Orbital stability: analysis meets geometry

Stephan de Bievre, François Genoud, Simona Rota Nodari

ORBITAL STABILITY: ANALYSIS MEETS GEOMETRY

STEPHAN DE BIÈVRE[1,2], FRANÇOIS GENOUD[3], AND SIMONA ROTA NODARI[1]

ABSTRACT. We present an introduction to the orbital stability of relative equilibria of Hamiltonian dynamical systems on (finite and infinite dimensional) Banach spaces. A convenient formulation of the theory of Hamiltonian dynamics with symmetry and the corresponding momentum maps is proposed that allows us to highlight the interplay between (symplectic) geometry and (functional) analysis in the proofs of orbital stability of relative equilibria via the so-called energy-momentum method. The theory is illustrated with examples from finite dimensional systems, as well as from Hamiltonian PDE's, such as solitons, standing and plane waves for the nonlinear Schrödinger equation, for the wave equation, and for the Manakov system.

CONTENTS

Date: July 22, 2014.

1. INTRODUCTION

The purpose of these notes is to provide an introduction to the theory of orbital stability of relative equilibria, a notion from the theory of (mostly Hamiltonian) dynamical systems with symmetry that finds its origins in the study of planetary motions [AM78]. In more recent times it has proven important in two new ways at least. It has on the one hand found an elegant reformulation in the modern framework of Hamiltonian mechanics of finite dimensional systems with symmetry in terms of symplectic geometry. It can indeed be phrased and studied in terms of the theory of momentum maps and of symplectic reduction [AM78, LM87, Pat92, LS98, Mon97, PRW04, RSS06, MRO11]. On the other hand, it also underlies the stability analysis of plane waves, of travelling wave solutions and of solitons in infinite dimensional nonlinear Hamiltonian PDE's, which has received considerable attention over the last thirty years or so, and continues to be an active area of research [CL82, SS85, Wei86, GSS87, GSS90, SS99, Caz03, HS04, Tao06, GH07a, GH07b, Stu08, GS08, AP09, LC09, Tao09, Gen09, Gen13]. We refer to the first chapter of [AP09] for an interesting historical overview of this aspect of the theory.

It is clear that in this field nonlinear analysis can be expected to meet geometry in interesting and beautiful ways. It nevertheless appears that in the literature on Hamiltonian PDE's, the simple and elegant geometric ideas underlying the proofs of orbital stability aren't emphasized. The goal of these notes is to provide a unified formulation of the theory in a sufficiently general but not too abstract framework that allows one to treat finite and infinite dimensional systems on the same footing. In this manner, one may hope to harness the geometric intuition readily gained from treating finite dimensional systems and use it as a guide when dealing with the infinite dimensional ones that are the main focus of our interest, but that demand more sophisticated technical tools from functional analysis and PDE theory. The text is of an introductory nature and suitable for young researchers wishing to familiarize themselves with the field. It is aimed at analysts not allergic to geometry and at geometers with a taste for analysis, and written in the hope such people exist.

There are many notions of stability for dynamical systems. One may in particular consider stability with respect to perturbations in the vector field generating the dynamics, or stability with respect to a variation in the initial conditions. It is the latter one we shall be considering here. For a sampling of possible definitions in this context, one can consult Section 6.3 of Abraham and Marsden [AM], who give nine different ones and mention there exist others still.... We start by introducing the ones of interest to us in these notes.

The simplest possible one is presumably the following. Let E be a normed vector space, d the corresponding metric on E, and X a vector field on E. Let $u \in E$ and $t \in \mathbb{R} \to u(t) \in E$ a flow line of X (*i.e.* $\dot{u}(t) = X(u(t))$, with $u(0) = u$). Let us assume the flow is well-defined globally, with $u(t) = \Phi_t^X(u)$. Then one says that the initial condition u is stable if for all $\varepsilon > 0$, there exists a $\delta > 0$ so that, for all $v \in E$,

$$\mathrm{d}(v,u) \leq \delta \Rightarrow \sup_{t\in\mathbb{R}} \mathrm{d}(v(t),u(t)) \leq \varepsilon. \tag{1.1}$$

Here $v(t) = \Phi_t^X(v)$. This can be paraphrased as follows: once close, forever not too far. Note that, if u is stable in this sense, then so is $u(t)$ for all $t \in \mathbb{R}$. There exists one situation where proving stability is straightforward. It is the case where $u = u_*$ is a fixed point of the dynamics, meaning $u(t) = u_*$, for all $t \in \mathbb{R}$, and where u_* is a local non-degenerate minimum of a constant of the motion, that is a function $\mathscr{L} : E \to \mathbb{R}$, referred to as a *Lyapunov function*, satisfying $\mathscr{L}(v(t)) = \mathscr{L}(v(0))$ for all $t \in \mathbb{R}$, and for all v in a neighbourhood of

u_*. Let us sketch the argument, which is classic. Supposing $\mathscr{L} \in C^2(E,E)$ and that $D^2_{u_*}\mathscr{L}$ is positive definite, one obtains from a Taylor expansion of $\mathscr{L}$ about u_* , an estimate of the type

$$c\mathrm{d}(v,u_*)^2 \leq \mathscr{L}(v) - \mathscr{L}(u_*) \leq C\mathrm{d}(v,u_*)^2, \tag{1.2}$$

for all v in a neighbourhood of u_*. Then, for v sufficiently close to u_*, one can easily show, using an argument by contradiction, that $v(t)$ stays in this neighbourhood and hence, for all t,

$$c\mathrm{d}(v(t),u_*)^2 \leq \mathscr{L}(v(t)) - \mathscr{L}(u_*) = \mathscr{L}(v) - \mathscr{L}(u_*) \leq C\mathrm{d}(v,u_*)^2, \tag{1.3}$$

from which (1.1) follows immediately. This approach is known as the Lyapunov method for proving stability.[1]

In Hamiltonian systems, at least one constant of the motion always exists, namely the Hamiltonian itself. The above argument leads therefore to the perfectly standard result that local minima of the Hamiltonian are stable fixed points of the dynamics. All orbital stability results that we shall discuss below are, *in fine*, based on this single argument, appropriately applied and combined with additional geometric properties of (Hamiltonian) systems with symmetry, and, of course, with an appropriate dose of (functional) analysis. Let us finally point out that when this approach does not work, and this is very often the case, one is condemned to resort to considerably more sophisticated techniques, involving the KAM theorem or Nekhoroshev estimates, for example.

A stronger version of stability than (1.1) is an asymptotic one, and goes as follows: there exists a $\delta > 0$ so that, for all $v \in E$,

$$\mathrm{d}(v,u) \leq \delta \Rightarrow \lim_{t\to+\infty} \mathrm{d}(v(t),u(t)) = 0.$$

This phenomenon can only occur in dissipative systems. When u is a fixed point of the dynamics, it corresponds to requiring it is attractive. If the flow line issued from u is periodic, one obtains a limit cycle. So in this second definition, the idea is that, if two points start close enough, they end up together. Since our focus here is on Hamiltonian systems, where such behaviour cannot occur (because volumes are preserved), we shall not discuss it further. Note, however, that another notion of "asymptotic stability" has been introduced and studied in the context of Hamiltonian nonlinear dispersive PDE's. This notion encompasses finer asymptotic properties of the dynamics than those studied here. A general heuristic discussion can be found in [Tao09], and recent results for the nonlinear Schrödinger equation in [Cuc11].

There are several cases when definition (1.1) is too strong, and a weaker notion is needed, referred to as *orbital stability*. The simplest definition of this notion goes as follows. Suppose $t \in \mathbb{R} \to u(t) \in E$ is a flow line of the dynamics and consider the dynamical orbit

$$\gamma = \{u(t) \mid t \in \mathbb{R}\}.$$

We say $u = u(0)$ is orbitally stable if the following holds. For all $\varepsilon > 0$, there exists $\delta > 0$, so that

$$\mathrm{d}(v,u) < \delta \Rightarrow \forall t \in \mathbb{R}, \mathrm{d}(v(t),\gamma) \leq \varepsilon. \tag{1.4}$$

The point here is that the new dynamical orbit $\tilde{\gamma} = \{v(t) \mid t \in \mathbb{R}\}$ stays close to the initial one, while possibly $v(t)$ can drift away from $u(t)$, for the same value of the time t. As we will see, this can be expected to be the rule since the nearby orbit may no longer be periodic even if the original one was, or have a different period. A simple example that

[1] Remark that $\mathscr{L}(v(t)) \leq \mathscr{L}(v)$ would suffice in (1.3). But in these notes we will exclusively work with constants of the motion.

can be understood without computation is this. Think of two satellites on circular orbits around the earth. Imagine the radii are very close. Then the periods of both motions will be close but different. Both satellites will eternally move on their respective circles, which are close, but they will find themselves on opposite sides of the earth after a long enough time, due to the difference in their angular speeds. In addition, a slight perturbation in the initial condition of one of the satellites will change its orbit, which will become elliptical, and again have a different period. But the new orbit will stay close to the original circle. So here the idea is this: if an initial condition v is chosen close to u, then at all later times t, $v(t)$ is close to *some* point on γ, but not necessarily close to $u(t)$, for the same value of t. We will treat this illustrative example in detail in Section 4.

The definition of orbital stability in (1.4) turns out to be too strong still for many applications, in particular in the presence of symmetries of the dynamics. This is notably the case in the study of solitons and standing or travelling wave solutions of nonlinear Hamiltonian differential or partial differential equations. We will therefore present an appropriate generalization of the notion of orbital stability in the presence of symmetries in Section 3. For that purpose, we introduce in Section 2 dynamical systems Φ_t^X, $t \in \mathbb{R}$ on Banach spaces E, which admit an invariance group G with an action $\Phi_g, g \in G$ on E, *i.e.* $\Phi_g \Phi_t^X = \Phi_t^X \Phi_g$. We then say $u \in E$ is a *relative equilibrium* if, for all $t \in \mathbb{R}$, $\Phi_t^X(u) \in \mathscr{O}_u$, where $\mathscr{O}_u = \Phi_G(u)$ is the group orbit of u under the action of G. As we will see, solitons, travelling waves and plane waves are relative equilibria. We say a relative equilibrium u is *orbitally stable* if initial conditions $v \in E$ close to u have the property that for all $t \in \mathbb{R}$, $\Phi_t^X(v)$ remains close to $\mathscr{O}_u$. Note that the larger the symmetry group G, the weaker the corresponding notion of stability.

The main goal of these notes is to present a general framework allowing to establish orbital stability of such *relative equilibria* of (both finite and infinite) dynamical systems with symmetry, using an appropriate generalization of the Lyapunov method sketched above. This approach to stability is often referred to as the "energy-momentum" method. In the process, we wish to clearly separate the part of the argument which is abstract and very general, from the part that is model-dependent. We will also indicate for which arguments one needs the dynamics to be Hamiltonian and which ones go through more generally.

In Section 4, we treat the illustrative example of the relative equilibria of the motion in a spherical potential, allowing us to present four variations of the proof of orbital stability, which are later extended to a very general setting in Section 7. The main hypothesis of the proofs, which work for general dynamical systems on Banach spaces, is the existence of a *coercive Lyapunov function* $\mathscr{L}$, which is a group-invariant constant of the motion satisfying an appropriately generalized coercive estimate of the type (1.2) (see (7.1)). In applications, the proof of orbital stability is thus reduced to the construction of such a function.

It is in this step that the geometry of Hamiltonian dynamical systems with symmetry plays a crucial role. Indeed, the construction of an appropriate Lyapunov function for such systems exploits the special link that exists between their constants of the motion F and their symmetries, as embodied in Noether's theorem and the theory of the momentum map. This is explained in Sections 5 and 6. The crucial observation is then that in Hamiltonian systems, relative equilibria tend to come in families $u_\mu \in E$, indexed by the value μ of the constants of the motion at u_μ. In fact, it turns out that $u_\mu \in E$ is a relative equilibrium of a Hamiltonian system if (and only if) u_μ is a critical point of the restriction of the Hamiltonian to the level surface $\Sigma_\mu = \{u \in E \mid F(u) = \mu\}$ of these constants of the motion (Theorem 6.1). This observation at once yields the candidate Lyapunov function $\mathscr{L}_\mu$ (see (6.5)).

We finally explain (Proposition 7.7) how the proof of the coercivity of the Lyapunov function can be obtained from a suitable lower bound on its second derivatives $D^2\mathscr{L}_\mu(w,w)$, with w restricted to an appropriate subspace of E, using familiar arguments from the theory of Lagrange multipliers (Section 7). This ends the very general, geometric and abstract part of the theory. To control $D^2\mathscr{L}_\mu(w,w)$ finally requires an often difficult, problem-dependent, and detailed spectral analysis of the Hessian of the Lyapunov function, as we will show in the remaining sections.

We illustrate the theory in Section 8 on a first simple example. We consider the plane waves $u_{\alpha,k}(t,x) = \alpha e^{-ikx}e^{i\xi t}$, $\xi \in \mathbb{R}$, $k \in 2\pi\mathbb{Z}$ and $\alpha \in \mathbb{R}$, which are solutions of the cubic nonlinear Schrödinger equation on the one-dimensional torus $\mathbb{T}$,

$$i\partial_t u(t,x) + \beta\partial_{xx}^2 u(t,x) + \lambda|u(t,x)|^2 u(t,x) = 0,$$

provided $\xi + \beta k^2 = \lambda|\alpha|^2$. This equation is (globally) well-posed on $E = H^1(\mathbb{T},\mathbb{C})$ and its dynamical flow is invariant under the globally Hamiltonian action Φ of the group $G = \mathbb{R}\times\mathbb{R}$ defined by $\big(\Phi_{a,\gamma}(u)\big)(x) = e^{i\gamma}u(x-a)$ (see Section 5.5). The plane waves $u_{\alpha,k}(t,x)$ are G-relative equilibria. We establish (Theorem 8.1) their orbital stability when $\beta(2\pi)^2 > 2\lambda|\alpha|^2$. Although the linear stability analysis for this model is sketched in many places, and the nonlinear (in)stability results seem to be known to many, we did not find a complete proof of nonlinear orbital stability in the literature. A brief comparison between our analysis and related results ([Zhi01, GH07b, GH07a]) ends Section 8. Note that the analysis of orbital stability of plane waves of the cubic nonlinear Schrödinger equation on a torus of dimension $d > 1$ is much more involved (see for example [FGL13]).

In Section 9 we will present orbital stability results pertaining to curves (*i.e.* one-dimensional families) of standing waves of nonlinear Schrödinger equations on $\mathbb{R}^d$ with a space-dependent coefficient f:

$$i\partial_t u(t,x) + \Delta u(t,x) + f(x,|u|^2(t,x))u(t,x) = 0. \tag{1.5}$$

Imposing a non-trivial spatial dependence has two major consequences. First, the space-translation symmetry of the equation is destroyed, and one is left with the reduced one-parameter symmetry group $G = \mathbb{R}$, acting on the Sobolev space $E = H^1(\mathbb{R}^d)$ via $\Phi_\gamma(u) = e^{i\gamma}u$. Note that the associated group orbits are of the simple form $\mathscr{O}_u = \{e^{i\gamma}u : \gamma \in \mathbb{R}\} \subset H^1(\mathbb{R}^d)$. Now, standing waves are, by definition, solutions of (1.5) of the form $u(x,t) = e^{i\xi t}w(x)$, which are therefore clearly relative equilibria. Such standing waves are sometimes referred to as "solitons" due to the spatial localization of the profile $w(x)$, and to their stability.

Second, constructing curves of standing wave solutions of (1.5) is now a hard problem, and we will outline the bifurcation theory developed in [GS08, Gen09, Gen10a, Gen13] to solve it. This powerful approach allows one to deal with power-type nonlinearities $f(x,|u|^2) = V(x)|u|^{\sigma-1}$ (under an approriate decay assumption on the coefficient $V : \mathbb{R}^d \to \mathbb{R}$) but also with more general nonlinearities, for instance the asymptotically linear $f(x,|u|^2) = V(x)\frac{|u|^{\sigma-1}}{1+|u|^{\sigma-1}}$. This will give a good illustration of how involved the detailed analysis of $D^2\mathscr{L}(w,w)$ required by the model can be. As we shall see, this analysis turns out to be deeply connected with the bifurcation behaviour of the standing waves.

In the pure power (space-independent) case $f(x,|u|^2) = |u|^{\sigma-1}$, the appropriate notion of stability is that associated with the action of the full group $G = \mathbb{R}^d\times\mathbb{R}$, $\big(\Phi_{a,\gamma}(u)\big)(x) = e^{i\gamma}u(x-a)$. The stability of standing waves in this context was proved in the seminal paper of Cazenave and Lions [CL82] for $1 < \sigma < 1 + \frac{4}{d}$, and this result is sharp (*i.e.* stability does not hold at $\sigma = 1 + \frac{4}{d}$). The contribution [CL82] is one of the first rigorous results on

orbital stability for nonlinear dispersive equations, and is based on variational arguments using the concentration-compactness principle (see for instance [Zhi01, HS04] for more recent results in this direction). This line of argument is conceptually very different from the energy-momentum approach developed here, so we shall not say more about it.

The modern treatment of Hamiltonian dynamical systems with symmetries uses the language of symplectic geometry, as for example in [AM78, Arn99, LM87, Sou97]. But we don't need the full power of this theory, since we will work exclusively with linear symplectic structures on (infinite dimensional) symplectic vector spaces. For the reader not familiar with Hamiltonian mechanics, Lie group theory and symplectic group actions, elementary self-contained introductions to these subjects sufficient for our purposes are provided in the Appendix.

Acknowledgments. This work was supported in part by the Labex CEMPI (ANR-11-LABX-0007-01). F.G. thanks CEMPI and the Lab. Paul Painlevé for their hospitality during his one-month visit to the Université Lille 1 in September 2013. He also acknowledges the support of the ERC Advanced Grant "Nonlinear studies of water flows with vorticity". The authors are grateful to V. Combet, A. De Laire, S. Keraani, G. Rivière, B. Tumpach and G. Tuynman for stimulating discussions on the subject matter of these notes.

2. Dynamical systems, symmetries and relative equilibria

2.1. **Dynamical systems on Banach spaces.**

Let E be a Banach space. A domain $\mathscr{D}$ is a dense subset of E; in the examples presented in these notes, it will be a dense linear subspace of E. We consider on E a *dynamical system*, by which we mean a separately continuous map

$$\Phi^X : (t,u) \in \mathbb{R} \times E \to \Phi_t^X(u) := \Phi^X(t,u) \in E, \tag{2.1}$$

with the following properties:

(i) For all $t,s \in \mathbb{R}$,

$$\Phi_t^X \circ \Phi_s^X = \Phi_{t+s}^X, \quad \Phi_0^X(u) = \mathrm{Id}_E. \tag{2.2}$$

(ii) For all $t \in \mathbb{R}$, $\Phi_t^X(\mathscr{D}) = \mathscr{D}$.

(iii) $X : \mathscr{D} \subset E \to E$ is a vector field that generates the dynamics in the sense that, when $u \in \mathscr{D}$, $\Phi_t^X(u) := u(t) \in \mathscr{D}$ is a solution of the differential equation

$$\dot{u}(t) = X(u(t)), \quad u(0) = u. \tag{2.3}$$

By this we mean that the curve $t \in \mathbb{R} \to u(t) \in E$ is differentiable as a map from $\mathbb{R}$ to E. In infinite dimensional problems, the vector fields are often only defined on a domain $\mathscr{D}$, where they may not even be continuous. But note that we always assume that the flows themselves are defined on all of E (or on an open subset of E). For examples illustrating these subtleties, see Section 2.4. Local flows can be defined analogously. In that case the domains are dense in some open subset of E, but we shall not deal with such situations in these notes since we will always assume the flows to be globally defined.

Suppose there exists a function $F : E \to \mathbb{R}^m$ so that

$$F \circ \Phi_t^X = F, \quad \forall t \in \mathbb{R}. \tag{2.4}$$

We then say that the vector field X or its associated flow Φ_t^X admits *m constants of the motion*, which are the components F_i of F. In that case, one may consider the restriction

of the flow Φ_t^X to the level sets of F: for $\mu \in \mathbb{R}^m$, we define

$$\Sigma_\mu = \{u \in E \mid F(u) = \mu\}, \tag{2.5}$$

and one has that $\Phi_t^X \Sigma_\mu = \Sigma_\mu$, for all $\mu \in \mathbb{R}^m$.

2.2. **Symmetries, reduced dynamics and relative equilibria**

We now define the notion of an invariance group for a dynamical system. For that purpose, we need to say a few words about group actions. Let G be a topological group acting on E. By this we mean there exists a separately continuous map

$$\Phi : (g,u) \in G \times E \to \Phi_g(u) \in E,$$

satisfying $\Phi_e = \mathrm{Id}$, $\Phi_{g_1 g_2} = \Phi_{g_1} \circ \Phi_{g_2}$. We will call

$$\mathcal{O}_u = \{\Phi_g(u) \mid g \in G\} \tag{2.6}$$

the orbit of G through $u \in E$. For later reference, we define the isotropy group of u, G_u, as follows

$$G_u = \{g \in G \mid \Phi_g(u) = u\}. \tag{2.7}$$

We can then introduce the notion of an invariance group for Φ_t^X.

Definition 2.1. We say G is an invariance group (or symmetry group) for the dynamical system Φ_t^X if, for all $g \in G$, and for all $t \in \mathbb{R}$,

$$\Phi_g \circ \Phi_t^X = \Phi_t^X \circ \Phi_g. \tag{2.8}$$

Remark that $G = \mathbb{R}$ is always an invariance group of the dynamical system, with action Φ_t^X on E. While this is correct, this is not of any particular use, as one can suspect from the start. Indeed, the flow Φ_t^X is in applications obtained by integrating a nonlinear differential or partial differential equation, and is not explicitly known. In fact, it is the object of study. "Useful" symmetries are those that help to simplify this study; they need to have a simple and explicit action on E. They are often of a clearcut geometric origin: translations, rotations, gauge transformations, *etc.* Several examples are provided in the following sections.

Finally, it should be noted we did not define "the" symmetry group for Φ_t^X, but "a" symmetry group. Depending on the problem at hand and the questions addressed, different symmetry groups may prove useful for the same dynamical system, as we shall also illustrate. In particular, any subgroup of an invariance group is also an invariance group, trivially.

It follows immediately from (2.6) and (2.8) that, for all $x \in E$,

$$\Phi_t^X \mathcal{O}_u = \mathcal{O}_{\Phi_t^X(u)}. \tag{2.9}$$

In other words, if G is an invariance group, then the dynamical system maps G-orbits into G-orbits. This observation lies at the origin of the following construction which is crucial for the definitions of relative equilibrium and orbital stability that we shall introduce. We give the general definitions here, and refer to the coming sections for examples. Defining an equivalence relation on E through

$$u \sim u' \Leftrightarrow \mathcal{O}_u = \mathcal{O}_{u'},$$

we consider the corresponding quotient space that we denote by $E_G = E/\sim$ and that we refer to as the reduced phase space. We will occasionally use the notation

$$\pi : u \in E \to \mathcal{O}_u \in E_G \tag{2.10}$$

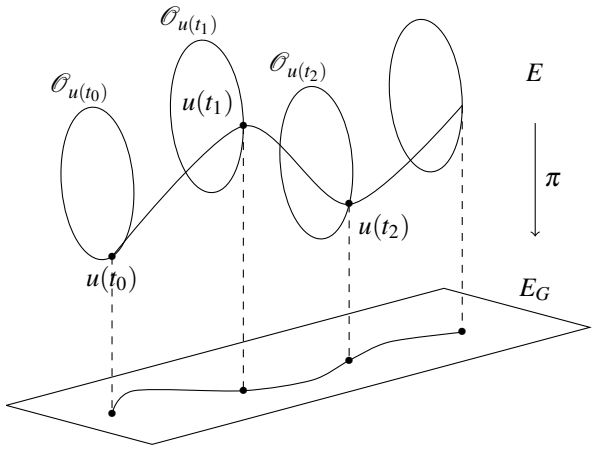

FIGURE 1. A dynamical orbit $t \to u(t)$ and its "attached" G-orbits, with the projection into E_G.

for the associated projection. So the elements of E_G are just the G-orbits in E. It is then clear from (2.9) that the dynamical system Φ_t^X on E naturally induces reduced dynamics on the orbit space E_G: it "passes to the quotient" in the usual jargon. We will use the same notation for these reduced dynamics and write $\Phi_t^X \mathcal{O} = \mathcal{O}(t)$ for any $\mathcal{O} \in E_G$. Note that $\Phi_t^X \mathcal{O}_u = \mathcal{O}_{u(t)}$ (See Fig. 1).

As a general rule of thumb, one may hope that the reduced dynamics are simpler than the original one, since they take place on a lower dimensional (or in some sense smaller) quotient space. This idea can sometimes provide a useful guideline, notably in the study of stability properties of fixed points or periodic orbits of the original dynamical system, as will be illustrated in the coming sections. Implementing it concretely can nevertheless be complicated, in particular because the quotient itself may be an unpleasant object to do analysis on, even in finite dimensions, as its topology or differential structure may be pathological and difficult to deal with. Conditions on G and on the action Φ are needed, for example, to ensure the quotient topology on E_G is Hausdorff, or that it has a differentiable structure [AM78, LM87, PRW04]. In addition, concrete computations on models are more readily done on E directly, than in the abstract quotient space, particularly in infinite dimensional problems. We will avoid these difficulties, in particular because we will work almost exclusively with *isometric* group actions. Their orbits have simplifying features that we will repeatedly use: see Proposition 2.3 below.

We are now in a position to introduce the notion of *relative equilibrium*, as follows.

Definition 2.2. Let $u \in E$. Let Φ_t^X be a dynamical system on E and let G be a symmetry group for E. We say u is a *G-relative equilibrium*[2] for Φ_t^X if, for all $t \in \mathbb{R}$, $u(t) \in \mathcal{O}_u$. Or, equivalently, if for all $t \in \mathbb{R}$, $\Phi_t^X \mathcal{O}_u = \mathcal{O}_u$. When there is no ambiguity about the dynamical system Φ_t^X and the group G considered, we will simply say u is a *relative equilibrium*.

With the language introduced, u is a relative equilibrium if $\mathcal{O}_u$ is a fixed point of the reduced dynamics on E_G. Again, we refer to the following sections for examples. We are interested in these notes in the stability of such relative equilibria. Roughly speaking, we

[2]In [LM87], the term *stationary motion* is used for this concept.

will say a relative equilibrium is orbitally stable if it is stable as a fixed point of the reduced dynamics; we give a precise definition in Section 3.

We end this section with two comments. First, the above terminology comes from the literature on Hamiltonian dynamical systems in finite dimensions. We will see in the following sections what the many specificities are of that situation. We refer to [Arn99, AM78, LM87] for textbook treatments and historical background and to [Pat92, Mon97, LS98, PRW04, RSS06, MRO11] for more recent developments. Second, we will often need to deal with the restriction of the dynamical systems under consideration to the level sets $\Sigma_\mu \subset E$ of a family of constants of the motion F, as defined in (2.5). Note that Σ_μ is a metric space. We define

$$G_{\Sigma_\mu} = \{g \in G \mid \forall u \in \Sigma_\mu, \Phi_g(u) \in \Sigma_\mu\}. \tag{2.11}$$

This is clearly a subgroup of G, which is a symmetry group of the dynamical system restricted to Σ_μ. We will often deal with isometric group actions on such Σ_μ, or on the full Banach space E. The following simple proposition collects some of the essential properties of their orbits that we shall repeatedly need and use. We first recall the definition of the Hausdorff metric. Let (Σ, d) be a metric space and let $S, S' \subset (\Sigma, \mathrm{d})$. Then

$$\Delta(S,S') = \max\{\sup_{u\in S} \mathrm{d}(u,S'), \sup_{u'\in S'} \mathrm{d}(S,u')\}. \tag{2.12}$$

Notice that this is only a pseudometric[3] and that $\Delta(S,S') = +\infty$ is possible.

Proposition 2.3. *Let G be a group, (Σ, d) a metric space and $\Phi : G \times \Sigma \to \Sigma$ an action of G on Σ. Suppose that for each $g \in G$, Φ_g is an isometry: $\forall u, u' \in \Sigma, \mathrm{d}(\Phi_g(u), \Phi_g(u')) = \mathrm{d}(u,u')$. Let $\mathscr{O}, \mathscr{O}'$ be two G-orbits in Σ. Then*

(i) $\forall u_1, u_2 \in \mathscr{O}, \forall u_1', u_2' \in \mathscr{O}'$, $\mathrm{d}(u_1, \mathscr{O}') = \mathrm{d}(u_2, \mathscr{O}')$, $\quad \mathrm{d}(u_1', \mathscr{O}) = \mathrm{d}(u_2', \mathscr{O})$,
(ii) $\forall u \in \mathscr{O}, u' \in \mathscr{O}'$, $\quad \mathrm{d}(u, \mathscr{O}') = \Delta(\mathscr{O}, \mathscr{O}') = \mathrm{d}(u', \mathscr{O})$,
(iii) $\forall u \in \mathscr{O}, u' \in \mathscr{O}'$, $\quad \Delta(\mathscr{O}, \mathscr{O}') \leq \mathrm{d}(u,u')$.

Proof. The first statement follows from the existence of $g \in G$ so that $\Phi_g(u_1) = u_2$. For the second, we proceed by contradiction. Suppose first that, $\forall u \in \mathscr{O}, u' \in \mathscr{O}'$, $\mathrm{d}(u, \mathscr{O}') < \mathrm{d}(u', \mathscr{O})$. Let $u \in \mathscr{O}, u' \in \mathscr{O}'$. Then we know there exists $v \in \mathscr{O}'$ (depending on u, u') so that $\mathrm{d}(u, \mathscr{O}') \leq \mathrm{d}(u,v) < \mathrm{d}(u', \mathscr{O})$. But since, by the first part of the proposition, $\mathrm{d}(v, \mathscr{O}) = \mathrm{d}(u', \mathscr{O})$, this implies $\mathrm{d}(u,v) < \mathrm{d}(v, \mathscr{O})$, which is a contradiction. So we conclude, using the first part again, that $\forall u \in \mathscr{O}, u' \in \mathscr{O}'$, $\mathrm{d}(u, \mathscr{O}') \geq \mathrm{d}(u', \mathscr{O})$. Repeating the argument with the roles of $\mathscr{O}, \mathscr{O}'$ inverted, the result follows. □

If the action is not isometric, it is quite possible for all the statements of the theorem to fail. For example, consider on $E = \mathbb{R}^2$ the action $\Phi_a(q,p) = (\exp(a)q, \exp(-a)p)$, $a \in \mathbb{R}$.

2.3. **Motion in a spherical potential**

In this section, we illustrate the preceding notions on a simple Hamiltonian mechanical system: a particle in a spherical potential. We will make free use of the concepts and notation of Appendices A.2 and A.3 that we invite the reader unfamiliar with Hamiltonian mechanics or Lie group theory to peruse.

By a spherical potential we mean a function $V : \mathbb{R}^3 \to \mathbb{R}$, satisfying $V(Rq) = V(q)$, for all $R \in \mathrm{SO}(3)$. With a slight abuse of notation, we write $V(q) = V(\|q\|)$, for a smooth function $V : \mathbb{R}^+ \to \mathbb{R}$. We consider on $E = \mathbb{R}^6$ the Hamiltonian

$$H(u) = H(q,p) = \frac{1}{2}p^2 + V(\|q\|) \tag{2.13}$$

[3] $\Delta(S,S') = 0$ does not imply $S = S'$. In particular, $\Delta(S, \overline{S}) = 0$.

and the corresponding Hamiltonian equations of motion

$$\dot{q} = p, \quad \dot{p} = -V'(\|q\|)\hat{q}, \tag{2.14}$$

where we introduce the notation $\hat{b} = \frac{b}{\|b\|}$ for any $b \in \mathbb{R}^3$. Integrating those, we obtain the Hamiltonian flow $\Phi_t^H(u) = u(t)$, where $u = (q,p) \in \mathbb{R}^6$. Introducing the angular momentum

$$L(q,p) = q \wedge p, \tag{2.15}$$

one checks immediately that, for any solution $t \in \mathbb{R} \to (q(t),p(t)) \in \mathbb{R}^6$, one has

$$\frac{\mathrm{d}}{\mathrm{d}t} L(q(t),p(t)) = 0. \tag{2.16}$$

In other words, angular momentum is conserved during the motion in a central potential: its three components are constants of the motion. This implies the familiar result that the motion takes place in the plane perpendicular to L and passing through 0.

We will now use Noether's Theorem (Theorem A.3.9) to show this system is SO(3)-invariant. We start with the following observations. First, the action of the group G =SO(3) on $E = \mathbb{R}^6$ given by

$$\Phi_R(u) = (Rq, Rp) \tag{2.17}$$

is easily checked to be globally Hamiltonian[4]. Indeed, for each $\xi \in \mathrm{so}(3)$,

$$\Phi_{\exp(t\xi)} = \Phi_t^{F_\xi},$$

where

$$F_\xi(q,p) = \xi \cdot L(q,p)$$

(recall that we can identify so(3) with $\mathbb{R}^3$ via (A.2.6)). In other words,"angular momentum generates rotations." Next, it is clear that the Hamiltonian satisfies $H \circ \Phi_R = H$. As a result, it follows from Theorem A.3.9 (iii) that the dynamical flow is rotationally invariant:

$$\Phi_t^H \circ \Phi_R = \Phi_R \circ \Phi_t^H, \qquad \forall t \in \mathbb{R},\ R \in \mathrm{SO}(3).$$

Note that, here and in what follows, we are using, apart from the symplectic, also the standard euclidean structure on $\mathbb{R}^6$.

We now wish to identify the relative equilibria of these systems. For that purpose, consider first $u \in \mathbb{R}^6$ with $L(u) = \mu \neq 0$. Then the ensuing dynamical trajectory $u(t)$ lies in the surface

$$\Sigma_\mu = \{u \in \mathbb{R}^6 \mid L(u) = \mu\}. \tag{2.18}$$

Now, if u is a relative equilibrium, then, for each t, there exists $R(t) \in \mathrm{SO}(3)$ so that $\Phi_{R(t)}u = u(t)$. Hence $\mu = L(u(t)) = L(\Phi_{R(t)}u) = R(t)L(u) = R(t)\mu$. In other words, $R(t)$ belongs to

$$G_\mu = \{R \in \mathrm{SO}(3) \mid R\mu = \mu\} \simeq \mathrm{SO}(2),$$

which is the subgroup of rotations about the μ-axis. It follows that $\|q(t)\| = \|q\|$, for all t. Since $q(t)$ is perpendicular to μ, this means that $q(t)$ lies on the circle of radius $\|q\|$ centered at 0 and perpendicular to μ. The orbit is therefore circular and, in particular, for all t, $q(t) \cdot p(t) = 0$. Conversely, it is clear that all circular dynamical orbits are relative equilibria. The initial conditions corresponding to such circular orbits are easily seen to be of the form

$$q = \rho_* \hat{q}, \quad p = \sigma_* \hat{p}, \quad \sigma_*^2 = \rho_* V'(\rho_*), \quad \hat{q} \cdot \hat{p} = 0, \tag{2.19}$$

with $\rho_*, \sigma_* > 0$ and hence $V'(\rho_*) > 0$. We will discuss in Section 4 under what conditions they are orbitally stable in the sense of (1.4).

[4]See Definition A.3.7.

Now, let $u=(q,p)\in\mathbb{R}^6$ be such that $L(u)=0$. In this case q and p are parallel and this remains true at all times. But if $p(t)\neq 0$ at any time t, u cannot be a relative equilibrium. Indeed, the motion is then along a straight line passing through the origin and such a straight line cannot lie in an $\mathrm{SO}(3)$ orbit since the $\mathrm{SO}(3)$ action preserves norms. If on the other hand $u=(\rho_*\hat{q},0)=u(t)$ is a fixed point of the dynamics, it is a fortiori a relative equilibrium. This occurs if and only if $V'(\rho_*)=0$ as is clear from the equations of motion. Note that these fixed points fill the sphere of radius ρ_*.

It is clear these fixed points cannot be stable in the sense of definition (1.1) or (1.4). Indeed, any initial condition u' close to such fixed point u, but with $p'\neq 0$ gives rise to a trajectory in the plane spanned by q' and p': when q' and p' are not parallel, the trajectory will wind around the origin in this plane, moving away from the initial condition. What we will prove in Section 4 is that, provided $V''(\rho_*)>0$, these trajectories all stay close to

$$\mathscr{O}_{\rho_*,0,0}=\{u\in\mathbb{R}^6\mid q\cdot q=\rho_*^2,\ p\cdot p=0,\ q\cdot p=0\},\tag{2.20}$$

which is the $\mathrm{SO}(3)$ orbit through the fixed point $u=(\rho_*\hat{q},0)$. Those fixed points are therefore $\mathrm{SO}(3)$-orbitally stable, in the sense of Definition 3.1 (i) below.

To end this section, we list, for later purposes, all $\mathrm{SO}(3)$-orbits in $E=\mathbb{R}^6$. Those are easily seen to be the hypersurfaces $\mathscr{O}_{\rho,\sigma,\alpha}$ of the form

$$\mathscr{O}_{\rho,\sigma,\alpha}=\{(q,p)\in\mathbb{R}^6\mid q\cdot q=\rho^2,\ p\cdot p=\sigma^2,\ q\cdot p=\alpha\},\tag{2.21}$$

with $\rho,\sigma\geq 0,\alpha\in\mathbb{R}$. Note that $|\alpha|\leq\rho\sigma$. Those orbits are three-dimensional smooth submanifolds of $\mathbb{R}^6$, except on the set where the angular momentum L vanishes, *i.e.* on

$$\Sigma_0=\{(q,p)\in\mathbb{R}^6\mid L(q,p)=0\}.$$

This surface (which is not a submanifold of E) is itself $\mathrm{SO}(3)$-invariant and foliated by group orbits as follows:

$$\Sigma_0=\bigcup_{\rho\sigma=|\alpha|}\mathscr{O}_{\rho,\sigma,\alpha}=\{(0,0)\}\cup\bigcup_{\substack{\rho\sigma=|\alpha|\\(\rho,\sigma)\neq(0,0)}}\mathscr{O}_{\rho,\sigma,\alpha}.$$

On the latter orbits, q and p are parallel, but do not both vanish, so that these orbits can be identified with two-dimensional spheres.

2.4. **The nonlinear Schrödinger equation**

An important example of an infinite dimensional dynamical system is the nonlinear Schrödinger equation

$$\begin{cases} i\partial_t u(t,x)+\Delta u(t,x)+f(x,u(t,x))=0,\\ u(0,x)=u_0(x),\end{cases}\tag{2.22}$$

with $u(t,x):\mathbb{R}\times\mathbb{R}^d\to\mathbb{C}$. Here Δ denotes the usual Laplace operator and f is a local nonlinearity. More precisely, consider $f:(x,u)\in\mathbb{R}^d\times\mathbb{R}^+\to f(x,u)\in\mathbb{R}$ such that f is measurable in x and continuous in u. Assume that

$$f(x,0)=0 \text{ a.e. in } \mathbb{R}^d\tag{2.23}$$

and that for every $K>0$ there exists $L(K)<+\infty$ such that

$$|f(x,u)-f(x,v)|\leq L(K)|u-v|\tag{2.24}$$

a.e. in $\mathbb{R}^d$ and for all $0 \le u, v \le K$. Assume further that

$$\begin{cases} L(\cdot) \in C^0([0,+\infty)) & \text{if } d = 1, \\ L(K) \le C(1+K^\alpha) \text{ with } 0 \le \alpha < \dfrac{4}{d-2} & \text{if } d \ge 2, \end{cases} \tag{2.25}$$

and extend f to the complex plane by setting

$$f(x,u) = \frac{u}{|u|} f(x,|u|), \tag{2.26}$$

for all $u \in \mathbb{C}$, $u \neq 0$.

Finally, let H be the Hamiltonian of the system defined by

$$H(u) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla u|^2(x)\,\mathrm{d}x - \int_{\mathbb{R}^d}\int_0^{|u|(x)} f(x,s)\,\mathrm{d}s\,\mathrm{d}x. \tag{2.27}$$

We now explain how the Schrödinger equation defines an infinite dimensional dynamical system with symmetries, within the framework of Sections 2.1 and 2.2. The sense in which the Schrödinger equation defines a Hamiltonian dynamical system will be explained in Section 5.

For that purpose, we need the following results on local and global existence of solutions to (2.22). First, concerning local existence, we have:

Theorem 2.4 ([Caz03])**.** *If f is as above, then for every $u_0 \in H^1(\mathbb{R}^d,\mathbb{C})$ there exist numbers $T_{\min}, T_{\max} > 0$ and a unique maximal solution $u : t \in (-T_{\min}, T_{\max}) \to u(t) \in H^1(\mathbb{R}^d,\mathbb{C})$ of* (2.22) *satisfying*

$$u \in C^0((-T_{\min}, T_{\max}), H^1(\mathbb{R}^d)) \cap C^1((-T_{\min}, T_{\max}), H^{-1}(\mathbb{R}^d)).$$

Moreover, u depends continuously on u_0 in the following sense: if $u_0^k \to u_0$ in $H^1(\mathbb{R}^d,\mathbb{C})$ and if u_k is the maximal solution of (2.22) *with the initial value u_0^k, then $u_k \to u$ in $C^0([-S,T], H^1(\mathbb{R}^d))$ for every interval $[-S,T] \subset (-T_{\min}, T_{\max})$. In addition, there is conservation of charge and energy, that is*

$$\|u(t)\|_{L^2} = \|u_0\|_{L^2}, \quad H(u(t)) = H(u_0) \tag{2.28}$$

for all $t \in (-T_{\min}, T_{\max})$.

For global existence of solutions, one needs a growth condition on f in its second variable.

Theorem 2.5 ([Caz03])**.** *Let f be as in Theorem 2.4. Suppose in addition that there exist $A \ge 0$ and $0 \le \nu < \frac{4}{d}$ such that*

$$\int_0^{|u|} f(x,s)\,\mathrm{d}s \le A|u|^2(1+|u|^\nu), \quad x \in \mathbb{R}^d,\ u \in \mathbb{C}. \tag{2.29}$$

It follows that for every $u_0 \in H^1(\mathbb{R}^d,\mathbb{C})$, the maximal strong H^1-solution u of (2.22) *given by Theorem 2.4 is global and $\sup_{t\in\mathbb{R}} \|u(t)\|_{H^1} < +\infty$.*

Note that the condition on f is always satisfied when f is negative. This result implies that one can define Φ_t^X on $E = H^1(\mathbb{R}^d,\mathbb{C})$ by $\Phi_t^X(u) = u(t) \in E$ and that Φ_t^X satisfies (2.1)–(2.2). Note however that, whereas the flow lines $t \to u(t) \in E$ are guaranteed to be continuous by the above theorems, they are C^1 only when viewed as taking values in $E^* = H^{-1}(\mathbb{R}^d,\mathbb{C})$. The following "propagation of regularity" theorem allows one to identify the appropriate domain $\mathscr{D}$ on which the stronger condition (2.3) holds.

Theorem 2.6 ([Caz03]). *Let f be as in Theorem 2.4, and consider $u_0 \in H^1(\mathbb{R}^d,\mathbb{C})$ and $u \in C^0((-T_{\min},T_{\max}),H^1(\mathbb{R}^d))$ the solution of the problem* (2.22) *given by Theorem 2.4. Then the following statements hold.*

(i) If $u_0 \in H^2(\mathbb{R}^d,\mathbb{C})$, then $u \in C^0((-T_{\min},T_{\max}),H^2(\mathbb{R}^d))$. If, in addition, $f(x,\cdot) \in C^1(\mathbb{C},\mathbb{C})$, then u depends continuously on u_0 in the following sense: if $u_0^k \to u_0$ in $H^2(\mathbb{R}^d,\mathbb{C})$ and if u_k is the maximal solution of (2.22) *with the initial value u_0^k, then $u_k \to u$ in $C^0([-S,T],H^2(\mathbb{R}^d))$ for every interval $[-S,T] \subset (-T_{\min},T_{\max})$.*

(ii) If $u_0 \in H^m(\mathbb{R}^d,\mathbb{C})$ for some integer $m > \max\left\{\frac{d}{2},2\right\}$ and if $f(x,\cdot) \in C^m(\mathbb{C},\mathbb{C})$, then $u \in C^0((-T_{\min},T_{\max}),H^m(\mathbb{R}^d))$. In addition, u depends continuously on u_0 in the following sense: if $u_0^k \to u_0$ in $H^m(\mathbb{R}^d,\mathbb{C})$ and if u_k is the maximal solution of (2.22) *with the initial value u_0^k, then $u_k \to u$ in $L^\infty([-S,T],H^m(\mathbb{R}^d))$ for every interval $[-S,T] \subset (-T_{\min},T_{\max})$.*

Note that the derivatives of f should be understood in the real sense here.

Remark 2.7. It follows from Theorem 2.6 that, if we take $\mathscr{D} = H^m(\mathbb{R}^d,\mathbb{C})$, with $m \geq 3$, then (2.3) is satisfied, and so the flow is differentiable as a map from $\mathbb{R}$ to $E = H^1(\mathbb{R}^d,\mathbb{C})$.

Example 2.8. A typical example of local nonlinearity which satisfies (2.23), (2.24), (2.25) and (2.26) is the pure power nonlinearity

$$f(u) = \lambda |u|^{\sigma-1}u \tag{2.30}$$

with

$$\begin{aligned} &1 \leq \sigma < +\infty && \text{for } d = 1, \\ &1 \leq \sigma < 1 + \frac{4}{d-2} && \text{for } d \geq 2, \end{aligned} \tag{2.31}$$

and $\lambda \in \mathbb{R}$. The standard "cubic" Schrödinger equation corresponds to $\sigma = 3$, which is an allowed value of σ only if $1 \leq d \leq 3$. The Hamiltonian is then given by

$$H(u) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla u|^2(x)\,\mathrm{d}x - \frac{\lambda}{\sigma+1}\int_{\mathbb{R}^d} |u|^{\sigma+1}(x)\,\mathrm{d}x. \tag{2.32}$$

In this case, the nonlinear Schrödinger equation reads

$$\begin{cases} i\partial_t u(t,x) + \Delta u(t,x) + \lambda |u|^{\sigma-1}(t,x)u(t,x) = 0, \\ u(0,x) = u_0(x). \end{cases} \tag{2.33}$$

Theorem 2.4 then ensures the existence of a local solution

$$u \in C^0((-T_{\min},T_{\max}),H^1(\mathbb{R}^d)) \cap C^1((-T_{\min},T_{\max}),H^{-1}(\mathbb{R}^d)) \tag{2.34}$$

and the conservation of the Hamiltonian energy H. To guarantee the existence of a global flow, we have to distinguish the focusing ($\lambda > 0$) and the defocusing case ($\lambda < 0$). More precisely, Theorem 2.5 implies the flow is globally defined on $H^1(\mathbb{R}^d,\mathbb{C})$, i.e.

$$\Phi^X : \mathbb{R} \times H^1(\mathbb{R}^d,\mathbb{C}) \to H^1(\mathbb{R}^d,\mathbb{C}), \tag{2.35}$$

if σ satisfies (2.31) in the defocusing case or if $1 \leq \sigma < 1 + \frac{4}{d}$ in the focusing case. Note that, in the latter situation, $\sigma = 3$ is allowed only if $d = 1$.

Next, we recall that

$$\begin{aligned} &\sigma \in \mathbb{N},\ \sigma \text{ odd } \Rightarrow f \in C^\infty(\mathbb{C},\mathbb{C}), \\ &\sigma \in \mathbb{N},\ \sigma \text{ even } \Rightarrow (f \in C^m(\mathbb{C},\mathbb{C}) \Leftrightarrow m \leq \sigma - 1), \\ &\sigma \notin \mathbb{N} \Rightarrow (f \in C^m(\mathbb{C},\mathbb{C}) \Leftrightarrow m \leq [\sigma - 1] + 1), \end{aligned}$$

and, in particular, $f \in C^1(\mathbb{C},\mathbb{C})$ for all $\sigma \geq 1$. Hence Theorem 2.6 applies and the flow can be restricted to $H^2(\mathbb{R}^d,\mathbb{C})$

$$\Phi^X : \mathbb{R} \times H^2(\mathbb{R}^d,\mathbb{C}) \to H^2(\mathbb{R}^d,\mathbb{C}),$$

whenever σ satisfies (2.31) in the defocusing case or $1 \leq \sigma < 1 + \frac{4}{d}$ in the focusing case. This, however, is not enough for our purposes, since it only guarantees the existence of the derivative of $t \to u(t)$ as a function in $L^2(\mathbb{R}^d,\mathbb{C})$, and not as a function in $E = H^1(\mathbb{R}^d,\mathbb{C})$. In other words, we cannot take $\mathscr{D} = H^2(\mathbb{R}^d,\mathbb{C})$ if we wish to satisfy (2.3). To obtain sufficient propagation of regularity, having in mind Remark 2.7, we state the following results.

In dimension $d = 1$ both in the defocusing case, for $3 \leq \sigma < +\infty$, and in the focusing case, for $3 \leq \sigma < 5$,

$$\Phi^X : \mathbb{R} \times H^3(\mathbb{R},\mathbb{C}) \to H^3(\mathbb{R},\mathbb{C}).$$

Hence, in these cases, using the notation introduced in Section 2.1, $E = H^1(\mathbb{R},\mathbb{C})$ and the domain $\mathscr{D}$ of the vector field X can be chosen to be the Sobolev space $H^3(\mathbb{R},\mathbb{C})$.

In dimension $d = 2,3$ and in the defocusing case, the global flow Φ^X can be defined on $E = H^1(\mathbb{R}^d,\mathbb{C})$ for all $3 \leq \sigma < 1 + \frac{4}{d-2}$. As before, the domain $\mathscr{D}$ of the vector field X can be chosen to be the Sobolev space $H^3(\mathbb{R}^d,\mathbb{C})$.

It follows in particular from what precedes that the cubic Schrödinger equation ($\sigma = 3$) fits in the framework of the previous section provided either $d = 1$ (with λ arbitrary) or $\lambda < 0$ and $d = 2,3$.

We now turn to the study of the symmetries of the nonlinear Schrödinger equation (2.33). Let $G = \mathrm{SO}(d) \times \mathbb{R}^d \times \mathbb{R}$ and define its action on $E = H^1(\mathbb{R}^d,\mathbb{C})$ via

$$\forall u \in H^1(\mathbb{R}^d), \quad \left(\Phi_{R,a,\gamma}(u)\right)(x) = e^{i\gamma} u(R^{-1}(x-a)). \tag{2.36}$$

Here the group law of G is

$$(R_1,a_1,\gamma_1)(R_2,a_2,\gamma_2) = (R_1R_2, a_1 + R_1a_2, \gamma_1+\gamma_2)$$

for all $R_1,R_2 \in \mathrm{SO}(d)$, $a_1,a_2 \in \mathbb{R}^d$ and $\gamma_1,\gamma_2 \in \mathbb{R}$. We claim that G is an invariance group (see Definition 2.1) for the dynamics Φ_t^X. Indeed, let $u(t,x) = (\Phi_t^X(u))(x)$ a solution to the nonlinear Schrödinger equation (2.33) and consider $\left((\Phi_{R,a,\gamma} \circ \Phi_t^X)(u)\right)(x) = e^{i\gamma}u(t,Rx-a)$. A straightforward calculation shows that $e^{i\gamma}u(t,R^{-1}(x-a))$ is again a solution to equation (2.33). More precisely,

$$\begin{aligned}
&i\partial_t(e^{i\gamma}u(t,R^{-1}(x-a))) + \Delta(e^{i\gamma}u(t,R^{-1}(x-a))) \\
&+ \lambda|e^{i\gamma}u(t,R^{-1}(x-a))|^{\sigma-1}(e^{i\gamma}u(t,R^{-1}(x-a))) \\
&= e^{i\gamma}\left(i(\partial_t u)(t,R^{-1}(x-a)) + (\Delta u)(t,R^{-1}(x-a)) + (\lambda|u|^{\sigma-1}u)(t,R^{-1}(x-a))\right) \\
&= 0
\end{aligned}$$

where we use the fact that the Laplace operator is invariant under space rotations, space translations and phase rotations. As a consequence,

$$\left((\Phi_{R,a,\gamma} \circ \Phi_t^X)(u)\right)(x) = \left((\Phi_t^X \circ \Phi_{R,a,\gamma})(u)\right)(x)$$

and G is an invariance group for the dynamics Φ_t^X. Moreover, we can easily prove that $H \circ \Phi_{R,a,\gamma} = H$. Indeed, using the definition of H given in (2.32), we have

$$\begin{aligned}
H \circ \Phi_{R,a,\gamma}(u) &= \frac{1}{2}\int_{\mathbb{R}^d} |\nabla u|^2(R^{-1}(x-a))\,\mathrm{d}x - \frac{\lambda}{\sigma+1}\int_{\mathbb{R}^d}|u|^{\sigma+1}(R^{-1}(x-a))\,\mathrm{d}x \\
&= H(u).
\end{aligned}$$

We will see later (in Section 5.3) why this is important.

Now, let us give some examples of G-relative equilibria of the nonlinear Schrödinger equation (2.33). First, consider the simplest case where $d = 1$ and $\sigma = 3$. The invariance group G reduces to $\mathbb{R} \times \mathbb{R}$ and the nonlinear Schrödinger equation becomes

$$i\partial_t u(t,x) + \partial_{xx}^2 u(t,x) + \lambda |u(t,x)|^2 u(t,x) = 0. \tag{2.37}$$

In the focusing case ($\lambda > 0$), there exists a two-parameters family of functions, the so-called *bright solitons*,

$$u_{\alpha,c}(t,x) = \alpha \sqrt{\frac{2}{\lambda}} \operatorname{sech}(\alpha(x-ct)) e^{i\left(\frac{c}{2}x + \left(\alpha^2 - \frac{c^2}{4}\right)t\right)}$$

that are solutions to (2.37) for all $(\alpha, c) \in \mathbb{R} \times \mathbb{R}$, with initial conditions

$$u_{\alpha,c}(x) = u_{\alpha,c}(0,x) = \alpha \sqrt{\frac{2}{\lambda}} \operatorname{sech}(\alpha x) e^{i\left(\frac{c}{2}x\right)} \in E = H^1(\mathbb{R}). \tag{2.38}$$

For each $(\alpha, c) \in \mathbb{R} \times \mathbb{R}$, $u_{\alpha,c}(x)$ is a G-relative equilibrium of (2.37). Indeed, the G-orbit of $u_{\alpha,c}(x)$ is given by

$$\mathscr{O}_{u_{\alpha,c}} = \left\{ e^{i\gamma} u_{\alpha,c}(x-a), (a,\gamma) \in \mathbb{R} \times \mathbb{R} \right\}. \tag{2.39}$$

Hence, it is clear that for all $t \in \mathbb{R}$, $u_{\alpha,c}(t,x) \in \mathscr{O}_{u_{\alpha,c}}$ and, by Definition 2.2, we can conclude that $u_{\alpha,c}(x)$ is a G-relative equilibrium of (2.37).

More generally, standing and travelling waves are examples of G-relative equilibria of the nonlinear Schrödinger equation (2.33). More precisely, standing waves are solutions to (2.33) of the form

$$u_{\mathrm{S}}(t,x) = e^{i\xi t} w_{\mathrm{S}}(x) \tag{2.40}$$

with $\xi \in \mathbb{R}$. For this to be the case, the profile w_{S} has to be a solution of the *stationary equation*

$$\Delta w + \lambda |w|^{\sigma-1} w = \xi w.$$

Bright solitons with $c = 0$ are examples of such standing waves, with $d = 1, \sigma = 3$. Standing waves of the one-dimensional Schrödinger equation with a spatially inhomogeneous nonlinearity, as well as their orbital stability, will be studied in Section 9. Travelling waves are solutions to (2.33) of the form

$$u_{\mathrm{TW}}(t,x) = e^{i\xi t} w_{\mathrm{TW}}(x - ct) \tag{2.41}$$

with $\xi \in \mathbb{R}$ and $c \in \mathbb{R}^d$. Now, the profile w_{TW} has to be a solution of

$$\Delta w + \lambda |w|^{\sigma-1} w = \xi w + ic \cdot \nabla w.$$

Bright solitons with $c \neq 0$ are examples of such travelling waves, with $d = 1, \sigma = 3$.

The G-orbit of the initial condition $w_{\mathrm{S}}(x)$ is given by

$$\mathscr{O}_{w_{\mathrm{S}}} = \left\{ e^{i\gamma} w_{\mathrm{S}}(R^{-1}(x-a)), (R,a,\gamma) \in G \right\} \tag{2.42}$$

and it is clear that $u_{\mathrm{S}}(t,x) \in \mathscr{O}_{w_{\mathrm{S}}}$ for all $t \in \mathbb{R}$. The same holds true for u_{TW} with w_{S} replaced by w_{TW}.

Another, closely related example of an infinite dimensional dynamical system is the cubic Schrödinger equation

$$\begin{cases} i\partial_t u(t,x) + \partial_{xx}^2 u(t,x) \pm |u(t,x)|^2 u(t,x) = 0 \\ u(0,x) = u_0(x) \end{cases} \tag{2.43}$$

in the space periodic setting $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ (the one dimensional torus). In [Bou93], the following theorem is proven.

Theorem 2.9 ([Bou93])**.** *The Cauchy problem* (2.43) *is globally well-posed for data* $u_0 \in H^s(\mathbb{T},\mathbb{C})$*,* $s \geq 0$ *and the solution* $u \in C^0(\mathbb{R},H^s(\mathbb{T}))$*. Moreover, if* u*,* v *are the solutions corresponding to data* $u_0, v_0 \in H^s(\mathbb{T},\mathbb{C})$*, there is the regularity estimate*

$$\|u(t) - v(t)\|_{H^s} \leq C^{|t|}\|u_0 - v_0\|_{H^s} \tag{2.44}$$

where C *depends on the* L^2*-size of the data, i.e.* $C = C(\|u_0\|_{L^2}, \|v_0\|_{L^2})$*.*

This ensures the existence of a global flow

$$\Phi^X : \mathbb{R} \times H^s(\mathbb{T},\mathbb{C}) \to H^s(\mathbb{T},\mathbb{C}).$$

for all $s \geq 1$. Hence, we can choose $E = H^1(\mathbb{T},\mathbb{C})$ and $\mathscr{D} = H^3(\mathbb{T},\mathbb{C})$ to ensure the conditions of Section 2.1 are satisfied.

As before, by using the invariance of Equation (2.43) under space translations and phase rotations, we can show that the dynamics defined by Φ_t^X are invariant under the action of the group $G = \mathbb{R} \times \mathbb{R}$ given by

$$(\Phi_{a,\gamma}(u))(x) = e^{i\gamma}u(x-a). \tag{2.45}$$

As an example of G-relative equilibria, we can consider the two-parameter family of plane waves

$$u_{\alpha,k}(t,x) = \alpha e^{-ikx}e^{i\xi t} \tag{2.46}$$

with $\alpha \in \mathbb{R}$ and $k \in \mathbb{Z}$ and $\xi = -k^2 \pm |\alpha|^2$. The G-orbit of the initial condition $u_{\alpha,k}(x) = \alpha e^{-ikx}$ is given by

$$\mathscr{O}_{u_{\alpha,k}} = \left\{ \alpha e^{i\gamma}e^{-ik(x-a)}, (a,\gamma) \in G \right\}.$$

As before, it is clear that $u_{\alpha,k}(t,x) \in \mathscr{O}_{u_{\alpha,k}}$ for all $t \in \mathbb{R}$. We will study the orbital stability of these relative equilibria in Section 8.

Remark that plane waves are the simplest elements of a family of solutions of the NLS equation of the form

$$u_{p,c}(t,x) = e^{i\xi t}e^{-ipx}U(x-ct),\ (t,x) \in \mathbb{R} \times \mathbb{R}$$

with $\xi, p, c \in \mathbb{R}$ and $U : \mathbb{R} \to \mathbb{C}$ a periodic function. This kind of solutions are called *quasi-periodic travelling waves* and their orbital stability has been studied in [GH07a, GH07b].

2.5. **The Manakov equation**

The Manakov equation [Man74, Gaz12] is a system of two coupled nonlinear Schrödinger equations which describe the evolution of nonlinear electric fields in optical fibers with birefringence, defined by

$$\begin{cases} i\partial_t u(t,x) + \Delta u(t,x) + \lambda |u(t,x)|^2 u(t,x) = 0 \\ u(0,x) = u_0(x) \end{cases} \tag{2.47}$$

with $u(t,x) = \begin{pmatrix} u_1(t,x) \\ u_2(t,x) \end{pmatrix} : \mathbb{R} \times \mathbb{R} \to \mathbb{C}^2$, $|u(t,x)|^2 = (|u_1(t,x)|^2 + |u_2(t,x)|^2)$ and $\lambda \in \mathbb{R}$.

With the same arguments as those used for the nonlinear Schrödinger equation (2.33), one can easily show that the flow is globally defined in $H^1(\mathbb{R},\mathbb{C}^2)$, i.e.

$$\Phi^X : \mathbb{R} \times H^1(\mathbb{R},\mathbb{C}^2) \to H^1(\mathbb{R},\mathbb{C}^2) \tag{2.48}$$

both in the focusing ($\lambda > 0$) and in the defocusing case ($\lambda < 0$). Moreover, thanks to the propagation of regularity, the flow preserves $H^3(\mathbb{R},\mathbb{C}^2)$ i.e.

$$\Phi^X : \mathbb{R} \times H^3(\mathbb{R},\mathbb{C}^2) \to H^3(\mathbb{R},\mathbb{C}^2) \tag{2.49}$$

as before. Hence, using the notation of Section 2.1, one can choose $E = H^1(\mathbb{R},\mathbb{C}^2)$ and the domain $\mathscr{D} = H^3(\mathbb{R},\mathbb{C}^2)$.

Now, let $(a,S) \in G = \mathbb{R} \times \mathrm{U}(2)$ act on $E = H^1(\mathbb{R},\mathbb{C}^2)$ via

$$\Phi_{a,S}(u) = Su(x-a). \tag{2.50}$$

Here the group law of G is $(a_1,S_1)(a_2,S_2) = (a_1+a_2, S_1S_2)$ for all $a_1,a_2 \in \mathbb{R}^d$ and $S_1,S_2 \in \mathrm{U}(2)$. A straightforward calculation proves that G is an invariance group for the dynamics Φ_t^X.

In the focusing case ($\lambda > 0$), there exists a family of solitons,

$$u_\nu(t,x) = \alpha\sqrt{\frac{2}{\lambda}}\mathrm{sech}(\alpha(x-ct))e^{i\left(\frac{c}{2}x+\left(\alpha^2-\frac{c^2}{4}\right)t\right)}\begin{pmatrix}\cos\theta e^{i\gamma_1}\\ \sin\theta e^{i\gamma_2}\end{pmatrix}$$

that are solutions to (2.47) for all $\nu = (\alpha,c,\theta,\gamma_1,\gamma_2) \in \mathbb{R}^5$, with initial condition

$$u_\nu(x) = u_\nu(0,x) = \alpha\sqrt{\frac{2}{\lambda}}\mathrm{sech}(\alpha x)e^{i\left(\frac{c}{2}x\right)}\begin{pmatrix}\cos\theta e^{i\gamma_1}\\ \sin\theta e^{i\gamma_2}\end{pmatrix} \in E = H^1(\mathbb{R},\mathbb{C}^2).$$

For each $\nu \in \mathbb{R}^5$, $u_\nu(x)$ is a G-relative equilibrium of (2.47). Indeed, the G-orbit of $u_\nu(x)$ is given by

$$\mathscr{O}_{u_\nu} = \{Su_\nu(x-a), (a,S) \in \mathbb{R} \times \mathrm{U}(2)\}.$$

Hence, it is clear that for all $t \in \mathbb{R}$, $u_\nu(t,x) \in \mathscr{O}_{u_\nu}$ and, by Definition 2.2, we can conclude that $u_\nu(x)$ is a G-relative equilibrium of (2.37).

2.6. **The nonlinear wave equation**

Let us consider the nonlinear wave equation

$$\begin{cases}\partial_{tt}^2 u(t,x) - \Delta u(t,x) + \lambda|u(t,x)|^{\sigma-1}u(t,x) = 0\\ u(0,x) = u_0(x), \partial_t u(0,x) = u_1(x)\end{cases} \tag{2.51}$$

with $u(t,x) : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, and, for simplicity, let us take $d = 1,2,3$. Moreover, we restrict our attention to the defocusing case, that in our notation corresponds to $\lambda > 0$ (because of the minus sign in front of the Laplacian), and to the so-called algebraic nonlinearities, which means $\sigma \in \mathbb{N}$ is odd. As a consequence the function $f(u) = |u|^{\sigma-1}u$ is smooth.

Let H defined by

$$H(u,\partial_t u) = \frac{1}{2}\int_{\mathbb{R}^d}|\nabla u|^2\,\mathrm{d}x + \frac{1}{2}\int_{\mathbb{R}^d}|\partial_t u|^2\,\mathrm{d}x + \frac{\lambda}{\sigma+1}\int_{\mathbb{R}^d}|u|^{\sigma+1}\,\mathrm{d}x \tag{2.52}$$

be the Hamiltonian of the system. As for the Schrödinger equation, we will explain in Section 5 how the nonlinear wave equation defines an infinite dimensional Hamiltonian dynamical system.

In the defocusing case and whenever $1 \le \sigma < +\infty$ for $d = 1$ or $1 \le \sigma < 1 + \frac{4}{d-2}$ for $d = 2,3$, we can define a global flow on $H^1(\mathbb{R}^d,\mathbb{R}) \times L^2(\mathbb{R}^d,\mathbb{R})$, i.e.

$$\begin{aligned}\Phi^X : \mathbb{R} \times (H^1(\mathbb{R}^d,\mathbb{R}) \times L^2(\mathbb{R}^d,\mathbb{R})) &\to H^1(\mathbb{R}^d,\mathbb{R}) \times L^2(\mathbb{R}^d,\mathbb{R})\\ (t,u(0),\partial_t u(0)) &\to (t,u(t),\partial_t u(t))\end{aligned} \tag{2.53}$$

with $u \in C(\mathbb{R}, H^1(\mathbb{R}^d)) \cap C^1(\mathbb{R}, L^2(\mathbb{R}^d))$ the unique solution to (2.51). Moreover the Hamiltonian energy (2.52) is conserved along the flow, i.e.

$$H(u(0), \partial_t u(0)) = H(u(t), \partial_t u(t))$$

for all $t \in \mathbb{R}$ (see [Tao06] and references therein). Furthermore, it follows from the integral form of (2.51) (see [Tao06, Ex. 2.18 and 2.22]) that $u \in C^2(\mathbb{R}, H^{-1}(\mathbb{R}^d))$.

In the algebraic case, thanks to the persistence of regularity, the flow can be restricted to $H^s(\mathbb{R}^d, \mathbb{R}) \times H^{s-1}(\mathbb{R}^d, \mathbb{R})$,

$$\Phi^X : \mathbb{R} \times (H^s(\mathbb{R}^d, \mathbb{R}) \times H^{s-1}(\mathbb{R}^d, \mathbb{R})) \to H^s(\mathbb{R}^d, \mathbb{R}) \times H^{s-1}(\mathbb{R}^d, \mathbb{R})$$

for all $s > \frac{d}{2}$. Hence, using the notation of Section 2.1, $E = H^1(\mathbb{R}^d, \mathbb{R}) \times L^2(\mathbb{R}^d, \mathbb{R})$ and the domain $\mathscr{D}$ of the vector field X can be chosen to be the Sobolev space $H^2(\mathbb{R}^d, \mathbb{R}) \times H^1(\mathbb{R}^d, \mathbb{R})$.

As for the nonlinear Schrödinger equation, by using the invariance of Equation (2.51) under space rotations, space translations and phase rotations, we can show that the dynamics defined by Φ_t^X are invariant under the action of the group $G = \mathrm{SO}(d) \times \mathbb{R}^d$ on $E = H^1(\mathbb{R}^d, \mathbb{R}) \times L^2(\mathbb{R}^d, \mathbb{R})$ defined by

$$\Phi_{R,a}(u, \partial_t u) = (u(R^{-1}(x-a)), \partial_t u(R^{-1}(x-a))).$$

Moreover, $H \circ \Phi_{R,a,\gamma} = H$ and we will explain in Section 5.3 the consequences of this fact.

2.7. **Generalized symmetries**

The nonlinear Schrödinger equation is often said to be invariant under Galilei transformations. This invariance is however of a slightly different nature than the one defined in Definition 2.1, as we now explain[5].

Recall that Newtonian mechanics is known to be invariant under coordinate changes between inertial frames. These include space and time translations, rotations, and changes to a moving frame, often referred to as Galilei boosts. All together, they form a group, the Galilei group G_{Gal}, which is a Lie group that can be defined formally as

$$G_{\mathrm{Gal}} = \mathrm{SO}(d) \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}$$

with composition law

$$(R', v', a', t')(R, v, a, t) = (R'R, R'v + v', R'a + a' + v't, t + t').$$

It acts naturally on space-time $(x,t) \in \mathbb{R}^d \times \mathbb{R}$ as follows:

$$(R', v', a', t')(x, t) = (R'x + a' + v't, t' + t).$$

Of course, the physical case corresponds to $d = 3$.

The statement that Newton's equations are invariant under boosts means for example that, if $t \to (q_1(t), q_2(t))$ is the solution of Newton's equations of motion for two particles moving in a spherically symmetric interaction potential V

$$m_1 \ddot{q}_1(t) = -\nabla_{q_1} V(\|q_1(t) - q_2(t)\|), \quad m_2 \ddot{q}_2(t) = -\nabla_{q_2} V(\|q_1(t) - q_2(t)\|),$$

with initial conditions

$$q_1(0) = q_1,\ q_2(0) = q_2,\ \dot{q}_1(0) = \frac{p_1}{m_1},\ \dot{q}_2(0) = \frac{p_2}{m_2},$$

then, for all $v \in \mathbb{R}^3$, $t \to (q_1(t) + vt, q_2(t) + vt)$ is also such a solution, with initial conditions

$$q_1(0) = q_1,\ q_2(0) = q_2,\ \dot{q}_1(0) = \frac{p_1}{m_1} + v,\ \dot{q}_2(0) = \frac{p_2}{m_2} + v.$$

[5]We will, in this section, make free use of the material of Appendices A.2 and A.3.

In a Hamiltonian description[6], the above equations of motion are associated to the Hamiltonian

$$H(q,p) = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + V(\|q_1 - q_2\|),$$

which generates a flow Φ_t^H that is clearly invariant under space translations and rotations. The situation for Galilei boosts, however, is different. Indeed, in this context they act on the phase space $E = \mathbb{R}^6 \times \mathbb{R}^6$ with symplectic transformations, as follows:

$$\forall v \in \mathbb{R}^3, \quad \Phi_v^K(q,p) = (q, p_1 - m_1 v, p_2 - m_2 v).$$

Here $K = m_1 q_1 + m_2 q_2$ and Φ_v^K is a shorthand notation for

$$\Phi_v^K = \Phi_{v_1}^{K_1} \circ \Phi_{v_2}^{K_2} \circ \cdots \circ \Phi_{v_n}^{K_n},$$

where each $\Phi_{v_i}^{K_i}$ is the hamiltonian flow of one component of K. But those do NOT commute with the dynamical flow Φ_t^H. Indeed, one easily checks that

$$\Phi_v^K \Phi_t^H \Phi_{-v}^K = \Phi_{vt}^P \Phi_t^H, \tag{2.54}$$

where $P = p_1 + p_2$ is the total momentum of the system, which generates translations: $\Phi_a^P(q_1,q_2,p_1,p_2) = (q_1 + a, q_2 + a, p_1, p_2)$. In that sense, the three dimensional commutative group of Galilei boosts is NOT an invariance group for the dynamical system according to Definition 2.1. To remedy this situation, one can proceed as follows. Define, on $E = \mathbb{R}^6 \times \mathbb{R}^6$, for each $g = (R, v, a, t) \in G_{\text{Gal}}$, the symplectic transformation

$$\Phi_g = \Phi_a^P \Phi_t^H \Phi_v^K \Phi_R,$$

where Φ_R is defined as in (2.17). It is then easily checked using (2.54) that the Φ_g define an action of G_{Gal} on E. It is clearly globally Hamiltonian (Definition A.3.7)[7]. It follows that the Galilei boosts are generalized symmetries for the dynamical system Φ_t^H, in the following sense:

Definition 2.10. Let G be a Lie group, and Φ an action of G on a Banach space E. Let Φ_t^X a dynamical system on E. We say G is a generalized symmetry group for Φ_t^X provided there exists $\xi \in \mathfrak{g}$ so that $\Phi_t^X = \Phi_{\exp(t\xi)}$.

For our purposes, an important difference between symmetries and generalized symmetries in Hamiltonian systems is that the latter do NOT give rise to constants of the motion. To illustrate this, remark that, although the Galilei boosts are generated by $K(q,p) = m_1 q_1 + m_2 q_2$, it is clear that K is not a constant of the motion of H:

$$\{K, H\} = P, \tag{2.55}$$

where $P = p_1 + p_2$ is the total momentum of the two-particle system. This is not a surprise: $K = MR$, where R is the center of mass of the two-particle system and $M = m_1 + m_2$ its mass. And of course, the center of mass moves: in fact, (2.55) implies it moves at constant velocity.

A similar situation occurs with the nonlinear Schrödinger equation. If $u(t,x)$ is a solution of (2.22) with a power law nonlinearity, then so is, for every $v \in \mathbb{R}^d$,

$$\tilde{u}(t,x) = \exp\big(-\tfrac{i}{2}(v \cdot x + \tfrac{v^2}{2}t)\big)\, u(t, x + vt), \tag{2.56}$$

as is readily checked. The function $\tilde{u}$ can be interpreted as the wave function in the moving frame, as can be seen from the shift $x \to x + vt$ in position and from the factor $\exp(-i\frac{v}{2} \cdot x)$,

[6] See Appendix A.3.

[7] It is however not Ad^*-equivariant.

which corresponds to a translation by $\frac{1}{2}v$ in momentum, in the usual quantum mechanical interpretation of the Schrödinger equation. Adopting the framework of Section 2.4, one observes that the maps

$$\widehat{\Psi}_v u(x) = \exp\left(-\tfrac{i}{2}(v\cdot x)\right) u(x),$$

defined for all $v \in \mathbb{R}^d$ on $E = H^1(\mathbb{R}^d)$ are not symmetries for the Schrödinger flow Φ^X defined in (2.35) but that

$$\widehat{\Psi}_v \Phi^X_t \widehat{\Psi}_{-v} = \Phi_{I,vt,-\frac{v^2}{4}t}\, \Phi^X_t, \tag{2.57}$$

where $\Phi_{I,vt,-\frac{v^2}{4}t}$ is defined in (2.36). This commutation relation is very similar to (2.54), except for the extra phase $\exp(-i\frac{v^2}{4}t)$. We note in passing that the boosts $\hat{\Psi}_v$ are unitary on L^2, but do not preserve the H^1 norm. They are nevertheless bounded operators on $E = H^1(\mathbb{R}^d)$.

As in classical mechanics, one can put together the above transformations with the representation of the Euclidean group in (2.36) to form a (projective) representation of the Galilei group showing that the Galilei boosts are generalized symmetries of the nonlinear Schrödinger equation with a power law nonlinearity. We will not work this out in detail here, but note for further use that

$$\Phi_{R,a,\gamma}\widehat{\Psi}_v = \exp(i\tfrac{v\cdot a}{2})\widehat{\Psi}_{Rv}\Phi_{R,a,\gamma}. \tag{2.58}$$

In particular $\Phi_{I,a,0}\widehat{\Psi}_v = \exp(i\frac{v\cdot a}{2})\widehat{\Psi}_v\Phi_{I,a,0}$ so that, in this setting, the boosts $\hat{\Psi}_v$ commute with translations only "up to a global phase" $\exp(i\frac{v\cdot a}{2})$, in the usual terminology of quantum mechanics. In contrast, in classical mechanics, Φ^K_v and Φ^P_a clearly commute.

Generalized symmetries do not provide constants of the motion via Noether's Theorem, and hence cannot quite play the same role as symmetries in the study of relative equilibria. We will now show how one may nevertheless use (2.57) in the analysis of the stability of the relative equilibria of the (non)linear Schrödinger equation.

We first remark that the $u_{\alpha,c}$, defined in (2.38), satisfy $u_{\alpha,c} = \widehat{\Psi}_{-c}u_{\alpha,0}$. We will show that, thanks to (2.57), if $u_{\alpha,0}$ is orbitally stable, then so is $u_{\alpha,c}$, for any $c \in \mathbb{R}$. We will only sketch the argument, leaving the details to the reader. Note first that $u_{\alpha,0}$ is orbitally stable, if and only if, for all $\varepsilon > 0$, there exists $\delta > 0$ so that, for all $w \in E$ with $\mathrm{d}(w, u_{\alpha,0}) \le \delta$, there exists, for all $t \in \mathbb{R}$, $a(t) \in \mathbb{R}, \gamma(t) \in \mathbb{R}$ so that

$$\|\Delta_t\| \le \varepsilon, \quad \text{where} \quad \Delta_t := \Phi^X_t w - \Phi_{I,a(t),\gamma(t)}u_{\alpha,0}.$$

Now suppose $u \in E$ is sufficiently close to $u_{\alpha,c}$, for some $c \in \mathbb{R}$. Then, since $\widehat{\Psi}_c$ is a bounded operator, $\widehat{\Psi}_c u = w$ is close to $u_{\alpha,0}$. Then, using (2.57) and (2.58), one finds

$$\begin{aligned}
\Phi^X_t u &= \Phi^X_t \widehat{\Psi}_{-c} w \\
&= \widehat{\Psi}_{-c}\Phi_{I,ct,-\frac{c^2}{4}t}\,\Phi^X_t w \\
&= \widehat{\Psi}_{-c}\Phi_{I,ct,-\frac{c^2}{4}t}\,\Phi_{I,a(t),\gamma(t)}u_{\alpha,0} + \widehat{\Psi}_{-c}\Phi_{I,ct,-\frac{c^2}{4}t}\Delta_t \\
&= \widehat{\Psi}_{-c}\Phi_{I,ct+a(t),-\frac{c^2}{4}t+\gamma(t)}u_{\alpha,0} + \widehat{\Psi}_{-c}\Phi_{I,ct,-\frac{c^2}{4}t}\Delta_t \\
&= \Phi_{I,ct+a(t),-\frac{c^2}{4}t+\gamma(t)+\frac{c(ct+a(t))}{2}}\widehat{\Psi}_{-c}u_{\alpha,0} + \widehat{\Psi}_{-c}\Phi_{I,ct,-\frac{c^2}{4}t}\Delta_t.
\end{aligned}$$

Since $u_{\alpha,c} = \widehat{\Psi}_{-c}u_{\alpha,0}$, and since $\widehat{\Psi}_{-c}$ is bounded, it is now clear that $\Phi^X_t u$ is at all times close to $\mathscr{O}_{u_{\alpha,c}}$, defined in (2.39).

The above argument shows, more generally, that the relative equilibria of the homogeneous NLS for $G = \mathrm{SO}(d) \times \mathbb{R}^d \times \mathbb{R}$ (see (2.36)) come in families $\widehat{\Psi}_{-c} u_0 = u_c$, indexed by $c \in \mathbb{R}^d$. Moreover, if u_0 is spherically symmetric and orbitally stable, then all u_c are orbitally stable.

3. Orbital stability: a general definition

We can now formulate the general definition of orbital stability that we shall study. In fact, several definitions appear naturally:

Definition 3.1. Let Φ_t^X be a dynamical system on a Banach space E and let G be a symmetry group for Φ_t^X.

(i) Let $u \in E$ and let $\mathscr{O}_u$ be the corresponding G-orbit . We say $u \in E$ is orbitally stable if

$$\forall \varepsilon > 0, \exists \delta > 0, \forall v \in E, \ \left(\mathrm{d}(v,u) \leq \delta \Rightarrow \forall t \in \mathbb{R}, \ \inf_{t' \in \mathbb{R}} \mathrm{d}(v(t), \mathscr{O}_{u(t')}) \leq \varepsilon \right).$$

(ii) Let $\mathscr{O}$ be a G-orbit in E. We say $\mathscr{O}$ is stable if each $u \in \mathscr{O}$ is orbitally stable in the sense of (i) above.

(iii) Let $\mathscr{O}$ be a G-orbit in E. We say $\mathscr{O}$ is uniformly stable if it is stable and δ in (i) does not depend on $u \in \mathscr{O}$. In other words, if $\forall \varepsilon > 0$, there exists $\delta > 0$ so that, $\forall u \in \mathscr{O}$, $\forall v \in E$,

$$\mathrm{d}(v,u) \leq \delta \Rightarrow \forall t \in \mathbb{R}, \ \inf_{t' \in \mathbb{R}} \mathrm{d}(v(t), \mathscr{O}_{u(t')}) \leq \varepsilon. \tag{3.1}$$

(iv) We say $\mathscr{O} \in E_G$ is Hausdorff orbitally stable if $\mathscr{O}$ satisfies: $\forall \varepsilon > 0$, there exists $\delta > 0$ so that, $\forall \mathscr{O}' \in E_G$

$$\Delta(\mathscr{O}, \mathscr{O}') \leq \delta \Rightarrow \forall t \in \mathbb{R}, \ \inf_{t'} \Delta(\mathscr{O}'(t), \mathscr{O}(t')) \leq \varepsilon. \tag{3.2}$$

The four definitions are subtly different.

Definition (i) requires that the dynamical orbit issued from the nearby initial condition v remains close to the orbit $\{\Phi_{t'}^X \Phi_g(u) \mid t' \in \mathbb{R}, g \in G\}$ of the larger group $\mathbb{R} \times G$. It is therefore a generalization of definition (1.4), which corresponds to the case $G = \{e\}$. This notion of orbital stability therefore depends on the choice of the group G and it is clear that, the larger G, the weaker it is. As we will see in the examples of Section 4 and Section 5.5, there are cases where definition (1.4) is not satisfied for some $u \in E$, but where the above definition holds for a suitable choice of G. As we will also see, the choice of G may depend on the point $u \in E$ considered and it is in particular not always necessary to use the largest symmetry group G available for Φ_t^X to obtain orbital stability.

The stability of the orbit $\mathscr{O}$ as defined in part (ii) simply requires the orbital stability of each point $u \in \mathscr{O}$, as defined in (i). Note that δ depends on u here. In part (iii) of the definition, uniformity is required.

Part (iv) requires that if two G-orbits $\mathscr{O}, \mathscr{O}' \subset E$ are initially close (in the sense of the Hausdorff metric) then, for all t, $\mathscr{O}'(t)$ is close to $\mathscr{O}(t')$ for some value of t'. It is the natural transcription of the definition of orbital stability in (1.4) from the original dynamical system on E to the reduced dynamics on E_G.

Parts (i), (ii) and (iii) are the most telling/interesting, since they give a statement directly on the phase space E, using the original distance d, rather than in the more abstract quotient space E_G. They do moreover not use the somewhat unpleasant Hausdorff metric. In applications, one really wants to prove (i), (ii) or (iii).

As shown in the lemma below, the four definitions in Definition 3.1 are equivalent when the group action is isometric. For many applications in infinite dimensional systems in particular, this is the case.

Lemma 3.2. *Let Φ_t^X be a dynamical system on E and let G be a symmetry group for Φ_t^X, acting isometrically. Let $u \in E$. Then the following statements are equivalent.*

(i) $u \in E$ is orbitally stable.
(ii) Each $v \in \mathscr{O}_u$ is orbitally stable.
(iii) $\mathscr{O}_u$ is uniformly stable.
(iv) $\mathscr{O}_u$ is Hausdorff orbitally stable.

In practice, one often proves (i) for a suitably chosen u on the orbit. This then automatically yields (iii). The statement in terms of the reduced dynamics in (iv) is intellectually satisfying but rarely encountered, it seems.

Proof. We prove $(i) \Leftrightarrow (ii)$ and $(i) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i)$.

$(i) \Rightarrow (iii)$ and $(i) \Rightarrow (ii)$: Let $v \in \mathscr{O}_u$ and $v' \in E$, $\mathrm{d}(v', v) \leq \delta$. Then there exists $g \in G$ so that $v = \Phi_g(u)$. Define $u' = \Phi_g^{-1}(v')$. Then, by the isometry of Φ_g, $\mathrm{d}(u', u) \leq \delta$ and hence, by hypothesis, for all t, there exists t' so that $\mathrm{d}(u'(t), \mathscr{O}_{u(t')}) < \varepsilon$. Hence

$$\mathrm{d}(v'(t), \mathscr{O}_{v(t')}) = \mathrm{d}(\Phi_g(u'(t)), \mathscr{O}_{u(t')}) = \mathrm{d}(u'(t), \mathscr{O}_{u(t')}) \leq \varepsilon.$$

This proves (iii) and, in particular, (ii). Since it is clear that $(ii) \Rightarrow (i)$, we obtain $(i) \Leftrightarrow (ii)$.

$(iii) \Rightarrow (iv)$: Suppose $\mathscr{O}_u$ is uniformly stable. Let $\mathscr{O}'$ be such that $\Delta(\mathscr{O}_u, \mathscr{O}') < \delta$. Let $u' \in \mathscr{O}'$ with $\mathrm{d}(u, u') \leq \delta$. Then (3.1), together with Proposition 2.3 (ii), imply $\Delta(\mathscr{O}'(t), \mathscr{O}_{u(t')}) \leq \varepsilon$.

$(iv) \Rightarrow (i)$: Suppose $\mathscr{O}_u$ is orbitally stable. Let $u' \in E$ so that $\mathrm{d}(u, u') \leq \delta$. Let $\mathscr{O}' = \mathscr{O}_{u'}$. Then, by Proposition 2.3 (iii), $\Delta(\mathscr{O}, \mathscr{O}') \leq \delta$. Hence, for all t, $\inf_{t'} \Delta(\mathscr{O}'(t), \mathscr{O}(t')) \leq \varepsilon$. Proposition 2.3 (ii) then implies (i). □

In many applications, especially in infinite dimensional problems, the Φ_g are both linear and norm-preserving: several examples were given in Section 2. In that case the action is of course isometric. In addition, all group orbits are then bounded. Note nevertheless that, if the Φ_g are norm-preserving, but not linear, the action is no longer isometric, while the group orbits are still bounded. Finally, isometric actions may have unbounded group orbits: think for example of translations on $E = \mathbb{R}^{2n}$.

4. Orbital stability in spherical potentials

Before presenting the general Lyapunov approach to the proof of orbital stability in Section 7, we show here the orbital stability of the relative equilibria in spherical potentials that we identified in Section 2.3. This simple example is instructive for several reasons. First, it permits one to appreciate the group theoretic and symplectic mechanisms underlying the construction of a suitable candidate Lyapunov function. Second, it nicely illustrates the various methods available to use this Lyapunov function in order to prove orbital stability via an appropriate "coercivity estimate" generalizing (1.2). We will present three such methods below.

4.1. Fixed points

The proof of the uniform orbital stability of $\mathscr{O}_{\rho_*,0,0}$ in (2.20) is straightforward, and can be done with H itself as the Lyapunov function, in close analogy with the proof sketched in the introduction.

Proposition 4.1. *Let $V \in C^2(\mathbb{R}^3)$ be a spherical potential and $H(u) = \frac{1}{2}p^2 + V(q)$ the corresponding Hamiltonian. Let $\rho_* > 0$ with $V'(\rho_*) = 0$, $V''(\rho_*) > 0$. Let $\mathscr{O}_{\rho_*,0,0} = \{(q,p) \in \mathbb{R}^6 \mid \|q\| = \rho_*, p = 0\}$ be the corresponding $SO(3)$ orbit. Then $\mathscr{O}_{\rho_*,0,0}$ is uniformly orbitally stable.*

This result is intuitively clear. Under the assumptions stated, the Hamiltonian reaches a local minimum at each of the fixed points of the dynamics that make up the sphere $\mathscr{O}_{\rho_*,0,0}$, and it increases quadratically in directions perpendicular to that sphere. Any nearby initial condition must therefore give rise to an orbit that stays close to the sphere: the potential acts locally as a potential well trapping the particle close to $\mathscr{O}_{\rho_*,0,0}$.

Proof. We know from Section 2.3 that the Hamiltonian H in (2.13) is an SO(3)-invariant constant of the motion, and that $D_u H = 0$ for all $u \in \mathscr{O}_{\rho_*,0,0}$, so that each such point is a fixed point of the dynamics. We will write $H_* = H(u), \forall u \in \mathscr{O}_{\rho_*,0,0}$. Moreover, for all $u = (q,0) \in \mathscr{O}_{\rho_*,0,0}$

$$D_u^2 H = \begin{pmatrix} V''(\rho_*)\hat{q}_i\hat{q}_j & 0 \\ 0 & \mathrm{I}_3 \end{pmatrix}.$$

Note that the Hessian is not positive definite. In fact, it vanishes on $w = (a,0)$, for $a \cdot q = 0$, which is the two-dimensional tangent space $T_u\mathscr{O}_{\rho_*,0,0}$ to the orbit. We can therefore not expect to obtain a coercive estimate as in (1.2). On the other hand, since $V''(\rho_*) > 0$, $D_u^2 H$ *is* positive definite on the four-dimensional orthogonal complement to the tangent space, given by

$$\left(T_u\mathscr{O}_{\rho_*,0,0}\right)^{\perp} = \{(\alpha\hat{q}, b) \in \mathbb{R}^6 \mid \alpha \in \mathbb{R}, b \in \mathbb{R}^3\}. \tag{4.1}$$

As a result, we can still show that there exist constants $c_*, \eta_* > 0$ with the property that

$$\forall u' \in E, \left(\mathrm{d}(u', \mathscr{O}_{\rho_*,0,0}) \leq \eta_* \Rightarrow H(u') - H_* \geq c_* \mathrm{d}(u', \mathscr{O}_{\rho_*,0,0})^2\right), \tag{4.2}$$

and this will suffice for the proof of orbital stability. To show (4.2), note first that setting $u' = (q', p')$ and taking $\eta_* < \rho_*/2$, one has $q' \neq 0$. Consider then $u = (\rho_* \hat{q}', 0) \in \mathscr{O}_{\rho_*,0,0}$ and remark that $\mathrm{d}(u', \mathscr{O}_{\rho_*,0,0}) = \|u' - u\|$. Now compute

$$\begin{aligned} H(u') - H_* &= H(u') - H(u) = D_v^2 H(u' - u, u' - u) + \mathrm{o}(\|u' - u\|^2) \\ &\geq \min\{1, V''(\rho_*)\}\mathrm{d}(u', \mathscr{O}_{\rho_*,0,0})^2 + \mathrm{o}(\mathrm{d}(u', \mathscr{O}_{\rho_*,0,0})^2). \end{aligned}$$

One can then conclude (4.2) holds by using that the term in $\mathrm{o}(\mathrm{d}(u', \mathscr{O}_{\rho_*,0,0})^2)$ is uniformly small in $u \in \mathscr{O}_{\rho_*,0,0}$ since H is SO(3)-invariant. We now prove that $\mathscr{O}_{\rho_*,0,0}$ is uniformly orbitally stable. Since the action of SO(3) is isometric, Lemma 3.2 shows it is enough to prove all $u \in \mathscr{O}_{\rho_*,0,0}$ are orbitally stable. Suppose that this is not true. Then there exists $u \in \mathscr{O}_{\rho_*,0,0}$ and $\varepsilon > 0$, and for each $n \in \mathbb{N}_*$, $u_n' \in E$, $t_n \in \mathbb{R}$ so that $\mathrm{d}(u_n', u) \leq \frac{1}{n}$ and $\mathrm{d}(u_n'(t_n), \mathscr{O}_{\rho_*,0,0}) = \varepsilon_0$. Since we can choose $\varepsilon < \eta_*$, we can apply (4.2) to write

$$H(u_n') - H(u) = H(u_n'(t_n)) - H_* \geq c_* \mathrm{d}(u_n'(t_n), \mathscr{O}_{\rho_*,0,0})^2 = c_* \varepsilon^2.$$

Taking $n \to +\infty$ leads to the desired contradiction. □

4.2. **Circular orbits**

Proving an appropriate notion of stability for the initial conditions in (2.19) giving rise to circular orbits of the dynamics turns out to be slightly less straightforward. Intuitively, as explained already in the introduction, one expects that, under a suitable condition on the potential, an initial condition close to a circular orbit will generate a dynamical orbit that stays close to this orbit. As a result, orbital stability is satisfied in the sense of (1.4). The following proposition gives a precise statement of this phenomenon.

Proposition 4.2. *Let $V \in C^2(\mathbb{R}^3)$ be a spherical potential and $H(u) = \frac{1}{2}p^2 + V(q)$ the corresponding Hamiltonian. Let $\rho_*, \sigma_* > 0$ with $V'(\rho_*)\rho_* = \sigma_*^2$. Consider $u_{\mu_*} = (q_*, p_*) = (\rho_*\hat{q}_*, \sigma_*\hat{p}_*)$, with $\hat{q}_* \cdot \hat{p}_* = 0$. Then u_{μ_*} is a relative equilibrium for the group $SO(2)$ of rotations about $\mu_* = q_* \wedge p_*$. If in addition,*

$$V''(\rho_*) > -3\sigma_*^2\rho_*^{-2}, \tag{4.3}$$

u_{μ_*} is orbitally stable in the sense of definition (1.4) *and of Definition 3.1 (i). In addition, u_{μ_*} is a local minimum of H_{μ_*}, the restriction of H to the level surface Σ_{μ_*}, defined in* (2.18)*.*

Note that the two definitions of orbital stability mentioned coincide in this particular case. Also, since the action of the rotation group is isometric, the result implies uniform orbital stability as well. Below, we will give three different arguments to prove the proposition, each of which can and has been used to treat various infinite dimensional problems.

The origin of the condition $V''(\rho_*) > -3\sigma_*^2\rho_*^{-2}$ can be understood as follows. In standard mechanics textbooks such as [Gol80], motion in a spherical potential is treated by fixing the angular momentum $q \wedge p = \mu_*$, and then using for q, p polar coordinates $(r, \theta, p_r, p_\theta)$ in the plane perpendicular to the angular momentum. The Hamiltonian then reads, in these coordinates,

$$H(r, \theta, p_r, p_\theta) = \frac{p_r^2}{2} + \frac{p_\theta^2}{2r^2} + V(r).$$

The equation of motions are

$$\dot{r} = p_r, \quad \dot{\theta} = \frac{p_\theta}{r^2} \quad \dot{p}_r(t) = \frac{p_\theta^2}{r^3} - V'(r), \quad \dot{p}_\theta = 0$$

and $\mu_* = p_\theta$. It follows that the radial motion is decoupled from the angular one, since $\ddot{r} = -V'_{\mu_*}(r)$ with $V_{\mu_*}(r) = V(r) + \frac{\mu_*^2}{2r^2}$. It is then clear that the circular orbits correspond to the critical points $r = \rho_*$ of the effective potential V_{μ_*} which are fixed points of the radial dynamics. By an argument as in the introduction, those are stable if the critical point is a local minimum of

$$H_{\mu_*}(r, p_r) = \frac{p_r^2}{2} + V_{\mu_*}(r),$$

and so in particular if $V''_{\mu_*}(\rho_*) > 0$, which is precisely condition (4.3). Note however that the preceding argument does not prove orbital stability of the circular orbits: it does not allow to consider initial conditions $u \in \mathbb{R}^6$ with $\mu \neq \mu_*$. This is actually the tricky part of the proof of the proposition.

Proof. To mimic the previous proof, we would like to find a constant of the motion $\mathscr{L}$ which is SO(2) invariant and so that $D\mathscr{L}$ vanishes on the orbit under consideration. We cannot use H for this, since clearly $D_{u_{\mu_*}}H \neq 0$, as we are not dealing with a fixed point of the dynamics. On the other hand, as we pointed out after the definition of relative equilibrium, when u_{μ_*} is a relative equilibrium, then there exists an element ξ of the Lie-algebra of the invariance group so that $X_H(u_{\mu_*}) = X_\xi(u_{\mu_*})$ or, equivalently, so that

$$D_{u_{\mu_*}}(H - F_\xi) = 0.$$

In the present case, the invariance group is a one-dimensional rotation group and the statement becomes: there exists $\xi \in \mathbb{R}$ so that

$$D_{u_{\mu_*}}(H - \xi\mu_* \cdot L) = 0, \tag{4.4}$$

since, as we saw, $\mu_* \cdot L$ generates rotations about the μ_*-axis. Since, for all $u \in \mathbb{R}^6$

$$D_u H = (V'(\|q\|)\hat{q}, p), \qquad D_u(\mu_* \cdot L) = (p \wedge \mu_*, \mu_* \wedge q),$$

one easily checks that (4.4) is satisfied iff $\xi = \rho_*^{-2}$. This suggests to define

$$\mathscr{L}(u) = H(u) - \rho_*^{-2} \mu_* \cdot L(u)$$

and to try using it as a Lyapunov function. $\mathscr{L}$ is often referred to as the "augmented Hamiltonian". Note that the theory of Lagrange multipliers implies that (4.4) is equivalent to the statement that the restriction H_{μ_*} of H to Σ_{μ_*} has u_{μ_*} as a critical point. Hence the circular orbits can be characterized as the critical points of H_{μ_*}. This is a general feature of relative equilibria of Hamiltonian systems, as shown in Theorem 6.1.

The main ingredient of the proof is the following statement:

$$\exists c > 0, \forall v \in \mathscr{O}_{u_{\mu_*}}, \forall w \in \left(T_v \mathscr{O}_{u_{\mu_*}}\right)^{\perp} \cap T_v \Sigma_{\mu_*}, \quad D_v^2 \mathscr{L}(w,w) \geq c\|w\|^2. \tag{4.5}$$

This is a lower bound on the Hessian of $\mathscr{L}$ restricted to the two-dimensional subspace of $\mathbb{R}^6$ spanned by the vectors tangent to Σ_{μ_*} (see (2.18)) and perpendicular to the dynamical orbit $\mathscr{O}_{u_{\mu_*}} \subset \Sigma_{\mu_*}$. It will allow us to show the following lower bound on the variation of the Lyapunov function, which is to be compared to (1.2):

$$\begin{aligned} &\exists \delta > 0,\, c > 0, \forall u' \in \Sigma_{\mu_*}, \\ &\qquad \left(\mathrm{d}(u', \mathscr{O}_{u_{\mu_*}}) \leq \delta \Rightarrow \mathscr{L}(u') - \mathscr{L}(u_{\mu_*}) \geq c\mathrm{d}^2(u', \mathscr{O}_{u_{\mu_*}})\right). \end{aligned} \tag{4.6}$$

Note that this immediately implies that H_{μ_*} attains a local minimum on $\mathscr{O}_{u_{\mu_*}}$.

To show (4.5), note that the three vectors

$$e_1 = \begin{pmatrix} p \\ -\left(\frac{\sigma^*}{\rho^*}\right)^2 q \end{pmatrix}, \quad e_2 = \begin{pmatrix} q \\ -p \end{pmatrix}, \quad e_3 = \begin{pmatrix} p \\ q \end{pmatrix}, \tag{4.7}$$

form an orthogonal basis of $T_v \Sigma_{\mu_*}$, for each point $v = (q,p) \in \mathscr{O}_{u_*}$; e_1 is easily seen to be tangent to $\mathscr{O}_{u_*}$, so that e_2 and e_3 span $(T_v \mathscr{O}_{u_*})^{\perp} \cap T_v \Sigma_{\mu_*}$. A simple but tedious computation then shows that the matrices of $D_v^2(\mu_* \cdot L)$ and of $D_v^2 H$ in this basis are

$$D_v^2(\mu_* \cdot L) = \begin{pmatrix} 2\sigma_*^4 & 0 & \mu_*^2\left[\left(\frac{\sigma_*}{\rho_*}\right)^2 - 1\right] \\ 0 & -2\mu_*^2 & 0 \\ \mu_*^2\left[\left(\frac{\sigma_*}{\rho_*}\right)^2 - 1\right] & 0 & -2\mu_*^2 \end{pmatrix}$$

and

$$D_v^2 H = \begin{pmatrix} V'(\rho_*)\rho_*^{-1}\sigma_*^2 & 0 & (V'(\rho_*)\rho_*^{-1} - 1)\sigma_*^2 \\ 0 & V''(\rho_*)\rho_*^2 + \sigma_*^2 & 0 \\ (V'(\rho_*)\rho_*^{-1} - 1)\sigma_*^2 & 0 & V'(\rho_*)\rho_*^{-1}\sigma_*^2 + \rho_*^2 \end{pmatrix}$$

The estimate (4.5) now follows immediately from the hypothesis that $V''(\rho_*)\rho_*^2 + 3\sigma_*^2 > 0$.

We now turn to the proof of (4.6). Let $u' \in \Sigma_{\mu_*}$. Then there exists $v' \in \mathscr{O}_{u_*}$ so that $\mathrm{d}(v', \mathscr{O}_{u_*}) = \|u' - v'\|$ and as a result, one has that $u' - v' \in (T_{v'} \mathscr{O}_{u_*})^{\perp}$. We can write

$$u' = u' - v' + v' = v' + (u' - v')_{\|} + (u' - v')_{\perp}.$$

Here $(u' - v')_{\perp}$ is perpendicular to $T_{v'}\Sigma_{\mu_*}$, and $(u' - v')_{\|}$ belongs to $T_{v'}\Sigma_{\mu_*}$ and is perpendicular to $T_{v'}\mathscr{O}_{u_*}$ since $u' - v'$ is. Now remark that, since $D_{v'}L((u' - v')_{\|}) = 0$, and since $u', v' \in \Sigma_{\mu_*}$,

$$0 = L(u') - L(v') = D_{v'}L((u' - v')_{\perp}) + \mathrm{O}(\|u' - v'\|^2). \tag{4.8}$$

It is easily checked that, for each $v' \in \mathscr{O}_{u_{\mu_*}}$, the restriction of $D_{v'}L$ to $(T_{v'}\Sigma_{\mu_*})^\perp$ is an isomorphism. It follows that there exists a constant C so that

$$\|(u'-v')_\perp\| \leq C\|(u'-v')\|^2. \tag{4.9}$$

Note that this constant is independent of $v' \in \mathscr{O}_{\mu_*}$ since, for all $R \in \mathrm{SO}(3)$, and for all $u \in \mathbb{R}^6$,

$$\Phi_R \circ D_u L \circ \Phi_{R^{-1}} = D_{\Phi_R u} L,$$

where Φ_R, defined in (2.17), is an isometry. Returning to (4.8), and using this last remark, we conclude there exists a constant c_0 so that, for $\|u'-v'\|$ small enough, one has

$$\|(u'-v')_\parallel\| \geq \|u'-v'\| - \|(u'-v')_\perp\| \geq c_0\|u'-v'\|. \tag{4.10}$$

We can now conclude the proof of (4.6) as follows, using (4.9), (4.10) and (4.5):

$$\begin{aligned}
\mathscr{L}(u') - \mathscr{L}(u_{\mu_*}) &= \mathscr{L}(u') - \mathscr{L}(v') \\
&= D_{v'}\mathscr{L}(u'-v') + \frac{1}{2}D^2_{v'}\mathscr{L}(u'-v',u'-v') + \mathrm{o}(\|u'-v'\|^2) \\
&= \frac{1}{2}D^2_{v'}\mathscr{L}((u'-v')_\parallel,(u'-v')_\parallel) + \mathrm{O}(\|u'-v'\|^3) + \mathrm{o}(\|u'-v'\|^2) \\
&= \frac{1}{2}D^2_{v'}\mathscr{L}((u'-v')_\parallel,(u'-v')_\parallel) + \mathrm{o}(\|u'-v'\|^2) \\
&\geq \frac{1}{2}c\|(u'-v')_\parallel\|^2 + \mathrm{o}(\|u'-v'\|^2) \\
&\geq \tilde{c}\|u'-v'\|^2 = \tilde{c}\,\mathrm{d}^2(u',\mathscr{O}_{u_{\mu_*}}).
\end{aligned}$$

Remark that as before, the constant c is independent of $v' \in \mathscr{O}_{\mu_*}$. This shows (4.6). Note that we used the boundedness of $D^2_{v'}\mathscr{L}$, uniformly in $v' \in \mathscr{O}_{u_{\mu_*}}$.

We can now prove orbital stability, namely:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall u' \in \mathbb{R}^6, \quad \left(\mathrm{d}(u',\mathscr{O}_{u_{\mu_*}}) \leq \delta \Rightarrow \forall t \in \mathbb{R}, \mathrm{d}(u'(t),\mathscr{O}_{u_{\mu_*}}) \leq \varepsilon\right). \tag{4.11}$$

For that purpose, we propose three different arguments.

First argument. We proceed by contradiction, as before. Suppose there exists $\varepsilon_0 > 0$ and for each $n \in \mathbb{N}$, $u'_n \in \mathbb{R}^6$ and $t_n \in \mathbb{R}$ such that $\mathrm{d}(u'_n,u_{\mu_*}) \leq \frac{1}{n}$ and $\mathrm{d}(u'_n(t_n),\mathscr{O}_{u_{\mu_*}}) = \varepsilon_0$. We can suppose, without loss of generality, that $2\varepsilon_0 < \delta$, where δ is given in (4.6). We know that $\mathscr{L}(u'_n(t_n)) = \mathscr{L}(u'_n)$, since $\mathscr{L}$ is a constant of the motion. Hence

$$\lim_{n\to+\infty} \mathscr{L}(u'_n(t_n)) = \mathscr{L}(u_{\mu_*}) = \mu_*.$$

Since the orbit $\mathscr{O}_{u_{\mu_*}}$ is bounded, and since $\mathrm{d}(u'_n(t_n),\mathscr{O}_{u_{\mu_*}}) = \varepsilon_0$, it follows that the sequence $u'_n(t_n)$ is bounded; we can therefore conclude that $\lim_{n\to+\infty}\mathrm{d}(u'_n(t_n),\Sigma_{\mu_*}) = 0$. (In other words $\mathscr{L}$ satisfies Hypothesis F, see Lemma 7.2.) As a consequence, there exist $w_n \in \Sigma_{\mu_*}$ so that $\|w_n - u'_n(t_n)\| \to 0$. We can now conclude. Since, for n large enough, $\frac{\varepsilon_0}{2} \leq \mathrm{d}(w_n,\mathscr{O}_{u_{\mu_*}}) \leq \frac{3}{2}\varepsilon_0$, we have

$$\begin{aligned}
\mathscr{L}(u'_n) - \mathscr{L}(u_{\mu_*}) &= \mathscr{L}(u'_n(t_n)) - \mathscr{L}(u_{\mu_*}) \\
&= \mathscr{L}(u'_n(t_n)) - \mathscr{L}(w_n) + \mathscr{L}(w_n) - \mathscr{L}(u_{\mu_*}) \\
&\geq \mathscr{L}(u'_n(t_n)) - \mathscr{L}(w_n) + c\mathrm{d}^2(w_n,\mathscr{O}_{u_{\mu_*}}).
\end{aligned}$$

The sequences $u'_n(t_n)$ and w_n are bounded. This, combined with the uniform continuity of $\mathscr{L}$ on bounded sets, leads again to a contradiction upon taking $n \to +\infty$.

Second argument. The second proof uses the fact that the relative equilibrium u_{μ_*}, which gives rise to a circular orbit, belongs to a continuous family $\mu \to u_\mu$ of such equilibria, defined on a neighbourhood $I \subset \mathbb{R}^3$ of μ_*. We will only sketch the argument, the general case is treated in Theorem 7.5. One first observes that, for μ belonging to a suitably small neighbourhood of μ_*, both (4.5) and (4.6) hold, with μ_* replaced by μ, and with μ-independent c and δ. This allows one to prove that the equilibria u_μ are orbitally stable with respect to perturbations of the initial condition *within* Σ_μ, that is:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall u' \in \Sigma_\mu, \quad \left(\mathrm{d}(u', \mathscr{O}_{u_\mu}) \leq \delta \Rightarrow \forall t \in \mathbb{R}, \mathrm{d}(u'(t), \mathscr{O}_{u_\mu}) \leq \varepsilon\right). \tag{4.12}$$

Indeed, suppose that this is not true. Then there exists $\varepsilon_0 > 0$, and for each $n \in \mathbb{N}^*$, $u_n' \in \Sigma_\mu$, $t_n \in \mathbb{R}$ so that $\mathrm{d}(u_n', u_\mu) \leq \frac{1}{n}$ and $\mathrm{d}(u_n'(t_n), \mathscr{O}_{u_\mu}) = \varepsilon_0$. Since we can choose $\varepsilon_0 < \delta$, we can apply (4.6) to write

$$\mathscr{L}(u_n') - \mathscr{L}(u_\mu) = \mathscr{L}(u_n'(t_n)) - \mathscr{L}(u_\mu) \geq c\mathrm{d}(u_n'(t_n), \mathscr{O}_{u_\mu})^2 = c_* \varepsilon_0^2.$$

Taking $n \to +\infty$ leads to the desired contradiction. It remains to prove (4.12) with "$\forall u' \in \Sigma_\mu$" replaced by "$\forall u \in \mathbb{R}^6$." For that purpose, note that, if $u' \in \mathbb{R}^6$ is close to u_{μ_*}, then $\mu = L(u')$ is close to μ_* and hence u_μ close to u_{μ_*}. So u' is close to u_μ. Hence $u'(t)$ remains close at all times to $\mathscr{O}_{u_\mu}$ by (4.12). Now, since $\mathscr{O}_\mu$ is close to $\mathscr{O}_{\mu_*}$, the result follows.

Third argument. If (4.5) had been valid for all $w \in (T_v \mathscr{O}_{u_{\mu_*}})^\perp$, the first argument above would have been slightly easier, since we could then have mimicked the proof of Proposition 4.1 directly. As it stands, we were able to first show (4.6), which is valid only for $v' \in \Sigma_{\mu_*}$ and which shows $\mathscr{L}$, restricted to Σ_{μ_*}, attains a local minimum on the orbit. This immediately implies an orbital stability result for perturbations u' of the initial condition u_{μ_*} that stay within Σ_{μ_*}, as is readily seen. But to obtain a stability result for arbitrary perturbations $u' \in \mathbb{R}^6$ of the initial condition u_{μ_*}, we had to work a little harder and invoke Hypothesis F, which is not necessarily easily verified in infinite dimensional problems. It turns out that (4.5) is *not* valid[8] for all $w \in (T_v \mathscr{O}_{u_{\mu_*}})^\perp$. However, it is possible to adjust the Lyapunov function $\mathscr{L}$ so that this *is* the case. Consider, for all $K > 0$,

$$\mathscr{L}_K(u) = \mathscr{L}(u) + K(L(u) - \mu_*)^2. \tag{4.13}$$

Note that the additional term vanishes on Σ_{μ_*}, where $\mathscr{L}_K$ reaches an absolute minimum. We now show

$$\exists \hat{c} > 0, K > 0, \forall v \in \mathscr{O}_{u_{\mu_*}}, \forall w \in \left(T_v \mathscr{O}_{u_{\mu_*}}\right)^\perp, \quad D_v^2 \mathscr{L}_K(w,w) \geq \hat{c}\|w\|^2. \tag{4.14}$$

For that purpose, introduce, for each $v = (q,p) \in \mathscr{O}_{u_{\mu_*}}$,

$$e_4 = \begin{pmatrix} \hat{q} \wedge \hat{p} \\ 0 \end{pmatrix}, \quad e_5 = \begin{pmatrix} 0 \\ \hat{q} \wedge \hat{p} \end{pmatrix}, \quad e_6 = \frac{1}{\sqrt{\rho_*^2 + \sigma_*^2}} \begin{pmatrix} \sigma_* \hat{q} \\ \rho_* \hat{p} \end{pmatrix}, \tag{4.15}$$

which, together with e_1, e_2, e_3 in (4.7) form an orthonormal basis of $\mathbb{R}^6$. Clearly, $D_v(L - \mu_*)^2(w) = 0$, for all $v \in \mathscr{O}_{u_{\mu_*}}$ and for all $w \in \mathbb{R}^6$. Moreover, if $\eta_1, \eta_2, \eta_3 \in \mathbb{R}^3$ form an orthonormal basis, then

$$D_v^2(L - \mu_*)^2(w,w) = 2\sum_{i=1}^{3} \left[D_v(\eta_i \cdot L)(w)\right]^2,$$

[8] This can be seen from a straightforward computation, which is most readily made in the basis e_i introduced in (4.7) and (4.15).

with

$$D_v(\eta_i \cdot L)(w) = w_1 \cdot (p \wedge \eta_i) + w_2 \cdot (\eta_i \wedge q), \quad w = (w_1, w_2) \in \mathbb{R}^6.$$

Now, writing $w = \sum_{j=2}^{6} \alpha_j e_j \in (T_v \mathscr{O}_{u_{\mu_*}})^\perp$ and using $\eta_1 = \hat{q}, \eta_2 = \hat{p}, \eta_3 = \hat{q} \wedge \hat{p}$, we find

$$\begin{aligned} D_v^2(L-\mu_*)^2(w,w) &= 2\left[\alpha_4^2\sigma_*^2 + \alpha_5^2\rho_*^2 + \alpha_6^2(\rho_*^2+\sigma_*^2)\right] \\ &\geq 2\min\{\sigma_*^2,\rho_*^2\}\left[\alpha_4^2+\alpha_5^2+\alpha_6^2\right]. \end{aligned} \tag{4.16}$$

We can now conclude the proof of (4.14) as follows. We write $w = w_A + w_B$ with $w_A = \alpha_2 e_2 + \alpha_3 e_3$ and $w_B = \alpha_4 e_4 + \alpha_5 e_5 + \alpha_6 e_6$. Then there exists a constant $C > 0$, independent of $v \in \mathscr{O}_{u_{\mu_*}}$, so that

$$\begin{aligned} D_v^2\mathscr{L}_K(w,w) &\geq D_v^2\mathscr{L}(w,w) + 2K\min\{\sigma_*^2,\rho_*^2\}\|w_B\|^2 \\ &\geq D_v^2\mathscr{L}(w_A,w_A) + 2K\min\{\sigma_*^2,\rho_*^2\}\|w_B\|^2 - C\left[\|w_A\|\|w_B\| + \|w_B\|^2\right]. \end{aligned}$$

Using (4.5), one finds that, for all $m > 0$,

$$D_v^2\mathscr{L}_K(w,w) \geq \left(c - \frac{Cm^2}{2}\right)\|w_A\|^2 + \left(2K\min\{\sigma_*^2,\rho_*^2\} - C - \frac{C}{2m^2}\right)\|w_B\|^2,$$

where we have applied Young's inequality to the term $\|w_A\|\|w_B\|$. Choosing m small enough and K large enough, one finds (4.14). We can now prove the following statement, which is to be compared to (4.6): $\exists \delta, c > 0$ so that, for all $u' \in \mathbb{R}^6$,

$$\mathrm{d}(u', \mathscr{O}_{u_{\mu_*}}) \leq \delta \Rightarrow \mathscr{L}_K(u') - \mathscr{L}_K(u_{\mu_*}) \geq c^2\mathrm{d}^2(u', \mathscr{O}_{u_{\mu_*}}). \tag{4.17}$$

Indeed, for all $u' \in \mathbb{R}^6$, there exists $v' \in \mathscr{O}_{u_{\mu_*}}$ so that $u' - v' \in (T_{v'}\mathscr{O}_{u_{\mu_*}})^\perp$. Hence

$$\mathscr{L}_K(u') - \mathscr{L}(u_{\mu_*}) = \mathscr{L}(u') - \mathscr{L}_K(v') \geq \frac{\hat{c}}{2}\|u'-v'\|^2 + \mathrm{O}(\|u'-v'\|^3).$$

This implies (4.17), from which orbital stability follows by the now familiar argument. □

We point out that the core ingredient of all three arguments in the proof is estimate (4.5). Its proof constitutes the only truly model-dependent part of the proofs of orbital stability via the energy-momentum method. This will become clear in Section 7 where we will show how a suitably adapted version of this estimate implies orbital stability in a general infinite dimensional setting as well (Theorem 7.4, Theorem 7.5, Theorem 7.10).

As a second remark, note that (4.6) allows one to prove immediately the orbital stability for perturbations of the initial condition that preserve the angular momentum. The three strategies of the proof above therefore concern three different methods for extending this result to arbitrary perturbations of the initial condition. The same structure of the proof will be apparent in the general situation treated in Section 7.

The first argument in the above proof is used in [GSS87] and [GSS90]. It has the disadvantage of using Hypothesis F, which, while obvious in finite dimensions, may not always be easy to check in infinite dimensional systems, notably when the group G_{μ_*} is not one-dimensional and in particular when it is not commutative (as in [GSS90]). It has the advantage of not using the fact that the relative equilibrium under consideration belongs to a continuous family.

The second argument is used for example in [Wei86], and in [GH07b, GH07a]. For this argument the existence of a continuous family of relative equilibria is needed but not Hypothesis F.

The third argument is commonly used in the literature on finite dimensional Hamiltonian systems [Pat92], and appears also in [Stu08] in the infinite dimensional case. It is not

universally useable, since it depends on the existence of a G_μ-invariant Euclidean structure on the dual of the Lie-algebra of G, as we will see in Section 7.

5. Hamiltonian dynamics in infinite dimension

The modern formulation of Hamiltonian dynamics has been adapted to the framework of infinite dimensional Banach manifolds in [CM74, MR94]. This approach is not well suited for our purposes for two reasons. First, we are interested in flows defined by the solutions to (nonlinear) partial differential equations that are defined on Banach (or even Hilbert) spaces, for which a general Banach manifold formulation is overly complex. In addition, the notions of "Hamiltonian vector field" and "Hamiltonian flow" introduced in [CM74] seem too general for the purpose of studying stability questions. We therefore present a simpler and more restricted framework that is well adapted to the analysis of the stability questions that are our main focus, including for nonlinear Schrödinger and wave equations.

Our main goal in this section is thus to give a workable and not too complex definition of "Hamiltonian dynamical system" or of "Hamiltonian flow" in the infinite dimensional Banach space setting (Section 5.2). The formalism allows us to easily obtain general results on the link between symmetries and conserved quantities for such systems, as in the finite dimensional case (Section 5.3). This link is indeed an essential ingredient for the identification of relative equilibria and the construction of coercive Lyapunov functions in Hamiltonian systems with symmetry, as we shall explain in Section 6. Several examples of Hamiltonian PDE's that fit in our framework are given in Section 5.5. Although this section is self-contained, the reader unfamiliar with finite dimensional Hamiltonian dynamical systems and their symmetries may find it useful to consult Appendix A.3 for a concise and self-contained treatment of this case. We will make regular use of the notation and concepts introduced there.

5.1. Symplectors, symplectic Banach triples, symplectic transformations, Hamiltonian vector fields

We first generalize the notion of symplectic form to the infinite dimensional setting and introduce the equivalent notion of *symplector* (Definition 5.2). It turns out that, in the infinite dimensional setting, it is convenient to treat the latter as the central object of the theory, rather than the symplectic form itself, as is customary in finite dimensions. As we will see, the two approaches are perfectly equivalent.

We need some preliminary terminology. Let E be a Banach space and $B : E \times E \to \mathbb{R}$ a bilinear continuous form. We can then define, in the usual manner, for all $u \in E$, $\mathscr{J}_B u \in E^*$ via

$$\mathscr{J}_B u(v) = B(u,v).$$

It follows easily that $\mathscr{J}_B : u \in E \to \mathscr{J}_B u \in E^*$ is linear and continuous, with $\|\mathscr{J}_B\| = \|B\|$. We will write $\mathscr{R}_{\mathscr{J}_B} = \operatorname{Ran} \mathscr{J}_B$. Conversely, given a continuous linear map $\mathscr{J} : E \to E^*$, one can construct $B_{\mathscr{J}}(u,v) = (\mathscr{J} u)(v)$. We introduce the following terminology:

Definition 5.1. A bilinear continuous form B is non-degenerate (or weakly non-degenerate) if $\mathscr{J}_B$ is injective. It is strongly non-degenerate if $\mathscr{J}_B$ is both injective and surjective. Similarly, a linear map $\mathscr{J} : E \to E^*$ is said to be (weakly) non-degenerate if it is injective, and strongly non-degenerate if it is a bijection.

Definition 5.2. We now introduce the notion of symplector.[9]

(i) A symplector or weak symplector is a continuous linear map $\mathscr{J} : E \to E^*$ that is injective and anti-symmetric, in the sense that

$$(\mathscr{J} u)(v) = -(\mathscr{J} v)(u).$$

If in addition $\mathscr{J}$ is surjective, we say it is a strong symplector.

(ii) A (strong) symplectic form ω is a (strongly) non-degenerate bilinear continuous form that is anti-symmetric.

(iii) When $\mathscr{J}$ is a (strong) symplector, we will say $(E, \mathscr{J})$ is a (strong) symplectic vector space, or simply that E is a (strong) symplectic vector space, when there is no ambiguity about the choice of $\mathscr{J}$.

There clearly is a one-to-one correspondence between (strong) symplectors and (strong) symplectic forms. Note that the definition implies that

$$\forall \alpha, \beta \in \mathscr{R}_{\mathscr{J}}, \quad \alpha(\mathscr{J}^{-1}\beta) = -\beta(\mathscr{J}^{-1}\alpha). \tag{5.1}$$

The following examples of (strong) symplectors cover all applications we have in mind in these notes. Let $\mathscr{K}$ be a real Hilbert space and set $E = \mathscr{K} \times \mathscr{K}$. Then

$$\mathscr{J} : (q,p) \in E \to (-p,q) \in E^*$$

is clearly a strong symplector. Here we wrote $u = (q,p) \in \mathscr{K} \times \mathscr{K}$ and used the Riesz identification of E with E^*. The corresponding strong symplectic form is

$$\omega_{\mathscr{J}}(u,u') = q \cdot p' - q' \cdot p,$$

where $\cdot$ denotes the inner product on $\mathscr{K}$. The analogy with (A.3.1) is self-evident: there $\mathscr{K} = \mathbb{R}^n$, where $\mathbb{R}^n$ is equipped with its standard Euclidean structure. Note that if Q is a bounded self-adjoint operator on $\mathscr{K}$ with $\mathrm{Ker} Q = \{0\}$, then

$$\mathscr{J} : (q,p) \in E \to (-Qp, Qq) \in E^*$$

is also a symplector with

$$\omega_{\mathscr{J}}(u,u') = q \cdot Qp' - p \cdot Qq'.$$

We will need the following straightforward generalization of the above construction. Let K^2 be a positive (possibly and typically unbounded) self-adjoint operator on $\mathscr{K}$, with domain $\mathscr{D}(K)$. Introduce, for all $s \in \mathbb{R}$, $\mathscr{K}_s = [\mathscr{D}(\langle K \rangle^s)]$, where $\langle K \rangle = \sqrt{1+K^2}$ and where $\langle K \rangle^s$ is defined by the functional calculus of self-adjoint operators. Here $[\mathscr{D}(\langle K \rangle^s)]$ denotes the closure of $\mathscr{D}(\langle K \rangle^s)$ in the topology induced by the Hilbert norm

$$\|u\|_s := \|\langle K \rangle^s u\|.$$

Note that, since $\langle K \rangle^s : (\mathscr{D}(\langle K \rangle^s), \|\cdot\|_s) \to (\mathscr{D}(\langle K \rangle^{-s}), \|\cdot\|)$ is an isometric bijection, it extends to a unitary map from $\mathscr{K}_s$ to $\mathscr{K}$ for which we still write $\langle K \rangle^s$. With these conventions, we can then make the usual identification between $\mathscr{K}_s^*$ and $\mathscr{K}_{-s}$: $\forall v \in \mathscr{K}_{-s}$, we define

$$u \in \mathscr{K}_s \to v \cdot u \in \mathbb{R},$$

by setting $v \cdot u := \langle K \rangle^{-s} v \cdot \langle K \rangle^s u$. Note that

$$\forall s, s' \in \mathbb{R}, \quad s \le s' \Rightarrow \mathscr{K}_{s'} \subset \mathscr{K}_s.$$

It is easy to see using the spectral theorem that this is an inclusion as sets, and we will therefore not introduce explicit identification operators to represent such inclusions which

[9] This object does not seem to have been blessed with a name in the literature, so we took the liberty to baptize it.

are moreover continuous for the respective Hilbert space topologies. The typical example of this construction to keep in mind is $K^2 = -\Delta$ on $\mathscr{K} = L^2(\mathbb{R}^d)$. We then have $\mathscr{K}_s = H^s(\mathbb{R}^d)$, the usual Sobolev spaces.

For $s = (s_1, s_2) \in \mathbb{R}^2$, we define $E_s = \mathscr{K}_{s_1} \times \mathscr{K}_{s_2}$. Defining a partial order relation by $s \preceq s'$ iff $s_1 \leq s_1'$ and $s_2 \leq s_2'$, we have

$$\forall s, s' \in \mathbb{R}^2, \quad s \preceq s' \Rightarrow E_{s'} \subset E_s.$$

Setting $\bar{s} = (s_2, s_1)$ we then define

$$\mathscr{J}_s : u = (q, p) \in E_s \to (-p, q) \in E_{\bar{s}}. \tag{5.2}$$

The following lemma is now immediate.

Lemma 5.3. *$\mathscr{J}_s$ is a weak symplector if and only if $s_1 \geq -s_2$. In that case*

$$\mathscr{J}_s : u = (q, p) \in E_s \to (-p, q) \in E_{\bar{s}} \subset E_{-s} = E_s^*.$$

We have $\mathscr{R}_s := \mathscr{R}_{\mathscr{J}_s} = E_{\bar{s}}$. And $\mathscr{J}_s^{-1} = \mathscr{J}_{-s|E_{\bar{s}}}$. If K^2 is unbounded, $\mathscr{J}_s$ is a strong symplector if and only if $s_1 = -s_2$.

Typical examples of this construction are the use of $E = E_{(1/2,-1/2)}$ or of $E = E_{(1,0)}$ with $\mathscr{K} = L^2(\mathbb{R}^d)$ and $K^2 = -\Delta$ to study the wave equation. For the Schrödinger equation, $E = E_{(1,1)}$ is a natural choice. We refer to Section 5.5 for the details of these examples. Note that of these three examples, only the first corresponds to a strong symplector and hence to a strong symplectic form. It is therefore clear that the use of weak symplectors is unavoidable in applications to PDE's.

We end our discussion of symplectors with a simple lemma that collects some of their essential properties.

Lemma 5.4. *Let E be a Banach space and $\mathscr{J} : E \to E^*$ be a bounded linear map. Then the following holds:*

(i) If $\mathscr{J}$ is a strong symplector, then $\mathscr{J}^{-1}$ is bounded.
(ii) If $\mathscr{J}$ is injective and (anti-)symmetric, and if E is reflexive, then $\mathscr{R}_{\mathscr{J}}$ is dense in E^*.
(iii) Suppose $\mathscr{J}$ is injective and (anti-)symmetric, and that its inverse is bounded on $\mathscr{R}_{\mathscr{J}}$. Suppose E is reflexive. Then $\mathscr{R}_{\mathscr{J}} = E^*$.

Proof. (i) This is a consequence of the open mapping theorem.
(ii) Suppose $v \in E$ satisfies $\mathscr{J}u(v) = 0$ for all $u \in E$. Then $\mathscr{J}v(u) = 0$ for all $u \in E$, by (anti-)symmetry. Hence $\mathscr{J}v = 0$ and hence, since $\mathscr{J}$ is injective, $v = 0$. Since E is reflexive, this means that, if $v \in E^{**}$ vanishes on $\mathscr{R}_{\mathscr{J}} \subset E^*$, then $v = 0$. This implies $\mathscr{R}_{\mathscr{J}}$ is dense (Hahn-Banach).
(iii) Since the inverse is bounded, $\mathscr{R}_{\mathscr{J}}$ is closed. The result then follows from (ii).
□

If E is not reflexive, a symplector may not have a dense range, as the following example[10] shows. Let

$$E = \{u \in L^1(\mathbb{R}, \mathrm{d}x) \mid \int_{\mathbb{R}} u(x)\mathrm{d}x = 0\} \subset L^1(\mathbb{R})$$

and define

$$\mathscr{J}u(x) = \int_{-\infty}^{x} u(y)\mathrm{d}y \in L^\infty(\mathbb{R}) \subset E^*.$$

[10] Due to S. Keraani.

This is clearly bounded, injective and antisymmetric. But it is clear that

$$\| \mathscr{J} u - 1 \|_\infty \geq 1,$$

for all $u \in E$. So the range is not dense in $L^\infty(\mathbb{R})$ and a fortiori not dense in E^*.

We are now ready to define what we mean by a symplectic transformation and by a Hamiltonian vector field. First we recall a very basic definition: when $F : E_1 \to E_2$ is a function between two Banach spaces E_1 and E_2, and when $u \in E_1$, one says that F is (Fréchet) differentiable at u if there exists $D_u F \in \mathscr{L}(E_1, E_2)$ so that

$$\lim_{w \to 0} \frac{\|F(u+w) - F(u) - D_u F(w)\|_{E_2}}{\|w\|_{E_1}} = 0.$$

Also, one says that $F : E_1 \to E_2$ is differentiable on some subset of E_1 if for all u in that subset, F is differentiable in the above sense.

In particular, if $E_1 = E, E_2 = \mathbb{R}$, and if F is differentiable at $u \in E$, we have $D_u F \in E^*$. And if $\mathscr{D}$ is a domain in E, saying that $F : E \to \mathbb{R}$ *is differentiable on* $\mathscr{D}$ means that F is differentiable at each $u \in \mathscr{D}$. In that case, one can define

$$u \in \mathscr{D} \subset E \to D_u F \in E^*.$$

As a last comment, we stress that, in these definitions, the only topology used is the one on E. This is important to keep in mind in the applications, where the domain $\mathscr{D}$ often carries a natural topology, stronger than the one induced by the norm on E, and for which $\mathscr{D}$ is closed. One can think of $E = H^1(\mathbb{R})$ and $\mathscr{D} = H^3(\mathbb{R})$. Such a topology is NOT used in the above statements, nor in the following general definition. We refer to the examples treated in Sections 5.4 and 5.5 for several illustrations of this last comment.

Definition 5.5. Let E be a Banach space, $\mathscr{D}$ a domain in E (See Section 2.1) and $\mathscr{J}$ a symplector.

(i) We will refer to $(E, \mathscr{D}, \mathscr{J})$ as a symplectic Banach triple.

(ii) Let $(E, \mathscr{D}, \mathscr{J})$ be a symplectic Banach triple and $\Phi \in C^0(E,E) \cap C^1(\mathscr{D},E)$. We say Φ is a symplectic transformation if

$$\forall u \in \mathscr{D}, \forall v, w \in E, (\mathscr{J} D_u\Phi(v))(D_u\Phi(w)) = (\mathscr{J} v)(w). \tag{5.3}$$

(iii) We say that a function $F : E \to \mathbb{R}$ has a $\mathscr{J}$-compatible derivative if F is differentiable on $\mathscr{D}$ and if, for all $u \in \mathscr{D}$, $D_u F \in \mathscr{R}_{\mathscr{J}}$. In that case we write $F \in \mathrm{Dif}(\mathscr{D}, \mathscr{J})$.

(iv) For each $F \in \mathrm{Dif}(\mathscr{D}, \mathscr{J})$, the Hamiltonian vector field $X_F : \mathscr{D} \subset E \to E$ associated to F is defined by

$$X_F(u) = \mathscr{J}^{-1} D_u F, \quad \forall u \in \mathscr{D}. \tag{5.4}$$

The analogy between (5.3) and (A.3.17) as well as between (5.4) and (A.3.10) is evident. Note however that, when dealing with *weak* symplectors, as is often the case in applications, the vector field X_F does not inherit the continuity or smoothness properties that F may enjoy. In particular, even if

$$D.F : \mathscr{D} \subset E \to E^*$$

is continuous, the same may not hold for X_F. We shall for that reason avoid making use of the vector fields X_F where possible and state all our hypotheses in terms of F directly. We finally point out that, here and in what follows, and unless otherwise specified, all functions we consider are globally defined[11] on E.

[11] This is a difference with [CM74], as we will explain in some detail in Section 5.4.

5.2. **Hamiltonian flows and constants of the motion**

Definition 5.6. Let $(E, \mathscr{D}, \mathscr{J})$ be a symplectic Banach triple. Let $F \in \mathrm{Dif}(\mathscr{D}, \mathscr{J})$. A Hamiltonian flow for F is a separately continuous map $\Phi^F : \mathbb{R} \times E \to E$ with the following properties:

(i) For all $t, s \in \mathbb{R}$, $\Phi^F_{t+s} = \Phi^F_t \circ \Phi^F_s$, $\Phi^F_0 = \mathrm{Id}$;
(ii) For all $t \in \mathbb{R}$, $\Phi^F_t(\mathscr{D}) = \mathscr{D}$;
(iii) For all $u \in \mathscr{D}$, the curve $t \in \mathbb{R} \to u(t) := \Phi^F_t(u) \in \mathscr{D} \subset E$ is differentiable and is the unique solution of

$$\mathscr{J}\dot{u}(t) = D_{u(t)}F, \quad u(0) = u. \tag{5.5}$$

Local Hamiltonian flows are defined in the usual way. We refer to (5.5) as the Hamiltonian differential equation associated to F (Compare to (A.3.11) and (A.3.4)) and to its solutions as Hamiltonian flow lines. Note that in this setting separate continuity implies continuity (See [CM74], Section 3.2). We refer to Section 5.5 for examples of PDE's generating Hamiltonian flows.

Observe that (5.5) implies that, for all $u \in \mathscr{D}$,

$$\forall \alpha \in \mathscr{R}_{\mathscr{J}}, \quad -\frac{\mathrm{d}}{\mathrm{d}t}\alpha(u(t)) = D_{u(t)}F(\mathscr{J}^{-1}\alpha), \tag{5.6}$$

which is a weak form of (5.5). With this in mind, one could think of changing Definition 5.6 by replacing (iii) by the following alternative statement[12]:

(iii') For all $u \in E$, the curve $t \in \mathbb{R} \to u(t) := \Phi^F_t(u) \in E$ belongs to $C(\mathbb{R}, E)$ and (5.6) holds.

This has the advantage of eliminating the introduction of the domain $\mathscr{D}$ (and therefore of condition (ii)) and is precisely the definition of "solution" to (5.5) used in [GSS87, GSS90]. In [Stu08], E is a Hilbert space and still a different formulation is adopted. Basically, the domain $\mathscr{D}$ is not introduced, the equation (5.5) is interpreted as an equation in E^* and the time derivative is understood as a strong derivative for E^*-valued functions. Those alternative formulations do not allow for a direct proof of the kind of natural "conservation theorems" such as Theorem 5.8 below, that are typical for Hamiltonian systems and that we need for the stability analysis. As a result, the conclusions of such conservation theorems are added as assumptions in the general setup of the cited works. It turns out that, in examples, the proof of such assumptions requires a stronger notion of "solution" than the ones used in [GSS87, GSS90, Stu08], so we found it more efficient to adopt from the start the stronger notion of Hamiltonian flow found in Definition 5.6.

Definition 5.7. Let $F, G \in \mathrm{Dif}(\mathscr{D}, \mathscr{J})$. Then the Poisson bracket of F and G is defined by

$$\{F, G\}(u) = D_u F(\mathscr{J}^{-1} D_u G), \qquad \forall u \in \mathscr{D}. \tag{5.7}$$

Equation (5.7) is the obvious transcription of (A.3.12) to the infinite dimensional setting. We now have the following crucial result, which is a simple form of Noether's Theorem in the Hamiltonian setting. A more complete form follows below (Theorem 5.11).

Theorem 5.8. *Let $(E, \mathscr{D}, \mathscr{J})$ be a symplectic Banach triple. Let $H, F : E \to \mathbb{R}$ be differentiable on E and suppose they have a $\mathscr{J}$-compatible derivative,* i.e. *$H, F \in \mathrm{Dif}(\mathscr{D}, \mathscr{J})$. Suppose there exist Hamiltonian flows Φ^H_t, Φ^F_t for H and F. Then:*

[12] Note that for this formulation one needs $F \in \mathrm{Dif}(E, \mathbb{R})$, but it is not necessary that it has a $\mathscr{J}$-compatible derivative.

(i) For all $u \in \mathscr{D}$, and for all $t \in \mathbb{R}$,

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\Phi_t^F(u)) = \{H,F\}(\Phi_t^F(u)). \tag{5.8}$$

(ii) The following three statements are equivalent:

(a) For all $u \in \mathscr{D}$, $\{F,H\}(u) = 0$.

(b) For all $u \in E$, and for all $t \in \mathbb{R}$,

$$(H \circ \Phi_t^F)(u) = H(u). \tag{5.9}$$

(c) For all $u \in E$, and for all $t \in \mathbb{R}$,

$$(F \circ \Phi_t^H)(u) = F(u). \tag{5.10}$$

In this result, the roles of H and F are interchangeable. But in practice, one of the flows, say Φ_t^F, is simple, explicitly known, and often linear, whereas Φ_t^H is obtained by integrating a possibly nonlinear PDE of some complexity, such as the nonlinear Schrödinger or wave equations. It is then often very easy to check by a direct computation that $H \circ \Phi_t^F$ is constant in time for all $u \in E$: one says that H is invariant under the flow Φ_t^F, or that the Φ_t^F are symmetries of H. The important conclusion of the theorem is that this implies that F is a constant of the motion for Φ_t^H. This is a strong statement, since in applications, the flow Φ_t^H is complex and poorly known. So being able to assert that it leaves the level surfaces of F invariant is a non-trivial piece of information. Several examples are given in Section 5.5.

Proof. (i) Let $u \in \mathscr{D}$. Then $t \in \mathbb{R} \to H(\Phi_t^F(u)) \in E$ is differentiable and the chain rule applies: writing $u(t) = \Phi_t^F(u)$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t} H(\Phi_t^F(u)) = D_{\Phi_t^F(u)} H(\dot{u}(t)),$$

which yields the first equality in (5.8) since $\mathscr{J}\dot{u}(t) = D_{u(t)}F$. Exchanging the roles of H and F, the second equality follows similarly.

(ii) That (5.9) or (5.10) imply $\{H,F\}(u) = 0$ for $u \in \mathscr{D}$ is immediate from (i). Conversely, it follows from (i) and the fact that $\{H,F\}(u) = 0$, for all $u \in \mathscr{D}$, that $(H \circ \Phi_t^F)(u) = H(u)$. Since $\mathscr{D}$ is dense in E, $H \in C(E,\mathbb{R})$ and $\Phi_t^F \in C(E,E)$, (b) now follows for all $u \in E$. Similarly for (c).

□

It should be noted that condition (ii) of Definition 5.6 is crucial here. We are assuming there is a common invariant domain for both flows. To obtain conservation theorems of the above type without such an assumption requires other technical conditions [CM74].

We end with some technical remarks. First, it follows from Theorem A.3.5 in the Appendix, that Hamiltonian flows Φ_t^F are symplectic as soon as $F \in C^2(E,E)$ and $\Phi^F \in C^2(E,E)$. But these two assumptions (especially the latter) are generally too strong to be of use in infinite dimensional dynamical systems generated by PDE's, except possibly when they are linear. Of course, one can conceive of weaker conditions that imply the result. For efforts in that direction, we refer to [CM74]. In other words, proving that Hamiltonian flows, as defined above, are symplectic, can be painful. A second, related issue is the following. In finite dimensional systems, we know that, if $\{F_1,F_2\} = 0$, with $F_1,F_2 \in C^2(E)$, then the corresponding Hamiltonian flows commute: see (A.3.14) and Lemma A.1.1. This is a very useful fact: indeed, computing a Poisson bracket is a routine matter of taking derivatives, and the information obtained about the flows is very strong. Again, this is not immediate in infinite dimensional systems under reasonable conditions. For our purposes,

and in particular for the proof of Theorem 5.11, the following analog of Lemma A.3.4 will suffice.

Lemma 5.9. *Let $(E,\mathscr{D},\mathscr{J})$ be a symplectic Banach triple. Let Φ be a C^1-diffeomorphism on E and suppose that $\Phi(\mathscr{D})=\mathscr{D}$ and that Φ is symplectic. Let $F\in C^1(E,\mathbb{R})\cap \mathrm{Dif}(\mathscr{D},\mathscr{J})$ and let X_F be its Hamiltonian vector field. (See Definition 5.5 (iv)). Then, for all $u\in\mathscr{D}$*

$$D_u\Phi(X_{F\circ\Phi}(u))=X_F(\Phi(u)). \tag{5.11}$$

Moreover, for all $t\in\mathbb{R}$,

$$\Phi\circ\Phi_t^{F\circ\Phi}\circ\Phi^{-1}=\Phi_t^F. \tag{5.12}$$

In particular, if $F\circ\Phi=F$, then Φ commutes with Φ_t^F, for all $t\in\mathbb{R}$. And if Φ commutes with Φ_t^F, for all $t\in\mathbb{R}$, then there exists $c\in\mathbb{R}$ so that $F\circ\Phi=F+c$.

The proof is identical to that of Lemma A.3.4 and we omit it. The point here is that we suppose Φ to be a symplectic transformation. As we just saw, that is a strong assumption. In practice, to avoid the difficulties just mentioned, we will always *assume* that the symmetry group of the system under consideration acts with symplectic transformations. Since the latter are often linear, that they are symplectic can then be checked through a direct computation. We finally point out that, if one wanted to exploit the presence of a formal constant of the motion with a nonlinear flow, such as in completely integrable systems, it could in general be difficult to prove it acts symplectically and commutes with the dynamics. This, in turn, makes it difficult to exploit such formal constants of the motion in the stability analysis that is our main interest here.

5.3. **Symmetries and Noether's Theorem**

When dealing with a symplectic Banach triple, the appropriate type of group action to consider is the following.

Definition 5.10. Let $(E,\mathscr{D},\mathscr{J})$ be a symplectic Banach triple. Let G be a Lie group and $\Phi:(g,x)\in G\times E\to\Phi_g(x)\in E$, an action of G on E. We will say Φ 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 $g\in G$, $\Phi_g(\mathscr{D})=\mathscr{D}$.
(iii) For all $\xi\in\mathfrak{g}$, there exists $F_\xi\in C^1(E,\mathbb{R})\cap\mathrm{Dif}(\mathscr{D},\mathscr{J})$ such that $\Phi_{\exp(t\xi)}=\Phi_t^{F_\xi}$, and the map $\xi\to F_\xi$ is linear.

This definition reduces to Definition A.3.7 in the Appendix, for finite dimensional spaces E: in that case $\mathscr{D}=E$ and the restriction that $F\in\mathrm{Dif}(\mathscr{D},\mathscr{J})$ is superfluous. We can now state the version of Noether's Theorem that we need. It links the invariance group of Hamiltonian dynamics to constants of the motion and is to be compared to the finite dimensional version given in the appendix (Theorem A.3.9). As in (A.3.22), we will identify $\mathfrak{g}$ and $\mathfrak{g}^*$ with $\mathbb{R}^m$ and view F as a map $F:E\to\mathbb{R}^m$ (See (A.3.24)). This allows us to write

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

where $\cdot$ refers to the canonical inner product on $\mathbb{R}^m$.

Theorem 5.11. *Let $(E,\mathscr{D},\mathscr{J})$ be a symplectic Banach triple. Let G be a Lie group and Φ a globally Hamiltonian action of G on E. Let $H\in C^1(E,\mathbb{R})\cap\mathrm{Dif}(\mathscr{D},\mathscr{J})$ and let Φ_t^H be the corresponding Hamiltonian flow. Suppose that*

$$\forall g\in G,\quad H\circ\Phi_g=H. \tag{5.13}$$

Then:

(i) For all $\xi \in \mathfrak{g}$, $\{H, F_\xi\} = 0$.
(ii) For all $t \in \mathbb{R}$, $F_\xi \circ \Phi_t^H = F_\xi$.
(iii) G *is an invariance group*[13] *for* Φ_t^H.

This is an immediate consequence of Theorem 5.8 and Lemma 5.9. In the applications, the result is used as follows. The action Φ of G is simple and well known. It is then easy to check (5.13) directly. One then concludes that (ii) and (iii) hold, which are the important pieces of information for the further analysis. In particular, the level surfaces Σ_μ, defined in (2.5) are invariant under the dynamics Φ_t^H. Examples are given in the next section. The result in [CM74] that is closest in spirit to Theorem 5.11 is Theorem 2 of Section 6.2.

5.4. **Linear symplectic flows**

Since invariance groups often act linearly on the symplectic Banach space $(E, \mathscr{J})$, and since the nonlinear dynamical flows studied often are perturbations of linear ones, it is important to have a good understanding of linear symplectic flows. Their study also sheds some light on the various technical difficulties mentioned above, and in particular on the role of the domain $\mathscr{D}$, the definition of Hamiltonian flow we adopted, etc.

Proposition 5.12 below (which corresponds to Theorem 2 in Section 2.3 of [CM74]) characterizes all strongly continuous linear symplectic one-parameter groups on a symplectic Banach space in terms of their generators. We adopt the following notation. Given a strongly continuous group of linear transformations on E, we denote its generator by A, with domain $\mathscr{D}(A)$. By the Hille-Yosida theorem, we then know that $t \in \mathbb{R} \to u(t) = \Phi_t u \in E$ satisfies

$$\dot{u}(t) = Y_A(u(t)), \tag{5.14}$$

provided $u \in \mathscr{D}(A)$, where we introduced the vector field

$$Y_A : u \in \mathscr{D}(A) \subset E \to Au \in E.$$

Note that Y_A is not continuous if A is an unbounded operator. Clearly, the Φ_t form a dynamical system as defined in Section 2. We introduce the function

$$H_A : u \in \mathscr{D}(A) \to H_A(u) = \tfrac{1}{2}\omega_{\mathscr{J}}(Au, u) \in \mathbb{R}.$$

Observe that H_A admits directional (or Gâteaux) derivatives $\delta_u H_A(v)$, for all $u, v \in \mathscr{D}(A)$:

$$\begin{aligned} \delta_u H_A(v) &= \lim_{t\to 0} \frac{1}{t}\left(H_A(u+tv) - H_A(u)\right) \\ &= \frac{1}{2}\left(\omega_{\mathscr{J}}(Av, u) + \omega_{\mathscr{J}}(Au, v)\right). \end{aligned}$$

Nevertheless, if A is an unbounded operator, H_A is not continuous since, for all $u, w \in \mathscr{D}(A)$

$$H_A(u+w) - H_A(u) = \omega_{\mathscr{J}}(Au, w) + \omega_{\mathscr{J}}(Aw, u) + \omega_{\mathscr{J}}(Aw, w)$$

and the last term in particular does not necessarily converge to 0 as $w \to 0$ in the topology of E. It follows that, a fortiori, H_A is not Fréchet differentiable.

Proposition 5.12. *Let* $(E, \mathscr{J})$ *be a symplectic vector space. Let* Φ_t *be a strongly continuous one-parameter group of bounded linear operators on* E. *Let* $(A, \mathscr{D}(A))$ *be the generator of* Φ_t. *Then the following are equivalent.*

(i) The Φ_t *are symplectic,* i.e. $\omega_{\mathscr{J}}(\Phi_t u, \Phi_t v) = \omega_{\mathscr{J}}(u, v)$ *for all* $u, v \in E$;

[13] See Definition 2.1

(ii) For all $u, v \in \mathscr{D}(A)$,

$$\omega_{\mathscr{J}}(Au, v) = -\omega_{\mathscr{J}}(u, Av);$$

(iii) For all $u \in \mathscr{D}(A)$, one has

$$\mathscr{J} Y_A(u) = \delta_u H_A \in E^*. \tag{5.15}$$

In this case, $\delta_u H_A(v) = \omega_{\mathscr{J}}(Au, v)$, $H_A(\Phi_t u) = H_A(u)$ for all $u \in \mathscr{D}(A)$ and for all $t \in \mathbb{R}$.

Proof. The three equivalences are obvious. To prove H_A is a constant of the motion, it suffices to remember that the Hille-Yosida theorem implies $A\Phi_t u = \Phi_t A u$ provided $u \in \mathscr{D}(A)$. □

In other words, when the Φ_t are symplectic, the equation of motion (5.14) can be rewritten

$$\mathscr{J}\dot{u}(t) = \delta_{u(t)} H_A, \tag{5.16}$$

which is to be compared to (5.5). Clearly, the symplectic linear flows considered here are NOT Hamiltonian in the sense of Definition 5.6. Still, (5.16) gives meaning to the idea that in infinite dimension as well, linear strongly continuous symplectic flows are of "Hamiltonian nature," with a quadratic Hamiltonian. Moreover, the Hamiltonian H_A is a constant of the motion for the flow Φ_t. But note that, whereas in (5.9), the conservation of energy holds for all $u \in E$, this makes no sense here, since H_A is only defined on $\mathscr{D}(A)$.

Generally, because of the appearance of the Gâteaux derivative rather than a Fréchet differential in the right hand side, it turns out that the above formulation is inadequate for various reasons. For example, the absence of a chain rule for Gâteaux derivatives does not allow one to compute derivatives such as $\frac{\mathrm{d}}{\mathrm{d}t} H_A(u(t))$ directly to prove H_A is constant along the motion. In fact, in the proof above, this result is proven using the Hille-Yosida theorem, and without computing a derivative at all. This approach cannot work for non-linear flows of course. Similar problems arise when dealing with other constants of the motion than the Hamiltonian himself, even in the linear case, due to various domain questions and the complications in defining commutators. Finally, for our purposes, we need to restrict the motion to the level sets of the constants of the motion, and to use their manifold structure. This requires sufficient smoothness, a property not guaranteed at all by Gâteaux differentiability alone. Again, as pointed out before, an approach to the resolution of these technical difficulties other than the one chosen here can be found in [CM74].

In applications to PDE's, the function spaces that occur naturally are often complex Hilbert spaces. To make the link with Hamiltonian dynamics, one then proceeds as follows. Let $\mathscr{H}$ be a complex Hilbert space and let us write $\langle \cdot, \cdot \rangle$ for its inner product. First, it is clear that $\mathscr{H}$ is a real Hilbert space for the real inner product defined by $\mathrm{Re}\langle \cdot, \cdot \rangle$, which induces the same topology on $\mathscr{H}$ as the original inner product since both inner products have the same associated norm. Let us write E for this real Hilbert space. We now identify E^* with E using the corresponding Riesz isomorphism. Note that this is not the same as identifying $\mathscr{H}^*$ with $\mathscr{H}$ through the Riesz isomorphism associated to $\langle \cdot, \cdot \rangle$ and that there is no natural identification between $\mathscr{H}^*$ and E^* as sets: each non-zero element of $\mathscr{H}^*$ necessarily takes complex values, whereas the elements of E^* take real values only.

On the real Hilbert space E, one checks readily that

$$\omega(u, v) = \mathrm{Im}\langle u, v \rangle \in \mathbb{R}$$

defines a strong symplectic form. Note in particular that ω is real bilinear, but not complex bilinear. To identify the corresponding symplector $\mathscr{J} : E \to E$ in a convenient manner[14],

[14] We identified E^* with E, so the symplector can be seen as a map from E to E.

one proceeds as follows:

$$\omega(u,v) = \mathrm{Re}\,\langle iu, v\rangle$$

so that $\mathscr{J}u = iu$. The reader should not let itself be confused by the fact that we write iu, while considering u as an element of the real vector space E. The way to see this is as follows: the real vector space E is, as a set, identical to $\mathscr{H}$. And on $\mathscr{H}$, multiplication by i is well defined and actually an isometric complex linear map. So multiplication by i is well defined on E as an isometric real linear map.

To sum up, we showed how to associate to a complex Hilbert space $(\mathscr{H}, \langle\cdot,\cdot\rangle)$ a real Hilbert space $(E, \langle\cdot,\cdot\rangle_E)$ with symplectic structure

$$\omega(u,v) = \langle \mathscr{J}u, v\rangle_E, \quad \mathscr{J}u = iu.$$

Now let us return to the linear symplectic flows. Suppose B is a self-adjoint operator on $\mathscr{H}$, with domain $\mathscr{D}(B)$. Then $U_t = \exp(-iBt)$ is a strongly continuous one-parameter group of unitaries[15] . The corresponding Hille-Yosida generator is $A = -iB$, with $\mathscr{D}(A) = \mathscr{D}(B)$. Clearly, each U_t is a symplectic transformation on E with the symplectic form ω. We are therefore in the setting of Proposition 5.12 and

$$H_A(u) = \frac{1}{2}\langle u, Bu\rangle.$$

It turns out that in the applications we have in mind, the one parameter subgroups of the symmetry group G act with such unitary groups on the relevant Hilbert space $\mathscr{H}$. But within this framework, as we pointed out above, the U_t are NOT Hamiltonian flows. To remedy this situation, one can, and we will, proceed along the following lines. First remark that the function H_A above is C^1 if we view it as a function on the Banach space E_B obtained by considering on $\mathscr{D}(|B|^{1/2})$ the graph norm. And that the flow U_t is strongly differentiable on $\mathscr{D} := \mathscr{D}(|B|^{3/2})$, viewed as a subset of E_B. So now we are in the setting of Definition 5.6, and U_t is a Hamiltonian flow on E_B, on which $\mathscr{J}$ still defines a weak symplector. The trouble with this reformulation so far is that now the Banach space E_B and the domain $\mathscr{D}$ depend on B. If the symmetry group is multi-dimensional, it will have several generators, and we need a common domain and Banach space on which to realize them all as Hamiltonian flows. We will see several examples where this formalism is implemented.

In practice, very often, $\mathscr{H} = \mathscr{K}^{\mathbb{C}} = \mathscr{K} \oplus i\mathscr{K}$, where $\mathscr{K}$ is a real Hilbert space. One has $u = q + ip \in \mathscr{H}$ with $q, p \in \mathscr{K}$. Then, clearly $E = \mathscr{K} \times \mathscr{K}$ with its natural Hilbert space structure. Moreover, identifying $u \in \mathscr{H}$ with $(q,p) \in \mathscr{K} \times \mathscr{K}$, clearly $\mathscr{J}(q,p) = (-p,q)$ and we are back to the examples of symplectors given in Section 5.1.

5.5. **Hamiltonian PDE's: examples**

In this section we give some examples of PDE's generating Hamiltonian flows in the sense of Definition 5.6.

Let $E = H^1(\mathbb{R}^d, \mathbb{C})$, $\mathscr{D} = H^3(\mathbb{R}^d, \mathbb{C})$ and consider the nonlinear Schrödinger equation

$$\begin{cases} i\partial_t u(t,x) + \Delta u(t,x) + \lambda |u(t,x)|^{\sigma-1} u(t,x) = 0 \\ u(0,x) = u_0(x) \end{cases} \tag{5.17}$$

introduced in Section 2.4, defined on $\mathbb{R}^d$, $d = 1,2,3$. For $d = 1$, suppose that $3 \le \sigma < +\infty$ in the defocusing case and $3 \le \sigma < 5$ in the focusing case. In dimension $d = 2,3$, consider only the defocusing case and assume $3 \le \sigma < 1 + \frac{4}{d-2}$. Let $\Phi_t^X : E \to E$ be the global flow

[15] By Stone's theorem, every strongly continuous one parameter group of unitaries is of this form.

defined in (2.35). Recall that the existence of Φ_t^X is ensured by Theorem 2.5 and, thanks to Theorem 2.6, $\Phi_t^X(\mathscr{D}) = \mathscr{D}$ for all $t \in \mathbb{R}$.

Our purpose is to show that Equation (5.17) is the Hamiltonian differential equation associated to the function H defined by (2.32) and $\Phi_t^X = \Phi_t^H$ for all $t \in \mathbb{R}$.

As explained in the end of Section 5.4, we usually identify $u = q + ip \in H^s(\mathbb{R}^d, \mathbb{C})$ with $(q,p) \in H^s(\mathbb{R}^d, \mathbb{R}) \times H^s(\mathbb{R}^d, \mathbb{R})$ for all $s \in \mathbb{R}$. Hence, let $(E, \mathscr{D}, \mathscr{J})$ be the symplectic Banach triple given by

$$\begin{aligned} E &= H^1(\mathbb{R}^d, \mathbb{R}) \times H^1(\mathbb{R}^d, \mathbb{R}), \\ \mathscr{D} &= H^3(\mathbb{R}^d, \mathbb{R}) \times H^3(\mathbb{R}^d, \mathbb{R}), \\ \mathscr{J}(q,p) &= (-p, q), \ \forall (q,p) \in E. \end{aligned}$$

Clearly $\mathscr{J} u = iu$ and $\mathscr{R}_{\mathscr{J}} = E \subset E^*$. Now consider

$$H(q,p) = \frac{1}{2}\left(\|\nabla q\|_{L^2}^2 + \|\nabla p\|_{L^2}^2\right) - \frac{\lambda}{\sigma+1}\int_{\mathbb{R}^d} (|q|^2 + |p|^2)^{\frac{\sigma+1}{2}},$$

and remark that if we write $u = q + ip$ with $(q,p) \in E$, $H(u) = H(q,p)$ is exactly the energy defined in (2.32). A straightforward calculation, using the Sobolev embedding theorem, shows that $H \in C^2(E, \mathbb{R})$. In particular,

$$D_{(q,p)}H = (-\Delta q, -\Delta p) - \lambda (|q|^2 + |p|^2)^{\frac{\sigma-1}{2}}(q,p) \in E^*$$

which can be written as

$$D_u H = -\Delta u - \lambda |u|^{\sigma-1} u$$

in terms of $u = q + ip$. Next, using the fact that the Sobolev space $H^3(\mathbb{R}^d)$ is an algebra for $d = 1, 2, 3$, we have $DH(\mathscr{D}) \subset \mathscr{R}_{\mathscr{J}}$ so that H has a $\mathscr{J}$-compatible derivative on $\mathscr{D}$.

Moreover, the curve $(q(t), p(t)) = \Phi_t^X(q,p)$ is the unique solution to

$$\mathscr{J}(\dot{q}(t), \dot{p}(t)) = (-\Delta q, -\Delta p) - \lambda (|q|^2 + |p|^2)^{\frac{\sigma-1}{2}}(q,p) = D_{(q(t),p(t))}H$$

that is Equation (5.5). As a consequence, Φ^X is a Hamiltonian flow for H in the sense of Definition 5.6, $\Phi_t^X = \Phi_t^H$ and the nonlinear Schrödinger equation (5.17) is a Hamiltonian differential equation.

In Section 2.4, we prove directly from the equation that $G = \mathrm{SO}(d) \times \mathbb{R}^d \times \mathbb{R}$ with the action defined by (2.36) is an invariance group for the dynamics. In general, the action of this group is not globally Hamiltonian. Nevertheless, let us consider the subgroup $\tilde{G} = \mathbb{R}^d \times \mathbb{R}$ and the restricted action

$$\begin{aligned} \Phi : \tilde{G} \times E &\to E \\ (a, \gamma, u) &\to \Phi_{a,\gamma}(u) = e^{i\gamma} u(x - a). \end{aligned} \tag{5.18}$$

For all $g \in \tilde{G}$, $\Phi_g \in C^1(E,E)$ is symplectic, $\Phi_g(\mathscr{D}) = \mathscr{D}$ and for all

$$\xi = (\xi_1, \ldots, \xi_d, \xi_{d+1}) \in \mathfrak{g},$$

point (iii) of Definition 5.10 is satisfied by taking $F_{\xi_j} = \xi_j F_j$ with

$$F_j(u) = -\frac{i}{2}\int_{\mathbb{R}^d} \bar{u}(x) \partial_{x_j} u(x)\, \mathrm{d}x \ \forall j = 1, \ldots, d, \tag{5.19}$$

$$F_{d+1}(u) = -\frac{1}{2}\int_{\mathbb{R}^d} \bar{u}(x) u(x)\, \mathrm{d}x. \tag{5.20}$$

As a consequence the action Φ of $\tilde{G}$ on E is globally Hamiltonian. Moreover, in Section 2.4, we showed that $H \circ \Phi_g = H$, hence we may apply Theorem 5.11 and conclude that $F_{\xi_j} \circ \Phi_t^H = F_{\xi_j}$ that means that each F_j is a constant of the motion.

Finally we show that the action $\Phi : (R,u) \in G \times E \to \Phi_R(u) = u(R^{-1}x) \in E$ of $G = \mathrm{SO}(d)$ on E is not globally Hamiltonian. For simplicity, let us consider $d = 2$ and let us identify a matrix $\xi \in \mathrm{so}(2)$ with $\xi \in \mathbb{R}$

$$\xi = \begin{pmatrix} 0 & \xi \\ -\xi & 0 \end{pmatrix}.$$

Then for each $\xi \in \mathbb{R}$, $\Phi_{\exp(t\xi)} = \Phi_t^{F_\xi}$ with $F_\xi = \xi F$ and

$$F(u) = -\frac{i}{2} \int_{\mathbb{R}^d} (x_1 \partial_{x_2} - x_2 \partial_{x_1}) u(x) \bar{u}(x) \, \mathrm{d}x.$$

The issue is that F is not even well-defined on the Banach space $H^1(\mathbb{R}^2)$!

Finally, let us remark that if we choose $\mathscr{D} = H^2(\mathbb{R}^d) \times H^2(\mathbb{R}^d)$, then $DH(\mathscr{D}) \subset L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \not\subset \mathscr{R}_{\mathscr{J}} = H^1(\mathbb{R}^d) \times H^1(\mathbb{R}^d)$ and H does not have a $\mathscr{J}$-compatible derivative for this new choice of $\mathscr{D}$. In the same way, if we take $E = L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ and $\mathscr{D} = H^1(\mathbb{R}^d) \times H^1(\mathbb{R}^d)$, the same function H is not even continuous.

Next, let $(E, \mathscr{D}, \mathscr{J})$ be the symplectic Banach triple given

$$\begin{aligned} E &= H^1(\mathbb{R}^d, \mathbb{R}) \times L^2(\mathbb{R}^d, \mathbb{R}), \\ \mathscr{D} &= H^2(\mathbb{R}^d, \mathbb{R}) \times H^1(\mathbb{R}^d, \mathbb{R}), \\ \mathscr{J}(q,p) &= (-p,q), \ \forall (q,p) \in E. \end{aligned}$$

and consider the nonlinear wave equation

$$\begin{cases} \partial_{tt}^2 u(t,x) - \Delta u(t,x) + \lambda |u(t,x)|^{\sigma-1} u(t,x) = 0 \\ u(0,x) = u_0(x), \partial_t u(0,x) = u_1(x) \end{cases} \tag{5.21}$$

introduced in Section 2.6, defined on $\mathbb{R}^d$, $d = 1,2,3$. Suppose $\lambda > 0$ and σ an odd integer such that $3 \leq \sigma < +\infty$ in dimension $d = 1$ and $3 \leq \sigma < 1 + \frac{4}{d-2}$ for $d = 2,3$. Let $\Phi_t^X : E \to E$ the global flow defined in (2.53). Thanks to the persistence of regularity, we have $\Phi_t^X(\mathscr{D}) = \mathscr{D}$ for all $t \in \mathbb{R}$ (see Section 2.6).

As before, our purpose is to show that Equation (5.21) is the Hamiltonian differential equation associated to the function H defined by (2.52) and $\Phi_t^X = \Phi_t^H$ for all $t \in \mathbb{R}$.

First of all, note that $\mathscr{R}_{\mathscr{J}} = L^2(\mathbb{R}^d) \times H^1(\mathbb{R}^d) \subset E^* = H^{-1}(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$. Next, consider

$$H(q,p) = \frac{1}{2} \left(\|\nabla q\|_{L^2}^2 + \|p\|_{L^2}^2 \right) + \frac{\lambda}{\sigma+1} \int_{\mathbb{R}^d} (|q|)^{\sigma+1},$$

and remark that if we write $q = u$ and $p = \partial_t u$ with $(q,p) \in E$, $H(u) = H(q,p)$ is exactly the energy defined in (2.52). As for the nonlinear Schrödinger equation, a straightforward calculation, using the Sobolev embedding theorem, shows that $H \in C^2(E, \mathbb{R})$. In particular,

$$D_{(q,p)} H = (-\Delta q + \lambda |q|^{\sigma-1} q, p) \in E^*.$$

Next, using the fact that the Sobolev space $H^2(\mathbb{R}^d)$ is an algebra for $d = 1,2,3$, we have $DH(\mathscr{D}) \subset \mathscr{R}_{\mathscr{J}}$ so that H has a $\mathscr{J}$-compatible derivative on $\mathscr{D}$.

Moreover, the curve $(u(t),\partial_t u(t)) = \Phi_t^X(u(0),\partial_t u(0))$ is the unique solution to (5.21). As a consequence, using $u = q$ and $\partial_t u = p$, we have that $(q(t),p(t)) = \Phi_t^X(q,p)$ is the unique solution to

$$\mathscr{J}(\dot{q}(t),\dot{p}(t)) = (-\Delta q + \lambda|q|^{\sigma-1}q, p) = D_{(q(t),p(t))}H,$$

that is, Equation (5.5). Finally, if $(q,p)\in\mathscr{D}$, the curve $t\to\Phi_t^H(q,p)\in C(\mathbb{R},\mathscr{D})\cap C^1(\mathbb{R},E)$. As a consequence, Φ^X is a Hamiltonian flow for H in the sense of Definition 5.6, $\Phi_t^X = \Phi_t^H$ and the nonlinear wave equation (5.21) is a Hamiltonian differential equation.

6. IDENTIFYING RELATIVE EQUILIBRIA

We now dispose of the necessary tools that will allow us to characterize the relative equilibria of Hamiltonian systems with symmetry and that will yield the candidate Lyapunov function to study their stability. Before stating the main result (Theorem 6.1), we recall some of the terminology used below, but refer to the appendices for details. First, for $\mu\in\mathfrak{g}^*$, we have (see (A.2.12)),

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

$\mathfrak{g},\mathfrak{g}_\mu$ are the Lie algebras of G and G_μ respectively, and $\mathfrak{g}^*,\mathfrak{g}_\mu^*$ their duals. We always identify $\mathfrak{g}^*$ with $\mathbb{R}^m$ (see (A.2.13)). Hence, if Φ is a globally Hamiltonian action, we think of its momentum map as a map $F: E\to\mathbb{R}^m$ and define, for all $\mu\in\mathbb{R}^m$,

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

We then know from Proposition A.3.11 that $G_\mu = G_{\Sigma_\mu}$ provided the momentum map is Ad^*-equivariant.

Theorem 6.1. *Let $(E,\mathscr{D},\mathscr{J})$ be a symplectic Banach triple. Let $H\in C^1(E,\mathbb{R})\cap\mathrm{Dif}(\mathscr{D},\mathscr{J})$ and suppose H has a Hamiltonian flow Φ_t^H. Let furthermore G be a Lie group, and Φ a globally Hamiltonian action on E with Ad^*-equivariant momentum map F. Suppose that,*

$$\forall g\in G,\quad H\circ\Phi_g = H. \tag{6.1}$$

(i) Then G is an invariance group for Φ_t^H.

(ii) Let $u\in E$ and let $\mu = F(u)\in\mathbb{R}^m\simeq\mathfrak{g}^*$. Consider the following statements:

(1) u is a relative G-equilibrium.

(2) u is a relative G_μ-equilibrium.

(3) There exists $\xi\in\mathfrak{g}_\mu$ so that, for all $t\in\mathbb{R}$,

$$\Phi_t^H(u) = \Phi_{\exp(t\xi)}(u). \tag{6.2}$$

(4) There exists $\xi\in\mathfrak{g}_\mu$ so that

$$D_uH - \xi\cdot D_uF = 0. \tag{6.3}$$

(5) There exists $\xi\in\mathfrak{g}$ so that

$$D_uH - \xi\cdot D_uF = 0. \tag{6.4}$$

Then (1) $\Leftrightarrow$ (2) $\Leftarrow$ (3).
If $u\in\mathscr{D}$, then (1) $\Leftrightarrow$ (2) $\Leftarrow$ (3) $\Leftrightarrow$ (4) $\Leftrightarrow$ (5).
If in addition, μ is a regular value of F (See Definition A.1.3), then
(1) $\Leftrightarrow$ (2) $\Leftarrow$ (3) $\Leftrightarrow$ (4) $\Leftrightarrow$ (5) $\Leftrightarrow$ (6), where (6) is the statement:
(6) u is a critical point of H_μ on Σ_μ, where $H_\mu = H|_{\Sigma_\mu}$.
In addition, ξ is then unique.

That (1) is equivalent to (2) is a particular feature of Hamiltonian systems. In fact, its statement makes no sense outside of the Hamiltonian setting. It implies that, if u is a G-relative equilibrium, it is automatically a relative equilibrium for the *smaller* group G_μ. So the relevant invariance group depends on the point u through the value $\mu = F(u)$ of the constants of the motion at u. This is important since, as we will see in Section 7, one then ends up showing u is G_μ-orbitally stable, which is a stronger result than G-orbital stability. We already saw examples of this mechanism in Section 4. The proof of the equivalence between (1) and (2), although very simple, uses the subtle relations between constants of the motion and symmetries for Hamiltonian systems explained in the previous section.

For our purposes, the most interesting information obtained in this result is the observation that if $u \in \mathscr{D}$ satisfies (6.3), sometimes referred to in the PDE literature as "the stationary equation", then it is a relative equilibrium. And that, if μ is a regular value of F, those solutions are precisely the critical values of H_μ. This means that, given a Hamiltonian system with symmetries, one can find relative equilibria by looking for critical points of the Hamiltonian H restricted to the surfaces Σ_μ. In practice, this can be done concretely by solving (6.4), which in applications to Hamiltonian PDE's often takes the form of a stationary PDE in which ξ is treated as a (vector valued) parameter. Examples are given in the following sections. See also Section 4 for examples in finite dimension.

One immediately suspects that the Lagrange theory of multipliers for the study of constrained extrema should be of relevance here. This is indeed the case: introducing, on E, the Lagrange function

$$\forall v \in E, \quad \mathscr{L}(v) = H(v) - \xi \cdot F(v), \tag{6.5}$$

one sees that (6.4) expresses the vanishing of its first variation at u: $D_u\mathscr{L} = 0$. Here, $\xi \in \mathfrak{g} \simeq \mathbb{R}^m$ plays the role of a Lagrange multiplier. From the experience gained with the examples given so far, one suspects that, to show u is a stable relative equilibrium, one could try proceeding in two steps. First, show u is not just a critical point, but actually a local minimum of H_μ by studying the second variation of the Lagrange function $\mathscr{L}$ on Σ_μ. Next, use the Lagrange function as Lyapunov function in the proof of stability. Indeed, $u \in \Sigma_\mu$ is a local minimum of H_μ if and only if

$$\exists \rho > 0, \forall v \in \Sigma_\mu, \quad \mathrm{d}(v,u) \leq \rho \Rightarrow H_\mu(v) - H_\mu(u) \geq 0,$$

which is equivalent to

$$\exists \rho > 0, \forall v \in \Sigma_\mu, \quad \mathrm{d}(v,u) \leq \rho \Rightarrow \mathscr{L}(v) - \mathscr{L}(u) \geq 0,$$

since F is constant on Σ_μ. This is clearly the strategy used in the proofs of Section 4. We will see in Section 7 how to implement it in a general setting and give examples from the nonlinear Schrödinger equation in Sections 8 and 9. This is the approach that goes by the name of *energy-momentum method*.

Proof. (i) This is an immediate consequence of Theorem 5.11 (iii).

(ii) (1) $\Leftrightarrow$ (2). If u is a relative G-equilibrium, then there exists, for each $t \in \mathbb{R}$, $g(t) \in G$ so that $\Phi_t^H(u) = \Phi_{g(t)}(u)$. Since $u \in \Sigma_\mu$, so is $\Phi_t^H(u)$, since F is a constant of the motion for H, by Theorem 5.11 (ii). Hence

$$\mu = F(u) = F(\Phi_t^H(u)) = F(\Phi_{g(t)}(u)) = \mathrm{Ad}^*_{g(t)}\mu.$$

It follows that $g(t) \in G_\mu$, which concludes the argument. The reverse implication is obvious.

(3) $\Rightarrow$ (2). Obvious from the definition.

Now suppose $u \in \mathscr{D}$.

(3) $\Leftrightarrow$ (4). Suppose (3) holds. Since $u \in \mathscr{D}$, this implies that $\mathscr{J}^{-1} D_u H = \mathscr{J}^{-1} D_u(\xi \cdot F)$, which implies (4). Now suppose (4) holds. Since $u \in \mathscr{D}$ and since $H \circ \Phi_t^H = H$ and $F_\xi \circ \Phi_t^H = F_\xi$ by Theorem 5.11 (ii), we have, for all $t \in \mathbb{R}$,

$$D_{\Phi_t^H(u)} H D_u \Phi_t^H = D_u H, \quad D_{\Phi_t^H(u)}(\xi \cdot F) D_u \Phi_t^H = D_u(\xi \cdot F).$$

Writing $u(t) = \Phi_t^H(u)$, this yields $D_{u(t)} H = D_{u(t)}(\xi \cdot F)$ so that $\mathscr{J}\dot{u}(t) = D_{u(t)}(\xi \cdot F)$, which shows $t \to u(t)$ is a flow line of the Hamiltonian $\xi \cdot F$, with initial condition u. Since the latter is unique, we find $u(t) = \Phi_t^{\xi \cdot F}(u)$, which concludes the argument since $\Phi_t^{\xi \cdot F} = \Phi_{\exp(t\xi)}$ (See Definition 5.10 (iii)).
(4) $\Leftrightarrow$ (5). We only need to establish that (5) implies (4). As above, (5) implies $u(t) = \Phi_{\exp(t\xi)}$. Hence

$$\mathrm{Ad}^*_{\exp(t\xi)} \mu = \mathrm{Ad}^*_{\exp(t\xi)} F(u) = (F \circ \Phi_{\exp(t\xi)})(u) = F(u(t)) = F(u) = \mu,$$

since $F_i \circ \Phi_t^H = F_i$. Hence $\xi \in \mathfrak{g}_\mu$.

Now suppose in addition μ is a regular value of F.
(4) $\Leftrightarrow$ (6). We remark that, since μ is a regular value of F, Σ_μ is a co-dimension m submanifold of E and (see (A.1.6))

$$T_u \Sigma_\mu = \{v \in E \mid D_u F(v) = 0\}.$$

Hence clearly (4) implies (6). Conversely, suppose $D_u H$ vanishes on $T_u \Sigma_\mu$. Since μ is a regular value of F, we know that $D_u F$ is onto $\mathbb{R}^m$. Let W be a subspace of E complementary to $T_u \Sigma$, so that $E = T_u \Sigma \oplus W$. It follows $\dim W = m$ and that the m one-forms $D_u F_i \in W^*$, $i = 1, \ldots m$ form a basis of W^*. Consequently, the restriction of $D_u H$ to W can be written uniquely as $D_u H = \sum_{i=1}^m \xi_i D_u F_i = D_u(\xi \cdot F)$. Since both sides vanish on $T_u \Sigma_\mu$, (4) follows.

□

We conclude this section with two technical remarks that can be skipped in a first reading.

Remark 6.2. First, we have seen that (3) implies (2). Under suitable technical conditions, the reverse is also true. This can be understood as follows. If $u \in \mathscr{D}$ is a G_μ-relative equilibrium then, for all $t \in \mathbb{R}$, there exists $g(t) \in G_\mu$ so that $u(t) = \Phi_t^H u = \Phi_{g(t)} u$. So the curve

$$t \in \mathbb{R} \to \Phi_t^H(u) \in G_\mu u := \{\Phi_g(u) \mid g \in G_\mu\} \subset E$$

is a smooth curve on the group orbit $G_\mu u$. Under appropriate topological conditionson G_μ and G_u (defined in (2.7)), and if the action Φ of the group G_μ is sufficiently smooth[16], this orbit is an immersed submanifold of Σ_μ that can be identified with the homogeneous space G_μ / G_u, and its tangent space at u is therefore

$$T_u(G_\mu u) = \{X_{F_\xi}(u) \mid \xi \in \mathfrak{g}_\mu\}.$$

We recall that X_{F_ξ} is the Hamiltonian vector field associated to the function $F_\xi = \xi \cdot F$. Since $X_H(u) = \frac{\mathrm{d}}{\mathrm{d}t} \Phi_t^H(u)_{|t=0} \in T_u(G_\mu u)$, it follows that there exists $\xi \in \mathfrak{g}_\mu$ so that

$$X_H(u) = X_{\xi \cdot F}(u),$$

which is equivalent to (6.3) and therefore implies (3). We refer to [AM78, LM87] for the detailed argument, in the finite dimensional setting. We shall not have a need for the

[16] See for example Section 4 of [AM78], and in particular Corollary 4.1.22.

implication (2) $\Rightarrow$ (3), but will point out that, "morally", there is a one-one relationship between the critical points of H_μ and the relative equilibria of the Hamiltonian flow Φ_t^H.

Remark 6.3. Second, what is the role of the condition that μ be a regular value of F? This has several consequences. First, it guarantees that Σ_μ is a co-dimension m submanifold of E and that $T_u\Sigma_\mu = \mathrm{Ker} D_uF$. This is convenient in the further stability analysis, as we will see. Second, if $u \in \mathscr{D}$ and $\mathrm{Rank}\, D_uF = m$, then $\xi \in \mathbb{R}^m \simeq \mathfrak{g} \to \Phi_1^{\xi\cdot F}(u) \in \mathscr{O}_u = Gu \subset E$ is a local immersion and the action is locally free, meaning that the isotropy group G_u of u is discrete. Hence $\xi \in \mathfrak{g}_\mu \to \Phi_1^{\xi\cdot F}(u) \in \mathscr{O}_u \cap \Sigma_\mu = G_\mu u \subset E$ is also a local immersion. This observation will be used in Lemma 7.9 in the next section. If μ is not regular, various additional technical difficulties arise in the stability analysis of the next section, even in finite dimensional settings, where they have been studied in [LS98, MRO11]. As an example of such a singular value μ, consider the action of $\mathrm{SO}(3)$ on $\mathbb{R}^6$ introduced in Section 2.3, on the level set $L(u) = \mu = 0$. The corresponding isotropy group G_μ is $\mathrm{SO}(3)$ itself in that case. Its action is not locally free, since G_u, for $u = (q, p)$, with q and p parallel, is the copy of $\mathrm{SO}(2)$ given by the rotations about the common axis of q and p. We will see another example of such a situation when treating the nonlinear Schrödinger equation on the torus in Section 8. In both these cases, the ensuing complication is easily dealt with on an ad hoc basis.

7. Orbital stability: an abstract proof

7.1. **Introduction: strategy**

We have seen that in many situations the relative equilibria of Hamiltonian systems with symmetry are precisely the critical points of the restriction H_μ of the Hamiltonian H to a level surface Σ_μ, for some $\mu \in \mathfrak{g}^*$, of the constants of the motion F associated to the symmetry group via the Noether Theorem. This at once explains why they tend to come in families u_μ, indexed by μ in some open subset of $\mathfrak{g}^* \simeq \mathbb{R}^m$. Indeed, considering equation (6.4), it is natural to think of it as an equation in which both ξ and u are unknown. And so, under suitable circumstances, one can hope to find a family of solutions u_ξ of (6.4) by letting ξ run through some neighbourhood inside $\mathfrak{g}$. Typically, as ξ changes, so does $\mu_\xi = F(u_\xi) \in \mathfrak{g}^*$. Depending on the situation, it may be more convenient to label the solutions by μ_ξ than by $\xi \in \mathfrak{g}$. In these notes, we use mostly μ as a parameter, except in Section 9 where ξ is used. The question of the existence of such families of relative equilibria – a problem related to bifurcation theory – is studied, in the finite dimensional setting, in [Mon97] and [LS98]. We already saw several examples of this phenomenon and more will be provided in Sections 8 and 9.

It remains to see how one can prove the orbital stability of those relative equilibria. The basic intuition is that – modulo technical problems – they should be stable if they are not just critical points, but actually local minima of H_μ. To understand the origin of this intuition, recall that, if $u_\mu \in \Sigma_\mu$ is a relative equilibrium of the Hamiltonian dynamics Φ_t^H, then the orbit $G_\mu u = \{\Phi_g(u) \mid g \in G_\mu\}$ of G_μ, viewed as an element of the orbit space Σ_μ/G_μ, is a fixed point of the reduced dynamics. And, since H_μ is invariant under the action of G_μ, it can be viewed as a function on this orbit space. If H_μ has a local minimum at u_μ, it thus has a local minimum at the orbit $G_\mu u \in \Sigma_\mu/G_\mu$. Finally, since H_μ is a constant of the motion for the reduced dynamics, we are precisely in the situation described in the introduction: $G_\mu u$ is a fixed point of the reduced dynamics, and H_μ is a constant of the motion for which $G_\mu u$ is a minimum. We can therefore hope to use the Lyapunov method

to prove the stability of $G_\mu u$. To do so, it would suffice to prove a coercive estimate of the type (1.2) for H_μ on Σ_μ/G_μ.

There are two obvious problems one has to face when trying to implement this strategy. First, even if one executes this program, one will have proven only that u_μ is orbitally stable with respect to perturbations v of u_μ with $v \in \Sigma_\mu$. But one would like to prove this is true for arbitrary perturbations $v \in E$. Second, it is difficult to work on the abstract quotient space Σ_μ/G_μ, which, even in finite dimensional systems, but particularly in infinite dimensional ones, may not have a nice topological or differentiable structure, so that analytical tools to prove estimates are not readily available. To deal with both these problems, the idea is to use the theory of constraint minimization and Lagrange multipliers. This has the obvious advantage that one can work in the ambient space E, which has the added redeeming feature of being linear. As already outlined in the dicussion following Theorem 6.1, it turns out that it is the Lagrange function

$$\mathscr{L}_\mu = H - \xi_\mu \cdot F$$

associated to the relative equilibrium u_μ (see (6.5)) that plays the role of Lyapunov function in the proofs. In practice, one uses a Taylor expansion to second order of $\mathscr{L}_\mu$ about points on the orbit $G_\mu u$ and controls the second derivative of $\mathscr{L}_\mu$ to prove it is a minimum; this in turn gives the necessary coercivity to conclude stability. The reader will have noticed that the above strategy was worked out in all detail in the simple example of motion in a spherical potential presented in Section 4.

In this section, we will provide a detailed implementation of the above strategy in the following general setup. We refer to Section 2 for the definitions of the objects introduced below.

HYPOTHESIS A

(i) E is a Banach space and $\mathscr{D}$ a domain in E.
(ii) Φ_t^X is a dynamical system on E with a vector field $X : \mathscr{D} \to E$.
(iii) $F \in C^2(E, \mathbb{R}^m)$ is a vector of constants of the motion for Φ_t^X with level surfaces $\Sigma_\mu, \mu \in \mathbb{R}^m$, as in (2.5).
(iv) Φ_t^X admits an invariance group G, with an action Φ of G on E.

Recall that if μ is a regular value for F then Σ_μ is a co-dimension m submanifold of E. In this setting, we consider relative equilibria of the following type.

Let $\mu \in \mathbb{R}^m$.

HYPOTHESIS Bμ

(i) There exists $u_\mu \in \Sigma_\mu$ which is a relative equilibrium of the dynamics for the group $G_{\Sigma_\mu} = \{g \in G \mid \Phi_g \Sigma_\mu = \Sigma_\mu\}$.
(ii) There exists $\mathscr{L}_\mu \in C(E, \mathbb{R})$ which is a G_{Σ_μ}- invariant constant of the motion.
(iii) There exist $\eta > 0, c > 0$ so that

$$\forall u \in \mathscr{O}_{u_\mu}, \forall u' \in \Sigma_\mu, \quad \mathrm{d}(u, u') \leq \eta \Rightarrow \mathscr{L}_\mu(u') - \mathscr{L}_\mu(u) \geq c\mathrm{d}^2(u', \mathscr{O}_{u_\mu}) \tag{7.1}$$

where

$$\mathscr{O}_{u_\mu} = \Phi_{G_{\Sigma_\mu}}(u) = \{\Phi_g(u) \mid g \in G_{\Sigma_\mu}\}. \tag{7.2}$$

Under the above conditions, we say $\mathscr{L}_\mu$ is a coercive Lyapunov function on $\mathscr{O}_{u_\mu}$ *along* Σ_μ. If the G_{Σ_μ}-action is isometric then it is enough to check (7.1) holds at one single point $u \in \mathscr{O}_{u_\mu}$. It will then hold everywhere, with the same η, c, as a result of the G_{Σ_μ}-invariance of $\mathscr{L}_\mu$. Isometric actions are common in applications and this is one of the places where

they provide a simplification. For what follows, the power 2 in the right hand side of (7.1) is of no consequence. One can generalize the definition by replacing the right hand side in (7.3) by $f(\mathrm{d}(u', \mathscr{O}_{u_\mu}))$, for some function $f : \mathbb{R}^+ \to \mathbb{R}^+$, $f(0) = 0$, $f(d) > 0$ if $d > 0$. In practice, as we will see below, one gets the lower bound in (7.1) from a Taylor expansion of $\mathscr{L}$, so that the square appears naturally. We point out that condition (iii) in Hypothesis Bμ implies (i). Indeed, if $u \in \mathscr{O}_{u_{\mu_*}}$ and $u' = u(t')$ for small enough t', then (iii) implies that

$$0 = \mathscr{L}_{\mu_*}(u(t')) - \mathscr{L}_{\mu_*}(u) \geq c\mathrm{d}^2(u(t'), \mathscr{O}_{u_{\mu_*}}).$$

We have however found it convenient to keep this redundancy in the statement of the hypothesis.

We point out that Hypotheses A and Bμ are formulated without imposing the dynamical system to be Hamiltonian. Nor do they impose any link between the symmetry group G, the constants of the motion F and the Lyapunov function $\mathscr{L}_\mu$. The first goal of this section is to formulate and prove very general abstract theorems establishing orbital stability under the above general assumptions and some extra technical conditions. The first such result, Theorem 7.1, is a general version of Proposition 4.1: it imposes a strong coercivity condition, but is nevertheless sometimes of use, as we will see in Section 8. Theorem 7.4 and Theorem 7.5 correspond essentially to the first two arguments proposed in the proof of Proposition 4.2. The proofs of these results are quite simple, as we shall see. These three results show that the essential ingredient in the proof of orbital stability is the coercivity condition in Hypothesis Bμ (iii).

It therefore remains to understand how to find a Lyapunov function satisfying in particular Hypothesis Bμ (iii). It is at this point that the Hamiltonian nature of the dynamical system plays an important role. We already saw in Section 6 that a candidate Lyapunov function arises naturally in that context. We will furthermore show in Proposition 7.7 how to obtain the coercivity condition Hypothesis Bμ (iii) from a lower bound on the Hessian of the Lyapunov function, in the case of Hamiltonian systems with symmetry. Combining this with Theorem 7.4 and Theorem 7.5 then yields a complete proof of orbital stability.

We will end this section with Theorem 7.10 which provides a slightly different proof of orbital stability of relative equilibria in Hamiltonian systems, and which is a generalization of the third argument proposed in the proof of Proposition 4.2. The argument uses Proposition 7.7 again, but combines it with the construction of an "augmented" Lyapunov function.

In applications of the theory developed in this section, the work is therefore reduced to solving (6.4) to identify the relative equilibria, and to proving a suitable lower bound on the Hessian of the corresponding Lyapunov function. This usually involves non-trivial (spectral) analysis, as one may expect. Illustrative examples are presented in Sections 8 and 9.

In conclusion, the theorems of this section isolate the "soft analysis" part of the proof of orbital stability of relative equilibria from the more concrete and model dependent estimates needed to prove coercivity.

7.2. **A simple case**

Before turning to the general results, we first formulate and prove a simple orbital stability result, under a stronger coercivity condition than (7.1).

Theorem 7.1. *Let Hypotheses A and Bμ_* (i)–(ii) for some $\mu_* \in \mathbb{R}^m$ be satisfied. Let $\mathscr{O}_{u_{\mu_*}}$ be as in* (7.2)*. Suppose there exist $\eta > 0, c > 0$ so that*

$$\forall u \in \mathscr{O}_{u_{\mu_*}}, \forall v \in E, \quad \mathrm{d}(v,u) \leq \eta \Rightarrow \mathscr{L}_{\mu_*}(v) - \mathscr{L}_{\mu_*}(u) \geq c\mathrm{d}^2(v, \mathscr{O}_{u_{\mu_*}}). \tag{7.3}$$

Then, all $u \in \mathscr{O}_{u_{\mu_*}}$ are orbitally stable $G_{\Sigma_{\mu_*}}$-relative equilibria.

We refer to Definition 3.1 for the definition of orbital stability. Observe that in (7.3) the coercivity estimate is imposed for all perturbations v in E, rather than only in Σ_μ, as in (7.1). So here we are assuming that the Lyapunov function reaches a local minimum at u, when viewed as a function on E, rather than only as a function on Σ_μ. This therefore constitutes a strengthening of Hypothesis Bμ_*(iii).The theorem can be used to prove orbital stability in some cases: for the fixed points in the spherical potentials treated in Section 4.1, for example, this is how we proceeded. Similarly, to establish the stability of the plane waves for the nonlinear Schrödinger equation on a one-dimensional torus, this theorem will also be sufficient, as we will see in Section 8. But we have already noticed in Section 4 that the coercivity imposed in (7.3) may be too strong a condition: we saw it is not satisfied for the natural choice of Lyapunov function for the circular orbits of Section 4.2, for example. It is too strong also in many situations involving the stability of solitons or standing waves. An example is treated in Section 9.

The proof is very simple, and based on the usual argument by contradiction.

Proof. Suppose there exists a point $u \in \mathscr{O}_{u_{\mu_*}}$ that is not orbitally stable. Then there exists $\varepsilon_0 > 0$ and for all $n \in \mathbb{N}^*$, there exists $v_n \in E$ so that $\mathrm{d}(v_n, u) \leq \frac{1}{n}$ and $\exists t_n \in \mathbb{R}$ so that $\mathrm{d}(v_n(t_n), \mathscr{O}_{u_{\mu_*}}) = \varepsilon_0$. We can suppose $\varepsilon_0 < \eta$. Then there exists $\tilde{v}_n \in \mathscr{O}_{u_{\mu_*}}$ so that $\mathrm{d}(v_n(t_n), \tilde{v}_n) \leq \eta$ and hence, since $\mathscr{L}_{\mu_*}$ is both a constant of the motion and constant on $\mathscr{O}_{u_{\mu_*}}$,

$$\mathscr{L}_{\mu_*}(v_n) - \mathscr{L}_{\mu_*}(u) = \mathscr{L}_{\mu_*}(v_n(t_n)) - \mathscr{L}_{\mu_*}(\tilde{v}_n) \geq c\mathrm{d}^2(v_n(t_n), \mathscr{O}_{u_{\mu_*}}) = c\varepsilon_0^2.$$

Since $\mathscr{L}_{\mu_*}$ is continuous, the left hand side tends to zero when $n \to +\infty$, which is a contradiction. □

7.3. **Coercivity implies stability I**

We now turn to the task of showing that Hypotheses A and Bμ_* imply the $G_{\Sigma_{\mu_*}}$-orbital stability of u_{μ_*}. For our first result, we need the following hypothesis.

HYPOTHESIS F Let $F : E \to \mathbb{R}^m$. Let $\mu \in \mathbb{R}^m$. We say F satisfies Hypothesis F at μ if, for any bounded sequence u_n in E,

$$\lim_n F(u_n) = \mu \Rightarrow \mathrm{d}(u_n, \Sigma_\mu) \to 0. \tag{7.4}$$

The following lemma gives sufficient conditions for this to be satisfied.

Lemma 7.2. *(a) Suppose* $\dim E < +\infty$. *Let $F \in C(E, \mathbb{R}^m)$. Then F satisfies Hypothesis F for all $\mu \in \mathbb{R}^m$.*

(b) Suppose $F \in C(E, \mathbb{R}^m)$ and that there exists $C > 0$ so that $\{u \in E \mid F(u)^2 \leq C^2\}$ is compact. Let $\mu \in \mathbb{R}^m$ with $\mu^2 < C^2$. Then F satisfies Hypothesis F at μ.

(c) Let $F : E \to \mathbb{R}$. Suppose that there exists $k \in \mathbb{R}^*$ so that, $\forall u \in \mathscr{D}$, for all $\lambda \in \mathbb{R}^*$, $F(\lambda u) = \lambda^k F(u)$. Suppose $\mu \neq 0$. Then F satisfies Hypothesis F at μ.

Proof. (a) Suppose there exists $\varepsilon_0 > 0$ and a bounded sequence u_n so that $F(u_n) \to \mu$, but $\mathrm{d}(u_n, \Sigma_\mu) \geq \varepsilon_0$. Then the boundedness of the sequence implies the existence of a convergent subsequence $u_{n_k} \to v \in E$. By continuity of F, it follows that $F(v) = \mu$ so that $v \in \Sigma_\mu$. So $\mathrm{d}(u_{n_k}, \Sigma_\mu) \to 0$. This is a contradiction.

(b) The proof is similar to the one in (a).

(c) Let $(u_n)_n$ be a bounded sequence satisfying $F(u_n) \to \mu \neq 0$. Then, for large enough n one has $\mu/F(u_n) > 0$ and we can define $v_n = \left(\frac{\mu}{F(u_n)}\right)^{1/k} u_n$. Then $F(v_n) = \mu$. Clearly $\|v_n - u_n\| \to 0$ so that $\mathrm{d}(u_n, \Sigma_\mu) \to 0$.

□

Remark 7.3. (i) The boundedness of the sequence is important, even in finite dimension. Indeed, consider on $\mathbb{R}^2$ the function $F(x,y) = \frac{y^2}{1+x^4}$, $\mu = 0$ and remark that $F(x,x) \to 0$ as $x \to +\infty$.
(ii) Condition (c) can be used for linear actions of one-parameter groups on a Hilbert space $(E, \langle\cdot,\cdot\rangle)$, which have a quadratic hamiltonian of the type

$$F(u) = \frac{1}{2}\langle u, Bu\rangle,$$

such as in (5.19).

Theorem 7.4. *Suppose Hypotheses A and Bμ_* (Section 7.1) are satisfied for some $\mu_* \in \mathbb{R}^m$. Then*

$$\forall u \in \mathscr{O}_{u_{\mu_*}}, \forall \varepsilon > 0, \exists \delta > 0,\ (\forall u' \in \Sigma_{\mu_*}, \mathrm{d}(u', u) \leq \delta \Rightarrow \sup_{t\in\mathbb{R}} \mathrm{d}(u'(t), \mathscr{O}_{u_{\mu_*}}) \leq \varepsilon). \tag{7.5}$$

If in addition,

(i) $\mathscr{L}_{\mu_*}$ is uniformly continuous on bounded sets,
(ii) $\mathscr{O}_{u_{\mu_*}}$ is bounded,
(iii) $F : E \to \mathbb{R}^m$ satisfies Hypothesis F,

then all $u \in \mathscr{O}_{u_{\mu_*}}$ are orbitally stable $G_{\Sigma_{\mu_*}}$-relative equilibria.

We point out that (7.5) is already an orbital stability result for all $u \in \mathscr{O}_{u_{\mu_*}} = G_{\Sigma_{\mu_*}} u$, but only with respect to perturbations of the initial condition u *inside* Σ_{μ_*}. The theorem asserts that, with the extra conditions (i)–(ii)–(iii), orbital stability with respect to all perturbations within E is obtained. It is the observation that coercivity *along* Σ_{μ_*} (Hypothesis Bμ (iii)) suffices to establish orbital stability that explains, *in fine*, the advantage of Theorem 7.4 over Theorem 7.1. This is already illustrated in Section 4.2 on a simple example. Note furthermore that conditions (i) and (iii) of the theorem are automatically satisfied in finite dimension. The boundedness of $\mathscr{O}_{u_{\mu_*}}$ (condition (ii)) is guaranteed for example when the group is compact, or when E is a Hilbert space and the group acts with unitary transformations, which is often the case in infinite dimensional systems.

The argument in the proof of Theorem 7.4 is extracted from the proof of Theorem 5.3 in [GSS87] and is used in [GSS90] as well. We point out however, that conditions (i) and (iii) are left implicit there. The first one is usually easy to check in examples, where the Lyapunov function tends at any rate to be uniformly Lipschitz on bounded sets. For the second one, we gave some sufficient conditions in Lemma 7.2.

Proof. We will prove (7.5) by contradiction, yet again. Let us therefore suppose there exists $u \in \mathscr{O}_{u_{\mu_*}}$ and $\varepsilon_0 > 0$ so that for all $n \in \mathbb{N}_*$, there exists $u_n \in \Sigma_{\mu_*}$ so that

$$\mathrm{d}(u_n, u) \leq \frac{1}{n}, \quad \text{and} \quad \exists \tilde{t}_n \in \mathbb{R} \text{ so that } \mathrm{d}(u_n(\tilde{t}_n), \mathscr{O}_{u_{\mu_*}}) > \varepsilon_0.$$

We can choose, without loss of generality, $\varepsilon_0 < \eta$, where η is defined in (7.1) and choose t_n the smallest value of t so that

$$\mathrm{d}(u_n, u) \leq \frac{1}{n}, \quad \text{and} \quad \mathrm{d}(u_n(t_n), \mathscr{O}_{u_{\mu_*}}) = \varepsilon_0 < \eta.$$

Consequently, there exists $y_n \in \mathscr{O}_{u_{\mu_*}}$ so that $\mathrm{d}(u_n(t_n), y_n) < \eta$. Note that $u_n(t_n) \in \Sigma_{\mu_*}$, since Σ_{μ_*} is invariant under the dynamical flow. Then, since $\mathscr{L}_{\mu_*}$ is a constant of the motion, and since it is constant and coercive on $\mathscr{O}_{u_{\mu_*}}$ along Σ_{μ_*},

$$\begin{aligned}\mathscr{L}_{\mu_*}(u_n) - \mathscr{L}_{\mu_*}(u) &= \mathscr{L}_{\mu_*}(u_n(t_n)) - \mathscr{L}_{\mu_*}(u) \\ &= \mathscr{L}_{\mu_*}(u_n(t_n)) - \mathscr{L}_{\mu_*}(y_n) \geq c\mathrm{d}^2(u_n(t_n), \mathscr{O}_{u_{\mu_*}}) = c\varepsilon_0^2.\end{aligned}$$

Since $\mathscr{L}_{\mu_*}$ is continuous, one obtains a contradiction by taking $n \to +\infty$. This shows (7.5).

To prove the last statement, suppose $\mathscr{O}_{u_{\mu_*}}$ is bounded and $\mathscr{L}_{\mu_*}$ uniformly continuous on bounded sets. We need to show that

$$\forall u \in \mathscr{O}_{u_{\mu_*}}, \forall \varepsilon > 0, \exists \delta > 0,\ (\forall u' \in E, \mathrm{d}(u', u) \leq \delta \Rightarrow \sup_{t\in\mathbb{R}} \mathrm{d}(u'(t), \mathscr{O}_{u_{\mu_*}}) \leq \varepsilon). \tag{7.6}$$

We proceed again by contradiction. Suppose there exists $u \in \mathscr{O}_{u_{\mu_*}}$ and $0 < \varepsilon_0 < \eta$ so that, for all $n \in \mathbb{N}$, there exists $u_n \in E$,

$$\mathrm{d}(u_n, u) \leq \frac{1}{n}, \quad \text{and} \quad \exists t_n \in \mathbb{R} \text{ so that } \mathrm{d}(u_n(t_n), \mathscr{O}_{u_{\mu_*}}) = \varepsilon_0 < \eta.$$

Note that, this time, $u_n \in E$ and $u_n(t_n) \in E$, not in Σ_{μ_*}. So we can't use the coercivity of $\mathscr{L}_{\mu_*}$ along Σ_{μ_*} directly. We do know, however, that $F(u_n(t_n)) = F(u_n)$, since F is a constant of the motion. Hence

$$\lim_{n\to+\infty} F(u_n(t_n)) = \mu_*.$$

Since the orbit $\mathscr{O}_{u_{\mu_*}}$ is bounded, and since $\mathrm{d}(u_n(t_n), \mathscr{O}_{u_{\mu_*}}) = \varepsilon_0$, it follows that the sequence $u_n(t_n)$ is bounded. Hypothesis F then implies there exist $z_n \in \Sigma_{\mu_*}$ so that $\|u_n(t_n) - z_n\| \to 0$.

We can now conclude. Since, for n large enough, $\frac{\varepsilon_0}{2} \leq \mathrm{d}(z_n, \mathscr{O}_{u_{\mu_*}}) \leq \eta$, we have

$$\begin{aligned}\mathscr{L}_{\mu_*}(u_n) - \mathscr{L}(u) &= \mathscr{L}_{\mu_*}(u_n(t_n)) - \mathscr{L}_{\mu_*}(u) \\ &= \mathscr{L}_{\mu_*}(u_n(t_n)) - \mathscr{L}_{\mu_*}(z_n) + \mathscr{L}_{\mu_*}(z_n) - \mathscr{L}_{\mu_*}(u) \\ &\geq \mathscr{L}_{\mu_*}(u_n(t_n)) - \mathscr{L}_{\mu_*}(z_n) + c\mathrm{d}^2(z_n, \mathscr{O}_{u_{\mu_*}}).\end{aligned}$$

Since the orbit $\mathscr{O}_{u_{\mu_*}}$ is bounded, the sequences $u_n(t_n)$ and z_n are bounded. This, combined with the uniform continuity of $\mathscr{L}_{\mu_*}$ on bounded sets, leads again to a contradiction upon taking $n \to +\infty$. □

We now give a third proof of orbital stability starting from a coercive Lyapunov function, along the lines of the third argument in the proof of Proposition 4.2. The point here is that we exploit the fact that the relative equilibria u_μ often come in families.

Theorem 7.5. *Suppose the following.*

(i) Hypothesis A holds.
(ii) There exists a continuous map $\mu \in U \subset \mathbb{R}^m \to u_\mu \in \Sigma_\mu \subset E$ so that Hypothesis Bμ is satisfied for all $\mu \in U$, with η and c in (7.1) *independent of μ.*
(iii) $\sup_{\mu\in U} \|u_\mu\| < +\infty$.
(iv) There exists $C > 0$ so that

$$\forall \mu \in U, \forall u' \in \Sigma_\mu, \quad \|u' - u_\mu\| \leq \eta \Rightarrow \mathscr{L}_\mu(u') - \mathscr{L}_\mu(u_\mu) \leq C\|u' - u_\mu\|. \tag{7.7}$$

(v) $\forall g \in G$, Φ_g is an isometry on E: $\forall u, u' \in E$, $\mathrm{d}(\Phi_g(u), \Phi_g(u')) = \mathrm{d}(u, u')$.

Then, any $u \in \mathscr{O}_{u_\mu}$ is an orbitally stable G_{Σ_μ}-relative equilbrium of the flow Φ_t^H.

Condition (iii) is not very restrictive. It is sufficient to take U bounded, for example. Condition (iv) follows if we know that $D_u\mathscr{L}_\mu$ is bounded for u in bounded sets. This is a reasonable condition. Condition (v) is commonly satisfied in PDE systems, but is quite restrictive, as we already explained. It implies we can use Proposition 2.3 and Lemma 3.2.

Proof. Let $\mu_* \in U$. As a result of Lemma 3.2, it is enough to show the orbital stability of u_{μ_*}. So we need to show that, for all $\varepsilon > 0$, there exists $\delta > 0$ so that, for all $u' \in E$, one has

$$\|u' - u_{\mu_*}\| \leq \delta \Rightarrow \forall t \in \mathbb{R}, \mathrm{d}(u'(t), \mathscr{O}_{u_{\mu_*}}) \leq \varepsilon. \tag{7.8}$$

For that purpose, we need three preliminary estimates. We first show that $\forall \varepsilon > 0$, there exists $\hat{\delta} > 0$ so that, for all $\mu \in U$, for all $u' \in \Sigma_\mu$,

$$\|u' - u_\mu\| \leq \hat{\delta} \Rightarrow \forall t \in \mathbb{R}, \mathrm{d}(u'(t), \mathscr{O}_{u_\mu}) \leq \varepsilon/2. \tag{7.9}$$

In other words, we first show that the u_μ are all orbitally stable for perturbations *within* Σ_μ. The method of proof – by contradiction – is the same as several times before, but we need to make sure to obtain the necessary uniformity in μ. If the above is not true, then there exists $\varepsilon_0 > 0$ so that for all $n \in \mathbb{N}^*$ there exist $\mu_n \in U$ and $u_n \in \Sigma_{\mu_n}$, $t_n \in \mathbb{R}$, so that

$$\|u_n - u_{\mu_n}\| \leq \frac{1}{n}, \quad \mathrm{d}(u_n(t_n), \mathscr{O}_{u_{\mu_n}}) = \frac{\varepsilon_0}{2} < \eta.$$

Here η is given in Hypothesis Bμ (iii) and we recall that it is independent of μ_n. Hence

$$\mathscr{L}_{\mu_n}(u_n) - \mathscr{L}_{\mu_n}(u_{\mu_n}) = \mathscr{L}_{\mu_n}(u_n(t_n)) - \mathscr{L}_{\mu_n}(u_{\mu_n}) \geq c\mathrm{d}^2(u_n(t_n), \mathscr{O}_{\mu_n}) = c\frac{\varepsilon_0^2}{4}.$$

Now, since the u_{μ_n} form a bounded set by hypothesis (iii) of the theorem, the same is true for the u_n. Hence, it follows from hypothesis (iv) of the theorem that

$$\mathscr{L}_{\mu_n}(u_n) - \mathscr{L}_{\mu_n}(u_{\mu_n}) \leq C\|u_n - u_{\mu_n}\|,$$

where C does not depend on n. Hence $C\|u_n - u_{\mu_n}\| \geq c\frac{\varepsilon_0^2}{4}$, so that, taking $n \to +\infty$, we obtain a contradiction. This proves (7.9).

As a second step, we show the following estimate. Let $\mu_* \in U$. Then, for all $\varepsilon > 0$, there exists $\hat{\rho} > 0$ so that,

$$\forall \mu \in U, \quad \left(\|\mu - \mu_*\| \leq \hat{\rho} \Rightarrow \forall v \in \mathscr{O}_{u_\mu}, \mathrm{d}(v, \mathscr{O}_{u_{\mu_*}}) \leq \frac{\varepsilon}{2}\right). \tag{7.10}$$

To see, this, note that hypothesis (i) of the theorem implies that there exists $\hat{\rho} > 0$ so that $\|\mu - \mu_*\| \leq \hat{\rho}$ implies $\|u_\mu - u_{\mu_*}\| \leq \varepsilon/2$. Hence $\mathrm{d}(u_\mu, \mathscr{O}_{u_{\mu_*}}) \leq \varepsilon/2$. The result then follows from Proposition 2.3, since we suppose the action Φ of G is isometric.

The third ingredient for the proof of (7.8) is the following:

$$\forall \hat{\delta} > 0, \forall \hat{\rho} > 0, \exists \delta > 0, \forall u' \in E,$$
$$\left(\|u' - u_{\mu_*}\| \leq \delta \Rightarrow \|\mu' - \mu_*\| \leq \hat{\rho}, \|u' - u_{\mu'}\| \leq \hat{\delta}\right), \tag{7.11}$$

where $\mu' = F(u')$. This follows immediately from the continuity of F and of $\mu \to u_\mu$ at μ_*.

We can now conclude. Let $\mu_* \in U$ and $\varepsilon > 0$. Choose $\hat{\delta}$ as in (7.9), $\hat{\rho}$ as in (7.10) and δ as in (7.11). Then, by (7.9) and (7.11), we find that

$$\forall u' \in E, \left(\|u' - u_{\mu_*}\| \leq \delta \Rightarrow \forall t \in \mathbb{R}, \mathrm{d}(u'(t), \mathscr{O}_{u_{\mu'}}) < \frac{\varepsilon}{2}\right).$$

Hence, for all $t \in \mathbb{R}$, there exists $v(t) \in \mathscr{O}_{u_{\mu'}}$, so that $\mathrm{d}(u'(t), v(t)) < \varepsilon/2$. Next, from (7.11) and (7.10) , there exists $w(t) \in \mathscr{O}_{u_{\mu_*}}$ so that $\mathrm{d}(v(t), w(t)) < \frac{\varepsilon}{2}$. Hence $\mathrm{d}(u'(t), \mathscr{O}_{u_{\mu_*}}) < \varepsilon$. This proves (7.8). □

7.4. **Sufficient condition for coercivity**

We now turn to the task of showing how one can obtain Hypothesis Bμ (iii) from an estimate on the Hessian of $\mathscr{L}_\mu$ (Proposition 7.7). We work in the following setting.

Let E be a Banach space, G a Lie group and Φ a G-action on E. Let $F \in C^2(E, \mathbb{R}^m)$. We recall that, for $\mu \in \mathbb{R}^m$,

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

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

$$\forall v, w \in E, \quad |\langle v, w \rangle| \leq \|v\| \|w\|,$$

where we recall that $\|\cdot\|$ is our notation for the Banach norm on E. This inner product induces a metric on E, that we shall denote by

$$\mathrm{d}_\mathrm{s}(v, w) = \langle v - w, v - w \rangle. \tag{7.12}$$

Clearly $\mathrm{d}_\mathrm{s}(v, w) \leq \mathrm{d}(v, w)$. We point out that we are not supposing E is a Hilbert space for this inner product, and that the only topology we will be using is the one induced by the Banach norm on E. We introduce this inner product since we need a notion of orthogonality as will become clear below.

As an example, if $E = H^1(\mathbb{R}^d, \mathbb{C})$ and depending on the problem considered, one may want to use either the L^2 inner product or the H^1 inner product: in Section 8 the first choice is made and in Section 9 the second one.

Let $\mu \in \mathbb{R}^m$ and $u_\mu \in \Sigma_\mu$. We need the following hypothesis.

HYPOTHESIS Cμ

(i) Φ_g is linear and preserves both the structure $\langle \cdot, \cdot \rangle$ and the norm $\|\cdot\|$ for all $g \in G$;
(ii) $\mathrm{Ad}^*_g \in \mathrm{O}(m)$ for all $g \in G_{\Sigma_\mu}$;
(iii) μ is a regular value of F;
(iv) u_μ is a C^1-vector for Φ and the map

$$\xi \in \mathfrak{g}_{\Sigma_\mu} \to \Phi_{\exp(\xi)} u_\mu \in E \tag{7.13}$$

is one to one in a neighbourhood of $\xi = 0$.

The meaning of condition (ii) of Hypothesis Cμ is explained in Remark A.2.1.

Remark 7.6. We say $u \in E$ is a C^1-vector for the action Φ if the map

$$g \in G \to \Phi_g(u) \in E$$

is C^1. Now, if $u' \in \mathscr{O}_u = \Phi_G(u)$, then u' is also a C^1-vector. Indeed, there exists $g' \in G$ so that $\Phi_{g'} u = u'$ and, since $g \to gg'$ is smooth, it follows that $g \to \Phi_{gg'} u$ is C^1.

Let $\tilde{G}$ be a subgroup of G, we can define, for all $u' \in \mathscr{O}_u = \Phi_{\tilde{G}}(u)$,

$$T_{u'} \mathscr{O}_u := \{w \in E \mid \exists \xi \in \mathfrak{g}, w = X_\xi(u')\}, \tag{7.14}$$

where we recall from (A.2.15) that

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

Note that hypothesis Cμ and the following proposition involve G and its action on E, as well as F, but not the dynamics Φ^X_t itself.

Proposition 7.7. *Let E be a Banach space and $\langle\cdot,\cdot\rangle$ be a continuous scalar product on E. Let G be a Lie group and Φ a G-action on E. Let $F \in C^2(E,\mathbb{R}^m)$. Let $\mu_* \in \mathbb{R}^m$ and $u_{\mu_*} \in \Sigma_{\mu_*}$. Let $\mathscr{L}_{\mu_*} \in C^2(E,\mathbb{R})$ be a $G_{\Sigma_{\mu_*}}$-invariant function. Suppose Hypothesis $C\mu_*$ holds and that, for all $u \in \mathscr{O}_{u_{\mu_*}}$ (defined in* (7.2)*),*

$$\forall j = 1,\dots,m\ \exists \nabla F_j(u) \in E \text{ such that } D_u F_j(w) = \langle \nabla F_j(u), w\rangle\ \forall w \in E. \tag{7.15}$$

If

(a) $D_u \mathscr{L}_{\mu_*}(w) = 0$ for all $u \in \mathscr{O}_{u_{\mu_*}}$ and $w \in E$;
(b) there exists $C > 0$ so that

$$\forall u \in \mathscr{O}_{u_{\mu_*}}, \forall w \in E,\ D_u^2 \mathscr{L}_{\mu_*}(w,w) \le C\|w\|^2;$$

(c) there exists $c > 0$ so that

$$\forall u \in \mathscr{O}_{u_{\mu_*}}, \forall w \in T_u\Sigma_{\mu_*} \cap (T_u\mathscr{O}_{u_{\mu_*}})^\perp,\ D_u^2 \mathscr{L}_{\mu_*}(w,w) \ge c\|w\|^2 \tag{7.16}$$

where

$$(T_w\mathscr{O}_u)^\perp = \{z \in E \mid \langle z, y\rangle = 0, \forall y \in T_w\mathscr{O}_u\}; \tag{7.17}$$

then Hypothesis $B\mu_*$ (iii) holds.

Condition (7.15) is automatically satisfied when E is a Hilbert space and $\langle\cdot,\cdot\rangle$ the Hilbert space inner product. But not in general. For example, let $E = H^1(\mathbb{R},\mathbb{C})$ and let $\langle u, v\rangle = \mathrm{Re}\int_{\mathbb{R}} \bar{u}(x)v(x)\mathrm{d}x$. Now, if $F_1(u) = \frac{1}{2i}\int \overline{u}(x)\partial_x u(x)\mathrm{d}x$, (7.15) is satisfied if $u \in H^2(\mathbb{R},\mathbb{C})$ but not arbitrary $u \in E$.

For the proof of this proposition, we need some simple technical results.

First, let V be a bounded open neighbourhood of e in a subgroup $\tilde{G}$ of G with the property that, for all $g \in \tilde{G}$, $gVg^{-1} = V$. Let us introduce

$$R_V(u) = \min\{\mathrm{d_s}(\Phi_g(u), u) \mid g \in \partial V\}.$$

It then follows that, for all $u' \in \mathscr{O}_u$, $R_V(u') = R_V(u)$. Indeed, there exists $g' \in \tilde{G}$ so that $\Phi_{g'}(u) = u'$. Hence

$$\begin{aligned} R_V(u') &= \min\{\mathrm{d_s}(\Phi_{gg'}(u), \Phi_{g'}(u)) \mid g \in \partial V\} \\ &= \min\{\mathrm{d_s}(\Phi_{g'^{-1}gg'}(u), u) \mid g \in \partial V\} = R_V(u), \end{aligned}$$

since ${g'}^{-1}\partial V g' = \partial V$.

We can now formulate the following simple but crucial technical result, which is a multi-dimensional version of Lemma 2.1 in [Stu08].

Lemma 7.8. *Let E be a Banach space and $\langle\cdot,\cdot\rangle$ be a continuous scalar product on E. Let $\tilde{G}$ be a Lie subgroup of G and Φ a linear G-action on E which preserves the inner product $\langle\cdot,\cdot\rangle$. Suppose $u \in E$ is a C^1-vector for Φ and let V be a bounded open neighbourhood of $e \in \tilde{G}$ which is conjugation invariant (*i.e. *$gVg^{-1} = V$, for all $g \in \tilde{G}$). Suppose $R_V(u) > 0$. Then, for all $v \in E$,*

$$\mathrm{d}(v, \mathscr{O}_u) < \frac{1}{3}R_V(u) \Rightarrow \exists w \in \mathscr{O}_u = \Phi_{\tilde{G}}(u), w - v \in (T_w\mathscr{O}_u)^\perp. \tag{7.18}$$

The lemma states that if v is not too far from the orbit $\mathscr{O}_u$, then there exists a point w on the orbit so that the segment from v to w is orthogonal to the orbit at w. This point does *not* necessarily realize the distance between v and the orbit, which can vanish.

Proof. Let $v \in E$ and $\mathrm{d}(v,\mathscr{O}_u) < \frac{1}{3}R_V(u)$. Then there exists $u' \in \mathscr{O}_u$ so that $\mathrm{d}(v,u') \leq \frac{1}{3}R_V(u) = \frac{1}{3}R_V(u')$ and hence $\mathrm{d_s}(v,u') \leq \frac{1}{3}R_V(u')$. Now consider

$$g \in \overline{V} \to \mathrm{d_s}^2(v,\Phi_g u') \in \mathbb{R}^+.$$

Since $\overline{V}$ is compact, this function reaches a minimum at some point $\tilde{g} \in \overline{V}$. We set $w = \Phi_{\tilde{g}} u' \in \mathscr{O}_u$ so that $\mathrm{d_s}(v,w) \leq \mathrm{d_s}(v,u') \leq \frac{1}{3}R_V(u')$. We now show that $\tilde{g}$ cannot belong to ∂V. Indeed, if $\tilde{g}$ were on the boundary of V, then, by the definition of $R_V(u')$, $\mathrm{d_s}(w,u') \geq R_V(u')$. But then

$$\mathrm{d_s}(w,v) \geq -\mathrm{d_s}(u',v) + \mathrm{d_s}(u',w) \geq R_V(u') - \frac{1}{3}R_V(u') = \frac{2}{3}R_V(u').$$

which is a contradiction because $\mathrm{d_s}(v,w) \leq \frac{1}{3}R_V(u')$. So $\tilde{g}$ belongs to V. Now choose $\xi \in \mathfrak{g}$ and consider

$$t \in \mathbb{R} \to \mathrm{d_s}^2(v,\Phi_{\exp(t\xi)\tilde{g}}(u')) \in \mathbb{R}^+,$$

which now reaches a local minimum at $t = 0$ since for small t, $\exp(t\xi)\tilde{g}$ belongs to V. Hence its derivative vanishes. So

$$\begin{aligned} 0 = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d_s}^2(v,\Phi_{\exp(t\xi)\tilde{g}}(u'))_{|t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\langle v - \Phi_{\exp(t\xi)}(w), v - \Phi_{\exp(t\xi)}(w)\rangle_{|t=0} \\ &= -2\langle X_\xi(w), v - w\rangle, \end{aligned}$$

which proves the result in view of (7.14). □

In the proof of Proposition 7.7, we will need to apply the previous lemma to the group $G_{\Sigma_{\mu_*}}$ for some $\mu_* \in \mathbb{R}^m$ and $u_{\mu_*} \in \Sigma_{\mu_*}$. The following lemma gives hypotheses for this to be possible. It appears in various guises in the literature, and can be referred to as a "modulation" argument.

Lemma 7.9. *Let E be a Banach space and $\langle\cdot,\cdot\rangle$ be a continuous scalar product on E. Let G be a Lie group and Φ a G-action on E. Let $F \in C^2(E,\mathbb{R}^m)$. Let $\mu_* \in \mathbb{R}^m$ and $u_{\mu_*} \in \Sigma_{\mu_*}$. Suppose Hypothesis $C\mu_*$ holds. Then, there exists $R > 0$ such that, for all $v \in E$,*

$$\mathrm{d}(v,\mathscr{O}_{u_{\mu_*}}) < R \Rightarrow \exists w \in \mathscr{O}_{u_{\mu_*}}, w - v \in \left(T_w\mathscr{O}_{u_{\mu_*}}\right)^\perp \tag{7.19}$$

where $\mathscr{O}_{u_{\mu_*}} = \Phi_{G_{\Sigma_{\mu_*}}} u_{\mu_*}$.

Proof. Thanks to Lemma 7.8, it is enough to prove that there exists V a bounded open neighbourhood of $e \in G_{\Sigma_{\mu_*}}$, which is conjugation invariant (*i.e.* $gVg^{-1} = V$, for all $g \in G_{\Sigma_{\mu_*}}$) and such that $R_V(u_{\mu_*}) > 0$.

First of all, we recall that the exponential map

$$\exp : \xi \in \mathfrak{g}_{\mu_*} \to \exp(\xi) \in G_{\Sigma_{\mu_*}}$$

is a local diffeomorphism from some neighbourhood of $0 \in \mathfrak{g}_{\mu_*}$ to a neighbourhood of $e \in G_{\Sigma_{\mu_*}}$. In other words, there exists $\delta > 0$ such that

$$\exp : \xi \in B_\delta(0) \subset \mathfrak{g}_{\mu_*} \to \exp(\xi) \in G_{\Sigma_{\mu_*}}$$

is a local diffeomorphism onto a bounded open neighbourhood $V := \exp(B_\delta(0))$ of e in G_{μ_*}. In particular, note that $\partial V = \exp(\partial B_\delta(0))$.

Since, thanks to Hypothesis Cμ_*(ii), $B_\delta(0)$ is Ad_g-invariant for all $g \in G_{\Sigma_{\mu_*}}$, V is conjugation invariant. Indeed, for all $\xi \in B_\delta(0)$ and all $g \in G_{\Sigma_{\mu_*}}$, we have that $g\exp(\xi)g^{-1} = \exp(\mathrm{Ad}_g\xi) \in V$.

Hence, it only remains to show that $R_V(u_{\mu_*}) > 0$, which is equivalent to $G_{u_{\mu_*}} \cap \partial V = \emptyset$. Thanks to Hypothesis C$\mu_*$(iv), there exists $\delta_0 > 0$ such that

$$\xi \in B_{\delta_0}(0) \to \Phi_{\exp(\xi)} u_{\mu_*} \in E$$

is one to one. As a conclusion, choosing $\delta < \delta_0$, we have $\partial V \subset \exp(B_{\delta_0}(0))$ which implies $\Phi_{\exp(\xi)} u_{\mu_*} \neq u_{\mu_*}$ for all $\exp(\xi) \in \partial V$. Hence, for all $\exp(\xi) \in \partial V$, $\exp(\xi) \notin G_{u_{\mu_*}}$. □

We can then conclude this section with the proof of Proposition 7.7.

Proposition 7.7. Recall that we have to prove there exist $\eta > 0, \tilde{c} > 0$ so that

$$\forall u \in \mathscr{O}_{u_{\mu_*}}, \forall u' \in \Sigma_{\mu_*}, \quad \mathrm{d}(u,u') \leq \eta \Rightarrow \mathscr{L}_{\mu_*}(u') - \mathscr{L}_{\mu_*}(u) \geq \tilde{c}\mathrm{d}^2(u', \mathscr{O}_{u_{\mu_*}}).$$

Let $u' \in \Sigma_{\mu_*}$, $\mathrm{d}(u', \mathscr{O}_{u_{\mu_*}}) < R$. Thanks to Lemma 7.9, there exists $v' \in \mathscr{O}_{u_{\mu_*}}$ such that $u' - v' \in \left(T_{v'}\mathscr{O}_{u_{\mu_*}}\right)^{\perp}$.

Next, let $W_{v'}$ be the subspace of E spanned by $\{\nabla F_j(v')\}_{j=1\ldots,m}$. It follows from (A.1.6) and hypothesis (7.15) that $T_{v'}\Sigma_{\mu_*} = (W_{v'})^{\perp}$. As a consequence, we can write $E = T_{v'}\Sigma_{\mu_*} \oplus W_{v'}$. Indeed, since $W_{v'}$ has finite dimension, it admits an orthonormal basis $\{e_1, \ldots, e_m\}$ w.r.t. $\langle \cdot, \cdot \rangle$. Hence, all $w \in E$ can be written as

$$w = \left(w - \sum_{j=1}^{m} \langle w, e_j \rangle e_j \right) + \sum_{j=1}^{m} \langle w, e_j \rangle e_j.$$

Clearly $w - \sum_{j=1}^{m} \langle w, e_j \rangle e_j \in (W_{v'})^{\perp} = T_{v'}\Sigma_{\mu_*}$, $\sum_{j=1}^{m} \langle w, e_j \rangle e_j \in W_{v'}$ and $W_{v'} \cap (W_{v'})^{\perp} = \{0\}$. Then,

$$u' - v' = (u' - v')_1 + (u' - v')_2$$

where $(u' - v')_1 \in T_{v'}\Sigma_{\mu_*}$ and $(u' - v')_2 \in W_{v'}$. Moreover, since $u' - v' \in \left(T_{v'}\mathscr{O}_{u_{\mu_*}}\right)^{\perp}$, we can easily show $(u' - v')_1 \in T_{v'}\Sigma_{\mu_*} \cap \left(T_{v'}\mathscr{O}_{u_{\mu_*}}\right)^{\perp}$ and $(u' - v')_2 \in W_{v'} \cap \left(T_{v'}\mathscr{O}_{u_{\mu_*}}\right)^{\perp}$. Now, Lemma A.1.4 ensures the existence of constants c_1, c_0 so that, for $\|u' - v'\|$ small enough, one has

$$\|(u' - v')_1\| \geq c_0 \|u' - v'\| \text{ and } \|(u' - v')_2\| \leq c_1 \|u' - v'\|^2. \tag{7.20}$$

Since the action Φ_g is linear and preserves both $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, the decomposition above is group invariant and the constant c_0 and c_1 do not depend on v'.

We can now conclude the proof as follows, using respectively conditions (a), (b) and (c), and (7.20):

$$\begin{aligned}
\mathscr{L}_{\mu_*}(u') - \mathscr{L}_{\mu_*}(u_{\mu_*}) &= \mathscr{L}_{\mu_*}(u') - \mathscr{L}_{\mu_*}(v') \\
&= D_{v'}\mathscr{L}_{\mu_*}(u' - v') + \frac{1}{2} D^2_{v'}\mathscr{L}_{\mu_*}(u' - v', u' - v') + \mathrm{o}(\|u' - v'\|^2) \\
&= \frac{1}{2} D^2_{v'}\mathscr{L}_{\mu_*}((u' - v')_1, (u' - v')_1) + \mathrm{O}(\|u' - v'\|^3) + \mathrm{o}(\|u' - v'\|^2) \\
&= \frac{1}{2} D^2_{v'}\mathscr{L}_{\mu_*}((u' - v')_1, (u' - v')_1) + \mathrm{o}(\|u' - v'\|^2) \\
&\geq \frac{c}{2} \|(u' - v')_1\|^2 + \mathrm{o}(\|u' - v'\|^2) \\
&\geq \tilde{c}\|u' - v'\|^2 \geq \tilde{c}\mathrm{d}^2(u', \mathscr{O}_{u_{\mu_*}}).
\end{aligned}$$

Remark that as before, the constant $\tilde{c}$ is independent of $v' \in \mathscr{O}_{\mu_*}$. □

7.5. **Coercivity implies stability II**

We can now state and prove a fourth theorem yielding orbital stability under slightly different technical assumptions. We will work in the Hamiltonian setting and in particular use the characterization of relative equilibria given by Theorem 6.1. Recall that in this context, for each $\mu \in \mathfrak{g}^* \simeq \mathbb{R}^m$, $G_{\Sigma_\mu} = G_\mu$ (Proposition A.3.11).

Theorem 7.10. *Let E be a Banach space and $\langle\cdot,\cdot\rangle$ be a continuous scalar product on E, $\mathscr{D}$ a domain in E and $\mathscr{J}$ a symplector. Let $H \in C^2(E,\mathbb{R}) \cap \mathrm{Dif}(\mathscr{D},\mathscr{J})$. Let G be a Lie group, and Φ a globally Hamiltonian G-action on E with Ad^*-equivariant momentum map F. Let $\mu_* \in \mathbb{R}^m \simeq \mathfrak{g}^*$ and $u_{\mu_*} \in \mathscr{D} \cap \Sigma_{\mu_*}$. Suppose that Hypothesis Cμ_*(i)–(iii) is satisfied, and $H \circ \Phi_g = H$ for all $g \in G$. Let $\mathscr{L}_{\mu_*} = H - \xi_{\mu_*} \cdot F$ with $\xi_{\mu_*} \in \mathfrak{g}_{\mu_*}$ given by Theorem 6.1 and assume $D_{u_{\mu_*}} \mathscr{L}_{\mu_*} = 0$. Suppose in addition that*

$$\forall\, j = 1,\dots,m \;\exists \nabla F_j(u_{\mu_*}) \in E \text{ such that } D_{u_{\mu_*}} F_j(w) = \langle \nabla F_j(u_{\mu_*}), w\rangle \;\forall w \in E. \tag{7.21}$$

and

(a) G_{μ_*} is commutative;
(b) there exists $C > 0$ so that

$$\forall w \in E,\; D^2_{u_{\mu_*}} \mathscr{L}_{\mu_*}(w,w) \le C\|w\|^2;$$

(c) there exists $c > 0$ so that

$$\forall w \in T_{u_{\mu_*}} \Sigma_{\mu_*} \cap (T_{u_{\mu_*}} \mathscr{O}_{u_{\mu_*}})^\perp,\; D^2_{u_{\mu_*}} \mathscr{L}_{\mu_*}(w,w) \ge c\|w\|^2.$$

Then all $u \in \mathscr{O}_{u_{\mu_*}}$ are orbitally stable G_{μ_*}-relative equilibria.

Hypothesis (a) in Theorem 7.10 is not very restrictive (see [DV69]).

Proof. Let $K > 0$ and define

$$\mathscr{L}_K(u) = \mathscr{L}_{\mu_*}(u) + K(F(u) - \mu_*)^2.$$

Here $(F(u)-\mu_*)^2 = (F(u)-\mu_*)\cdot(F(u)-\mu_*)$ where, $\cdot$ is the G_{μ_*}-invariant inner product described in Remark A.2.1. It follows that $\mathscr{L}_K$ is a G_{μ_*}-invariant constant of the motion. Indeed, for all $g \in G_{\mu_*}$ and for all $u \in E$,

$$\begin{aligned}
\mathscr{L}_K(\Phi_g u) &= H(\Phi_g u) - \xi_{\mu_*} \cdot F(\Phi_g u) + K(F(\Phi_g u) - \mu_*)^2 \\
&= H(u) - \xi_{\mu_*} \cdot \mathrm{Ad}^*_g F(u) + K(\mathrm{Ad}^*_g F(u) - \mathrm{Ad}^*_g \mu_*)^2 \\
&= H(u) - \mathrm{Ad}_g \xi_{\mu_*} \cdot F(u) + K(F(u) - \mu_*)^2 && \text{as } \mathrm{Ad}^*_g \in \mathrm{O}(m) \\
&= H(u) - \xi_{\mu_*} \cdot F(u) + K(F(u) - \mu_*)^2 && \text{as } G_{\mu_*} \text{ is commutative} \\
&= \mathscr{L}_K(u).
\end{aligned}$$

The main idea is to prove that the hypotheses of Proposition 7.7 are satisfied by $\mathscr{L}_K$ and then use its proof to conclude that all $u \in \mathscr{O}_{u_{\mu_*}}$ are orbitally stable G_{μ_*}-relative equilibria.

First, note that in this setting Hypothesis Cμ_*(iv) follows from Remark 6.3.

Next, we claim that $D_u \mathscr{L}_K(w) = 0$ for all $u \in \mathscr{O}_{u_{\mu_*}}$ and for all $w \in E$. Indeed, it is clear that $D_u(F(u)-\mu_*)^2 = 2(F(u)-\mu_*)\cdot D_u F = 0$ for all $u \in \mathscr{O}_{u_{\mu_*}}$ and, thanks to the fact that $D_{u_{\mu_*}} \mathscr{L}_{K_{\mu_*}}(w) = 0$, we obtain $D_{u_{\mu_*}} \mathscr{L}_K(w) = 0$ for all $w \in E$. Next, let $u \in \mathscr{O}_{u_{\mu_*}}$ and $g \in G_{\mu_*}$ such that $u = \Phi_g(u_{\mu_*})$, then

$$D_u \mathscr{L}_K(w) = [D_{\Phi_g(u_{\mu_*})} \mathscr{L}_K \circ \Phi_{g^{-1}}](w) = [D_{u_{\mu_*}} \mathscr{L}_K \circ D_{\Phi_g(u_{\mu_*})} \Phi_{g^{-1}}](w) = 0.$$

Using the fact that Φ_g is linear and preserves both $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$, we can easily show, as a consequence of hypothesis (c), that

$$\forall u \in \mathscr{O}_{u_{\mu_*}},\ D_u^2\mathscr{L}_{\mu_*}(w,w) \geq c\|w\|^2, \tag{7.22}$$

for all $w \in T_u\Sigma_{\mu_*} \cap (T_u\mathscr{O}_{u_{\mu_*}})^\perp$. Indeed, for all $u \in \mathscr{O}_{u_{\mu_*}}$ and $w \in T_u\Sigma_{\mu_*} \cap (T_u\mathscr{O}_{u_{\mu_*}})^\perp$,

$$\begin{aligned} D_u^2\mathscr{L}_{\mu_*}(w,w) &= D^2_{\Phi_g u_{\mu_*}}(\mathscr{L}_{\mu_*}\circ\Phi_{g^{-1}})(w,w) \\ &= D^2_{u_{\mu_*}}\mathscr{L}_{\mu_*}(D_u\Phi_{g^{-1}}w, D_u\Phi_{g^{-1}}w) + D_{u_{\mu_*}}\mathscr{L}_{\mu_*}(D_u^2\Phi_{g^{-1}}(w,w)) \\ &= D^2_{u_{\mu_*}}\mathscr{L}_{\mu_*}(\Phi_{g^{-1}}w, \Phi_{g^{-1}}w) \geq c\|\Phi_{g^{-1}}w\|^2 = c\|w\|^2 \end{aligned}$$

because $\Phi_{g^{-1}}w \in T_{u_{\mu_*}}\Sigma_{\mu_*} \cap (T_{u_{\mu_*}}\mathscr{O}_{u_{\mu_*}})^\perp$.

Similarly, using hypothesis (b), we prove that

$$D_u^2\mathscr{L}_{\mu_*}(w,w) \leq C\|w\|^2$$

for all $u \in \mathscr{O}_{u_{\mu_*}}$ and $w \in E$.

Next, by a straightforward calculation, we obtain for all $u \in \mathscr{O}_{u_{\mu_*}}$ and $w \in E$, $D_u^2(F-\mu_*)^2(w,w) = 2D_uF(w)\cdot D_uF(w)$, and

$$\begin{aligned} D_uF(w) &= [D_{\Phi_g u_{\mu_*}}F\circ\Phi_g\circ\Phi_{g^{-1}}](w) = [D_{u_{\mu_*}}F\circ\Phi_g](D_u\Phi_{g^{-1}}w) \\ &= [D_{u_{\mu_*}}\mathrm{Ad}_g^*\circ F](\Phi_{g^{-1}}w) = \mathrm{Ad}_g^*(D_{u_{\mu_*}}F(\Phi_{g^{-1}}w)). \end{aligned} \tag{7.23}$$

As a consequence, since $\mathrm{Ad}_g^* \in \mathrm{O}(m)$,

$$D_u^2(F-\mu_*)^2(w,w) = 2D_{u_{\mu_*}}F(\Phi_{g^{-1}}w)\cdot D_{u_{\mu_*}}F(\Phi_{g^{-1}}w). \tag{7.24}$$

It is then clear that $D_u^2(F-\mu_*)^2(w,w) \leq C_{\mu_*}\|w\|^2$ for all $u \in \mathscr{O}_{u_{\mu_*}}$ and $w \in E$, and hypothesis (b) of Proposition 7.7 is satisfied by $\mathscr{L}_K$. In addition (7.23) together with the fact that the Φ_g preserve the inner product $\langle\cdot,\cdot\rangle$ shows that (7.21) implies (7.15).

Now let $w \in (T_u\mathscr{O}_{u_{\mu_*}})^\perp$ and write $w = w_1 + w_2$ with $w_1 \in T_u\Sigma_{\mu_*} \cap (T_u\mathscr{O}_{u_{\mu_*}})^\perp$ and $w_2 \in W_u \cap (T_u\mathscr{O}_{u_{\mu_*}})^\perp$. Then

$$\begin{aligned} D_u^2\mathscr{L}_K(w,w) &= D_u^2\mathscr{L}_{\mu_*}(w,w) + 2KD_{u_{\mu_*}}F(\Phi_{g^{-1}}w_2)\cdot D_{u_{\mu_*}}F(\Phi_{g^{-1}}w_2) \\ &\geq D_u^2\mathscr{L}_{\mu_*}(w_1,w_1) - C(\|w_1\|\|w_2\| + \|w_2\|^2) \\ &\quad + 2KD_{u_{\mu_*}}F(\Phi_{g^{-1}}w_2)\cdot D_{u_{\mu_*}}F(\Phi_{g^{-1}}w_2) \\ &\geq c\|w_1\|^2 - C(\|w_1\|\|w_2\| + \|w_2\|^2) + Kc_{\mu_*}\|w_2\|^2, \end{aligned}$$

where in the last line we use the fact that $\dim W_{u_{\mu_*}} = m$ and $D_{u_{\mu_*}}F_{|W_{u_{\mu_*}}} : W_{u_{\mu_*}} \to \mathbb{R}^m$ is an isomorphism. Finally, thanks to Young's inequality, there exists $\varepsilon > 0$ so that

$$D_u^2\mathscr{L}_K(w,w) \geq \left(c - \frac{C\varepsilon}{2}\right)\|w_1\|^2 + \left(Kc_{\mu_*} - C - \frac{C}{2\varepsilon}\right)\|w_2\|^2 \geq \tilde{c}\|w\|^2$$

with $\tilde{c} > 0$ provided that $K > 0$ is chosen large enough. As a consequence, using the same arguments as in the proof of Proposition 7.7, we conclude that there exist $\eta > 0, c > 0$ so that

$$\forall u \in \mathscr{O}_{u_{\mu_*}}, \forall v \in E, \quad \mathrm{d}(u,v) \leq \eta \Rightarrow \mathscr{L}_K(v) - \mathscr{L}_K(u) \geq c\mathrm{d}^2(v, \mathscr{O}_{u_\mu})$$

which implies, thanks to Theorem 7.1, that all $u \in \mathscr{O}_{u_{\mu_*}}$ are orbitally stable G_{μ_*}-relative equilibria. □

8. Plane wave stability on the torus for NLS

In this section we will illustrate the general theory described above on a simple example, that is the orbital stability of plane waves of the cubic focusing and defocusing nonlinear Schrödinger equation on the one-dimensional torus. More precisely, let us consider the cubic Schrödinger equation

$$i\partial_t u(t,x) + \beta \partial_{xx}^2 u(t,x) + \lambda |u(t,x)|^2 u(t,x) = 0 \tag{8.1}$$

in the space periodic setting $\mathbb{T}_L$, the one-dimensional torus of length $L > 0$, and with $u(t,x) \in \mathbb{C}$. The constants β and λ are parameters of the model; $\beta\lambda < 0$ corresponds to the defocusing case and $\beta\lambda > 0$ to the focusing one. In what follows, we fix $\beta > 0$.

Using the same arguments as in Section 5.5, we can show that Equation (8.1) is the Hamiltonian differential equation associated to the function H defined by

$$H(u) = \frac{1}{2}\left(\beta \int_0^L |\partial_x u(x)|^2 \,\mathrm{d}x - \frac{\lambda}{2}\int_0^L |u(x)|^4 \,\mathrm{d}x\right). \tag{8.2}$$

As before the symplectic Banach triple is given by $(E, \mathscr{D}, \mathscr{J})$ with $E = H^1(\mathbb{T}_L, \mathbb{C})$, $\mathscr{D} = H^3(\mathbb{T}_L, \mathbb{C})$, both viewed as real Hilbert spaces, and $\mathscr{J} u = iu$ (see Section 5.4 to understand how a complex Hilbert space can be viewed as a real Hilbert space with symplectic structure). We recall that the scalar product on $E = H^1(\mathbb{T}_L, \mathbb{C})$ is

$$(u,v)_E = \mathrm{Re}\int_0^L (\partial_x u(x)\partial_x \bar{v}(x) + u(x)\bar{v}(x))\,\mathrm{d}x \quad u, v \in E, \tag{8.3}$$

and the dual space E^* can be identified with $H^{-1}(\mathbb{T}_L, \mathbb{C})$ through the pairing

$$\langle u, v\rangle = \mathrm{Re}\int_0^L u(x)\bar{v}(x)\,\mathrm{d}x, \quad u \in E^*,\ v \in E. \tag{8.4}$$

Moreover, since the action Φ of the group $G = \mathbb{R}\times\mathbb{R}$ defined by $\Phi_{a,\gamma}(u) = e^{i\gamma}u(x-a)$ is globally Hamiltonian (see Section 5.5) and $H \circ \Phi_g = H$ (see Section 2.4), the quantities

$$F_1(u) = -\frac{i}{2}\int_0^L \bar{u}(x)\partial_x u(x)\,\mathrm{d}x, \tag{8.5}$$

$$F_2(u) = -\frac{1}{2}\int_0^L |u(x)|^2\,\mathrm{d}x \tag{8.6}$$

are constants of the motion.

As pointed out in Section 2.4, the two-parameter family of plane waves

$$u_{\alpha,k}(t,x) = \alpha e^{-ikx}e^{i\xi t} \tag{8.7}$$

with $\xi \in \mathbb{R}$, $k \in \frac{2\pi}{L}\mathbb{Z}$ and $\alpha \in \mathbb{R}$ are G-relative equilibria of (8.1) whenever ξ, k and α satisfy the dispersion relation

$$\xi + \beta k^2 = \lambda |\alpha|^2. \tag{8.8}$$

In the notation of the previous sections, $u_{\alpha,k} = u_{\mu_{\alpha,k}}$ with $\mu_{\alpha,k} \in \mathbb{R}^2$ given by

$$\mu_{\alpha,k} = \begin{pmatrix} F_1(u_{\alpha,k}) \\ F_2(u_{\alpha,k}) \end{pmatrix} = -\frac{\alpha^2}{2} L \begin{pmatrix} k \\ 1 \end{pmatrix}.$$

Remark that in this case $\mu_{\alpha,k}$ is not a regular value of $F = (F_1, F_2)$ as is readily checked (see Definition A.1.3).

The G-orbit of the initial condition $u_{\mu_{\alpha,k}}(x) = \alpha e^{-ikx}$ is given by

$$\mathscr{O}_{u_{\mu\,\alpha,k}} = \left\{\alpha e^{i\gamma}e^{-ik(x-a)}, (a,\gamma) \in G\right\}. \tag{8.9}$$

Our goal is to investigate the orbital stability of these particular solutions by applying the general arguments presented above. Our main result is the following theorem showing the orbital stability of plane waves in the defocusing case ($\lambda < 0$) as well as in the focusing case provided $0 < 2\lambda|\alpha|^2 < \beta\left(\frac{2\pi}{L}\right)^2$.

Theorem 8.1. *If $\beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2 > 0$, then all $u \in \mathcal{O}_{u_{\mu_{\alpha,k}}}$ are orbitally stable relative equilibria.*

Furthermore, in the case $\beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2 < 0$, we can investigate the linear stability of the plane waves and we obtain the following theorem.

Theorem 8.2. *Let the plane wave $u_{\alpha,k}(t,x) = \alpha e^{i(\xi t - kx)}$ be a solution to* (8.1) *and $\beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2 < 0$. Then the spectrum of the linearization of* (8.1) *around $u_{\alpha,k}$ in $L^2(\mathbb{T}_L)$ has eigenvalues with strictly positive real part. Consequently, this wave is spectrally unstable in $L^2(\mathbb{T}_L)$.*

This second result follows from a rather straightforward computation that we do not reproduce here.

As discussed in the introduction, the nonlinear (in)stability of plane waves for the cubic focusing and defocusing nonlinear Schrödinger equation in a one-dimensional space is a result known to the experts in the field (see the introduction of [GH07b, GH07a], for example). We did not however find a complete proof of it in the literature, so we furnish one here as an illustration of the general theory presented in the previous sections.

In [Zhi01], a related but slightly different analysis is proposed. The cubic nonlinear Schrödinger equation is defined on the entire line $\mathbb{R}$ and not on the one-dimensional torus $\mathbb{T}_L$. Using the Galilean invariance of the equation (see Section 2.7), the stability of any plane wave is equivalent to that of $u(t,x) = \alpha e^{i\lambda|\alpha|^2 t}$. The main result on stability of plane waves of [Zhi01] is given in Theorem III.3.1. It states that, in the defocusing case ($\lambda < 0$), the plane wave $u(t,x) = \alpha e^{i\lambda|\alpha|^2 t}$ is orbitally stable under small perturbations in $H^1(\mathbb{R})$.

Our approach is different: we focus on the Schrödinger equation on a one-dimensional torus. Our functions live on a torus and the perturbations too. In other words, our definition of stability is with respect to perturbations within $H^1(\mathbb{T}_L) = H^1_{\text{per}}([0,L])$. Moreover in Zhidkov's book nothing is said about the (in)stability of plane waves in the focusing case, a situation we cover partially.

8.1. **Orbital Stability**

To study the stability of $u_{\mu_{\alpha,k}}(x)$, it is useful to write the solutions of (8.1) in the form

$$u(t,x) = e^{-ikx}U(t,x) \tag{8.10}$$

where $U(t,x)$ is a function which satisfies the evolution equation

$$i\partial_t U + \beta\partial_{xx}^2 U - 2i\beta k\partial_x U + \lambda|U|^2 U - \beta k^2 U = 0. \tag{8.11}$$

Equation (8.11) is the Hamiltonian differential equation associated to the function $\tilde{H}$ defined by

$$\tilde{H}(U) = H(U) - 2\beta k F_1(U) - \beta k^2 F_2(U). \tag{8.12}$$

As before, the action Φ of the group $G = \mathbb{R}\times\mathbb{R}$ defined by $\Phi_{a,\gamma}(u) = e^{i\gamma}u(x-a)$ is globally Hamiltonian, $\tilde{H}\circ\Phi_g = \tilde{H}$ and the quantities F_1, F_2 defined by (8.5) and (8.6) are constants of the motion.

If ξ, k and α satisfy the dispersion relation (8.8), $U_{\mu_\alpha}(t,x) = \alpha e^{i\xi t}$ is a solution to (8.11). Moreover, $U_{\mu_\alpha}(x) = U_{\mu_\alpha}(0,x) = \alpha$ is a one-parameter family of G-relative equilibria and

our goal is to study their stability. Here $\mu_\alpha = -\frac{\alpha^2}{2}L\begin{pmatrix}0\\1\end{pmatrix}$ and, as above, μ_α is not a regular value of $F=(F_1,F_2)$.

Recall that the G-orbit of $U_{\mu_\alpha}(x)=\alpha$ is

$$\mathscr{O}_{U_{\mu_\alpha}} = \left\{e^{i\gamma}\alpha, \gamma\in[0,2\pi)\right\}. \tag{8.13}$$

and, by definition, $U\in\mathscr{O}_{U_{\mu\alpha}}$ is orbitally stable if

$$\forall\varepsilon,\exists\delta,\forall W\in E,\ \left(\mathrm{d}(W,U)\le\delta\Rightarrow\forall t\in\mathbb{R},\ \mathrm{d}(W(t,\cdot),\mathscr{O}_{U_{\mu\alpha}})\le\varepsilon\right)$$

(see Definition 3.1).

Proposition 8.3. *Let $\beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2 > 0$. Then every $U\in\mathscr{O}_{U_{\mu\alpha}}$ is orbitally stable.*

Our stability result in Theorem 8.1 is an immediate consequence of the previous statement since the change of variables $u\to U$ is bounded in E.

Now, to prove this proposition, we would like to apply the general results given in the previous section and more precisely Theorem 7.1. The idea is to construct a Lyapunov function $\mathscr{L}$ which is a group invariant constant of the motion and such that $D\mathscr{L}$ vanishes on $\mathscr{O}_{U_{\mu\alpha}}$. Moreover, we require $\mathscr{L}$ to be coercive on $\mathscr{O}_{U_{\mu\alpha}}$, which means here that there exist $\delta>0$ and $c>0$, depending only on β,L,λ and $|\alpha|^2$, such that, for all $W\in E$,

$$\mathrm{d}(W,\mathscr{O}_{U_{\mu\alpha}})\le\delta\Rightarrow\mathscr{L}(W)-\mathscr{L}(\alpha)\ge c\mathrm{d}(W,\mathscr{O}_{U_{\mu\alpha}})^2. \tag{8.14}$$

Note that, since the group action is unitary in the present setting, this is equivalent to (7.3). Since U_{μ_α} is a G-relative equilibrium, Theorem 6.1 ensures that it satisfies

$$D_{U_{\mu\alpha}}\tilde{H} - \tilde{\xi}\cdot D_{U_{\mu\alpha}}F = 0$$

for some $\tilde{\xi}\in\mathbb{R}^2$. As a consequence, $\tilde{H}-\tilde{\xi}\cdot F$ is a good candidate to be a Lyapunov function. Nevertheless, since $D_{U_{\mu\alpha}}F_1=0$, μ_α is not a regular value of F, and the choice of $\tilde{\xi}\in\mathbb{R}^2$ is not unique. A convenient choice for $\mathscr{L}_{\mu_\alpha}$ is

$$\mathscr{L}_{\mu_\alpha}(U) = H(U) - (\xi+\beta k^2)F_2(U), \tag{8.15}$$

which corresponds to $\tilde{\xi}=\begin{pmatrix}-2\beta k\\ \xi\end{pmatrix}$. By construction, $D_U\mathscr{L}_{\mu_\alpha}$ vanishes for $U\in\mathscr{O}_{U_{\mu\alpha}}$. Indeed, since $D_U\mathscr{L}_{\mu_\alpha}\in E^*$, $D_U\mathscr{L}_{\mu_\alpha}(V)=\langle D_U\mathscr{L}_{\mu_\alpha},V\rangle$ with

$$D_U\mathscr{L}_{\mu_\alpha} = -\beta\partial_{xx}^2U - \lambda|U|^2U + (\xi+\beta k^2)U\in H^{-1}(\mathbb{T}_L,\mathbb{C}), \tag{8.16}$$

so clearly $D_U\mathscr{L}_{\mu_\alpha}=0$ if $U\in\mathscr{O}_{U_{\mu\alpha}}$. Furthermore, the bilinear form $D_U^2\mathscr{L}_{\mu_\alpha}:E\times E\to\mathbb{R}$ is given by $D_U^2\mathscr{L}(V,V)=\langle\nabla^2\mathscr{L}_{\mu_\alpha}(U)V,V\rangle$ with

$$\nabla^2\mathscr{L}_{\mu_\alpha}(U)V = -\beta\partial_{xx}^2V - \lambda|U|^2V - \lambda(|U|^2V+\bar{V}U^2) + (\xi+\beta k^2)V\in H^{-1}(\mathbb{T}_L;\mathbb{C}); \tag{8.17}$$

in particular, for all $U\in E$, $\nabla^2\mathscr{L}_{\mu_\alpha}(U)$ is a bounded linear operator from E to E^* and the expression above makes sense.

Now, as in Section 4.2, to prove (8.14), the main ingredient is the property:

$$\exists c>0,\forall V\in\left(T_{U_{\mu\alpha}}\mathscr{O}_{U_{\mu\alpha}}\right)^\perp,\ D^2_{U_{\mu\alpha}}\mathscr{L}_{\mu_\alpha}(W,W)\ge c\|W\|^2,$$

where

$$\left(T_{U_{\mu\alpha}}\mathscr{O}_{U_{\mu\alpha}}\right)^\perp = \{W\in E,\langle i,W\rangle=0\}.$$

This is proven in the following proposition, from which coercivity is deduced in Proposition 8.6.

Proposition 8.4. *If* $\beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2 > 0$ *then, given* $c_\lambda = \min\left\{\beta\left(\frac{2\pi}{L}\right)^2, \beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2\right\}$,

$$D^2_{U_{\mu_\alpha}}\mathscr{L}_{\mu_\alpha}(W,W) = \langle \nabla^2\mathscr{L}_{\mu_\alpha}(U_{\mu_\alpha})V, V\rangle \geq \frac{c_\lambda}{2}\|V\|^2 \tag{8.18}$$

for all $V \in \left(T_{U_{\mu_\alpha}}\mathscr{O}_{U_{\mu_\alpha}}\right)^\perp$.

Proof. We will show that the linear operator $\nabla^2\mathscr{L}_{\mu_\alpha}(\alpha)$ has a basis of eigenvectors belonging to E. Let $\zeta \in \mathbb{C}$ be an eigenvalue of $\nabla^2\mathscr{L}_{\mu_\alpha}(\alpha)$. Hence, by definition, there exists $V = v_1 + iv_2 = (v_1, v_2) \in E$, such that

$$\zeta\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} -\beta\partial^2_{xx} - 2\lambda|\alpha|^2 & 0 \\ 0 & -\beta\partial^2_{xx} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

is satisfied. Now, v_1 and v_2 are real functions on the torus and we can write them in Fourier representation, namely,

$$v_1(x) = \frac{a_0(v_1)}{2} + \sum_{n=1}^{\infty} a_n(v_1)\cos\left(\frac{2\pi}{L}nx\right) + b_n(v_1)\sin\left(\frac{2\pi}{L}nx\right),$$
$$v_2(x) = \frac{a_0(v_2)}{2} + \sum_{n=1}^{\infty} a_n(v_2)\cos\left(\frac{2\pi}{L}nx\right) + b_n(v_2)\sin\left(\frac{2\pi}{L}nx\right).$$

As a consequence, for $n = 0$ we obtain

$$\begin{cases} \zeta\dfrac{a_0(v_1)}{2} = -2\lambda|\alpha|^2\dfrac{a_0(v_1)}{2}, \\ \zeta\dfrac{a_0(v_2)}{2} = 0, \end{cases}$$

and, for all $n \geq 1$, we have the following system:

$$\begin{cases} \zeta a_n(v_1) = \left(\beta\left(\dfrac{2\pi}{L}n\right)^2 - 2\lambda|\alpha|^2\right)a_n(v_1), \\ \zeta b_n(v_1) = \left(\beta\left(\dfrac{2\pi}{L}n\right)^2 - 2\lambda|\alpha|^2\right)b_n(v_1), \\ \zeta a_n(v_2) = \beta\left(\dfrac{2\pi}{L}n\right)^2 a_n(v_2), \\ \zeta b_n(v_2) = \beta\left(\dfrac{2\pi}{L}n\right)^2 b_n(v_2). \end{cases}$$

It follows that the set of the eigenvalues of $\nabla^2\mathscr{L}_{\mu_\alpha}(\alpha)$ is given by

$$\left\{\beta\left(\frac{2\pi}{L}n\right)^2,\ n \in \mathbb{N}\right\} \cup \left\{\beta\left(\frac{2\pi}{L}n\right)^2 - 2\lambda|\alpha|^2,\ n \in \mathbb{N}\right\}.$$

It is clear that this is a subset of $\mathbb{R}_+$ whenever $\beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2 \geq 0$.

Moreover, if $\beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2 > 0$, the kernel of $\nabla^2\mathscr{L}_{\mu_\alpha}(\alpha)$ is given by

$$\ker(\nabla^2\mathscr{L}_{\mu_\alpha}(\alpha)) = \operatorname{span}_{\mathbb{R}}\left\{\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\} = \operatorname{span}_{\mathbb{R}}\{i\}$$

and the coercivity property of $\nabla^2\mathscr{L}_{\mu_\alpha}(\alpha)$ on $(\ker(\nabla^2\mathscr{L}_{\mu_\alpha}(\alpha)))^\perp = \left(T_{U_{\mu_\alpha}}\mathscr{O}_{U_{\mu_\alpha}}\right)^\perp$ follows easily. □

The following lemma gives a representation of the elements of E which are close to the G-orbit $\mathscr{O}_{U_{\mu_\alpha}}$. It is used in the proof of Proposition 8.6 and is a special case of Lemma 7.8. We give a direct proof in the current simple setting.

Lemma 8.5. *There exists $\delta > 0$ such that any $W \in E$ with $\mathrm{d}(W, \mathscr{O}_{U_{\mu_\alpha}}) \leq \delta$ can be represented as*

$$e^{i\gamma} W = \alpha + V \tag{8.19}$$

with $\gamma = \gamma(W) \in [0, 2\pi)$ and $V \in \left(T_{U_{\mu_\alpha}} \mathscr{O}_{U_{\mu_\alpha}}\right)^{\perp}$. Moreover, there exists a positive constant C such that

$$\mathrm{d}(W, \mathscr{O}_{U_{\mu_\alpha}}) \leq \|V\| \leq C \mathrm{d}(W, \mathscr{O}_{U_{\mu_\alpha}}). \tag{8.20}$$

Proof. Let $W \in E$ such that $\mathrm{d}(W, \mathscr{O}_{U_{\mu_\alpha}}) < \delta$ with $\delta > 0$ sufficiently small. Hence there exists $\tilde{\gamma}$, which depends on W, such that

$$\|e^{i\tilde{\gamma}} W - \alpha\| \leq 2 \inf_{\lambda \in [0, 2\pi)} \|W - e^{i\lambda} \alpha\| \leq 2\delta$$

Next, consider the functional

$$\begin{aligned} \mathscr{F} : E \times \mathbb{R} &\to \mathbb{R} \\ (v, \phi) &\to \langle e^{i\phi} v, i \rangle = -\mathrm{Re} \int_0^L i e^{i\phi} v(x) \, \mathrm{d}x. \end{aligned}$$

Since $\mathscr{F}(\alpha, 0) = 0$ and $\partial_\phi \mathscr{F}(\alpha, 0) = \alpha L \neq 0$, by means of the implicit function theorem, we can conclude that there exists $\Lambda : \mathscr{V} \to (-\varepsilon, \varepsilon)$ with $\mathscr{V}$ a neighbourhood of α in E and $\varepsilon > 0$ sufficiently small, such that if $v \in \mathscr{V}$ then there exists a unique $\phi = \Lambda(v) \in (-\varepsilon, \varepsilon)$ for which we have $\langle e^{i\phi} v, i \rangle = 0$.

As a consequence, since $\|e^{i\tilde{\gamma}} W - \alpha\| < 2\delta$, if we choose $\delta > 0$ sufficiently small then there exists $\phi \in \mathbb{R}$ such that $\langle e^{i(\phi + \tilde{\gamma})} W, i \rangle = 0$. By taking $\gamma = \tilde{\gamma} + \phi$ modulo 2π, we obtain (8.19). Indeed, $E = T_{U_{\mu_\alpha}} \mathscr{O}_{U_{\mu_\alpha}} \oplus \left(T_{U_{\mu_\alpha}} \mathscr{O}_{U_{\mu_\alpha}}\right)^{\perp}$ and $T_{U_{\mu_\alpha}} \mathscr{O}_{U_{\mu_\alpha}} = \mathrm{span}_{\mathbb{R}} \{i\}$. Hence,

$$e^{i\gamma} W - \alpha = ai + V$$

with $a \in \mathbb{R}$ and $V \in \left(T_{U_{\mu_\alpha}} \mathscr{O}_{U_{\mu_\alpha}}\right)^{\perp}$. As a consequence,

$$0 = \langle e^{i\gamma} W - \alpha, i \rangle = a \langle i, i \rangle$$

and a has to be equal to 0.

Estimate (8.20) follows directly from the definition of V. □

Finally, the following proposition proves the coercivity of $\mathscr{L}_{\mu_\alpha}$ on $\mathscr{O}_{U_\alpha}$.

Proposition 8.6. *Let $\beta \left(\frac{2\pi}{\beta}\right)^2 - 2\lambda |\alpha|^2 > 0$, $\mathscr{L}_{\mu_\alpha}$ be defined as in* (8.15)*, and let $c_\lambda = \min \left\{ \beta \left(\frac{2\pi}{L}\right)^2, \beta \left(\frac{2\pi}{L}\right)^2 - 2\lambda |\alpha|^2 \right\}$. Then there exists $\delta > 0$ such that*

$$\mathscr{L}_{\mu_\alpha}(W) - \mathscr{L}_{\mu_\alpha}(U_{\mu_\alpha}) \geq \frac{c_\lambda}{4} \mathrm{d}(W, \mathscr{O}_{U_{\mu_\alpha}})^2 \tag{8.21}$$

for all $W \in E$, such that $\mathrm{d}(W, \mathscr{O}_{U_{\mu_\alpha}}) \leq \delta$.

Proof. Let $W \in E$ such that $\mathrm{d}(W, \mathscr{O}_{U_{\mu_\alpha}}) \leq \delta$ with $\delta > 0$ sufficiently small. By Lemma 8.5, there exists $\gamma \in [0, 2\pi]$ such that $e^{i\gamma} W - \alpha = V$ with $V \in \left(T_{U_{\mu_\alpha}} \mathscr{O}_{U_{\mu_\alpha}}\right)^{\perp}$ and $\|V\| \leq$

$C\mathrm{d}(W,\mathscr{O}_{U_{\mu_\alpha}})$. As a consequence, since $D_\alpha \mathscr{L}_{\mu_\alpha} = 0$,

$$\begin{aligned}\mathscr{L}_{\mu_\alpha}(W) - \mathscr{L}_{\mu_\alpha}(\alpha) =& \mathscr{L}_{\mu_\alpha}(e^{i\gamma}W) - \mathscr{L}_{\mu_\alpha}(\alpha)\\ =& \frac{1}{2} D^2_\alpha \mathscr{L}_{\mu_\alpha}(e^{i\gamma}W-\alpha, e^{i\gamma}W-\alpha) + \mathrm{o}(\|e^{i\gamma}W-\alpha\|^2)\\ =& \frac{1}{2} D^2_\alpha \mathscr{L}_{\mu_\alpha}(V,V) + \mathrm{o}(\|V\|^2)\\ \geq& \frac{1}{2} c_\lambda \|V\|^2 + \mathrm{o}(\|V\|^2)\end{aligned}$$

with c_λ defined in (8.18). Therefore for $\mathrm{d}(W,\mathscr{O}_{U_{\mu_\alpha}})$ small,

$$\mathscr{L}_{\mu_\alpha}(W) - \mathscr{L}_{\mu_\alpha}(\alpha) \geq \frac{c_\lambda}{4} \mathrm{d}(W,\mathscr{O}_{U_{\mu_\alpha}})^2.$$

□

Now, a straightforward application of the proof of Theorem 7.1 with $\mathscr{L}_{\mu_\alpha}$ as Lyapunov function allows us to conclude that $\mathscr{O}_{U_{\mu_\alpha}}$ is orbitally stable under small perturbations in E whenever $\beta\left(\frac{2\pi}{L}\right)^2 - 2\lambda|\alpha|^2 > 0$.

9. Orbital stability for inhomogeneous NLS

This section is concerned with an NLS equation of the form

$$i\partial_t u + \Delta u + f(x,|u|^2)u = 0, \quad u = u(t,x) : \mathbb{R}\times\mathbb{R}^d \to \mathbb{C}. \tag{9.1}$$

We consider standing wave solutions $u(t,x) = e^{i\xi t}w(x)$, where $w : \mathbb{R}^d \to \mathbb{R}$ is localized[17] – typically $w \in H^1(\mathbb{R}^d)$ and $w(x) \to 0$ exponentially as $|x| \to \infty$. Such a solution exists if and only if

$$\Delta w - \xi w + f(x,w^2)w = 0, \tag{9.2}$$

which is precisely the "stationary equation" (6.3). Note that the notation for the nonlinearity in (9.1) is slightly different than in Section 2.4, and automatically ensures that (2.26) holds, for all $u \in \mathbb{C}\setminus\{0\}$.

The existence of solutions of (9.2) can be obtained under various hypotheses on f, the easiest case being the pure power nonlinearity, $f(x,w^2) = |w|^{\sigma-1}$, $\sigma > 1$. Note that, unlike in the case of periodic boundary conditions studied in the previous section, it is crucial here that the nonlinearity be focusing for standing waves to exist. The stationary equation (9.2) has no solutions if, for instance, $f(x,w^2) = -|w|^{\sigma-1}$. In the sequel, we will indeed suppose that the nonlinearity is focusing, which in the context of (9.1) means that $f(x,s)$ is positive and increasing in $s > 0$.

The purpose of this section is to further illustrate the general stability theory developed in Section 7. Orbital stability results for standing waves of (9.1) have been obtained in [GS08, Gen09, Gen10a, Gen13] and will be summarized here. The stability analysis in these papers benefits from having solution curves $\xi \to w_\xi$. In the setting of Section 7, they can be seen as an application of Theorem 7.5. The approach used in [GS08, Gen09, Gen10a, Gen13] was to apply the celebrated Theorem 2 of Grillakis, Shatah, Strauss [GSS87]. This result essentially relies on the set of spectral conditions (S1)–(S3), formulated below in the context of (9.1), together with a convexity condition, which here takes the form (9.14). In the framework developed in these notes, the role of Theorem 2 of [GSS87] can be interpreted as follows. It will be shown in Proposition 9.8 that the conditions (S1)–(S3) and (9.14) ensure that the coercivity property (7.16) required

[17] Note that we focus here on situations where the wave profile $w(x)$ is real-valued.

by Proposition 7.7 is satisfied at the relative equilibrium w_ξ. Theorem 7.5 can then be applied. As already mentioned in the introduction to Section 7, and explained in more detail after the proof of Proposition 9.8, the relative equilibria of (9.1) can be parametrized equivalently by the parameter ξ appearing in (9.2), or by the corresponding value $\mu = \frac{1}{2}\|w_\xi\|_{L^2}^2$ of the constant of the motion. It turns out that using ξ is more convenient here. Note that, since this constant of the motion satisfies Hypothesis F, one could also apply Theorem 7.4 instead of Theorem 7.5.

The notion of orbital stability we shall be concerned with here is that corresponding to the group action (5.18) of Section 5.5. Note however that the explicit spatial dependence in (9.1) breaks the invariance under translations, and one rather needs to consider the restricted action Φ_γ on the phase space $E = H^1(\mathbb{R}^d, \mathbb{C})$,

$$\Phi_\gamma(u) = e^{i\gamma}u(x), \quad u \in E,\ \gamma \in \mathbb{R}. \tag{9.3}$$

The standing waves corresponding to solutions w_ξ of the stationary equation (9.2) are then relative equilibria for the dynamics of (9.1), with respect to the action Φ_γ.

Remark 9.1. If f does not depend on x then the full group action (5.18) is to be considered, and the standing waves of (9.1) are in general not orbitally stable in the sense of (9.3). Orbital stability in the sense of the full group action (5.18) was proved by Cazenave and Lions [CL82] by variational arguments.

We will only consider here situations where the coefficient f explicitly depends on the space variable $x \in \mathbb{R}^d$ – (9.1) is then often referred to as an *inhomogeneous NLS* –, and decays as $|x| \to \infty$, in a sense that will be made more precise below. We shall also suppose that $f(x, w^2) \sim V(x)|w|^{\sigma-1}$ as $w \to 0$. Conditions relating the function V and the power $\sigma > 1$ will be given for stability of standing waves to hold. In particular our assumptions will imply $\sigma < 1 + \frac{4}{d-2}$, so that local existence in $H^1(\mathbb{R}^d)$ for the Cauchy problem associated with (9.1) is ensured by the results of Section 2.4. Two cases will be considered:

(PT) the power-type nonlinearity $f(x, w^2) = V(x)|w|^{\sigma-1}$;

(AL) the asymptotically linear case $f(x, w^2) \to V(x)$ as $|w| \to \infty$
(e.g. with $f(x, w^2) = V(x)\frac{|w|^{\sigma-1}}{1+|w|^{\sigma-1}}$).

We will give a short account of the main arguments used in [GS08, Gen09, Gen10a, Gen13] to establish the stability of standing waves along a global solution curve. We will also briefly sketch the bifurcation analysis yielding a smooth branch of non-trivial solutions of (9.2) emerging from the trivial solution $w = 0$. This part of the argument is crucial since, in the approach originally developed in [GS08], the spectral properties and the condition (9.14) required to obtain the coercivity of an appropriate Lyapunov functional are derived by continuation from the limit $w_\xi \to 0$. It is worth emphasizing here that the verification of these hypotheses is precisely that part of the stability analysis which strongly relies on the model considered. Once the required coercivity properties are established, the orbital stability can be deduced from the abstract results of Section 7.

9.1. **Hamiltonian setting**

Similarly to Section 8, we work here with

$$E = H^1(\mathbb{R}^d, \mathbb{C}), \quad (u, v)_E = \mathrm{Re}\int_{\mathbb{R}^d} \nabla u(x) \cdot \nabla \bar{v}(x) + u(x)\bar{v}(x)\,\mathrm{d}x.$$

The Hamiltonian and the charge are respectively defined by $H, Q : E \to \mathbb{R}$,

$$H(u) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla u|^2 \,\mathrm{d}x - \frac{1}{2}\int_{\mathbb{R}^d}\int_0^{|u|^2} f(x,s)\,\mathrm{d}s\,\mathrm{d}x, \quad Q(u) = \frac{1}{2}\int_{\mathbb{R}^d} |u|^2\,\mathrm{d}x, \quad u \in E. \tag{9.4}$$

In the notation of Section 5.5, $Q(u) \equiv -F_{d+1}(u)$, but we will keep the customary notation Q here. Under our assumptions, $H, Q \in C^2(E,\mathbb{R})$.

Now (9.1) can precisely be written in the form

$$\mathscr{J}\dot{u}_t = D_{u_t} H \tag{9.5}$$

considered in Section 5, with $E = H^1(\mathbb{R}^d,\mathbb{C}) \simeq H^1(\mathbb{R}^d,\mathbb{R}) \times H^1(\mathbb{R}^d,\mathbb{R})$ and

$$\mathscr{J} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$$

with $I : H^1 \hookrightarrow H^{-1}$ the (dense) injection. That is, $\mathscr{J}(q,p) = (-p,q) \in E^*$, for all $(q,p) \in E$, as in Section 5.5. Note that we use the identification

$$H^1(\mathbb{R}^d,\mathbb{R}) \subset L^2(\mathbb{R}^d,\mathbb{R}) = L^2(\mathbb{R}^d,\mathbb{R})^* \subset H^{-1}(\mathbb{R}^d,\mathbb{R}).$$

In this setting a solution of (9.1) is a function $u \in C^1((-T_{\min}, T_{\max}), E)$, for some $T_{\min}, T_{\max} > 0$ (depending on $u(0)$), satisfying (9.5) for all $t \in (-T_{\min}, T_{\max})$. Standing waves are particular solutions of the form $u(t) = \Phi_{(\xi t)} w$, $w \in E$, and the stationary equation (9.2) now reads

$$D_w H + \xi D_w Q = 0. \tag{9.6}$$

Hence, the discussion in Sections 6 and 7 indicates that

$$\mathscr{L}_\xi = H + \xi Q \tag{9.7}$$

is the natural candidate for the Lyapunov function. Furthermore, the invariance of H and Q under the action of Φ_γ implies that

$$D_{\Phi_\gamma(w)} H + \xi D_{\Phi_\gamma(w)} Q = 0, \quad \gamma \in \mathbb{R}. \tag{9.8}$$

Finally, note that the isometric action (9.3) can equivalently be expressed as

$$\Phi_\gamma \begin{pmatrix} \operatorname{Re} u \\ \operatorname{Im} u \end{pmatrix} = \begin{pmatrix} \cos\gamma & -\sin\gamma \\ \sin\gamma & \cos\gamma \end{pmatrix} \begin{pmatrix} \operatorname{Re} u \\ \operatorname{Im} u \end{pmatrix}, \quad u \in E,\ \gamma \in \mathbb{R}.$$

9.2. **Bifurcation results**

In this section we present bifurcation results ensuring the existence of smooth curves of solutions of (9.2). From a bifurcation-theoretic viewpoint the peculiarity of these results is that, in both the (PT) and (AL) cases, bifurcation occurs from the essential spectrum of the linearization of (9.2), namely

$$\Delta w = \xi w,$$

this linear problem set on $\mathbb{R}^d$ having no eigenvalues.

We start with the power-type case (PT), that is, we first consider the problem

$$\Delta w(x) + V(x)|w(x)|^{\sigma-1} w(x) = \xi w(x), \quad w \in H^1(\mathbb{R}^d,\mathbb{R}), \tag{9.9}$$

where $d \geq 1$ and $V : \mathbb{R}^d \to \mathbb{R}$ satisfies:

(V1) $V \in C^1(\mathbb{R}^d)$;

(V2) there exists $b \in (0,2)$ ($b \in (0,1)$ if $d = 1$) such that

$$1 < \sigma < \tfrac{4-2b}{d-2} \quad \text{if } d \geq 3, \quad 1 < \sigma < \infty \quad \text{for } d = 1,2,$$

$$\lim_{|x|\to\infty} |x|^b V(x) = 1 \quad \text{and} \quad \lim_{|x|\to\infty} |x|^b [x \cdot \nabla V(x) + bV(x)] = 0;$$

(V3) V is radial with $V(r) > 0$ and $V'(r) < 0$ for $r > 0$;

(V4) $r\dfrac{V'(r)}{V(r)}$ is decreasing in $r > 0$ (and so $\to -b$ by (V2)).

Note that $V(x) = (1+|x|^2)^{-b/2}$ satisfies all of the above assumptions.

Theorem 9.2. *Suppose that the hypotheses (V1) to (V4) hold. Then there exists a curve $w \in C^1\big((0,\infty), H^1(\mathbb{R}^d)\big)$ such that, for all $\xi \in (0,\infty)$, $w_\xi \equiv w(\xi)$ is the unique positive radial solution of* (9.9)*, $w_\xi \in C^2(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$, and w_ξ is strictly radially decreasing, with $w_\xi(x), |\nabla w_\xi(x)| \to 0$ exponentially as $|x| \to \infty$. Furthermore, the asymptotic behaviour of the curve reads*

$$\lim_{\xi\to 0} \|w_\xi\|_{H^1} = \begin{cases} 0 & \text{if} \quad 1 < \sigma < 1 + \frac{4-2b}{d}, \\ \infty & \text{if} \quad 1 + \frac{4-2b}{d} < \sigma < 1 + \frac{4-2b}{d-2}, \end{cases}$$

and

$$\lim_{\xi\to\infty} \|w_\xi\|_{H^1} = \infty \quad \text{for all} \quad 1 < \sigma < 1 + \tfrac{4-2b}{d-2}.$$

This theorem has been proved in [Gen10b] by a combination of variational and analytical arguments. It provides a global continuation, in the radial case, of the local curve of solutions of (9.9) obtained in [GS08] (parametrized by $\xi \in (0,\xi_0)$, with $\xi_0 > 0$ small) under the much weaker assumptions (V1) and (V2). Note in particular that (V2) only requires the problem to be focusing at infinity, no further sign restrictions being imposed on V. The orbital stability of the solutions w_ξ, $\xi \in (0,\xi_0)$, is also discussed in [GS08], and it is found that they are stable provided

$$1 < \sigma < 1 + \tfrac{4-2b}{d}, \tag{9.10}$$

and unstable if $1 + \frac{4-2b}{d} < \sigma < 1 + \frac{4-2b}{d-2}$.

Remark 9.3. In fact, more information about the asymptotic behaviour as $\xi \to 0$ is obtained in [GS08]. In particular,

$$\lim_{\xi\to 0} \|w_\xi\|_{L^2} = \begin{cases} 0 & \text{if} \quad 1 < \sigma < 1 + \frac{4-2b}{d}, \\ \infty & \text{if} \quad 1 + \frac{4-2b}{d} < \sigma < 1 + \frac{4-2b}{d-2}, \end{cases}$$

whereas

$$\lim_{\xi\to 0} \|\nabla w_\xi\|_{L^2} = 0 \quad \text{for all } 1 < \sigma < 1 + \tfrac{4-2b}{d-2}.$$

We now state a global bifurcation result similar to Theorem 9.2, for (9.2) in dimension $d = 1$, in the asymptotically linear case (AL). That is, we consider

$$w''(x) + f(x, w(x)^2)w(x) = \xi w(x), \quad w \in H^1(\mathbb{R},\mathbb{R}), \tag{9.11}$$

where, to fix the ideas[18], we let

$$f(x,w^2) = V(x)\frac{|w|^{\sigma-1}}{1+|w|^{\sigma-1}}. \tag{9.12}$$

In the asymptotically linear case, one cannot expect to find positive solutions of (9.11)–(9.12) for large values of $\xi > 0$. Heuristically, letting $u \to \infty$ in (9.11)–(9.12) leads to the so-called asymptotic linearization

$$w''(x) + V(x)w(x) = \xi w(x), \tag{9.13}$$

[18]More general assumptions on the coefficient f in (AL) can be given, under which the bifurcation and stability results presented here still hold, see [Gen13].

having a ray of positive eigenfunctions $\{\mu w_\infty : \mu > 0\}$ corresponding to a principal eigenvalue $\xi_\infty > 0$. This has been put on rigorous grounds in [Gen11], where it is shown that positive even solutions of (9.11)–(9.12) only exist for $\xi < \xi_\infty$, and satisfy $\|w_\xi\|_{H^1} \to \infty$ as $\xi \to \xi_\infty$.

Theorem 9.4. *Suppose (V1) to (V3) and $1 < \sigma < 5 - 2b$. Then there exists a curve $w \in C^1\big((0,\xi_\infty), H^1(\mathbb{R})\big)$ such that, for all $\xi \in (0,\xi_\infty)$, w_ξ is the unique positive even solution of (9.11)–(9.12), $w_\xi \in C^2(\mathbb{R}) \cap H^2(\mathbb{R})$ with $w_\xi'(x) < 0$ for $x > 0$, and $w_\xi(x), w_\xi(x)' \to 0$ exponentially as $|x| \to \infty$. Furthermore, there holds*

$$\lim_{\xi \to 0} \|w_\xi\|_{H^1(\mathbb{R})} = 0 \quad \textit{and} \quad \lim_{\xi \to \xi_\infty} \|w_\xi\|_{H^1(\mathbb{R})} = \infty.$$

Remark 9.5. The reader might wonder why (V4) is not needed for Theorem 9.4. It turns out that this assumption is essential in the proof of Theorem 9.2, where it ensures uniqueness of positive radial solutions of (9.9), for any fixed $\xi > 0$. In the one-dimensional problem (9.11)–(9.12), uniqueness can be proved without invoking (V4).[19] However, we will see in the next section that this hypothesis is crucial to the stability analysis, in both the (PT) and (AL) cases.

Remark 9.6. Thanks to the form of the nonlinearity in (9.12) the global branch of Theorem 9.4, bifurcating from the trivial solution $u = 0$ at $\xi = 0$, is obtained by perturbation from the (PT) nonlinearity dealt with in Theorem 9.2. In fact, the case where asymptotic bifurcation occurs at $\xi = 0$, corresponding in dimension $d = 1$ to $5 - 2b < \sigma < \infty$, could also be extended to the (AL) case, where instability could be inferred, in the limit $\xi \to 0$. We refrain from going in this direction here since we were only able so far to extend the discussion to a global branch in the stable case. We shall therefore assume (9.10) from now on, both for (PT) and (AL).

9.3. **Stability**

In dimension $d = 1$, assuming that $1 < \sigma < 5 - 2b$, the global curves of standing wave solutions given by Theorems 9.2 and 9.4 are stable. This has been proved in [Gen10a] for the (PT) case and in [Gen13] for the (AL) case. The proofs rely on the theory of orbital stability in [GSS87] and we will now outline the main arguments.

We shall start by convincing the reader that, in the context of (9.1), one cannot hope for stability in the usual sense (1.1). Indeed, suppose $\xi_n \to \xi$ and consider

$$u_\xi(t,x) = e^{i\xi t} w_\xi(x) \quad \text{and} \quad u_n(t,x) = e^{i\xi_n t} w_{\xi_n}(x).$$

Then

$$\forall \delta > 0 \ \exists N_\delta \in \mathbb{N},\ n \geq N_\delta \Rightarrow \|u_n(0,\cdot) - u_\xi(0,\cdot)\|_{H^1} = \|w_{\xi_n} - w_\xi\|_{H^1} \leq \delta.$$

However,

$$\begin{aligned} \|u_n(t,\cdot) - u_\xi(t,\cdot)\|_{H^1} &\geq \big| |e^{i\xi t} - e^{i\xi_n t}| \|w_\xi\|_{H^1} - \|w_{\xi_n} - w_\xi\|_{H^1} \big| \\ &\Rightarrow \sup_{t \geq 0} \|u_n(t) - u_\xi(t)\|_{H^1} \geq 2\|w_\xi\|_{H^1} - \delta,\ n \geq N_\delta. \end{aligned}$$

Therefore, for n large enough, the initial datum $u_n(0)$ may be chosen δ-close to $u_\xi(0)$, $u_n(t)$ will nevertheless drift at least $2\|w_\xi\|_{H^1} - \delta$ far away from $u_\xi(t)$.

[19] Note that the main reason for restricting the discussion to $d = 1$ in Theorem 9.4 is the lack of uniqueness results in higher dimensions for the (AL) case.

Theorem 9.7. *Suppose that $d = 1$ and the hypotheses (V1) to (V4) are satisfied. Then the standing waves $u_\xi(t,x) = e^{i\xi t} w_\xi(x)$ of (9.1) given by either Theorem 9.2 or Theorem 9.4 are orbitally stable.*

The proofs of Theorem 9.7 given in [Gen10a, Gen13] used Theorem 2 of [GSS87], and so relied upon verifying Assumptions 1–3 of [GSS87], as well as the condition

$$\|w_\xi\|_{L^2} \text{ is strictly increasing in } \xi > 0. \tag{9.14}$$

The latter is often referred to as *the slope condition* or the *Vakhitov-Kolokolov condition*. It seems to have indeed first appeared in the paper [VK73] of Vakhitov and Kolokolov (1968), in the context of nonlinear optical waveguides.[20]

Assumption 1 of [GSS87] is about the well-posedness of the Cauchy problem for (9.1) which, under our hypotheses, follows from Section 2.4. Assumption 2 pertains to the existence of smooth solution curves and is ensured by Theorem 9.2/9.4. It is this property which allows us to apply Theorem 7.5 of Section 7.

We will see that Assumption 3 of [GSS87], together with the slope condition (9.14), ensure the required coercivity property of the Lyapunov function $\mathscr{L}_\xi$ introduced in (9.7). In order to formulate Assumption 3 in the present context, consider the bounded linear operator $D^2_{w_\xi}\mathscr{L}_\xi : E \to E^*$,

$$D^2_{w_\xi}\mathscr{L}_\xi = D^2_{w_\xi}H + \xi D^2_{w_\xi}Q, \quad \xi > 0. \tag{9.15}$$

We define the *spectrum* of $D^2_{w_\xi}\mathscr{L}_\xi$ as the following subset of $\mathbb{R}$:

$$\sigma(D^2_{w_\xi}\mathscr{L}_\xi) = \big\{\lambda \in \mathbb{R} : D^2_{w_\xi}\mathscr{L}_\xi - \lambda\widetilde{R} : E \to E^* \text{ is not an isomorphism}\big\}, \tag{9.16}$$

where $\widetilde{R} = \operatorname{diag}(R,R)$ and $R = -\frac{\mathrm{d}^2}{\mathrm{d}x^2} + 1 : H^1(\mathbb{R},\mathbb{R}) \to H^{-1}(\mathbb{R},\mathbb{R})$ is the Riesz isomorphism. Under the hypotheses of Theorem 9.2/9.4, $\widetilde{R}^{-1}D^2_{w_\xi}\mathscr{L}_\xi : E \to E$ is a bounded self-adjoint Schrödinger operator, and its spectrum coincides with $\sigma(D^2_{w_\xi}\mathscr{L}_\xi)$. The motivation for this definition of the spectrum of $D^2_{w_\xi}\mathscr{L}_\xi$ will be discussed in Remark 9.9.

A straightforward calculation shows that $D^2_{w_\xi}\mathscr{L}_\xi$ is explicitly given by

$$D^2_{w_\xi}\mathscr{L}_\xi = \begin{pmatrix} -\dfrac{\mathrm{d}^2}{\mathrm{d}x^2} + \xi - [f(x,w_\xi^2) + 2\partial_2 f(x,w_\xi^2)w_\xi^2] & 0 \\ 0 & -\dfrac{\mathrm{d}^2}{\mathrm{d}x^2} + \xi - f(x,w_\xi^2) \end{pmatrix}, \tag{9.17}$$

and the spectral conditions formulated in Assumption 3 of [GSS87] are:

(S1) $\exists \alpha_\xi \in \mathbb{R}$ such that $\sigma(D^2_{w_\xi}\mathscr{L}_\xi) \cap (-\infty, 0) = \{-\alpha_\xi^2\}$ and $\ker(D^2_{w_\xi}\mathscr{L}_\xi + \alpha_\xi^2\widetilde{R})$ is one-dimensional;
(S2) $\ker D^2_{w_\xi}\mathscr{L}_\xi = \operatorname{span}\{iw_\xi\}$;
(S3) $\sigma(D^2_{w_\xi}\mathscr{L}_\xi) \setminus \{-\alpha_\xi^2, 0\}$ is bounded away from zero.

The fact that $iw_\xi \in \ker D^2_{w_\xi}\mathscr{L}_\xi$ directly follows by differentiating (9.8) with respect to γ at $\gamma = 0$. So (S2) really only states that $\ker D^2_{w_\xi}\mathscr{L}_\xi$ is one-dimensional.

We now explain how hypotheses (S1)–(S3), together with (9.14), imply the coercivity property (7.16) in Proposition 7.7. In order to explicitly write down condition (7.16), let us

[20] The mathematical theory of NLS has been intimately connected to nonlinear optics from its early days. See [Gen10a] for additional references on this.

first observe that we parametrized the standing waves by the "frequency" ξ here, whereas in Section 7 the relative equilibria are rather labelled using the value μ of the constraint. In the present context, $\mu = \mu(\xi) = Q(w_\xi)$, and we only deal with situations where μ is a smooth, strictly increasing function of ξ, so both parametrizations are equivalent. Now the level surface

$$\Sigma_{Q(w_\xi)} = \{u \in E \mid Q(u) = Q(w_\xi)\}$$

and, given a standing wave $u_\xi(t) = \Phi_{(\xi t)} w_\xi$ we have, for any $u = e^{-i\gamma(u)} w_\xi \in \mathscr{O}_{u_\xi}$,

$$T_u \Sigma_{Q(w_\xi)} = \{v \in E \mid \langle e^{-i\gamma(u)} D_{w_\xi} Q, v\rangle = 0\}.$$

On the other hand, $T_u \mathscr{O}_{u_\xi} = \operatorname{span}\{e^{-i\gamma(u)} i w_\xi\}$, so that

$$T_u \Sigma_{Q(w_\xi)} \cap (T_u \mathscr{O}_{u_\xi})^\perp = \{v \in E \mid \langle e^{-i\gamma(u)} D_{w_\xi} Q, v\rangle = (e^{-i\gamma(u)} i w, v)_E = 0\}.$$

Next, differentiating

$$D_{w_\xi} H + \xi D_{w_\xi} Q = 0$$

with respect to ξ yields

$$D^2_{w_\xi} \mathscr{L}_\xi \chi_\xi = -D_{w_\xi} Q, \quad \text{where} \quad \chi_\xi := \frac{\mathrm{d} w_\xi}{\mathrm{d}\xi}, \tag{9.18}$$

so that

$$\langle D^2_{w_\xi} \mathscr{L}_\xi \chi_\xi, \chi_\xi\rangle = -\langle D_{w_\xi} Q, \chi_\xi\rangle = -\frac{\mathrm{d}}{\mathrm{d}\xi} Q(w_\xi) < 0 \tag{9.19}$$

by (9.14).

Proposition 9.8. *Suppose that (S1) to (S3) hold, as well as* (9.14)*. Then there exists $c > 0$ such that*

$$\forall u \in \mathscr{O}_{u_\xi}, \forall v \in T_u \Sigma_{Q(w_\xi)} \cap (T_u \mathscr{O}_{u_\xi})^\perp,\ D^2_u \mathscr{L}_\xi(v,v) \geq c\|v\|^2_E.$$

Proof. Let $u = e^{-i\gamma(u)} w_\xi \in \mathscr{O}_{u_\xi}$. First remark that, by the invariance of $\mathscr{L}$ on the orbit $\{\Phi_\gamma w_\xi \mid \gamma \in \mathbb{R}\}$, we have

$$D^2_u \mathscr{L}_\xi = D^2_{\Phi_{\gamma(u)} w_\xi} \mathscr{L}_\xi = D^2_{w_\xi}(\mathscr{L}_\xi \circ \Phi_{-\gamma(u)}) = D^2_{w_\xi} \mathscr{L}_\xi.$$

Therefore, we need only prove the result at $u = w_\xi$, i.e. that there exists $c > 0$ such that

$$\forall v \in E,\ \langle D_{w_\xi} Q, v\rangle = (i w_\xi, v)_E = 0 \Rightarrow D^2_{w_\xi} \mathscr{L}_\xi(v,v) \geq c\|v\|^2_E.$$

Introducing the bounded self-adjoint operator $S_\xi := \widetilde{R}^{-1} D^2_{w_\xi} \mathscr{L}_\xi : E \to E$, this is equivalent to

$$\forall v \in E,\ (S_\xi \chi_\xi, v)_E = (i w_\xi, v)_E = 0 \Rightarrow (S_\xi v, v)_E \geq c\|v\|^2_E.$$

Now by (9.19) we see that $(S_\xi \chi_\xi, \chi_\xi)_E < 0$, and the result readily follows from Lemma 5.3 in [Stu08]. □

The verification of properties (S1)–(S3) and of the slope condition (9.14) in [Gen10a, Gen13] is intimately connected with the behaviour as $\xi \to 0$ of the solutions given by Theorem 9.2/9.4. The main idea is to show that the required properties hold true for a limiting problem obtained by letting $\xi \to 0$ in the stationary equation (9.2) (in suitably rescaled variables), and then to deduce them for the original problem by perturbation and continuation along the global curve given by Theorem 9.2/9.4. In other words, it is first shown that (S1)–(S3) and (9.14) hold for small values of $\xi > 0$, and then that these properties cannot change along the global curve. It is worth noting here that, in both Theorem 9.2

and Theorem 9.4, it can be shown that $\|w_\xi\|_{L^\infty} \to 0$ as $\xi \to 0$ (see Section 9.3.1 below). Therefore, case (AL) can be seen as a perturbation of (PT), in the limit of small ξ, and the stability properties of standing waves are the same in both cases for small $\xi > 0$.

The remainder of this section is devoted to the proof of Theorem 9.7. We will sketch the arguments yielding the local stability results close to $\xi = 0$, and the continuation procedure extending these to the whole curves of solutions in Theorem 9.2 and Theorem 9.4. For the local results, we shall only consider case (PT), the details of the perturbation argument one has to go through to deal with (AL) being cumbersome and not very enligthening (see [Gen09] for more details). We will however present the global continuation procedure for both cases in a unified manner. For this we will use the general notation of (9.1)–(9.2) rather than the particular form of f in each case, and we will merely write $\xi > 0$ throughout, of course really meaning $0 < \xi < \xi_\infty$ in the (AL) case.

9.3.1. *Local stability by bifurcation.* We consider here (9.2) in dimension $d = 1$, and with $f(x,s^2) = V(x)|s|^{\sigma-1}$. The scaling

$$\xi = k^2, \quad u(x) = k^{\frac{2-b}{\sigma-1}} v(y), \quad y := kx, \quad k > 0, \tag{9.20}$$

yields

$$v'' - v + k^{-b} V(y/k)|v|^{\sigma-1} v = 0, \quad k > 0. \tag{9.21}$$

Then, by (V2),

$$\lim_{k\to 0} k^{-b} V(y/k) = |y|^{-b} |y/k|^b V(y/k) = |y|^{-b} \quad \forall y \neq 0,$$

which suggests considering the limit problem

$$v'' - v + |y|^{-b}|v|^{\sigma-1} v = 0. \tag{9.22}$$

It turns out [Gen10a] that (9.22) has a unique positive radial solution $v_0 \in H^1(\mathbb{R})$. This solution can be shown to have a variational characterization, from which it bears the name *ground state* of (9.22).

The advantage of the scaling is that, in the new variables (k,v), one can now obtain solutions by perturbation of (9.22), which is non-degenerate. More precisely, one can apply a version of the implicit function theorem to the function $F : \mathbb{R} \times H^1(\mathbb{R}) \to H^{-1}(\mathbb{R})$ defined by

$$F(k,v) = \begin{cases} v'' - v + |k|^{-b} V(y/|k|)|v|^{\sigma-1} v, & k \neq 0, \\ v'' - v + |y|^{-b}|v|^{\sigma-1} v, & k = 0, \end{cases}$$

at the point $(k,v) = (0,v_0) \in \mathbb{R} \times H^1(\mathbb{R})$, where $D_2 F(0,v_0) : H^1(\mathbb{R}) \to H^{-1}(\mathbb{R})$ is an isomorphism (see [Gen10a, Proposition 2.1]). This provides a small $k_0 > 0$ and a local C^1 curve of solutions $\{(k,v_k) : |k| < k_0\} \subset \mathbb{R} \times H^1(\mathbb{R})$ of $F(k,v) = 0$. The local bifurcation in Theorem 9.2 can then be obtained by going back to the original variables using (9.20), which yields a local C^1 curve of solutions

$$\{(\xi, w_\xi) : 0 < \xi < k_0^2\} \subset \mathbb{R} \times H^1(\mathbb{R})$$

of (9.2). The various solution norms in the two sets of variables are related by

$$\|w_\xi\|_{L^2}^2 = \xi^{\alpha-1} \|v_{\xi^{1/2}}\|_{L^2}^2, \quad \|\nabla w_\xi\|_{L^2}^2 = \xi^{\alpha} \|\nabla v_{\xi^{1/2}}\|_{L^2}^2,$$

$$\|w_\xi\|_{L^\infty} = \xi^{\frac{2-b}{2(\sigma-1)}} \|v_{\xi^{1/2}}\|_{L^\infty}, \quad \text{where } \alpha = \tfrac{4-2b+(\sigma-1)}{2(\sigma-1)}.$$

The behaviour of w_ξ as $\xi \to 0$ follows readily from these relations and the fact that $v_k \to v_0$ both in $H^1(\mathbb{R})$ and in $L^\infty(\mathbb{R})$ (see [Gen10a, Proposition 3.1]).

The slope condition. Let us now explain how the slope condition (9.14) can be derived from this analysis, for small $\xi > 0$. We show that $\frac{\mathrm{d}}{\mathrm{d}\xi}\|w_\xi\|_{L^2}^2 > 0$ for $\xi > 0$ small enough. Observe that

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\|w_\xi\|_{L^2}^2 = \frac{1}{2k}\frac{\mathrm{d}}{\mathrm{d}k}\|w_{k^2}\|_{L^2}^2 = \frac{1}{2k}\frac{\mathrm{d}}{\mathrm{d}k}\{k^\beta\|v_k\|_{L^2}^2\}$$

where

$$\beta = \frac{4-2b-(\sigma-1)}{\sigma-1} = 2(\alpha-1). \tag{9.23}$$

Now

$$\frac{\mathrm{d}}{\mathrm{d}k}\{k^\beta\|v_k\|_{L^2}^2\} = \beta k^{\beta-1}\|v_k\|_{L^2}^2 + k^\beta 2\big\langle v_k, \frac{\mathrm{d}}{\mathrm{d}k}v_k\big\rangle_{L^2}$$

$$= k^{\beta-1}\big\{\beta\|v_k\|_{L^2}^2 + 2k\big\langle v_k, \frac{\mathrm{d}}{\mathrm{d}k}v_k\big\rangle_{L^2}\big\}.$$

Since $\|v_k\|_{L^2}^2 \to \|v_0\|_{L^2}^2 > 0$ as $k \to 0$, we have that

$$\operatorname{sgn}\{\frac{\mathrm{d}}{\mathrm{d}\xi}\|w_\xi\|_{L^2}^2\} = \operatorname{sgn}\{\alpha-1\} \text{ for } \xi = k^2 \text{ small}, \tag{9.24}$$

provided

$$k\big\langle v_k, \frac{\mathrm{d}}{\mathrm{d}k}v_k\big\rangle_{L^2} \to 0 \text{ as } k \to 0. \tag{9.25}$$

On the other hand,

$$F(k,v_k) = 0 \Rightarrow D_kF(k,v_k) + D_vF(k,v_k)\frac{\mathrm{d}}{\mathrm{d}k}v_k = 0$$

$$\Rightarrow k\frac{\mathrm{d}}{\mathrm{d}k}v_k = -D_vF(k,v_k)^{-1}kD_kF(k,v_k) = -D_vF(k,v_k)^{-1}k^{-b}W(y/k)v_k^\sigma,$$

where $W(x) := x\cdot V'(x) + bV(x)$ appears in hypothesis (V2). Then, using (V2), it is not difficult to show that

$$k^{-b}W(y/k)v_k^\sigma \to 0 \text{ in } H^{-1} \text{ as } k \to 0.$$

Finally, it follows from the open mapping theorem that

$$D_vF(k,v_k)^{-1} \to D_vF(0,v_0)^{-1} \text{ in } B(H^{-1},H^1) \text{ as } k \to 0,$$

and we conclude that $k\frac{\mathrm{d}}{\mathrm{d}k}v_k \to 0$ in H^1 as $k \to 0$, from which (9.25) follows. Recalling our assumption that $1 < \sigma < 5-2b$, the slope condition (9.14) now readily follows from (9.23) and (9.24).

The spectral assumptions. Regarding the verification of (S1)–(S3), we shall not give as much detail as for the slope condition. That the solutions w_ξ indeed give rise to a Hessian $D^2_{w_\xi}\mathscr{L}_\xi : E \to E^*$ with the appropriate spectral structure also follows from the properties of the limit problem (9.22) through the perturbation procedure outlined above. The crucial point is the variational characterization of the ground state v_0, which can be shown to minimize the functional

$$\widetilde{\mathscr{L}_0}(v) = \frac{1}{2}\int_{\mathbb{R}}(v')^2 + v^2\,\mathrm{d}x - \frac{1}{\sigma+1}\int_{\mathbb{R}}|x|^{-b}|v|^{\sigma+1}\,\mathrm{d}x$$

on an appropriate codimension 1 submanifold N of $H^1(\mathbb{R})$. Note that the direct method of the calculus of variations cannot be applied to the functional $\widetilde{\mathscr{L}_0}$ since it is not coercive. In fact it turns out that v_0 is a saddle-point of $\widetilde{\mathscr{L}_0}$. More precisely, v_0 is a critical point of $\widetilde{\mathscr{L}_0}$

(i.e. $D_{v_0}\widetilde{\mathscr{L}_0} = 0$), and the quadratic form $D^2_{v_0}\widetilde{\mathscr{L}_0} : H^1 \times H^1 \to \mathbb{R}$ is positive definite tangentially to N, and negative along the ray spanned by v_0, transverse to N. This information – together with some Schrödinger operator theory – precisely implies that $D^2_{v_0}\widetilde{\mathscr{L}_0}$ enjoys the properties (S1)–(S3). Furthermore, if w_ξ and v_k are related by the change of variables (9.20), a straightforward calculation shows that

$$\mathscr{L}_\xi(w_\xi) = k^{\frac{3-2b+\sigma}{(\sigma-1)}}\,\widetilde{\mathscr{L}_0}(v_k),$$

where $\mathscr{L}_\xi$ is the Lyapunov function defined in (9.7). However, it is by no means trivial to verify that the spectral properties of $D^2_{v_0}\widetilde{\mathscr{L}_0}$ are carried through to $D^2_{w_\xi}\mathscr{L}_\xi$, for $\xi > 0$ small, in the perturbation procedure. This was shown in [GS08] in arbitrary dimension.

Note that, if the solutions w_ξ are themselves saddle-points of $\mathscr{L}_\xi$, the perturbation procedure can be dispensed of, and the spectral properties of the Hessian $D^2_{w_\xi}\mathscr{L}_\xi$ derived directly from this variational characterization. This is in fact the case for the solutions obtained in Theorem 9.2, but it is not known in the (AL) case, where the variational structure is much less transparent.

Remark 9.9. When verifying assumptions (S1)–(S3) in the context of (9.1)–(9.2) (which are set on the whole of $\mathbb{R}^d$) one has to deal with the continuous spectrum of $D^2_{w_\xi}\mathscr{L}_\xi$ in addition to the negative eigenvalue lying at the bottom of the spectrum. The standard approach to tackle this is *via* the theory of Schrödinger operators applied to the self-adjoint operator $\widetilde{R}^{-1}D^2_{w_\xi}\mathscr{L}_\xi : E \to E$. This motivates the definition of $\sigma(D^2_{w_\xi}\mathscr{L}_\xi)$ given in (9.16). On the other hand, the problem considered in Section 8 (set on a compact manifold) only gives rise to discrete spectrum in the linearization, and so can be handled with a more elementary spectral analysis, not requiring to introduce the Riesz isomorphism $\widetilde{R} : E \to E^*$ explicitly.

9.3.2. *Global continuation.* In this section we show how both the slope condition (9.14) and the spectral properties (S1)–(S3) extend from the previous local analysis to the global curve given by either Theorem 9.2 or Theorem 9.4. We will handle the two cases in a unified approach, using the general notation $f(x,w^2)w$ for the nonlinearity. As earlier, we will often merely write $\xi > 0$, really meaning $\xi \in (0,\infty)$ in the (PT) case and $\xi \in (0,\xi_\infty)$ in the (AL) case. Again, we only consider here the case $d = 1$.

The slope condition. From the previous analysis, (9.14) holds for $\xi > 0$ small enough. Hence we need only verify that

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\int_{\mathbb{R}} w_\xi^2\,\mathrm{d}x \neq 0 \quad \forall \xi > 0.$$

First notice that, since the solutions w_ξ are even,

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\int_{\mathbb{R}} w_\xi^2\,\mathrm{d}x = 2\int_{\mathbb{R}} w_\xi \frac{\mathrm{d}}{\mathrm{d}\xi}w_\xi\,\mathrm{d}x = 4\int_0^\infty w_\xi \chi_\xi,$$

where $\chi_\xi = \frac{\mathrm{d}w_\xi}{\mathrm{d}\xi}$ satisfies

$$\chi_\xi'' + \{f(x,w_\xi^2) + 2\partial_2 f(x,w_\xi^2)w_\xi^2\}\,\chi_\xi = \xi\chi_\xi + w_\xi.$$

To simplify the notation, we will drop the subscript ξ in the remainder of the argument. It can be shown [Gen10a, Gen13] that

$$\int_0^\infty \left\{2f(x,w^2) + x\partial_1 f(x,w^2) - \partial_2 f(x,w^2)w^2\right\} w\chi\,\mathrm{d}x = 2\xi\int_0^\infty w\chi\,\mathrm{d}x \tag{9.26}$$

and that there exists $x_0 > 0$ such that

$$\chi > 0 \text{ on } (0,x_0), \quad \chi(x_0) = 0, \quad \chi < 0 \text{ for } x > x_0.$$

Supposing by contradiction that $\int_0^\infty w\chi\,\mathrm{d}x = 0$, we can write (9.26) as

$$\int_0^\infty \Big\{ \frac{2f(x,w^2) + x\partial_1 f(x,w^2)}{\partial_2 f(x,w^2)w^2} - 1 \Big\} \partial_2 f(x,w^2) w^3 \chi\,\mathrm{d}x = 0.$$

Denoting by $\zeta(x)$ the function in the curly brackets, this becomes

$$\int_0^\infty \zeta(x)\partial_2 f(x,w^2) w^3 \chi\,\mathrm{d}x = 0.$$

Now using the unique zero x_0 of χ, we can rewrite this identity as

$$\int_0^\infty \{\zeta(x) - \zeta(x_0)\}\partial_2 f(x,w^2) w^3 \chi\,\mathrm{d}x + \zeta(x_0) \int_0^\infty \partial_2 f(x,w^2) w^3 \chi\,\mathrm{d}x = 0.$$

Moreover, multiplying the equation for w by χ, the equation for χ by w, subtracting and integrating, yields

$$\int_0^\infty w^2\,\mathrm{d}x = 2\int_0^\infty \partial_2 f(x,w^2) w^3 \chi\,\mathrm{d}x,$$

and so

$$\int_0^\infty \partial_2 f(x,w^2) w^3 \{\zeta(x) - \zeta(x_0)\}\chi\,\mathrm{d}x + \frac{\zeta(x_0)}{2}\int_0^\infty w^2\,\mathrm{d}x = 0. \tag{9.27}$$

Now,

$$\partial_2 f(x,w^2) w^3 = \begin{cases} \frac{\sigma-1}{2} V(x) w^\sigma & \text{in the (PT) case,} \\ \frac{\sigma-1}{2} V(x) \frac{w^\sigma}{(1+w^{\sigma-1})^2} & \text{in the (AL) case,} \end{cases}$$

hence $\partial_2 f(x,w^2)w^3 > 0$ on $(0,\infty)$ in any case. On the other hand,

$$\zeta(x) = \begin{cases} \frac{2}{\sigma-1}[x\frac{V'(x)}{V(x)} + \frac{5-\sigma}{2}] & \text{(PT)} \\ \frac{2}{\sigma-1}[x\frac{V'(x)}{V(x)} + \frac{5-\sigma}{2}] + \frac{2}{\sigma-1}[x\frac{V'(x)}{V(x)} + 2]w^{\sigma-1} & \text{(AL)} \end{cases}$$

and we claim that ζ is positive and decreasing in any case, which immediately leads to a contradiction with (9.27). To conclude, the claim follows from our hypotheses since

$$x \to x\frac{V'(x)}{V(x)} \text{ decreasing}, \quad x\frac{V'(x)}{V(x)} \geq -b \quad \text{and} \quad \sigma < 5 - 2b$$

$$\Rightarrow x\frac{V'(x)}{V(x)} + \frac{5-\sigma}{2} > 0 \text{ and decreasing}$$

(note that hypothesis (V4) is crucial here). Furthermore,

$$w > 0 \text{ and decreasing } \Rightarrow \underbrace{\Big[x\frac{V'(x)}{V(x)} + 2\Big]}_{\geq -b+2>0} w^{\sigma-1} > 0 \text{ and decreasing},$$

so that ζ is indeed positive and decreasing in any case.

The spectral conditions. The spectral conditions (S1)–(S3) can be reformulated in terms of the self-adjoint operators $L_\xi^+, L_\xi^- : H^2(\mathbb{R}) \subset L^2(\mathbb{R}) \to L^2(\mathbb{R})$ defined by

$$L_\xi^+ v = -v'' + \xi v - [f(x,w_\xi^2) + 2\partial_2 f(x,w_\xi^2) w_\xi^2] v,$$
$$L_\xi^- v = -v'' + \xi v - f(x,w_\xi^2) v.$$

Then (S1)–(S3) are equivalent to

(C1) $\inf \sigma_{\mathrm{ess}}(L_\xi^+) > 0, \quad M(L_\xi^+) = 1, \quad \ker L_\xi^+ = \{0\},$

(C2) $\inf \sigma_{\mathrm{ess}}(L_\xi^-) > 0, \quad 0 = \inf \sigma(L_\xi^-), \quad \ker L_\xi^- = \mathrm{vect}\{w_\xi\},$

where $\sigma_{\mathrm{ess}}(A)$ denotes the *essential spectrum* of a self-adjoint operator A, and $M(A)$ its *Morse index*, i.e. the dimension of the larger subspace where A is negative definite.

A first step toward verifying that (C1) and (C2) hold for all $\xi > 0$ is to show that all eigenvalues of L_ξ^+, L_ξ^- are simple, which follows by standard ODE arguments. Then, since

$$\lim_{|x|\to\infty} f(x, w_\xi(x)^2) = \lim_{|x|\to\infty} 2\partial_2 f(x, w_\xi(x)^2) w_\xi(x)^2 = 0,$$

it follows from the spectral theory of Schrödinger operators (see e.g. [Stu98]) that

$$\inf \sigma_{\mathrm{ess}}(L_\xi^+) = \inf \sigma_{\mathrm{ess}}(L_\xi^-) = \xi > 0.$$

Furthermore, applying ODE comparison arguments to the equations $L_\xi^+ v = 0$ and (9.2), it can be seen that $\ker L_\xi^+ = \{0\}$. On the other hand, since $w_\xi > 0$ is a solution of (9.2), it follows again from standard spectral theory that

$$\ker L_\xi^- = \mathrm{span}\{w_\xi\} \text{ and } 0 = \inf \sigma(L_\xi^-).$$

It remains to show that L_ξ^+ has exactly one negative eigenvalue. As discussed earlier, the local bifurcation analysis close to $\xi = 0$ shows that $M(L_\xi^+) = 1$ for $\xi > 0$ small enough. By perturbation theory, the eigenvalues of L_ξ^+ depend continuously on $\xi > 0$. Since $\ker L_\xi^+ = \{0\}$ for all $\xi > 0$, the eigenvalues cannot cross zero as ξ varies. Therefore, $M(L_\xi^+) = 1$ for all $\xi > 0$, which completes the proof of conditions (C1) and (C2).

Appendix

The goal of this appendix is to present those very basic notions from differential geometry, Lie group theory and Hamiltonian mechanics that are indispensable to follow the treatment of the main text and that are not necessarily familiar to all. The only prerequisites for this part are a good grasp of differential calculus on finite dimensional normed vector spaces not going much beyond a fluent mastery of the chain rule for differentiation and an intuitive grasp of what a submanifold of such spaces is.

Appendix A.1. Differential geometry: the basics

We first recall some elementary notions of differential geometry and dynamical systems on a normed vector space E. For the general theory on differentiable manifolds, one may for example consult [AM78, LM87, Spi79].

By a vector field on E we will mean a smooth map $X : E \to E$. Given $u \in E$, one should think of $X(u)$ as a "tangent vector to E at u". With this idea in mind, a vector field naturally determines a differential equation

$$\dot{u}(t) = X(u(t)), \quad u_0 = u,$$

the solutions of which induce a flow on E defined as $\Phi_t^X(u) = u(t)$. For ease of discussion, we will suppose throughout the appendix that all solutions are global and hence all flows complete. Most results carry over even if the flow exists only locally in time.

The diffeomorphisms[21] Φ of E act naturally on vector fields as follows. First note that, when Φ is a diffeomorphism, and $\gamma : t \in (a,b) \to E$ a curve with $\gamma(0) = u, \dot{\gamma}(0) = v$, then we can consider the curve $\tilde{\gamma} : t \in (a,b) \to E$ defined by $\tilde{\gamma}(t) = \Phi(\gamma(t))$. This is the curve γ, "pushed forward" by Φ: we invite the reader to draw a picture. This new curve satisfies $\tilde{\gamma}(0) = \Phi(u)$, so it passes through $\Phi(u)$. What is its tangent vector at that point? The chain rule yields immediately

$$\dot{\tilde{\gamma}}(0) = D_u\Phi(v),$$

where $D_y\Phi$ is our notation for the Fréchet derivative of Φ at $y \in E$, which is a continuous linear map from E to E. This equality gives a geometric interpretation to the purely analytical object $D_u\Phi(v)$: it is the tangent vector at $\Phi(u)$ to the curve $\tilde{\gamma}$ at $t = 0$. With this in mind, given a vector field X, we can now define a new vector field $\Phi_* X$, the *push forward* of the vector field X by the diffeomorphism Φ, as follows:

$$\Phi_* X(\Phi(u)) := D_u\Phi(X(u)).$$

Note that, with the above interpretation of the "push forward" of a vector at u, $D_u\Phi(X(u))$ is a vector "at $\Phi(u)$", which explains why $\Phi(u)$ appears in the argument in the left hand side. Of course, we can write

$$\Phi_* X(u) = D_{\Phi^{-1}(u)}\Phi(X(\Phi^{-1}(u))). \tag{A.1.1}$$

We will make little use of this notation from differential geometry, preferring to write out the explicit expression $D_u\Phi(X(u))$ whenever needed.

Diffeomorphisms also act naturally on flows, as follows. Given a diffeomorphism $\Phi : E \to E$, one has, for all $u \in E$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\Phi \circ \Phi_t^X)(u) = D_{\Phi_t(u)}\Phi(X(\Phi_t(u))).$$

[21] We mean $\Phi \in C^1(E,E)$ with a $C^1(E,E)$ inverse.

From this and (A.1.1), one concludes

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\Phi\circ\Phi_t^X\circ\Phi^{-1})(u) &= D_{\Phi_t^X(\Phi^{-1}(u))}\Phi(X(\Phi_t^X(\Phi^{-1}(u))))\\ &= \Phi_* X(\Phi\circ\Phi_t^X\circ\Phi^{-1}(u)).\end{aligned}$$

In other words, the flow $\Phi\circ\Phi_t^X\circ\Phi^{-1}$ is generated by the pushed forward vector field $\Phi_* X$.

It follows from the above and an application of the chain rule that, if X,Y are two vector fields on E, then, for all $u\in E$,

$$\begin{aligned}\frac{\partial^2}{\partial s\partial t}\Phi_s^Y\circ\Phi_t^X\circ\Phi_{-s}^Y(u)_{|s=0=t} &= \frac{\mathrm{d}}{\mathrm{d}s}D_{\Phi_{-s}^Y(u)}\Phi_s^Y(X(\Phi_{-s}^Y(u)))_{s=0}\\ &= \frac{\mathrm{d}}{\mathrm{d}s}X(\Phi_{-s}^Y(u))_{s=0}+\frac{\mathrm{d}}{\mathrm{d}s}D_x\Phi_s^Y(X(u))_{s=0}\\ &= [X,Y](u), \qquad (A.1.2)\end{aligned}$$

where the commutator $[X,Y]$ of two vector fields is defined as follows:

$$[X,Y](u)=D_uY(X(u))-D_uX(Y(u)).$$

This definition is justified by the following observation. Given a vector field X and a C^1 function $F:E\to\mathbb{R}$, one can define a differential operator

$$\widehat{X}(F)(u)=D_uF(X(u)), \qquad (A.1.3)$$

which is – geometrically – nothing but the directional derivative of F at u in the direction $X(u)$. A simple computation shows readily that

$$[\widehat{X},\widehat{Y}]=\widehat{[X,Y]}. \qquad (A.1.4)$$

The following is then well known:

Lemma A.1.1. *The following are equivalent:*

(i) For all $s,t\in\mathbb{R}$, $\Phi_t^X\circ\Phi_s^Y=\Phi_s^Y\circ\Phi_t^X$;
(ii) $[X,Y]=0$.

Proof. That (i) implies (ii) follows immediately from the preceding computation. The proof of the converse is slightly more involved, for a simple argument we refer to [Spi79]. □

Remark A.1.2. Note that, if $X(u)=Au, Y(u)=Bu$, where $A,B:E\to E$ are linear, then, with our convention, $[X,Y](u)=-[A,B]u$. Here $[A,B]=AB-BA$ is the standard commutator of linear maps.

Definition A.1.3. Let $F\in C^k(E,\mathbb{R}^m)$ for some $k\geq 1$. For each $\mu\in\mathbb{R}^m$ we define a level set of F by

$$\Sigma_\mu=\{u\in E\mid F(u)=\mu\}. \qquad (A.1.5)$$

We will say $u\in E$ is a regular point of F if $D_uF:E\to\mathbb{R}^m$ is surjective. We will say μ is a regular value of F, if $\Sigma_\mu\neq\emptyset$ and all $u\in\Sigma_\mu$ are regular points of F.

If μ is a regular value of F, then Σ_μ is a co-dimension m submanifold of E [BER99, Theorem 6.3.34]. In that case, the tangent space to Σ_μ at u is defined as follows:

$$T_u\Sigma_\mu=\{w\in E\mid D_uF(w)=0\}=\mathrm{Ker}(D_uF). \qquad (A.1.6)$$

We point out that if $r=\mathrm{Rank}(D_uF)$ is constant on Σ_μ, then Σ_μ is a co-dimension r submanifold. We will need the following simple result in Section 7.4.

Lemma A.1.4. *Let $F \in C^k(E,\mathbb{R}^m)$ for some $k \geq 2$. Let $\mu \in \mathbb{R}^m$ be a regular value of F. Let $u \in \Sigma_\mu$ and let W_u be a subspace of E so that $E = T_u\Sigma_\mu \oplus W_u$. Then, for all $v \in \Sigma_\mu$,*

$$\|(v-u)_2\| \leq \mathrm{O}(\|v-u\|^2),$$

and there exist $\delta, C > 0$ such that

$$\|v-u\| \leq \delta \Rightarrow \|(v-u)_1\| \geq C\|v-u\|,$$

where $(v-u) = (v-u)_1 + (v-u)_2 \in T_u\Sigma_\mu \oplus W_u$.

Note that both δ and C depend on u and on the decomposition of E chosen.

Proof. Write $u - v = w_1 + w_2$, with $w_1 \in T_u\Sigma_\mu$ and $w_2 \in W_u$. Then, using that $D_uF(w_1) = 0$, we have

$$0 = F(v) - F(u) = D_uF(w_2) + \mathrm{O}(\|v-u\|^2).$$

Now, since D_uF is a diffeomorphism from W_u to $\mathbb{R}^m$, there exists $c > 0$ so that

$$\|D_uF(w_2)\| \geq c\|w_2\|, \quad \text{hence} \quad \mathrm{O}(\|v-u\|^2) \geq c\|w_2\|.$$

Finally

$$\|w_1\| = \|u - v - w_2\| \geq \|u-v\| - \|w_2\| \geq \|u-v\| - \mathrm{O}(\|v-u\|^2),$$

from which the result follows. □

Appendix A.2. Lie algebras, Lie groups and their actions

In general, a Lie algebra is a vector space V equipped with a bilinear composition law $(u,v) \in V \times V \to [u,v] \in V$, called a Lie bracket, which is anti-symmetric and satisfies the Jacobi identity, meaning that for all $u, v, w \in V$:

$$[[u,v],w] + [[v,w],u] + [[w,u],v] = 0. \tag{A.2.1}$$

The basic example of this structure is given by spaces of matrices or, more generally, of linear operators on vector spaces, where the Lie bracket is given by the usual commutator. Two other examples play an important role in these notes, namely the space of vector fields on a normed vector space with the commutator defined in (A.1.2) and the space of all smooth functions on a symplectic vector space, where the Lie bracket is given by the Poisson bracket, as explained in Section A.3 below. The validity of the Jacobi identity follows in all these examples from a direct computation, whereas the bilinearity and the anti-symmetry are obvious. Lie algebras are intimately linked to Lie groups, as the terminology strongly suggests, and as we now further explain.

In general, a Lie group is a group equipped with a compatible manifold structure. For our purposes, it is however enough to define a Lie group G to be a subgroup of $\mathrm{GL}(\mathbb{R}^N)$, such that G is also a submanifold of $\mathbb{R}^{N^2}$ (*i.e.* for our purposes, typically the level surface of a vector-valued function). As such, $\mathrm{GL}(\mathbb{R}^N)$ itself, which is an open subset of $\mathbb{R}^{N^2}$, is a Lie group. So are the rotation group

$$\mathrm{SO}(N) = \{R \in \mathrm{GL}(N,\mathbb{R}) \mid R^TR = \mathrm{I}_N\}$$

and the symplectic group

$$\mathrm{Sp}(2N) = \{S \in \mathrm{GL}(2N,\mathbb{R}) \mid S^TJS = J\}, \quad \text{with} \quad J = \begin{pmatrix} 0 & \mathrm{I}_N \\ -\mathrm{I}_N & 0 \end{pmatrix}. \tag{A.2.2}$$

A simple verification shows that $\mathrm{Sp}(2) = \mathrm{SL}(2,\mathbb{R})$, the space of two by two matrices of determinant one. The dimension of a Lie group is by definition its dimension as a manifold. For $\mathrm{SO}(N)$, it is $N(N-1)/2$, and for $\mathrm{Sp}(2N)$, it is $N(2N+1)$, as is readily checked. The

group $\mathbb{R}^n$ is also a Lie group in this sense. Indeed, putting $N = n+1$, and defining, for each $a \in \mathbb{R}^n$,

$$A(a) = \begin{pmatrix} \mathrm{I}_n & a \\ 0 & 1 \end{pmatrix}$$

one readily sees that $A(a)A(b) = A(a+b)$, so that one can view $\mathbb{R}^n$ as a subgroup of $\mathrm{GL}(n+1,\mathbb{R})$.

We recall that, in general, an action of a group G on a set Σ is a map $\Phi : (g,x) \in G \times \Sigma \to \Phi_g(x) \in \Sigma$ which satisfies $\Phi_e(x) = x$, for all $x \in \Sigma$, and $\Phi_{g_1} \circ \Phi_{g_2} = \Phi_{g_1 g_2}$. In these notes, we consider actions that are defined on a normed vector space E. If the Φ_g are linear, one says Φ is a representation of the group. This will *not* always be the case in these notes: actions may be nonlinear. Furthermore, all actions considered will be at least continuous, and very often they will have additional smoothness properties. In this appendix, where we deal with finite dimensional systems only, the actions are supposed to be separately C^1 in each of their two variables $g \in G$ and $u \in E$. Appropriate technical conditions to deal with infinite dimensional spaces E are given in the main part of the text as needed.

By definition, *the* Lie algebra $\mathfrak{g}$ of a Lie group G is the tangent space to the manifold G at the unit element $e \in G$:

$$\mathfrak{g} = T_e G.$$

In other words, for each $\xi \in \mathfrak{g}$, there exists $\gamma : t \in \mathbb{R} \to G$, a smooth curve with $\gamma(0) = e = \mathrm{I}_N$, and $\dot\gamma(0) = \xi$. Note that one should think of ξ as a matrix, since for each t, $\gamma(t)$ is one. In addition, it turns out that, given $\xi \in \mathfrak{g}$,

$$\exp(t\xi) \in G,$$

for all $t \in \mathbb{R}$ where $\exp(t\xi)$ is to be understood as the exponential of the matrix $t\xi$. Indeed, given ξ and γ as above, for all $n \in \mathbb{N}$, $\gamma(\frac{t}{n}) \in G$ and so $\gamma(\frac{t}{n})^n \in G$. Taking $n \to +\infty$, the result follows. A one-parameter subgroup of G is, by definition, a smooth curve $\gamma : t \in \mathbb{R} \to \gamma(t) \in G$, which is also a group diffeomorphism: $\gamma(t+s) = \gamma(t)\gamma(s)$. What precedes shows that any such one-parameter group is of the form $t \to \exp(t\xi)$. So there is a one-to-one correspondence between the one-parameter subgroups of G and its Lie-algebra, which starts to explain the importance of this latter notion. In addition, it turns out that, if $\xi, \eta \in T_e G$, then so is their commutator (seen as matrices)

$$[\xi,\eta] = \xi\eta - \eta\xi,$$

which justifies calling $T_e G$ a Lie algebra. Indeed, consider, for each $s \in \mathbb{R}$, the curve

$$\gamma : t \in \mathbb{R} \to \exp(s\eta)\exp(t\xi)\exp(-s\eta) \in G.$$

Clearly $\gamma(0) = \mathrm{I}_N$ and $\dot\gamma(0) = \exp(s\eta)\xi\exp(-s\eta) \in T_e G$. So we have a curve

$$s \in \mathbb{R} \to \exp(s\eta)\xi\exp(-s\eta) \in T_e G.$$

Taking the derivative with respect to s yields $[\eta,\xi] \in T_e G$:

$$\frac{\mathrm{d}}{\mathrm{d}s}\exp(s\eta)\xi\exp(-s\eta)_{|s=0} = [\eta,\xi]. \tag{A.2.3}$$

As an example, the Lie algebra of $\mathrm{SO}(N)$, denoted by $\mathrm{so}(N)$, is given by

$$\mathrm{so}(N) = \{A \in \mathscr{M}(N,\mathbb{R}) \mid A^T + A = 0\},$$

which is the space of all anti-symmetric $N \times N$ matrices. This is easily established by writing $\exp(tA^T)\exp(tA) = \mathrm{I}_N$ and taking a t-derivative at $t = 0$. And it is obvious that the

commutator of two anti-symmetric matrices is anti-symmetric. A basis for $\mathrm{so}(3)$ is

$$e_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},\ e_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix},\ e_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{A.2.4}$$

and one readily checks that

$$[e_1, e_2] = e_3, \quad [e_2, e_3] = e_1, \quad [e_3, e_1] = e_2. \tag{A.2.5}$$

One then identifies $\xi \in \mathrm{so}(3)$ with $\xi \in \mathbb{R}^3$ via

$$\xi = \sum_{i=1}^{3} \xi_i e_i = \begin{pmatrix} 0 & -\xi_3 & \xi_2 \\ \xi_3 & 0 & -\xi_1 \\ -\xi_2 & \xi_1 & 0 \end{pmatrix}. \tag{A.2.6}$$

Similarly, a basis for $\mathrm{sl}(2,\mathbb{R})$, the Lie algebra of $\mathrm{SL}(2,\mathbb{R})$, is

$$e_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ e_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ e_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \tag{A.2.7}$$

and one has

$$[e_0, e_+] = 2e_+, \quad [e_0, e_-] = -2e_-, \quad [e_-, e_+] = -e_0. \tag{A.2.8}$$

In general, if e_i, $i = 1, \ldots, m$ is a basis of $\mathfrak{g}$, there exists constants c_{ij}^k so that

$$[e_i, e_j] = c_{ij}^k e_k, \tag{A.2.9}$$

where the summation over k is understood; the c_{ij}^k are called the structure constants of $\mathfrak{g}$.

There exists a natural *linear* action of G on its Lie algebra, called the *adjoint action* or *adjoint representation*, defined as follows, for all $g \in G, \xi \in T_eG$:

$$\mathrm{Ad}_g \xi = g \xi g^{-1}.$$

Clearly $\mathrm{Ad}_{g_1 g_2} = \mathrm{Ad}_{g_1} \mathrm{Ad}_{g_2}$. Note that for a commutative Lie group G, such as $\mathbb{R}^n$, it is trivial: $\mathrm{Ad}_g \xi = \xi$. It is instructive to compute some non-trivial adjoint actions explicitly. For $\mathrm{SO}(3)$, one finds, with the above (somewhat abusive) notation

$$\mathrm{Ad}_R \xi = \begin{pmatrix} 0 & -(R\xi)_3 & (R\xi)_2 \\ (R\xi)_3 & 0 & -(R\xi)_1 \\ -(R\xi)_2 & (R\xi)_1 & 0 \end{pmatrix} = R\xi. \tag{A.2.10}$$

We invite the reader to do the analogous computation for $\mathrm{sl}(2,\mathbb{R})$, determining the matrix of Ad_g in the basis given above.

The dual of the Lie algebra $\mathfrak{g}$ (as a vector space) is denoted by $\mathfrak{g}^*$. It appears very naturally in the study of symplectic group actions arising in the study of Hamiltonian systems with symmetry, as we will see in Section A.3.2. Given a basis e_i of $\mathfrak{g}$, we denote by e_i^* the dual basis defined by $e_i^*(e_j) = \delta_{ij}$.

Moreover, there is a natural action of G on $\mathfrak{g}^*$, obtained by dualization as follows. For all $\mu \in \mathfrak{g}^*$, for all $\xi \in \mathfrak{g}$, we define

$$\mathrm{Ad}_g^* \mu(\xi) = \mu(\mathrm{Ad}_{g^{-1}} \xi). \tag{A.2.11}$$

This is called the co-adjoint action of G. For later purposes, we define, for all $\mu \in \mathfrak{g}^*$,

$$G_\mu = \{ g \in G \mid \mathrm{Ad}_g^* \mu = \mu \}, \tag{A.2.12}$$

the so-called *stabilizer* or *isotropy group* of $\mu \in \mathfrak{g}^*$.

As above, given a basis e_i of $\mathfrak{g}$, one identifies $\mu \in \mathfrak{g}^*$ with $\mu = (\mu_1, \dots, \mu_m) \in \mathbb{R}^m$ by writing

$$\mu = \sum_{i=1}^{m} \mu_i e_i^* \text{ so that } \mu(\xi) = \sum_{i=1}^{m} \mu_i \xi_i. \tag{A.2.13}$$

Let $\mu \in \mathrm{so}(3)^*$; we write $\mu(\xi) = \sum_{i=1}^{3} \mu_i \xi_i$ and identify $\mu \in \mathrm{so}(3)^*$ with $\mu = (\mu_1, \mu_2, \mu_3) \in \mathbb{R}^3$. Again, one readily checks that

$$\mathrm{Ad_R}^* \mu = R\mu. \tag{A.2.14}$$

Remark A.2.1. It is often useful to suppose there exists an Euclidian structure on $\mathfrak{g}$ that is preserved by Ad_g for all $g \in G$. This is equivalent to supposing that there exists a basis e_i of $\mathfrak{g}$ so that the matrix of Ad_g in e_i belongs to $\mathrm{O}(m)$. We will simply write $\mathrm{Ad}_g \in \mathrm{O}(m)$ in this case. It follows that the matrix of Ad_g^* in the dual basis e_i^* belongs to $\mathrm{O}(m)$ as well. This implies that the natural Euclidian structure induced on $\mathfrak{g}^*$ by the one on $\mathfrak{g}$ is preserved by Ad_g^* for all $g \in G$. Such a structure always exists if the group G is compact.

Suppose now we have a C^1-action $\Phi : (g,u) \in G \times E \to \Phi_g(u) \in E$ of a Lie group G on a normed vector space E. Then, for all $\xi \in T_eG$, one can define the vector field X_ξ on E via

$$X_\xi(u) = \frac{\mathrm{d}}{\mathrm{d}t} \Phi_{\exp(t\xi)}(u)_{|t=0}. \tag{A.2.15}$$

Lemma A.2.2. *If Φ is a C^2-action, then for all $g \in G$, $\xi, \eta \in \mathfrak{g}$, for all $u \in E$, one has*

$$[X_\xi, X_\eta] = -X_{[\xi,\eta]}, \tag{A.2.16}$$

$$X_{\mathrm{Ad}_g \xi}(\Phi_g(u)) = D_u \Phi_g(X_\xi(u)). \tag{A.2.17}$$

Proof. It follows from (A.1.2) that

$$\frac{\partial^2}{\partial s \partial t} \Phi_{\exp(s\eta)\exp(t\xi)\exp(-s\eta)_{|s=0=t}} = \left[X_\xi, X_\eta\right].$$

Now, by definition,

$$X_{\exp(s\eta)\xi\exp(-s\eta)} = \frac{\mathrm{d}}{\mathrm{d}t} \Phi_{\exp(s\eta)\exp(t\xi)\exp(-s\eta)_{|t=0}}$$

and furthermore

$$\frac{\mathrm{d}}{\mathrm{d}s} X_{\exp(s\eta)\xi\exp(-s\eta)_{|s=0}} = X_{[\eta,\xi]}.$$

This proves (A.2.16). For (A.2.17), note that the chain rule implies

$$\frac{\mathrm{d}}{\mathrm{d}t} \Phi_g(\Phi_{\exp(t\xi)}(u))_{|t=0} = D_u \Phi_g(X_\xi(u)).$$

On the other hand, $\Phi_{g\exp(t\xi)}(u) = \Phi_{g\exp(t\xi)\ g^{-1}}(\Phi_g(u))$. Hence

$$\frac{\mathrm{d}}{\mathrm{d}t} \Phi_g(\Phi_{\exp(t\xi)}(u))_{|t=0} = X_{\mathrm{Ad}_g \xi}.$$

□

Lemma A.2.2 shows that the map $\xi \in \mathfrak{g} \to X_\xi$ is a Lie algebra anti-homomorphism.

APPENDIX A.3. HAMILTONIAN DYNAMICAL SYSTEM WITH SYMMETRY IN FINITE DIMENSION

We now turn to a very short description of Hamiltonian dynamical systems and their symmetries on a finite dimensional normed vector space E. We present the theory in a simple but slightly abstract formalism that is well-suited for the generalization to the infinite dimensional situation needed for the main body of the text and presented in Section 5. The modern theory of finite dimensional Hamiltonian dynamical systems finds its natural setting in the theory of (finite dimensional) symplectic geometry [AM78, Arn99, LM87, Sou97]. We shall however have no need for this more general formulation in these notes.

A.3.1. **Hamiltonian dynamical systems**

The central object of the theory in its usual formulation is a symplectic form, that we now define. Let $\omega : E \times E \to \mathbb{R}$ be a bilinear form which is anti-symmetric, meaning

$$\forall u, u' \in E,\ \omega(u,u') = -\omega(u',u),$$

and non-degenerate, meaning that, for all $u \in E$,

$$\big(\forall u' \in E,\ \omega(u,u') = 0\big) \Rightarrow u = 0.$$

Such a form is called a symplectic form. The standard example is $E = \mathbb{R}^n \times \mathbb{R}^n$ with $u = (q,p)$ and

$$\omega(u,u') = q \cdot p' - q' \cdot p, \tag{A.3.1}$$

where $\cdot$ indicates the standard inner product on $\mathbb{R}^n$. Given a C^1-function $F : E \to \mathbb{R}$, one defines the *Hamiltonian vector field X_F associated to F* as follows: for all $u \in E$,

$$\omega(X_F(u), u') = D_u F(u'), \quad \forall u' \in E. \tag{A.3.2}$$

We recall that $D_u F \in E^*$ is our notation for the Frechet derivative of F at u. Observe that one can think of the map $u \in E \to D_u F \in E^*$ as a differential one-form on E. The vector field X_F is well-defined and unique, thanks to the non-degeneracy of the symplectic form. If ω were symmetric, rather than anti-symmetric, it would define an inner product on E, rather than a symplectic form, and (A.3.2) would actually define the gradient of F; in analogy, one sometimes refers to X_F as the symplectic gradient of F. We will see it has radically different features from the gradient.

For later reference, we point out that

$$X_F = 0 \Rightarrow \exists c \in \mathbb{R},\ \forall u \in E,\ F(u) = c. \tag{A.3.3}$$

The flow of the Hamiltonian vector field X_F, for which we shall write Φ_t^F, is obtained by integrating the differential equation

$$\dot{u}(t) = X_F(u(t)), \quad u_0 = u, \tag{A.3.4}$$

referred to as the Hamiltonian equation of motion. One writes $\Phi_t^F(u) = u(t)$. In this section we suppose that (A.3.4) admits a unique and global solution and that, for all $t \in \mathbb{R}$, $\Phi_t \in C(E,E)$.

As a typical example from elementary mechanics, let $V \in C^1(\mathbb{R}^3;\mathbb{R})$ and define the function

$$H(q,p) = \tfrac{1}{2}p^2 + V(q) \tag{A.3.5}$$

on $E = \mathbb{R}^6$, with the symplectic form as above. The equations of motion corresponding to H are then

$$\dot{q}(t) = p(t), \quad \dot{p}(t) = -\nabla V(q(t)). \tag{A.3.6}$$

Note that they lead to Newton's force law in the form $\ddot{q}(t) = -\nabla V(q(t))$. More generally, in the example above, with $E = \mathbb{R}^{2n}$, one finds

$$X_F(q,p) = \begin{pmatrix} \partial_p F(q,p) \\ -\partial_q F(q,p) \end{pmatrix},$$

which leads to the familiar Hamiltonian equations of motion:

$$\dot{q}(t) = \partial_p F(q(t),p(t)), \qquad \dot{p}(t) = -\partial_q F(q(t),p(t)).$$

We give several other explicit examples of such flows in the main part of these notes.

Let us return to the general situation. Given two functions $F_1, F_2 : E \to \mathbb{R}$, one defines their Poisson bracket $\{F_1, F_2\}$ via

$$\{F_1,F_2\} = \omega(X_{F_1}, X_{F_2}) = -\{F_2,F_1\}. \tag{A.3.7}$$

Observe that, with the notation from (A.1.3), we have

$$\widehat{X}_{F_1}(F_2) = DF_2(X_{F_1}) = \omega(X_{F_2}, X_{F_1}) = \{F_2,F_1\}, \tag{A.3.8}$$

i.e. for all $u \in E$,

$$\widehat{X}_{F_1}(F_2)(u) = D_u F_2(X_{F_1}(u)) = \omega(X_{F_2}(u), X_{F_1}(u)) = \{F_2,F_1\}(u).$$

It is then immediate from what precedes that, for all $u \in E$,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(F_2 \circ \Phi_t^{F_1})(u) &= D_{\Phi_t^{F_1}(u)} F_2(X_{F_1}(\Phi_t^{F_1}(u))) \\ &= \{F_2,F_1\}(\Phi_t^{F_1}(u)) \end{aligned}$$

which in turn yields:

Theorem A.3.1. *Let $F_1, F_2 \in C^1(E,\mathbb{R})$. Then $F_1 \circ \Phi_t^{F_2} = F_1$ for all t iff $F_2 \circ \Phi_t^{F_1} = F_2$ for all t, iff $\{F_1,F_2\} = 0$.*

When $F_1 \circ \Phi_t^{F_2} = F_1$ for all t, one says either that the $\Phi_t^{F_2}$ form a symmetry group[22] for F_1 or that F_1 is a constant of the motion[23] for the flow $\Phi_t^{F_2}$. The theorem, which is a Hamiltonian version of Noether's theorem (See [AM78, Arn99, LM87, Sou97] for a general treatment), can therefore be paraphrased by saying that F_2 is a constant of the motion for the flow $\Phi_t^{F_1}$ iff the flow $\Phi_t^{F_2}$ of F_2 forms a group of symmetries for F_1. Several instances and applications of this result appear in the main body of the text. It is typically used in the following manner. One wishes to study the dynamical flow $\Phi_t^{F_1}$. One has a simple and well-known one parameter group $\Phi_t^{F_2}$ for which one readily establishes with an explicit computation that $F_1 \circ \Phi_t^{F_2} = F_1$. From this, one can then conclude that F_2 is a constant of the motion for the dynamical group $\Phi_t^{F_1}$. We will elaborate on this point in Section A.3.2.

The radical difference between the properties of the symplectic gradient and the "usual" gradient is now apparent. The anti-symmetry of the Poisson bracket implies $\widehat{X}_F(F) = 0$, that is, the symplectic gradient is *tangent* to the level surfaces of F (See (A.1.6)), rather than orthogonal. Hence its flow Φ_t^F preserves these surfaces rather than moving points to increasing values of F as does the usual gradient. These features, together with the Jacobi identity, are at the origin of all special properties of Hamiltonian systems.

[22] See Definition 2.1.

[23] Defined in (2.4).

To prepare for the treatment of Hamiltonian dynamical systems in infinite dimension (see Section 5), we reformulate the above as follows. Given a symplectic form ω on a finite dimensional normed vector space E, one can define a bijective linear map

$$\mathscr{J} : u \in E \to \mathscr{J} u \in E^*$$

by $\mathscr{J} u(v) = \omega(u,v)$. It is clear that

$$\mathscr{J} u(v) = -\mathscr{J} v(u). \tag{A.3.9}$$

With this notation, we find that

$$X_F = \mathscr{J}^{-1} DF, \quad \text{or} \quad \mathscr{J} X_F = DF \tag{A.3.10}$$

so that the Hamiltonian equations of motion (A.3.4) can be equivalently rewritten as

$$\mathscr{J} \dot{u}(t) = D_{u(t)} F. \tag{A.3.11}$$

This formulation is the one that we carry over to the infinite dimensional setting in the main body of these notes. Note that the Poisson bracket of two functions can now be written as

$$\{F,G\} = DF(\mathscr{J}^{-1} DG). \tag{A.3.12}$$

The point to make is that all objects of the theory can be expressed in terms of $\mathscr{J}$. This is illustrated in the proof of the following result.

Lemma A.3.2. *If $F_1, F_2, F_3 \in C^2(E,\mathbb{R})$, then the Jacobi identity holds:*

$$\{\{F_1,F_2\},F_3\} + \{\{F_2,F_3\},F_1\} + \{\{F_3,F_1\},F_2\} = 0 \tag{A.3.13}$$

If $F_1, F_2 \in C^2(E,\mathbb{R})$, then

$$X_{\{F_1,F_2\}} = -[X_{F_1}, X_{F_2}]. \tag{A.3.14}$$

Proof. To prove (A.3.13), one first easily checks that

$$\begin{aligned} &\{\{F_1,F_2\},F_3\}(u) = \\ = \;& D_u^2 F_1(\mathscr{J}^{-1} D_u F_2, \mathscr{J}^{-1} D_u F_3) + D_u F_1(\mathscr{J}^{-1} D_u^2 F_2(\cdot, \mathscr{J}^{-1} D_u F_3)) \\ = \;& D_u^2 F_1(\mathscr{J}^{-1} D_u F_2, \mathscr{J}^{-1} D_u F_3) - D_u^2 F_2(\mathscr{J}^{-1} D_u F_1, \mathscr{J}^{-1} D_u F_3), \end{aligned}$$

where we used (A.3.9). The result is then immediate. To prove (A.3.14) we use (A.1.3)–(A.1.4) to write

$$\begin{aligned} \widehat{[X_{F_1}, X_{F_2}]}(F_3) &= \widehat{X}_{F_1}(\widehat{X}_{F_2}(F_3)) - \widehat{X}_{F_2}(\widehat{X}_{F_1}(F_3)) \\ &= \widehat{X}_{F_1}(\{F_3,F_2\}) - \widehat{X}_{F_2}(\{F_3,F_1\})) \\ &= \{\{F_3,F_2\},F_1\} - \{\{F_3,F_1\},F_2\} \\ &= \{\{F_1,F_2\},F_3\} = -\widehat{X}_{\{F_1,F_2\}}(F_3), \end{aligned}$$

where we used the Jacobi identity in the last line. □

For the case where $E = \mathbb{R}^{2n}$ with the standard symplectic structure, one readily finds

$$\{F_1,F_2\} = \partial_q F_1 \cdot \partial_p F_2 - \partial_p F_1 \cdot \partial_q F_2. \tag{A.3.15}$$

The above lemma then follows from a direct computation.

The lemma implies that the vector space $C^\infty(E,\mathbb{R})$, equipped with the Poisson bracket, is a Lie algebra. In addition, it follows that the constants of the motion of a given function $F \in C^\infty(E,\mathbb{R})$ form a Lie subalgebra. Indeed, introducing the space of constants of the motion of F,

$$\mathscr{C}_F = \{G \in C^\infty(E,\mathbb{R}) \mid G \circ \Phi_t^F = G, \forall t \in \mathbb{R}\}, \tag{A.3.16}$$

which is clearly a vector space, it follows immediately from (A.3.13) that

$$G_1, G_2 \in \mathscr{C}_F \Rightarrow \{G_1, G_2\} \in \mathscr{C}_F,$$

so that $\mathscr{C}_F$ is a Lie subalgebra of $C^\infty(E,E)$.

We finally need to introduce symplectic transformations.

Definition A.3.3. A symplectic transformation on a symplectic space (E,ω) is a C^1 diffeomorphism $\Phi : E \to E$ so that, for all $u,v,w \in E$

$$\omega(D_u\Phi(v), D_u\Phi(w)) = \omega(v,w). \tag{A.3.17}$$

This is often paraphrased by the statement that "Φ preserves the symplectic structure." To understand what this means, one should recall the interpretation of $D_u\Phi(v)$ as the "push forward" of v by Φ, explained in Section A.1. Equation (A.3.17) states that a diffeomorphism is symplectic if the symplectic form is left invariant by the "push forward" operation of its arguments. Note that, if Φ is linear, (A.3.17) reduces to $\omega(\Phi(v),\Phi(w)) = \omega(v,w)$. And if $E = \mathbb{R}^{2n}$ with its standard symplectic structure, this then means that $\Phi \in \mathrm{Sp}(2n)$, defined in (A.2.2).

Lemma A.3.4. *Let $F \in C^1(E,\mathbb{R})$ and let $\Phi \in C^1(E,E)$ be a symplectic transformation. Then, for all $u \in E$,*

$$D_u\Phi(X_{F\circ\Phi}(u)) = X_F(\Phi(u)). \tag{A.3.18}$$

Moreover, for all $t \in \mathbb{R}$,

$$\Phi \circ \Phi_t^{F\circ\Phi} \circ \Phi^{-1} = \Phi_t^F. \tag{A.3.19}$$

In particular, if $F \circ \Phi = F$, then Φ commutes with Φ_t^F, for all $t \in \mathbb{R}$. And if Φ commutes with Φ_t^F, for all $t \in \mathbb{R}$, then there exists $c \in \mathbb{R}$ so that $F \circ \Phi = F + c$.

Equation (A.3.18) asserts that the push forward of the vector field $X_{F\circ\Phi}$ by Φ is X_F.

Proof. For all $u,v \in E$, one has

$$\begin{aligned} \omega(X_{F\circ\Phi}(u), v) &= D_u(F\circ\Phi)(v) = D_{\Phi(u)}F(D_u\Phi(v)) \\ &= \omega(X_F(\Phi(u)), D_u\Phi(v)). \end{aligned}$$

Hence, since Φ is symplectic and since $D_{\Phi(u)}\Phi^{-1} D_u\Phi = \mathrm{Id}_E = D_u\Phi D_{\Phi(u)}\Phi^{-1}$,

$$\omega(D_u\Phi(X_{F\circ\Phi}(u)), v) = \omega(X_{F\circ\Phi}(u), D_{\Phi(u)}\Phi^{-1}(v)) = \omega(X_F(\Phi(u)), v)$$

which yields (A.3.18). Next, for all $u \in E$, one finds from the chain rule and (A.3.18)

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\Phi(\Phi_t^{F\circ\Phi}(\Phi^{-1}(u))) &= D_{\Phi_t^{F\circ\Phi}(\Phi^{-1}(u))}\Phi\left(X_{F\circ\Phi}(\Phi_t^{F\circ\Phi}(\Phi^{-1}(u)))\right) \\ &= X_F((\Phi\circ\Phi_t^{F\circ\Phi}\circ\Phi^{-1})(u)). \end{aligned}$$

This shows $t \in \mathbb{R} \to (\Phi\circ\Phi_t^{F\circ\Phi}\circ\Phi^{-1})(u) \in E$ is a flow line of X_F. Since the latter are unique, (A.3.19) follows. □

We end with a proof of a basic fact about Hamiltonian flows: if they are smooth, they are symplectic.

Theorem A.3.5. *Let $F \in C^2(E,\mathbb{R})$. Suppose that the corresponding Hamiltonian flow $\Phi^F : \mathbb{R}\times E \to E$ is of class C^2. Then, for all $t \in \mathbb{R}$, Φ_t^F is a symplectic transformation.*

Proof. It will be sufficient to show that, for all $u, v, w \in E$, and for all $t \in \mathbb{R}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathscr{J} D_u\Phi_t^F v)(D_u\Phi_t^F w) = 0.$$

Using the group property of the flow, one sees it is enough to show this at $t = 0$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathscr{J} D_u\Phi_t^F v)(D_u\Phi_t^F w)_{|t=0} = \mathscr{J}(\mathscr{J}^{-1}D_u^2F(v,\cdot))(w) + (\mathscr{J} v)(\mathscr{J}^{-1}D_u^2F(w,\cdot)).$$

where we used the continuity of $\mathscr{J}$, the Schwarz Lemma (exchange of partial derivatives) and the observation that

$$\mathscr{J}\frac{\partial\Phi^F}{\partial t}(u) = D_{u(t)}F \in E^*,$$

and hence, at $t = 0$,

$$\mathscr{J} D_u\left(\frac{\partial\Phi^F}{\partial t}\right)(u) = \left(D_u^2F\right),$$

which means that, for all $v \in E$,

$$\mathscr{J} D_u\left(\frac{\partial\Phi^F}{\partial t}\right)(u)v = \left(D_u^2F\right)(v,\cdot).$$

Note that both sides are elements of E^* since $u \in E \to \frac{\partial\Phi^F}{\partial t} \in E$ so that $D_u\left(\frac{\partial\Phi^F}{\partial t}\right)(u) \in \mathscr{B}(E,E)$. Using the anti-symmetry of $\mathscr{J}$, one then finds

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathscr{J} D_u\Phi_t^F v)(D_u\Phi_t^F w)_{|t=0} = D_u^2F(v,w) - D_u^2F(w,v) = 0.$$

□

Remark A.3.6. We point out that the proof, as it stands, is valid in infinite dimensional systems. Remark however that the conditions imposed on the flow Φ_t^F are very strong for systems in infinite dimension. Too strong actually to be of much use in that context. We use/need those conditions to apply the Schwarz Lemma at several points in the proof. Also, it is known that Hamiltonian flows in infinite dimension need not always be symplectic. In the framework of Section 5 it is possible to give sufficient smoothness conditions on the restriction of the flow to $\mathscr{D}$ that will guarantee the result, but we shall not need this. For a different set of technical conditions guaranteeing the symplecticity of the flow, we refer to [CM74].

A.3.2. **Symmetries and constants of the motion**

Hamiltonian dynamical systems have many special features, but the one important to us here is that there exists for them a special link between the symmetries of the dynamics and the constants of the motion. This link takes the form of a Hamiltonian version of Noether's Theorem, of which we already gave a simple version in Theorem A.3.1, and has far-reaching consequences, some of which we further explore in this section. Again, a general treatment can for example be found in [AM78, LM87]; we give just those few elements needed in these notes.

We start with some notions on Hamiltonian Lie group actions on a symplectic vector space.

Definition A.3.7. Let G be a Lie group and $\Phi : (g,x) \in G \times E \to \Phi_g(x) \in E$, an action of G on E with $\Phi_g \in C^1(E,E)$. We will say Φ is globally Hamiltonian if Φ_g is symplectic for all $g \in G$ and if, for all $\xi \in \mathfrak{g}$, there exists $F_\xi \in C^2(E,\mathbb{R})$ so that $\Phi_{\exp(t\xi)} = \Phi_t^{F_\xi}$.

In other words, an action is globally Hamiltonian if all Φ_g are symplectic and if all one parameter groups are realized by Hamiltonian flows. In the notation of the previous sections this means that

$$X_\xi = X_{F_\xi}.$$

Here, the left hand side is the generator of the action, defined in (A.2.15) and the right hand side is the Hamiltonian vector field associated to F_ξ.

Remark A.3.8. In view of Theorem A.3.5, if $g = \exp(\xi)$ for some $\xi \in \mathfrak{g}$ and Φ_g can be written as $\Phi_{\exp(\xi)} = \Phi_1^{F_\xi}$ for some $F_\xi \in C^2(E,\mathbb{R})$ such that $\Phi_1^{F_\xi}$ is C^2, then Φ_g is symplectic. This will obviously hold as well for all g that can be written as a finite product of elements of the form $\exp(\xi)$, which is the case for all g in the connected component of G containing $e \in G$ (See [LM87], page 145, Proposition 2.10). So the assumption that Φ_g is symplectic is only needed for elements g that are not connected to $e \in G$. In infinite dimensional systems, as indicated in Remark A.3.6 at the end of the previous section, the condition that all Φ_g must be symplectic is more restrictive. In practice, one often works with linear actions of the symmetry group, for which the symplectic property can be checked directly.

The above definition is a special case of the more general definition of *globally Hamiltonian action* for infinite dimensional systems that we introduced in Definition 5.10. It suffices to take $\mathscr{D} = E$ in the latter to obtain the definition here.

We shall now continue with the abstract theory where, in particular, we will see through a version of Noether's theorem that, if the Hamiltonian is invariant under a globally Hamiltonian action Φ as above, then the functions $F_\xi \in C^2(E,\mathbb{R})$ are constants of the motion. The theory will be illustrated in Example A.3.13 at the end of the section, in the simple case where $E = \mathbb{R}^6$ and $G = \mathrm{SO}(3)$.

Theorem A.3.9. *Let G be a Lie group and Φ a globally Hamiltonian action of G on a symplectic vector space E. Let $H \in C^1(E,\mathbb{R})$ and let Φ_t^H be the corresponding Hamiltonian flow. Suppose that*

$$\forall g \in G, \quad H \circ \Phi_g = H. \tag{A.3.20}$$

Then the following statements hold.

(i) For all $\xi \in \mathfrak{g}$, $\{H, F_\xi\} = 0$.
(ii) For all $t \in \mathbb{R}$, $F_\xi \circ \Phi_t^H = F_\xi$.
(iii) G is an invariance group[24] for Φ_t^H.

Proof. This is an immediate consequence of Theorem A.3.1 and of Lemma A.3.4. □

This result is useful because it is often easy to check (A.3.20), whereas the conclusions (ii) and (iii) are statements about the flow Φ_t^H, which is usually not explicitly known, and are therefore hard to check directly. In particular, (iii) says that if the Hamiltonian H is G-invariant as a function, then G is an invariance group of the dynamics[25]. And (ii) ascertains that the group generators F_ξ are then constants of the motion for Φ_t^H.

Let us point out that (iii) implies neither (i), (ii) or (A.3.20) (See Lemma A.3.4.)

So the hypothesis that the Hamiltonian is invariant under the group action is strictly stronger than the statement that the Hamiltonian flow is invariant under G. The map

$$\xi \in \mathfrak{g} \to F_\xi \in C^2(E,\mathbb{R}) \tag{A.3.21}$$

[24] See Definition 2.1

[25] This is the point in the proof where the symplectic nature of the Φ_g is used, via Lemma A.3.4.

can be chosen to be linear. Indeed, if e_i, $i = 1, \ldots, d$ is a basis of $\mathfrak{g}$, if we choose $F_i = F_{e_i}$, and if we write $\xi = \sum_i \xi_i e_i$, we can define

$$F_\xi = \sum_i \xi_i F_i, \tag{A.3.22}$$

by linearity. This allows one to define the *momentum map* for the action Φ, as follows:

$$\mathscr{F} : u \in E \to \mathscr{F}(u) \in \mathfrak{g}^*, \quad \mathscr{F}(u)(\xi) = F_\xi(u). \tag{A.3.23}$$

This, of course, is just a rewriting of (A.3.21). In the main body of the text we shall always assume a basis has been chosen for $\mathfrak{g}$, as above, so that we can identify $\mathfrak{g} \simeq \mathbb{R}^m$. And we shall simply write

$$F : u \in E \to (F_1(u), \cdots, F_m(u)) \in \mathbb{R}^m \simeq \mathfrak{g}^*. \tag{A.3.24}$$

We shall refer to $\mathscr{F}$ or to F as a momentum map for the action, indifferently.

Definition A.3.10. Let Φ be a globally Hamiltonian action of G on E, with momentum map F. One says the momentum map is Ad^*-equivariant if, for all $g \in G$, for all $\xi \in \mathfrak{g}$,

$$F_\xi \circ \Phi_g = F_{\mathrm{Ad}_{g^{-1}}\xi}. \tag{A.3.25}$$

The terminology comes from the following observation. If (A.3.25) holds, then it follows from (A.3.23) and (A.2.11) that

$$\mathscr{F} \circ \Phi_g = \mathrm{Ad}^*_g \circ \mathscr{F}. \tag{A.3.26}$$

Since we identify $\mathfrak{g}^* \simeq \mathbb{R}^m$, this can be written

$$F \circ \Phi_g = \mathrm{Ad}^*_g F. \tag{A.3.27}$$

We can now formulate the final result from the theory of invariant Hamiltonian systems that we need. It is an immediate consequence of (A.3.27) or, for the reader weary of duals, of (A.3.25).

Proposition A.3.11. *Let Φ be a globally Hamiltonian, Ad^*-equivariant action of a Lie group G on a symplectic vector space E. Let $\mu \in \mathfrak{g}^* \simeq \mathbb{R}^m$ and define*

$$\Sigma_\mu = \{u \in E \mid F(u) = \mu\} \tag{A.3.28}$$

Then $G_\mu = G_{\Sigma_\mu}$, where G_μ is the stabilizer of μ, defined in (A.2.12) *and G_{Σ_μ} is defined in* (2.11)*.*

The situation we have in mind is the one where G is such that $H \circ \Phi_g = H$, for all $g \in G$. By Theorem A.3.9, the functions F_i are then constants of the motion for the flow Φ^H_t and hence the surfaces Σ_μ are Φ^H_t invariant. We can therefore consider the dynamical system (Σ_μ, Φ^H_t), which has G_μ as an invariance group (G_μ leaves invariant both Σ_μ and the flow Φ^H_t). This viewpoint will prove useful in the study of orbital stability in several situations.

Definition A.3.12. Let Φ be a globally Hamiltonian action of a Lie group G on a symplectic vector space E. Let $\mu \in \mathfrak{g}^*$. We say μ is a regular point of the momentum map F if, for all $u \in \Sigma_\mu$, $D_u F$ is surjective.

This definition simply guarantees that Σ_μ is a co-dimension m submanifold of E, where m is the dimension of $\mathfrak{g}$.

Example A.3.13. For the simple Hamiltonian system with spherical potentials considered in Section 2.3 and Section 4, one has $E = \mathbb{R}^6$, $G = \mathrm{SO}(3)$, and it is not difficult to check

that, for all $u(q,p) \in \mathbb{R}^6$, $F(u) = L(q,p) \in \mathbb{R}^3 \simeq \mathrm{so}(3)^*$ and $F_\xi(q,p) = \xi \cdot L(q,p)$, where we use the identifications (A.2.6) and (A.2.13). Furthermore, for all $R \in \mathrm{SO}(3)$,

$$L(Rq,Rp) = RL(q,p),$$

which shows the action is Ad^*-invariant, in view of (A.2.14).

We end this section with some comments on the Poisson brackets of the components of the momentum map. Remark first that the momentum map of a globally Hamiltonian action is not unique since, for any choice of $\lambda \in \mathfrak{g}^*$, $\tilde{F}_\xi = F_\xi + \lambda(\xi)$ also satisfies $X_\xi = X_{\tilde{F}_\xi}$. Note furthermore that, in view of (A.2.16) and (A.3.14), the momentum map satisfies, for all $\xi, \eta \in \mathfrak{g}$,

$$X_{F_{[\xi,\eta]}} = X_{[\xi,\eta]} = X_{\{F_\xi, F_\eta\}}.$$

It then follows from (A.3.3) that, for all $\xi, \eta \in \mathfrak{g}$, there exists a constant $c(\xi,\eta)$ so that

$$F_{[\xi,\eta]} = \{F_\xi, F_\eta\} + c(\xi,\eta).$$

The following lemma is useful and an easy consequence of (A.3.27):

Lemma A.3.14. *Let Φ be a globally Hamiltonian action of G on E, with momentum map F. If F is Ad^*-equivariant, then, for all $\xi, \eta \in \mathfrak{g}$,*

$$F_{[\xi,\eta]} = \{F_\xi, F_\eta\}. \tag{A.3.29}$$

Conversely, if (A.3.29) *holds, then* (A.3.25) *holds for all $g \in G$ of the form $g = \exp(\eta)$, for some $\eta \in \mathfrak{g}$ and then for all g in the connected component of e.*

What one has to remember here is this. In applications, we often wish to assure (A.3.25) holds. The preceding lemma states this is essentially guaranteed by (A.3.29), at least for all $g = \exp \eta$, which, for many Lie groups, means all of G. Finally, (A.3.29) is guaranteed by

$$\{F_i, F_j\} = c_{ij}^k F_k, \tag{A.3.30}$$

where we used the notation introduced in (A.2.9) and (A.3.22). As an example, one may remark that the components of the angular momentum vector L satisfy the commutation relations of the Lie algebra of $\mathrm{SO}(3)$, namely

$$\{L_i, L_j\} = \varepsilon_{ijk} L_k, \quad i,j,k = 1,2,3.$$

One may therefore show that an action is Ad^*-equivariant by showing (A.3.30) holds. However, in infinite dimension, this is not immediate since the necessary smoothness properties of the F_i's and even of the corresponding Hamiltonian vector fields are not readily verified.

REFERENCES

[AM78] Ralph Abraham and Jerrold 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 Raţiu and Richard Cushman.

[AP09] Jaime 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.

[Arn99] V. I. Arnold. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 199? Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition.

[BER99] M. Salah Baouendi, Peter Ebenfelt, and Linda Preiss Rothschild. *Real submanifolds in complex space and their mappings*. Princeton Mathematical Series. Princeton University Press, 1999.

[Bou93] Jean Bourgain. Fourier transform restriction phenomena for certain lattice subsets and application to nonlinear evolution equations. *Geom. Funct. Anal.*, 3(2), 1993.

[Caz03] Thierry Cazenave. *Semilinear Schrödinger Equations*. Courant Lecture Notes. American Mathematical Society, 2003.

[CL82] Thierry Cazenave and Pierre-Louis Lions. Orbital stability of standing waves for some nonlinear Schrödinger equations. *Comm. Math. Phys.*, 85:549–561, 1982.

[CM74] Paul R. Chernoff and Jerrold E. Marsden. *Properties of infinite dimensional Hamiltonian systems*. Lecture Notes in Mathematics, Vol. 425. Springer-Verlag, Berlin, 1974.

[Cuc11] Scipio Cuccagna. The hamiltonian structure of the nonlinear Schrödinger equation and the asymptotic stability of its ground states. *Comm. Math. Phys.*, 305(2):279–331, 2011.

[DV69] Michel Duflo and Michel Vergne. Une propriété de la représentation coadjointe d'une algèbre de Lie. *C.R. Acad. Sc. Paris*, 268(A):583–585, 1969.

[FGL13] Erwan Faou, Ludwig Gauckler, and Christian Lubich. Sobolev stability of plane wave solutions to the cubic nonlinear Schrödinger equation on a torus. *Communications in Partial Differential Equations*, 38(7):1123–1140, 2013.

[Gaz12] Maxime Gazeau. *Analyse de modèles mathématiques pour la propagation de la lumière dans les fibres optiques en présence de biréfringence aléatoire*. PhD thesis, École Polytechnique, 2012.

[Gen09] François Genoud. Existence and orbital stability of standing waves for some nonlinear Schrödinger equations, perturbation of a model case. *J. Differential Equations*, 246:1921–1943, 2009.

[Gen10a] François Genoud. Bifurcation and stability of travelling waves in self-focusing planar waveguides. *Adv. Nonlinear Stud.*, 10:357–400, 2010.

[Gen10b] François Genoud. A smooth global branch of solutions for a semilinear elliptic equation on $\mathbb{R}^n$. *Calc. Var. Partial Differential Equations*, 38:207–232, 2010.

[Gen11] François Genoud. Bifurcation from infinity for an asymptotically linear problem on the half-line. *Nonlinear Anal.*, 74:4533–4543, 2011.

[Gen13] François Genoud. Orbitally stable standing waves for the asymptotically linear one-dimensional NLS. *Evol. Equ. Control Theory*, 2:81–100, 2013.

[GH07a] Thierry Gallay and Mariana Hărăgus. Orbital stability of periodic waves for the nonlinear Schrödinger equation. *J. Dynam. Differential Equations*, 19(4):825–865, December 2007.

[GH07b] Thierry Gallay and Mariana Hărăgus. Stability of small periodic waves for the nonlinear Schrödinger equation. *J. Differential Equations*, 234(2):544–581, March 2007.

[Gol80] Herbert Goldstein. *Classical mechanics*. Addison-Wesley Publishing Co., Reading, Mass., second edition, 1980. Addison-Wesley Series in Physics.

[GS08] François Genoud and Charles A. Stuart. Schrödinger equations with a spatially decaying nonlinearity: existence and stability of standing waves. *Discrete Contin. Dyn. Syst.*, 21:137–186, 2008.

[GSS87] Manoussos Grillakis, Jalal Shatah, and Walter Strauss. Stability theory of solitary waves in the presence of symmetry. I. *J. Funct. Anal.*, 74(1):160–197, 1987.

[GSS90] Manoussos Grillakis, Jalal Shatah, and Walter Strauss. Stability theory of solitary waves in the presence of symmetry. II. *J. Funct. Anal.*, 94(2):308–348, 1990.

[HS04] Hichem Hajaiej and Charles A. Stuart. On the variational approach to the stability of standing waves for the nonlinear Schrödinger equation. *Adv. Nonlinear Stud.*, 4:469–501, 2004.

[LC09] Stefan Le Coz. *Analytical and numerical aspects of partial differential equations*, chapter Standing waves in nonlinear Schrödinger equations, pages 151–192. Walter de Gruyter, Berlin, 2009.

[LM87] Paulette Libermann and Charles-Michel 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.

[LS98] Eugene M. Lerman and Stephanie F. Singer. Stability and persistence of relative equilibria at singular values of the moment map. *Nonlinearity*, 11(6):1637–1649, 1998.

[Man74] Sergei V. Manakov. On the theory of two-dimensional stationary self-focusing of electromagnetic waves. *Sov. Phys. JETP*, 38(2):248–253, 1974.

[Mon97] James Montaldi. Persistence and stability of relative equilibria. *Nonlinearity*, 10(2):449–466, 1997.

[MR94] Jerrold E. Marsden and Tudor S. Ratiu. *Introduction to mechanics and symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1994. A basic exposition of classical mechanical systems.

[MRO11] James Montaldi and Miguel Rodríguez-Olmos. On the stability of Hamiltonian relative equilibria with non-trivial isotropy. *Nonlinearity*, 24(10):2777–2783, 2011.

[Pat92] George 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.

[PRW04] George W. Patrick, Mark Roberts, and Claudia Wulff. Stability of Poisson equilibria and Hamiltonian relative equilibria by energy methods. *Arch. Ration. Mech. Anal.*, 174(3):301–344, 2004.

[RSS06] Mark Roberts, Tanya Schmah, and Cristina Stoica. Relative equilibria in systems with configuration space isotropy. *J. Geom. Phys.*, 56(5):762–779, 2006.

[Sou97] Jean-Marie Souriau. *Structure of dynamical systems: a symplectic view of physics*, volume 149 of *Progress in Mathematics*. Springer, 1997.

[Spi79] Michael Spivak. *A comprehensive introduction to differential geometry. Vol. I*. Publish or Perish Inc., Wilmington, Del., second edition, 1979.

[SS85] Jalal Shatah and Walter Strauss. Instability of nonlinear bound states. *Comm. Math. Phys.*, 100:173–190, 1985.

[SS99] Catherine Sulem and Pierre-Louis Sulem. *The nonlinear Schrödinger equation. Self-focusing and wave collapse.* Springer-Verlag, New York, 1999.

[Stu98] Charles A. Stuart. An introduction to elliptic equations on $\mathbb{R}^n$. In *Nonlinear Functional Analysis and Applications to Differential Equations (Trieste, 1997)*, pages 237–285. World Science Publishing, River Edge, 1998.

[Stu08] Charles A. Stuart. Lectures on the orbital stability of standing waves and application to the nonlinear Schrödinger equation. *Milan j. math.*, 76:329–399, 2008.

[Tao06] Terence Tao. Cbms regional conf. ser. math. In *Nonlinear Dispersive Equations. Local and Global Analysis.* American Mathematical Society, 2006.

[Tao09] Terence Tao. Why are solitons stable? *Bull. Amer. Math. Soc.*, 46(1):1–33, 2009.

[VK73] Nazib Vakhitov and Aleksandr A. Kolokolov. Stationary solutions of the wave equation in a medium with nonlinearity saturation. *Radiophys. Quantum Electron.*, 16, 1973.

[Wei86] Michael I. Weinstein. Lyapunov stability of ground states of nonlinear dispersive evolution equations. *Comm. Pure Appl. Math.*, 39(1):51–67, 1986.

[Zhi01] Peter E. Zhidkov. *Korteweg-de Vries and Nonlinear Schrödinger Equations: Korteweg-de Vries and Nonlinear Schrödinger Equations: Qualitative Theory*. Lecture Notes in Mathematics. Springer, 2001.

[1]LABORATOIRE PAUL PAINLEVÉ, CNRS, UMR 8524 ET UFR DE MATHÉMATIQUES, UNIVERSITÉ LILLE 1, SCIENCES ET TECHNOLOGIES F-59655 VILLENEUVE D'ASCQ CEDEX, FRANCE.

E-mail address: Stephan.De-Bievre@math.univ-lille1.fr

E-mail address: Simona.Rota-Nodari@math.univ-lille1.fr

[2] 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.

[3]FACULTY OF MATHEMATICS, UNIVERSITY OF VIENNA, OSKAR-MORGENSTERN-PLATZ 1, 1090 VIENNA, AUSTRIA.

E-mail address: francois.genoud@univie.ac.at