Nonlinearity 8 (1995) 1087–1113. Printed in the UK

Approximations with curves of relative equilibria in Hamiltonian systems with dissipation

Gianne Derks†§, Debra Lewis‡‖ and Tudor Ratiu‡¶

† Department of Mathematics and Statistics, Simon Fraser University, Burnaby, BC, Canada V5A 1S6

‡ Department of Mathematics, University of California, Santa Cruz, CA 95064, USA

Received 15 September 1994, in final form 7 August 1995
Recommended by J M Sanz-Serna

Abstract. In this paper we will investigate the relevance of a stable family of relative equilibria in a dissipative Hamiltonian system with symmetry. We are interested in relative equilibria of the Hamiltonian system, whose stability follows from the fact that they are local extrema of the energy–momentum function which is a combination of the Hamiltonian and a conserved quantity of the Hamiltonian system, induced by the momentum map related to the symmetry group.

Although the dissipative perturbation is equivariant under the action of the symmetry group, it will destroy the conservation law associated with the symmetry group. We will specify its dissipative properties in terms of the induced time behaviour of the momentum map and quasi-static attractive properties of the relative equilibria. By analysing the time behaviour of the previously mentioned energy–momentum function we derive sufficient conditions such that solutions of the dissipative system which are initially close to a relative equilibrium can be approximated by a (long) curve of relative equilibria. At the end we illustrate the method by analysing the example of a rigid body in a rotational symmetric field with dissipative rotation-like perturbation added.

AMS classification scheme numbers: 58F30, 34D10, 34D35, 34E10

1. Introduction

The behaviour of solutions of Hamiltonian systems with symmetry has long been a subject of intensive research. The analysis of relative equilibria plays a key role in this research. Relative equilibria are equilibria modulo symmetries. For example, if the symmetry group is a rotation group, then the relative equilibria are uniformly rotating states. Relative equilibria form a highly structured class of motions, which makes them accessible for detailed analysis. Two (related) systematic ways to analyse the stability of relative equilibria are the energy-Casimir method (see Holm *et al* [7], Krishnaprasad and Marsden [8] and related papers) and the (reduced) energy–momentum method (see Simo *et al* [13], Simo *et al* [14] and references therein). The key to both these methods is the characterization of

§ Research supported by a NATO research science fellowship of the Netherlands Organization for Scientific Research (NWO).

‖ Research supported by the National Science Foundation, DMS-9122708, and UCSC Faculty Research Grant 503128–09523.

¶ Research supported by the National Science Foundation, DMS-9122708, the Department of Energy, DE–FG03–92ER25129, the Office of Naval Research, N00014–92–J–1545, and UCSC Faculty Research Grant 503185–19900.

relative equilibria as critical points of the so-called energy–momentum function, which is determined by the Hamiltonian and the momentum map, or related induced functions. If the relative equilibria are local extremals of the energy–momentum function modulo certain symmetries, then they are orbitally stable.

However, purely Hamiltonian systems seldom occur in reality. Often (small) perturbations that destroy the Hamiltonian structure are present. In this paper we consider the relevance of the relative equilibria in the presence of weak dissipation. If momentum is dissipated, most relative equilibria are not preserved, even modulo symmetries. Any trajectory must pass through the appropriate momentum level sets and will eventually leave the neighbourhood of the relative equilibrium it initially approximated and deviate far from this initial neighbourhood. Thus it is not generally useful to talk about the stability of a single relative equilibrium, but rather a *long curve* of relative equilibria. The basic question we address now is the following.

> If a solution of the dissipative system starts near a relative equilibrium of the unperturbed system, can one sharply approximate it by a *time-dependent curve of relative equilibria*?

We shall see that under some reasonable hypotheses, it is possible to characterize a curve of relative equilibria with dissipating momentum as being attracting. In Derks and Valkering [5] this is shown for finite-dimensional mechanical systems with one cyclic coordinate and uniform friction. An extension to more general Hamiltonian systems, but still with only one-dimensional symmetry groups, is given in Derks [3] and Derks and van Groesen [4]. They consider the approximation of solutions of the uniformly damped periodic Korteweg–de Vries equation with a curve of cnoidal waves and show that it can be approximated by a curve of cnoidal waves (which are solitary wave-like solutions of the periodic KdV equation). The approximation is sharp in the sense that as the solutions tend to zero as $t \to \infty$, their difference tends to zero in a norm that sharply picks out differences in shape. Roughly speaking, this means that the solution converges to the solitary waves at the same rate as the dissipation causes it to disappear.

In this paper we will generalize the work of [5], by considering a finite-dimensional symplectic manifold and a compact (possibly non-Abelian) group of Hamiltonian symmetries defining the momentum map $\boldsymbol{J}$. We assume that the unperturbed Hamiltonian system possesses a smooth manifold of relative equilibria which are stable according to the criteria of the energy–momentum method. This assumption is motivated by the work of Bloch *et al* [2]. They consider a relative equilibrium for which the energy–momentum method predicts formal instability. By adding a small, momentum preserving damping to the Hamiltonian system, the relative equilibrium becomes unstable. It seems unlikely that such relative equilibria are stable under perturbations which do not preserve the momentum map, hence our assumption on the stability of the relative equilibria. This assumption allows us to use the energy–momentum function to estimate distances on a neighbourhood of the relative equilibria.

We assume that the perturbation is smooth, dissipative, and equivariant for the action of the symmetry group. Furthermore, there are three technical hypotheses on the perturbation. First, the influence of the perturbation on the momentum map has to be such that the value of the momentum map $\boldsymbol{J}(u(t))$ of any solution $u(t)$ has a limit for $t \to \infty$. This assumption allows us to provide an asymptotic prediction of the behaviour of the system. Secondly, the manifold of relative equilibria need not be invariant under the perturbed dynamics, but the effects of the perturbation that push trajectories away from the manifold of relative equilibria should not be strong for a long time. Finally, every relative equilibrium has to be attractive in a 'quasi-static' context. This means that the linearization of the perturbation in directions tangential to the level set of the momentum map is attractive towards the relative equilibria.

For an approach proving the stability of the curve of relative equilibria, we look at the conservative case. Can the traditional variational analysis of the relative equilibria be used in a dissipative setting? The answer is, to a great extent, yes. Once the stability of the relative equilibria in the absence of dissipation and the asymptotic properties of the evolution of the momentum map itself under the influence of dissipation are known, we have most of the information in hand that is required to analyse the stability of the dissipative trajectories. The approximation result is the following.

> Let $u(t)$ be a solution of the perturbed system and let $\mu(t) = J(u(t))$. If the initial distance between $u(0)$ and the manifold of relative equilibria is small, say of order $\mathcal{O}(\varepsilon)$, then the distance between $u(t)$ and the relative equilibria on the level set $J = \mu(t)$ is of order $\mathcal{O}(\varepsilon\, \mathrm{e}^{-\kappa t})$, for all $t \geqslant 0$, where the constant κ depends only on the perturbation.

Our strategy for the proof of this estimate is as follows: a familiar approach to the stability analysis of relative equilibria in the absence of dissipation is to use the energy–momentum function as a Lyapunov function. In the presence of dissipation, it is natural to hope that one can estimate the time derivative of this Lyapunov function based on the dissipative equations. However, it turns out that this estimate is not sufficiently sharp because the family of relative equilibria is not invariant for the perturbed system. To sharpen the estimate one needs to define a slightly different Hamiltonian. We construct a small (of the order of the perturbation) addition to the Hamiltonian and show that this new Hamiltonian has relative equilibria that satisfy the perturbed equation to one higher order. (A similar idea for this construction is used in [3–5] and in Lebovitz and Neishtadt [9].)

As an application of the general ideas, we study the example of a rigid body in a rotationally symmetric field with a dissipative rotation-like perturbation. This example, with configuration space $\mathbb{R}^3 \times SO(3)$, is intended to illustrate some of the geometric considerations: the phase space is a nonlinear manifold, the group does not act freely on the limiting relative equilibrium, and the subgroup of momentum-preserving symmetries is not constant. Nonetheless, the analysis can be carried out and the approximation with the curve of relative equilibria can be verified.

Appropriate manipulations of the symmetries of the system will be a recurrent theme in our analysis. Hence we briefly discuss some of the most important concerns. It is a well known aspect of the study of Hamiltonian systems with symmetry that relative equilibria are fixed points of the induced dynamics on an appropriate orbit manifold. Many of the techniques for analysing such equilibria are formulated on quotient manifolds. Symplectic reduction, in which the reduced manifold is the quotient of a level set of the momentum map by the subgroup of symmetries preserving that level set, and Poisson reduction, in which the reduced phase space is the quotient of the original phase space by the full symmetry group, are both well known. See, for example, Meyer [12] and Marsden and Weinstein [11].

Such an approach has several essential limitations in the present context. Symplectic reduction involves restriction to the momentum level set and determination of the quotient with respect to the subgroup of momentum preserving group elements. However, as momentum is dissipated, the momentum level set clearly changes; in many cases the isotropy subgroup changes as well. Thus application of symplectic reduction seems inherently inappropriate in this context and the assumption of a fixed isotropy subgroup would involve a significant restriction of the applicability of the technique.

Another possible quotient would be the Poisson reduced space, that is, the quotient of the full space by the full symmetry group. However, this would weaken the results for non-Abelian group actions, since we could only show that the perturbed trajectory remains near the full group orbit of the manifold of relative equilibria, not the manifold itself. Even in the Abelian case, it would not significantly reduce the work that has to be done. In many cases, it can be very hard to determine the Poisson reduced space explicitly.

Our approach closely resembles slice techniques. We choose a representative curve of relative equilibria and select elements in the isotropy subgroup orbit of the perturbed trajectory which are close to this curve of relative equilibria. In conclusion, the motivation for our approach to the problem is based on tools used in reduction theory, but for deriving explicit estimates on the time behaviour, the full space seems to be more convenient.

To derive our estimates on the energy–momentum function, we use charts for the local analysis. We do not insist that the metrics or charts used in the analysis be equivariant. This is motivated by practical considerations—since various coefficients need to be explicitly computed in the charts, we want as much flexibility as possible in choosing convenient charts. While this occasionally leads to more complicated proofs, we believe that it is important to see that the results can be obtained in a very general setting. Thus, while the dynamics are equivariant and our final result is phrased in terms of orbits, we do not insist that equivariance be maintained at every step of the proof.

We note that most of the hypotheses required to show stability of the dissipative trajectory are related to those needed to show stability of the relative equilibria of the original conservative system. We have attempted to formulate our analysis in such a way that as little work as possible needs to be done to translate stability results from the conservative context to the dissipative one.

At the end of this introduction, we give a short description of each section. Section 2 contains a detailed description of the class of systems under consideration and provides the estimate that is the ultimate goal of the paper. In this section we introduce hypotheses to provide sufficient conditions for the existence of the manifold of relative equilibria and the previously described desired behaviour of the perturbation. It should be mentioned that these conditions are definitely not necessary. One can prove a number of related theorems using slight variations of the hypotheses.

In section 3 we specify some properties of the chart maps, that we will use in the estimates. In the section 4 we will show the existence of a curve of 'improved relative equilibria' and introduce the Lyapunov function that will allow us to derive the desired estimates on the distance to the manifold of relative equilibria in the last part of section 4.

Finally, in section 5 we consider the previously mentioned example of a rigid body in a rotationally symmetric potential field with dissipation as an application of the general results.

2. The Hamiltonian system with dissipation

Let $(\mathcal{M}, \omega, G, J, H)$ be a finite-dimensional symplectic G-space. This means that $(\mathcal{M}, \omega)$ is a symplectic manifold together with the symplectic action of a compact finite-dimensional Lie group G on $\mathcal{M}$, an equivariant momentum map $J : \mathcal{M} \to \mathfrak{g}^*$ and a G-invariant Hamiltonian $H : \mathcal{M} \to \mathbb{R}$. The symbols $\mathfrak{g}$ and $\mathfrak{g}^*$ denote the Lie algebra, respectively the dual Lie algebra, of the Lie group G. The pairing between the Lie algebra and its dual is denoted by $\langle \cdot, \cdot \rangle$. Furthermore, the norms on $\mathfrak{g}$ and $\mathfrak{g}^*$ are denoted by $|\cdot|$.

The symplectic structure on $\mathcal{M}$ induces an invertible Poisson structure $\Gamma : T^*\mathcal{M} \to$

$T\mathcal{M}$ given by

$$\omega(\delta u, \Delta u) = \langle \Gamma^{-1}(u)\delta u, \Delta u\rangle_{\mathcal{M}}$$

for all $u \in \mathcal{M}$ and $\delta u, \Delta u \in T_u\mathcal{M}$. The symbol $\langle\cdot,\cdot\rangle_{\mathcal{M}}$ denotes the pairing between $T^*\mathcal{M}$ and $T\mathcal{M}$. In the following we will no longer write the subscript $\mathcal{M}$ explicitly.

The Hamiltonian vector field is denoted by $X_H : \mathcal{M} \to T\mathcal{M}$. For every $\xi \in \mathfrak{g}$ we define the function $J_\xi : \mathcal{M} \to \mathbb{R}$ induced by the momentum map as

$$J_\xi(u) = \langle J(u), \xi\rangle \qquad u \in \mathcal{M}.$$

These functions are conserved quantities for the (unperturbed) Hamiltonian system $\dot{u} = X_H(u)$ and they will be used to form the augmented Hamiltonian.

The level sets of the momentum map are denoted by $\mathcal{M}_\mu$, that is,

$$\mathcal{M}_\mu = \{u \in \mathcal{M} \mid J(u) = \mu\}$$

for any $\mu \in \mathfrak{g}^*$. Noether's theorem implies that these level sets are invariant under the flow of the Hamiltonian system $\dot{u} = X_H(u)$. See Abraham and Marsden [1] or Guillemin and Sternberg [6] for more information on symplectic G-spaces.

We consider the following dynamical system on $\mathcal{M}$

$$\dot{u} = X_H(u) + \varepsilon\, P(u) = \Gamma(u)\, DH(u) + \varepsilon\, P(u). \tag{1}$$

In this expression $P : \mathcal{M} \to T\mathcal{M}$ is a smooth perturbation which is equivariant for the action of the group G and ε is a small parameter which measures the strength of the perturbation.

Remark 1. There is no real difference in the analysis in the case where we consider a perturbation of the form $\varepsilon\, P(u, \varepsilon)$, with $P(u, \varepsilon)$ bounded uniform in u for ε small. If one considers this case, at some points one has to make sure that the desired behaviour is uniform in ε.

In the following we will make additional hypotheses on the system (1). These hypotheses guarantee the existence of stable relative equilibria in the Hamiltonian system (hence for $\varepsilon = 0$). Furthermore, the hypotheses specify the influence of the perturbation on the momentum map and on the relative equilibria. One aspect of the perturbation specified by the hypotheses is that the perturbation has a certain dissipative behaviour. Further on we will define this behaviour which we call G_μ-orbit dissipative.

2.1. Relative equilibria of the Hamiltonian system

For $\varepsilon = 0$, the system (1) is Hamiltonian. If the following three hypotheses are satisfied, then this Hamiltonian system possesses a smooth family of stable relative equilibria. For more details about relative equilibria, see Abraham and Marsden [1], Marsden [10], and Simo *et al* [13].

(H1) There exists a smooth connected manifold of relative equilibria, denoted by MRE. This manifold has the property that there exists a subset $\mathfrak{g}^*_{\mathrm{MRE}} \subset \mathfrak{g}^*$ such that for every value of $\mu \in \mathfrak{g}^*_{\mathrm{MRE}}$ there exists at least one relative equilibrium with momentum μ. If $\bar{u}$ is a relative equilibrium in MRE with $\mu = J(\bar{u})$, then the set MRE_μ of relative equilibria in MRE with momentum μ is equal to the G_μ-orbit of $\bar{u}$. Specifically, there are smooth maps $\bar{u} : \mathfrak{g}^*_{\mathrm{MRE}} \to \mathrm{MRE}$ and $\bar{\xi} : \mathfrak{g}^*_{\mathrm{MRE}} \to \mathfrak{g}$ such that for each $\mu \in \mathfrak{g}^*_{\mathrm{MRE}}$ the relative equilibrium $\bar{u}(\mu)$ has momentum μ and generator $\bar{\xi}(\mu)$, i.e. $\bar{u}(\mu)$ is a critical point of the augmented Hamiltonian or energy–momentum function $H_\mu = H - J_{\bar{\xi}(\mu)}$. Note that

$$\mathrm{MRE} = \bigcup_{\mu\in\mathfrak{g}^*_{\mathrm{MRE}}} \mathrm{MRE}_\mu = \bigcup_{\mu\in\mathfrak{g}^*_{\mathrm{MRE}}} G_\mu \cdot \bar{u}(\mu).$$

(H2) For every $\bar{u} \in$ MRE the derivative of the momentum map $DJ(\bar{u}) : T_{\bar{u}}\mathcal{M} \to \mathfrak{g}^*$ is surjective. This implies that the group G acts freely in a neighbourhood of MRE and that all points in this neighbourhood are regular points of J. In particular, $T_{\bar{u}}\mathcal{M}_\mu = \ker[DJ(\bar{u})]$ for any $\bar{u} \in \text{MRE}_\mu$.

(H3) For all $\mu \in \mathfrak{g}^*_{\text{MRE}}$ the second derivative of the augmented Hamiltonian H_μ at the point $\bar{u}(\mu)$ is positive semi-definite on $T_{\bar{u}}\mathcal{M}_\mu = \ker[DJ(\bar{u}(\mu))]$, with kernel $\mathfrak{g}_\mu \cdot \bar{u} = \{X_{J_\xi}(\bar{u}) \mid \xi \in \mathfrak{g}_\mu\}$.

This hypothesis implies that all relative equilibria $\bar{u}(\mu)$ are G_μ-orbitally stable; see Simo *et al* [13]. We define the Lyapunov function $L(u)$ by

$$L(u) = H(u) - H(\bar{u}(\mu)) = H_\mu(u) - H_\mu(\bar{u}(\mu)) \qquad \text{where} \quad \mu = J(u).$$

Note that for convenience, we frequently do not indicate the explicit μ, ε, or t dependence of functions, e.g. $\bar{y}$ rather than $\bar{y}(\mu)$.

For $u \in \mathcal{M}_\mu$ sufficiently near MRE $L(u)$ provides an estimate of the distance between the G_μ-orbit of u and $\bar{u}(\mu)$. Let $\mu \in \mathfrak{g}^*_{\text{MRE}}$ and let d be a distance function on $\mathcal{M}$ which is compatible with the Euclidean norm on $\mathbb{R}^{2n}$. Define the following G_μ-orbit distance on $\mathcal{M}_\mu$:

$$\rho_d(u_1, u_2) = \min_{g \in G_\mu} d(g \cdot u_1, u_2) \qquad u_1, u_2 \in \mathcal{M}_\mu.$$

Then there exist constants $0 < c(\mu) \leqslant C(\mu)$ such that for all $u \in \mathcal{M}_\mu$ in a neighbourhood of MRE_μ we have

$$c(\mu)\, \max\{\rho_d(u, \bar{u}(\mu)), \rho_d(\bar{u}(\mu), u)\} \leqslant \sqrt{L(u)} \leqslant C(\mu)\, \min\{\rho_d(u, \bar{u}(\mu)), \rho_d(\bar{u}(\mu), u)\}.$$

(See lemma 8 for a proof of a generalized version of this result.) The explicit μ dependence of the constants $c(\mu)$ and $C(\mu)$ can lead to complications as the trajectory moves through the momentum level sets. We shall see in our example that $c(\mu)$ approaches zero as the trajectory approaches its limiting value. Hence we shall modify this result, replacing the fixed distance function d with a family of μ-dependent distance functions. In this way, although our 'control' over some of the variables grows increasingly weak as we approach the limit, we still have good estimates for most of the variables.

2.2. *Dissipation of the momentum map*

The next hypotheses are related to the perturbation P. First we specify the dissipative influence of P on the evolution of the momentum map.

For a solution $u(t)$, the time evolution of $\mu(t) = J(u(t))$ is given by

$$\dot{\mu} = \varepsilon DJ(u) \cdot P(u). \tag{2}$$

Hence μ is a function of a slow time variable $\tau = \varepsilon t$. We are interested in the case that the function $\mu(t)$ has an asymptotic value.

(H4) For any solution $u(t)$ of (1), the curve $\mu(t) = J(u(t))$ stays in $\mathfrak{g}^*_{\text{MRE}}$ and this curve has a limit for $t \to \infty$, say μ_∞. Furthermore, $\lim_{t\to\infty} \bar{u}(\mu(t))$ exists and the integral $\int_0^\infty |\dot{\bar{\xi}}(\mu(t))| \mathrm{d}t$ exists and can be bounded independently of ε.

In the case that $\mu_\infty \notin \mathfrak{g}^*_{\text{MRE}}$, we additionally assume some uniformity in the properties (H2) and (H3), which will be specified in (H5).

It follows from hypothesis (H4) that the closure of the curve $(\bar{u}(\mu(t)))_{t\geqslant 0}$ is compact. For every compact subset of the manifold $\mathcal{M}$ there exist a finite number of chart maps covering a neighbourhood of this subset; hence we have the following property.

Property 1. *There exist a finite number, say N, of chart maps $\varphi_i : \mathcal{U}_i \subset \mathcal{M} \to \mathbb{R}^{2n}$, $i = 1, \ldots, N$ such that $\mathcal{U} = \cup_{i=1}^{N} \mathcal{U}_i$ is a full neighbourhood of the closure of the curve $(\bar{u}(\mu(t)))_{t\geqslant 0}$.*

To avoid excessive notation, we will no longer indicate the index i explicitly.

The Hamiltonian, the Lie algebra and the perturbation vector field induce functions and a vector field on $\mathbb{R}^{2n}$ through the chart maps. Furthermore, we can locally define a symplectic structure on $\mathbb{R}^{2n}$ that is compatible with the symplectic structure of $\mathcal{M}$.

Definition 2. *Let $\mu \in \mathfrak{g}^*_{\mathrm{MRE}}$. We define $\mathcal{Y} = \varphi(\mathcal{U}) \subset \mathbb{R}^{2n}$. We define the functions $h_\mu : \mathcal{Y} \to \mathbb{R}$, $j : \mathcal{Y} \to \mathfrak{g}^*$, and the vector field $p : \mathcal{Y} \to \mathbb{R}^{2n}$ as the push forwards of the energy–momentum function, etc, by the chart map φ. Specifically, if $y = \varphi(u) \in \mathcal{Y}$, then*

$$h_\mu(y) = H_\mu(u) \qquad j(y) = J(u) \qquad \text{and} \qquad p(y) = D\varphi(u)P(u)\,.$$

For $\eta \in \mathfrak{g}$, define $j_\eta(y) = \langle j(y), \eta \rangle = J_\eta(u)$. The induced Poisson structure $\gamma : T^*\mathcal{Y} \to T\mathcal{Y}$ is defined $\gamma = \varphi^*\Gamma$, i.e. $\gamma(y) = D\varphi(u)\,\Gamma(u)\,(D\varphi(u))^*$, for $y = \varphi(u) \in \mathcal{Y}$. The points $\bar{y}(\mu) = \varphi(\bar{u}(\mu))$ are critical points of h_μ.

From this definition and hypothesis (H3), it follows that for every $\mu \in \mathfrak{g}^*_{\mathrm{MRE}}$ $D^2h_\mu(\bar{y}(\mu))$ is positive semi-definite on $\ker[Dj(\bar{y}(\mu))]$, with kernel $\mathfrak{g}_\mu \cdot \bar{y}(\mu) = \{X_{j_\xi}(\bar{y}(\mu)) \mid \xi \in \mathfrak{g}_\mu\}$. We wish to be able to consider cases in which the trajectory $\mu(t)$ tends towards a limiting value outside the set $\mathfrak{g}^*_{\mathrm{MRE}}$ such that the constant $c(\mu)$ approaches zero. To be able to deal with this case, we will work with a scaled metric and a scaled distance function.

Definition 3. *Let $B_\mu : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, $\mu \in \mathfrak{g}^*_{\mathrm{MRE}}$, be a family of invertible linear transformations. Define μ-dependent inner products $(\ ,\ )_\mu$ on $\mathbb{R}^{2n}$ by*

$$(y_1, y_2)_\mu = (B_\mu y_1, B_\mu y_2) \qquad y_1, y_2 \in \mathbb{R}^{2n}.$$

Define the associated norms $|\,|_\mu$, the gradient ∇_μ determined by the inner product $(\ ,\ )_\mu$ (i.e. $(\nabla_\mu f(y), v)_\mu = Df(y)\cdot v$ for all differentiable functions $f : \mathbb{R}^{2n} \to \mathbb{R}$, all y and $v \in \mathbb{R}^{2n}$), the orthogonal complement $s(y)$ to $\mathfrak{g}_\mu \cdot y$ in $\ker[Dj(y)]$ with respect to $(\ ,\ )_\mu$, for $y \in \mathbb{R}^{2n}$.

Let $\tilde{d} : \mathcal{M} \times \mathcal{M} \to \mathbb{R}$ be a smooth function that is compatible with $|\,|_\mu$ in the sense that there exist positive constants c_d and C_d satisfying

$$c_d\,\tilde{d}(u_1, u_2) \leqslant |\varphi(u_1) - \varphi(u_2)|_\mu \leqslant C_d\,\tilde{d}(u_1, u_2) \qquad \text{for all} \quad u_1, u_2 \in \mathcal{U} \cap \mathcal{M}_\mu, \mu \in \mathfrak{g}^*\,.$$

The corresponding orbit distance function is

$$\rho(u_1, u_2) = \min_{g \in G_\mu} \tilde{d}(g \cdot u_1, u_2) \qquad u_1, u_2 \in \mathcal{M}_\mu,\ \mu \in \mathfrak{g}^*\,.$$

The next hypothesis assures a uniform behaviour of the functions h and and j and the vector field $\hat{p}(y) = \mathbb{P}_y p(y)$, where $\mathbb{P}_y$ denotes the projection onto the $(\ ,\ )_\mu$ orthogonal complement to $\mathfrak{g}_\mu \cdot y = \{X_{j_\xi}(y) \mid \xi \in \mathfrak{g}_\mu\}$, with respect to the norm $|\,|_\mu$. In the case that $B_\mu = Id$, this hypothesis is satisfied if H is C^3 and $\hat{p}$ is C^0. (Smoothness of the maps j and X_{j_ξ} follows immediately from the assumption of a smooth action on $\mathcal{M}$ and smooth chart maps.) We introduce the notation $LC^\ell_\mu(\mathcal{D}, \mathcal{R}; B)$ to denote the set of functions with domain $\mathcal{D} \subset \mathbb{R}^{2n}$, range $\mathcal{R}$ and Lipschitz continuous ℓth derivative with respect to the norm $|\,|_\mu$, with Lipschitz constant bounded by B. If the bound B is not specified, the Lipschitz constant is of order one, i.e. μ and ε independent. Note that distances between vectors are measured in the $|\,|_\mu$ norm, distances between covectors are measured in the operator norm determined by $|\,|_\mu$, and distances between elements of the algebra and the dual of the algebra are measured with respect to the standard adjoint and coadjoint invariant inner products. For example, $f \in LC^1_\mu(\mathcal{D}, \mathbb{R})$ if there exists some constant C such that $|(Df(y_1) - Df(y_2))\delta y| \leqslant C\,|y_1 - y_2|_\mu\,|\delta y|_\mu$ for all y_1, $y_2 \in \mathcal{D}$ and all $\delta y \in \mathbb{R}^{2n}$.

(H5) There exist positive constants c_h, C_h, c_j, C_j, and functions $\mathrm{corr}(\mu)$ and $\mathrm{d}(\mu) \geqslant \mathrm{corr}(\mu)$ such that for all $\mu \in (\mu(t))_{t\geqslant 0}$ and $\bar{y} = \bar{y}(\mu)$ if we define $\mathcal{D}_\mu = \{y \in \mathbb{R}^{2n} : |y - \bar{y}|_\mu \leqslant \mathrm{d}(\mu)\}$, then

(i) $h_\mu \in LC^2_\mu(\mathcal{D}_\mu, \mathbb{R}; C_h/\mathrm{corr}(\mu))$, $X_{h_\mu} \in LC^0_\mu(\mathcal{D}_\mu, \mathbb{R}^{2n})$, $\hat{p} \in LC^0_\mu(\mathcal{D}_\mu, \mathbb{R}^{2n})$, $j \in LC^1_\mu(\mathcal{D}_\mu, \mathfrak{g}^*)$, and $X_{j_\xi} \in LC^0_\mu(\mathcal{D}_\mu, \mathbb{R}^{2n}; C_j\,|\xi|)$ for all $\xi \in \mathfrak{g}$.

(ii) $c_h|\delta y|^2_\mu \leqslant D^2h_\mu(\bar{y})(\delta y, \delta y) \leqslant C_h|\delta y|^2_\mu$ for all $\delta y \in s(\bar{y})$.

(iii) For all $\xi \in \mathfrak{g}$ and $y \in \mathcal{D}_\mu \cap j^{-1}(\mu)$

(a) $c_j\,\mathrm{corr}(\mu)\,|\nabla_\mu j_\xi(y)|_\mu \leqslant |D\mathrm{j}(\hat{y})\nabla_\mu j_\xi(y)| \leqslant C_j\,\mathrm{corr}(\mu)\,|\nabla_\mu j_\xi(y)|_\mu$

(b) $c_j\,\mathrm{corr}(\mu)\,|\xi| \leqslant |X_{j_\xi}(y)|_\mu$.

If hypothesis (H5) holds, then we can uniformly estimate the Lyapunov function $L(u)$ in terms of the orbit distance ρ. In section 3, lemma 8, we shall show that there exists constants $0 < c \leqslant C$ such that for all u in a cone shaped neighbourhood of the MRE we have

$$c\,\max\{\rho(u, \bar{u}(\mu)), \rho(\bar{u}(\mu), u)\} \leqslant \sqrt{L(u)} \leqslant C\,\min\{\rho(u, \bar{u}(\mu)), \rho(\bar{u}(\mu), u)\} \quad \text{where} \quad \mu = J(u)\,.$$

2.3. *Influence of the perturbation on the relative equilibria*

We continue with the hypotheses on the dissipative behaviour of the perturbation with respect to the manifold of relative equilibria. We are not interested in motions along the G_μ-orbits, hence the part of the perturbation that causes such motions is not relevant for our purposes. In other words, we are only interested in $\hat{p}(y)$, the part of the perturbation which is $(,)_\mu$ orthogonal to $\mathfrak{g}_\mu \cdot y$.

First we consider the value of the perturbation at the MRE. If the component of the perturbation orthogonal to $\mathfrak{g}_\mu \cdot \bar{y}$, known as the residual, equals zero, then the perturbation at the MRE is a tangent to the MRE and the MRE is an invariant manifold for the perturbed system, as well as for the unperturbed Hamiltonian system. However, in general $\mathrm{res}(\bar{y}, \varepsilon) = \mathbb{P}_{\bar{y}}(\varepsilon\, p(\bar{y}) - \dot{\bar{y}}) \neq 0$; the residual acts as a forcing on the evolution of the solution curve starting at the relative equilibrium $\bar{u}$, causing it to leave the MRE.

Hypothesis (H6) controls the strength of the forcing taking trajectories away from the MRE. To formulate this hypothesis, we first introduce some additional notation. For each $\mu \in \mathfrak{g}^*_{\mathrm{MRE}}$, define the co-residual at $\bar{y} = \bar{y}(\mu)$ by

$$r(\bar{y}, \varepsilon) = B_\mu^{-1}\,\gamma(\bar{y})^{-1}\,\mathrm{res}(\bar{y}, \varepsilon)\,.$$

Now we can formulate the hypothesis.

(H6) For any initial condition $u(0)$ in an order ε B_μ-neighbourhood of the MRE (i.e. $\rho(u(0), \bar{u}(\mu(0))) = \mathcal{O}(\varepsilon)$), the solution $u(t)$ of the differential equation (2) has associated curves $\mu(t)$, $\bar{u}(t) = \bar{u}(\mu(t))$, and $\bar{y}(t) = \varphi(\bar{u}(t))$ of momenta and relative equilibria, for which the functions $|\hat{p}(\bar{y})|_\mu$, $|\mathrm{res}(\bar{y}, \varepsilon)|_\mu$, and $|r(\bar{y}, \varepsilon)|$ are integrable. To be specific, we assume the existence of an $\alpha \geqslant 0$ such that $\mathrm{e}^{-\alpha\varepsilon t} = \mathcal{O}(\mathrm{corr}(\mu(t)))$ (i.e. $\alpha\varepsilon t + \log(\mathrm{corr}(\mu(t)))$ is bounded below for all $t \geqslant 0$) and an integrable function k (meaning $\int_0^\infty k(\tau)\,\mathrm{d}\tau$ is finite) such that $|\hat{p}(\bar{y})|_\mu \leqslant \varepsilon\,\mathrm{e}^{-\alpha\varepsilon t}$ and $\max\{|\mathrm{res}(\bar{y}, \varepsilon)|_\mu, |r(\bar{y}, \varepsilon)|\} \leqslant \varepsilon\, k(\varepsilon t)\,\mathrm{e}^{-\alpha\varepsilon t}$ for all $\varepsilon > 0$ and $t \geqslant 0$. This implies that $|\mathbb{P}_{\bar{y}}\dot{\bar{y}}|_\mu \leqslant \varepsilon\,\mathrm{e}^{-\alpha\varepsilon t}$.

Next we focus on the behaviour of the perturbation near the relative equilibria. This behaviour has to be dissipative to compensate for the forcing at the relative equilibria. We want to construct a modification $\tilde{H}_\mu(\cdot, \varepsilon)$ of the energy–momentum function, with critical points $\tilde{u}(\mu, \varepsilon)$ near $\bar{y}(\mu)$, such that $\tilde{L}(u, \varepsilon) = \tilde{H}_\mu(u, \varepsilon) - \tilde{H}_\mu(\tilde{u}(\mu, \varepsilon), \varepsilon)$ bounds the distance between u and the G_μ orbit of $\tilde{u}(\mu, \varepsilon)$ and satisfies the dissipation relation $\mathrm{d}/\mathrm{d}t\, \tilde{L}(u, \varepsilon) \leqslant -2\varepsilon\beta\tilde{L}(u, \varepsilon) +$ 'small terms' for some positive constant β. The remaining two hypotheses guarantee the existence of such a function.

Hypothesis (H7) takes care of the dissipative character of the perturbation. We present two different versions of this hypothesis. The first version (H7) is more general, but may require more work to verify in applications; the second version (H7A) is simpler and more intuitive, but requires additional smoothness of the energy–momentum function and the dissipative perturbation with respect to the $||_\mu$ norm.

(H7) There exists a positive constant β and an integrable function κ such that $\mu = \mu(t)$ satisfies

$$(Dh_\mu(y) - Dh_\mu(\check{y}))(\hat{p}(y) - \frac{1}{\varepsilon}\mathrm{res}(\bar{y}, \varepsilon)) \leqslant -2\beta(h_\mu(y) - h_\mu(\check{y}) - Dh_\mu(\check{y})(y - \check{y})) + \kappa(\varepsilon t)(\varepsilon \mathrm{e}^{-\alpha\varepsilon t} + |y - \check{y}|_\mu)|y - \check{y}|_\mu$$

for all $y, \check{y} \in j^{-1}(\mu)$ satisfying $|y - \bar{y}|_\mu = \mathcal{O}(|r(\bar{y}, \varepsilon)|)$, $|\check{y} - \bar{y}|_\mu = \mathcal{O}(|r(\bar{y}, \varepsilon)|)$.

This hypothesis assures that $\varepsilon\hat{p}(y) - \mathrm{res}(\bar{y}, \varepsilon)$ acts dissipatively on a neighbourhood of $\bar{y}$ in $\varphi(\mathcal{M} \cap \mathcal{U})$. This is what we mean when we say that the perturbation is G_μ-orbit-dissipative with respect to the relative equilibrium $\bar{u}(\mu)$ for variations tangential to the level set $\mathcal{M}_\mu$. This effect of the perturbation drives a solution back to an ε neighbourhood of the MRE.

If the second derivative of h_μ and the first derivative of $\hat{p}$ are uniformly Lipschitz, then we can give a more intuitive approach to the dissipative character of the perturbation, leading to the alternative hypothesis (H7A). We define for all $\mu \in \mathfrak{g}^*_{\mathrm{MRE}}$ the dissipation coefficient $\beta(\mu)$ and the tangential dissipation coefficient $\beta_{\mathrm{T}}(\mu)$ at the relative equilibrium $\bar{u}(\mu)$. The dissipation coefficient $\beta(\mu)$ is the largest number β such that for all $\delta y \in \mathbb{R}^{2n}$

$$D^2h_\mu(\bar{y}(\mu))\,(\delta y, D\hat{p}(\bar{y}(\mu))\,\delta y) \leqslant -\beta\, D^2h_\mu(\bar{y}(\mu))\,(\delta y, \delta y).$$

The tangential dissipation coefficient $\beta_{\mathrm{T}}(\mu)$ is

$$\beta_{\mathrm{T}}(\mu) = -\max_{\delta y \in s(\bar{y}(\mu))} \frac{D^2h_\mu(\bar{y}(\mu))\,(\delta y, D\hat{p}(\bar{y}(\mu))\,\delta y)}{D^2h_\mu(\bar{y}(\mu))\,(\delta y, \delta y)}.$$

To explain the term dissipation coefficient, we consider the case in which there exists some curve $g(t) \in G$ such that $g(t) \cdot \bar{u}(\mu)$ is a solution curve of the perturbed system (1) and $DL_{g(t)^{-1}}\dot{g}(t) = \bar{\xi}(\mu(t)) - \varepsilon\,\eta(t)$, with $\eta(t) \in \mathfrak{g}_{\mu(t)}$. (Here L_g denotes left translation by g.) This implies that $\bar{y}(\mu)$ is a solution curve of the time-dependent vector field

$$X_{h_\mu} + \varepsilon\,[p - X_{j_{\eta(t)}}]. \tag{3}$$

If the submanifold of relative equilibria which are equilibria of the vector field (3) is strongly attractive, then the eigenvalues of the linearization of the vector field (3) at $\bar{u}(\mu)$ have a negative real part, except for the zero eigenvalues in the direction of the infinitesimal generators of the Lie algebra $\mathfrak{g}_\mu$. Because $\hat{p}(\bar{y}(\mu)) = \mathbb{P}_{\bar{y}}(p(\bar{y}(\mu)) - X_{j_\eta}(\bar{y}(\mu)))$, a sufficient condition for this property is that the dissipation coefficient $\beta(\mu)$ is positive. The dissipation coefficient $-\beta(\mu)$ is always larger than or equal to the largest real part of the relevant eigenvalues of the linearization. If $-\beta(\mu)$ is equal to this largest real part, then $\mathrm{e}^{-\varepsilon\beta(\mu)t}$ is a sharp estimate for the attractive behaviour of the G_μ-orbit of $\bar{u}(\mu)$. Even if $\bar{u}(\mu)$ is not an equilibrium of (3) for any $\eta \in \mathfrak{g}_\mu$, the dissipation coefficient still measures the dissipative part of the perturbation. For more information about this aspect of the analysis, see Derks [3].

Because the evolution along the level sets is described by the μ equation, it would seem that only the tangential dissipation coefficient $\beta_{\mathrm{T}}(\mu)$ is relevant. However, to avoid possible problems with higher order terms we sometimes use the full dissipation coefficient $\beta(\mu)$. In hypothesis (H7A) we set restrictions on the behaviour of the dissipation coefficients in both cases. Define $\beta_{\mathrm{T}} = \lim_{t\to\infty} \beta_{\mathrm{T}}(\mu(t))$. If $\mathrm{e}^{-\min(\alpha,\beta_{\mathrm{T}})\tau}/\mathrm{corr}(\mu(\tau/\varepsilon))$ is an integrable function, then set $\beta = \beta_{\mathrm{T}}$ and define the function

$$b(\tau) = |\beta_{\mathrm{T}}(\mu(\tau/\varepsilon)) - \beta| + \mathrm{e}^{-\min(\alpha,\beta_{\mathrm{T}})\tau}/\mathrm{corr}(\mu(\tau/\varepsilon)) .$$

Otherwise, set $\beta = \lim_{t\to\infty} \beta(\mu(t))$ and $b(\tau) = |\beta(\mu(\tau/\varepsilon)) - \beta|$.

(H7A)

(i) The asymptotic dissipation coefficient β defined above is positive for every $\mu \in \mathfrak{g}^*_{\mathrm{MRE}}$. The function b is integrable with respect to τ and hence can be bounded by an ε-independent constant. Furthermore, $\varepsilon\beta t + \log(\mathrm{corr}(\mu(t)))$ is bounded below for all $t \geqslant 0$.

(ii) $h_\mu \in LC^2_\mu(\mathcal{D}_\mu, \mathbb{R})$, i.e. $D^2 h_\mu$ has Lipschitz constant of order one, and $\hat{p} \in LC^1_\mu(\mathcal{D}_\mu, \mathbb{R}^{2n})$.

In section 3, we will show that hypothesis (H7A) implies hypothesis (H7).

In general, we can expect competition between a forcing which drives the solution away from the MRE and a dissipation which attracts the solution towards the MRE. The hypothesis (H6) guarantees that the forcing does not dominate this competition.

Our last hypothesis is quite technical. To verify the approximation with the relative equilibria, we have to use a better approximation of the solutions of the perturbed system than our original curve of relative equilibria. In order to do this, we use a μ-dependent function on $\mathcal{Y}$ with the property that its Hamiltonian vector field at $\bar{y}$ is approximately equal to the residual. The hypothesis (H8) asserts the existence of such a function and specifies some of its behaviour.

(H8) For $\mu \in (\mu(t))_{t\geqslant 0}$, $\varepsilon > 0$, there exists a full tubular neighbourhood $\mathcal{V}$ of the MRE and a smooth function $F_\mu(\cdot, \varepsilon) : \mathcal{V} \to \mathbb{R}$ that is G_μ invariant on $\mathcal{M}_\mu$. The push forward f_μ of F_μ by the chart map satisfies

(i) $|X_{f_\mu}(\bar{y}(\mu), \varepsilon) - \frac{1}{\varepsilon}\mathrm{res}(\bar{y}(\mu), \varepsilon)|_\mu = \mathcal{O}(|r(\bar{y}, \varepsilon)|)$,

(ii) There exists an integrable function k_f such that for all $\mu = \mu(t)$

(a) $f_\mu \in LC^1_\mu(\mathcal{D}_\mu, \mathbb{R}; \mathcal{O}(k_f(\varepsilon t)))$.

(b) $X_{f_\mu} \in LC^0_\mu(\mathcal{D}_\mu, \mathbb{R}^{2n}; \mathcal{O}(\min(\mathrm{e}^{-\varepsilon\alpha t}/\mathrm{corr}(\mu), k_f(\varepsilon t))))$.

(c) $\dot{f}_\mu = D_\mu f_\mu \dot{\mu} \in LC^0_\mu(\mathcal{D}_\mu, \mathbb{R}; \mathcal{O}(\varepsilon|r(\bar{y}(\mu), \varepsilon)|))$.

Remark 2. The last four hypotheses are expressed in terms of the chart maps φ. However, for our purposes the specific choice of the charts is irrelevant—if the hypotheses are satisfied in one set of charts, they will be satisfied in any other charts.

2.4. The result

After stating the hypotheses on the system, we formulate the result that we will prove in the following sections.

Theorem 4. *Let $u(t)$ be a solution of the dissipative Hamiltonian system (1) and define $\mu(t) = J(u(t))$. Let hypotheses 1–8 be satisfied. If $u(0)$ is close to the set $\mathrm{MRE}_{\mu(0)}$ of* *relative equilibria in the level set* $\mathcal{M}_{\mu(0)}$, *then for all* $t \geqslant 0$, $u(t)$ *stays close to the set* $\mathrm{MRE}_{\mu(t)}$ *of relative equilibria in the level set* $\mathcal{M}_{\mu(t)}$. *To be precise, if* ε *is sufficiently small and*

$$\rho(u(0), \bar{u}(\mu(0))) = \mathcal{O}(\varepsilon) \qquad \text{then} \qquad \rho(u(t), \bar{u}(\mu(t))) = \mathcal{O}(\varepsilon\, \mathrm{e}^{-\min(\alpha,\beta)\varepsilon t})$$

for all $t \geqslant 0$. *An equivalent formulation is*

$$L(u(0)) = \mathcal{O}(\varepsilon^2) \Rightarrow L(u(t)) = \mathcal{O}(\varepsilon^2\, \mathrm{e}^{-2\min(\alpha,\beta)\varepsilon t}) \qquad t \geqslant 0\,.$$

Remark 3. If we also know that $L(u(t))$ decays exponentially to 0 for $t \to \infty$ (which is suggested by the hypothesis (H7)), the result seems trivial. However, this is not true. A counterexample is the function $L(t) = (L(0) + t^k)\, \mathrm{e}^{-\varepsilon(\beta+\delta)t}$, with $\delta, k > 0$. Indeed, $L(t) = \mathcal{O}(\mathrm{e}^{-\varepsilon\beta t})$ for all $t \geqslant 0$, but for $t = \varepsilon^{-1}$ we have $L(\varepsilon^{-1}) = [\,L(0) + \varepsilon^{-k}\,]\,\mathrm{e}^{-\beta+\delta}$, hence it is of order ε^{-k} instead of order ε!

To prove theorem 4, we derive an estimate for the dynamical behaviour of the distance function $L(u(t))$. To do this, we first have to make a local approximation in the charts to derive a relation between L and the distance functions in the charts.

3. A local approximation in the charts

To estimate the evolution of the Lyapunov function L, we work in the charts. In general the solution $u(t)$ itself need not be an element of $\mathcal{U}$, even if $L(u)$ is very small. It is only true that if $L(u)$ is small enough, there exist $g \in G_\mu$ such that $g \cdot u \in \mathcal{U}$. Therefore we first have to let G_μ act on the solution and then make a transformation to the charts. We will do this in a specific way to gain some additional properties.

Conceptually, we work modulo the current momentum isotropy subgroup G_μ. We do so, not by working directly on the (varying) quotient spaces, but by defining appropriate representatives of the orbits and estimating the distances between those representatives. In particular, we can map the solution $u(t)$ onto the charts, if $u(t)$ is sufficiently close to the MRE, and use invariant functions to estimate orbital distances.

We note that the closure of the set $(\bar{u}(\mu(t)))_{t\geqslant 0}$ is compact. Hence there exists a neighbourhood $\hat{\mathcal{U}}$ around this set and some $\delta_1 > 0$ such that for all $u_* \in \hat{\mathcal{U}}$ and for all $u \in \mathcal{M}$ with $J(u) = J(u_*)$

$$\max\{\tilde{d}(u_*, u), \tilde{d}(u, u_*)\} < \delta_1 \Rightarrow u \in \mathcal{U}\,.$$

(See also definition 3.)

Lemma 5. Let $u_* \in \hat{\mathcal{U}}$ and set $y_* = \varphi(u_*)$, $\mu = J(u_*)$. For all $u \in \mathcal{M}_\mu$ close enough to u_*, i.e. $\rho(u, u_*) < \delta_1 \min\{1, 1/C_d\}$, there exists $y(u) \in \mathcal{Y}$ with the following properties:

(i) There exists some $g \in G_\mu$ such that $g \cdot u \in \mathcal{U}$ and $y = \varphi(g \cdot u)$.
(ii) $(y - y_*, X_{j_\eta}(y))_\mu = 0$ for all $\eta \in \mathfrak{g}_\mu$.
(iii) $c_d\, \rho(u, u_*) \leqslant |y - y_*|_\mu \leqslant C_d\, \rho(u, u_*)$.

Proof. Let $u \in \mathcal{M}_\mu$, with $\rho(u, u_*) < \delta_1 \min\{1, 1/C_d\}$; there exists some $g_* \in G_\mu$ such that $|\varphi(g_* \cdot u) - y_*|_\mu < \delta_1$. Define

$$\mathcal{Y}_*(u) = \varphi((G_\mu \cdot u) \cap \tilde{B}_{\delta_1}(u_*)) \cap B^\mu_{\delta_1}(y_*)$$

where $\tilde{B}_{\delta_1}(u_*))$ is the $\tilde{d}$-sphere around u_* and $B^\mu_{\delta_1}(y_*)$ the $|\,|_\mu$-sphere around y_* with radius δ_1. The set $\mathcal{Y}_*(u)$ is compact; the condition $\rho(u, u_*) < \delta_1 \min\{1, 1/C_d\}$ implies that it is non-empty. Hence there is an element $y(u)$ of $\mathcal{Y}_*(u)$ satisfying

$$|y(u) - y_*|_\mu \doteq \min_{y\in\mathcal{Y}_*(u)} |y - y_*|_\mu\,.$$

If $y(u)$ is at the boundary of $\mathcal{Y}_*(u)$, then either $\tilde{d}(\varphi^{-1}(y(u)), u_*) = \delta_1$ or $|y(u) - y_*|_\mu = \delta_1$ and the condition $|y(u) - y_*|_\mu \leqslant C_d\, \rho(u, u_*) < \delta_1$ is violated. Hence $y(u)$ is in the interior of $\mathcal{Y}_*(u)$ and $y(u)$ is a critical point of the function $\tilde{f} : \mathcal{Y}_*(u) \to \mathbb{R}$ given by $\tilde{f}(y) = \frac{1}{2}|y - y_*|_\mu^2$. For any $\eta \in \mathfrak{g}_\mu$ and θ sufficiently small, the curve $v(\theta) = \varphi(\exp(\theta\,\eta) \cdot \varphi^{-1}(y(u)))$ lies in $\mathcal{Y}_*(u)$. Thus

$$0 = D\tilde{f}(y(u)) \cdot v'(0) = (y(u) - y_*, X_{j_\eta}(y(u)))_\mu .$$

There exists some element $\check{u}$ of the G_μ orbit of u such that $\tilde{d}(\check{u}, u_*) = \rho(u, u_*)$. If we set $\hat{u} = \varphi^{-1}(y(u))$, then

$$c_d\, \tilde{d}(\check{u}, u_*) \leqslant c_d\, \tilde{d}(\hat{u}, u_*) \leqslant |y(u) - y_*|_\mu \leqslant |\varphi(\check{u}) - y_*|_\mu \leqslant C_d\, \tilde{d}(\check{u}, u_*)$$

and hence $c_d\, \rho(u, u_*) \leqslant |y(u) - y_*|_\mu \leqslant C_d\, \rho(u, u_*)$. □

The main part of the vector $(y(u) - y_*)$ is in the space $s(y_*)$, i.e. if $\Pi_{y_*} : \mathbb{R}^{2n} \to s(y_*)$ denotes the projection onto $s(y_*)$, then $|(1 - \Pi_{y_*})(y(u) - y_*)|_\mu = \mathcal{O}(|\Pi_{y_*}(y(u) - y_*)|_\mu^2/\mathrm{corr}(\mu))$. This is a corollary of the following lemma.

Lemma 6. *Let $u_* \in \hat{\mathcal{U}}$, $\mu = J(u_*)$ and $y_* = \varphi(u_*)$. A $(\ ,\ )_\mu$ orthogonal decomposition for $\mathbb{R}^{2n}$ is given by*

$$\mathbb{R}^{2n} = \ker[D\jmath(y_*)] \oplus r(y_*) = s(y_*) \oplus \mathfrak{g}_\mu \cdot y_* \oplus r(y_*)$$

where $r(y)$ is defined for any $y \in \mathcal{Y}$ by $r(y) = \{\nabla_\mu j_\xi(y) \mid \xi \in \mathfrak{g}\}$. In other words, for any $y \in \mathcal{Y}$ we can write

$$y - y_* = \delta y + X_{j_\eta}(y_*) + \nabla_\mu j_\xi(y_*) \tag{4}$$

for unique $\delta y \in s(y_*)$, $\eta \in \mathfrak{g}_\mu$ and $\xi \in \mathfrak{g}$.

If $\jmath(y) = \mu$ and $y_* \in \mathcal{D}_\mu$, then

$$|\nabla_\mu j_\xi(y_*)|_\mu = \mathcal{O}(|y - y_*|_\mu^2/\mathrm{corr}(\mu)) = \mathcal{O}((|\delta y|_\mu^2 + |X_{j_\eta}(y_*)|_\mu^2)/\mathrm{corr}(\mu)) . \tag{5}$$

If $u \in \mathcal{U} \cap \mathcal{M}_\mu$ with $\rho(u, u_*) < \delta_1 \min\{1, 1/C_d\}$, $y_* \in \mathcal{D}_\mu$, and $y(u)$ is given by lemma 5, then

$$|(1 - \Pi_{y_*})(y(u) - y_*)|_\mu = \mathcal{O}(|\Pi_{y_*}(y(u) - y_*)|_\mu^2/\mathrm{corr}(\mu)) . \tag{6}$$

Proof. The identity $(\nabla_\mu j_\xi(y_*), \delta y)_\mu = \langle D\jmath(y_*)\delta y, \xi\rangle = 0$ for any $\xi \in \mathfrak{g}$ and $\delta y \in \ker[D\jmath(y_*)]$ implies that $\ker[D\jmath(y_*)]$ and $\{\nabla_\mu j_\xi(y_*) \mid \xi \in \mathfrak{g}\}$ are orthogonal with respect to the inner product $)_\mu$. Because $\dim\{\nabla_\mu j_\xi(y_*) \mid \xi \in \mathfrak{g}\} = \dim \mathrm{Range}(D\jmath(y_*))$, this shows that the first decomposition of the lemma holds. The other decomposition is an immediate consequence of the definitions of $s(y_*)$ and the inner product $(\ ,\ )_\mu$.

Furthermore, hypothesis (H5.1) implies that $|\jmath(y) - \jmath(y_*) - D\jmath(y_*)(y - y_*)| = \mathcal{O}\left(|y - y_*|^2\right)$. If $y \in \mathcal{Y}$, with $\jmath(y) = \mu$, then using (H5.3) and (4) we can see that

$$c_j\, \mathrm{corr}(\mu)\, |\nabla_\mu j_\xi(y_*)|_\mu \leqslant |D\jmath(y_*)\, \nabla_\mu j_\xi(y_*)| = \mathcal{O}\left(|y - y_*|_\mu^2\right) .$$

Let $u \in \mathcal{U} \cap \mathcal{M}_\mu$ with $\rho(u, u_*) < \delta_1 \min\{1, 1/C_d\}$ and define $y = y(u)$. Lemma 5.2 implies that $(y - y_*, X_{j_\zeta}(y_*))_\mu = \mathcal{O}(|y - y_*|_\mu^2 |\zeta|)$, for any $\zeta \in \mathfrak{g}_\mu$. Using the decomposition (4) for $(y - y_*)$ and choosing $\zeta = \eta$, it follows that $|X_{j_\eta}(y_*)|_\mu^2 = \mathcal{O}(|y - y_*|_\mu^2 |\eta|)$. Hypothesis (H5.3) implies that the algebra element η satisfies $|\eta| = \mathcal{O}(|X_{j_\eta}(y_*)|_\mu/\mathrm{corr}(\mu))$; hence $|X_{j_\eta}(y_*)|_\mu = \mathcal{O}(|y - y_*|_\mu^2/\mathrm{corr}(\mu))$. Combining this estimate with the estimate (5), we obtain $|\nabla j_\xi(y_*)|_\mu = \mathcal{O}(|y - y_*|_\mu^2/\mathrm{corr}(\mu)) = \mathcal{O}(|\Pi_{y_*}(y - y_*)|_\mu^2/\mathrm{corr}(\mu))$. □

We are now able to show that hypothesis (H7A) can replace (H7).

Lemma 7. *If hypotheses (H1)–(H6) and (H7A) hold, then (H7) holds as well.*

Proof. Let $y, \check{y} \in j^{-1}(\mu)$ satisfying $|y - \bar{y}|_\mu = \mathcal{O}(|r(\bar{y}, \varepsilon)|)$, $|\check{y} - \bar{y}|_\mu = \mathcal{O}(|r(\bar{y}, \varepsilon)|)$. We write $\delta y = y - \check{y}$, then $|\delta y|_\mu = \mathcal{O}(|r(\bar{y}, \varepsilon)|)$. Hypothesis (H5) implies that

$$\begin{aligned}
(Dh_\mu(y) - Dh_\mu(\check{y}))&\left(\hat{p}(y) - \hat{p}(\bar{y}) + \frac{1}{\varepsilon}\dot{\bar{y}}\right) \\
&= (Dh_\mu(y) - Dh_\mu(\check{y}))\left(D\hat{p}(\bar{y})\,(y - \bar{y}) + \frac{1}{\varepsilon}\dot{\bar{y}}\right) + \mathcal{O}(|\delta y|_\mu |y - \bar{y}|_\mu^2) \\
&= D^2 h_\mu(\bar{y})\left(D\hat{p}(\bar{y})\,\delta y + \frac{1}{\varepsilon}\dot{\bar{y}}, \delta y\right) + \mathcal{O}((|\delta y|_\mu + |\check{y} - \bar{y}|_\mu)|\delta y|_\mu).
\end{aligned}$$

Differentiating the relation $Dh_\mu(\bar{y}(\mu)) = 0$ with respect to t and then applying lemma 6 yields

$$|D^2 h_\mu(\bar{y})(\delta y, \dot{\bar{y}})| = |Dj_{\dot{\xi}}(\bar{y})\,\delta y| = \mathcal{O}\left(|\dot{\xi}|\,|\delta y|_\mu^2\right)$$

since $j(y) = j(\check{y})$ implies that $\delta y = y - \check{y}$ is mainly tangent to the momentum level set. Similarly,

$$D^2 h_\mu(\bar{y})(\delta y, D\hat{p}(\bar{y})\delta y) \leqslant -\beta_T(\mu)\, D^2 h_\mu(\bar{y})(\delta y, \delta y) + \mathcal{O}(|\delta y|_\mu^3/\mathrm{corr}(\mu)).$$

If $\mathrm{e}^{-\min(\alpha, \beta_T(\mu))\varepsilon t}/\mathrm{corr}(\mu)$ is integrable, then we will use this estimate, since $|\delta y|_\mu = \mathcal{O}(\mathrm{e}^{-\min(\alpha,\beta)\varepsilon t})$. Otherwise we use the estimate

$$D^2 h_\mu(\bar{y})(\delta y, D\hat{p}(\bar{y})\delta y) \leqslant -\beta(\mu)\, D^2 h_\mu(\bar{y})(\delta y, \delta y).$$

In either case,

$$\begin{aligned}
D^2 h_\mu(\bar{y})(\delta y, D\hat{p}(\bar{y})\delta y) &\leqslant -\beta\, D^2 h_\mu(\bar{y})(\delta y, \delta y) + b(\mu)\mathcal{O}(|\delta y|_\mu^2) \\
&= -2\beta\,(h_\mu(y) - h_\mu(\check{y}) - Dh_\mu(\check{y})\delta y) + \kappa(\mu)|\delta y|_\mu^2
\end{aligned}$$

for some $\kappa(\mu) = \mathcal{O}\big(|\delta y|_\mu + |\check{y} - \bar{y}|_\mu + |\dot{\xi}| + b(\mu)\big) = \mathcal{O}\big(|r(\bar{y}, \varepsilon)| + |\dot{\xi}| + b(\mu)\big)$, which is integrable. □

One of the consequences of the following lemma is that the Lyapunov function L acts as a measure for the G_μ-orbital distance to the relative equilibria. This measure is compatible on the momentum level sets with the measure induced by $|y(u) - \bar{y}(\mu)|_\mu^2$. We prove these facts in a more general setting so as to be able to state similar facts for another Lyapunov function in section 4.

Lemma 8. *Assume that there exists a neighbourhood $\mathcal{W}$ of the* MRE, *positive constants B_η and ε_0, and a compact set $\mathfrak{g}^*_{\mathcal{W}} \subset \mathfrak{g}^*$ such that for any $(\mu, \varepsilon) \in \mathfrak{g}^*_{\mathcal{W}} \times [0, \varepsilon_0)$ there exist differentiable functions N_μ^ε, points $u_\mu^\varepsilon \in \mathcal{W}_\mu \cap \hat{\mathcal{U}}$, where $\mathcal{W}_\mu = \mathcal{W} \cap \mathcal{M}_\mu$, and $\eta(\mu, \varepsilon) \in \mathfrak{g}$ satisfying*

(i) $D\tilde{H}_\mu^\varepsilon(u_\mu^\varepsilon) = DJ_{\eta(\mu,\varepsilon)}(u_\mu^\varepsilon)$, where $\tilde{H}_\mu^\varepsilon = H_\mu + \varepsilon N_\mu^\varepsilon$.
(ii) $y_\mu^\varepsilon = \varphi(u_\mu^\varepsilon) \in \hat{\mathcal{D}}_\mu$ and $|\eta(\mu, \varepsilon)| \leqslant \varepsilon B_\eta$.
(iii) The restriction of N_μ^ε to $\mathcal{W}_\mu$ is G_μ-invariant and $n_\mu^\varepsilon = N_\mu^\varepsilon \circ \varphi^{-1} \in LC_\mu^1(\varphi(\mathcal{W} \cap \mathcal{U}), \mathbb{R})$.

Then there exist μ-independent positive constants ε_1, δ, c and C such that for any $0 \leqslant \varepsilon < \varepsilon_1$, any $u \in \mathcal{W}_\mu$ satisfying $\rho(u, u_\mu^\varepsilon) < \delta\,\mathrm{corr}(\mu)$ also satisfies

$$c|y(u) - y_\mu^\varepsilon|_\mu^2 \leqslant \tilde{H}_\mu^\varepsilon(u) - \tilde{H}_\mu^\varepsilon(u_\mu^\varepsilon) \leqslant C|y(u) - y_\mu^\varepsilon|_\mu^2 \qquad (7)$$

where y is the map associated to y_μ^ε given by lemma 5.

Proof. Given $u \in \mathcal{W}_\mu$ with $\rho(u, u_\mu) < \delta_1 \min\{1, 1/C_d\}$, let $\Delta y = y(u) - y^\varepsilon_\mu = \delta y + X_{j_\eta}(y^\varepsilon_\mu) + \nabla_\mu j_\xi(y^\varepsilon_\mu)$ be the decomposition given by lemma 6. Lemma 6 and equations (5) and (6) imply that there exist positive constants δ_2, κ, and κ' such that $|\Delta y|_\mu < \delta_2 \operatorname{corr}(\mu)$ implies that $|\nabla_\mu j_\xi(y^\varepsilon_\mu)|_\mu \leqslant \kappa |\delta y|^2_\mu/\operatorname{corr}(\mu) \leqslant \kappa\delta_2|\delta y|_\mu$ and $|\Delta y|_\mu = \mathcal{O}(|\delta y|_\mu(1 + |\delta y|_\mu/\operatorname{corr}(\mu))) \leqslant \kappa'|\delta y|_\mu$. Then

$$|D^2 h_\mu(y^\varepsilon_\mu)(\Delta y, \Delta y) - D^2 h_\mu(y^\varepsilon_\mu)(\delta y, \delta y)| \leqslant \tilde{C}_h |\delta y|^2_\mu$$

where $\tilde{C}_h = C_h \kappa \delta_2 (2 + \kappa\delta_2)$, if $|\Delta y|_\mu < \delta_2 \operatorname{corr}(\mu)$. If we define $c_2 = (c_h - \tilde{C}_h)/(\kappa')^2$ and $C_2 = (C_h + \tilde{C}_h)/(\kappa')^2$, then $y^\varepsilon_\mu \in \hat{\mathcal{D}}_\mu$ and $|\Delta y|_\mu < \delta_2 \operatorname{corr}(\mu)$ imply

$$c_2 |\Delta y|^2_\mu \leqslant c_2 (\kappa')^2 |\delta y|^2_\mu \leqslant D^2 h_\mu(y^\varepsilon_\mu)(\Delta y, \Delta y) \leqslant C_2 |\Delta y|^2_\mu .$$

We assume that δ_2 has been chosen to be sufficiently small that c_2 is positive.

Define $\tilde{h}^\varepsilon_\mu = \tilde{H}^\varepsilon_\mu \circ \varphi^{-1}$. Because y^ε_μ is a critical point of $\tilde{h}^\varepsilon_\mu - j_{\eta(\mu,\varepsilon)}$, (H5.3) and $j(y(u)) = j(y^\varepsilon_\mu)$ imply that there exists $B_j > 0$ such that

$$|D\tilde{h}^\varepsilon_\mu(y^\varepsilon_\mu)\Delta y| = |D j_{\eta(\mu,\varepsilon)}(y^\varepsilon_\mu)\Delta y| \leqslant B_j |\eta(\mu, \varepsilon)| |\Delta y|^2_\mu .$$

Hence if $|\Delta y|_\mu \leqslant \delta_3 \operatorname{corr}(\mu)$ the estimate

$$\begin{aligned}
&\left|\tilde{h}^\varepsilon_\mu(y(u)) - \tilde{h}^\varepsilon_\mu(y^\varepsilon_\mu) - \tfrac{1}{2} D^2 h_\mu(y^\varepsilon_\mu)(\Delta y, \Delta y)\right| \\
&\qquad \leqslant \varepsilon |n^\varepsilon_\mu(y(u)) - n^\varepsilon_\mu(y^\varepsilon_\mu) - D n^\varepsilon_\mu(y^\varepsilon_\mu)\Delta y| + B_j |\eta(\mu, \varepsilon)| |\Delta y|^2_\mu + \tfrac{1}{2} C_h |\Delta y|^3_\mu/\operatorname{corr}(\mu) \\
&\qquad \leqslant \left(\varepsilon(B_j B_\eta + B_n) + \tfrac{1}{2} C_h \delta_3\right) |\Delta y|^2_\mu
\end{aligned}$$

with B_n some Lipschitz constant, follows from Lipschitz continuity of $D^2 h_\mu$ and $D n^\varepsilon_\mu$. By taking δ_3 and ε sufficiently small, we can guarantee that $c_2 > c_c = \delta_3 C_h + 2\varepsilon(B_j B_\eta + B_n)$. Hence, if we set $c = \frac{1}{2}(c_2 - c_c)$ and $C = \frac{1}{2}(C_2 + c_c)$, then (7) holds if $\rho(u, e^\varepsilon_\mu) < \delta \operatorname{corr}(\mu)$, where $\delta = \min\{1, \delta_1, \delta_2, \delta_3\} \min\{1, 1/c_d\}$. □

Lemmas 5 and 8 imply that there exist positive constants δ_0, c_0, and C_0 such that for $u \in \bigcup_{t \geqslant 0} \mathcal{M}_{\mu(t)}$ which satisfy $L(u) < \delta_0^2 \operatorname{corr}(\mu)^2$ the map $y(u)$ is well defined and, for $\mu = J(u)$,

$$c_0^2 \, |y(u) - \bar{y}(\mu)|^2_\mu \leqslant L(u) \leqslant C_0^2 \, |y(u) - \bar{y}(\mu)|^2_\mu . \tag{8}$$

This can be seen by setting $u_* = \bar{u}(\mu)$ in lemma 5 and using the compactness of the closure of the curve $(\mu(t))_{t\geqslant 0}$ to derive the existence of the map $y(u)$ in a uniform neighbourhood of the MRE. The estimates (8) follow from lemma 8 by taking $\mathcal{W} = \mathcal{M}$, $\mathfrak{g}^*_{\mathcal{W}} = \overline{(\mu(t))_{t\geqslant 0}}$, $N^\varepsilon_\mu = 0$, $\eta(\mu, \varepsilon) = 0$, and $u^\varepsilon_\mu = \bar{u}(\mu)$.

Unfortunately, applying these results directly to the Lyapunov function L does not yield sufficiently sharp bounds to accurately estimate the evolution of the perturbed trajectory. If $u(t)$ is a solution of the system (1), then the following estimate on the time derivative of L can be derived:

$$\frac{\mathrm{d}}{\mathrm{d}t} L(u(t)) \leqslant -2\varepsilon\,(\beta - \bar{C}\,[\,|\hat{p}(\bar{y})|_\mu + b(\varepsilon t)\,])\, L(u) + 2\varepsilon\,\bar{C}\,|\hat{p}(\bar{y})|_\mu \sqrt{L(u)}$$

for some constant $\bar{C}$ and $L(u) < \delta_0^2 \operatorname{corr}(\mu)^2$. Applying differential inequalities such as Gronwall's Lemma to this estimate will give

$$\sqrt{L(u(t))} \leqslant \tilde{C}\,\mathrm{e}^{-\varepsilon\beta t} \sqrt{L(u(0))} + \tilde{C}\,\mathrm{e}^{-\varepsilon\beta t} \int_0^t |\hat{p}(\bar{y}(\mu(\tau)))|_\mu \,\mathrm{e}^{\varepsilon\beta\tau}\,\mathrm{d}\tau$$

for some constant $\tilde{C}$, which does not imply that $L(u) = \mathcal{O}(\varepsilon)$ even if $L(u(0)) = 0$. Hence this straightforward estimate of $L(u(t))$ does not work. The residual at the relative equilibria is of order ε and this gives the order 1 term in the estimate. (See also Derks and van Groesen [4], Derks and Valkering [5], and Lebovitz and Neishtadt [9].) In order to avoid this technical problem, we have to use the curve $(\tilde{y}(\mu, \varepsilon))_{|\varepsilon|\leqslant\varepsilon_0(\mu)}$, which has a co-residual of order ε^2.

4. Estimate of the Lyapunov function

In the previous section we have seen that we need a curve $(\tilde{y}(\mu, \varepsilon))_{|\varepsilon|\leqslant\varepsilon_0(\mu)} \in \mathcal{Y}$ with a co-residual of order ε^2. This curve can be interpreted as determining an improved MRE for the dissipative equation, because the forcing at this new curve is smaller. To explain this, we first define the residual of an differential equation. Let $y(t)$ be a curve in the charts. The forcing on this curve is the residual of the differential equation at this curve, defined by

$$\operatorname{res}(y(t), \varepsilon) = \mathbb{P}_y(X_h(y) + \varepsilon\, p(y) - \dot{y})$$

where $\mathbb{P}_y$ denotes the $(\,,\,)_\mu$-orthogonal projection onto $s(y)\oplus r(y)$, as discussed in lemma 6. Intuitively, to find a curve $(\tilde{y}(\mu, \varepsilon))_{|\varepsilon|\leqslant\varepsilon_0(\mu)}$, we should solve an equation of the form

$$Dh_\mu(\tilde{y}) - Dj_\eta(\tilde{y}) + \gamma(\tilde{y})^{-1}\operatorname{res}(\tilde{y}, \varepsilon) = \mathcal{O}(\varepsilon^2)$$

for $\tilde{y} \in j^{-1}(\mu)$ and $\eta \in \mathfrak{g}$.

To establish the existence of the curve $\tilde{y}(\mu, \varepsilon)$, we apply the Implicit Function Theorem to the functions induced by $\tilde{H}_\mu = H_\mu + \varepsilon F_\mu$ on the symplectically reduced manifolds $\mathcal{P}_\mu = J^{-1}(\mu)/G_\mu$ associated to the momentum value μ and the tubular neighbourhood $\mathcal{V}$ of the MRE, because $\varepsilon Df_\mu(\bar{y}, \varepsilon) = \gamma(\bar{y})^{-1}\operatorname{res}(\bar{y}, \varepsilon) + \mathcal{O}(\varepsilon^2)$.

Lemma 9. *For all $\mu \in \mathfrak{g}_{MRE}$ there exist some $\varepsilon_0(\mu) > 0$ and a curve $(\tilde{y}(\mu, \varepsilon),$ $\eta(\mu, \varepsilon))_{|\varepsilon|\leqslant\varepsilon_0(\mu)} \in \mathcal{Y}\times\mathfrak{g}$ depending smoothly on μ that satisfies*

$$D(h_\mu + \varepsilon f_\mu - j_\eta)(\tilde{y}) = 0 \qquad andtqsj(\tilde{y}) = \mu\,.$$

Furthermore,

$$|\tilde{y} - \bar{y}|_\mu = \mathcal{O}(|r(\bar{y}, \varepsilon)|) = O(\varepsilon k(\varepsilon t)\,\mathrm{e}^{-\varepsilon\alpha t}) \qquad \text{and} \qquad |\eta| = \mathcal{O}(\varepsilon k(\varepsilon t))\,.$$

Proof. Since both H_μ and F_μ are G_μ invariant on $\mathcal{M}_\mu$, they determine functions $\hat{H}_\mu$ and $\hat{F}_\mu$ on $\mathcal{P}_\mu$. Hypothesis (H3) implies that $D^2\hat{H}_\mu([\bar{u}(\mu)])$ is positive definite, since the tangent space of $\mathcal{P}_\mu$ at the equivalence class $[\bar{u}(\mu)]$ is isomorphic to $\ker[DJ(\bar{u}(\mu))]$ modulo $\mathfrak{g}_\mu\cdot\bar{u}(\mu)$. Hence the implicit function theorem implies that for sufficiently small ε, there exists a curve $[\tilde{u}(\mu, \varepsilon)]$ satisfying $D(\hat{H}_\mu + \varepsilon\hat{F}_\mu)([\tilde{u}(\mu, \varepsilon)]) = 0$.

We now use lemma 5 to map the equivalence class $[\tilde{u}(\mu, \varepsilon)]$ onto an element of $\mathcal{V}$ near $\bar{u}(\mu)$. Specifically, if we set $y_* = \bar{y}(\mu)$, then the map y given in lemma 5 determines a map $\hat{y}$ on a neighbourhood of $[\bar{y}(\mu)]$ in $\mathcal{P}_\mu$, since $y(u) = y(g\cdot u)$ if $u \in \mathcal{M}_\mu$ and $g \in G_\mu$. We can choose a representative $\tilde{u}(\varepsilon)$ for each equivalence class $[\tilde{u}(\mu, \varepsilon)]$ by defining $\tilde{y}(\mu, \varepsilon) = \hat{y}([\tilde{u}(\mu, \varepsilon)])$ and $\tilde{u}(\mu, \varepsilon) = \varphi^{-1}(\tilde{y}(\mu, \varepsilon))$. The implicit function theorem implies that the distance between $[\bar{u}(\mu)]$ and $[\tilde{u}(\mu, \varepsilon)]$ is of order $|r(\bar{y}, \varepsilon)|$; hence the bound

$$\begin{aligned} c_0|\tilde{y}(\mu, \varepsilon) - \bar{y}(\mu)|_\mu^2 &= H_\mu(\tilde{u}(\mu, \varepsilon)) - H_\mu(\bar{u}(\mu)) = \hat{H}_\mu([\tilde{u}(\mu, \varepsilon)]) - \hat{H}_\mu([\bar{u}(\mu)]) \\ &= \mathcal{O}(|r(\bar{y}, \varepsilon)|^2) \end{aligned}$$

given by lemma 8 for some constant c_0 implies that $|\tilde{y}(\mu, \varepsilon) - \bar{y}(\mu)|_\mu = \mathcal{O}(|r(\bar{y}, \varepsilon)|)$.

The Lagrange multiplier theorem implies that there exists a curve $\eta(\mu,\varepsilon) \in \mathfrak{g}$ such that $\tilde{u}(\mu,\varepsilon)$ is a critical point of $\tilde{H}_\mu + J_{\eta(\mu,\varepsilon)}$ and hence $\tilde{y}(\mu,\varepsilon)$ is a critical point of $\tilde{h}_\mu + j_{\eta(\mu,\varepsilon)}$. Hypotheses (H5.2) and (H8) imply that

$$c_j\,|\eta|\,\mathrm{corr}(\mu) \leqslant |X_{j_\eta}(\tilde{y})|_\mu = |X_{\tilde{h}_\mu}(\tilde{y})|_\mu \leqslant |X_{\tilde{h}_\mu}(\tilde{y}) - X_{\bar{h}_\mu}(\bar{y})|_\mu + \varepsilon |X_{f_\mu}(\tilde{y},\varepsilon)|_\mu$$

and hence

$$|\eta| = \mathcal{O}((|\tilde{y}-\bar{y}|_\mu + \varepsilon(|\mathrm{res}(\bar{y},\varepsilon)|_\mu + |r(\bar{y},\varepsilon)|))/\mathrm{corr}(\mu)) = \mathcal{O}(\varepsilon k(\varepsilon t))$$

since $|\tilde{y}-\bar{y}|_\mu = \mathcal{O}(|r(\bar{y},\varepsilon)|) = \mathcal{O}(\varepsilon k(\varepsilon t)\mathrm{e}^{-\varepsilon\alpha t})$ and $\mathrm{e}^{-\varepsilon\alpha t} = \mathcal{O}(\mathrm{corr}(\mu))$. □

To derive estimates for the distance between a solution and the approximation $\tilde{u}(\mu,\varepsilon)$, we define a new Lyapunov function centred around $\tilde{u}(\mu,\varepsilon)$.

Definition 10. *Define the G_μ-invariant Lyapunov function $\tilde{L}$ as*

$$\tilde{L}(u,\varepsilon) = \tilde{H}_\mu(u,\varepsilon) - \tilde{H}_\mu(\tilde{u}(\mu,\varepsilon),\varepsilon) \tag{9}$$

with $\mu = J(u)$.

The new Lyapunov function $\tilde{L}(u)$ determines a G_μ-orbit distance function.

Lemma 11. *Let $\hat{y}$ be the map given in lemma 5 for $u_* = \tilde{u}(\mu,\varepsilon)$. There exist constants $0 < \tilde{\delta}$, $0 < \tilde{c} \leqslant \tilde{C} < \infty$ and $0 < \check{c} \leqslant \check{C} < \infty$, independent of ε and u, such that*

$$\tilde{c}\,\rho(u,\tilde{u}(\mu,\varepsilon))^2 \leqslant \check{c}\,|\hat{y}(u,\varepsilon) - \tilde{y}(\mu,\varepsilon)|_\mu^2 \leqslant \tilde{L}(u) \leqslant \check{C}\,|\hat{y}(u,\varepsilon) - \tilde{y}(\mu,\varepsilon)|_\mu^2$$
$$\leqslant \tilde{C}\,\rho(u,\tilde{u}(\mu,\varepsilon))^2 \tag{10}$$

for $u \in \mathcal{M}_\mu$ satisfying $\rho(u,\tilde{u}(\mu,\varepsilon)) \leqslant \tilde{\delta}\,\mathrm{corr}(\mu)$.

Proof. We will apply lemma 8, taking $\mathcal{W} = \mathcal{V}$, $N_\mu^\varepsilon = F_\mu(,\varepsilon)$, $\eta(\mu,\varepsilon)$ as given, and $u_\mu^\varepsilon = \tilde{u}(\mu,\varepsilon)$. The restriction of $F_\mu(,\varepsilon)$ to $\mathcal{M}_\mu$ is G_μ-invariant. Because $r(\bar{y},\varepsilon) = \mathcal{O}(\mathrm{corr}(\mu))$, the points $\tilde{y}(\mu,\varepsilon) \in \hat{\mathcal{D}}_\mu$. The functions $|r(\bar{y},\varepsilon)|$ and $k(\varepsilon t)$ are bounded; hence B_n and B_η exist. Thus the conditions of lemma 8 are satisfied and (10) holds. □

We can rewrite the time derivative of $\tilde{L}$, by using the appropriate functions on $\mathcal{Y}$. Define $\tilde{h}_\mu$ to be the push forward of $\tilde{H}_\mu$ by φ.

Lemma 12. *If $u(t)$ is a solution of (1), $\mu(t) = J(u(t))$ and $y(t) = \hat{y}(u(t),\varepsilon)$, then*

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{L}(u(t),\varepsilon) = \varepsilon\,D\tilde{h}_\mu(y,\varepsilon)\,v(y) + \langle\dot{\mu},\eta(\mu,\varepsilon)\rangle + \varepsilon[D_\mu f_\mu(y,\varepsilon) - D_\mu f_\mu(\tilde{y}(\mu,\varepsilon),\varepsilon)]\cdot\dot{\mu}$$
$$\text{where}\qquad v(y) = \hat{p}(y) - X_{f_\mu}(y,\varepsilon)\,. \tag{11}$$

Proof. By definition

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{L}(u(t),\varepsilon) = \frac{\mathrm{d}}{\mathrm{d}t}[\tilde{H}_{\mu(t)}(u(t),\varepsilon) - \tilde{H}_{\mu(t)}(\tilde{u}(\mu(t),\varepsilon),\varepsilon)]\,.$$

Using $D\tilde{H}_\mu(\tilde{u}(\mu,\varepsilon),\varepsilon) = DJ_{\eta(\mu,\varepsilon)}(\tilde{u}(\mu,\varepsilon))$ and $J(u) = J(\tilde{u}(\mu,\varepsilon))$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{L}(u(t),\varepsilon) = D\tilde{H}_\mu(u,\varepsilon)\,\dot{u} - DJ_{\eta(\mu,\varepsilon)}(\tilde{u})\,\dot{\tilde{u}} + \varepsilon\,[D_\mu F_\mu(u,\varepsilon) - D_\mu F_\mu(\tilde{u}(\mu,\varepsilon),\varepsilon)]\,\dot{\mu}\,.$$

Equation (1) and the definition of $\tilde{H}_\mu$ yield

$$\dot{u} = X_{\tilde{H}_\mu}(u) + \varepsilon\,P(u) - \varepsilon X_{F_\mu}(u,\varepsilon) - X_{J_\xi}(u)\,.$$

Using the identity $DJ_{\eta(\mu,\varepsilon)}(\tilde{u})\,\dot{\tilde{u}} = \langle\dot{\mu},\eta(\mu,\varepsilon)\rangle$ and the fact that all functions and vector fields involved are G_μ invariant, c.q. G_μ equivariant on $\mathcal{M}_\mu$, we see that (11) holds. □

Now we are ready to prove that the curve $\tilde{u}(\mu(t),\varepsilon)$ is a good approximation for the solution $u(t)$, if $u(0)$ is near $\{g\cdot\tilde{u}(\mu(0),\varepsilon) \mid g \in G_{\mu(0)}\}$.

Proposition 13. *Let $u(t)$ be a solution of the dissipative Hamiltonian system (1). If ε is sufficiently small and the initial condition $u(0)$ is such that $\rho(u(0), \tilde{u}(\mu(0), \varepsilon)) = \mathcal{O}(\varepsilon)$, then there exists an ε-independent constant $\tilde{C}$ such that*

$$\rho(u(t), \tilde{u}(\mu(t), \varepsilon)) \leqslant \tilde{C}\, \mathrm{e}^{-\varepsilon\beta t}\, \rho(u(0), \tilde{u}(\mu(0), \varepsilon)) + \varepsilon\, \tilde{C}\, \mathrm{e}^{-\min(\alpha,\beta)\varepsilon t}$$

for all $t \geqslant 0$. Again, this is equivalent to the following: If $\tilde{L}(u(0), \varepsilon) = \mathcal{O}(\varepsilon^2)$ then there exists an ε-independent constant $\breve{C}$ such that

$$\tilde{L}(u(t), \varepsilon) \leqslant \breve{C}\, \mathrm{e}^{-2\varepsilon\beta t}\, \tilde{L}(u(0), \varepsilon) + \varepsilon^2\, \breve{C}\, \mathrm{e}^{-2\varepsilon \min(\alpha,\beta) t}$$

for all $t \geqslant 0$.

Proof. Let $u(t)$ be a solution of the dissipative Hamiltonian system (1). If ε is sufficiently small and the initial condition $u(0)$ satisfies $\tilde{L}(u(0), \varepsilon) = \mathcal{O}(\varepsilon^2)$ then there is some $T > 0$ such that $u(t) \in \mathcal{U}$ and $\tilde{L}(u(t), \varepsilon) \leqslant \mathrm{e}^{-2\min(\alpha,\beta)\varepsilon t}$ for $0 \leqslant t \leqslant T$. Define $y(t) = \hat{y}(u(t), \varepsilon)$ for $0 \leqslant t \leqslant T$. Lemma 12 states that

$$\frac{1}{\varepsilon}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{L}(u(t), \varepsilon) = D\tilde{h}_\mu(y, \varepsilon)\, v(y) - \frac{1}{\varepsilon}\langle \dot{\mu}, \eta(\mu, \varepsilon)\rangle + [D_\mu f_\mu(y, \varepsilon) - D_\mu f_\mu(\tilde{y}(\mu, \varepsilon), \varepsilon)]\,\dot{\mu}\,. \tag{12}$$

We expand $D\tilde{h}_\mu(\tilde{y}(\mu, \varepsilon), \varepsilon)\, v(y) - \frac{1}{\varepsilon}\langle \dot{\mu}, \eta(\mu, \varepsilon)\rangle$ in terms of $\delta y = y - \tilde{y}(\mu, \varepsilon)$.

We can write

$$\begin{aligned} D\tilde{h}_\mu(y, \varepsilon)\, v(y) - \frac{1}{\varepsilon}\langle \dot{\mu}, \eta(\mu, \varepsilon)\rangle &= (D\tilde{h}_\mu(y, \varepsilon) - D\tilde{h}_\mu(\tilde{y}, \varepsilon))\, v(y) - (Dj_\eta(y) \\ &\quad -Dj_\eta(\tilde{y}))\, \hat{p}(y) - Dj_\eta(\tilde{y})(X_{f_\mu}(y, \varepsilon) - X_{f_\mu}(\tilde{y}, \varepsilon)) \end{aligned}$$

using the identity $D\tilde{h}_\mu(\tilde{y}, \varepsilon) = Dj_\eta(\tilde{y})$, its corollary

$$\varepsilon\, Dj_\eta(\tilde{y})\, X_{f_\mu}(\tilde{y}, \varepsilon) = Dj_\eta(\tilde{y})\, (X_{j_\eta}(\tilde{y}) - X_{h_\mu}(\tilde{y})) = 0$$

and the momentum evolution equation $\dot{\mu} = \varepsilon Dj(y)\, \hat{p}(y)$. The bound $|X_{f_\mu}(\tilde{y}, \varepsilon) - \frac{1}{\varepsilon}\mathrm{res}(\tilde{y}, \varepsilon)|_\mu = \mathcal{O}(|r(\tilde{y}, \varepsilon)|)$ and hypothesis (H8.2) imply that

$$\begin{aligned} \left| v(y) - \hat{p}(y) + \frac{1}{\varepsilon}\mathrm{res}(\tilde{y}, \varepsilon)\right|_\mu &\leqslant |X_{f_\mu}(y, \varepsilon) - X_{f_\mu}(\tilde{y}, \varepsilon)|_\mu + \mathcal{O}(|r(\tilde{y}, \varepsilon)|) \\ &= \mathcal{O}(k_f(\varepsilon t)(|\delta y|_\mu + |\tilde{y} - \bar{y}|_\mu) + |r(\tilde{y}, \varepsilon)|)\,. \end{aligned}$$

Hypotheses (H5.1), (H7), and (H8.2) imply that

$$\begin{aligned} &(D\tilde{h}_\mu(y, \varepsilon) - D\tilde{h}_\mu(\tilde{y}, \varepsilon))\left(\hat{p}(y) - \frac{1}{\varepsilon}\mathrm{res}(\tilde{y}, \varepsilon)\right) \\ &\quad= (Dh_\mu(y) - Dh_\mu(\tilde{y}))\left(\hat{p}(y) - \frac{1}{\varepsilon}\mathrm{res}(\tilde{y}, \varepsilon)\right) \\ &\qquad+\mathcal{O}\left(\varepsilon k_f(\varepsilon t)\left(|\delta y|_\mu + |\tilde{y} - \bar{y}|_\mu + \frac{1}{\varepsilon}|\mathbb{P}_{\tilde{y}}\dot{\tilde{y}}|\right)|\delta y|_\mu\right) \\ &\quad\leqslant -2\beta(h_\mu(y) - h_\mu(\tilde{y}) - Dh_\mu(\tilde{y})\delta y) \\ &\qquad+\mathcal{O}\left(\varepsilon k_f(\varepsilon t)\left(|\delta y|_\mu + |\tilde{y} - \bar{y}|_\mu + \frac{1}{\varepsilon}|\mathbb{P}_{\tilde{y}}\dot{\tilde{y}}| + \kappa(\varepsilon t)(\varepsilon \mathrm{e}^{-\alpha\varepsilon t} + |\delta y|_\mu)\right)|\delta y|_\mu\right). \end{aligned}$$

Hypotheses (H5) and (H8) yield

$$\begin{aligned} &h_\mu(y) - h_\mu(\tilde{y}) - Dh_\mu(\tilde{y})\delta y \\ &\quad= \tilde{h}_\mu(y) - \tilde{h}_\mu(\tilde{y}) - Dj_\eta(\tilde{y})\delta y + \varepsilon(f_\mu(y, \varepsilon) - f_\mu(\tilde{y}, \varepsilon) - Df_\mu(\tilde{y}, \varepsilon)\delta y) \\ &\quad= \tilde{L}(u, \varepsilon) + \mathcal{O}((\varepsilon k_f(\varepsilon t) + |\eta|)|\delta y|_\mu^2) \end{aligned}$$

and

$$|(Dj_\eta(y) - Dj_\eta(\tilde{y}))\,\hat{p}(y)| + |Dj_\eta(\tilde{y})(X_{f_\mu}(y,\varepsilon) - X_{f_\mu}(\tilde{y},\varepsilon))|$$
$$= \mathcal{O}(\,|\eta|\,|\delta y|_\mu(|\hat{p}(\tilde{y})|_\mu + |\delta y|_\mu + |\tilde{y} - \bar{y}|_\mu + \mathrm{e}^{-\varepsilon\alpha t}))\,.$$

Substituting these estimates into (12), we see that

$$\frac{1}{\varepsilon}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{L}(u(t),\varepsilon) \leqslant -2\beta\,\tilde{L}(u,\varepsilon) + \kappa_1(\varepsilon t)|\delta y|_\mu + \kappa_2(\varepsilon t)|\delta y|_\mu^2 \tag{13}$$

where

$$\kappa_1(\tau) = \mathcal{O}(k_f(\tau)|\mathbb{P}_{\bar{y}}\dot{\bar{y}}|_\mu + |r(\bar{y},\varepsilon)| + (\varepsilon\kappa(\tau) + |\eta|)\mathrm{e}^{-\alpha\tau})$$

and

$$\kappa_2(\tau) = \mathcal{O}(k_f(\tau) + \kappa(\tau) + |\eta|)\,.$$

Together with lemma 11 and (H6), this inequality implies that there exists an integrable function $k_1(\tau)$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{L}(u(t),\varepsilon) \leqslant -2\varepsilon\,\beta\,\tilde{L}(u,\varepsilon) + \varepsilon^2 k_1(\varepsilon t)\,\mathrm{e}^{-\alpha\varepsilon t}\sqrt{\tilde{L}(u,\varepsilon)} + \varepsilon\,k_1(\varepsilon t)\,\tilde{L}(u,\varepsilon)\,.$$

Integration of this expression gives the statement of the lemma for $0 \leqslant t \leqslant T$. But from this estimate we can conclude that for ε sufficiently small $T = \infty$. □

Using proposition 13, it is easy to prove the curve $\bar{u}(\mu(t))$ is also a good approximation for a solution $u(t)$, which starts in a neighbourhood of the MRE.

Proof of theorem 4. Let $u(t)$ be a solution of the perturbed system (1) and let $\mu(t) = J(u(t))$. Hypothesis (H8) implies that $d(\tilde{u}(\mu,\varepsilon),\bar{u}(\mu)) = \mathcal{O}(|r(\bar{y}(\mu),\varepsilon)|)$.

If $L(u(0)) = \mathcal{O}(\varepsilon^2)$, then $\rho(u(0),\bar{u}(\mu(0))) = \mathcal{O}(\varepsilon)$; hence the triangle inequality implies that $\rho(u(0),\tilde{(}\mu(0),\varepsilon)) = \mathcal{O}(\varepsilon)$. Applying proposition 13, we see that

$$\rho(u(t),\tilde{u}(\mu(t),\varepsilon)) = \mathcal{O}(\varepsilon\,\mathrm{e}^{-\min(\alpha,\beta)\varepsilon t})$$

for all $t \geqslant 0$. The triangle inequality now yields $\rho(u(t),\bar{u}(\mu(t))) = \mathcal{O}(\varepsilon\,\mathrm{e}^{-\min(\alpha,\beta)\varepsilon t})$.

Although we have suppressed this fact in our notation, one should realize that the curve $\bar{u}(\mu(t))$ passes through a finite number of charts $(\varphi_i,\mathcal{U}_i)$, $i = 1,\ldots,N$. Thus we can find real numbers $0 = t_1 < \cdots < t_N$ such that $t \in [t_j,t_{j+1}]$ implies that $\bar{u}(\mu(t)) \in \mathcal{U}_i$ and $(G_{\mu(t)}\cdot u(t)) \cap \mathcal{U}_i \neq \emptyset$. The orbital distance ρ from $u(t)$ to $\bar{u}(\mu(t))$ can be estimated in each time interval using the argument given above. This completes the proof. □

5. A rigid body with dissipation

As an application of the previous general theory, we consider a simple mechanical system consisting of a spherical rigid body placed in a rotational symmetric potential field. The position of the centre of mass is denoted by $q \in \mathbb{R}^3$ and the rotation of the body around its centre of mass is denoted by $\Lambda \in SO(3)$. The potential is given by a smooth function $V(\frac{1}{2}|q|^2)$. Furthermore, there is a dissipative perturbation that acts on the body. We will specify this dissipation later.

The configuration manifold is $Q = SO(3) \times \mathbb{R}^3$, hence $TQ \equiv Q \times so(3) \times \mathbb{R}^3$. We will identify $so(3)$ with $\mathbb{R}^3$ using the following identification. Given $\zeta \in \mathbb{R}^3$, let $\hat{\zeta}$ denote the skew matrix satisfying $\hat{\zeta}\,y = \zeta \times y$ for all $y \in \mathbb{R}^3$. Let $\Lambda \in SO(3)$, then

$$\delta\Lambda \in T_\Lambda SO(3) \equiv so(3) \Longleftrightarrow \exists_{\delta\theta\in\mathbb{R}^3}[\delta\Lambda = \Lambda\,\widehat{\delta\theta}]\,.$$

This implies that the phase space $\mathcal{M} = T^*Q \equiv SO(3) \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3$. We will denote an element of $\mathcal{M}$ by $u = (\Lambda, q\,;\, \Pi, p)$ and identify $T_u\mathcal{M}$ with $\mathbb{R}^{12}$. The Poisson structure is

$$\Gamma(u) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & \widehat{\Pi} & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

The Hamiltonian for the rigid body in the rotational symmetric potential field is

$$H(u) = \frac{1}{2n}\,|\Pi|^2 + \frac{1}{2m}\,|p|^2 + V\left(\frac{1}{2}|q|^2\right)$$

with Hamiltonian vector field

$$X_H(u) = \left(\frac{1}{n}\,\Lambda\,\hat{\Pi}, \frac{1}{m}\,p\,;\, 0, -V'\left(\frac{1}{2}|q|^2\right) q\right). \tag{14}$$

We assume that $V'(x)$ is strictly positive for all non-negative values of x. The symmetry group for this system is $G = SO(3) \times SO(3)$, where the action of $g = (R_1, R_2) \in G$ on $(\Lambda, q) \in Q$ is $g \cdot (\Lambda, q) = (R_1 \Lambda R_2^{\mathrm{T}}, R_1 q)$, with the induced action

$$g \cdot u = (R_1 \Lambda R_2^{\mathrm{T}}, R_1 q\,;\, R_2 \Pi, R_1 p) \tag{15}$$

on $\mathcal{M} = T^*Q$. The infinitesimal generator $(\xi, \omega)_{\mathcal{M}} : \mathcal{M} \to T\mathcal{M}$ associated to an algebra element $(\xi, \omega) \in \mathfrak{g} \equiv \mathbb{R}^3 \times \mathbb{R}^3$ is

$$(\xi, \omega)_{\mathcal{M}}(u) = (\Lambda^T \xi - \omega, \xi \times q\,;\, \omega \times \Pi, \xi \times p)\,.$$

The momentum map $J : \mathcal{M} \to \mathfrak{g}^* \equiv \mathbb{R}^3 \times \mathbb{R}^3$ associated to the G action is

$$J(u) = (q \times p + \Lambda \Pi\,, -\Pi)\,. \tag{16}$$

We consider the G-equivariant dissipative perturbation

$$P(u) = -\tfrac{1}{2}(0, q; 2\Pi, p) - \sigma(q \times p - \Lambda\Pi, 0)_{\mathcal{M}}(u)\,. \tag{17}$$

The first component of the perturbation consists of uniform damping in the (q, p)-variables and friction in Π. The σ-component is an infinitesimal spatial rotation about $q \times p - \Lambda\,\Pi$.

The dynamical equation for this dissipatively perturbed Hamiltonian system is

$$\dot{u} = X_H(u) + \varepsilon\, P(u) \qquad \text{or} \qquad \begin{cases} \dot{\Lambda} = \dfrac{1}{n}\,\Lambda\,\hat{\Pi} - \varepsilon\big(0 + \sigma(\widehat{q \times p}\,\Lambda - \Lambda\,\hat{\Pi})\big) \\ \dot{q} = \dfrac{1}{m}\,p - \varepsilon\left(\dfrac{1}{2}\,q + \sigma(q \times p - \Lambda\Pi) \times q\right) \\ \dot{\Pi} = 0 - \varepsilon(\Pi + \sigma 0) \\ \dot{p} = -V'\big(\tfrac{1}{2}|q|^2\big)\,q - \varepsilon\big(\tfrac{1}{2}\,p + \sigma(q \times p - \Lambda\Pi) \times p\big)\,. \end{cases} \tag{18}$$

5.1. *The hypotheses*

First we check the hypotheses on the relative equilibria of the unperturbed Hamiltonian system. The first observation is the existence of a stable family of relative equilibria. In general, this family is defined on a subset of $\mathfrak{g}^*$. We shall show below that there exists a relative equilibrium with momentum (μ, ν) if and only if there exists a positive number $\bar{x}$ satisfying the equation

$$4\,m\,\bar{x}^2\,V'(\bar{x}) = (|\mu| - |\nu|)^2\,. \tag{19}$$

To satisfy hypothesis (H2), we need to set conditions on $\mathfrak{g}^*_{\mathrm{MRE}}$ that guarantee that the relative equilibria with momentum values in $\mathfrak{g}^*_{\mathrm{MRE}}$ have trivial isotropy. Therefore we define the G-invariant subset $\mathfrak{g}^*_{\mathrm{MRE}}$ of $\mathfrak{g}^*$ by

$$\mathfrak{g}^*_{\mathrm{MRE}} = \left\{(\mu, \nu) \in \mathfrak{g}^* \mid \exists_{\bar{x}>0}\left[4\, m\, \bar{x}^2\, V'(\bar{x}) = (|\mu| - |\nu|)^2 \neq 0 \neq |\mu||\nu|\right]\right\}.$$

Let $\underline{\mu} = (\mu, \nu) \in \mathfrak{g}^*_{\mathrm{MRE}}$. Let $\bar{\Lambda}(\underline{\mu})$ be an element of $SO(3)$ satisfying $\mu \cdot \bar{\Lambda}\nu = -|\mu||\nu|$ and $\bar{x}(\underline{\mu}) > 0$ be a solution of (19). Finally, let $\bar{q}(\underline{\mu}) \in \mathbb{R}^3$ be a vector of length $\sqrt{2\bar{x}}$ in the plane orthogonal to μ. Define

$$\bar{\Pi}(\underline{\mu}) = -\nu \qquad \text{and} \qquad \bar{p}(\underline{\mu}) = \tfrac{|\mu|-|\nu|}{2\bar{x}\,|\mu|}\,(\mu \times \bar{q}). \tag{20}$$

It is an immediate consequence of (19) and (20) that the lengths of $\bar{p}$ and $\bar{q}$ satisfy the relation

$$|\bar{p}|^2 = 2\, m\, \bar{x}\, V'(\bar{x}) = m\, |\bar{q}|^2 V'(|\bar{q}|^2/2). \tag{21}$$

Straightforward calculations show that $J(\bar{u}(\underline{\mu})) = \underline{\mu}$ and $DH(\bar{u}) = DJ_{\underline{\xi}(\underline{\mu})}(\bar{u})$, with

$$\underline{\xi}(\underline{\mu}) = (\bar{\xi}(\underline{\mu}), \bar{\omega}(\underline{\mu})) = \frac{1}{2\, m\, \bar{x}}\left(\left(1 - \frac{|\nu|}{|\mu|}\right)\mu, \left(1 - \frac{|\mu|}{|\nu|} + \frac{2\, m\, \bar{x}}{n}\right)\nu\right). \tag{22}$$

This implies that $\bar{u}(\underline{\mu}) = (\bar{\Lambda}(\underline{\mu}), \bar{q}(\underline{\mu}), \bar{\Pi}(\underline{\mu}), \bar{p}(\underline{\mu}))$ is a relative equilibrium with generator $\underline{\xi}(\underline{\mu})$. (See, for example, Abraham and Marsden 1978.)

Equivariance implies that any element of the $G_{\underline{\mu}}$-orbit of $\bar{u}(\underline{\mu})$ is a relative equilibrium with the same generator. In other words, for every $\varphi, \theta \in \mathbb{R}$, $\bar{u}(\underline{\mu}; \varphi, \theta) = (\exp(\varphi\, \mu), \exp(\theta\, \nu)) \cdot \bar{u}(\underline{\mu})$ is a relative equilibrium with generator $\underline{\xi}(\underline{\mu})$ given by (22). Hence $\mathrm{MRE}_{\underline{\mu}}$ is the $G_{\underline{\mu}}$ orbit of $\bar{u}(\underline{\mu})$ and the hypotheses (H1) is satisfied.

It is an immediate consequence of (15) that if $u = (\Lambda, q; \Pi, p)$ is fixed by any non-trivial element of the symmetry group, then q and p must be parallel, which implies that $J(u) = (\Lambda\Pi, -\Pi) \notin \mathfrak{g}^*_{\mathrm{MRE}}$. Hence all elements of $\mathfrak{g}^*_{\mathrm{MRE}}$ are regular values of J and the hypothesis (H2) that relative equilibria taking momentum values in $\mathfrak{g}^*_{\mathrm{MRE}}$ have trivial isotropy holds.

We now specify a condition on the potential V which guarantees that the relative equilibria with momentum in $\mathfrak{g}^*_{\mathrm{MRE}}$ are stable. Specifically, if the function $x^2\, V'(x)$ is monotone increasing, then the relative equilibria $\bar{u}(\underline{\mu}; \varphi, \theta)$ are orbitally stable for all $\varphi, \theta \in \mathbb{R}$. Note that this implies that $||\mu| - |\nu|| = |\bar{q} \times \bar{p}|$ increases monotonically with $|\bar{q}|$.

Lemma 14. *If*

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[x^2\, V'(x)\right] = x\,(x\, V''(x) + 2\, V'(x)) > 0 \tag{23}$$

for all positive x, then the relative equilibria $\bar{u}(\underline{\mu}; \varphi, \theta)$ are $G_{\underline{\mu}}$-orbitally stable. In particular, $D^2 H_{\underline{\mu}}(\bar{u})$ is positive semi-definite on $T_{\bar{u}}\mathcal{M}_{\underline{\mu}}$ with kernel $\mathfrak{g}_{\underline{\mu}} \cdot \bar{u}$.

Proof. Let $\underline{\mu}$ be an element of $\mathfrak{g}^*_{\mathrm{MRE}}$ and set $\bar{u} = \bar{u}(\underline{\mu}; \varphi, \theta)$ for some choice of φ and θ. (Note that if $\bar{u}$ is stable, then equivariance implies that any point in the $G_{\underline{\mu}}$-orbit of $\bar{u}$ is also stable.)

To prove orbital stability of $\bar{u}$, it is sufficient to show that $D^2 H_{\underline{\mu}}(\bar{u})$ is definite on the orthogonal complement $\mathcal{S}(\bar{u})$ to $\mathfrak{g}_{\underline{\mu}} \cdot \bar{u}$ in $\ker[DJ(\bar{u})]$. We can explicitly describe the basis with respect to which the restriction of $D^2H_{\underline{\mu}}(\bar{u})$ to $\mathcal{S}(\bar{u})$ diagonalizes. The second variation of the energy–momentum function is written in block matrix form as

$$D^2H_{\underline{\mu}}(u) = \begin{pmatrix} -\Lambda^T\bar{\xi}\cdot D^2\exp(0)(\Lambda^T\cdot,\Lambda^T\cdot)\Pi & 0 & \widehat{\Lambda^T\bar{\xi}} & 0 \\ 0 & V'(x)\mathbf{1}+V'(x)q\otimes q & 0 & \hat{\bar{\xi}} \\ -\widehat{\Lambda^T\bar{\xi}} & 0 & \frac{1}{n}\mathbf{1} & 0 \\ 0 & -\hat{\bar{\xi}} & 0 & \frac{1}{m}\mathbf{1} \end{pmatrix} \tag{24}$$

where $\exp : so(3) \to SO(3)$ denotes the exponential map. The identity $-\bar{\Lambda}^T\bar{\xi}\cdot D^2\exp(0)(\zeta,\eta)\bar{\Pi} = (\bar{\Lambda}^T\bar{\xi}\times\zeta)\cdot(\bar{\Pi}\times\eta)$ implies that $D^2H_{\underline{\mu}}(\bar{u})(v_i,v_j) = \delta_{ij}\lambda_i|v_i|^2$, where

$$\begin{aligned}
\lambda_1 &= \frac{2\,(\bar{x}V''(\bar{x})+2\,V'(\bar{x}))}{1+mV'(\bar{x})} & v_1 &= (0,\bar{q}\,;\,0,-\bar{p}) \\
\lambda_2 &= \frac{1}{m}+V'(\bar{x}) & v_2 &= (0,\bar{p};\,0,\bar{q}) \\
\lambda_3 &= \frac{|\mu||\nu|}{m\left(2\bar{x}\left(1+|\nu|^2\right)+|\nu|^2\right)} & v_3 &= (\bar{\Lambda}^T\bar{q},0\,;0,-\bar{\Lambda}\bar{\Pi}) \\
\lambda_4 &= \frac{|\mu||\nu|}{2m\,\bar{x}\left(1+|\nu|^2\right)+|\nu|^2/V'(\bar{x})} & v_4 &= (\bar{\Lambda}^T\bar{p},\bar{\Lambda}\bar{\Pi}\,;0,0)\,.
\end{aligned} \tag{25}$$

This can be seen by straightforward calculations involving repeated use of the equilibrium relations.

All of the eigenvalues except λ_1 are guaranteed to be positive, since the equilibrium conditions imply that $V'(\bar{x})$ is positive. Since $\mathcal{S}(\bar{u}) = \operatorname{span}\{v_1,v_2,v_3,v_4\}$, the restriction of $D^2H_{\underline{\mu}}(\bar{u})$ to $\mathcal{S}(\bar{u})$ is positive-definite, and $\bar{u}$ is orbitally stable, if (23) holds. □

Lemma 14 shows that hypothesis (H3) is satisfied. Orbital stability of the relative equilibrium $\bar{u}$ implies that the functional $L(u) = H(u) - H(\bar{u}(J(u)))$ is a measure for the orbital distance between u and $\bar{u}(J(u))$. The functional L depends on the variables Π and Λ only through the momentum $J(u)$. Specifically, let $J(u) = \underline{\mu} = (\mu,\nu) \in \mathfrak{g}^*_{\mathrm{MRE}}$ and let $\bar{u}$ be a relative equilibrium with momentum $\underline{\mu}$. Then $\Pi = -\nu = \bar{\Pi}$ and thus

$$\begin{aligned}
L(u) &= \frac{1}{2m}(|p|^2-|\bar{p}|^2)+V(|q|^2/2)-V(|\bar{q}|^2/2) \\
&= \frac{1}{2m}\,|p|^2+V(|q|^2/2)-V(\bar{x})-\bar{x}\,V'(\bar{x})\,.
\end{aligned}$$

Given a solution $u(t)$ of the dynamical system (18), we define the projected curve in the family of relative equilibria as follows. Define $\underline{\mu}(t) = J(u(t))$ and write $\underline{\mu}_0 = \underline{\mu}(0)$. Solving the differential equations for μ and ν, we find that

$$\underline{\mu}(t) = \mathrm{e}^{-\varepsilon t}\,(R(t)\,\mu_0,\nu_0)$$
$$\text{where} \qquad R(t) = \exp(\sigma(\mathrm{e}^{-\varepsilon t}-1)\,[\mu_0 - 2\,\Lambda(0)\Pi(0)]) \in SO(3)\,. \tag{26}$$

Let $\bar{\Lambda}_0$ be some element in $SO(3)$ such that $\mu_0\cdot\bar{\Lambda}_0\,\nu_0 = -|\mu_0||\nu_0|$. Let u_0 be some unit vector perpendicular to μ_0 and let $\bar{x}(t)$ satisfy the relation (19) for $\mu = \mu(t)$ and $\nu = \nu(t)$. We define the curve $\bar{u} : \mathbb{R}^+ \to \mathrm{MRE}$ of relative equilibria by

$$\bar{u}(t) = (R(t),\mathrm{Id})\left(\bar{\Lambda}_0,\sqrt{2\bar{x}(t)}\,u_0\,;\,-\nu(t),\,\frac{\mathrm{e}^{-\varepsilon t}(|\mu_0|-|\nu_0|)}{\sqrt{2\bar{x}(t)}|\mu_0|}\,\mu_0\times u_0\right). \tag{27}$$

The expression (26) for the momentum map implies that $|\underline{\mu}(t)| = e^{-\varepsilon t}|\underline{\mu}_0|$. Thus if $J(u(t))$ is initially in the subset $\mathfrak{g}^*_{\mathrm{MRE}}$ of $\mathfrak{g}^*$, then $J(u(t))$ stays in this subset for all positive times t, but $\mu_\infty = 0 \notin \mathfrak{g}^*_{\mathrm{MRE}}$. If we consider the evolution of the eigenvalues (25), we see that λ_3 and λ_4 decay to zero as $t \to \infty$. Thus we are indeed in the case in which we need a scaled metric to describe the behaviour of the Lyapunov function. Since $|\underline{\mu}(t)| \to 0$, $\bar{x}(t)$ must converge either to zero or to an extremum of the function V. We have assumed that V' is positive on $\mathbb{R}^+$; hence $\bar{x} \to 0$ as $t \to \infty$.

If equation (23) is satisfied, then it follows from the implicit definition (19) of $\bar{x}$ and the relation (21) between $|\bar{q}|$ and $|\bar{p}|$ that

$$|\bar{q}(\underline{\mu})| = \mathcal{O}(|\underline{\mu}|^{1/2}) = \mathcal{O}(e^{-\varepsilon t/2}) \qquad \text{and} \qquad |\bar{p}(\underline{\mu})| = \mathcal{O}(|\underline{\mu}|^{1/2}) = \mathcal{O}(e^{-\varepsilon t/2}) \quad \text{as} \quad t \to \infty. \tag{28}$$

Specifically, if $c_V^2 = m \min_{0\leqslant x\leqslant \bar{x}(0)} V'(x)$ and $C_V^2 = m \max_{0\leqslant x\leqslant \bar{x}(0)} V'(x)$ (with c_V and C_V positive), then

$$\frac{1}{C_V} \leqslant \frac{|\bar{q}|^2}{||\mu| - |\nu||} \leqslant \frac{1}{c_V} \qquad \text{and} \qquad c_V \leqslant \frac{|\bar{p}|^2}{||\mu| - |\nu||} \leqslant C_V\,.$$

Thus $\lim_{t\to\infty} \bar{u}(\underline{\mu}(t)) = (\bar{\Lambda}_\infty, 0; 0, 0)$ for some $\bar{\Lambda}_\infty \in SO(3)$. Furthermore, by differentiating the relation (19) with respect to t, we obtain

$$\frac{\dot{\bar{x}}}{\bar{x}} = \frac{-2\varepsilon V'(\bar{x})}{2V'(\bar{x}) + \bar{x}V'(\bar{x})} = \varepsilon(c(\underline{\mu}) - 1) \qquad \text{where} \quad c(\underline{\mu}) = \frac{\bar{x}V''(\bar{x})}{2V'(\bar{x}) + \bar{x}V'(\bar{x})}\,. \tag{29}$$

The estimate $\bar{x} = \mathcal{O}(|\underline{\mu}|)$ follows from the expression (19); hence $c(\underline{\mu}(t)) = \mathcal{O}(e^{-\varepsilon t})$. Differentiation of the expression (22) gives

$$\frac{1}{\varepsilon}\dot{\underline{\xi}} = -c(\underline{\mu})\,\underline{\xi} - \left(\frac{\sigma}{m\bar{x}}\left(1 - \frac{|\nu|}{|\mu|}\right)\mu \times \Pi,\ \frac{\dot{\bar{x}}}{n\bar{x}}\nu\right).$$

Hence we conclude that $|\dot{\underline{\xi}}| = \mathcal{O}(\varepsilon e^{-\varepsilon t})$ and (H4) is satisfied.

Equation (25) implies that two of the eigenvalues (λ_3 and λ_4) of the restricted second variation of the energy–momentum function are of order $|\underline{\mu}|$ as $\underline{\mu} \to 0$, while the other two are of order one. Thus, while the energy–momentum function can be used as a Lyapunov function, the estimates obtained using this function become increasingly weak as $\underline{\mu} \to 0$. Specifically, if we define the orbital distance functional

$$d^2_{\underline{\mu}}(u, u') = \min_{g\in G_{\underline{\mu}}}[\,|g\cdot q - q'|^2 + |g\cdot p - p'|^2\,] + |\Lambda\nu \times \Lambda'\nu|^2/|\nu|^4$$

on $\mathcal{M}_{\underline{\mu}}$, then Lyapunov stability arguments involving $H_{\underline{\mu}}$ yield $c_h(\underline{\mu}(t))\, d^2_{\underline{\mu}(t)}(u(t), \bar{u}(\underline{\mu}(t))) \leqslant L(u(t))$ only if $c_h(\underline{\mu}) = \mathcal{O}(|\underline{\mu}|)$ as $\underline{\mu} \to 0$.

While we cannot sharpen this estimate, in the sense of finding a better-behaved function c_h, we can replace the estimate with a more informative one by choosing a different distance function. The distance function $d_{\underline{\mu}}$ takes equally into account the influence of all of the components of u. However, $|\bar{q}(\underline{\mu})|$ and $|\bar{p}(\underline{\mu})|$ are of order $|\underline{\mu}|^{1/2}$ as $\underline{\mu} \to 0$. Thus, we would need $d_{\underline{\mu}}(u, \bar{u}) = \mathcal{O}(|\underline{\mu}|^{1/2})$ to show that the relative distances $|q - \bar{q}|/|\bar{q}|$ and $|p - \bar{p}|/|\bar{p}|$ (assume that $\bar{u}$ is optimally rotated) are even bounded as $\underline{\mu} \to 0$. However, the orbital distance between Λ and $\bar{\Lambda}$ does not tend towards zero with $\underline{\mu}$. We can explicitly integrate the differential equation for Λ, obtaining $\Lambda(t) = R(t)\Lambda(0)\exp\!\left(\frac{(e^{-\varepsilon t}-1)}{\varepsilon n}\nu(0)\right)$, where $R(t)$ is given by (26). Thus $|\Lambda\nu \times \bar{\Lambda}\nu|/|\nu|^2$ is time independent. Therefore we define a weighted distance functional that progressively discounts Λ as $\underline{\mu} \to 0$. (It does not seem reasonable to expect much better control than this—in the limit $\underline{\mu} = 0$, the momentum isotropy subgroup is the full group G, so the limiting equilibrium is only stable modulo arbitrary rotations.)

Definition 15. *Let* $\varphi : \mathcal{M} \to \mathbb{R}^{12}$ *be a chart map such that* $\varphi(\Lambda, q; \Pi, p) = (\Theta(\Lambda), q; \Pi, p)$ *for some chart map* Θ *on the appropriate neighbourhood of* Λ *in* $SO(3)$, *e.g.* $\Theta(\Lambda) = exp^{-1}(\Lambda_*^T \Lambda)$ *on a neighbourhood of some matrix* Λ_*, *and define the map* $B_{\underline{\mu}} : \mathbb{R}^{12} \to \mathbb{R}^{12}$ *by*

$$B_{\underline{\mu}}(\varphi(\Lambda, q; \Pi, p)) = \left(|\nu|^{1/2}\Theta(\Lambda), q;\ |\nu|^{-1/2}\Pi, p\right).$$

Define the weighted distance functional $\tilde{d}$ *on* $\mathcal{M}_{\underline{\mu}}$ *by*

$$\tilde{d}^2(u, u') = [\,|q - q'|^2 + |p - p'|^2\,] + |\Lambda\nu \times \Lambda'\nu|^2/|\nu|^3 \qquad u, u' \in \mathcal{M}_{\underline{\mu}}.$$

Then $\tilde{d}$ *and* $|\,|_{\underline{\mu}}$ *are compatible.*

Note that since $|\nu(t)|$ decreases with time, any function that does not depend on the component of y corresponding to the Λ variable on the original manifold is differentiable, Lipschitz continuous, etc with respect to $||_{\underline{\mu}}$ if it has that property with respect to the standard Euclidean metric.

We now check hypothesis (H5). We can write $h_{\underline{\mu}}$ as the sum: $h_{\underline{\mu}} = \hat{h}_{\underline{\mu}} + \check{h}_{\underline{\mu}}$ for

$$\hat{h}_{\underline{\mu}}(y) = \frac{1}{2m}|p|^2 + V(|q|^2/2) - \bar{\xi}(\underline{\mu}) \cdot q \times p$$

and

$$\check{h}_{\underline{\mu}}(y) = \frac{1}{2n}|\Pi|^2 - (\Lambda^T \xi(\underline{\mu}) - \omega(\underline{\mu})) \cdot \Pi$$

where $\Lambda \in SO(3)$ is the rotational component of $\varphi^{-1}(y)$. Note that $\hat{h}_{\underline{\mu}}$ is the energy–momentum function of the point mass system with potential V, while $\check{h}_{\underline{\mu}}$ is the energy–momentum function of the spherical free rigid body. The function $\hat{h}_{\underline{\mu}}$ is smooth and independent of θ; hence $\hat{h}_{\underline{\mu}} \in LC^{\ell}_{\underline{\mu}}(\mathcal{D}_{\underline{\mu}}, \mathbb{R})$ for any ℓ. Inspection of the expression for $\check{h}_{\underline{\mu}}$ and the Λ and Π components of (24) show that

$$\begin{aligned} D^3\check{h}_{\underline{\mu}}(\theta, q; \Pi, p)(\delta y_1, \delta y_2, \delta y_3) &= |\nu|^{-1/2} D^3\check{h}_{\underline{\mu}}(\theta, q; \Pi/|\nu|, p)(B_{\underline{\mu}}\delta y_1, B_{\underline{\mu}}\delta y_2, B_{\underline{\mu}}\delta y_3) \\ &\leqslant |\nu|^{-1/2} C_{\check{h}} |\delta y_1|_{\underline{\mu}} |\delta y_2|_{\underline{\mu}} |\delta y_3|_{\underline{\mu}} \end{aligned}$$

for some constant $C_{\check{h}}$ and all $(\theta, q; \Pi, p) \in \mathcal{D}_{\underline{\mu}}$, where $\mathrm{d}(\underline{\mu}) = \sqrt{|\nu|}$, since $\check{h}_{\underline{\mu}}$ is smooth with respect to the Euclidean norm. Hence $h_{\underline{\mu}} \in LC^2_{\underline{\mu}}(\mathcal{D}_{\underline{\mu}}, \mathbb{R}; C_h/\mathrm{corr}(\underline{\mu}))$ for some constant C_h if $\mathrm{corr}(\underline{\mu}) = \sqrt{|\nu|}$.

The second variation of the momentum map satisfies

$$D^2 j(\theta, q; \Pi, p)(\delta y, \delta y) = D^2 j(\theta, q; \Pi/|\nu|, p)\left(B_{\underline{\mu}}\delta y, B_{\underline{\mu}}\delta y\right) = \mathcal{O}\left(|\delta y|^2_{\underline{\mu}}\right)$$

for all $\delta y \in \mathbb{R}^{2n}$ and all $(\theta, q; \Pi, p) \in \mathcal{D}_{\underline{\mu}}$. Thus $j \in LC^1_{\underline{\mu}}(\mathcal{D}_{\underline{\mu}}, \mathfrak{g}^*)$. It follows immediately from the expression (14) that $X_{h_{\underline{\mu}}} \in LC^0_{\underline{\mu}}(\mathcal{D}_{\underline{\mu}}, \mathbb{R})$.

The vector field $\hat{p}$ satisfies

$$\begin{aligned} |\hat{p}(y_1) - \hat{p}(y_2)|_{\underline{\mu}} &\leqslant |y_1 - y_2|_{\underline{\mu}} + |\mathbb{P}_{y_1}(X_{j(\Lambda_1\Pi_1,0)}(y_1)) - \mathbb{P}_{y_2}(X_{j(\Lambda_1\Pi_1,0)}(y_2))|_{\underline{\mu}} \\ &\quad + |\mathbb{P}_{y_2}(X_{j(\Lambda_1\Pi_1 - \Lambda_2\Pi_2, 0)}(y_2))|_{\underline{\mu}} \\ &= \mathcal{O}(|y_1 - y_2|_{\underline{\mu}}) \end{aligned}$$

for $y_1,\ y_2 \in \mathcal{D}_{\underline{\mu}}$, since the orthogonal projections $\mathbb{P}_y$ and the vector field $X_{j(\Lambda_1\Pi_1,0)}$ are Lipschitz continuous with respect to $|\,|_{\underline{\mu}}$. Hence $\hat{p} \in LC^0_{\underline{\mu}}(\mathcal{D}_{\underline{\mu}}, \mathbb{R}^{2n})$, as required, and (H5.1) is satisfied.

The condition (H5.2) follows immediately from the definition of the Hamiltonian and equation (25). A straightforward calculation shows that for $\underline{\xi} = (\xi, \omega) \in \mathfrak{g}$

$$|\nabla_{\underline{\mu}} j_{\underline{\xi}}(y)|^2_{\underline{\mu}} = |(|\nu|^{-1/2}(D_\theta \jmath(y))^T \xi, p \times \xi; |\nu|^{1/2}(\Lambda^T \xi - \omega), \xi \times q)|^2 = |\nu| \, \underline{\xi} \cdot M(y) \underline{\xi}$$

where $D_\theta \jmath(y) \in L(\mathbb{R}^3, \mathbb{R}^3)$ is determined by the relationship $D_\theta \jmath(y)\delta\theta = (D\Theta^{-1}(\theta)\delta\theta)\Pi$, and

$$M(y) = \begin{pmatrix} 1 - (\hat{q}\hat{q} + \hat{p}\hat{p})/|\nu| - D_\theta \jmath(y) D_\theta \jmath(y)^T/|\nu|^2 & -\Lambda \\ -\Lambda^T & \mathbf{1} \end{pmatrix}.$$

If we set $c_j(\underline{\mu}) = \min_{y \in \mathcal{D}_{\underline{\mu}}} \lambda_{\min}(y)$ and $C_j(\underline{\mu}) = \max_{y \in \mathcal{D}_{\underline{\mu}}} \lambda_{\max}(y)$, where $\lambda_{\min}(y)$ (respectively $\lambda_{\max}(y)$) denotes the minimum (respectively maximum) eigenvalue of $M(y)$, and $c_j = \min_{\underline{\mu} \in \overline{(\underline{\mu}(t))}} c_j(\underline{\mu})$, $C_j = \max_{\underline{\mu} \in \overline{(\underline{\mu}(t))}} C_j(\underline{\mu})$ then

$$c_j \, |\nu| \, |\underline{\xi}|^2 \leqslant |\nabla_{\underline{\mu}} j_{\underline{\xi}}(y)|^2_{\underline{\mu}} \leqslant C_j \, |\nu| \, |\underline{\xi}|^2$$

for all $y \in \mathcal{D}_{\underline{\mu}}$ and all $\underline{\mu} \in \overline{(\underline{\mu}(t))}$. Note that $c_j > 0$ because $|\bar{q}|/|\nu|^{1/2}$ and $|\bar{p}|/|\nu|^{1/2}$ are bounded away from zero. Furthermore, $D\jmath(y)\nabla_{\underline{\mu}} j_{\underline{\xi}}(y) = |\nu| \, M(y)\underline{\xi}$. Altogether this implies

$$\frac{c_j}{\sqrt{C_j}} \sqrt{|\nu|} |\nabla_{\underline{\mu}} j_{\underline{\xi}}(\bar{y})|_{\underline{\mu}} \leqslant |D\jmath(\bar{y})\nabla_{\underline{\mu}} j_{\underline{\xi}}(\bar{y})| \leqslant \frac{C_j}{\sqrt{c_j}} \sqrt{|\nu|} |\nabla_{\underline{\mu}} j_{\underline{\xi}}(\bar{y})|_{\underline{\mu}}.$$

The infinitesimal generator $X_{j_\xi}(y)$ satisfies

$$|X_{j_\xi}(y)|^2_{\underline{\mu}} = |\nu| \, (\xi, \omega) \cdot \left(M(y) + \mathrm{diag}(D_\theta \jmath(y) D_\theta \jmath(y)^T, -D_\theta \jmath(y) D_\theta \jmath(y)^T)\right) (\xi, \omega)^T$$

so that $c_j \, |\nu| \, |\underline{\xi}|^2 \leqslant |X_{j_\xi}(\bar{y})|^2_{\underline{\mu}}$ and the condition (H5.3) is satisfied for $\mathrm{corr}(\underline{\mu}) = \sqrt{|\nu|}$. Hence we conclude that (H5) is satisfied.

To verify (H6), we differentiate the equations for the relative equilibria $\bar{u}$ with respect to time, obtaining

$$\dot{\bar{u}} = \varepsilon\sigma(2\Lambda\Pi(t) - \mu, 0)_{\mathcal{M}}(\bar{u}) - \frac{1}{2}\left(0, -\frac{\dot{\bar{x}}}{\bar{x}}\bar{q}; 2\varepsilon\,\bar{\Pi}, \left(2\varepsilon + \frac{\dot{\bar{x}}}{\bar{x}}\right)\bar{p}\right). \tag{30}$$

Comparing (30) with the expression (17) for the perturbation P, we see that

$$\mathrm{res}(\bar{y}, \varepsilon) = \frac{1}{2}\left(\varepsilon + \frac{\dot{\bar{x}}}{\bar{x}}\right)(0, -\bar{q}; 0, \bar{p}) - \varepsilon \mathbb{P}_{\bar{y}}(X_{j_{(\zeta,0)}}(\bar{y})) = \varepsilon \, X_{\bar{f}_{\underline{\mu}}}(\bar{y}) \tag{31}$$

where

$$\zeta = 2\sigma(\bar{\Lambda}\bar{\Pi} - \Lambda(t)\Pi(t)) = 2\sigma(\Lambda(t) - \bar{\Lambda})\nu$$

and

$$\bar{f}_{\underline{\mu}}(y) = -\tfrac{1}{2}\, c(\underline{\mu})\,(q, p) - j_{(\zeta,0)}(y)$$

for $c(\mu)$ as given in (29). Because $c(\underline{\mu}) = \mathcal{O}(|\underline{\mu}|)$, $|r(\bar{y}, \varepsilon)| = \mathcal{O}(\varepsilon|\underline{\mu}|^{3/2})$. Similarly, $|\hat{p}(\bar{y})|_{\underline{\mu}} = \mathcal{O}(\varepsilon|\underline{\mu}|^{1/2})$. Hence we can conclude that (H6) is satisfied for $\alpha = -\frac{1}{2}$ and $k(\tau) = \mathrm{e}^{-\tau}$.

Next we consider the dissipation coefficient. The energy–momentum function $h_{\underline{\mu}} \notin LC^2_{\underline{\mu}}(\mathcal{D}_{\underline{\mu}}, \mathbb{R})$; hence (H7A) is not valid and we must work directly with hypothesis ($\tilde{\mathrm{H}}$7). First we show that the contribution of the rotational components of the dissipation and the residual is neglible. We see that

$$\left|\hat{p}(y) - \frac{1}{\varepsilon}\mathrm{res}(\bar{y}, \varepsilon) + \frac{1}{2}(0, q - \bar{q}; 0, p - \bar{p}) - \frac{1}{\varepsilon}\mathbb{P}_{\bar{y}}\dot{\bar{y}}\right|_{\underline{\mu}} = |\mathbb{P}_y(X_{j_{(\zeta,0)}}(y)) - \mathbb{P}_{\bar{y}}(X_{j_{\zeta,0}}(\bar{y}))|_{\underline{\mu}}$$
$$= \mathcal{O}(|\zeta|\mathrm{corr}(\mu)).$$

The bound $|\zeta| = |(\Lambda - \bar{\Lambda})\nu| = \mathcal{O}(|\nu|^{1/2}|y - \bar{y}|_{\underline{\mu}})$ and the Lipschitz continuity of $Dh_{\underline{\mu}}$ imply that

$$(Dh_{\underline{\mu}}(y) - Dh_{\underline{\mu}}(\tilde{y}))\left(\hat{p}(y) - \frac{1}{\varepsilon}\mathrm{res}(\tilde{y}, \varepsilon)\right)$$
$$= (Dh_{\underline{\mu}}(y) - Dh_{\underline{\mu}}(\tilde{y}))\left(-\frac{1}{2}(0, \delta q; 0, \delta p) + \frac{1}{\varepsilon}\mathbb{P}_{\tilde{y}}\dot{\tilde{y}}\right)$$
$$+\mathcal{O}((|\nu||\delta y|_{\underline{\mu}} + |\tilde{y} - \bar{y}|_{\underline{\mu}})|\delta y|_{\underline{\mu}}). \tag{32}$$

We make use of the decomposition $h_{\underline{\mu}} = \hat{h}_{\underline{\mu}} + \check{h}_{\underline{\mu}}$, estimating the terms associated to $\hat{h}_{\underline{\mu}}$ and $\check{h}_{\underline{\mu}}$ separately. Applying Taylor's theorem to $\hat{h}_{\underline{\mu}}$ yields

$$\hat{h}_{\underline{\mu}}(y) - \hat{h}_{\underline{\mu}}(\tilde{y}) - D\hat{h}_{\underline{\mu}}(\tilde{y})\delta y = \tfrac{1}{2}D^2\hat{h}_{\underline{\mu}}(\tilde{y})(\delta y, \delta y) + \mathcal{O}(|\delta y|^3_{\underline{\mu}}) \tag{33}$$

and

$$(D\hat{h}_{\underline{\mu}}(y) - D\hat{h}_{\underline{\mu}}(\tilde{y}))\left(-\frac{1}{2}(0, q - \bar{q}; 2\Pi, p - \bar{p}) + \frac{1}{\varepsilon}\mathbb{P}_{\tilde{y}}\dot{\tilde{y}}\right) - \frac{1}{2}D^2\hat{h}_{\underline{\mu}}(\tilde{y})(\delta y, \delta y)$$
$$+\frac{1}{\varepsilon}D^2\hat{h}_{\underline{\mu}}(\tilde{y})(\dot{\tilde{y}}, \delta y) + \mathcal{O}((|\hat{p}(y)|_{\mu} + |\mathbb{P}_{\tilde{y}}\dot{\tilde{y}}|_{\underline{\mu}})|y - \bar{y}|_{\underline{\mu}}|\delta y|_{\underline{\mu}}). \tag{34}$$

The relationship $D^2h_{\underline{\mu}}(\tilde{y})(\dot{\tilde{y}}, \delta y) = Dj_{\dot{\underline{\xi}}}(\tilde{y})\delta y$ implies that

$$D^2\hat{h}_{\underline{\mu}}(\tilde{y})(\dot{\tilde{y}}, \delta y) = D^2h_{\underline{\mu}}(\tilde{y})(\dot{\tilde{y}}, (0, \delta q; 0, \delta p)) = Dj_{\dot{\underline{\xi}}}(\tilde{y})(0, \delta q; 0, \delta p).$$

The estimate $D_\theta \jmath(\tilde{y})\delta\theta = \mathcal{O}(|\delta\theta||\bar{\Pi}|) = \mathcal{O}(|\nu|^{1/2}|\delta y|_{\underline{\mu}})$ leads to the bound

$$Dj_{\dot{\underline{\xi}}}(\tilde{y})(0, \delta q; 0, \delta p) = Dj_{\dot{\underline{\xi}}}(\tilde{y})\delta y + \mathcal{O}\left(\left|\dot{\underline{\xi}}\right| |\nu|^{1/2}|\delta y|_{\underline{\mu}}\right) = \mathcal{O}\left(\left|\dot{\underline{\xi}}\right| (|\nu|^{1/2} + |\delta y|_{\underline{\mu}})|\delta y|_{\underline{\mu}}\right).$$

We now bound the contribution of the function $\check{h}_{\underline{\mu}}$ using the estimates

$$\check{h}_{\underline{\mu}}(y) - \check{h}_{\underline{\mu}}(\tilde{y}) + \frac{1}{\varepsilon}(D\check{h}_{\underline{\mu}}(y) - D\check{h}_{\underline{\mu}}(\tilde{y}))\mathbb{P}_{\tilde{y}}\dot{\tilde{y}} \leqslant |\xi \cdot (D_\theta \jmath(y) - D_\theta \jmath(\tilde{y}))\zeta|$$
$$= \mathcal{O}(|\nu|^{1/2}|\zeta||\delta y|_{\underline{\mu}}) \tag{35}$$

and

$$D\check{h}_{\underline{\mu}}(\tilde{y})\delta y = (D\check{h}_{\underline{\mu}}(\tilde{y}) - D\check{h}_{\underline{\mu}}(\bar{y}))\delta y = \mathcal{O}(|\tilde{y} - \bar{y}|_{\underline{\mu}}|\delta y|_{\underline{\mu}}). \tag{36}$$

Combining (32)–(36), we see that

$$(Dh_{\underline{\mu}}(y) - Dh_{\underline{\mu}}(\tilde{y}))\left(\hat{p}(y) - \frac{1}{\varepsilon}\mathrm{res}(\tilde{y}, \varepsilon)\right) + h_{\underline{\mu}}(y) - h_{\underline{\mu}}(\tilde{y}) - Dh_{\underline{\mu}}(\tilde{y}))\delta y$$
$$= \mathcal{O}(\mathrm{e}^{-\varepsilon t}(\varepsilon \mathrm{e}^{-\varepsilon t/2} + |\delta y|_{\underline{\mu}})|\delta y|_{\underline{\mu}})$$

and hence (H7) holds for $\beta = \frac{1}{2}$ and $\kappa(\tau) = \kappa_0 \mathrm{e}^{-\tau}$ for some constant κ_0.

Finally, we turn to the technical hypotheses (H8). The function $\bar{f}_{\underline{\mu}}$ does not pull back to a $G_{\underline{\mu}}$ invariant function on $\mathcal{M}_{\underline{\mu}}$, due to the algebra element $(\zeta, 0)$. Hence we drop that term and set

$$f_{\underline{\mu}}(y) = -c(\underline{\mu})(q, p) = F_{\underline{\mu}}(u)$$

which is invariant under the full group action. Furthermore, $\mathrm{res}(\tilde{y}, \varepsilon) = \varepsilon X_{f_{\underline{\mu}}}(\tilde{y}) - \varepsilon X_{j_{(\zeta,0)}}(\tilde{y})$. Because $|\zeta| = \mathcal{O}(\varepsilon \mathrm{e}^{-\varepsilon t})$, we have $\varepsilon|X_{j_{(\zeta,0)}}(\tilde{y})|_{\underline{\mu}} = \mathcal{O}(\varepsilon|r(\tilde{y}, \varepsilon)|)$, so condition (H8.1) is satisfied. Clearly $f_{\underline{\mu}} \in LC^1_{\underline{\mu}}(\mathcal{D}_{\underline{\mu}}, \mathbb{R}; c(\underline{\mu}))$ and $X_{f_{\underline{\mu}}} \in LC^0_{\underline{\mu}}(\mathcal{D}_{\underline{\mu}}, \mathbb{R}^{12}; c(\underline{\mu}))$. Because $\mathrm{e}^{-\alpha\varepsilon t}/\mathrm{corr}(\mu) = \mathcal{O}(1)$, this implies that (H8.2a) is satisfied for $k_f(\tau) = c(\underline{\mu}(\tau/\varepsilon)) = \mathcal{O}(\mathrm{e}^{-\tau})$. Straightforward calculations show that $\frac{\mathrm{d}}{\mathrm{d}t}c(\underline{\mu}(t)) = \mathcal{O}(\varepsilon|\underline{\mu}|)$. Thus (H8.2c) is satisfied.

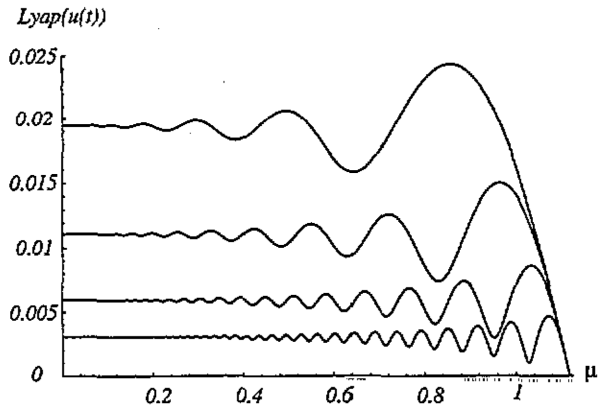

Figure 1. Square root of the scaled Lyapunov function $e^{\varepsilon t/2}\sqrt{L(u(t))}$ against the absolute value of the momentum $|\underline{\mu}(t)|$. We started at a relative equilibrium with $|\mu| = 1$ and $|\nu| = \frac{1}{2}$. We have depicted the evolution for $\varepsilon = 0.02, 0.04, 0.08$ and 0.16. Note the ratio 2 of the ε-values that comes into the picture. Furthermore, there is a striking period doubling if ε is divided by 2. We do not have a good explanation for this phenomenon yet.

5.2. Approximation with relative equilibria

Having verified all of the hypotheses, we can apply theorem 4 to show that the curve (27) of relative equilibria is a good approximation of the curve $u(t)$ if $u(0)$ is sufficiently close to the relative equilibrium $\bar{u}(0)$. In figure 1 we have illustrated this theorem by depicting the value of $\sqrt{e^{\varepsilon t}\, L(u(t))}$ against $|\underline{\mu}(t)|$ as it follows from numerical simulations for various values of ε and with an initial condition being a relative equilibrium.

Theorem 16. *Let $u(t)$ be a solution of the perturbed system (18) with potential V satisfying (23), initial condition $u(0)$ such that $L(u(0), \varepsilon) = \mathcal{O}(\varepsilon^2)$ for sufficiently small ε and $\underline{\mu}(0) \in \mathfrak{g}^*_{\mathrm{MRE}}$. Then $L(u(t), \varepsilon) = \mathcal{O}(\varepsilon^2\, e^{-\varepsilon t})$, for all $t \geqslant 0$. Hence if we let $(\bar{q}_{\mathrm{opt}}(t), \bar{p}_{\mathrm{opt}}(t))$ denote the optimal $G_{\underline{\mu}(t)}$-rotation of $(\bar{q}(t), \bar{p}(t))$, then*

$$|\Lambda(t)\nu \times \bar{\Lambda}(t)\nu|/|\nu|^2 = \mathcal{O}(\varepsilon) \qquad |q(t) - \bar{q}_{\mathrm{opt}}(t)| = \mathcal{O}(\varepsilon\, e^{-\varepsilon t/2})$$
$$|p(t) - \bar{p}_{\mathrm{opt}}(t)| = \mathcal{O}(\varepsilon\, e^{-\varepsilon t/2}) .$$

Note that the relative distances $|q(t) - \bar{q}_{\mathrm{opt}}(t)|/|\bar{q}_{\mathrm{opt}}(t)|$ and $|p(t) - \bar{p}_{\mathrm{opt}}(t)|/|\bar{p}_{\mathrm{opt}}(t)|$ are of order ε. As was indicated previously, $|\Lambda(t)\nu(t) \times \bar{\Lambda}(t)\nu(t)|/|\nu(t)|^2 = \text{constant} = \mathcal{O}(\varepsilon)$. Hence the estimate for $\Lambda(t)$ is sharp.

Acknowledgments

This research was started while GD was a visitor at the University of California at Santa Cruz and continued while the authors were visiting the Mathematical Sciences Research Institute in Berkeley. We thank both institutes for their hospitality.

References

[1] Abraham R and Marsden J E 1978 *Foundations of Mechanics* 2nd edn (Reading, MA: Benjamin Cummings)

[2] Bloch A, Krishnaprasad P S, Marsden J E and Ratiu T 1994 Dissipation induced instabilities *Ann. Inst. H. Poincaré* **11** 37–90

[3] Derks G 1992 Coherent structures in the dynamics of perturbed Hamiltonian systems *PhD Thesis* University of Twente

[4] Derks G and van Groesen E 1995 Dissipation in Hamiltonian systems: decaying cnoidal waves *SIAM J. Math. Anal.* to appear

[5] Derks G and Valkering T P 1992 Approximation in a damped Hamiltonian system by successive relative equilibria *Japan. J. Indus. Appl. Math.* 9 141–61

[6] Guillemin V and Sternberg S 1984 *Symplectic Techniques in Physics* (Cambridge: Cambridge University Press)

[7] Holm D D, Marsden J E, Ratiu T and Weinstein A 1985 Nonlinear stability of fluid and plasma equilibria *Phys. Rep.* **123** 1–116

[8] Krishnaprasad P S and Marsden Jerrold E 1987 Hamiltonian structures and stability for rigid bodies with flexible attachments *Arch. Rat. Mech. Anal.* **98** 137–58

[9] Lebovitz N R and Neishtadt A 1994 Slow evolution in perturbed Hamiltonian systems *Stud. Appl. Math.* **92** 127–44

[10] Marsden J E 1992 *Lectures on Mechanics* (*London Mathematical Society Lecture Note Series*) vol 174 (Cambridge: Cambridge University Press)

[11] Marsden J E and Weinstein A 1974 Reduction of symplectic manifolds with symmetry *Rep. Math. Phys.* **5** 121–30

[12] Meyer K R 1973 Symmetries and integrals in mechanics *Dynamical Systems* ed M M Peixoto (New York: Academic) p 259–72

[13] Simo J C, Lewis D and Marsden J E 1991 Stability of relative equilibria. Part I: the reduced energy momentum method *Arch. Rat. Mech. Anal.* **115** 15–60

[14] Simo J C, Posbergh T A and Marsden J E 1991 Stability of relative equilibria. Part II: application to nonlinear elasticity *Arch. Rat. Mech. Anal.* **115** 60–100