2 + \delta |\Phi_2|^2) \Phi_1 \\ -\Delta \Phi_2 - \omega_2 \Phi_2 = (\delta |\Phi_1|^2 + \gamma |\Phi_2|^2) \Phi_2 \end{cases} \quad (8.21)$$

with  $\omega_1, \omega_2 \in \mathbb{R}$ ,  $\omega_1, \omega_2 < 0$ , are  $G$ -relative equilibria of (8.14). Indeed, if we define  $\mathcal{L}_\xi$  by

$$\mathcal{L}_\xi(u) = H(u) - \xi_1 F_1(u) - \xi_2 F_2(u) - \sum_{j=1}^d \xi_{j+2} F_j(u), \quad (8.22)$$

we can easily verify that  $u_{c, \omega_1, \omega_2}$  is a solution to the stationary equation  $D_u \mathcal{L}_\xi = 0$  with  $\xi = (\omega_1 - \frac{|c|^2}{4}, \omega_2 - \frac{|c|^2}{4}, c)$ .

In particular, if  $\omega_1 = \omega_2 = \omega_* < 0$ , then

$$\Phi_{\omega_*, \omega_*}(x) = u_{\omega_*}(x) \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \quad (8.23)$$

with

$$\alpha \zeta_1^2 + \delta \zeta_2^2 = 1, \quad \delta \zeta_1^2 + \gamma \zeta_2^2 = 1 \quad (8.24)$$

and  $u_{\omega_*}$  the unique positive solution to

$$-\Delta \varphi - \omega_* \varphi = |\varphi|^2 \varphi \quad (8.25)$$

is a solution to (8.21). As a consequence,  $u_{\xi_*}(x) = e^{i \frac{c}{2} \cdot x} \Phi_{\omega_*, \omega_*}(x)$  with  $\xi_* = (\omega_* - \frac{|c|^2}{4}, \omega_* - \frac{|c|^2}{4}, c)$  is a  $G$ -relative equilibrium of (8.14) and our goal is to investigate its orbital stability by means of Theorem 6.3. Note that this kind of solution exists only if  $\delta \notin [\min(\alpha, \gamma), \max(\alpha, \gamma)]$  or if  $\delta = \alpha = \gamma$  which corresponds to the integrable case. In what follows, we will assume  $\delta < \min(\alpha, \gamma)$  or  $\delta > \max(\alpha, \gamma)$ .

First of all, to apply Theorem 6.3, we have to show the existence of a family of solutions to  $D_{u_\xi} \mathcal{L}_\xi = 0$  for each  $\xi$  in a neighbourhood of  $\xi_*$ . The existence of such a family of  $G$ -relative equilibria is obtained using Theorem 7.1.

As before, Hypotheses A (by taking again  $\mathcal{D} = H^3(\mathbb{R}^d, \mathbb{C}^2)$  and the  $L^2$ -scalar product on  $E = H^1(\mathbb{R}^d, \mathbb{C}^2)$ ) and B are clearly satisfied.

A straightforward calculation gives

$$\begin{aligned}
D_u^2 \mathcal{L}_\xi(v, w) = & \operatorname{Re} \left[ \int_{\mathbb{R}^d} (-\Delta w_1) v_1^* + \int_{\mathbb{R}^d} (-\Delta w_2) v_2^* \right. \\
& - \int_{\mathbb{R}^d} (\alpha |u_1|^2 + \delta |u_2|^2) w_1 v_1^* - \int_{\mathbb{R}^d} (\delta |u_1|^2 + \gamma |u_2|^2) w_2 v_2^* \\
& - \int_{\mathbb{R}^d} (\alpha (u_1 w_1^* + u_1^* w_1) + \delta (u_2 w_2^* + u_2^* w_2)) u_1 v_1^* \\
& - \int_{\mathbb{R}^d} (\delta (u_1 w_1^* + u_1^* w_1) + \gamma (u_2 w_2^* + u_2^* w_2)) u_2 v_2^* \\
& \left. - \xi_1 \int_{\mathbb{R}^d} w_1 v_1^* - \xi_2 \int_{\mathbb{R}^d} w_2 v_2^* - \sum_{j=1}^d \xi_{j+2} \int_{\mathbb{R}^d} \left( \frac{1}{i} \partial_{x_j} w_1 \right) v_1^* + \left( \frac{1}{i} \partial_{x_j} w_2 \right) v_2^* \right].
\end{aligned}$$

Writing  $v(x) = e^{i\frac{x}{2}} \tilde{v}(x)$  and  $w(x) = e^{i\frac{x}{2}} \tilde{w}(x)$  and using the particular form of  $u_\xi$ , we obtain

$$\begin{aligned}
D_{u_\xi}^2 \mathcal{L}_\xi(v, w) = & \operatorname{Re} \left[ \int_{\mathbb{R}^d} (-\Delta \tilde{w}_1) \tilde{v}_1^* + \int_{\mathbb{R}^d} (-\Delta \tilde{w}_2) \tilde{v}_2^* \right. \\
& - \int_{\mathbb{R}^d} (\alpha |\Phi_1|^2 + \delta |\Phi_2|^2) \tilde{w}_1 \tilde{v}_1^* - \int_{\mathbb{R}^d} (\delta |\Phi_1|^2 + \gamma |\Phi_2|^2) \tilde{w}_2 \tilde{v}_2^* \\
& - \int_{\mathbb{R}^d} (\alpha (\Phi_1 \tilde{w}_1^* + \Phi_1^* \tilde{w}_1) + \delta (\Phi_2 \tilde{w}_2^* + \Phi_2^* \tilde{w}_2)) \Phi_1 \tilde{v}_1^* \\
& - \int_{\mathbb{R}^d} (\delta (\Phi_1 \tilde{w}_1^* + \Phi_1^* \tilde{w}_1) + \gamma (\Phi_2 \tilde{w}_2^* + \Phi_2^* \tilde{w}_2)) \Phi_2 \tilde{v}_2^* \\
& \left. - \omega_1 \int_{\mathbb{R}^d} \tilde{w}_1 \tilde{v}_1^* - \omega_2 \int_{\mathbb{R}^d} \tilde{w}_2 \tilde{v}_2^* \right].
\end{aligned}$$

Hence, by using the definition of  $u_{\xi_*}$ , we obtain

$$\begin{aligned}
D_{u_{\xi_*}}^2 \mathcal{L}_{\xi_*}(v, w) = & \operatorname{Re} \left[ \int_{\mathbb{R}^d} (-\Delta \tilde{w}_1) \tilde{v}_1^* + \int_{\mathbb{R}^d} (-\Delta \tilde{w}_2) \tilde{v}_2^* - \int_{\mathbb{R}^d} u_{\omega_*}^2 \tilde{w}_1 \tilde{v}_1^* - \int_{\mathbb{R}^d} u_{\omega_*}^2 \tilde{w}_2 \tilde{v}_2^* \right. \\
& - \int_{\mathbb{R}^d} (\alpha \zeta_1 (\tilde{w}_1^* + \tilde{w}_1) + \delta \zeta_2 (\tilde{w}_2^* + \tilde{w}_2)) \zeta_1 u_{\omega_*}^2 \tilde{v}_1^* \\
& - \int_{\mathbb{R}^d} (\delta \zeta_1 (\tilde{w}_1^* + \tilde{w}_1) + \gamma \zeta_2 (\tilde{w}_2^* + \tilde{w}_2)) \zeta_2 u_{\omega_*}^2 \tilde{v}_2^* \\
& \left. - \omega_* \int_{\mathbb{R}^d} \tilde{w}_1 \tilde{v}_1^* - \omega_* \int_{\mathbb{R}^d} \tilde{w}_2 \tilde{v}_2^* \right] = \langle L \tilde{w}, \tilde{v} \rangle = \langle \nabla^2 \mathcal{L}_{\xi_*}(u_{\xi_*}) w, v \rangle
\end{aligned}$$

with

$$L \tilde{w} = \begin{pmatrix} L_+ & 0_2 \\ 0_2 & L_- \end{pmatrix} \tilde{w}, \quad (8.26)$$

$L = U^* \nabla^2 \mathcal{L}_{\xi_*}(u_{\xi_*}) U$  and  $U$  a unitary matrix, and  $L_+$ ,  $L_-$  given by

$$L_+ = \begin{pmatrix} -\Delta - u_{\omega_*}^2 - \omega_* - 2\alpha \zeta_1^2 u_{\omega_*}^2 & -2\delta \zeta_1 \zeta_2 u_{\omega_*}^2 \\ -2\delta \zeta_1 \zeta_2 u_{\omega_*}^2 & -\Delta - u_{\omega_*}^2 - \omega_* - 2\gamma \zeta_2^2 u_{\omega_*}^2 \end{pmatrix}, \quad (8.27)$$

$$L_- = \begin{pmatrix} -\Delta - u_{\omega_*}^2 - \omega_* & 0 \\ 0 & -\Delta - u_{\omega_*}^2 - \omega_* \end{pmatrix}. \quad (8.28)$$

Note that  $\mathcal{D}(\nabla^2 \mathcal{L}_{\xi_*}(u_{\xi_*})) = H^2(\mathbb{R}^d)$ . Next, since  $u_{\omega_*}$  is the unique positive solution of (8.21),  $L_- = -\Delta - u_{\omega_*}^2 - \omega_*$  is a nonnegative operator and  $\operatorname{Ker}(L_-) = \operatorname{span}\{u_{\omega_*}\}$ . Hence,

$$\operatorname{Ker}(L_-) = \operatorname{span} \left\{ \begin{pmatrix} u_{\omega_*} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ u_{\omega_*} \end{pmatrix} \right\}.$$

Next, to analyze the spectrum of  $L_+$ , it is convenient to perform the orthogonal transformation defined by

$$P = \frac{1}{\sqrt{\zeta_1^2 + \zeta_2^2}} \begin{pmatrix} \zeta_1 & \zeta_2 \\ \zeta_2 & -\zeta_1 \end{pmatrix} \quad (8.29)$$

which leads to

$$PL_+P = \begin{pmatrix} -\Delta - 3u_{\omega_*}^2 - \omega_* & 0 \\ 0 & -\Delta - (3 - 2\delta(\zeta_1^2 + \zeta_2^2))u_{\omega_*}^2 - \omega_* \end{pmatrix} := \begin{pmatrix} L_+ & 0 \\ 0 & L_\beta \end{pmatrix}.$$

We know that the operator  $L_+ = -\nabla - 3u_{\omega_*}^2 - \omega_*$  has exactly one negative eigenvalue and that  $\text{Ker}(L_+) = \text{span}\{\partial_{x_1}u_{\omega_*}, \dots, \partial_{x_d}u_{\omega_*}\}$ . Moreover, from [39, Lemma 4.1], we can deduce that  $\text{Ker}(L_\delta) = \{0\}$  except if  $\delta = 0$  or  $\delta = \alpha = \gamma$ . As a consequence, if  $\delta \in \mathbb{R}_+^* \setminus [\min(\alpha, \gamma), \max(\alpha, \gamma)]$ ,

$$\text{Ker}(L_+) = \text{span} \left\{ P \begin{pmatrix} \partial_{x_j}u_{\omega_*} \\ 0 \end{pmatrix}, j = 1, \dots, d \right\} = \text{span} \left\{ \partial_{x_j}u_{\omega_*} \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}, j = 1, \dots, d \right\}. \quad (8.30)$$

This implies  $\text{Ker}(D_{u_{\xi_*}}^2 \mathcal{L}_{\xi_*}) = T_{u_{\xi_*}} \mathcal{O}_{u_{\xi_*}}$ .

Now we have to count the negative eigenvalues of  $L_\delta$ . First of all, we remark that

$$L_\delta = L_- - 2(1 - \delta(\zeta_1^2 + \zeta_2^2))u_{\omega_*}^2.$$

If  $1 - \delta(\zeta_1^2 + \zeta_2^2) < 0$ ,  $L_\delta$  is clearly a positive operator. Then  $n(L_\delta) = 0$  and  $n(D_{u_{\xi_*}}^2 \mathcal{L}_{\xi_*}) = 1$ . This corresponds to the case  $\delta > \max(\alpha, \gamma)$ . If  $\delta < \min(\alpha, \gamma)$ , then  $1 - \delta(\delta_1^2 + \delta_2^2) > 0$  and

$$\langle L_\delta u_{\omega_*}, u_{\omega_*} \rangle = -2(1 - \delta(\zeta_1^2 + \zeta_2^2)) \int_{\mathbb{R}^d} |u_{\omega_*}|^4 < 0$$

so that  $n(L_\delta) \geq 1$ . Since  $L_\delta = L_+ + 2\delta(\zeta_1^2 + \zeta_2^2)u_{\omega_*}^2$ , it is clear, by means of a min-max type argument, that  $n(L_\delta) \leq 1$ . Hence,  $n(L_\delta) = 1$  and  $n(D_{u_{\xi_*}}^2 \mathcal{L}_{\xi_*}) = 2$ . In both cases  $n(D_{u_{\xi_*}}^2 \mathcal{L}_{\xi_*}) < +\infty$  and Theorem 7.1 leads to the existence of a family of  $G_{\mu_\xi}$ -relative equilibria.

Next, to apply Theorem 6.3, we have to show that  $n(D_{u_{\xi_*}}^2 \mathcal{L}_{\xi_*}) = p(D_{\xi_*}^2 W)$  where  $D_{\xi_*}^2 W$  is the notation for the Hessian of  $W$  evaluated at  $\xi_*$ .

A straightforward calculation gives

$$\begin{aligned} W(\xi) &= H(u_\xi) - \xi_1 F_1(u_\xi) - \xi_2 F_2(u_\xi) - \sum_{j=1}^d \xi_{j+1} F_{j+2}(u_\xi) \\ &= \frac{1}{2} \int_{\mathbb{R}^d} (|\nabla \Phi_1(x)|^2 + |\nabla \Phi_2(x)|^2) dx \\ &\quad - \frac{1}{4} \int_{\mathbb{R}^d} (\alpha |\Phi_1(x)|^4 + 2\delta |\Phi_1(x)|^2 |\Phi_2(x)|^2 + \gamma |\Phi_2(x)|^4) dx \\ &\quad - \frac{\omega_1}{2} \int_{\mathbb{R}^d} |\Phi_1(x)|^2 dx - \frac{\omega_2}{2} \int_{\mathbb{R}^d} |\Phi_2(x)|^2 dx. \end{aligned} \quad (8.31)$$

Hence, using the fact that  $\Phi_{\omega_1, \omega_2}$  is a solution to (8.21), for each  $k = 1, \dots, d+2$ , we obtain

$$\begin{aligned} \frac{\partial W}{\partial \xi_k}(\xi) &= -\frac{\partial \omega_1}{\partial \xi_k}(\xi) \frac{1}{2} \int_{\mathbb{R}^d} |\Phi_1(x)|^2 dx - \frac{\partial \omega_2}{\partial \xi_k}(\xi) \frac{1}{2} \int_{\mathbb{R}^d} |\Phi_2(x)|^2 dx \\ &= -\frac{\partial \omega_1}{\partial \xi_k}(\xi) F_1(\Phi_{\omega_1, \omega_2}) - \frac{\partial \omega_2}{\partial \xi_k}(\xi) F_2(\Phi_{\omega_1, \omega_2}). \end{aligned}$$

Recalling  $\omega_1(\xi) = \xi_1 + \frac{|\hat{\xi}|^2}{4}$  and  $\omega_2(\xi) = \xi_2 + \frac{|\hat{\xi}|^2}{4}$ , we have

$$\begin{cases} \frac{\partial W}{\partial \xi_1}(\xi) = -F_1(\Phi_{\omega_1, \omega_2}) \\ \frac{\partial W}{\partial \xi_2}(\xi) = -F_2(\Phi_{\omega_1, \omega_2}) \\ \frac{\partial W}{\partial \xi_k}(\xi) = -\frac{\xi_k}{2}(F_1(\Phi_{\omega_1, \omega_2}) + F_2(\Phi_{\omega_1, \omega_2})) \quad \text{for } k = 3, \dots, d+2 \end{cases}$$

Next, a straightforward computation gives, for all  $\ell = 1, \dots, d+2$ ,

$$\begin{cases} \frac{\partial^2 W}{\partial \xi_\ell \partial \xi_1}(\xi) = -\frac{\partial}{\partial \xi_\ell} F_1(\Phi_{\omega_1, \omega_2}) \\ \frac{\partial^2 W}{\partial \xi_\ell \partial \xi_2}(\xi) = -\frac{\partial}{\partial \xi_\ell} F_2(\Phi_{\omega_1, \omega_2}) \end{cases}$$

and

$$\frac{\partial^2 W}{\partial \xi_\ell \partial \xi_k}(\xi) = -\frac{\xi_k}{2} \left( \frac{\partial}{\partial \xi_\ell} F_1(\Phi_{\omega_1, \omega_2}) + \frac{\partial}{\partial \xi_\ell} F_2(\Phi_{\omega_1, \omega_2}) \right) - \frac{\delta_{k\ell}}{2} (F_1(\Phi_{\omega_1, \omega_2}) + F_2(\Phi_{\omega_1, \omega_2}))$$

for  $k = 3, \dots, d+2$ .

Note that for all  $k = 3, \dots, d+2$  and for all  $\ell = 1, \dots, d+2$ ,

$$\frac{\partial^2 W}{\partial \xi_\ell \partial \xi_k}(\xi) - \frac{\xi_k}{2} \left( \frac{\partial^2 W}{\partial \xi_\ell \partial \xi_1}(\xi) + \frac{\partial^2 W}{\partial \xi_\ell \partial \xi_2}(\xi) \right) = -\frac{\delta_{k\ell}}{2} (F_1(\Phi_{\omega_1, \omega_2}) + F_2(\Phi_{\omega_1, \omega_2})).$$

Hence, for all  $\lambda \in \mathbb{R}$ ,

$$\det(D_\xi^2 W - \lambda \mathbb{I}) = \left( -\frac{1}{2} (F_1(\Phi_{\omega_1, \omega_2}) + F_2(\Phi_{\omega_1, \omega_2})) - \lambda \right)^d \det \left( \begin{pmatrix} \frac{\partial^2 W}{\partial \xi_1^2} & \frac{\partial^2 W}{\partial \xi_1 \partial \xi_2} \\ \frac{\partial^2 W}{\partial \xi_2 \partial \xi_1} & \frac{\partial^2 W}{\partial \xi_2^2} \end{pmatrix} - \lambda \mathbb{I} \right)$$

This means that  $D_\xi^2 W$  has at least  $d$  negatives eigenvalues and  $p(D_\xi^2 W) \leq 2$ . More precisely, the two remaining eigenvalues are the eigenvalues of the matrix

$$M = \begin{pmatrix} \frac{\partial^2 W}{\partial \xi_1^2} & \frac{\partial^2 W}{\partial \xi_1 \partial \xi_2} \\ \frac{\partial^2 W}{\partial \xi_2 \partial \xi_1} & \frac{\partial^2 W}{\partial \xi_2^2} \end{pmatrix} = \begin{pmatrix} -\frac{\partial F_1}{\partial \omega_1} & -\frac{\partial F_1}{\partial \omega_2} \\ -\frac{\partial F_2}{\partial \omega_1} & -\frac{\partial F_2}{\partial \omega_2} \end{pmatrix}$$

On the one hand, if  $\frac{\partial F_1}{\partial \omega_1} \frac{\partial F_2}{\partial \omega_2} - \frac{\partial F_1}{\partial \omega_2} \frac{\partial F_2}{\partial \omega_1} < 0$ , the matrix  $M$  has exactly one positive eigenvalue and  $p(D_\xi^2 W) = 1$ . On the other hand, if  $\frac{\partial F_1}{\partial \omega_1} \frac{\partial F_2}{\partial \omega_2} - \frac{\partial F_1}{\partial \omega_2} \frac{\partial F_2}{\partial \omega_1} > 0$  and  $\frac{\partial F_1}{\partial \omega_1} + \frac{\partial F_2}{\partial \omega_2} < 0$ , the matrix  $M$  has two positive eigenvalues and  $p(D_\xi^2 W) = 2$ .

By applying Theorem 6.3, we obtain to the following result.

**Theorem 8.3.** (1) Let  $\delta > \max(\alpha, \gamma)$ . If

$$\frac{\partial F_1}{\partial \omega_1}(\Phi_{\omega_*, \omega_*}) \frac{\partial F_2}{\partial \omega_2}(\Phi_{\omega_*, \omega_*}) - \frac{\partial F_1}{\partial \omega_2}(\Phi_{\omega_*, \omega_*}) \frac{\partial F_2}{\partial \omega_1}(\Phi_{\omega_*, \omega_*}) < 0$$

then  $u_{\omega_*, \omega_*, c}$  is such that the local coercivity estimate (6.5) is satisfied.

(2) Let  $\delta < \min(\alpha, \gamma)$ . If

$$\frac{\partial F_1}{\partial \omega_1}(\Phi_{\omega_*, \omega_*}) \frac{\partial F_2}{\partial \omega_2}(\Phi_{\omega_*, \omega_*}) - \frac{\partial F_1}{\partial \omega_2}(\Phi_{\omega_*, \omega_*}) \frac{\partial F_2}{\partial \omega_1}(\Phi_{\omega_*, \omega_*}) > 0 \quad \text{and} \quad \frac{\partial F_1}{\partial \omega_1}(\Phi_{\omega_*, \omega_*}) + \frac{\partial F_2}{\partial \omega_2}(\Phi_{\omega_*, \omega_*}) < 0$$

then  $u_{\omega_*, \omega_*, c}$  is such that the local coercivity estimate (6.5) is satisfied.

A lengthy but straightforward computation gives

$$\frac{\partial F_1}{\partial \omega_1}(\Phi_{\omega_*, \omega_*}) \frac{\partial F_2}{\partial \omega_2}(\Phi_{\omega_*, \omega_*}) - \frac{\partial F_1}{\partial \omega_2}(\Phi_{\omega_*, \omega_*}) \frac{\partial F_2}{\partial \omega_1}(\Phi_{\omega_*, \omega_*}) = \frac{\delta_1^2 \delta_2^2}{2\omega_*} \left( 1 - \frac{d}{2} \right) \left( \int_{\mathbb{R}^d} u_{\omega_*}^2 \right) \int_{\mathbb{R}^d} u_{\omega_*} L_\delta^{-1} u_{\omega_*}$$

and

$$\frac{\partial F_1}{\partial \omega_1}(\Phi_{\omega_*, \omega_*}) + \frac{\partial F_2}{\partial \omega_2}(\Phi_{\omega_*, \omega_*}) = \frac{\zeta_1^4 + \zeta_2^4}{\zeta_1^2 + \zeta_2^2} \frac{(1 - \frac{d}{2})}{2\omega_*} \int_{\mathbb{R}^d} u_{\omega_*}^2 + \frac{2\zeta_1^2 \zeta_2^2}{\zeta_1^2 + \zeta_2^2} \int_{\mathbb{R}^d} u_{\omega_*} L_\delta^{-1} u_{\omega_*}.$$

Recalling that if  $\delta > \max(\alpha, \gamma)$  then  $L_\delta$  is a positive operator, together with Proposition 2.3 and the results of [10], we obtain the following corollary.

**Theorem 8.4.** *Let  $d = 1$ .*

- (1) *If  $\delta > \max(\alpha, \gamma)$  then  $u_{\omega_*, \omega_*, c}$  is an orbitally stable relative equilibrium.*
- (2) *If  $\delta < \min(\alpha, \gamma)$  and  $\int_{\mathbb{R}^d} u_{\omega_*} L_\delta^{-1} u_{\omega_*} < 0$  then  $u_{\omega_*, \omega_*, c}$  is an orbitally stable relative equilibrium.*

Note that a proof of orbital of  $u_{\omega_*, \omega_*, 0}$  in dimension  $d = 1$  and for  $\alpha = \gamma = 1$  was given in [25] using concentration-compactness arguments.

### 8.3. Stability of plane waves for a system of coupled nonlinear Schrödinger equations.

We consider a system of two coupled nonlinear Schrödinger equations on a one-dimensional torus  $\mathbb{T} = \mathbb{T}^1$  is the one-dimensional torus of length  $L > 0$ . The system is given by

$$\begin{cases} i\partial_t u_1(t, x) + \beta \Delta u_1(t, x) + (\alpha |u_1(t, x)|^2 + \delta |u_2(t, x)|^2) u_1(t, x) = 0 \\ i\partial_t u_2(t, x) + \beta \Delta u_2(t, x) + (\delta |u_1(t, x)|^2 + \gamma |u_2(t, x)|^2) u_2(t, x) = 0 \\ u(0, x) = u(x) \end{cases} \quad (8.32)$$

with  $u(t, x) = \begin{pmatrix} u_1(t, x) \\ u_2(t, x) \end{pmatrix} : \mathbb{R} \times \mathbb{T} \rightarrow \mathbb{C}^2$ . The constants  $\alpha, \gamma, \delta \in \mathbb{R}$  and  $\beta \in \mathbb{R}_+^*$  are parameters of the model.

As already mentioned, this system is of relevance in nonlinear optics. Linear instability is in this context referred to as *modulational instability* and was studied for various parameter ranges in [2, 9, 13], among others. We will use the methods exposed in this paper to show that, in the parameter regimes where linear stability can be established, orbital stability also holds.

The four-parameter family of plane waves

$$\tilde{u}_\xi(t, x) = \begin{pmatrix} \zeta_1 e^{ik_1 \cdot x} e^{-i\xi_1 t} \\ \zeta_2 e^{ik_2 \cdot x} e^{-i\xi_2 t} \end{pmatrix} \quad (8.33)$$

with  $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$ ,  $(k_1, k_2) \in \frac{2\pi}{L}\mathbb{Z}^2$  and  $\alpha = (\alpha_1, \alpha_2) \in \mathbb{R}^2$  such that

$$\begin{cases} \xi_1 = \beta k_1^2 - (\alpha \zeta_1^2 + \delta \zeta_2^2) \\ \xi_2 = \beta k_2^2 - (\delta \zeta_1^2 + \gamma \zeta_2^2). \end{cases} \quad (8.34)$$

are solution to the equation (8.32) and we are interested in study their orbital stability. Using Galilean invariance of the equation (see [10]), the stability of these plane waves is seen to be equivalent to that of

$$\tilde{u}_\xi(t, x) = \begin{pmatrix} \zeta_1 e^{ikx} e^{-i\xi_1 t} \\ \zeta_2 e^{-ikx} e^{-i\xi_2 t} \end{pmatrix} \quad (8.35)$$

with  $k = k_1 - k_2$ . Furthermore, we can easily remark that  $\tilde{u}_\xi(t, x)$  can be written in the form

$$\tilde{u}_\xi(t, x) = \begin{pmatrix} e^{ikx} & 0 \\ 0 & e^{-ikx} \end{pmatrix} u_\xi(t, x)$$

with

$$u_\xi(t, x) = \begin{pmatrix} \zeta_1 e^{-i\xi_1 t} \\ \zeta_2 e^{-i\xi_2 t} \end{pmatrix} \quad (8.36)$$

a solution to the system of coupled nonlinear Schrödinger equations

$$\begin{cases} i\partial_t u_1 + \beta\Delta u_1 + 2\beta ik\nabla u_1 + (\alpha|u_1|^2 + \delta|u_2|^2)u_1 - \beta k^2 u_1 = 0 \\ i\partial_t u_2 + \beta\Delta u_2 - 2\beta ik\nabla u_2 + (\delta|u_1|^2 + \gamma|u_2|^2)u_2 - \beta k^2 u_2 = 0 \\ u(0, x) = u(x) \end{cases} \quad (8.37)$$

It is easy to show that the Cauchy problem (8.37) is globally well-posed in  $H^1(\mathbb{T}, \mathbb{C}^2)$  (since we consider here only the dimension  $d = 1$ ).

Equation (8.37) is the Hamiltonian differential equation associated to the function  $H$  defined by

$$\begin{aligned} H(u) &= \frac{\beta}{2} \int_0^L (|(\nabla + ik)u_1(x)|^2 + |(\nabla - ik)u_2(x)|^2) dx \\ &\quad - \frac{1}{4} \int_0^L (\alpha|u_1(x)|^4 + 2\delta|u_1(x)|^2|u_2(x)|^2 + \gamma|u_2(x)|^4) dx. \end{aligned} \quad (8.38)$$

Let  $G = \mathbb{R} \times \mathbb{R}$  and define its action on  $E = H^1(\mathbb{T}, \mathbb{C}^2)$  via

$$\forall u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \in H^1(\mathbb{T}, \mathbb{C}^2), \quad (\Phi_{\gamma_1, \gamma_2}(u))(x) = \begin{pmatrix} e^{-i\gamma_1} u_1(x) \\ e^{-i\gamma_2} u_2(x) \end{pmatrix}. \quad (8.39)$$

The group  $G$  is an invariance group for the dynamics and the quantities

$$F_1(u) = \frac{1}{2} \int_0^L |u_1(x)|^2 dx, \quad F_2(u) = \frac{1}{2} \int_0^L |u_2(x)|^2 dx$$

are the corresponding constants of the motion.

The two-parameter family of plane waves

$$u_\xi(x) = \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix} \quad (8.40)$$

with  $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$  and  $\zeta_1, \zeta_2 \in \mathbb{R} \setminus \{0\}$  such that

$$\begin{cases} \xi_1 = \beta k^2 - (\alpha\zeta_1^2 + \delta\zeta_2^2) \\ \xi_2 = \beta k^2 - (\delta\zeta_1^2 + \gamma\zeta_2^2). \end{cases} \quad (8.41)$$

are solutions to the stationary equation (2.9). As a consequence,  $u_\xi$  is  $G_{\mu_\xi}$ -relative equilibria of (8.37) and our goal is to investigate the orbital stability of these plane wave solutions by applying Theorem 6.3.

From now, assume that  $\alpha\gamma \neq \delta^2$ . This is the necessary and sufficient condition for the map

$$\mu = \left( \frac{L}{2} |\zeta_1|^2, \frac{L}{2} |\zeta_2|^2 \right) \rightarrow (\xi_1, \xi_2)$$

defined in (8.41) to be invertible. Its inverse is  $\hat{F}$ , which is a diffeomorphism. Note that this condition corresponds to the case in which the system is not completely integrable.

As before, hypotheses A (by taking the  $L^2$ - scalar product on  $E = H^1(\mathbb{T}, \mathbb{C}^2)$ ), B and C are clearly satisfied. By using the dispersion relation (8.34), we have

$$\begin{aligned} W(\xi) &= H(u_\xi) - \xi_1 F_1(u_\xi) - \xi_2 F_2(u_\xi) \\ &= \frac{L}{4(\alpha\gamma - \delta^2)} [\gamma(\xi_1 - \beta k^2)^2 - 2\delta(\xi_1 - \beta k^2)(\xi_2 - \beta k^2) + \alpha(\xi_2 - \beta k^2)^2] \end{aligned} \quad (8.42)$$

As a consequence,

$$D_\xi^2 W = \frac{L}{2(\alpha\gamma - \delta^2)} \begin{pmatrix} \gamma & -\delta \\ -\delta & \alpha \end{pmatrix} \quad (8.43)$$

It is clear that  $D_\xi^2 W$  is non-degenerate. Moreover a straightforward calculation shows that

- (1) If  $\alpha\gamma - \delta^2 > 0$  and  $\min(\alpha, \gamma) > 0$ , then  $p(D_\xi^2 W) = 2$ ;



- (2) If  $\alpha\gamma - \delta^2 > 0$  and  $\max(\alpha, \gamma) < 0$ , then  $p(D_\xi^2 W) = 0$ ;  
(3) If  $\alpha\gamma - \delta^2 < 0$ , then  $p(D_\xi^2 W) = 1$ .

Next, we have to compute  $D_{u_\xi}^2 \mathcal{L}_\xi(v, v)$  with  $\mathcal{L}_\xi(u) = H(u) - \xi_1 F_1(u) - \xi_2 F_2(u)$ . A straightforward calculation gives  $D_{u_\xi}^2 \mathcal{L}_\xi(v, v) = \langle \nabla^2 \mathcal{L}_\xi(u_\xi) v, v \rangle$  with

$$\nabla^2 \mathcal{L}_\xi(u_\xi) = \begin{pmatrix} -\beta\Delta - 2\alpha\zeta_1^2 & -2\delta\zeta_1\zeta_2 & 2\beta k\nabla & 0 \\ -2\delta\zeta_1\zeta_2 & -\beta\Delta - 2\gamma\zeta_2^2 & 0 & -2\beta k\nabla \\ -2\beta k\nabla & 0 & -\beta\Delta & 0 \\ 0 & 2\beta k\nabla & 0 & -\beta\Delta \end{pmatrix}. \quad (8.44)$$

In particular, in this functional space setting the hypotheses of Lemma 6.1 are clearly satisfied by  $D_{u_\xi}^2 \mathcal{L}_\xi(\cdot, \cdot)$ .

Using Fourier series, we can show that the eigenvalues of  $\nabla^2 \mathcal{L}_\xi(u_\xi)$ , for all  $n \in \mathbb{N}$ , are of the form

$$\lambda_{\pm, n}^+ = \beta \left( \frac{2\pi}{L} n \right)^2 + \frac{1}{2} \left( C_+ \pm \sqrt{C_+^2 + 16\beta^2 k^2 \left( \frac{2\pi}{L} n \right)^2} \right) \quad (8.45)$$

$$\lambda_{\pm, n}^- = \beta \left( \frac{2\pi}{L} n \right)^2 + \frac{1}{2} \left( C_- \pm \sqrt{C_-^2 + 16\beta^2 k^2 \left( \frac{2\pi}{L} n \right)^2} \right) \quad (8.46)$$

with

$$C_\pm = -(\alpha\zeta_1^2 + \gamma\zeta_2^2) \pm \sqrt{(\alpha\zeta_1^2 + \gamma\zeta_2^2)^2 - 4\zeta_1^2\zeta_2^2(\alpha\gamma - \delta^2)},$$

By analyzing the sign of the eigenvalues for  $n = 0$ , we obtain the following situation

- (1) If  $\alpha\gamma - \delta^2 > 0$  and  $\min(\alpha, \gamma) > 0$ , then  $\lambda_{-, 0}^+$  and  $\lambda_{-, 0}^-$  are both negative and  $n(D_{u_\xi}^2 \mathcal{L}_\xi) \geq 2$ .  
(2) If  $\alpha\gamma - \delta^2 > 0$  and  $\min(\alpha, \gamma) < 0$ , then  $\lambda_{+, 0}^+$  and  $\lambda_{+, 0}^-$  are both positive. This implies  $\lambda_{+, n}^+ > 0$  and  $\lambda_{+, n}^- > 0$  for all  $n \in \mathbb{N}$ .  
(3) If  $\alpha\gamma - \delta^2 < 0$ , then  $\lambda_{-, 0}^- < 0$ ,  $\lambda_{+, 0}^+ > 0$  and  $n(D_{u_\xi}^2 \mathcal{L}_\xi) \geq 1$ . As a consequence  $\lambda_{+, n}^+ > 0$  for all  $n \in \mathbb{N}$ .

In all the cases, the two remaining eigenvalues are both 0 with purely imaginary eigenvectors which implies that

$$T_{u_\xi} \mathcal{O}_{u_\xi} = \text{span} \left\{ \begin{pmatrix} i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ i \end{pmatrix} \right\} \subset \text{Ker} (D_{u_\xi}^2 \mathcal{L}_\xi).$$

and  $\lambda_{+, n}^\pm > 0$  for all  $n \in \mathbb{N}^*$ .

Next, a straightforward calculation shows that if we assume

$$C_\pm^2 + 16\beta^2 k^2 \left( \frac{2\pi}{L} \right)^2 > 16\beta^2 k^4 \quad (8.47)$$

then  $\lambda_{\pm, n}^\pm$  is increasing as a function of  $n$  for all  $n \in \mathbb{N}^*$ . Hence, it enough to suppose that  $\lambda_{-, 1}^\pm > 0$ , that is

$$\beta \left( \frac{2\pi}{L} \right)^2 + C_\pm > 4\beta k^2, \quad (8.48)$$

to conclude that  $p(D_\xi^2 W) = n(D_{u_\xi}^2 \mathcal{L}_\xi)$ . Note that condition (8.48) implies condition (8.47). Moreover, since  $C_+ \geq C_-$ , it is enough to assume that

$$\beta \left( \frac{2\pi}{L} \right)^2 + C_- > 4\beta k^2.$$

Moreover, it is then clear that

$$\inf(\sigma(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap (0, +\infty)) > 0.$$

and

$$\text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) = \text{Ker}(\nabla^2 \mathcal{L}_\xi(u_\xi)) = \text{span} \left\{ \begin{pmatrix} i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ i \end{pmatrix} \right\} = T_{u_\xi} \mathcal{O}_{u_\xi}.$$

Hence Theorem 6.3 applies and, together with Proposition 2.3 and the results of [10], leads to the following result.

**Proposition 8.5.** *Let  $k \in \frac{2\pi}{L}\mathbb{Z}$ ,  $\zeta_1, \zeta_2 \in \mathbb{R}^*$  and  $\gamma_1, \gamma_2 \in \mathbb{R}$ ,  $\gamma_{12} \in \mathbb{R}^*$  such that  $\alpha\gamma \neq \delta^2$ . If*

$$\beta \left( \frac{2\pi}{L} \right)^2 + C_- > 4\beta k^2 \quad (8.49)$$

then  $u_\xi = \begin{pmatrix} \zeta_1 \\ \zeta_2 \end{pmatrix}$  is an orbitally stable  $G_{\mu_\xi}$ -relative equilibrium.

We note that in the present example, one does not strictly need to use the Vakhitov-Kolokolov condition since a direct study of the Hessian of  $D_{u_\xi}^2 \mathcal{L}_\xi$  restricted to  $T_{u_\xi} \Sigma_{\mu_\xi} \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp$  could also be performed starting from (8.44).

8.3.1. *Case  $k = 0$ .* If  $k = 0$ , condition (8.49) reads as

$$\beta \left( \frac{2\pi}{L} \right)^2 - (\alpha\zeta_1^2 + \gamma\zeta_2^2) - \sqrt{(\alpha\zeta_1^2 + \gamma\zeta_2^2)^2 - 4\zeta_1^2\zeta_2^2(\alpha\gamma - \delta^2)} > 0 \quad (8.50)$$

We remark that condition (8.50) is a necessary and sufficient condition for linear stability. To see this, note that the linearization of (8.37) around  $u_\xi = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$  is given by the system

$$\partial_t \begin{pmatrix} \text{Re}(v_1) \\ \text{Re}(v_2) \\ \text{Im}(v_1) \\ \text{Im}(v_2) \end{pmatrix} = \mathbb{L} \begin{pmatrix} \text{Re}(v_1) \\ \text{Re}(v_2) \\ \text{Im}(v_1) \\ \text{Im}(v_2) \end{pmatrix}, \quad \mathbb{L} = \begin{pmatrix} -2\beta k \nabla & 0 & -\beta \Delta & 0 \\ 0 & 2\beta k \nabla & 0 & -\beta \Delta \\ \beta \Delta + 2\alpha\zeta_1^2 & 2\delta\zeta_1\zeta_2 & -2\beta k \nabla & 0 \\ 2\delta\zeta_1\zeta_2 & \beta \Delta + 2\gamma\zeta_2^2 & 0 & 2\beta k \nabla \end{pmatrix}. \quad (8.51)$$

A solution to (8.37) is said to be linearly stable if all the eigenvalues of  $\mathbb{L}$  are purely imaginary. By using Fourier series, the eigenvalues of (8.51) can be seen to be the zeros of the characteristic polynomial

$$\begin{aligned} P_n(\lambda) &= \lambda^4 - 2\lambda^2 \beta n_L^2 (-\beta n_L^2 + (\alpha\zeta_1^2 + \gamma\zeta_2^2) - 4\beta k^2) + i\lambda 8\beta n_L^2 \beta k n_L (\alpha\zeta_1^2 - \gamma\zeta_2^2) \\ &\quad + (\beta n_L^2)^3 (\beta n_L^2 - 2(\alpha\zeta_1^2 + \gamma\zeta_2^2)) + 4(\beta n_L^2)^2 \zeta_1^2 \zeta_2^2 (\alpha\gamma - \delta^2) \\ &\quad + 8\beta n_L^2 \beta^2 k^2 n_L^2 (-\beta n_L^2 + (\alpha\zeta_1^2 + \gamma\zeta_2^2) + 2\beta k^2) \end{aligned}$$

with  $n_L = \left(\frac{2\pi}{L}n\right)$ . So, whenever  $k = 0$ ,  $P_n(\lambda)$  reduces to

$$\begin{aligned} P_n(\lambda) &= \lambda^4 - 2\lambda^2 \beta n_L^2 (-\beta n_L^2 + (\alpha\zeta_1^2 + \gamma\zeta_2^2)) \\ &\quad + (\beta n_L^2)^3 (\beta n_L^2 - 2(\alpha\zeta_1^2 + \gamma\zeta_2^2)) + 4(\beta n_L^2)^2 \zeta_1^2 \zeta_2^2 (\alpha\gamma - \delta^2). \end{aligned}$$

and, for all  $n \in \mathbb{N}$  the eigenvalues of  $\mathbb{L}$  are

$$\tilde{\lambda}_{\pm, n}^2 = \beta n_L^2 \left( -\beta n_L^2 + (\alpha\zeta_1^2 + \gamma\zeta_2^2) \pm \sqrt{(\alpha\zeta_1^2 - \gamma\zeta_2^2)^2 + 4\zeta_1^2 \zeta_2^2 \delta^2} \right).$$

Now it is clear that  $\tilde{\lambda}_{+, 1}^2 < 0$  if and only if condition (8.50) holds. In that case, for all  $n \in \mathbb{N}^*$ ,  $\tilde{\lambda}_{+, n}^2 \leq \tilde{\lambda}_{+, 1}^2$ . Moreover,  $\tilde{\lambda}_{-, n}^2 \leq \tilde{\lambda}_{+, n}^2$  for all  $n \in \mathbb{N}^*$ . As a consequence, all the eigenvalues of  $\mathbb{L}$  are purely imaginary and the corresponding plane wave is linearly stable if and only if (8.50) holds.

8.3.2. *Case  $k \neq 0$ .* If  $k \neq 0$ , in the particular case  $\alpha\zeta_1^2 = \gamma\zeta_2^2$  which is a generalization of the set of parameters treated in [9], we can show that condition (8.49) is a necessary and sufficient condition for linear stability.

Indeed, in this case,  $P_n(\lambda)$  the characteristic polynomial of (8.51) reduces to

$$\begin{aligned} P_n(\lambda) = & \lambda^4 - 2\lambda^2\beta n_L^2(-\beta n_L^2 + (\alpha\zeta_1^2 + \gamma\zeta_2^2) - 4\beta k^2) + i\lambda 8\beta n_L^2\beta k n_L(\alpha\zeta_1^2 - \gamma\zeta_2^2) \\ & + (\beta n_L^2)^3(\beta n_L^2 - 2(\alpha\zeta_1^2 + \gamma\zeta_2^2)) + 4(\beta n_L^2)^2\zeta_1^2\zeta_2^2(\alpha\gamma - \delta^2) \\ & + 8(\beta n_L^2)^2\beta k^2(-\beta n_L^2 + (\alpha\zeta_1^2 + \gamma\zeta_2^2) + 2\beta k^2) \end{aligned}$$

and, for all  $n \in \mathbb{N}$ , the eigenvalues are given by

$$\tilde{\lambda}_{\pm,n}^2 = \beta n_L^2 \left[ -\beta n_L^2 + (\alpha\zeta_1^2 + \gamma\zeta_2^2) - 4\beta k^2 \pm \sqrt{4\zeta_1^2\zeta_2^2\delta^2 + 16\beta k^2(\beta n_L^2 - (\alpha\zeta_1^2 + \gamma\zeta_2^2))} \right].$$

Now, a tedious but straightforward calculation shows that  $\tilde{\lambda}_{+,n}^2 < 0$  for all  $n \in \mathbb{N}^*$  if and only if condition (8.49) holds. Moreover,  $\tilde{\lambda}_{-,n}^2 \leq \tilde{\lambda}_{+,n}^2$  for all  $n \in \mathbb{N}^*$ . As a consequence, all the eigenvalues of  $\mathbb{L}$  are purely imaginary and the corresponding plane wave is linearly stable if and only if (8.49) holds.

8.3.3. *Physical interpretation.* To sum up, we can conclude that, given  $k = k_1 - k_2 \in \frac{2\pi}{L}\mathbb{Z}$ ,  $\zeta_1, \zeta_2 \in \mathbb{R}^*$  and  $\alpha, \gamma, \delta \in \mathbb{R}^*$  such that  $\alpha\gamma \neq \delta^2$ , if

$$\beta \left( \frac{2\pi}{L} \right)^2 - (\alpha\zeta_1^2 + \gamma\zeta_2^2) - \sqrt{(\alpha\zeta_1^2 + \gamma\zeta_2^2)^2 - 4\zeta_1^2\zeta_2^2(\alpha\gamma - \delta^2)} > 4\beta k^2 \quad (8.52)$$

then the plane waves given by (8.33) are orbitally stable  $G_{\mu_\xi}$ -relative equilibria. Moreover, for  $k = 0$ , if this condition is not satisfied the plane wave is unstable (at least linearly). For  $k \neq 0$  this remains true whenever  $\alpha\zeta_1^2 = \gamma\zeta_2^2$ .

We know from [10] that plane waves solutions to a cubic defocusing nonlinear Schrödinger equation on the one-dimensional torus are orbitally stable. This means that whenever,  $\delta = 0$ ,  $\alpha < 0$  and  $\gamma < 0$ , all the plane waves of the form

$$u_\xi(t, x) = \begin{pmatrix} \zeta_1 e^{ik_1 \cdot x} e^{-i\xi_1 t} \\ \zeta_2 e^{ik_2 \cdot x} e^{-i\xi_2 t} \end{pmatrix}$$

are orbitally stable. It is natural to ask what happens if  $|\delta| \neq 0$ . We have two different situations:  $k = 0$  (*i.e.* the plane waves have the same wave number  $k_1 = k_2$ ) and  $k \neq 0$ . If  $k = 0$  and  $\delta^2 < \alpha\gamma$ , which means that the coupling is weak, then  $C_- > 0$  and condition (8.52) remains true. This means that the plane waves with  $k = 0$  then remain stable. If  $k = 0$  and  $\delta^2 > \alpha\gamma$ , which means that the coupling is strong, then  $C_- < 0$  and condition (8.52) fails at least if  $L$  is large enough. Then the plane waves considered become unstable.

In the case  $k \neq 0$ , note that condition (8.52) can be satisfied only in the case (2) above, namely when  $\alpha\gamma - \delta^2 > 0$ , and  $\max(\alpha, \gamma) < 0$ , since otherwise  $C_- < 0$ . This corresponds to a relatively small perturbation of two uncoupled defocusing Schrödinger equations with orbitally stable plane wave solutions. Condition (8.52) can then be satisfied for a finite number of values of  $k$ , provided  $C_-$  is large enough, but it fails for larger ones. The size of  $C_-$  depends in particular on the ‘‘power’’ of the plane wave, determined by  $|\zeta_1|$  and  $|\zeta_2|$ . For larger values of  $k$ , the plane wave becomes linearly unstable, on the other hand, even at weak coupling. In other words, high  $k$  plane waves show modulational instability, even at arbitrarily low  $\delta$ .

## 9. ON THE LINK WITH GRILLAKIS-SHATAH-STRAUSS

We will now compare the results in this paper to [18]. As we have already pointed out, in [18] a proof of orbital stability is proposed with respect to an a priori different subgroup of  $G$  and under similar but nevertheless different conditions. Both in order to understand the general structure of

the theory and with an eye towards further applications, it is important to understand the relations between the two approaches.

**9.1. The main coercivity estimate of Grillakis-Shatah-Strauss.** Since in [18] the phase space  $E$  on which the dynamics takes place is taken to be a Hilbert space, we place ourselves for this discussion in the Hilbert space setting of Section 3 and consider the situation described by (3.1)-(3.6).

To state the coercivity estimate of [18] which is the analog of our Theorem 3.1, we need some additional notation. We define

$$\tilde{W} : \Omega \cap \mathfrak{g}_\xi \rightarrow \mathcal{L}_\xi(u_\xi) \in \mathbb{R}, \quad (9.1)$$

which is the restriction of the  $W$ -function (3.3) to the sub-Lie-algebra  $\mathfrak{g}_\xi$  of  $\mathfrak{g}$ , defined in (4.5). Also

$$\tilde{\mathcal{O}}_{u_\xi} = \Phi_{G_\xi}(u_\xi), \quad (9.2)$$

is the  $G_\xi$  orbit through  $u_\xi$ . Since a priori  $G_\xi$  differs from  $G_{\mu_\xi}$ , one should not confuse  $\tilde{\mathcal{O}}_{u_\xi}$  with  $\mathcal{O}_{u_\xi}$ , which is the  $G_{\mu_\xi}$ -orbit through  $u_\xi$ . We introduce furthermore

$$\tilde{\Sigma}_\xi = \{v \in E \mid \eta \cdot F(v) = \eta \cdot \mu_\xi, \forall \eta \in \mathfrak{g}_\xi\}. \quad (9.3)$$

In other words,  $\tilde{\Sigma}_\xi$  is the constraint surface corresponding to the constants of the motion  $\eta \cdot F$  for  $\eta \in \mathfrak{g}_\xi$ . Note that  $\Sigma_{\mu_\xi} \subset \tilde{\Sigma}_\xi$ . In fact, when the moment map is regular at  $\mu_\xi$ , then  $\tilde{\Sigma}_\xi$  is a submanifold of  $E$  of co-dimension  $\dim \mathfrak{g}_\xi$  which contains the submanifold  $\Sigma_{\mu_\xi}$ , itself of codimension  $\dim \mathfrak{g}$ . The following theorem, which is the analog of Theorem 3.1 above, can be inferred from the proof of Theorem 4.1 in [18].

**Theorem 9.1.** *Suppose Hypotheses A and C hold. Let  $\xi \in \Omega$  and suppose*

- (i)  $D_\xi^2 \tilde{W}$  is non-degenerate, i.e.  $\text{Ker}(D_\xi^2 \tilde{W}) = \{0\}$ ,
- (ii)  $\text{Ker} D_{u_\xi}^2 \mathcal{L}_\xi = Z_\xi$ ,
- (iii)  $\inf(\sigma(\nabla^2 \mathcal{L}_\xi(u_\xi)) \cap (0, +\infty)) > 0$ ,
- (iv)  $p(D_\xi^2 \tilde{W}) = n(D_{u_\xi}^2 \mathcal{L}_\xi)$ .

Then there exists  $\delta > 0$  such that

$$\forall v \in T_{u_\xi} \tilde{\Sigma}_\xi \cap \left(T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi}\right)^\perp, \quad D_{u_\xi}^2 \mathcal{L}_\xi(v, v) \geq \delta \|v\|^2. \quad (9.4)$$

It is clear that, when the invariance group  $G$  is one-dimensional, i.e.  $\dim \mathfrak{g} = 1$ , this theorem is identical to Theorem 3.1. Indeed, then  $G = G_\xi = G_{\mu_\xi}$  and hence  $W = \tilde{W}$  so that both the assumptions and the conclusions of both theorems are identical. This is the situation studied in [17] and [34]. The same conclusions hold true more generally when the group  $G$  is abelian, since then again,  $G_\xi = G_{\mu_\xi} = G$ . In general, however, the groups  $G_\xi$  and  $G_{\mu_\xi}$  may be distinct, and so may therefore be the orbits  $\tilde{\mathcal{O}}_{u_\xi}$  and  $\mathcal{O}_{u_\xi}$ . Hence, a priori, the two approaches could yield different coercivity estimates and hence different stability results. Their comparison therefore needs to be done with care, a task we turn to in the next subsection.

*Remark 9.2.*

- (1) A proof of Theorem 9.1 can be given along the same lines as the proof of Theorem 3.1 in Section 5 and we don't reproduce it here. We point out that in fact only the bound  $D_{u_\xi}^2 \mathcal{L}_\xi(v, v) \geq 0$  is shown in [18]; the argument leading from that bound to (9.4) is the same as in the proof of Theorem 3.1 above.
- (2) The theorem actually only necessitates a slightly weakened version of Hypotheses A and C. Indeed, once a  $\xi \in \mathfrak{g}$  is found satisfying the stationary equation (3.2), only the subgroup  $G_\xi$  of  $G$  is still of relevance to its assumptions, its statement and its proof. In particular, it is sufficient to establish persistence of the relative equilibrium on an open subset  $\Omega$  of  $\mathfrak{g}_\xi$ , for the same fixed value of  $\xi$ . Of course, whenever the persistence result of Section 7 applies, this is not a real gain.

- (3) The proof of Theorem 4.1 in [18] uses Theorem 3.1 of that same paper. We point out that the latter necessitates the unstated assumption that  $\xi' \in \Omega \cap \mathfrak{g}_\xi \rightarrow u_{\xi'}$  has an injective derivative at  $\xi$ . In Theorem 9.1 above, this assumption follows from hypothesis (i) and Lemma 4.1. Note also that, if persistence of the relative equilibrium is shown as in Section 7, the assumption follows from the construction (Theorem 7.1).

**9.2. Comparing Theorem 3.1 to Theorem 9.1.** Let us first compare the respective conclusions (3.7) and (9.4) as follows. Writing

$$\mathcal{C}_- = \{u \in E \mid D_{u_\xi}^2 \mathcal{L}(u, u) < 0\} \quad (9.5)$$

for the negative cone of  $D_{u_\xi}^2 \mathcal{L}_\xi$ , we see that they imply that

$$T_{u_\xi} \Sigma_{\mu_\xi} \cap \mathcal{C}_- = \emptyset, \quad \text{respectively} \quad T_{u_\xi} \tilde{\Sigma}_\xi \cap \mathcal{C}_- = \emptyset, \quad (9.6)$$

meaning that  $T_{u_\xi} \Sigma$ , respectively  $T_{u_\xi} \tilde{\Sigma}_\xi$  are positive subspaces of  $E$  for  $D_{u_\xi}^2 \mathcal{L}$ . Since  $T_{u_\xi} \Sigma_{\mu_\xi} \subset T_{u_\xi} \tilde{\Sigma}_\xi$  the second of these statements implies the first and should in general be harder to obtain. Indeed, the cone  $\mathcal{C}_-$  may avoid  $T_{u_\xi} \Sigma_{\mu_\xi}$  but have a non-trivial intersection with  $T_{u_\xi} \tilde{\Sigma}_\xi$ . This is further reflected in the fact that (9.6) implies that

$$\text{codim}(T_{u_\xi} \Sigma_{\mu_\xi}) \geq n(D_{u_\xi}^2 \mathcal{L}_\xi) \quad \text{respectively} \quad \text{codim}(T_{u_\xi} \tilde{\Sigma}_\xi) \geq n(D_{u_\xi}^2 \mathcal{L}_\xi).$$

When  $u_\xi$  is a regular relative equilibrium, one has  $\dim \mathfrak{g} = \text{codim}(T_{u_\xi} \Sigma_{\mu_\xi}) \geq \dim \mathfrak{g}_\xi = \text{codim}(T_{u_\xi} \tilde{\Sigma}_\xi)$ .

To understand how the stronger conclusion comes about, one may note that condition (iv) of Theorem 9.1 has a more limited range of applicability than condition (iv) of Theorem 3.1 since in general

$$p(D_\xi^2 \tilde{W}) \leq p(D_\xi^2 W) \leq n(D_{u_\xi}^2 \mathcal{L}_\xi). \quad (9.7)$$

In particular, condition (iv) of Theorem 9.1 cannot be satisfied when  $p(D_\xi^2 \tilde{W}) < p(D_\xi^2 W)$ . To illustrate this phenomenon, we will give below a simple finite dimensional example where indeed

$$p(D_\xi^2 \tilde{W}) < p(D_\xi^2 W) = n(D_{u_\xi}^2 \mathcal{L}_\xi),$$

so that Theorem 3.1 applies, but Theorem 9.1 does not.

The following corollary further clarifies the link between the two results.

**Corollary 9.3.** *Suppose the hypotheses of Theorem 9.1 are satisfied. Then*

$$T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi} = T_{u_\xi} \mathcal{O}_{u_\xi}$$

so that there exists  $\delta > 0$  such that

$$\forall v \in T_{u_\xi} \tilde{\Sigma}_\xi \cap (T_{u_\xi} \mathcal{O}_{u_\xi})^\perp, \quad D_{u_\xi}^2 \mathcal{L}_\xi(v, v) \geq \delta \|v\|^2. \quad (9.8)$$

Moreover hypotheses (ii), (iii) and (iv) of Theorem 3.1 are satisfied. If, in addition,  $u_\xi$  is a regular relative equilibrium, then  $\mathfrak{g}_{\mu_\xi} = \mathfrak{g}_\xi$ .

We can conclude from the previous discussion and the corollary that, under the non-degeneracy hypothesis  $\text{Ker}(D_\xi^2 W) = \{0\}$ , Theorem 3.1 provides the desired coercivity estimate (3.7) under weaker conditions than Theorem 9.1. As a result, to find a situation where Theorem 9.1 does apply, whereas Theorem 3.1 does not, one has to suppose  $\text{Ker}(D_\xi^2 W) \neq \{0\}$ , whereas  $\text{Ker}(D_\xi^2 \tilde{W}) = \{0\}$ . We did not find an example of such a situation.

*Proof.* It follows from Lemma 4.2 (6) that  $T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi} \subset T_{u_\xi} \mathcal{O}_\xi$ . Also, Lemma 4.2 (4) implies that  $T_{u_\xi} \mathcal{O}_{u_\xi} \subset \text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi \mid T_{u_\xi} \Sigma_{\mu_\xi})$ . So, if  $v \in T_{u_\xi} \mathcal{O}_{u_\xi} \subset T_{u_\xi} \Sigma_{\mu_\xi} \subset T_{u_\xi} \tilde{\Sigma}_\xi$ , then  $D_{u_\xi}^2 \mathcal{L}_\xi(v, v) = 0$ . Writing

$$v = v_\parallel + v_\perp, \quad v_\parallel \in T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi}, \quad v_\perp \in (T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi})^\perp$$

we have,

$$\begin{aligned} 0 &= D_{u_\xi}^2 \mathcal{L}(v, v) = D_{u_\xi}^2 \mathcal{L}_\xi(v_{\parallel}, v_{\parallel}) + D_{u_\xi}^2 \mathcal{L}_\xi(v_{\perp}, v_{\perp}) + 2D_{u_\xi}^2 \mathcal{L}_\xi(v_{\parallel}, v_{\perp}) \\ &= D_{u_\xi}^2 \mathcal{L}_\xi(v_{\perp}, v_{\perp}), \end{aligned}$$

since  $v_{\parallel} \in \text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi)$ . It follows from (9.4) that  $v_{\perp} = 0$  so that  $v \in T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi}$ . We conclude that  $T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi} = T_{u_\xi} \mathcal{O}_{u_\xi}$  and hence (9.8) follows from (9.4). It also follows that hypothesis (ii) of Theorem 3.1 is satisfied. Hypothesis (iii) is the same in both theorems and hypothesis (iv) of Theorem 9.1, together with (9.7), implies hypothesis (iv) of Theorem 3.1.

To prove the last statement, recall from Lemma 4.2 (6) that  $\mathfrak{g}_\xi \subset \mathfrak{g}_{\mu_\xi}$ . Now let  $\eta \in \mathfrak{g}_{\mu_\xi}$ . Since  $T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi} = T_{u_\xi} \mathcal{O}_{u_\xi}$  there exists  $\eta' \in \mathfrak{g}_\xi$  so that  $X_\eta(u_\xi) = X_{\eta'}(u_\xi)$ . Since  $F$  is regular at  $u_\xi$ , this implies  $\eta' = \eta$ , so that  $\eta \in \mathfrak{g}_\xi$ , proving the result.  $\square$

To complete our comparative analysis of those two theorems, we further analyse the conditions on the kernel of  $D_{u_\xi}^2 \mathcal{L}_\xi$  they impose. Similarly to the non-degeneracy condition (i), those conditions are also not in a clear logical relation, in particular because they refer to two a priori different subgroups of  $G$ , namely  $G_\xi$  and  $G_{\mu_\xi}$ . The following lemma sheds further light on the situation.

**Lemma 9.4.** *Suppose Hypotheses A and C are satisfied. Let  $\xi \in \Omega$  and suppose  $\text{Ker}(D_\xi^2 W) = \{0\}$ . Then*

$$\text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) = \text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi \mid \text{Ker} D_{u_\xi} F). \quad (9.9)$$

In addition, the following two statements are equivalent:

- (i)  $\text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) = T_{u_\xi} \tilde{\mathcal{O}}_\xi$ ;
- (ii)  $\text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) = T_{u_\xi} \mathcal{O}_\xi$  and  $\mathfrak{g}_\xi = \mathfrak{g}_{\mu_\xi}$ .

The lemma shows that, if  $u_\xi$  is a non-degenerate (see definition at page 10), and hence regular, relative equilibrium, then the condition on the kernel of  $D_{u_\xi}^2 \mathcal{L}_\xi$  of Theorem 9.1 implies not only the kernel condition in Theorem 3.1, but in addition that  $\mathfrak{g}_\xi = \mathfrak{g}_{\mu_\xi}$ . This statement is independent of the other spectral conditions of these theorems on  $D_{u_\xi}^2 \mathcal{L}_\xi$ .

*Proof.* It follows from  $\text{Ker}(D_\xi^2 W) = \{0\}$ , together with Lemma 4.2 (1) and Lemma 4.1 that  $E = \mathcal{U}_\xi \oplus T_{u_\xi} \Sigma_{\mu_\xi}$ . Hence (4.9) implies that

$$\text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) = \text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi \mid \mathcal{U}_\xi) \oplus \text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi \mid \text{Ker} D_{u_\xi} F) \quad (9.10)$$

On the other hand, it follows from (4.8) that

$$\text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi \mid \mathcal{U}_\xi) = D_\xi \tilde{u}(\text{Ker}(D_\xi^2 W)) = \{0\},$$

so that the first statement of the Lemma follows.

(i)  $\Rightarrow$  (ii) From Lemma 4.2 (4), together with (9.9) and hypothesis (i), we conclude that

$$T_{u_\xi} \mathcal{O}_{u_\xi} \subset \text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi \mid \text{Ker} D_{u_\xi} F) = \text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) = T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi} \subset T_{u_\xi} \Sigma_{\mu_\xi}.$$

On the other hand, as in the proof of Corollary 9.3, we have  $T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi} \subset T_{u_\xi} \mathcal{O}_{u_\xi}$ . Hence  $T_{u_\xi} \tilde{\mathcal{O}}_{u_\xi} = T_{u_\xi} \mathcal{O}_{u_\xi}$  and the first statement of (ii) follows. The second statement is now proven as in the proof of Corollary 9.3.

(ii)  $\Rightarrow$  (i) This is obvious.  $\square$

**9.3. Proving orbital stability.** The above coercivity estimate (9.4) (or, equivalently (9.8)) is used in [18] as an essential input to show the  $G_\xi$ -orbital stability of  $u_\xi$ . Note that this distinguishes their approach from the rest of the literature on orbital stability, including this paper and [10], where instead  $G_{\mu_\xi}$ -stability is proven. The argument given in [18] (and also in [17]) leading from the above coercivity estimate to  $G_\xi$ -orbital stability of  $u_\xi$  is however based on an implicit assumption on  $F$ , referred to as Hypothesis F in [10]. It was explained in [10] how, starting from a coercivity estimate, this condition is used in the cited works to obtain orbital stability for general perturbations of the

relative equilibrium from orbital stability for perturbations within the constraint surface  $\tilde{\Sigma}_\mu$ : see Section 8.3 and Theorem 9 of [10]. However, it was also explained in that last paper that Hypothesis F is typically not satisfied when  $F$  takes values in  $\mathbb{R}^m$ , with  $m > 1$ : it is therefore insufficient to deal with the situations under consideration in this paper as well as in [18]. As recalled in Section 2 above, it is instead possible to use arguments provided in [10] to prove  $G_{\mu_\xi}$ -relative stability, using (3.7) (or, a fortiori (9.8)) as a starting point.

**9.4. An example.** We end this section with a simple but illustrative example in  $E = \mathbb{R}^6$  where Theorem 3.1 applies, but Theorem 9.1 does not. As already pointed out,  $G$  must be non-commutative for this to happen. Consider the  $\text{SO}(3)$ -invariant Hamiltonian

$$H_\alpha(q, p) = H_0(q, p) - \alpha F^2(q, p), \quad \text{with} \quad H_0(q, p) = \frac{\|p\|^2}{2} + V(\|q\|), \quad (9.11)$$

and  $F(q, p) = q \wedge p$ . Note that  $H_0$  is the Hamiltonian of a particle in a central potential  $V$  and that the components of the angular momentum vector  $F$  generate rotations. Since  $H_0$  Poisson commutes with  $F^2$  and since  $F^2$  generates rotations about  $q \wedge p$ , it is easy to see that the circular orbits of  $H_0$  are also flow lines of  $H_\alpha$  and that they are relative equilibria. These are the ones whose orbital stability we shall study. Consider for that purpose the stationary equation, with  $\xi \in \mathbb{R}^3 \simeq \mathfrak{so}(3)$ ,

$$D_{u_\xi} H_\alpha - \xi \cdot D_{u_\xi} F = 0,$$

where  $u_\xi = (q_\xi, p_\xi)$ . A simple computation shows that any solution  $u_\xi$  is of the form  $u_\xi = (q_\xi, p_\xi) = (\rho_\xi \hat{q}_\xi, \sigma_\xi \hat{p}_\xi)$  with

$$\sigma_\xi^2 = \rho_\xi V'(\rho_\xi), \quad \hat{q}_\xi \cdot \hat{p}_\xi = 0.$$

Here  $\rho_\xi > 0, \sigma_\xi > 0$  and we write  $\hat{a} = a/\|a\|$  for each  $a \in \mathbb{R}^3$ . We have

$$\mu_\xi = F(u_\xi) = q_\xi \wedge p_\xi = \rho_\xi \sigma_\xi \hat{q}_\xi \wedge \hat{p}_\xi, \quad \text{with} \quad \xi = \eta_{\alpha, \xi} \mu_\xi, \quad \eta_{\alpha, \xi} = \frac{1 - 2\alpha \rho_\xi^2}{\rho_\xi^2}.$$

Clearly, in this situation  $G_\xi = G_{\mu_\xi} \simeq \text{SO}(2)$ . One has

$$\begin{aligned} \hat{F}(\xi) &= \rho_\xi \sigma_\xi \hat{\xi} = \rho_\xi^{3/2} [V'(\rho_\xi)]^{1/2} \hat{\xi} \\ \|\xi\| &= \rho_\xi \sigma_\xi \frac{|1 - 2\alpha \rho_\xi^2|}{\rho_\xi^2} = \rho_\xi^{3/2} [V'(\rho_\xi)]^{1/2} \frac{|1 - 2\alpha \rho_\xi^2|}{\rho_\xi} = \left[ \rho_\xi^{-1} V'(\rho_\xi) \right]^{1/2} |1 - 2\alpha \rho_\xi^2|. \end{aligned}$$

For illustrative purposes, it is sufficient to consider  $V(\rho_\xi) = \frac{1}{2} \omega \rho_\xi^2$ . Then it is clear that, provided  $2\alpha > \rho_\xi^{-2}$ ,  $\hat{F}$  is a local diffeomorphism and

$$\|\xi\| = \omega^{1/2} (2\alpha \rho_\xi^2 - 1).$$

A simple computation then yields

$$W(\xi) = H_\alpha(u_\xi) - \xi \cdot F(u_\xi) = \alpha \omega \rho_\xi^4 = \frac{1}{4} \frac{\omega}{\alpha} \left( 1 + \frac{\|\xi\|}{\sqrt{\omega}} \right)^2,$$

and furthermore that

$$D_\xi^2 W(v, v) = \frac{1}{2\alpha} \left[ 1 + \frac{\sqrt{\omega}}{\|\xi\|} \right] v^2 - \frac{\sqrt{\omega}}{2\alpha \|\xi\|} (v \cdot \hat{\xi})^2 \geq \frac{v^2}{2\alpha},$$

which is positive definite so that  $p(D_\xi^2 W) = 3$ . Hence,  $n(D^2 \mathcal{L}_\xi) \geq 3$ . A further lengthy but straightforward computation shows that

$$n(D^2 \mathcal{L}_\xi) = 3,$$

and that  $\text{Ker}(D_{u_\xi}^2 \mathcal{L}_\xi) = T_{u_\xi} \mathcal{O}_{u_\xi}$ . Hence Theorem 3.1 implies that the circular orbits are orbitally stable.

On the other hand the assumptions of [18] are too strong to apply in this simple example. Indeed, in [18] the authors consider the Hessian of function  $W$  restricted to  $\mathfrak{g}_\xi$ . The main hypothesis of

their stability theorem is that  $p(D_\xi^2 \tilde{W}) = n(D_{u_\xi}^2 \mathcal{L}_{\alpha, \xi})$ . In the present situation this condition is not satisfied. Indeed, since  $\mathfrak{g}_\xi = \mathfrak{so}(2)$ ,  $p(D_\xi^2 \tilde{W}) \leq 1$ . In fact, it is easy to see it is equal to 1. As a consequence,  $p(D_\xi^2 \tilde{W}) < n(D_{u_\xi}^2 \mathcal{L}_{\alpha, \xi}) = 3$  and so the hypotheses of Theorem 9.1 are not satisfied.

**Conflict of Interest:** The authors declare that they have no conflict of interest.

#### REFERENCES

- [1] R. Abraham and J. E. Marsden. *Foundations of Mechanics*. Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978. Second edition, revised and enlarged, With the assistance of Tudor Rațiu and Richard Cushman.
- [2] G. P. Agrawal. *Nonlinear fiber optics*. Elsevier, 2006.
- [3] J. Angulo Pava. *Nonlinear Dispersive Equations*, volume 156 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2009. Existence and stability of solitary and periodic travelling wave solutions.
- [4] X. Antoine and R. Duboscq. Modeling and computation of Bose-Einstein condensates: stationary states, nucleation, dynamics, stochasticity. In *Nonlinear optical and atomic systems*, volume 2146 of *Lecture Notes in Math.*, pages 49–145. Springer, Cham, 2015.
- [5] J. L. Bona, P. E. Souganidis, and W. A. Strauss. Stability and instability of solitary waves of korteweg-de vries type. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 411(1841):395–412, 1987.
- [6] T. Cazenave. *An introduction to nonlinear Schrödinger equations*, volume 26 of *Textos de Métodos Matemáticos*. Universidade Federal do Rio de Janeiro, 1996.
- [7] T. Cazenave. *Semilinear Schrödinger Equations*. Courant Lecture Notes. American Mathematical Society, 2003.
- [8] T. Cazenave and P.-L. Lions. Orbital stability of standing waves for some nonlinear Schrödinger equations. *Comm. Math. Phys.*, 85:549–561, 1982.
- [9] C. De Angelis, M. Santagiustina, and S. Trillo. Four-photon homoclinic instabilities in nonlinear highly birefringent media. *Phys. Rev. A*, 51:774 – 791, 1995.
- [10] S. De Bièvre, F. Genoud, and S. Rota Nodari. Orbital stability: analysis meets geometry. In *Nonlinear optical and atomic systems*, volume 2146 of *Lecture Notes in Math.*, pages 147–273. Springer, Cham, 2015.
- [11] M. Duflo and M. Vergne. Une propriété de la représentation coadjointe d’une algèbre de Lie. *C.R. Acad. Sc. Paris*, 268(A):583–585, 1969.
- [12] L. Fanelli and E. Montefusco. On the blow-up threshold for weakly coupled nonlinear schrödinger equations. *J. Phys. A: Math. Theor.*, 40:14139, 2007.
- [13] M. G. Forest, D. W. McLaughlin, D. J. Muraki, and O. C. Wright. Nonfocusing instabilities in coupled, integrable nonlinear Schrödinger pdes. *J. Nonlinear Sci.*, 10(3):291–331, 2000.
- [14] F. Genoud. Bifurcation and stability of travelling waves in self-focusing planar waveguides. *Adv. Nonlinear Stud.*, 10:357–400, 2010.
- [15] F. Genoud. Orbitally stable standing waves for the asymptotically linear one-dimensional NLS. *Evol. Equ. Control Theory*, 2:81–100, 2013.
- [16] M. Ghimenti, S. Le Coz, and M. Squassina. On the stability of standing waves of klein-gordon equations in a semiclassical regime. *Discrete and Continuous Dynamical Systems - Series A (DCDS-A)*, pages 2389 – 2401, 2013.
- [17] M. Grillakis, J. Shatah, and W. Strauss. Stability theory of solitary waves in the presence of symmetry. I. *J. Funct. Anal.*, 74(1):160–197, 1987.
- [18] M. Grillakis, J. Shatah, and W. Strauss. Stability theory of solitary waves in the presence of symmetry. II. *J. Funct. Anal.*, 94(2):308–348, 1990.
- [19] I. Ianni and S. Le Coz. Orbital stability of standing waves of semiclassical nonlinear schrödinger-poisson equation. *Adv. Differential Equations*, 14(7/8):717–748, 2009.
- [20] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, second edition, 1980.
- [21] E. Lerman and S. F. Singer. Stability and persistence of relative equilibria at singular values of the moment map. *Nonlinearity*, 11(6):1637–1649, 1998.
- [22] P. Libermann and C.-M. Marle. *Symplectic Geometry and Analytical Mechanics*, volume 35 of *Mathematics and its Applications*. D. Reidel Publishing Co., Dordrecht, 1987. Translated from the French by Bertram Eugene Schwarzbach.
- [23] J. Montaldi. Persistence and stability of relative equilibria. *Nonlinearity*, 10(2):449–466, 1997.
- [24] J. Montaldi and M. Rodríguez-Olmos. On the stability of Hamiltonian relative equilibria with non-trivial isotropy. *Nonlinearity*, 24(10):2777–2783, 2011.



- [25] M. Ohta. Stability of solitary waves for coupled nonlinear Schrödinger equations. *Nonlinear Anal Theory Methods Appl*, 26(5):933–939, 1996.
- [26] M. Onorato, A. R. Osborne, and M. Serio. Modulational instability in crossing sea states: A possible mechanism for the formation of freak waves. *Phys. Rev. Lett.*, 96, 2006.
- [27] J.-P. Ortega and T. S. Ratiu. Stability of Hamiltonian relative equilibria. *Nonlinearity*, 12(3):693–720, 1999.
- [28] G. Patrick. Relative equilibria of hamiltonian systems with symmetry: Linearization, smoothness, and drift. *J. Nonlinear Sci.*, 5:373 – 418, 1995.
- [29] G. W. Patrick. Relative equilibria in Hamiltonian systems: the dynamic interpretation of nonlinear stability on a reduced phase space. *J. Geom. Phys.*, 9(2):111–119, 1992.
- [30] G. W. Patrick, M. Roberts, and C. Wulff. Stability of Poisson equilibria and Hamiltonian relative equilibria by energy methods. *Arch. Ration. Mech. Anal.*, 174(3):301–344, 2004.
- [31] L. Pitaevskii and S. Stringari. *Bose-Einstein Condensation*. Clarendon Press, Oxford, 2003.
- [32] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. I: Functional Analysis*. Academic Press, Inc., 1980.
- [33] M. Roberts, T. Schmah, and C. Stoica. Relative equilibria in systems with configuration space isotropy. *J. Geom. Phys.*, 56(5):762–779, 2006.
- [34] C. A. Stuart. Lectures on the orbital stability of standing waves and application to the nonlinear Schrödinger equation. *Milan J. Math.*, 76:329–399, 2008.
- [35] T. Tao. *Nonlinear Dispersive Equations. Local and Global Analysis*. CBMS Regional Conf. Ser. Math., American Mathematical Society, 2006.
- [36] G. Teschl. *Mathematical Methods in Quantum Mechanics. With Applications to Schrödinger Operators*, volume 99 of *Graduate Studies in Mathematics*. American Mathematical Society, 2009.
- [37] N. Vakhitov and A. A. Kolokolov. Stationary solutions of the wave equation in a medium with nonlinearity saturation. *Radiophys. Quantum Electron.*, 16, 1973.
- [38] C. E. Wayne and M. I. Weinstein. *Dynamics of Partial Differential Equations*, volume 3 of *Frontiers in Applied Dynamical Systems: Reviews and Tutorials*. Springer International Publishing, 2015.
- [39] J. Wei. On the construction of single-peaked solutions to a singularly perturbed semilinear dirichlet problem. *J. Differential Equations*, 129:315–333, 1996.
- [40] M. I. Weinstein. Lyapunov stability of ground states of nonlinear dispersive evolution equations. *Comm. Pure Appl. Math.*, 39(1):51–67, 1986.

<sup>1</sup>LABORATOIRE PAUL PAINLEVÉ, CNRS, UMR 8524 ET UFR DE MATHÉMATIQUES, UNIVERSITÉ DE LILLE, F-59655 VILLENEUVE D’ASCQ CEDEX, FRANCE.

*E-mail address:* [Stephan.De-Bievre@math.univ-lille1.fr](mailto:Stephan.De-Bievre@math.univ-lille1.fr)

<sup>2</sup> EQUIPE-PROJET MEPHYSTO, CENTRE DE RECHERCHE INRIA FUTURS, PARC SCIENTIFIQUE DE LA HAUTE BORNE, 40, AVENUE HALLEY B.P. 70478, F-59658 VILLENEUVE D’ASCQ CEDEX, FRANCE.

<sup>3</sup> INSTITUT DE MATHÉMATIQUES DE BOURGOGNE (IMB), CNRS, UMR 5584, UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ, F-21000 DIJON, FRANCE.

*E-mail address:* [simona.rota-nodari@u-bourgogne.fr](mailto:simona.rota-nodari@u-bourgogne.fr)