Physica D 428 (2021) 133036

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

Second-order fast–slow dynamics of non-ergodic Hamiltonian systems: Thermodynamic interpretation and simulation

Matthias Klar [a,*], Karsten Matthies [a], Celia Reina [b], Johannes Zimmer [c]

[a] Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom
[b] Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia PA 19104, USA
[c] Department of Mathematics, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany

ARTICLE INFO

Article history:
Received 9 March 2021
Received in revised form 30 July 2021
Accepted 5 September 2021
Available online 29 September 2021
Communicated by Gary Froyland

Keywords:
Two-scale Hamiltonian
Asymptotic expansion
Coarse-graining
Far-from-equilibrium
Many-degrees-of-freedom interaction

ABSTRACT

A class of fast–slow Hamiltonian systems with potential U_ε describing the interaction of non-ergodic fast and slow degrees of freedom is studied. The parameter ε indicates the typical timescale ratio of the fast and slow degrees of freedom. It is known that the Hamiltonian system converges for $\varepsilon \to 0$ to a homogenised Hamiltonian system. We study the situation where ε is small but positive.

First, we rigorously derive the second-order corrections to the homogenised (slow) degrees of freedom. They can be decomposed into explicitly given terms that oscillate rapidly around zero and terms that trace the average motion of the corrections, which are given as the solution to an inhomogeneous linear system of differential equations.

Then, we analyse the energy of the fast degrees of freedom expanded to second-order from a thermodynamic point of view. In particular, we define and expand to second-order a temperature, an entropy and external forces and show that they satisfy to leading-order, as well as on average to second-order, thermodynamic energy relations akin to the first and second law of thermodynamics.

Finally, we analyse for a specific fast–slow Hamiltonian system the second-order asymptotic expansion of the slow degrees of freedom from a numerical point of view. Their approximation quality for short and long time frames and their total computation time are compared with those of the solution to the original fast–slow Hamiltonian system of similar accuracy.

1. Introduction

Many scientists in physics, chemistry and materials science resort to computer simulations to study real-world dynamical processes. These simulations open up the possibility to quickly and inexpensively iterate through different experimental setups, thus hugely reducing cost in the form of time and labour and allow a level of insight into small- and large-scale processes that were out of reach decades ago. In chemical physics or materials science, for example, scientists frequently analyse large scale molecular dynamics simulations to predict properties of large dynamical systems based on mathematical models that aim to describe the dynamical evolution of the constituent particles. In simulating these systems, one typically encounters two problems that severely impede their scalability. Firstly, the simulation of molecular structures requires a step size in the numerical integration scheme that ranges in the order of 10^{-15} seconds to accurately replicate the fast molecular vibrations in the system. Secondly, even small macroscopic systems of interest require the integration of a potentially large number of particles. Even more, the two problems often compound and pose a challenging obstacle in the scalability and utility of molecular dynamic simulations.

From an applications point of view, one is often not interested in analysing the exact evolution of the fast molecular vibrations, but in the slow conformal motion that embodies the macroscale dynamics of the system. Here lies an opportunity to bypass at least partly the scalability issues by advancing the understanding of these systems and a subsequent development of numerical integration schemes that describe only the average evolution of the dynamical system without fully resolving the small-scale vibrations.

Fast–slow Hamiltonian systems provide a simplified fundamental description of large-scale interacting particles systems, where the system's degrees of freedom evolve on different scales in time and space. They can be used, for example, to model the evolution

* Corresponding author.
E-mail addresses: M.Klar@bath.ac.uk (M. Klar), K.Matthies@bath.ac.uk (K. Matthies), creina@seas.upenn.edu (C. Reina), jz@ma.tum.de (J. Zimmer).

https://doi.org/10.1016/j.physd.2021.133036
0167-2789/

of molecules where the slow degrees of freedom represent the conformal motion and the fast degrees of freedom represent the high-frequency molecular vibrations [1]. There is a vast body of literature for averaging general dynamical systems, not necessarily of Hamiltonian type, for example, using Young measures [2]. Applications of such fast–slow systems arise, for example, in models of plasticity [3]. Recent work on averaging of Hamiltonian fast–slow systems and connections with adiabatic invariants include [4–6] and references therein. Similar averaging techniques are also relevant to understand equilibration in springy billiards [7]. For general references to averaging, we refer the reader to [8–10].

With a mathematical description of a dynamical system in the form of a fast–slow Hamiltonian system, we are able to derive the conformal motion through homogenisation in time. The theory laid out by Bornemann in [1] enables us to derive the homogenised evolution of a specific class of fast–slow Hamiltonian systems. More precisely, Bornemann considers a family of mechanical systems, parametrised by a scale parameter ε, whose Lagrangian is of the form

$$\mathscr{L}_\varepsilon(x,\dot{x}) = \tfrac{1}{2}\langle \dot{x},\dot{x}\rangle - W_\varepsilon(x), \qquad \dot{x}\in T_xM, \tag{1}$$

on a Euclidean configuration space $M=\mathbb{R}^m$ with a potential given by

$$W_\varepsilon(x) = V(x) + \varepsilon^{-2}U(x).$$

Here, the potential U characterises the fast dynamics of the system. By splitting the coordinates according to $x=(y,z)\in\mathbb{R}^n\times\mathbb{R}^r=\mathbb{R}^m$, where y represents the slow and z the fast degrees of freedom, Bornemann showed that system (1) converges as $\varepsilon\to 0$ to a system on a slow manifold $N=U^{-1}(0)$, governed by the Lagrangian

$$\mathscr{L}_{\mathrm{hom}}(x,\dot{x}) = \tfrac{1}{2}\langle \dot{x},\dot{x}\rangle - V(x) - U_{\mathrm{hom}}(x), \qquad \dot{x}\in T_xN, \tag{2}$$

where U_{hom} can be derived from the Hessian of U and the initial conditions of x.

System (2) describes the slow, leading-order dynamics of the original system (1). As such, it allows for the integration of the corresponding equations of motion with a larger step size than what would usually be required for the integration of the original system. This can speed up the numerical integration significantly.

A crucial aspect of approximating the fast–slow solution of system (1) by a slow solution of system (2) is that the approximation error depends on the scale parameter ε. This scale parameter is a critical element in the dynamics described by (1). It is determined by the underlying true natural system that the model aims to represent and indicates the ratio of the typical timescales of the fast and slow degrees of freedom. In the case of a very small ε, a description of the fast–slow solution of system (1) by a slow solution of system (2) might be an acceptable trade-off in order to deal with the scalability issue mentioned earlier. However, a problem arises if ε is small, so that the microscale oscillations severely affect the numerical integration, but not small enough so that the dynamics of system (1) cannot be sufficiently approximated by the homogenised dynamics given by system (2). In this case, the dynamics of the fast degrees of freedom contribute much more to the evolution of the whole system than in the case of a very small ε. For example, in [1, Chapter III §2], the author applies the homogenisation process to derive the conformal motion of a butane molecule, where in the united atom representation, the scale parameter is $\varepsilon\approx 0.25$, which cannot be considered as small. It is thus natural to extend the theory presented in [1] to describe the slow dynamics of the original system on a finer scale, potentially revealing microscale properties in the case of a scale parameter away from the limit $\varepsilon\to 0$. This line of research begins already in [1], where formal asymptotic expansions are derived in Appendix C.

A further step was developed in [11], where the authors derive a second-order asymptotic expansion to the solution of system (1) in the case of one fast and one slow degree of freedom, i.e., $n=r=1$. Although the model in [11] is rather simple and the fast subsystem is ergodic, the fast–slow character is sufficient to derive properties of the fine-scale dynamics that are characteristic for thermodynamic processes. More precisely, for $V(x)\equiv 0$ and $U(x)=\frac{1}{2}\omega^2(y)z^2$, where $\omega>0$ is a smooth frequency function, the thermodynamic character of the model in [11] becomes evident by analysing the fast subsystem, which models the dynamics of the fast degree of freedom z, as a motion that is perturbed by the evolution of the slow degree of freedom y. This setting allows an interpretation of the fast subsystem from a thermodynamic point of view. By applying the thermodynamic theory for ergodic Hamiltonian systems, first developed by Boltzmann and Gibbs [12], and later specified by Hertz [13], one derives expressions for temperature, entropy and external force in the fast subsystem. Utilising the second-order asymptotic expansion, one can determine the leading-order terms of these thermodynamic expressions and show that they satisfy a thermodynamic energy relation akin to the first and second law of thermodynamics. It turns out that the entropy expression to *leading-order* is *constant*, suggesting an interpretation of the leading-order dynamics as an adiabatic thermodynamic process. Remarkably, although away from the limit $\varepsilon\to 0$, one finds a similar energy relation for the averaged second-order terms of the expansion. Most importantly, the entropy expression to second-order is *not* constant. The dynamics to second-order can therefore be interpreted as a non-adiabatic thermodynamic process.

In this article, we carry out a comparable study for the case of more than one fast and slow degrees of freedom, with the important difference that the higher dimensional fast subsystem is non-ergodic. We extend the theory presented in [1] and derive the second-order asymptotic expansion to the solution of system (1) in the case of an arbitrary finite number of fast and slow degrees of freedom, i.e., $n,r\in\mathbb{N}$. Specifically, we analyse the mechanical system (1) with a smooth potential $V=V(y)$ and $U(x)=\frac{1}{2}\langle H(y)z,z\rangle$, where $H(y)=\mathrm{diag}(\omega_1^2(y),\dots,\omega_r^2(y))$ for smooth frequency functions $\omega_\lambda>0$ ($\lambda=1,\dots,r$). Unlike in [11], we have to impose certain non-resonance conditions to derive the second-order asymptotic expansion. Following the strategy presented in [11], a key element in the derivation of the second-order asymptotic expansion is a transformation of the fast degrees of freedom into action–angle variables. By using weak convergence methods we show that the second-order asymptotic expansion of the ε-dependent transformed variables is given, for instance in the case of y_ε, as $y_\varepsilon = y_0+\varepsilon^2(\bar{y}_2+[y_2]^\varepsilon)+\varepsilon^2 y_3^\varepsilon$, where $y_3^\varepsilon\to 0$ in $C([0,T],\mathbb{R}^n)$ as $\varepsilon\to 0$. Here, the function y_0 is the leading-order term derived from system (2), the function $\bar{y}_2$ is the slow component of the second-order correction, which can be derived as the solution to an inhomogeneous linear system of differential equations, and the function $[y_2]^\varepsilon$ is the fast component of the second-order correction, which consists of explicitly given rapidly oscillating terms that converge weakly* to zero in $L^\infty([0,T],\mathbb{R}^n)$.

Furthermore, we interpret the dynamics of the fast subsystem, which is composed of the fast degrees of freedom, from a thermodynamic point of view. This is based on the thermodynamic theory for Hamiltonian systems formalised by Hertz [13] and used in [11]. More precisely, by decomposing the total energy E_ε into the energies $E_\varepsilon^{\parallel}$ and $E_\varepsilon^{\perp}$ such that $E_\varepsilon = E_\varepsilon^{\parallel}+E_\varepsilon^{\perp}$, we consider

$E_\varepsilon^\perp(z_\varepsilon, \dot{z}_\varepsilon; y_\varepsilon)$ as the energy describing the evolution of the fast degrees of freedom z_ε under a slow, external influence described by the dynamics of y_ε. As the fast subsystem is not necessarily ergodic, we follow along the lines of [14] and replace time averages in the thermodynamic theory by ensemble averages, i.e., averages over uniformly distributed initial values on the energy surface. With this modification, we apply Hertz' thermodynamic formalism and derive a temperature T_ε, an entropy S_ε and an external force F_ε for the fast subsystem. By applying the asymptotic expansion results from the first part of this article, we similarly expand the energy $E_\varepsilon^\perp = E_0^\perp + \varepsilon[E_1^\perp]^\varepsilon + \varepsilon^2\left(\bar{E}_2^\perp + [E_2^\perp]^\varepsilon\right) + \varepsilon^2 E_3^{\perp\varepsilon}$, the temperature $T_\varepsilon = T_0 + \mathcal{O}(\varepsilon)$, the entropy $S_\varepsilon = S_0 + \varepsilon[S_1]^\varepsilon + \varepsilon^2\left(\bar{S}_2 + [S_2]^\varepsilon\right) + \varepsilon^2 S_3^\varepsilon$ and the external force $F_\varepsilon = F_0 + \mathcal{O}(\varepsilon)$, where $E_3^{\perp\varepsilon}, S_3^\varepsilon \to 0$ in $C([0,T])$. We find that to leading-order the thermodynamic quantities satisfy an energy relation akin to the first and second law of thermodynamics (in the sense of Carathéodory [15])

$$dE_0^\perp = \sum_{j=1}^{n} F_0^j dy_0^j + T_0 dS_0.$$

In contrast to the work in [11], the leading-order entropy S_0 can be *constant* or *non-constant*, depending on the characteristics of the weighted frequency ratios $\theta_*^\lambda \omega_\lambda(y_0)/\omega_\mu(y_0)$ $(\lambda, \mu = 1, \dots, r)$. Here, we use the definition of the entropy as the logarithm of the phase space volume where the latter does not have to change slowly; the analysis shows that even in this situation, a meaningful thermodynamic setting exists. As a consequence, we interpret the dynamics to leading-order as an adiabatic or non-adiabatic thermodynamic process, respectively. Moreover, by considering the average dynamics to second-order for fixed y_0 and $p_0 = \dot{y}_0$, we similarly find, although away from the limit $\varepsilon \to 0$, a comparable energy relation of the form

$$d\bar{E}_2^\perp = \sum_{j=1}^{n} F_0^j d\bar{y}_2^j + T_0 d\bar{S}_2.$$

Likewise, with a non-constant second-order entropy expression $\bar{S}_2$, we can interpret the averaged second-order dynamics as a non-adiabatic thermodynamic process.

Finally, we analyse the viability of the second-order asymptotic expansion as a suitable approximation to the slow degrees of freedom of system (1) from a numerical point of view. More precisely, we choose a specific model from the class of fast–slow Hamiltonian systems represented by (1) and compare the numerical solution of y_ε with $y_0 + \varepsilon^2(\bar{y}_2 + [y_2]^\varepsilon)$ in terms of its short- and long-term approximation quality and computation time. The maximal time frame for which an approximation of y_ε can be considered sufficiently accurate significantly increases by using $y_0 + \varepsilon^2(\bar{y}_2 + [y_2]^\varepsilon)$ instead of y_0 alone. Moreover, we show that the computation of $y_0 + \varepsilon^2(\bar{y}_2 + [y_2]^\varepsilon)$ is up to two orders of magnitude faster (depending on the scale parameter ε) than a computation of y_ε to comparable accuracy as a solution to system (1). As described earlier, the reason is that fast oscillations severely affect the runtime for numerically computing y_ε from (1). In contrast, the problematic oscillatory term at second-order $[y_2]^\varepsilon$ is given explicitly, and the derivation of y_0 and $\bar{y}_2$ only require a numerical integration of two slow systems of differential equations, which can be solved, in parallel, using a relatively large step size.

An application of the theory presented in this article may not only improve large-scale molecular dynamics simulations. It can also find applications in cases where the homogenisation theory outlined above and related work as in [16] are applicable. Some examples are given by the description of quantum–classical models in quantum-chemistry [1], the problem of deriving the guiding centre motion in plasma physics [17] or, more recently, the derivation of a coarse-grained description of the coupled thermoelastic behaviour from an atomistic model in materials science [18].

Finally, we want to point out other thermodynamic analyses based on the fast–slow system governed by the Lagrangian (1). In [19], the authors extend system (1) by coupling the fast and slow degrees of freedom to an external Nosé–Hoover thermostat and analyse the thermodynamic equilibration of the system on the fast and slow scale. In a similar line of thought, the authors in [20] expand system (1) by embedding it into an external heat bath and subsequently analysing the resulting slow dynamics, analogous to the homogenisation procedure introduced above, in the limit $\varepsilon \to 0$.

1.1. Outline of the paper

In Section 2 we introduce the model problem, which establishes the foundation for the analysis in this article, and state necessary non-resonance conditions, which ensure that the subsequently derived second-order expansion of the solution to the model problem is well-defined. A summary of our main results is provided in Section 3. We start the analysis of the model problem by introducing a transformation of the fast degrees of freedom into action–angle variables in Section 4, where we also prove the existence and uniqueness of a solution of the transformed system. In Section 5 we introduce some notation that simplifies the governing equations of motion and derive the second-order asymptotic expansion for the transformed degrees of freedom. Subsequently, in Section 6, we define expressions for the temperature, the entropy and the external force for the fast subsystem and interpret the model from a thermodynamic point of view. For a test model, the global error for approximating y_ε by $y_0 + \varepsilon^2(\bar{y}_2 + [y_2]^\varepsilon)$ are analysed on short and long time intervals in Section 7, where we also compare the runtimes for computing y_ε, y_0 and $\bar{y}_2 + [y_2]^\varepsilon$. Section 8 provides a short conclusion of this article. In Appendix A we summarise how the thermodynamic expressions can be derived for the fast subsystem. Finally, in Appendix B we present some data on the computation times corresponding to the maximal step sizes used in the numerical simulations presented in this article.

2. The model problem

For a small scale parameter $0 < \varepsilon < \varepsilon_0 < \infty$, we study the family of mechanical systems given by the Lagrangian

$$\mathscr{L}_\varepsilon(x, \dot{x}) = \tfrac{1}{2}\langle \dot{x}, \dot{x}\rangle - W_\varepsilon(x), \qquad \dot{x} \in T_x M, \tag{3}$$

on a Euclidean configuration space $M = \mathbb{R}^m$. Here and in the following, $\langle \cdot, \cdot \rangle$ denotes Euclidean inner products and $|\cdot|$ denotes Euclidean norms. Splitting the coordinates according to $x = (y, z) \in \mathbb{R}^n \times \mathbb{R}^r = \mathbb{R}^m$, we specify, following [1], the potential $W_\varepsilon = V + \varepsilon^{-2} U$ by a smooth potential $V = V(y)$, which is assumed to be bounded from below and

$$U(x) = \tfrac{1}{2} \langle H(y)z, z \rangle \qquad \text{with} \qquad H(y) = \operatorname{diag}(\omega_1^2(y), \dots, \omega_r^2(y)). \tag{4}$$

We assume that the smooth functions $\omega_\lambda \in C^\infty(\mathbb{R}^n)$ are uniformly positive, i.e., there exists a constant $\omega_* > 0$ such that

$$\omega_\lambda(y) \geq \omega_*, \qquad y \in \mathbb{R}^n, \quad \lambda = 1, \dots, r. \tag{5}$$

A componentwise formulation of the equations of motion for the ε-dependent coordinates y_ε and z_ε in (3) yields

$$\ddot{y}_\varepsilon^j = -\partial_j V(y_\varepsilon) - \tfrac{1}{2}\varepsilon^{-2} \left\langle \partial_j H(y_\varepsilon) z_\varepsilon, z_\varepsilon \right\rangle, \qquad j = 1, \dots, n, \tag{6a}$$

$$\ddot{z}_\varepsilon = -\varepsilon^{-2} H(y_\varepsilon) z_\varepsilon. \tag{6b}$$

Moreover, we consider the ε-independent initial values

$$y_\varepsilon(0) = y_*, \qquad \dot{y}_\varepsilon(0) = p_*, \qquad z_\varepsilon(0) = 0, \qquad \dot{z}_\varepsilon(0) = u_*. \tag{7}$$

We notice that the energy E_ε of the system is independent of ε due to the particular choice $z_\varepsilon(0) = 0$,

$$E_\varepsilon = \tfrac{1}{2} |\dot{y}_\varepsilon|^2 + \tfrac{1}{2} |\dot{z}_\varepsilon|^2 + V(y_\varepsilon) + \varepsilon^{-2} U(y_\varepsilon, z_\varepsilon) = \tfrac{1}{2} |p_*|^2 + \tfrac{1}{2} |u_*|^2 + V(y_*) = E_*. \tag{8}$$

Remark. For the equations in (6)–(8) and below, we will simultaneously make use of the vector notation for the coordinates $y_\varepsilon \in \mathbb{R}^n$ and $z_\varepsilon \in \mathbb{R}^r$ (and related expressions) as well as their componentwise representation y_ε^j ($j = 1, \dots, n$) and z_ε^λ ($\lambda = 1, \dots, r$). The index in the superscript should not be confused with an exponent.

We are primarily interested in the evolution of the slow degrees of freedom y_ε^j ($j = 1, \dots, n$). The following theorem by Bornemann shows that y_ε converges in the limit $\varepsilon \to 0$ to a function y_0 which is given as the solution to a second-order differential equation.

Theorem 2.1 (*Bornemann, [1]*)**.** *For*

$$U_{\text{hom}}(y_0) = \sum_{\lambda=1}^{r} \theta_*^\lambda \omega_\lambda(y_0), \qquad \text{where} \qquad \theta_*^\lambda = \frac{|u_*^\lambda|^2}{2\omega_\lambda(y_*)}, \qquad \lambda = 1, \dots, r,$$

let y_0 be the solution to the second-order differential equation

$$\ddot{y}_0^j = -\partial_j V(y_0) - \partial_j U_{\text{hom}}(y_0), \qquad j = 1, \dots, n, \tag{9}$$

with initial values $y_0(0) = y_*$, $\dot{y}_0(0) = p_*$. Then, for every finite time interval $[0, T]$, we obtain the strong convergence

$$y_\varepsilon \to y_0 \quad \text{in} \quad C^1([0, T], \mathbb{R}^n)$$

and the weak convergences $\varepsilon^{-1} z_\varepsilon \overset{*}{\rightharpoonup} 0$ and $\dot{z}_\varepsilon \overset{*}{\rightharpoonup} 0$ in $L^\infty([0, T], \mathbb{R}^r)$.*

Theorem 2.1 shows that the family of mechanical systems (3) converges as $\varepsilon \to 0$ to a mechanical system which is again Hamiltonian.

2.1. Non-resonance conditions

As the interaction of multiple oscillating degrees of freedom can lead to resonance effects in the system, we will, similar to [1], impose suitable non-resonance conditions on the frequencies ω_λ to ensure that the second-order asymptotic expansions, which we will derive in Section 5, are well-defined. We say, referring to the definition stated for example in [21, Section 14.6], that a resonance of order $j \in \mathbb{N}$ at $y \in \mathbb{R}^n$ is given by the relation

$$\gamma_1 \omega_1(y) + \dots + \gamma_r \omega_r(y) = 0, \qquad |\gamma_1| + \dots + |\gamma_r| = j, \tag{10}$$

with integer coefficients $\gamma_\lambda \in \mathbb{Z}$ for $\lambda = 1, \dots, r$. Note that the non-degeneracy condition (5) implies that there is no resonance of order one.

Assumption 1. We assume that the homogenised solution in Theorem 2.1 is non-resonant of order two, i.e., we assume that

$$\gamma_1 \omega_1(y_0(t)) + \dots + \gamma_r \omega_r(y_0(t)) \neq 0, \qquad |\gamma_1| + \dots + |\gamma_r| = 2,$$

for all $t \in [0, T]$.

Assumption 2. Moreover, we assume that the homogenised solution in Theorem 2.1 is not flatly resonant up to order three. More precisely, we assume that

$$\frac{d}{dt} \left(\gamma_1 \omega_1(y_0(t_i)) + \dots + \gamma_r \omega_r(y_0(t_i)) \right) \neq 0, \qquad |\gamma_1| + \dots + |\gamma_r| \leq 3,$$

for all impact times $t_i \in [0, T]$ ($i \in I \subset \mathbb{N}$, I finite) such that the non-resonance condition (10) holds at $y_0(t_i)$.

We remark that Assumption 1 is intentionally chosen to simplify the derivation of the second-order asymptotic expansions (see remark following Lemma 5.9). Under these simplifications the assumption also ensures, that the second-order asymptotic expansions derived in Section 5 are well-defined. Assumption 2 is, analogous to [1], a necessary prerequisite for the theory developed below. It ensures that rapidly oscillating functions of the form $\exp\left(\pm i\varepsilon^{-1}\left(\omega_\lambda(y_\varepsilon)-\omega_\mu(y_\varepsilon)\right)\right)$ and $\exp\left(\pm i\varepsilon^{-1}\left(\omega_\lambda(y_\varepsilon)+\omega_\nu(y_\varepsilon)-\omega_\mu(y_\varepsilon)\right)\right)$ where $\lambda,\mu,\nu=1,\dots,r$, $\lambda\neq\mu$ converge weakly* to zero in $L^\infty([0,T])$. In [1] these functions appear due to interactions of the fast degrees of freedom caused by the structure of a more general potential $U(x)$ as well as a more general metric $\langle\cdot,\cdot\rangle$ and Assumption 2 is used to derive the leading-order asymptotic expansion of the system's degrees of freedom. Here, however, these functions appear only due to small-scale interactions in the second-order asymptotic expansions.

3. Summary of the main results

The goal of this article is to extend the theory developed in [1] by deriving the second-order asymptotic expansion rigorously for the solution of the equations of motion (6) and interpret the corresponding second-order asymptotic expansion of the energy (8) from a thermodynamic point of view. Note that the mechanical system (3) is not a classical thermodynamic system. In particular, the fast subsystem, consisting of the fast degrees of freedom z_ε^λ ($\lambda=1,\dots,r$), which we will consider in Section 6 as the thermodynamic part of the whole system, is in general not ergodic. Finally, we will discuss the numerical implications of the second-order asymptotic expansion of y_ε in terms of its approximation error and computational cost.

Our main findings in this article can be summarised as follows.

1. After transforming the rapidly oscillating degrees of freedom into action–angle variables $(z_\varepsilon,\dot z_\varepsilon)\mapsto(\theta_\varepsilon,\phi_\varepsilon)$, which also involves a transformation of the generalised momentum $\dot y_\varepsilon\mapsto p_\varepsilon$, we derive the second-order asymptotic expansion of $y_\varepsilon,p_\varepsilon,\theta_\varepsilon,\phi_\varepsilon$. This takes the form

$$\begin{aligned}
y_\varepsilon &= y_0+\varepsilon[\bar y_1]^\varepsilon+\varepsilon^2[\bar y_2]^\varepsilon+\varepsilon^2 y_3^\varepsilon,\\
p_\varepsilon &= p_0+\varepsilon[\bar p_1]^\varepsilon+\varepsilon^2[\bar p_2]^\varepsilon+\varepsilon^2 p_3^\varepsilon,\\
\theta_\varepsilon &= \theta_*+\varepsilon[\bar\theta_1]^\varepsilon+\varepsilon^2[\bar\theta_2]^\varepsilon+\varepsilon^2\theta_3^\varepsilon,\\
\phi_\varepsilon &= \phi_0+\varepsilon[\bar\phi_1]^\varepsilon+\varepsilon^2[\bar\phi_2]^\varepsilon+\varepsilon^2\phi_3^\varepsilon,
\end{aligned}$$

where for $i\in\{1,2\}$,

$$\begin{aligned}
&[\bar y_i]^\varepsilon := \bar y_i+[y_i]^\varepsilon\overset{*}{\rightharpoonup}\bar y_i \quad\text{in}\quad L^\infty([0,T],\mathbb{R}^n), && y_3^\varepsilon\to 0 \quad\text{in}\quad C([0,T],\mathbb{R}^n),\\
&[\bar p_i]^\varepsilon := \bar p_i+[p_i]^\varepsilon\overset{*}{\rightharpoonup}\bar p_i \quad\text{in}\quad L^\infty([0,T],\mathbb{R}^n), && p_3^\varepsilon\to 0 \quad\text{in}\quad C([0,T],\mathbb{R}^n),\\
&[\bar\theta_i]^\varepsilon := \bar\theta_i+[\theta_i]^\varepsilon\overset{*}{\rightharpoonup}\bar\theta_i \quad\text{in}\quad L^\infty([0,T],\mathbb{R}^r), && \theta_3^\varepsilon\to 0 \quad\text{in}\quad C([0,T],\mathbb{R}^r),\\
&[\bar\phi_i]^\varepsilon := \bar\phi_i+[\phi_i]^\varepsilon\overset{*}{\rightharpoonup}\bar\phi_i \quad\text{in}\quad L^\infty([0,T],\mathbb{R}^r), && \phi_3^\varepsilon\to 0 \quad\text{in}\quad C([0,T],\mathbb{R}^r).
\end{aligned}$$

In other words, for each degree of freedom the second-order asymptotic expansion is characterised – to leading-order by the theory developed in [1] (Theorem 2.1) – to ith order by a decomposition into a slow term, indicated by an overbar, which constitutes the average motion of the ith order expansion, and a fast term, indicated by square brackets, which oscillates rapidly and converges weakly* to zero — and by a residual term, indicated with a subscript three, which converges uniformly to zero. In particular, we show that

$$[\bar y_1]^\varepsilon=0,\qquad [\bar p_1]^\varepsilon=0,\qquad [\bar\theta_1]^\varepsilon=[\theta_1]^\varepsilon,\qquad [\bar\phi_1]^\varepsilon=0,$$

and that $(\bar\phi_2,\bar\theta_2,\bar y_2,\bar p_2)$ is given as the solution to an inhomogeneous linear system of differential equations (Theorem 5.2). Moreover, the rapidly oscillating functions $[\theta_1]^\varepsilon$, $[y_2]^\varepsilon$, $[p_2]^\varepsilon$, $[\theta_2]^\varepsilon$ and $[\phi_2]^\varepsilon$ are explicitly given in Definition 5.1.

2. In [13], Hertz formalises a thermodynamic theory for fast Hamiltonian systems which are perturbed by slow external agents. We regard the fast subsystem $(z_\varepsilon,\dot z_\varepsilon)$ as such a thermodynamic system, perturbed by the slow motion of $(y_\varepsilon,\dot y_\varepsilon)$. Since the fast subsystem is not ergodic, we follow along the lines of [14, Chapter 1.10] and replace the time average, which is an essential component in the thermodynamic theory, by the ensemble average, i.e., the average over uniformly distributed initial values on the energy surface (see Appendix A), and define, based on Hertz' formulation, a temperature T_ε, an entropy S_ε and an external force F_ε for the fast subsystem.

   In combination with the analytic result discussed under 1, we decompose the total energy E_ε into the energy associated with the fast subsystem $E_\varepsilon^\perp$ and its residual energy $E_\varepsilon^\parallel=E_\varepsilon-E_\varepsilon^\perp$, and expand, similar to above $E_\varepsilon^\perp$, $E_\varepsilon^\parallel$, T_ε, S_ε and F_ε into the form

$$\begin{aligned}
E_\varepsilon^\perp &= E_0^\perp+\varepsilon[\bar E_1^\perp]^\varepsilon+\varepsilon^2[\bar E_2^\perp]^\varepsilon+\varepsilon^2 E_3^{\perp\varepsilon},\\
E_\varepsilon^\parallel &= E_0^\parallel+\varepsilon[\bar E_1^\parallel]^\varepsilon+\varepsilon^2[\bar E_2^\parallel]^\varepsilon+\varepsilon^2 E_3^{\parallel\varepsilon},\\
S_\varepsilon &= S_0+\varepsilon[\bar S_1]^\varepsilon+\varepsilon^2[\bar S_2]^\varepsilon+\varepsilon^2 S_3^\varepsilon,\\
T_\varepsilon &= T_0+\mathcal{O}(\varepsilon),\\
F_\varepsilon &= F_0+\mathcal{O}(\varepsilon),
\end{aligned}$$

where for $i\in\{1,2\}$,

$$\begin{aligned}
&[\bar E_i^\perp]^\varepsilon := \bar E_i^\perp+[E_i^\perp]^\varepsilon\overset{*}{\rightharpoonup}\bar E_i^\perp \quad\text{in}\quad L^\infty([0,T]), && E_3^{\perp\varepsilon}\to 0 \quad\text{in}\quad C([0,T]),\\
&[\bar E_i^\parallel]^\varepsilon := \bar E_i^\parallel+[E_i^\parallel]^\varepsilon\overset{*}{\rightharpoonup}\bar E_i^\parallel \quad\text{in}\quad L^\infty([0,T]), && E_3^{\parallel\varepsilon}\to 0 \quad\text{in}\quad C([0,T]),\\
&[\bar S_i]^\varepsilon := \bar S_i+[S_i]^\varepsilon\overset{*}{\rightharpoonup}\bar S_i \quad\text{in}\quad L^\infty([0,T]), && S_3^\varepsilon\to 0 \quad\text{in}\quad C([0,T]).
\end{aligned}$$

The characterisation of the ith order expansion is similar to 1 and is already discussed, for the case of $n=r=1$, in [11].

In Section 6 we interpret these asymptotic expansions from a thermodynamic point of view. In particular, we show that, to leading-order, the dynamics can be interpreted as a thermodynamic process characterised by the energy relation

$$dE_0^{\perp} = \sum_{j=1}^{n} F_0^j dy_0^j + T_0 dS_0.$$

In contrast to the analysis in [11], we find, provided that $\theta_*^{\lambda} \neq 0$ for at least one $\lambda = 1, \dots, r$, that the entropy expression to leading-order is constant, $dS_0 = 0$, *if and only if* all pairwise weighted frequency ratios $\theta_*^{\lambda}\omega_\lambda(y_0)/\omega_\mu(y_0)$ $(\lambda, \mu = 1, \dots, r)$ are constant. In this case, the leading-order dynamics can be interpreted as an adiabatic thermodynamic process. Yet, if any of the weighted frequency ratios is non-constant, the entropy is non-constant and thus the leading-order dynamics can be interpreted as a non-adiabatic thermodynamic process. Here we use the definition of entropy given by Hertz in a context where the entropy is not necessarily the logarithm of an adiabatic invariant. Nevertheless, we show that a meaningful thermodynamic interpretation can be given.
Furthermore, we show that the averaged second-order dynamics, i.e., the dynamics in the weak* limit of the second-order terms, indicated by an overbar, represents for fixed (y_0, p_0) a non-adiabatic thermodynamic process with an averaged *non-constant entropy*, $d\bar{S}_2 \neq 0$, which also satisfies relations akin to equilibrium thermodynamics, despite being beyond the limit $\varepsilon \to 0$,

$$d\bar{E}_2^{\perp} = \sum_{j=1}^{n} F_0^j d\bar{y}_2^j + T_0 d\bar{S}_2.$$

Finally, we show in Theorem 6.1 that the evolution of $(\bar{y}_2, \bar{p}_2)$ is governed by equations which resemble Hamilton's canonical equations,

$$\frac{d\bar{y}_2}{dt} = \frac{\partial \bar{E}_2}{\partial p_0}, \qquad \frac{\partial \bar{p}_2}{dt} = -\frac{\partial \bar{E}_2}{\partial y_0},$$

for $\bar{E}_2 = \bar{E}_2^{\perp} + \bar{E}_2^{\parallel}$, which are complemented by the ε-independent initial values

$$\bar{y}_2(0) = -[y_2]^{\varepsilon}(0), \qquad \bar{p}_2(0) = -[p_2]^{\varepsilon}(0).$$

3. Finally, we compare in numerical simulations the second-order asymptotic expansion of the slow degrees of freedom $y_0 + \varepsilon^2(\bar{y}_2 + [y_2]^{\varepsilon})$ with simulations for y_ε of the original system (6). The latter is computationally expensive, as it requires a numerical integration of the fast degrees of freedom z_ε. To this end, we derive numerically the slow motion y_0 of the leading-order system (9) and the average motion $\bar{y}_2$ of the second-order system (23) and combine them with the explicitly given rapidly oscillating components $[y_2]^{\varepsilon}$ of the second-order expansion as specified in Definition 5.1. We find, depending on the value of the scale parameter ε, that the computation time for the second-order expansion is up to two orders of magnitude faster than the computation time for the slow degrees of freedom of the original system. Moreover, we show that $y_0 + \varepsilon^2(\bar{y}_2 + [y_2]^{\varepsilon})$ provides an approximation of y_ε which has significantly better global error bounds on long time intervals than an approximation by y_0 alone.

4. The model problem in action–angle variables

To study the dynamics of y_ε and z_ε on different scales, a detailed asymptotic analysis is required. Such an analysis was already presented for the model problem as introduced in Section 2 in the case of one fast and one slow degree of freedom (i.e., $n = r = 1$) in [11], which extends the analysis given in [1, Appendix C]. To derive the second-order asymptotic expansion of the solution to the model problem for arbitrary $n, r \in \mathbb{N}$, we analogously start by rephrasing the governing system of Newtonian equations (6) by transforming the fast degrees of freedom $(z_\varepsilon, \dot{z}_\varepsilon)$ into action–angle variables $(\theta_\varepsilon, \phi_\varepsilon)$.

We denote the canonical momenta corresponding to the positions $(y_\varepsilon, z_\varepsilon)$ as $(\eta_\varepsilon, \zeta_\varepsilon)$. Then, the equations of motion (6), together with the velocity relations

$$\dot{y}_\varepsilon = \eta_\varepsilon, \qquad \dot{z}_\varepsilon = \zeta_\varepsilon,$$

are given by the canonical equations of motion belonging to the energy function

$$E_\varepsilon = \frac{1}{2}|\eta_\varepsilon|^2 + \frac{1}{2}|\zeta_\varepsilon|^2 + V(y_\varepsilon) + \frac{1}{2}\varepsilon^{-2}\sum_{\lambda=1}^{r}\omega_\lambda^2(y_\varepsilon)(z_\varepsilon^{\lambda})^2.$$

The transformation $(z_\varepsilon, \zeta_\varepsilon) \mapsto (\theta_\varepsilon, \phi_\varepsilon)$ can be found by the theory of generating functions [22] as presented in [1, Appendix C]. For fixed y_ε, the generating function is given by

$$S_0(z_\varepsilon, \phi_\varepsilon; y_\varepsilon) = \frac{1}{2\varepsilon}\sum_{\lambda=1}^{r}\omega_\lambda(y_\varepsilon)(z_\varepsilon^{\lambda})^2 \cot(\varepsilon^{-1}\phi_\varepsilon^{\lambda}),$$

via $\zeta_\varepsilon = \partial S_0/\partial z_\varepsilon$ and $\theta_\varepsilon = -\partial S_0/\partial \phi_\varepsilon$. With this transformation, the fast degrees of freedom $(z_\varepsilon, \zeta_\varepsilon)$ can be written as

$$z_\varepsilon^{\lambda} = \varepsilon\sqrt{\frac{2\theta_\varepsilon^{\lambda}}{\omega_\lambda(y_\varepsilon)}}\sin(\varepsilon^{-1}\phi_\varepsilon^{\lambda}), \qquad \zeta_\varepsilon^{\lambda} = \sqrt{2\theta_\varepsilon^{\lambda}\omega_\lambda(y_\varepsilon)}\cos(\varepsilon^{-1}\phi_\varepsilon^{\lambda}).$$

It turns out, however, that the transformation $(z_\varepsilon, \zeta_\varepsilon) \mapsto (\theta_\varepsilon, \phi_\varepsilon)$ is symplectic only for fixed y_ε. To derive a transformation that preserves the symplectic structure on the whole phase-space, one introduces the generalised momenta p_ε through another transformation

$\eta_\varepsilon \mapsto p_\varepsilon$. To this end, we define the extended generating function $S(y_\varepsilon, p_\varepsilon, z_\varepsilon, \phi_\varepsilon) = p_\varepsilon^T y_\varepsilon + S_0(z_\varepsilon, \phi_\varepsilon; y_\varepsilon)$ which does not transform the position $y_\varepsilon = \partial S/\partial p_\varepsilon$, but changes the momentum η_ε such that the transformation remains symplectic on the whole phase-space. The missing transformation of the momentum η_ε is given componentwise for $j = 1, \dots, n$ by

$$\eta_\varepsilon^j = \frac{\partial S}{\partial y_\varepsilon^j} = p_\varepsilon^j + \varepsilon \sum_{\lambda=1}^{r} \frac{\theta_\varepsilon^\lambda \cdot \partial_j \omega_\lambda(y_\varepsilon)}{2\omega_\lambda(y_\varepsilon)} \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda).$$

By construction, the resulting transformation $(y_\varepsilon, \eta_\varepsilon; z_\varepsilon, \zeta_\varepsilon) \mapsto (y_\varepsilon, p_\varepsilon; \phi_\varepsilon, \theta_\varepsilon)$ is symplectic.

The energy can be expressed in the new coordinates as

$$\begin{aligned} E_\varepsilon = {} & \frac{1}{2}|p_\varepsilon|^2 + V(y_\varepsilon) + \sum_{\lambda=1}^{r} \theta_\varepsilon^\lambda \omega_\lambda(y_\varepsilon) + \varepsilon \sum_{j=1}^{n}\sum_{\lambda=1}^{r} \frac{\theta_\varepsilon^\lambda p_\varepsilon^j \cdot \partial_j \omega_\lambda(y_\varepsilon)}{2\omega_\lambda(y_\varepsilon)} \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ & + \frac{\varepsilon^2}{8} \sum_{j=1}^{n} \left(\sum_{\lambda=1}^{r} \frac{\theta_\varepsilon^\lambda \cdot \partial_j \omega_\lambda(y_\varepsilon)}{\omega_\lambda(y_\varepsilon)} \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \right)^2. \end{aligned}$$

Thus, by the canonical formalism, the equations of motion take the form

$$\dot\phi_\varepsilon^\lambda = \frac{\partial E_\varepsilon}{\partial \theta_\varepsilon^\lambda}, \qquad \dot\theta_\varepsilon^\lambda = -\frac{\partial E_\varepsilon}{\partial \phi_\varepsilon^\lambda}, \qquad \dot y_\varepsilon^j = \frac{\partial E_\varepsilon}{\partial p_\varepsilon^j}, \qquad \dot p_\varepsilon^j = -\frac{\partial E_\varepsilon}{\partial y_\varepsilon^j},$$

for $\lambda = 1, \dots, r$ and $j = 1, \dots, n$. After some calculations, we find that these equations are given by

$$\begin{aligned} \dot\phi_\varepsilon^\lambda = {} & \omega_\lambda(y_\varepsilon) + \varepsilon \sum_{j=1}^{n} \frac{p_\varepsilon^j \cdot \partial_j \omega_\lambda(y_\varepsilon)}{2\omega_\lambda(y_\varepsilon)} \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ & + \frac{\varepsilon^2}{8} \sum_{j=1}^{n}\sum_{\mu=1}^{r} \frac{\theta_\varepsilon^\mu \cdot \partial_j \omega_\mu(y_\varepsilon) \cdot \partial_j \omega_\lambda(y_\varepsilon)}{\omega_\mu(y_\varepsilon)\omega_\lambda(y_\varepsilon)} \left(\cos\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu - \phi_\varepsilon^\lambda\right)\right) - \cos\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu + \phi_\varepsilon^\lambda\right)\right)\right), \end{aligned} \tag{11a}$$

$$\begin{aligned} \dot\theta_\varepsilon^\lambda = {} & -\sum_{j=1}^{n} \frac{\theta_\varepsilon^\lambda p_\varepsilon^j \cdot \partial_j \omega_\lambda(y_\varepsilon)}{\omega_\lambda(y_\varepsilon)} \cos\left(2\varepsilon^{-1}\phi_\varepsilon^\lambda\right) \\ & - \frac{\varepsilon}{4} \sum_{j=1}^{n}\sum_{\mu=1}^{r} \frac{\theta_\varepsilon^\mu \theta_\varepsilon^\lambda \cdot \partial_j \omega_\mu(y_\varepsilon) \cdot \partial_j \omega_\lambda(y_\varepsilon)}{\omega_\mu(y_\varepsilon)\omega_\lambda(y_\varepsilon)} \left(\sin\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu - \phi_\varepsilon^\lambda\right)\right) + \sin\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu + \phi_\varepsilon^\lambda\right)\right)\right), \end{aligned} \tag{11b}$$

$$\dot y_\varepsilon^j = p_\varepsilon^j + \varepsilon \sum_{\lambda=1}^{r} \frac{\theta_\varepsilon^\lambda \cdot \partial_j \omega_\lambda(y_\varepsilon)}{2\omega_\lambda(y_\varepsilon)} \sin\left(2\varepsilon^{-1}\phi_\varepsilon^\lambda\right), \tag{11c}$$

$$\begin{aligned} \dot p_\varepsilon^j = {} & -\partial_j V(y_\varepsilon) - \sum_{\lambda=1}^{r} \theta_\varepsilon^\lambda \cdot \partial_j \omega_\lambda(y_\varepsilon) - \varepsilon \sum_{k=1}^{n}\sum_{\lambda=1}^{r} \frac{\theta_\varepsilon^\lambda p_\varepsilon^k}{2} \left(\frac{\partial_j \partial_k \omega_\lambda(y_\varepsilon)}{\omega_\lambda(y_\varepsilon)} - \frac{\partial_k \omega_\lambda(y_\varepsilon) \cdot \partial_j \omega_\lambda(y_\varepsilon)}{\omega_\lambda^2(y_\varepsilon)} \right) \sin\left(2\varepsilon^{-1}\phi_\varepsilon^\lambda\right) \\ & - \frac{\varepsilon^2}{8} \sum_{k=1}^{n}\sum_{\lambda=1}^{r}\sum_{\mu=1}^{r} \frac{\theta_\varepsilon^\lambda \theta_\varepsilon^\mu \cdot \partial_k \omega_\mu(y_\varepsilon)}{\omega_\mu(y_\varepsilon)} \left(\frac{\partial_j \partial_k \omega_\lambda(y_\varepsilon)}{\omega_\lambda(y_\varepsilon)} - \frac{\partial_k \omega_\lambda(y_\varepsilon) \cdot \partial_j \omega_\lambda(y_\varepsilon)}{\omega_\lambda^2(y_\varepsilon)} \right) \\ & \times \left(\cos\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu - \phi_\varepsilon^\lambda\right)\right) - \cos\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu + \phi_\varepsilon^\lambda\right)\right)\right). \end{aligned} \tag{11d}$$

The initial values as given in (7) transform to

$$\phi_\varepsilon(0) = 0, \qquad \theta_\varepsilon^\lambda(0) = \theta_*^\lambda = \frac{|u_*^\lambda|^2}{2\omega_\lambda(y_*)}, \qquad y_\varepsilon(0) = y_*, \qquad p_\varepsilon(0) = p_*. \tag{12}$$

4.1. Existence and uniqueness of a solution to the transformed model problem

Let us denote the right-hand side of (11) as $\mathcal{F}_\varepsilon \colon \mathbb{R}^{2m} \to \mathbb{R}^{2m}$. By assumption $\omega_\lambda \in C^\infty(\mathbb{R}^n)$ for $\lambda = 1, \dots, r$ and therefore $\mathcal{F}_\varepsilon \in C^\infty(\mathbb{R}^{2m}, \mathbb{R}^{2m})$ for $0 < \varepsilon < \varepsilon_0 < \infty$. In particular, $\mathcal{F}_\varepsilon$ is locally Lipschitz continuous. Hence, by the standard existence and uniqueness theory for ordinary differential equations (see for example [23]), there exists a $T > 0$ such that for fixed $0 < \varepsilon < \varepsilon_0 < \infty$ the initial value problem (11)–(12) has a unique solution

$$(\phi_\varepsilon, \theta_\varepsilon, y_\varepsilon, p_\varepsilon) \in C^\infty([0, T], \mathbb{R}^{2m}). \tag{13}$$

5. Asymptotic expansion

In this section, we rigorously derive the second-order asymptotic expansion of ϕ_ε, θ_ε, y_ε and p_ε. We will see, that the leading-order expansion follows directly from the evolution equations (11). To simplify these equations for the subsequent analysis, we introduce in Section 5.2 some suitable new notation. We then derive the first- and second-order asymptotic expansion in Section 5.3.

5.1. Leading-order expansion

We consider a sequence of solutions (13) for $\varepsilon \to 0$. The right-hand side of the evolution equations (11) is oscillatory and has rapidly oscillating terms of leading-order. As a consequence, the sequences $\{\dot\phi_\varepsilon\}$ and $\{\theta_\varepsilon\}$ are bounded in $C^{0,1}([0,T],\mathbb{R}^r)$, and the sequences $\{\dot y_\varepsilon\}$ and $\{\dot p_\varepsilon\}$ are bounded in $C^{0,1}([0,T],\mathbb{R}^n)$, while sequences of higher-order derivatives (in particular $\{\ddot\theta_\varepsilon\}$, which will thus require special attention in the later part of this analysis) become unbounded as $\varepsilon \to 0$. It follows from the extended Arzelà–Ascoli theorem [1, Chapter I §1] that we can extract a subsequence, not relabelled, and functions $\theta_0 \in C^{0,1}([0,T],\mathbb{R}^r)$, $\phi_0 \in C^{1,1}([0,T],\mathbb{R}^r)$ and $y_0, p_0 \in C^{1,1}([0,T],\mathbb{R}^n)$, such that

$$\phi_\varepsilon \to \phi_0 \quad \text{in} \quad C^1([0,T],\mathbb{R}^r), \qquad \ddot\phi_\varepsilon \overset{*}{\rightharpoonup} \ddot\phi_0 \quad \text{in} \quad L^\infty([0,T],\mathbb{R}^r), \tag{14a}$$

$$\theta_\varepsilon \to \theta_0 \quad \text{in} \quad C([0,T],\mathbb{R}^r), \qquad \dot\theta_\varepsilon \overset{*}{\rightharpoonup} \dot\theta_0 \quad \text{in} \quad L^\infty([0,T],\mathbb{R}^r), \tag{14b}$$

$$y_\varepsilon \to y_0 \quad \text{in} \quad C^1([0,T],\mathbb{R}^n), \qquad \ddot y_\varepsilon \overset{*}{\rightharpoonup} \ddot y_0 \quad \text{in} \quad L^\infty([0,T],\mathbb{R}^n), \tag{14c}$$

$$p_\varepsilon \to p_0 \quad \text{in} \quad C^1([0,T],\mathbb{R}^n), \qquad \ddot p_\varepsilon \overset{*}{\rightharpoonup} \ddot p_0 \quad \text{in} \quad L^\infty([0,T],\mathbb{R}^n). \tag{14d}$$

By taking the limit $\varepsilon \to 0$ in Eqs. (11a), (11c) and (11d) and the weak* limit in (11b) we deduce that

$$\dot\phi_0^\lambda = \omega_\lambda(y_0), \qquad \dot\theta_0^\lambda = 0, \qquad \dot y_0^j = p_0^j, \qquad \dot p_0^j = -\partial_j V(y_0) - \sum_{\lambda=1}^r \theta_*^\lambda \cdot \partial_j \omega_\lambda(y_0),$$

for $\lambda = 1,\dots,r$ and $j = 1,\dots,n$, and in particular that $\theta_0^\lambda \equiv \theta_*^\lambda$ (compare with (12)). Moreover, since the right-hand side of the limit equation

$$\ddot y_0^j = -\partial_j V(y_0) - \sum_{\lambda=1}^r \theta_*^\lambda \cdot \partial_j \omega_\lambda(y_0)$$

does not depend on a chosen subsequence, we can discard the extraction of a subsequence altogether (see [1, Principle 5, Chapter I §1]). Note that the above convergence results extend Theorem 2.1.

5.2. Reformulation of the governing equations

It will be convenient to introduce some notation to simplify the system of differential equations (11). To this end, we define for $f \in C^\infty(\mathbb{R}^n)$, where $f = f(y)$ and $y \in C^\infty([0,T],\mathbb{R}^n)$, the expression

$$D_t^k D_j^l f := \frac{d^k}{dt^k}\frac{\partial^l f}{\partial y_j^l},$$

for $k, l \in \mathbb{N}_0$ and $j = 1,\dots,n$. We will often apply this notation in combination with the function

$$L_\varepsilon^\lambda := \log(\omega_\lambda(y_\varepsilon)),$$

where $\lambda = 1,\dots,r$. Then, we can conveniently write, for instance,

$$DL_\varepsilon^\lambda = \sum_{j=1}^n D_j L_\varepsilon^\lambda \cdot e_j \qquad \text{or} \qquad D_t L_\varepsilon^\lambda = \left\langle \dot y_\varepsilon, DL_\varepsilon^\lambda \right\rangle = \sum_{j=1}^n \dot y_\varepsilon^j \cdot D_j L_\varepsilon^\lambda, \tag{15}$$

with e_j as the jth standard basis vector in $\mathbb{R}^n$. With these definitions, the equations in (11) read

$$\begin{aligned}\dot\phi_\varepsilon^\lambda &= \omega_\lambda(y_\varepsilon) + \frac{\varepsilon}{2}\left\langle p_\varepsilon, DL_\varepsilon^\lambda\right\rangle \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ &\quad + \frac{\varepsilon^2}{8}\sum_{\mu=1}^r \theta_\varepsilon^\mu \left\langle DL_\varepsilon^\lambda, DL_\varepsilon^\mu\right\rangle \left(\cos\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu - \phi_\varepsilon^\lambda\right)\right) - \cos\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu + \phi_\varepsilon^\lambda\right)\right)\right),\end{aligned} \tag{16a}$$

$$\begin{aligned}\dot\theta_\varepsilon^\lambda &= -\theta_\varepsilon^\lambda \left\langle p_\varepsilon, DL_\varepsilon^\lambda\right\rangle \cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ &\quad - \frac{\varepsilon}{4}\sum_{\mu=1}^r \theta_\varepsilon^\mu\theta_\varepsilon^\lambda \left\langle DL_\varepsilon^\lambda, DL_\varepsilon^\mu\right\rangle \left(\sin\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu - \phi_\varepsilon^\lambda\right)\right) + \sin\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu + \phi_\varepsilon^\lambda\right)\right)\right),\end{aligned} \tag{16b}$$

$$\dot y_\varepsilon^j = p_\varepsilon^j + \frac{\varepsilon}{2}\sum_{\lambda=1}^n \theta_\varepsilon^\lambda \cdot D_j L_\varepsilon^\lambda \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda), \tag{16c}$$

$$\begin{aligned}\dot p_\varepsilon^j &= -D_j V(y_\varepsilon) - \sum_{\lambda=1}^r \theta_\varepsilon^\lambda \cdot D_j\omega_\lambda(y_\varepsilon) - \frac{\varepsilon}{2}\sum_{\lambda=1}^r \theta_\varepsilon^\lambda \left\langle p_\varepsilon, DD_j L_\varepsilon^\lambda\right\rangle \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ &\quad - \frac{\varepsilon^2}{8}\sum_{\lambda=1}^r\sum_{\mu=1}^r \theta_\varepsilon^\lambda\theta_\varepsilon^\mu \left\langle DL_\varepsilon^\mu, DD_j L_\varepsilon^\lambda\right\rangle \left(\cos\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu - \phi_\varepsilon^\lambda\right)\right) - \cos\left(2\varepsilon^{-1}\left(\phi_\varepsilon^\mu + \phi_\varepsilon^\lambda\right)\right)\right).\end{aligned} \tag{16d}$$

Moreover, solving (16c) with respect to p^j_ε and inserting the result into (16a), (16b) and (16d) brings the equations of motion to their final form

$$\dot\phi^\lambda_\varepsilon = \omega_\lambda(y_\varepsilon) + \frac{\varepsilon}{2} D_t L^\lambda_\varepsilon \sin(2\varepsilon^{-1}\phi^\lambda_\varepsilon), \tag{17a}$$

$$\dot\theta^\lambda_\varepsilon = -\theta^\lambda_\varepsilon \cdot D_t L^\lambda_\varepsilon \cos(2\varepsilon^{-1}\phi^\lambda_\varepsilon), \tag{17b}$$

$$\dot y^j_\varepsilon = p^j_\varepsilon + \frac{\varepsilon}{2}\sum_{\lambda=1}^r \theta^\lambda_\varepsilon \cdot D_j L^\lambda_\varepsilon \sin(2\varepsilon^{-1}\phi^\lambda_\varepsilon), \tag{17c}$$

$$\dot p^j_\varepsilon = -D_j V(y_\varepsilon) - \sum_{\lambda=1}^r \theta^\lambda_\varepsilon \cdot D_j\omega_\lambda(y_\varepsilon) - \frac{\varepsilon}{2}\sum_{\lambda=1}^r \theta^\lambda_\varepsilon \cdot D_t D_j L^\lambda_\varepsilon \sin(2\varepsilon^{-1}\phi^\lambda_\varepsilon). \tag{17d}$$

5.3. First- and second-order expansion

We now define functions that will appear throughout this work and then state the first main result 1.

Definition 5.1. Let $(\phi_\varepsilon, \theta_\varepsilon, y_\varepsilon, p_\varepsilon)$ be the solution to (11)–(12) and $(\phi_0, \theta_0, y_0, p_0)$ be as in (14). With Assumption 1 and the notation introduced above we define for $\lambda = 1, \dots, r$ and $j = 1, \dots, n$ the functions

$$\theta_1^{\lambda\varepsilon} := \frac{\theta^\lambda_\varepsilon - \theta^\lambda_*}{\varepsilon}, \qquad \phi_2^{\lambda\varepsilon} := \frac{\phi^\lambda_\varepsilon - \phi^\lambda_0}{\varepsilon^2}, \qquad y_2^{j\varepsilon} := \frac{y^j_\varepsilon - y^j_0}{\varepsilon^2}, \qquad p_2^{j\varepsilon} := \frac{p^j_\varepsilon - p^j_0}{\varepsilon^2}, \qquad \theta_2^{\lambda\varepsilon} := \frac{\theta_1^{\lambda\varepsilon} - [\theta_1^\lambda]^\varepsilon}{\varepsilon},$$

$$[\theta_1^\lambda]^\varepsilon := -\frac{\theta^\lambda_* \cdot D_t L^\lambda_0}{2\omega_\lambda(y_0)} \sin(2\varepsilon^{-1}\phi^\lambda_0), \qquad [\phi_2^\lambda]^\varepsilon := -\frac{D_t L^\lambda_0}{4\omega_\lambda(y_0)} \cos(2\varepsilon^{-1}\phi^\lambda_0),$$

$$[y_2^j]^\varepsilon := -\sum_{\lambda=1}^r \frac{\theta^\lambda_* \cdot D_j L^\lambda_0}{4\omega_\lambda(y_0)} \cos(2\varepsilon^{-1}\phi^\lambda_0), \qquad [p_2^j]^\varepsilon := \sum_{\lambda=1}^r \frac{d}{dt}\left(\frac{\theta^\lambda_* \cdot D_j L^\lambda_0}{4\omega_\lambda(y_0)}\right) \cos(2\varepsilon^{-1}\phi^\lambda_0),$$

and

$$\begin{aligned}[\theta_2^\lambda]^\varepsilon := {} & \sum_{\mu=1}^r \frac{\theta^\lambda_*\theta^\mu_* \langle D\omega_\mu(y_0), DL^\lambda_0\rangle}{4\omega_\lambda^2(y_0)} \cos(2\varepsilon^{-1}\phi^\lambda_0) - \frac{\theta^\lambda_* \langle D^2 L^\lambda_0 \dot y_0, \dot y_0\rangle}{4\omega_\lambda^2(y_0)} \cos(2\varepsilon^{-1}\phi^\lambda_0) + \frac{\theta^\lambda_* (D_t L^\lambda_0)^2}{4\omega_\lambda^2(y_0)} \cos(2\varepsilon^{-1}\phi^\lambda_0) \\ & + \frac{(\theta^\lambda_*)^2 |DL^\lambda_0|^2}{16\omega_\lambda(y_0)} \cos(4\varepsilon^{-1}\phi^\lambda_0) - \frac{\theta^\lambda_* \cdot D_t L^\lambda_0}{\omega_\lambda(y_0)} \bar\phi_2^\lambda \cos(2\varepsilon^{-1}\phi^\lambda_0) + \frac{\theta^\lambda_* \langle DV(y_0), DL^\lambda_0\rangle}{4\omega_\lambda^2(y_0)} \cos(2\varepsilon^{-1}\phi^\lambda_0) \\ & + \sum_{\substack{\mu=1\\ \mu\neq\lambda}}^r \frac{\theta^\lambda_*\theta^\mu_* \langle DL^\mu_0, DL^\lambda_0\rangle}{8} \left\{ \frac{\cos(2\varepsilon^{-1}(\phi^\mu_0 - \phi^\lambda_0))}{\omega_\mu(y_0) - \omega_\lambda(y_0)} + \frac{\cos(2\varepsilon^{-1}(\phi^\mu_0 + \phi^\lambda_0))}{\omega_\mu(y_0) + \omega_\lambda(y_0)} \right\}. \end{aligned}$$

Theorem 5.2. *The functions specified in Definition 5.1 satisfy*

$$\theta^\varepsilon_1 - [\theta_1]^\varepsilon \to 0 \quad in \quad C([0,T],\mathbb{R}^r) \qquad \frac{d}{dt}\left(\theta^\varepsilon_1 - [\theta_1]^\varepsilon\right) \overset{*}{\rightharpoonup} 0 \quad in \quad L^\infty([0,T],\mathbb{R}^r), \tag{18}$$

$$\phi^\varepsilon_2 - [\phi_2]^\varepsilon \to \bar\phi_2 \quad in \quad C([0,T],\mathbb{R}^r) \qquad \frac{d}{dt}\left(\phi^\varepsilon_2 - [\phi_2]^\varepsilon\right) \overset{*}{\rightharpoonup} \frac{d\bar\phi_2}{dt} \quad in \quad L^\infty([0,T],\mathbb{R}^r), \tag{19}$$

$$y^\varepsilon_2 - [y_2]^\varepsilon \to \bar y_2 \quad in \quad C([0,T],\mathbb{R}^n) \qquad \frac{d}{dt}\left(y^\varepsilon_2 - [y_2]^\varepsilon\right) \overset{*}{\rightharpoonup} \frac{d\bar y_2}{dt} \quad in \quad L^\infty([0,T],\mathbb{R}^n), \tag{20}$$

$$p^\varepsilon_2 - [p_2]^\varepsilon \to \bar p_2 \quad in \quad C([0,T],\mathbb{R}^n) \qquad \frac{d}{dt}\left(p^\varepsilon_2 - [p_2]^\varepsilon\right) \overset{*}{\rightharpoonup} \frac{d\bar p_2}{dt} \quad in \quad L^\infty([0,T],\mathbb{R}^n), \tag{21}$$

and

$$\theta^\varepsilon_2 - [\theta_2]^\varepsilon \to \bar\theta_2 \quad in \quad C([0,T],\mathbb{R}^r), \tag{22}$$

where $(\bar\phi_2, \bar\theta_2, \bar y_2, \bar p_2)$ is the unique solution to the inhomogeneous linear system of differential equations

$$\frac{d\bar\phi^\lambda_2}{dt} = \langle D\omega_\lambda(y_0), \bar y_2\rangle + \frac{\theta^\lambda_* |DL^\lambda_0|^2}{8} - \frac{(D_t L^\lambda_0)^2}{8\omega_\lambda(y_0)}, \tag{23a}$$

$$\frac{d\bar\theta^\lambda_2}{dt} = \frac{d}{dt}\frac{\theta^\lambda_* (D_t L^\lambda_0)^2}{8\omega_\lambda^2(y_0)}, \tag{23b}$$

$$\frac{d\bar y^j_2}{dt} = \bar p^j_2 - \sum_{\lambda=1}^r \frac{\theta^\lambda_* \cdot D_j L^\lambda_0 \cdot D_t L^\lambda_0}{4\omega_\lambda(y_0)}, \tag{23c}$$

$$\frac{d\bar p^j_2}{dt} = -\langle \bar y_2, DD_j V(y_0)\rangle - \sum_{\lambda=1}^r \bar\theta^\lambda_2 \cdot D_j\omega_\lambda(y_0) - \sum_{\lambda=1}^r \theta^\lambda_* \langle \bar y_2, DD_j\omega_\lambda(y_0)\rangle - \sum_{\lambda=1}^r \frac{(\theta^\lambda_*)^2 \langle DL^\lambda_0, DD_j L^\lambda_0\rangle}{8} + \sum_{\lambda=1}^r \frac{\theta^\lambda_* \cdot D_t D_j L^\lambda_0 \cdot D_t L^\lambda_0}{4\omega_\lambda(y_0)}, \tag{23d}$$

for $\lambda = 1, \dots, r$ and $j = 1, \dots, n$, with ε-independent initial values

$$\bar{\phi}_2(0) = -[\phi_2]^\varepsilon(0), \qquad \bar{\theta}_2(0) = -[\theta_2]^\varepsilon(0), \qquad \bar{y}_2(0) = -[y_2]^\varepsilon(0), \qquad \bar{p}_2(0) = -[p_2]^\varepsilon(0). \tag{24}$$

5.4. Proof of Theorem 5.2

The proof of Theorem 5.2 will use the following Lemmas 5.3 to 5.12. We start by sketching the general strategy of the proof.

Theorem 5.2 states that the first- and second-order asymptotic expansions of ϕ_ε, θ_ε, y_ε and p_ε can be decomposed into rapidly oscillating terms $[\theta_1]^\varepsilon$, $[\phi_2]^\varepsilon$, $[y_2]^\varepsilon$, $[p_2]^\varepsilon$ and $[\theta_2]^\varepsilon$, which converge weakly* to zero, and slowly evolving terms $\bar{\phi}_2$, $\bar{y}_2$, $\bar{p}_2$ and $\bar{\theta}_2$, which describe the average motion of the second-order expansions and are given as the solution to an inhomogeneous linear system of ordinary differential equations.

To derive these second-order asymptotic expansions, we specified in Definition 5.1 the scaled first-order residual function θ_1^ε and the scaled second-order residual functions ϕ_2^ε, y_2^ε, p_2^ε and θ_2^ε by subtracting the leading- and first-order asymptotic expansion terms from the original solution to the model problem and by scaling these residual terms to appropriate order. The functions ϕ_2^ε, y_2^ε, p_2^ε and θ_2^ε carry all the information about the system's second-order asymptotic expansion in their leading-order expression. We thus analyse the limit $\varepsilon \to 0$ of these terms.

In the proof of Theorem 5.2, we will repeatedly integrate by parts, which requires us to regularly divide by $\dot{\phi}_\varepsilon^\lambda$ and $\dot{\phi}_\varepsilon^\lambda - \dot{\phi}_\varepsilon^\mu$ ($\lambda \neq \mu$). Lemma 5.3 ensures that the resulting terms are well-defined, provided that the scale parameter ε is small enough.

As the model problem is highly oscillatory, the interacting degrees of freedom can exhibit resonances of different types. Lemmas 5.4–5.6 clarify how the interaction of a generic function u_ε with a rapidly oscillating function $\exp(i\varepsilon^{-1}\psi_\varepsilon)$ affects their interaction in the limit $\varepsilon \to 0$. Here, u_ε and ψ_ε are representatives of functions that appear throughout the proof of Theorem 5.2. Lemmas 5.4–5.6 are used in the derivation of the weak* limit of specific rapidly oscillating functions under the non-resonance Assumptions 1 and 2.

Similarly, Lemma 5.7 provides information about the uniform convergence of the term $u_\varepsilon \exp(i\varepsilon^{-1}\psi_\varepsilon) - u_0 \exp(i\varepsilon^{-1}\psi_0)$, which is a representation of functions that appear throughout the proof of Theorem 5.2. Here, $u_\varepsilon \exp(i\varepsilon^{-1}\psi_\varepsilon)$ is rapidly oscillating at leading-order. By subtracting the leading-order term $u_0 \exp(i\varepsilon^{-1}\psi_0)$, their difference converges uniformly under certain convergence assumptions on u_ε and ψ_ε.

In Lemmas 5.8 and 5.9 we show that the sequences of scaled residual functions $\{\theta_1^\varepsilon\}$, $\{\phi_2^\varepsilon\}$ and $\{\theta_2^\varepsilon\}$ are bounded in $L^\infty([0,T], \mathbb{R}^r)$, and $\{y_2^\varepsilon\}$ and $\{p_2^\varepsilon\}$ are bounded in $L^\infty([0,T], \mathbb{R}^n)$. This is a necessary prerequisite for the analysis of the first- and second-order asymptotic expansion.

In general, the rapidly oscillating terms $[\theta_1]^\varepsilon$, $[\phi_2]^\varepsilon$, $[y_2]^\varepsilon$, $[p_2]^\varepsilon$ and $[\theta_2]^\varepsilon$, which do *not* converge in the limit $\varepsilon \to 0$, can be found through integration by parts. To find the evolution equation for the averaged second-order expansion terms $\bar{\phi}_2$, $\bar{y}_2$, $\bar{p}_2$ and $\bar{\theta}_2$, we analyse in Lemmas 5.10 and 5.11 the time derivatives of the terms $\phi_2^\varepsilon - [\phi_2]^\varepsilon$, $y_2^\varepsilon - [y_2]^\varepsilon$, $p_2^\varepsilon - [p_2]^\varepsilon$ and $\theta_2^\varepsilon - [\theta_2]^\varepsilon$. They carry information about the time derivative of $\bar{\phi}_2$, $\bar{y}_2$, $\bar{p}_2$ and $\bar{\theta}_2$ in their leading-order asymptotic expansion. Alaoglu's theorem [1, Principle 3] and the extended Arzelà–Ascoli theorem [1, Principle 4] justify the extraction of a subsequence such that in the weak* limit an evolution equation for the average dynamics at second-order emerges. However, since the evolution equation has a unique solution, Lemma 5.12 implies that the extraction of a subsequence can be discarded altogether, meaning the limit holds for the whole sequence.

The following lemmas collectively proof Theorem 5.2. They are stated separately for reference but should be understood in the context of Theorem 5.2. As mentioned earlier, the problem presented in Section 2 extends the model in [11]. More precisely, it generalises the model in [11] in two ways. Firstly, by describing the interaction of r fast and n slow degrees of freedom ($n, r \in \mathbb{N}$) instead of the interaction of one fast and one slow degree of freedom. This requires us to impose certain non-resonance conditions. Secondly, the model in this article includes a slow potential $V = V(y)$ which is absent in [11]. These generalisations make the following proof much more involved, yet it mimics at its core the proof as presented in [11]. As such, some of the following preparatory lemmas, with model-specific alterations, can be found in [11]. Nevertheless, we will state and prove these lemmas here for the reader's convenience.

Lemma 5.3 (*Similar to Lemma 3.4 in [11]*)**.** *There exist constants $0 < C < \infty$ and $0 < \varepsilon_0 < \infty$ where $\varepsilon_0 = \varepsilon_0(\phi_*, \theta_*, y_*, p_*, \omega, C)$ such that $0 < C \leq \dot{\phi}_\varepsilon^\lambda$ for $\lambda = 1, \dots, r$ and $0 < C \leq |\dot{\phi}_\varepsilon^\lambda - \dot{\phi}_\varepsilon^\mu|$ for $\lambda, \mu = 1, \dots, r$, $\lambda \neq \mu$, for all $0 < \varepsilon < \varepsilon_0$ small enough.*

Proof. The claim follows directly from Assumption 1 and (17a). □

Remark. Henceforth, we assume that $0 < \varepsilon < \varepsilon_0$ is small enough so that the statements of Lemma 5.3 apply.

Lemma 5.4 (*Lemma 3.5 in [11]*)**.** *Let $\{u_\varepsilon\}$ be a bounded sequence in $C^{0,1}([0,T])$ and $\{\psi_\varepsilon\}$ be a bounded sequence in $C^{1,1}([0,T])$ with $0 < C \leq \dot{\psi}_\varepsilon$. Then, for all $a, b \in [0,T]$:*

$$\int_a^b u_\varepsilon \sin(\varepsilon^{-1}\psi_\varepsilon)\,\mathrm{d}t = \mathcal{O}(\varepsilon), \qquad \int_a^b u_\varepsilon \cos(\varepsilon^{-1}\psi_\varepsilon)\,\mathrm{d}t = \mathcal{O}(\varepsilon).$$

Proof. Integration by parts gives for $0 < \varepsilon < \varepsilon_0$ small enough

$$\left|\int_a^b u_\varepsilon \exp\left(\frac{i\psi_\varepsilon}{\varepsilon}\right)\,\mathrm{d}t\right| \leq \varepsilon \left|\frac{u_\varepsilon(a)}{\dot{\psi}_\varepsilon(a)}\right| + \varepsilon \left|\frac{u_\varepsilon(b)}{\dot{\psi}_\varepsilon(b)}\right| + \varepsilon \left|\int_a^b \frac{d}{dt}\left(\frac{u_\varepsilon}{\dot{\psi}_\varepsilon}\right) \exp\left(\frac{i\psi_\varepsilon}{\varepsilon}\right)\,\mathrm{d}t\right| = \mathcal{O}(\varepsilon).$$

The claim follows by considering the real and imaginary parts separately and the isometric isomorphism $C^{k-1,1}([0,T]) \cong W^{k,\infty}([0,T])$ (see [24, p. 154]). □

Lemma 5.5. *Let $u_0 \in C^2([0,T])$ and $\psi_0 \in C^3([0,T])$. Let $\{u_\varepsilon\}$ be a sequence in $C^2([0,T])$ and $\{\psi_\varepsilon\}$ be a sequence in $C^3([0,T])$ such that the sequences $\{\varepsilon^{-1}(u_\varepsilon - u_0)\}$, $\{\varepsilon^{-2}(\psi_\varepsilon - \psi_0)\}$ are bounded in $L^\infty([0,T])$. Moreover, let $t_i \in [0,T]$ be an impact point in time (see Assumption 2) with $\dot\psi_0(t_i) = 0$ and $\ddot\psi_0(t_i) \neq 0$. Then, for all $a, b \in [0,T]$:*

$$\int_a^b u_\varepsilon \sin(\varepsilon^{-1}\psi_\varepsilon)\,\mathrm{d}t = \mathcal{O}(\varepsilon^{1/2}), \qquad \int_a^b u_\varepsilon \cos(\varepsilon^{-1}\psi_\varepsilon)\,\mathrm{d}t = \mathcal{O}(\varepsilon^{1/2}).$$

Proof. We treat the real and imaginary parts separately and write

$$u_\varepsilon \exp(i\varepsilon^{-1}\psi_\varepsilon) = (u_\varepsilon - u_0)\exp(i\varepsilon^{-1}\psi_\varepsilon) - u_0 \exp(i\varepsilon^{-1}\psi_0)\left(1 - \exp\left(i\varepsilon^{-1}(\psi_\varepsilon - \psi_0)\right)\right) + u_0 \exp(i\varepsilon^{-1}\psi_0).$$

Since the sequences $\{\varepsilon^{-1}(u_\varepsilon - u_0)\}$ and $\{\varepsilon^{-2}(\psi_\varepsilon - \psi_0)\}$ are bounded in $L^\infty([0,T])$, the claim is satisfied for the first two terms on the right-hand side. Moreover, let $\eta \in C_0^\infty([0,T])$ and U_{t_i}, V_{t_i} be small neighbourhoods around t_i such that $\operatorname{supp}\eta = V_{t_i}$, $U_{t_i} \subset V_{t_i}$ and $\eta = 1$ in U_{t_i} and write

$$\int_a^b u_0 \exp(i\varepsilon^{-1}\psi_0)\,\mathrm{d}t = \int_{V_{t_i}} u_0 \exp(i\varepsilon^{-1}\psi_0)\eta\,\mathrm{d}t + \int_{[a,b]\setminus U_{t_i}} u_0 \exp(i\varepsilon^{-1}\psi_0)(1-\eta)\,\mathrm{d}t.$$

For the second integral we can apply Lemma 5.4 since $t_i \notin [a,b] \setminus U_{t_i}$, and obtain an error of order $\mathcal{O}(\varepsilon)$. For the first integral we use the method of stationary phase to derive

$$\int_{V_{t_i}} u_0 \exp(i\varepsilon^{-1}\psi_0)\eta\,\mathrm{d}t = \mathcal{O}(\varepsilon^{1/2}).$$

A detailed description of the method of stationary phase can be found, for example, in [25, §1, Proposition 3], where smoothness of u_0 and ψ_0 is assumed. Here, we are only interested in the leading-order asymptotics, for which $u_0 \in C^2([0,T])$ and $\psi_0 \in C^3([0,T])$ is sufficient. □

Lemma 5.6 (*Generalisation of Lemma 3.6 in [11]*). *Let $u \in C^2(\mathbb{R}^r_{>0} \times \mathbb{R}^{r+2n})$ and $(\phi_\varepsilon, \theta_\varepsilon, y_\varepsilon, p_\varepsilon)$ be the solution to (11)–(12). Then, the sequence of functions $\{u_\varepsilon\}$ where $u_\varepsilon := u(\dot\phi_\varepsilon, \theta_\varepsilon, \dot y_\varepsilon, y_\varepsilon)$ satisfies for all $a, b \in [0,T]$ and $k = 1, 2$:*

$$\int_a^b \dot u_\varepsilon \cos(2k\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t \to \frac{2-k}{2}\int_a^b D_t L_0^\lambda\left(\omega_\lambda(y_0)\cdot\partial_\lambda u_0 - \theta_*^\lambda \cdot \partial_{r+\lambda}u_0\right) + \theta_*^\lambda \sum_{j=1}^n D_j\omega_\lambda(y_0)\cdot\partial_{2r+j}u_0\,\mathrm{d}t$$

and

$$\int_a^b \dot u_\varepsilon \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t = \mathcal{O}(\varepsilon).$$

Proof. The equations in (17) imply

$$\begin{aligned}
\dot u_\varepsilon = &\sum_{\mu=1}^r \partial_\mu u_\varepsilon \cdot \left(D_t\omega_\mu(y_\varepsilon) + \frac{\varepsilon}{2}D_t^2 L_\varepsilon^\mu \sin(2\varepsilon^{-1}\phi_\varepsilon^\mu) + D_t L_\varepsilon^\mu \cdot \dot\phi_\varepsilon^\mu \cos(2\varepsilon^{-1}\phi_\varepsilon^\mu)\right)\\
&- \sum_{\mu=1}^r \partial_{r+\mu}u_\varepsilon \cdot \theta_\varepsilon^\mu \cdot D_t L_\varepsilon^\mu \cos(2\varepsilon^{-1}\phi_\varepsilon^\mu)\\
&- \sum_{j=1}^n \partial_{2r+j}u_\varepsilon \cdot \left(D_j V(y_\varepsilon) + \sum_{\mu=1}^r \theta_\varepsilon^\mu \cdot D_j\omega_\mu(y_\varepsilon) - \sum_{\mu=1}^r \theta_\varepsilon^\mu \cdot D_j\omega_\mu(y_\varepsilon)\cos(2\varepsilon^{-1}\phi_\varepsilon^\mu)\right)\\
&+ \sum_{j=1}^n \partial_{2r+n+j}u_\varepsilon \cdot \dot y_\varepsilon^j.
\end{aligned}$$

The claim follows from the uniform convergence results in (14), Lemmas 5.4 and 5.5 and the trigonometric identities

$$2\cos(x)\cos(y) = \cos(x+y) + \cos(x-y), \qquad 2\cos(x)\sin(y) = \sin(x+y) - \sin(x-y).$$ □

Lemma 5.7 (*Similar to Lemma 3.9 in [11]*). *Let $u_0, \psi_0 \in C^1([0,T])$ and let $\{u_\varepsilon\}, \{\psi_\varepsilon\}$ be sequences in $C^1([0,T])$ such that the sequences $\{\dot u_\varepsilon\}$, $\{\varepsilon^{-1}(u_\varepsilon - u_0)\}$, $\{\varepsilon^{-2}(\psi_\varepsilon - \psi_0)\}$ and $\{\varepsilon^{-1}(\dot\psi_\varepsilon - \dot\psi_0)\}$ are bounded in $L^\infty([0,T])$. Then, for v_ε such that*

$$v_\varepsilon := u_\varepsilon \exp(i\varepsilon^{-1}\psi_\varepsilon) - u_0 \exp(i\varepsilon^{-1}\psi_0),$$

the sequence $\{\varepsilon^{-1}v_\varepsilon\}$ is bounded in $L^\infty([0,T],\mathbb{C})$ and in particular

$$v_\varepsilon \to 0 \quad \text{in} \quad C([0,T],\mathbb{C}), \qquad \dot v_\varepsilon \overset{*}{\rightharpoonup} 0 \quad \text{in} \quad L^\infty([0,T],\mathbb{C}).$$

Proof. By writing

$$v_\varepsilon = (u_\varepsilon - u_0)\exp(i\varepsilon^{-1}\psi_\varepsilon) - u_0\exp(i\varepsilon^{-1}\psi_0)\left(1 - \exp\left(i\varepsilon^{-1}(\psi_\varepsilon - \psi_0)\right)\right)$$

and

$$\dot{v}_\varepsilon = \left((\dot{u}_\varepsilon - \dot{u}_0) + i\varepsilon^{-1}\dot{\psi}_\varepsilon(u_\varepsilon - u_0) + i\varepsilon^{-1}u_0(\dot{\psi}_\varepsilon - \dot{\psi}_0)\right)\exp(i\varepsilon^{-1}\psi_\varepsilon) - \left(\dot{u}_0 + i\varepsilon^{-1}\dot{\psi}_0 u_0\right)\left(1 - \exp\left(i\varepsilon^{-1}\left(\psi_\varepsilon - \psi_0\right)\right)\right)\exp(i\varepsilon^{-1}\psi_0),$$

the assumptions imply that the sequences $\{\varepsilon^{-1}v_\varepsilon\}$ and $\{\dot{v}_\varepsilon\}$ are bounded in $L^\infty([0,T],\mathbb{C})$. This implies directly the uniform convergence of v_ε to zero. The weak* convergence of $\dot{v}_\varepsilon$ follows from [1, Principle 1]. □

Lemma 5.8. *The sequences $\{\theta_1^\varepsilon\}$ and $\{\phi_2^\varepsilon\}$ are uniformly bounded in $L^\infty([0,T],\mathbb{R}^r)$, and the sequences $\{y_2^\varepsilon\}$ and $\{p_2^\varepsilon\}$ are uniformly bounded in $L^\infty([0,T],\mathbb{R}^n)$.*

Proof. In this proof, the constant $0 < C < \infty$ depends on T but is independent of ε and can take different values from line to line. Let $t \in [0,T]$. For $0 < \varepsilon < \varepsilon_0$ small enough and $\lambda = 1,\dots,r$ let

$$M_1^\lambda := \sup_{0<\varepsilon<\varepsilon_0}\sup_{h\in[0,1]} \|D\omega_\lambda((1-h)y_\varepsilon + hy_0)\|_{L^\infty([0,T],\mathbb{R}^n)},$$
$$M_2^\lambda := \sup_{0<\varepsilon<\varepsilon_0}\sup_{h\in[0,1]} \left\|D^2\omega_\lambda((1-h)y_\varepsilon + hy_0)\right\|_{L^\infty([0,T],\mathbb{R}^{n\times n})},$$
$$M_3 := \sup_{0<\varepsilon<\varepsilon_0}\sup_{h\in[0,1]} \left\|D^2V((1-h)y_\varepsilon + hy_0)\right\|_{L^\infty([0,T],\mathbb{R}^{n\times n})}.$$

For $\lambda = 1,\dots,r$ and $j = 1,\dots,n$ we apply Lemma 5.4 to Eq. (17a),

$$\left|\phi_2^{\lambda\varepsilon}(t)\right| = \left|\int_0^t \dot{\phi}_2^{\lambda\varepsilon}\,\mathrm{d}s\right| \le \frac{1}{\varepsilon^2}\left|\int_0^t \omega_\lambda(y_\varepsilon) - \omega_\lambda(y_0)\,\mathrm{d}s\right| + \frac{1}{2\varepsilon}\left|\int_0^t D_s L_\varepsilon^\lambda \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}s\right| \le M_1^\lambda \sum_{j=1}^n \int_0^t \left|y_2^{j\varepsilon}\right|\,\mathrm{d}s + C, \tag{25}$$

then to Eq. (17b),

$$\left|\theta_1^{\lambda\varepsilon}(t)\right| = \left|\int_0^t \dot{\theta}_1^{\lambda\varepsilon}\,\mathrm{d}s\right| = \frac{1}{\varepsilon}\left|\int_0^t \theta_\varepsilon^\lambda \cdot D_s L_\varepsilon^\lambda \cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}s\right| \le C, \tag{26}$$

to Eq. (17c),

$$\left|y_2^{j\varepsilon}(t)\right| = \left|\int_0^t \dot{y}_2^{j\varepsilon}\,\mathrm{d}s\right| \le \int_0^t \left|p_2^{j\varepsilon}\right|\,\mathrm{d}s + \frac{1}{2\varepsilon}\sum_{\lambda=1}^r \left|\int_0^t \theta_\varepsilon^\lambda \cdot D_j L_\varepsilon^\lambda \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}s\right| \le \int_0^t \left|p_2^{j\varepsilon}\right|\,\mathrm{d}s + C, \tag{27}$$

and finally to Eq. (17d),

$$\begin{aligned}\left|p_2^{j\varepsilon}(t)\right| &= \left|\int_0^t \dot{p}_2^{j\varepsilon}(s)\,\mathrm{d}s\right| \\ &\le \frac{1}{\varepsilon^2}\left|\int_0^t D_j V(y_\varepsilon) - D_j V(y_0)\,\mathrm{d}s\right| + \frac{1}{\varepsilon^2}\sum_{\lambda=1}^r \left|\int_0^t \theta_\varepsilon^\lambda \cdot D_j\omega_\lambda(y_\varepsilon) - \theta_*^\lambda \cdot D_j\omega_\lambda(y_0)\,\mathrm{d}s\right| \\ &\quad + \frac{1}{2\varepsilon}\sum_{\lambda=1}^r \left|\int_0^t \theta_\varepsilon^\lambda \cdot D_t D_j L_\varepsilon^\lambda \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}s\right| \\ &\le M_3 \sum_{k=1}^n \int_0^t \left|y_2^{k\varepsilon}\right|\,\mathrm{d}s + \frac{1}{\varepsilon}\sum_{\lambda=1}^r \left|\int_0^t \theta_1^{\lambda\varepsilon} \cdot D_j\omega_\lambda(y_\varepsilon)\,\mathrm{d}s\right| + \langle\theta_*, M_2\rangle \sum_{k=1}^n \int_0^t \left|y_2^{k\varepsilon}\right|\,\mathrm{d}s + C.\end{aligned} \tag{28}$$

After integrating by parts, Eq. (17b) and Lemmas 5.4 and 5.6 imply that

$$\begin{aligned}\frac{1}{\varepsilon}\sum_{\lambda=1}^r \left|\int_0^t \theta_1^{\lambda\varepsilon} \cdot D_j\omega_\lambda(y_\varepsilon)\,\mathrm{d}s\right| &= \frac{1}{\varepsilon}\sum_{\lambda=1}^r \left|\int_0^t D_j\omega_\lambda(y_\varepsilon)\int_0^s \dot{\theta}_1^{\lambda\varepsilon}\,\mathrm{d}r\,\mathrm{d}s\right| \\ &= \frac{1}{\varepsilon^2}\sum_{\lambda=1}^r \left|\int_0^t D_j\omega_\lambda(y_\varepsilon)\int_0^s \theta_\varepsilon^\lambda \cdot D_r L_\varepsilon^\lambda \cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right| \\ &\le \frac{1}{2\varepsilon}\sum_{\lambda=1}^r \left|\int_0^t \frac{\theta_\varepsilon^\lambda \cdot D_s L_\varepsilon^\lambda \cdot D_j\omega_\lambda(y_\varepsilon)}{\dot{\phi}_\varepsilon^\lambda}\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}s\right| \\ &\quad + \frac{1}{2\varepsilon}\sum_{\lambda=1}^r \left|\int_0^t D_j\omega_\lambda(y_\varepsilon)\int_0^s \frac{d}{dr}\left(\frac{\theta_\varepsilon^\lambda \cdot D_r L_\varepsilon^\lambda}{\dot{\phi}_\varepsilon^\lambda}\right)\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}r\,\mathrm{d}s\right| \le C.\end{aligned} \tag{29}$$

By combining the inequalities (27)–(29) we obtain

$$\left|y_2^{j\varepsilon}(t)\right| \le C + (M_3 + \langle\theta_*, M_2\rangle)\int_0^T \int_0^T \sum_{j=1}^n \left|y_2^{j\varepsilon}\right|\,\mathrm{d}s$$

and thus

$$\sum_{j=1}^{n}\left|y_2^{j\varepsilon}(t)\right| \le nC + n\left(M_3 + \langle\theta_*, M_2\rangle\right)\int_0^T\int_0^T\sum_{j=1}^{n}\left|y_2^{j\varepsilon}\right| \mathrm{d}s.$$

Finally, a variation of the classical Gronwall inequality (see [26, p. 383]) implies that

$$\sum_{j=1}^{n}\left|y_2^{j\varepsilon}(t)\right| \le nC\exp\left(n\left(M_3 + \langle\theta_*, M_2\rangle\right)T^2\right),$$

for $t \in [0,T]$, which together with (25)–(29) yields the uniform bound for $\{\theta_1^\varepsilon\}$, $\{\phi_2^\varepsilon\}$, $\{y_2^\varepsilon\}$ and $\{p_2^\varepsilon\}$. □

Lemma 5.9. *The sequence $\{\theta_2^\varepsilon\}$ is uniformly bounded in $L^\infty([0,T],\mathbb{R}^r)$.*

Proof. We write θ_1^ε componentwise as

$$\theta_1^{\lambda\varepsilon} = \frac{1}{\varepsilon}\int_0^{\cdot}\dot\phi_\varepsilon^\lambda \,\mathrm{d}t = -\int_0^{\cdot}\frac{\theta_\varepsilon^\lambda\cdot D_tL_\varepsilon^\lambda}{2\dot\phi_\varepsilon^\lambda}\frac{d}{dt}\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t$$

and integrate by parts to derive $\theta_2^{\lambda\varepsilon} = \varepsilon^{-1}\left(\theta_1^{\lambda\varepsilon} - [\theta_1^\lambda]^\varepsilon\right) = \theta_{21}^{\lambda\varepsilon} + \theta_{22}^{\lambda\varepsilon}$, where

$$\theta_{21}^{\lambda\varepsilon} := \frac{1}{\varepsilon}\left(\frac{\theta_*^\lambda\cdot D_tL_0^\lambda}{2\omega_\lambda(y_0)}\sin(2\varepsilon^{-1}\phi_0^\lambda) - \frac{\theta_\varepsilon^\lambda\cdot D_tL_\varepsilon^\lambda}{2\dot\phi_\varepsilon^\lambda}\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\right),\quad \theta_{22}^{\lambda\varepsilon} := \frac{1}{\varepsilon}\int_0^{\cdot}\frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda\cdot D_tL_\varepsilon^\lambda}{2\dot\phi_\varepsilon^\lambda}\right)\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t. \tag{30}$$

The claim then follows from Assumption 1 and Lemmas 5.6 and 5.7. □

Remark. Lemmas 5.7 and 5.9 imply the convergence (18). Moreover, without Assumption 1, the sequence $\{\theta_{22}^{\lambda\varepsilon}\}$ is not necessarily bounded in $L^\infty([0,T])$. In fact, by Lemma 5.5 and the definition of $\theta_{227}^{\lambda\varepsilon}$ in the proof of Lemma 5.11, we would have in this case $\{\theta_{22}^{\lambda\varepsilon}\} = \mathcal{O}(\varepsilon^{-1/2})$.

Lemma 5.10. *There exist a subsequence $\{\varepsilon'\}$ and functions $\bar\phi_2 \in C^{0,1}([0,T],\mathbb{R}^r)$, $\bar y_2, \bar p_2 \in C^{0,1}([0,T],\mathbb{R}^n)$ such that the convergences (19)–(21) hold.*

Proof. By taking the time derivative of $\phi_2^{\lambda\varepsilon} - [\phi_2^\lambda]^\varepsilon$, $y_2^{j\varepsilon} - [y_2^j]^\varepsilon$ and $p_2^{j\varepsilon} - [p_2^j]^\varepsilon$ for $\lambda = 1,\dots,r$ and $j = 1,\dots,n$ we obtain

$$\frac{d}{dt}\left(\phi_2^{\lambda\varepsilon} - [\phi_2^\lambda]^\varepsilon\right) = \frac{\omega_\lambda(y_\varepsilon)-\omega_\lambda(y_0)}{\varepsilon^2} + \frac{d}{dt}\left(\frac{D_tL_\varepsilon^\lambda}{4\dot\phi_\varepsilon^\lambda}\right)\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - \frac{d}{dt}\left([\phi_2^\lambda]^\varepsilon + \frac{D_tL_\varepsilon^\lambda}{4\dot\phi_\varepsilon^\lambda}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\right), \tag{31}$$

$$\frac{d}{dt}\left(y_2^{j\varepsilon} - [y_2^j]^\varepsilon\right) = \frac{p_\varepsilon^j - p_0^j}{\varepsilon^2} + \sum_{\lambda=1}^{r}\frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda\cdot D_jL_\varepsilon^\lambda}{4\dot\phi_\varepsilon^\lambda}\right)\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - \frac{d}{dt}\left([y_2^j]^\varepsilon + \sum_{\lambda=1}^{r}\frac{\theta_\varepsilon^\lambda\cdot D_jL_\varepsilon^\lambda}{4\dot\phi_\varepsilon^\lambda}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\right), \tag{32}$$

$$\begin{aligned}\frac{d}{dt}\left(p_2^{j\varepsilon} - [p_2^j]^\varepsilon\right) &= -\frac{D_jV(y_\varepsilon)-D_jV(y_0)}{\varepsilon^2} - \sum_{\lambda=1}^{r}\theta_*^\lambda\frac{D_j\omega_\lambda(y_\varepsilon)-D_j\omega_\lambda(y_0)}{\varepsilon^2} - \sum_{\lambda=1}^{r}\frac{\theta_1^{\lambda\varepsilon}-[\theta_1^\lambda]^\varepsilon}{\varepsilon}D_j\omega_\lambda(y_\varepsilon)\\ &\quad - \sum_{\lambda=1}^{r}\frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda\cdot D_tD_jL_\varepsilon^\lambda}{4\dot\phi_\varepsilon^\lambda}\right)\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - \frac{d}{dt}\left([p_2^j]_1^\varepsilon - \sum_{\lambda=1}^{r}\frac{\theta_\varepsilon^\lambda\cdot D_tD_jL_\varepsilon^\lambda}{4\dot\phi_\varepsilon^\lambda}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\right)\\ &\quad + \sum_{\lambda=1}^{r}\frac{d}{dt}\left(\frac{\theta_*^\lambda\cdot D_tL_0^\lambda\cdot D_j\omega_\lambda(y_\varepsilon)}{4\omega^2(y_0)}\right)\cos(2\varepsilon^{-1}\phi_0^\lambda) - \frac{d}{dt}\left([p_2^j]_2^\varepsilon + \sum_{\lambda=1}^{r}\frac{\theta_*^\lambda\cdot D_tL_0^\lambda\cdot D_j\omega_\lambda(y_\varepsilon)}{4\omega_\lambda^2(y_0)}\cos(2\varepsilon^{-1}\phi_0^\lambda)\right),\end{aligned} \tag{33}$$

where we used $[p_2^j]^\varepsilon = [p_2^j]_1^\varepsilon + [p_2^j]_2^\varepsilon$ with

$$[p_2^j]_1^\varepsilon := \sum_{\lambda=1}^{r}\frac{\theta_*^\lambda\cdot D_tD_jL_0^\lambda}{4\omega_\lambda(y_0)}\cos(2\varepsilon^{-1}\phi_0^\lambda),\qquad [p_2^j]_2^\varepsilon := -\sum_{\lambda=1}^{r}\frac{\theta_*^\lambda\cdot D_tL_0^\lambda\cdot D_jL_0^\lambda}{4\omega_\lambda(y_0)}\cos(2\varepsilon^{-1}\phi_0^\lambda).$$

For the derivation of Eq. (33) note that in the evaluation of $\dot p_2^{j\varepsilon}$ we need to evaluate the expression

$$\sum_{\lambda=1}^{r}\frac{\theta_\varepsilon^\lambda\cdot D_j\omega_\lambda(y_\varepsilon) - \theta_*^\lambda\cdot D_j\omega_\lambda(y_0)}{\varepsilon^2} = \sum_{\lambda=1}^{r}\frac{\theta_\varepsilon^\lambda-\theta_*^\lambda}{\varepsilon^2}D_j\omega_\lambda(y_\varepsilon) + \sum_{\lambda=1}^{r}\theta_*^\lambda\frac{D_j\omega_\lambda(y_\varepsilon)-D_j\omega_\lambda(y_0)}{\varepsilon^2},$$

in which we rewrite the first term on the right-hand side by introducing $[\theta_1^\lambda]^\varepsilon$, i.e.,

$$\begin{aligned}\frac{\theta_\varepsilon^\lambda-\theta_*^\lambda}{\varepsilon^2}D_j\omega_\lambda(y_\varepsilon) &= \frac{\theta_1^{\lambda\varepsilon}-[\theta_1^\lambda]^\varepsilon}{\varepsilon}D_j\omega_\lambda(y_\varepsilon) + \frac{d}{dt}\left(\frac{\theta_*^\lambda\cdot D_tL_0^\lambda\cdot D_j\omega_\lambda(y_\varepsilon)}{4\omega_\lambda^2(y_0)}\cos(2\varepsilon^{-1}\phi_0^\lambda)\right)\\ &\quad - \frac{d}{dt}\left(\frac{\theta_*^\lambda\cdot D_tL_0^\lambda\cdot D_j\omega_\lambda(y_\varepsilon)}{4\omega^2(y_0)}\right)\cos(2\varepsilon^{-1}\phi_0^\lambda).\end{aligned}$$

By Lemmas 5.7–5.9 the sequence $\{\phi_2^\varepsilon - [\phi_2]^\varepsilon\}$ is bounded in $C^{0,1}([0,T],\mathbb{R}^r)$ and the sequences $\{y_2^\varepsilon - [y_2]^\varepsilon\}$ and $\{p_2^\varepsilon - [p_2]^\varepsilon\}$ are bounded in $C^{0,1}([0,T],\mathbb{R}^n)$. The claim follows after successive applications of [1, Principle 4]. □

Lemma 5.11. *There exists a further subsequence $\{\varepsilon'\}$ and a function $\bar{\theta}_2 \in C^\infty([0,T],\mathbb{R}^r)$ such that the convergence (22) holds. Moreover, the component functions $\bar{\theta}_2^\lambda$ $(\lambda = 1,\dots,r)$ satisfy (23b).*

Proof. We write $\theta_2^{\lambda\varepsilon} = \sum_{i=1}^2 \theta_{2i}^{\lambda\varepsilon}$ as in the proof of Lemma 5.9 and $[\theta_2^\lambda]^\varepsilon = \sum_{i=1}^2 [\theta_2^\lambda]_i^\varepsilon$, where $[\theta_2^\lambda]_i^\varepsilon$ $(i = 1,2)$ will be defined later in this proof. We then show that there exist component functions $\bar{\theta}_2^\lambda := \sum_{i=1}^2 \bar{\theta}_{2i}^\lambda$ and a subsequence $\{\varepsilon'\}$, not relabelled, such that for $i = 1$

$$\theta_{21}^{\lambda\varepsilon} - [\theta_2^\lambda]_1^\varepsilon \to \bar{\theta}_{21}^\lambda \quad \text{in} \quad C([0,T]) \tag{34}$$

and for $i = 2$

$$\theta_{22}^{\lambda\varepsilon} - [\theta_2^\lambda]_2^\varepsilon \to \bar{\theta}_{22}^\lambda \quad \text{in} \quad C([0,T]), \qquad \frac{d}{dt}\left(\theta_{22}^{\lambda\varepsilon} - [\theta_2^\lambda]_2^\varepsilon\right) \overset{*}{\rightharpoonup} \frac{d\bar{\theta}_{22}^\lambda}{dt} \quad \text{in} \quad L^\infty([0,T]). \tag{35}$$

The convergence result (22) then follows immediately.

Part $i = 1$: To prove (34), we expand $\theta_{21}^{\lambda\varepsilon}$ in (30), by replacing $\theta_\varepsilon^\lambda \to \theta_*^\lambda + \left(\theta_\varepsilon^\lambda - \theta_*^\lambda\right)$, $D_t L_\varepsilon^\lambda \to D_t L_0^\lambda + \left(D_t L_\varepsilon^\lambda - D_t L_0^\lambda\right)$ and $\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \to \sin(2\varepsilon^{-1}\phi_0^\lambda) + \left(\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - \sin(2\varepsilon^{-1}\phi_0^\lambda)\right)$, and assign the resulting terms to the functions $\theta_{21j}^{\lambda\varepsilon}$ $(j = 1,\dots,5)$. That is, we derive $\theta_{21}^{\lambda\varepsilon} = \sum_{j=1}^5 \theta_{21j}^{\lambda\varepsilon}$, where

$$\begin{aligned}
\theta_{211}^{\lambda\varepsilon} &:= \frac{1}{\varepsilon}\left(\frac{1}{\omega_\lambda(y_0)} - \frac{1}{\dot{\phi}_\varepsilon^\lambda}\right)\frac{\theta_*^\lambda \cdot D_t L_0^\lambda}{2}\sin(2\varepsilon^{-1}\phi_0^\lambda),\\
\theta_{212}^{\lambda\varepsilon} &:= -\frac{1}{\varepsilon}\frac{\left(\theta_\varepsilon^\lambda - \theta_*^\lambda\right)\cdot D_t L_0^\lambda}{2\dot{\phi}_\varepsilon^\lambda}\sin(2\varepsilon^{-1}\phi_0^\lambda),\\
\theta_{213}^{\lambda\varepsilon} &:= -\frac{1}{\varepsilon}\frac{\theta_*^\lambda \cdot \left(D_t L_\varepsilon^\lambda - D_t L_0^\lambda\right)}{2\dot{\phi}_\varepsilon^\lambda}\sin(2\varepsilon^{-1}\phi_0^\lambda),\\
\theta_{214}^{\lambda\varepsilon} &:= -\frac{1}{\varepsilon}\frac{\theta_*^\lambda \cdot D_t L_0^\lambda}{2\dot{\phi}_\varepsilon^\lambda}\left(\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - \sin(2\varepsilon^{-1}\phi_0^\lambda)\right),\\
\theta_{215}^{\lambda\varepsilon} &:= -\frac{1}{\varepsilon}\frac{\left(\theta_\varepsilon^\lambda - \theta_*^\lambda\right)\cdot \left(D_t L_\varepsilon^\lambda - D_t L_0^\lambda\right)}{2\dot{\phi}_\varepsilon^\lambda}\sin(2\varepsilon^{-1}\phi_0^\lambda)\\
&\quad - \frac{1}{\varepsilon}\frac{\left(\theta_\varepsilon^\lambda - \theta_*^\lambda\right)\cdot D_t L_0^\lambda}{2\dot{\phi}_\varepsilon^\lambda}\left(\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - \sin(2\varepsilon^{-1}\phi_0^\lambda)\right)\\
&\quad - \frac{1}{\varepsilon}\frac{\theta_*^\lambda \cdot \left(D_t L_\varepsilon^\lambda - D_t L_0^\lambda\right)}{2\dot{\phi}_\varepsilon^\lambda}\left(\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - \sin(2\varepsilon^{-1}\phi_0^\lambda)\right)\\
&\quad - \frac{1}{\varepsilon}\frac{\left(\theta_\varepsilon^\lambda - \theta_*^\lambda\right)\cdot \left(D_t L_\varepsilon^\lambda - D_t L_0^\lambda\right)}{2\dot{\phi}_\varepsilon^\lambda}\left(\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - \sin(2\varepsilon^{-1}\phi_0^\lambda)\right).
\end{aligned}$$

Notice that by (17) and Lemma 5.8, we have $\theta_{21j}^{\lambda\varepsilon} = \mathcal{O}(1)$ for $j = 1,\dots,4$ and $\theta_{215}^{\lambda\varepsilon} = \mathcal{O}(\varepsilon)$. The function $\theta_{21}^{\lambda\varepsilon} = \sum_{j=1}^5 \theta_{21j}^{\lambda\varepsilon}$ is composed of oscillatory and non-oscillatory (averaged) terms. For the proof of (34), we therefore define the corresponding oscillatory term $[\theta_2^\lambda]_1^\varepsilon := \sum_{j=1}^4 [\theta_2^\lambda]_{1j}^\varepsilon$, where

$$\begin{aligned}
[\theta_2^\lambda]_{11}^\varepsilon &:= -\frac{\theta_*^\lambda (D_t L_0^\lambda)^2}{8\omega_\lambda^2(y_0)}\cos(4\varepsilon^{-1}\phi_0^\lambda),\\
[\theta_2^\lambda]_{12}^\varepsilon &:= -\frac{\theta_*^\lambda (D_t L_0^\lambda)^2}{8\omega_\lambda^2(y_0)}\cos(4\varepsilon^{-1}\phi_0^\lambda),\\
[\theta_2^\lambda]_{13}^\varepsilon &:= \frac{(\theta_*^\lambda)^2\left|DL_0^\lambda\right|^2}{8\omega_\lambda(y_0)}\cos(4\varepsilon^{-1}\phi_0^\lambda) - \sum_{\substack{\mu=1\\ \mu\neq\lambda}}^r \frac{\theta_*^\lambda\theta_*^\mu\left\langle DL_0^\mu, DL_0^\lambda\right\rangle}{4\omega_\lambda(y_0)}\sin(2\varepsilon^{-1}\phi_0^\lambda)\sin(2\varepsilon^{-1}\phi_0^\mu),\\
[\theta_2^\lambda]_{14}^\varepsilon &:= -\frac{\theta_*^\lambda \cdot D_t L_0^\lambda}{\omega_\lambda(y_0)}\bar{\phi}_2^\lambda\cos(2\varepsilon^{-1}\phi_0^\lambda) + \frac{\theta_*^\lambda (D_t L_0^\lambda)^2}{8\omega_\lambda^2(y_0)}\cos(4\varepsilon^{-1}\phi_0^\lambda).
\end{aligned}$$

We now define the averaged functions $\bar{\theta}_{21j}^\lambda$ $(j = 1,\dots,4)$ such that for $\bar{\theta}_{21}^\lambda := \sum_{j=1}^4 \bar{\theta}_{21j}^\lambda$ the statement in (34) holds. More precisely, we will show that in the case of $j = 1,2,3$

$$\theta_{21j}^{\lambda\varepsilon} - [\theta_2^\lambda]_{1j}^\varepsilon \to \bar{\theta}_{21j}^\lambda \quad \text{in} \quad C([0,T]), \qquad \frac{d}{dt}\left(\theta_{21j}^{\lambda\varepsilon} - [\theta_2^\lambda]_{1j}^\varepsilon\right) \overset{*}{\rightharpoonup} \frac{d\bar{\theta}_{21j}^\lambda}{dt} \quad \text{in} \quad L^\infty([0,T]), \tag{36}$$

and in the case of $j = 4$

$$\theta_{214}^{\lambda\varepsilon} - [\theta_2^\lambda]_{14}^\varepsilon \to \bar{\theta}_{214}^\lambda \quad \text{in} \quad C([0,T]). \tag{37}$$

In the following, we give the detailed proof of the convergences in (36) for the case $j = 1$. The other cases follow along a similar line of arguments.

Case $j = 1$: We start by defining $\bar{\theta}^{\lambda}_{221}$, i.e.,

$$\bar{\theta}^{\lambda}_{211} := \frac{\theta^{\lambda}_{*}(D_t L^{\lambda}_0)^2}{8\omega^2_{\lambda}(y_0)},$$

for $\lambda = 1, \dots, r$ and use the trigonometric identity $1 - \cos(4x) = 2\sin^2(2x)$ to derive

$$\bar{\theta}^{\lambda}_{211} + [\theta^{\lambda}_2]^{\varepsilon}_{11} = \frac{\theta^{\lambda}_{*}(D_t L^{\lambda}_0)^2}{8\omega^2_{\lambda}(y_0)} - \frac{\theta^{\lambda}_{*}(D_t L^{\lambda}_0)^2}{8\omega^2_{\lambda}(y_0)}\cos(4\varepsilon^{-1}\phi^{\lambda}_0) = \frac{\theta^{\lambda}_{*}(D_t L^{\lambda}_0)^2}{4\omega^2_{\lambda}(y_0)}\sin^2(2\varepsilon^{-1}\phi^{\lambda}_0). \tag{38}$$

Moreover, with Eq. (17a) we can write

$$\begin{aligned}\theta^{\lambda\varepsilon}_{211} &= \frac{1}{\varepsilon}\left(\frac{\dot{\phi}^{\lambda}_{\varepsilon} - \omega_{\lambda}(y_0)}{\omega_{\lambda}(y_0)\dot{\phi}^{\lambda}_{\varepsilon}}\right)\frac{\theta^{\lambda}_{*}\cdot D_t L^{\lambda}_0}{2}\sin(2\varepsilon^{-1}\phi^{\lambda}_0)\\ &= \left(\frac{1}{\varepsilon}\frac{\omega_{\lambda}(y_{\varepsilon}) - \omega_{\lambda}(y_0)}{\dot{\phi}^{\lambda}_{\varepsilon}} + \frac{D_t L^{\lambda}_{\varepsilon}}{2\dot{\phi}^{\lambda}_{\varepsilon}}\sin(2\varepsilon^{-1}\phi^{\lambda}_{\varepsilon})\right)\frac{\theta^{\lambda}_{*}\cdot D_t L^{\lambda}_0}{2\omega_{\lambda}(y_0)}\sin(2\varepsilon^{-1}\phi^{\lambda}_0).\end{aligned} \tag{39}$$

Now, we use Eqs. (38) and (39) to write

$$\theta^{\lambda\varepsilon}_{211} - [\theta^{\lambda}_2]^{\varepsilon}_{11} - \bar{\theta}^{\lambda}_{211} = \left(u^{\lambda\varepsilon}_1 + v^{\lambda\varepsilon}_1\right) w^{\lambda\varepsilon}_1, \tag{40}$$

where

$$u^{\lambda\varepsilon}_1 := \frac{1}{\varepsilon}\frac{\omega_{\lambda}(y_{\varepsilon}) - \omega_{\lambda}(y_0)}{\dot{\phi}^{\lambda}_{\varepsilon}}, \quad v^{\lambda\varepsilon}_1 := \frac{D_t L^{\lambda}_{\varepsilon}}{2\dot{\phi}^{\lambda}_{\varepsilon}}\sin(2\varepsilon^{-1}\phi^{\lambda}_{\varepsilon}) - \frac{D_t L^{\lambda}_0}{2\omega_{\lambda}(y_0)}\sin(2\varepsilon^{-1}\phi^{\lambda}_0), \quad w^{\lambda\varepsilon}_1 := \frac{\theta^{\lambda}_{*}\cdot D_t L^{\lambda}_0}{2\omega_{\lambda}(y_0)}\sin(2\varepsilon^{-1}\phi^{\lambda}_0).$$

It follows from Lemmas 5.7 and 5.8, and the system of differential equations (17) that the sequences $\{\varepsilon^{-1}u^{\lambda\varepsilon}_1\}$, $\{\dot{u}^{\lambda\varepsilon}_1\}$, $\{\varepsilon^{-1}v^{\lambda\varepsilon}_1\}$, $\{\dot{v}^{\lambda\varepsilon}_j\}$, $\{w^{\lambda\varepsilon}_1\}$ and $\{\varepsilon\dot{w}^{\lambda\varepsilon}_1\}$ are bounded in $L^{\infty}([0,T])$. This implies the uniform convergence, and after an application of [1, Principle 1], the weak* convergence in (36).

For the cases $j = 2, 3$, we only summarise the equations corresponding to (40), from which $u^{\lambda\varepsilon}_j$, $v^{\lambda\varepsilon}_3$ and $w^{\lambda\varepsilon}_j$ can be read off. The convergences as in (36) are then proven similarly to the case $j = 1$, by applying Lemmas 5.7 and 5.8 (and 5.9 in the case $j = 2$). The case $j = 4$ requires more explanation and is thus again described in greater detail.

Case $j = 2$: With

$$\bar{\theta}^{\lambda}_{212} := \frac{\theta^{\lambda}_{*}(D_t L^{\lambda}_0)^2}{8\omega^2_{\lambda}(y_0)},$$

we can write

$$\theta^{\lambda\varepsilon}_{212} - [\theta^{\lambda}_2]^{\varepsilon}_{12} - \bar{\theta}^{\lambda}_{212} = u^{\lambda\varepsilon}_2 w^{\lambda\varepsilon}_2,$$

where

$$u^{\lambda\varepsilon}_2 := \frac{[\theta^{\lambda}_1]^{\varepsilon}}{\omega_{\lambda}(y_0)} - \frac{\theta^{\lambda\varepsilon}_1}{\dot{\phi}^{\lambda}_{\varepsilon}}, \qquad w^{\lambda\varepsilon}_2 := \frac{D_t L^{\lambda}_0}{2}\sin(2\varepsilon^{-1}\phi^{\lambda}_0).$$

Case $j = 3$: Analogously, with

$$\bar{\theta}^{\lambda}_{213} := -\frac{(\theta^{\lambda}_{*})^2\left|DL^{\lambda}_0\right|^2}{8\omega_{\lambda}(y_0)},$$

we write

$$\theta^{\lambda\varepsilon}_{213} - [\theta^{\lambda}_2]^{\varepsilon}_{13} - \bar{\theta}^{\lambda}_{213} = \left(u^{\lambda\varepsilon}_3 + v^{\lambda\varepsilon}_3\right) w^{\lambda\varepsilon}_3;$$

here

$$u^{\lambda\varepsilon}_3 := -\frac{1}{\varepsilon}\frac{\langle p_{\varepsilon}, DL^{\lambda}_{\varepsilon}\rangle - \langle p_0, DL^{\lambda}_0\rangle}{\dot{\phi}^{\lambda}_{\varepsilon}}, \qquad w^{\lambda\varepsilon}_3 := \frac{\theta^{\lambda}_{*}}{2}\sin(2\varepsilon^{-1}\phi^{\lambda}_0),$$

$$v^{\lambda\varepsilon}_3 := \sum_{\mu=1}^{r}\frac{\theta^{\mu}_{*}\langle DL^{\mu}_0, DL^{\lambda}_0\rangle}{2\omega_{\lambda}(y_0)}\sin(2\varepsilon^{-1}\phi^{\mu}_0) - \frac{\theta^{\mu}_{\varepsilon}\langle DL^{\mu}_{\varepsilon}, DL^{\lambda}_{\varepsilon}\rangle}{2\dot{\phi}^{\lambda}_{\varepsilon}}\sin(2\varepsilon^{-1}\phi^{\mu}_{\varepsilon}).$$

Case $j = 4$: For this final case we first define

$$\bar{\theta}^{\lambda}_{214} := \frac{\theta^{\lambda}_{*}(D_t L^{\lambda}_0)^2}{8\omega^2_{\lambda}(y_0)}$$

and use $\phi^{\lambda}_{\varepsilon} = \phi^{\lambda}_0 + \varepsilon^2\phi^{\lambda\varepsilon}_2$ (see Definition 5.1) with a trigonometric identity to write

$$\sin(2\varepsilon^{-1}\phi^{\lambda}_{\varepsilon}) = \sin(2\varepsilon^{-1}\phi^{\lambda}_0)\cos(2\varepsilon\phi^{\lambda\varepsilon}_2) + \cos(2\varepsilon^{-1}\phi^{\lambda}_0)\sin(2\varepsilon\phi^{\lambda\varepsilon}_2).$$

This allows us to derive the equation

$$\theta^{\lambda\varepsilon}_{214} - [\theta^{\lambda}_2]^{\varepsilon}_{14} - \bar{\theta}^{\lambda}_{214} = u^{\lambda\varepsilon}_{41}w^{\lambda\varepsilon}_{41} + u^{\lambda\varepsilon}_{42}w^{\lambda\varepsilon}_{42},$$

where

$$u_{41}^{\lambda\varepsilon} := \frac{[\phi_2^\lambda]^\varepsilon + \bar{\phi}_2^\lambda}{\omega_\lambda(y_0)} - \frac{1}{\varepsilon}\frac{\sin(2\varepsilon\phi_2^{\lambda\varepsilon})}{2\dot{\phi}_\varepsilon^\lambda}, \qquad w_{41}^{\lambda\varepsilon} := \theta_*^\lambda \cdot D_t L_0^\lambda \cos(2\varepsilon^{-1}\phi_0^\lambda),$$

$$u_{42}^{\lambda\varepsilon} := \frac{1}{\varepsilon}\frac{1-\cos(2\varepsilon\phi_2^{\lambda\varepsilon})}{2\dot{\phi}_\varepsilon^\lambda}, \qquad w_{42}^{\lambda\varepsilon} := \theta_*^\lambda \cdot D_t L_0^\lambda \sin(2\varepsilon^{-1}\phi_0^\lambda).$$

By (17), (19) and Lemma 5.8 we obtain

$$|u_{41}^{\lambda\varepsilon}| \le \left|\frac{[\phi_2^\lambda]^\varepsilon + \bar{\phi}_2^\lambda}{\omega_\lambda(y_0)} - \frac{\phi_2^{\lambda\varepsilon}}{\dot{\phi}_\varepsilon^\lambda}\right| + \left|\frac{\phi_2^{\lambda\varepsilon}}{\dot{\phi}_\varepsilon^\lambda}\sum_{k=1}^{\infty}\frac{(-1)^k}{(2k+1)!}(2\varepsilon\phi_2^{\lambda\varepsilon})^{2k}\right| \to 0 \quad \text{in} \quad C([0,T])$$

and

$$|u_{42}^{\lambda\varepsilon}| \le \left|\frac{\phi_2^{\lambda\varepsilon}}{\dot{\phi}_\varepsilon^\lambda}\sum_{k=1}^{\infty}\frac{(-1)^k}{(2k)!}(2\varepsilon\phi_2^{\lambda\varepsilon})^{2k-1}\right| = \mathcal{O}(\varepsilon),$$

which implies the uniform convergence in (37).

Part $i = 2$: To prove (35) we expand $\theta_{22}^{\lambda\varepsilon}$ in (30), by writing out the time derivative and using the equations

$$\ddot{\phi}_\varepsilon^\lambda = D_t\omega_\lambda(y_\varepsilon) + \frac{\varepsilon}{2}D_t^2 L_\varepsilon^\lambda \sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda) + D_t L_\varepsilon^\lambda \cdot \dot{\phi}_\varepsilon^\lambda \cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)$$

and

$$D_t^2 L_\varepsilon^\lambda = \left\langle D^2 L_\varepsilon^\lambda \dot{y}_\varepsilon, \dot{y}_\varepsilon\right\rangle - \left\langle DV(y_\varepsilon), DL_\varepsilon^\lambda\right\rangle - \sum_{\mu=1}^{r}\theta_\varepsilon^\mu\left\langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\right\rangle + \sum_{\mu=1}^{r}\theta_\varepsilon^\mu\left\langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\right\rangle\cos(2\varepsilon^{-1}\phi_\varepsilon^\mu).$$

In this way, we can write $\theta_{22}^{\lambda\varepsilon} = \sum_{j=1}^{8}\theta_{22j}^{\lambda\varepsilon}$, where

$$\theta_{221}^{\lambda\varepsilon} := -\frac{1}{\varepsilon}\int_0^\cdot \frac{\theta_\varepsilon^\lambda (D_t L_\varepsilon^\lambda)^2}{\dot{\phi}_\varepsilon^\lambda}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t,$$

$$\theta_{222}^{\lambda\varepsilon} := -\frac{1}{\varepsilon}\int_0^\cdot \frac{\theta_\varepsilon^\lambda \cdot D_t L_\varepsilon^\lambda \cdot D_t\omega_\lambda(y_\varepsilon)}{2(\dot{\phi}_\varepsilon^\lambda)^2}\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t,$$

$$\theta_{223}^{\lambda\varepsilon} := \frac{1}{\varepsilon}\int_0^\cdot \frac{\theta_\varepsilon^\lambda \left\langle D^2 L_\varepsilon^\lambda \dot{y}_\varepsilon, \dot{y}_\varepsilon\right\rangle}{2\dot{\phi}_\varepsilon^\lambda}\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t,$$

$$\theta_{224}^{\lambda\varepsilon} := -\frac{1}{\varepsilon}\int_0^\cdot \frac{\theta_\varepsilon^\lambda \left\langle DV(y_\varepsilon), DL_\varepsilon^\lambda\right\rangle}{2\dot{\phi}_\varepsilon^\lambda}\sin(2\varepsilon^{-1}\phi_0^\lambda)\,\mathrm{d}t,$$

$$\theta_{225}^{\lambda\varepsilon} := -\frac{1}{\varepsilon}\sum_{\mu=1}^{r}\int_0^\cdot \frac{\theta_\varepsilon^\lambda\theta_\varepsilon^\mu \left\langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\right\rangle}{2\dot{\phi}_\varepsilon^\lambda}\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t,$$

$$\theta_{226}^{\lambda\varepsilon} := \frac{1}{\varepsilon}\int_0^\cdot \frac{(\theta_\varepsilon^\lambda)^2 \left\langle D\omega_\lambda(y_\varepsilon), DL_\varepsilon^\lambda\right\rangle}{2\dot{\phi}_\varepsilon^\lambda}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t,$$

$$\theta_{227}^{\lambda\varepsilon} := \frac{1}{\varepsilon}\sum_{\substack{\mu=1\\ \mu\neq\lambda}}^{r}\int_0^\cdot \frac{\theta_\varepsilon^\lambda\theta_\varepsilon^\mu \left\langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\right\rangle}{2\dot{\phi}_\varepsilon^\lambda}\cos(2\varepsilon^{-1}\phi_\varepsilon^\mu)\sin(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t,$$

$$\theta_{228}^{\lambda\varepsilon} := -\int_0^\cdot \frac{\theta_\varepsilon^\lambda \cdot D_t L_\varepsilon^\lambda \cdot D_t^2 L_\varepsilon^\lambda}{4(\dot{\phi}_\varepsilon^\lambda)^2}\sin^2(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\,\mathrm{d}t.$$

Again, the function $\theta_{22}^{\lambda\varepsilon} = \sum_{j=1}^{8}\theta_{22j}^{\lambda\varepsilon}$ consists of oscillatory and non-oscillatory terms. To derive the statement in (35), we therefore define the corresponding oscillatory term $[\theta_2^\lambda]_2^\varepsilon := \sum_{j=1}^{7}[\theta_2^\lambda]_{2j}^\varepsilon$, where

$$[\theta_2^\lambda]_{21}^\varepsilon := \frac{\theta_*^\lambda (D_t L_0^\lambda)^2}{8\omega_\lambda^2(y_0)}\cos(4\varepsilon^{-1}\phi_0^\lambda),$$

$$[\theta_2^\lambda]_{22}^\varepsilon := \frac{\theta_*^\lambda (D_t L_0^\lambda)^2}{4\omega_\lambda^2(y_0)}\cos(2\varepsilon^{-1}\phi_0^\lambda),$$

$$[\theta_2^\lambda]_{23}^\varepsilon := -\frac{\theta_*^\lambda \left\langle D^2 L_0^\lambda \dot{y}_0, \dot{y}_0\right\rangle}{4\omega_\lambda^2(y_0)}\cos(2\varepsilon^{-1}\phi_0^\lambda),$$

$$[\theta_2^\lambda]_{24}^\varepsilon := \frac{\theta_*^\lambda \left\langle DV(y_0), DL_0^\lambda\right\rangle}{4\omega_\lambda^2(y_0)}\cos(2\varepsilon^{-1}\phi_0^\lambda),$$

$$[\theta_2^\lambda]_{25}^\varepsilon := \sum_{\mu=1}^{r} \frac{\theta_*^\lambda \theta_*^\mu \left\langle D\omega_\mu(y_0), DL_0^\lambda \right\rangle}{4\omega_\lambda^2(y_0)} \cos(2\varepsilon^{-1}\phi_0^\lambda),$$

$$[\theta_2^\lambda]_{26}^\varepsilon := -\frac{(\theta_*^\lambda)^2 |DL_0^\lambda|^2}{16\omega_\lambda(y_0)} \cos(4\varepsilon^{-1}\phi_0^\lambda),$$

$$[\theta_2^\lambda]_{27}^\varepsilon := \sum_{\substack{\mu=1 \\ \mu\neq\lambda}}^{r} \frac{\theta_*^\lambda \theta_*^\mu \left\langle D\omega_\mu(y_0), DL_0^\lambda \right\rangle}{8\omega_\lambda(y_0)} \left\{ \frac{\cos\left(2\varepsilon^{-1}(\phi_0^\mu - \phi_0^\lambda)\right)}{\omega_\mu(y_0) - \omega_\lambda(y_0)} - \frac{\cos\left(2\varepsilon^{-1}(\phi_0^\mu + \phi_0^\lambda)\right)}{\omega_\mu(y_0) + \omega_\lambda(y_0)} \right\}.$$

We show that for a subsequence $\{\varepsilon'\}$ (not relabelled) there exist non-oscillatory functions $\bar\theta_{22j}^\lambda$ $(j = 1, \dots, 8)$ such that for $\bar\theta_{22}^\lambda := \sum_{j=1}^8 \bar\theta_{22j}^\lambda$, the statement in (35) holds. More precisely, we will prove that for $j = 1, \dots, 7$

$$\theta_{22j}^{\lambda\varepsilon} - [\theta_2^\lambda]_{2j}^\varepsilon \to \bar\theta_{22j}^\lambda \quad \text{in} \quad C([0,T]), \qquad \frac{d}{dt}\left(\theta_{22j}^{\lambda\varepsilon} - [\theta_2^\lambda]_{2j}^\varepsilon\right) \overset{*}{\rightharpoonup} \frac{d\bar\theta_{22j}^\lambda}{dt} \quad \text{in} \quad L^\infty([0,T]), \tag{41}$$

and for $j = 8$

$$\theta_{228}^{\lambda\varepsilon} \to \bar\theta_{228}^\lambda \quad \text{in} \quad C([0,T]), \qquad \frac{d\theta_{228}^{\lambda\varepsilon}}{dt} \overset{*}{\rightharpoonup} \frac{d\bar\theta_{228}^\lambda}{dt} \quad \text{in} \quad L^\infty([0,T]). \tag{42}$$

Note that the scaling in $\theta_{228}^{\lambda\varepsilon}$ is different from the scaling in $\theta_{22j}^{\lambda\varepsilon}$ $(j = 1, \dots, 7)$. As a consequence, there is no non-converging oscillatory component that we would have to subtract from $\theta_{228}^{\lambda\varepsilon}$ in order to analyse the limit $\varepsilon \to 0$.

We now give a detailed proof of the convergences in (41) for the case $j = 1$. The other cases are dealt with similarly.

Case $j = 1$: For $\lambda = 1, \dots, r$ we start by writing

$$\begin{aligned}\frac{d\theta_{221}^{\lambda\varepsilon}}{dt} &= -\frac{1}{\varepsilon}\frac{\theta_\varepsilon^\lambda (D_t L_\varepsilon^\lambda)^2}{2\dot\phi_\varepsilon^\lambda} \sin(4\varepsilon^{-1}\phi_\varepsilon^\lambda) = \frac{\theta_\varepsilon^\lambda (D_t L_\varepsilon^\lambda)^2}{8(\dot\phi_\varepsilon^\lambda)^2}\frac{d}{dt}\cos(4\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ &= \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda (D_t L_\varepsilon^\lambda)^2}{8(\dot\phi_\varepsilon^\lambda)^2}\cos(4\varepsilon^{-1}\phi_\varepsilon^\lambda)\right) - \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda (D_t L_\varepsilon^\lambda)^2}{8(\dot\phi_\varepsilon^\lambda)^2}\right)\cos(4\varepsilon^{-1}\phi_\varepsilon^\lambda),\end{aligned}$$

which we use to derive

$$\begin{aligned}\frac{d}{dt}\left(\theta_{221}^{\lambda\varepsilon} - [\theta_2^\lambda]_{21}^\varepsilon\right) &= \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda (D_t L_\varepsilon^\lambda)^2}{8(\dot\phi_\varepsilon^\lambda)^2}\cos(4\varepsilon^{-1}\phi_\varepsilon^\lambda) - [\theta_2^\lambda]_{21}^\varepsilon\right) - \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda (D_t L_\varepsilon^\lambda)^2}{8(\dot\phi_\varepsilon^\lambda)^2}\right)\cos(4\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ &= \dot v_1^{\lambda\varepsilon} - \dot u_1^{\lambda\varepsilon}\cos(\varepsilon^{-1}\psi_1^{\lambda\varepsilon}).\end{aligned}$$

Here, we identified $v_1^{\lambda\varepsilon} := u_1^{\lambda\varepsilon}\cos(\varepsilon^{-1}\psi_1^{\lambda\varepsilon}) - u_1^{\lambda 0}\cos(\varepsilon^{-1}\psi_1^{\lambda 0})$, where

$$u_1^{\lambda\varepsilon} := \frac{\theta_\varepsilon^\lambda (D_t L_\varepsilon^\lambda)^2}{8(\dot\phi_\varepsilon^\lambda)^2}, \qquad \psi_1^{\lambda\varepsilon} := 4\phi_\varepsilon^\lambda,$$

according to Lemma 5.7. The necessary assumptions on $u_1^{\lambda\varepsilon}$ and $\psi_1^{\lambda\varepsilon}$ are satisfied by (13) and Lemma 5.8. Consequently, it follows that $\dot v_1^{\lambda\varepsilon} \overset{*}{\rightharpoonup} 0$ in $L^\infty([0,T])$. Moreover, since $\{\dot u_1^{\lambda\varepsilon}\cos(\varepsilon^{-1}\psi_1^{\lambda\varepsilon})\}$ is a bounded sequence in $L^\infty([0,T])$, [1, Chapter I. Lemma 1] implies the equivalence of the weak* convergence of the sequence $\{\dot u_1^{\lambda\varepsilon}\cos(\varepsilon^{-1}\psi_1^{\lambda\varepsilon})\}$ and the integral convergence as in Lemma 5.6. Hence, we reason that $\dot u_1^{\lambda\varepsilon}\cos(\varepsilon^{-1}\psi_1^{\lambda\varepsilon}) \overset{*}{\rightharpoonup} 0$ in $L^\infty([0,T])$. We therefore conclude that

$$\frac{d}{dt}\left(\theta_{221}^{\lambda\varepsilon} - [\theta_2^\lambda]_{21}^\varepsilon\right) \overset{*}{\rightharpoonup} \frac{d\bar\theta_{221}^\lambda}{dt} := 0 \quad \text{in} \quad L^\infty([0,T]).$$

Finally, with $\{\dot v_1^{\lambda\varepsilon}\}$ and $\{\dot u_1^{\lambda\varepsilon}\}$ being bounded sequences in $L^\infty([0,T])$, the convergence in (41) follows for a subsequence $\{\varepsilon'\}$ from [1, Principle 4], i.e., an extended version of the Arzelà–Ascoli theorem.

For $j = 2, \dots, 6$ we only summarise the results, since the arguments follow along the same lines as in the case $j = 1$ above. In particular, we always identify terms $\dot v_j^{\lambda\varepsilon} \overset{*}{\rightharpoonup} 0$ in $L^\infty([0,T])$, according to Lemma 5.7, and terms $\dot u_j^{\lambda\varepsilon}\cos(\varepsilon^{-1}\psi_j^{\lambda\varepsilon}) \overset{*}{\rightharpoonup} \bar{\dot\theta}_{22j}^\lambda$ $(j = 2, \dots, 6)$ according to Lemma 5.6. The cases $j = 7, 8$ require some different reasoning and are thus explained in more detail.

Case $j = 2$:

$$\begin{aligned}\frac{d}{dt}\left(\theta_{222}^{\lambda\varepsilon} - [\theta_2^\lambda]_{22}^\varepsilon\right) &= \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda \cdot D_t L_\varepsilon^\lambda \cdot D_t\omega_\lambda(y_\varepsilon)}{4(\dot\phi_\varepsilon^\lambda)^3}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - [\theta_2^\lambda]_{22}^\varepsilon\right) - \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda \cdot D_t L_\varepsilon^\lambda \cdot D_t\omega_\lambda(y_\varepsilon)}{4(\dot\phi_\varepsilon^\lambda)^3}\right)\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ &\overset{*}{\rightharpoonup} \frac{d\bar\theta_{222}^\lambda}{dt} := \frac{\theta_*^\lambda (D_t L_0^\lambda)^3}{2\omega_\lambda^2(y_0)} - \frac{(\theta_*^\lambda)^2 |DL_0^\lambda|^2 \cdot D_t L_0^\lambda}{4\omega_\lambda(y_0)} \quad \text{in} \quad L^\infty([0,T]).\end{aligned}$$

Case $j = 3$:

$$\begin{aligned}\frac{d}{dt}\left(\theta_{223}^{\lambda\varepsilon} - [\theta_2^\lambda]_{23}^\varepsilon\right) &= -\frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda \left\langle D^2 L_\varepsilon^\lambda \dot y_\varepsilon, \dot y_\varepsilon\right\rangle}{4(\dot\phi_\varepsilon^\lambda)^2}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) + [\theta_2^\lambda]_{23}^\varepsilon\right) + \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda \left\langle D^2 L_\varepsilon^\lambda \dot y_\varepsilon, \dot y_\varepsilon\right\rangle}{4(\dot\phi_\varepsilon^\lambda)^2}\right)\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) \\ &\overset{*}{\rightharpoonup} \frac{d\bar\theta_{223}^\lambda}{dt} := \frac{(\theta_*^\lambda)^2 \left\langle D_t DL_0^\lambda, DL_0^\lambda\right\rangle}{4\omega_\lambda(y_0)} - \frac{3\theta_*^\lambda \left\langle D^2 L_0^\lambda \dot y_0, \dot y_0\right\rangle \cdot D_t L_0^\lambda}{8\omega_\lambda^2(y_0)} \quad \text{in} \quad L^\infty([0,T]).\end{aligned}$$

Case $j = 4$:

$$\begin{aligned}\frac{d}{dt}\left(\theta_{224}^{\lambda\varepsilon} - [\theta_2^\lambda]_{24}^\varepsilon\right) &= \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda \langle DV(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{4(\dot\phi_\varepsilon^\lambda)^2}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - [\theta_2^\lambda]_{24}^\varepsilon\right) - \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda \langle DV(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{4(\dot\phi_\varepsilon^\lambda)^2}\right)\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\\ &\overset{*}{\rightharpoonup} \frac{d\bar\theta_{224}^\lambda}{dt} := \frac{3\theta_*^\lambda \langle DV(y_0), DL_0^\lambda\rangle \cdot D_t L_0^\lambda}{8\omega_\lambda^2(y_0)} \quad \text{in} \quad L^\infty([0,T]).\end{aligned}$$

Case $j = 5$:

$$\begin{aligned}\frac{d}{dt}\left(\theta_{225}^{\lambda\varepsilon} - [\theta_2^\lambda]_{25}^\varepsilon\right) &= \frac{d}{dt}\left(\sum_{\mu=1}^r \frac{\theta_\varepsilon^\lambda\theta_\varepsilon^\mu \langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{4(\dot\phi_\varepsilon^\lambda)^2}\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda) - [\theta_2^\lambda]_{25}^\varepsilon\right)\\ &\quad - \sum_{\mu=1}^r \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda\theta_\varepsilon^\mu \langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{4(\dot\phi_\varepsilon^\lambda)^2}\right)\cos(2\varepsilon^{-1}\phi_\varepsilon^\lambda)\\ &\overset{*}{\rightharpoonup} \frac{d\bar\theta_{225}^\lambda}{dt} := \sum_{\mu=1}^r \frac{3\theta_*^\lambda\theta_*^\mu \langle D\omega_\mu(y_0), DL_0^\lambda\rangle \cdot D_t L_0^\lambda}{8\omega_\lambda^2(y_0)} + \frac{(\theta_*^\lambda)^2 |DL_0^\lambda|^2 \cdot D_t L_0^\lambda}{8\omega_\lambda(y_0)} \quad \text{in} \quad L^\infty([0,T]).\end{aligned}$$

Case $j = 6$:

$$\begin{aligned}\frac{d}{dt}\left(\theta_{226}^{\lambda\varepsilon} - [\theta_2^\lambda]_{26}^\varepsilon\right) &= -\frac{d}{dt}\left(\frac{(\theta_\varepsilon^\lambda)^2 \langle D\omega_\lambda(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{16(\dot\phi_\varepsilon^\lambda)^2}\cos(4\varepsilon^{-1}\phi_\varepsilon^\lambda) + [\theta_2^\lambda]_{26}^\varepsilon\right)\\ &\quad + \frac{d}{dt}\left(\frac{(\theta_\varepsilon^\lambda)^2 \langle D\omega_\lambda(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{16(\dot\phi_\varepsilon^\lambda)^2}\right)\cos(4\varepsilon^{-1}\phi_\varepsilon^\lambda)\\ &\overset{*}{\rightharpoonup} \frac{d\bar\theta_{226}^\lambda}{dt} := 0 \quad \text{in} \quad L^\infty([0,T]).\end{aligned}$$

Case $j = 7$: Analogous to the previous cases we write

$$\begin{aligned}\frac{d}{dt}\left(\theta_{227}^{\lambda\varepsilon} - [\theta_2^\lambda]_{27}^\varepsilon\right) &= \frac{d}{dt}\left(\sum_{\substack{\mu=1\\ \mu\neq\lambda}}^r \frac{\theta_\varepsilon^\lambda\theta_\varepsilon^\mu \langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{8\dot\phi_\varepsilon^\lambda}\left\{\frac{\cos\left(2\varepsilon^{-1}(\phi_\varepsilon^\mu - \phi_\varepsilon^\lambda)\right)}{\dot\phi_\varepsilon^\mu - \dot\phi_\varepsilon^\lambda} - \frac{\cos\left(2\varepsilon^{-1}(\phi_\varepsilon^\mu + \phi_\varepsilon^\lambda)\right)}{\dot\phi_\varepsilon^\mu + \dot\phi_\varepsilon^\lambda}\right\} - [\theta_2^\lambda]_{27}^\varepsilon\right)\\ &\quad + \sum_{\substack{\mu=1\\ \mu\neq\lambda}}^r \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda\theta_\varepsilon^\mu \langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{8\dot\phi_\varepsilon^\lambda(\dot\phi_\varepsilon^\mu + \dot\phi_\varepsilon^\lambda)}\right)\cos\left(2\varepsilon^{-1}(\phi_\varepsilon^\mu + \phi_\varepsilon^\lambda)\right)\\ &\quad - \sum_{\substack{\mu=1\\ \mu\neq\lambda}}^r \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda\theta_\varepsilon^\mu \langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{8\dot\phi_\varepsilon^\lambda(\dot\phi_\varepsilon^\mu - \dot\phi_\varepsilon^\lambda)}\right)\cos\left(2\varepsilon^{-1}(\phi_\varepsilon^\mu - \phi_\varepsilon^\lambda)\right).\end{aligned}$$

We identify, similar to the case $j = 1$, the first term on the right-hand side with $\dot v_7^{\lambda\varepsilon}$. Then, it follows from Lemma 5.7 that $\dot v_7^{\lambda\varepsilon} \overset{*}{\rightharpoonup} 0$ in $L^\infty([0,T])$. Moreover, we identify the summands in the remaining two sums on the right-hand side with functions

$$\dot u_{7\pm}^{\lambda\mu\varepsilon}\cos(\varepsilon^{-1}\psi_{7\pm}^{\lambda\mu\varepsilon}) := \frac{d}{dt}\left(\frac{\theta_\varepsilon^\lambda\theta_\varepsilon^\mu \langle D\omega_\mu(y_\varepsilon), DL_\varepsilon^\lambda\rangle}{8\dot\phi_\varepsilon^\lambda(\dot\phi_\varepsilon^\mu \pm \dot\phi_\varepsilon^\lambda)}\right)\cos\left(2\varepsilon^{-1}(\phi_\varepsilon^\mu \pm \phi_\varepsilon^\lambda)\right),$$

for $\lambda, \mu = 1, \dots, r$, $\lambda \neq \mu$. Together with (17) we expand the time derivative in $\dot u_{7\pm}^{\lambda\mu\varepsilon}$ and find, based on the non-resonance Assumptions 1 and 2, and Lemmas 5.4 and 5.5, that $\dot u_{7\pm}^{\lambda\mu\varepsilon}\cos(\varepsilon^{-1}\psi_{7\pm}^{\lambda\mu\varepsilon}) \overset{*}{\rightharpoonup} 0$ in $L^\infty([0,T])$. All together, we conclude that

$$\frac{d}{dt}\left(\theta_{227}^{\lambda\varepsilon} - [\theta_2^\lambda]_{27}^\varepsilon\right) \overset{*}{\rightharpoonup} \frac{d\bar\theta_{227}^\lambda}{dt} := 0 \quad \text{in} \quad L^\infty([0,T]).$$

Finally, we use [1, Principle 4] to derive the uniform convergence for a subsequence $\{\varepsilon'\}$ in (41).

Case $j = 8$: The convergences in (42) follow for a subsequence $\{\varepsilon'\}$ (not relabelled) from [1, Principle 4] and

$$\frac{d\theta_{228}^{\lambda\varepsilon}}{dt} = -\frac{\theta_\varepsilon^\lambda \cdot D_t L_\varepsilon^\lambda \cdot D_t^2 L_\varepsilon^\lambda}{8(\dot\phi_\varepsilon^\lambda)^2} + \frac{\theta_\varepsilon^\lambda \cdot D_t L_\varepsilon^\lambda \cdot D_t^2 L_\varepsilon^\lambda}{8(\dot\phi_\varepsilon^\lambda)^2}\cos(4\varepsilon^{-1}\phi_\varepsilon^\lambda) \overset{*}{\rightharpoonup} \frac{d\bar\theta_{228}^\lambda}{dt} := -\frac{\theta_*^\lambda \cdot D_t L_0^\lambda \cdot D_t^2 L_0^\lambda}{8\omega_\lambda^2(y_0)} \quad \text{in} \quad L^\infty([0,T]),$$

where we used Lemma 5.4 for the weak* convergence.

Now, by combining the cases $j = 1, \dots, 8$ notice that

$$\sum_{j=1}^{8} \frac{d\bar{\theta}^{\lambda}_{22j}}{dt} = \frac{\theta^{\lambda}_{*}(D_t L^{\lambda}_0)^3}{2\omega^2_{\lambda}(y_0)} - \frac{\theta^{\lambda}_{*} \cdot D_t L^{\lambda}_0 \cdot D^2_t L^{\lambda}_0}{2\omega^2_{\lambda}(y_0)} + \frac{(\theta^{\lambda}_{*})^2 \left\langle D_t D L^{\lambda}_0, D L^{\lambda}_0 \right\rangle}{4\omega_{\lambda}(y_0)} - \frac{(\theta^{\lambda}_{*})^2 |D L^{\lambda}_0|^2 \cdot D_t L^{\lambda}_0}{8\omega_{\lambda}(y_0)}$$

$$= -\frac{d}{dt}\left(\frac{\theta^{\lambda}_{*}(D_t L^{\lambda}_0)^2}{4\omega^2_{\lambda}(y_0)} - \frac{(\theta^{\lambda}_{*})^2 \left| D L^{\lambda}_0 \right|^2}{8\omega_{\lambda}(y_0)} \right),$$

where we used

$$D^2_t L^{\lambda}_0 = \left\langle D^2 L^{\lambda}_0 \dot{y}_0, \dot{y}_0 \right\rangle - \left\langle DV(y_0), D L^{\lambda}_0 \right\rangle - \sum_{\mu=1}^{r} \theta^{\mu}_{*} \left\langle D\omega_{\mu}(y_0), D L^{\lambda}_0 \right\rangle.$$

Finally, Eq. (23b) follows with

$$\frac{d\bar{\theta}^{\lambda}_{2}}{dt} = \sum_{i=1}^{2} \frac{d\bar{\theta}^{\lambda}_{2i}}{dt} = \sum_{j=1}^{4} \frac{d\bar{\theta}^{\lambda}_{21j}}{dt} + \sum_{j=1}^{8} \frac{d\bar{\theta}^{\lambda}_{22j}}{dt}. \quad \square \tag{43}$$

Remark. It would be desirable to complement the uniform convergence result in (37) by a weak* convergence result of the form

$$\frac{d}{dt}\left(\theta^{\lambda\varepsilon}_{214} - [\theta^{\lambda}_{2}]^{\varepsilon}_{14} \right) \overset{*}{\rightharpoonup} \frac{d\bar{\theta}^{\lambda}_{214}}{dt} \quad \text{in} \quad L^{\infty}([0,T]).$$

This would allow us to extend the uniform convergence result in (22) by a weak* convergence as in (18)–(21). To this end one would need to show in the proof of Lemma 5.11, part $i = 1$, case $j = 4$, that $u^{\lambda\varepsilon}_4 = \mathcal{O}(\varepsilon)$. To do so one would need to extend Lemma 5.8 and show that the sequence $\{\varepsilon^{-1}\left(\phi^{\lambda\varepsilon}_2 - \left(\bar{\phi}^{\lambda}_2 + [\phi^{\lambda}_2]^{\varepsilon}\right)\right)\}$ is bounded in $L^{\infty}([0,T])$. This would require more notation and would significantly increase the complexity of this article. We therefore do not pursue this analysis further.

Lemma 5.12. *The extraction of a subsequence in Lemmas 5.10 and 5.11 can be discarded altogether and $(\bar{\phi}_2, \bar{\theta}_2, \bar{y}_2, \bar{p}_2)$ is the unique solution to the initial value problem (23)–(24).*

Proof. The differential equations (23a), (23c) and (23d) follow from (31)–(33) by taking the weak* limit in combination with Lemmas 5.6 and 5.7, and [1, Lemma 1]. Formula (23b) follows from (43). The initial values (24) can be derived from the uniform convergences in (19)–(22). Furthermore, since the right-hand side of (23) — and therefore the solution $(\bar{\phi}_2, \bar{\theta}_2, \bar{y}_2, \bar{p}_2) \in C^{\infty}([0,T], \mathbb{R}^{2m})$ — does not depend on the chosen subsequence, [1, Principle 5] allows us to discard the extraction of a subsequence altogether. $\square$

5.5. Higher-order asymptotic expansion and restrictions on the timescale

In the following we summarise how to derive higher-order asymptotic expansions of the solution to (11)–(12). Let us assume that we know the asymptotic expansion up to order $k-1$ and we want to derive the asymptotic expansion to kth order, i.e., for u_ε representing the functions ϕ_ε, θ_ε, y_ε or p_ε, we are looking for an asymptotic expansion of the form

$$u_\varepsilon = u_0 + \sum_{\ell=1}^{k-1} \varepsilon^{\ell} [\bar{u}_\ell]^{\varepsilon} + \varepsilon^{k} [\bar{u}_k]^{\varepsilon} + \varepsilon^{k} u^{\varepsilon}_{k+1},$$

where for $\ell = 1, \dots, k$,

$$[\bar{u}_\ell]^{\varepsilon} := \bar{u}_\ell + [u_\ell]^{\varepsilon} \overset{*}{\rightharpoonup} \bar{u}_\ell \quad \text{in} \quad L^{\infty}([0,T]), \qquad u^{\varepsilon}_{k+1} \to 0 \quad \text{in} \quad C([0,T]).$$

Two approaches can be used to derive the function $[\bar{u}_k]^{\varepsilon}$. They both rely on analysing the leading-order asymptotic expansion of

$$u^{\varepsilon}_k := \frac{u_\varepsilon - u_0}{\varepsilon^k} - \sum_{\ell=1}^{k-1} \varepsilon^{\ell-k} [\bar{u}_\ell]^{\varepsilon}.$$

The first approach relies on deriving $[\bar{u}_k]^{\varepsilon}$ directly from u^{ε}_k by applying the fundamental theorem of calculus to the function $u_\varepsilon - u_0$ and subsequently integrating the oscillatory component of the integrand $\dot{u}_\varepsilon - \dot{u}_0$ by parts to lower the exponent of the denominator ε^k. After k iterations by parts and corresponding expansions of the resulting terms, $\bar{u}_k$ and $[u_k]^{\varepsilon}$ can then be derived such that

$$u^{\varepsilon}_k - [u_k]^{\varepsilon} \to \bar{u}_k \quad \text{in} \quad C([0,T]).$$

This method was used to derive the leading-order asymptotic expansion of $\theta^{\varepsilon}_{21}$ in Lemma 5.11.

Another approach for the derivation of $[\bar{u}_k]^{\varepsilon}$ is based on an application of the extended Arzelà–Ascoli theorem. Analogous to Lemma 5.8, one shows first that the sequence $\{u^{\varepsilon}_k\}$ is uniformly bounded in $L^{\infty}([0,T])$. Then, by Alaoglu's theorem [1, Principle 3], there exists a subsequence $\{\varepsilon'\}$ such that $u^{\varepsilon}_k \overset{*}{\rightharpoonup} \bar{u}_k$ in $L^{\infty}([0,T])$. To determine $\bar{u}_k$, one chooses $[u_k]^{\varepsilon}$ such that the sequence $\{u^{\varepsilon}_k - [u_k]^{\varepsilon}\}$ is uniformly bounded in $C^{0,1}([0,T])$. Then, according to the extended Arzelà–Ascoli theorem [1, Chapter I §1], there exists a subsequence such that

$$u^{\varepsilon}_k - [u_k]^{\varepsilon} \to \bar{u}_k \quad \text{in} \quad C([0,T]), \qquad \frac{d}{dt}\left(u^{\varepsilon}_k - [u_k]^{\varepsilon} \right) \overset{*}{\rightharpoonup} \frac{d\bar{u}_k}{dt} \quad \text{in} \quad L^{\infty}([0,T]),$$

from which $\bar{u}_k$ can be determined as the solution to a system of differential equations. This approach was used to derive the leading-order asymptotic expansion of ϕ_2^ε, y_2^ε, p_2^ε in Lemma 5.10 and θ_{22}^ε in Lemma 5.11.

Remark. Theorem 5.2 provides immediately quantitative estimates on the difference between the original system $(\phi_\varepsilon, \theta_\varepsilon, y_\varepsilon, p_\varepsilon)$ and the limit system $(\phi_0, \theta_0, y_0, p_0)$ of order $\mathcal{O}(\varepsilon)$ for times up to arbitrary, but fixed T. With the second-order asymptotic expansions $\phi_0 + \varepsilon^2(\bar{\phi}_2 + [\phi_2])$, $\theta_0 + \varepsilon[\theta_1] + \varepsilon^2(\bar{\theta}_2 + [\theta_2])$, $y_0 + \varepsilon^2(\bar{y}_2 + [y_2])$ and $p_0 + \varepsilon^2(\bar{p}_2 + [p_2])$ the result provides error estimates of order better than $\mathcal{O}(\varepsilon^2)$ over the same timescale. There are other averaging approaches which deal with the differential equation only. A formal expansion in ε can also derive the equations for the averaged second-order corrections $\bar{\phi}_2, \bar{\theta}_2, \bar{y}_2, \bar{p}_2$, then error estimates need to be obtained separately, e.g., using some Gronwall and integration by parts arguments as described in [8]. The general restriction to finite timescales cannot be avoided unless the averaged correction terms vanish [8, Chap. 2], which does not hold in our situation.

6. Thermodynamic interpretation

We now give a thermodynamic interpretation of the analytic result presented in Theorem 5.2. The model problem in Section 2 describes the interaction of r (in general non-ergodic) fast and n slow degrees of freedom ($n, r \in \mathbb{N}$). A simplified model of one fast (hence, ergodic) and one slow degree of freedom was already studied in [11], where the authors similarly interpret a fast–slow system of the kind presented in Section 2 from a thermodynamic point of view. Since the thermodynamic interpretation of the model studied in [11] includes arguments that are similarly applicable to the more general model considered in this article, we will focus here on the differences and refer the interested reader for a detailed thermodynamic discussion to [11].

Fundamental in the theory of classical equilibrium thermodynamics is the transfer of energy in the form of work and heat in thermodynamic processes. This energy transfer is described by the energy relation

$$dE = dW + dQ = \sum_{j=1}^{n} F^j dy^j + TdS. \tag{44}$$

In more detail, let E be the energy of a generic thermodynamic system composed of many fast particles, such as gas particles trapped in a container with a piston. Then, the change of the system's energy dE is the sum of external work done on the system, $dW = \sum_{j=1}^{n} F^j dy^j$, where F^j are external forces exerted on the system by infinitesimal displacements of some external slow variables dy^j, and a change of heat, $dQ = TdS$, where T is the system's temperature and dS a change of entropy. Classical statistical mechanics provides the derivation of thermodynamic quantities such as temperature, entropy and external forces as the slow, average macroscale observations from the microscale dynamics in the system.

The energy transfer within a thermodynamic system in the form of work and heat also applies to mechanical systems which evolve within an environment that allow for thermodynamic interactions. A suitable thermodynamic theory for such mechanical systems was developed by L. Boltzmann and later refined by G. W. Gibbs [12], which was subsequently rederived by Hertz [13]. We will follow Hertz' line of thought. His formalisation is based on fast Hamiltonian systems that are slowly perturbed by external agents. In this setting, his theory describes how to define temperature, entropy and external forces such that the fundamental thermodynamic energy relation (44) is satisfied.

Applying Hertz' thermodynamic formalism to the model problem introduced in Section 2, we regard, similar to [11], the subsystem composed of the fast degrees of freedom z_ε^λ ($\lambda = 1, \dots, r$) as a thermodynamic system that is slowly perturbed by the interactions with the slow subsystem composed of y_ε^j ($j = 1, \dots, n$). Note that the ergodicity assumption for thermodynamic systems is not given for the fast subsystem. Nevertheless, one can still derive thermodynamic properties if one replaces time-averages with ensemble-averages (see [14]), which can be derived by averaging the trajectories not only with respect to time but also with respect to initial values assumed to be uniformly distributed over the energy surface. A more detailed explanation can be found in Appendix A.

In contrast to classical thermodynamic theory, which mainly focuses on the thermodynamic analysis of some fast dynamics that experiences some slow external influence, such as gas particles trapped in a container with a piston, our focus lies in analysing the slow dynamics that experiences some external thermodynamic effects through its interaction with the fast subsystem. This focus is motivated, for instance, by the conformal motion of a molecule in a solvent.

We will focus our attention on the energy associated to the fast degrees of freedom $E_\varepsilon^\perp$ and the residual energy $E_\varepsilon^\parallel$, which are given by

$$E_\varepsilon^\perp = \frac{1}{2}|\dot{z}_\varepsilon|^2 + \frac{1}{2}\varepsilon^{-2}\sum_{\lambda=1}^{r}\omega_\lambda^2(y_\varepsilon)(z_\varepsilon^\lambda)^2, \qquad E_\varepsilon^\parallel = E_\varepsilon - E_\varepsilon^\perp. \tag{45}$$

The evolution of the fast degrees of freedom is governed by the energy $E_\varepsilon^\perp = E_\varepsilon^\perp(z_\varepsilon, \dot{z}_\varepsilon; y_\varepsilon)$ which is subject to slowly varying external parameters given by y_ε^j ($j = 1, \dots, n$). As pointed out in [11], this framework allows us to interpret the model problem from a thermodynamic point of view by applying the thermodynamic theory of Hertz [13].

By applying Hertz' thermodynamic formalism to the fast subsystem, which is governed by the energy function $E_\varepsilon^\perp$, we derive in Appendix A, provided that $\theta_*^\lambda \neq 0$ for at least one $\lambda = 1, \dots, r$, the following expressions for the temperature T_ε, the entropy S_ε and the external force F_ε:

$$T_\varepsilon = \frac{1}{r}\sum_{\lambda=1}^{r}\theta_\varepsilon^\lambda\omega_\lambda(y_\varepsilon), \qquad S_\varepsilon = \sum_{\lambda=1}^{r}\log\left(\sum_{\mu=1}^{r}\theta_\varepsilon^\mu\frac{\omega_\mu(y_\varepsilon)}{\omega_\lambda(y_\varepsilon)}\right), \qquad F_\varepsilon = T_\varepsilon\sum_{\lambda=1}^{r}DL_\varepsilon^\lambda, \tag{46}$$

where, according to the notation introduced in (15), the vector DL_ε^λ represents the gradient of $L_\varepsilon^\lambda = \log(\omega_\lambda(y_\varepsilon))$ with respect to $y_\varepsilon \in \mathbb{R}^n$. The classical thermodynamic concepts of temperature and entropy are commonly described for systems in or near thermodynamic equilibrium, i.e., for an infinite separation of timescales, so in the limit $\varepsilon \to 0$. It is noteworthy that we derive these expressions

for finite but non-zero ε. Note that the assumption on θ_*^λ ensures, that the system exhibits a genuine scale-separation into fast and slow dynamics. Moreover, note that the temperature is the arithmetic mean of the frequencies $\omega_\lambda(y_\varepsilon)\,(\lambda = 1, \dots, r)$ weighted by their corresponding actions $\theta_\varepsilon^\lambda$. The entropy provides a measure for the pairwise weighted frequency ratios $\theta_\varepsilon^\lambda \omega_\lambda(y_\varepsilon)/\omega_\mu(y_\varepsilon)\,(\lambda, \mu = 1, \dots, r)$, while the external force primarily indicates the change of $\log(\omega_\lambda(y_\varepsilon))$ with respect to the slow coordinates y_ε.

In combination with the second-order expansion derived in Theorem 5.2 we can expand T_ε, S_ε and F_ε, and thus determine their asymptotic properties, i.e., $T_\varepsilon = T_0 + \mathcal{O}(\varepsilon)$, $F_\varepsilon = F_0 + \mathcal{O}(\varepsilon)$ and $S_\varepsilon = S_0 + \varepsilon[\bar{S}_1]^\varepsilon + \varepsilon^2[\bar{S}_2]^\varepsilon + \varepsilon^2 S_3^\varepsilon$ with $S_3^\varepsilon \to 0$ in $C([0,T])$, where

$$T_0 := \frac{1}{r}\sum_{\lambda=1}^{r} \theta_*^\lambda \omega_\lambda(y_0), \quad F_0 := T_0 \sum_{\lambda=1}^{r} DL_0^\lambda, \quad S_0 := \sum_{\lambda=1}^{r} \log\left(\sum_{\mu=1}^{r} \theta_*^\mu \frac{\omega_\mu(y_0)}{\omega_\lambda(y_0)}\right), \quad [\bar{S}_1]^\varepsilon := \frac{1}{T_0}\sum_{\lambda=1}^{r} [\theta_1^\lambda]^\varepsilon \omega_\lambda(y_0) \tag{47}$$

and

$$\begin{aligned}[\bar{S}_2]^\varepsilon := \frac{1}{T_0}\sum_{\lambda=1}^{r} \left(\bar{\theta}_2^\lambda + [\theta_2^\lambda]^\varepsilon\right)\omega_\lambda(y_0) + \frac{1}{T_0}\sum_{\lambda=1}^{r} \theta_*^\lambda \left\langle D\omega_\lambda(y_0), \bar{y}_2 + [y_2]^\varepsilon\right\rangle \\ - \sum_{\lambda=1}^{r} \left\langle DL_0^\lambda, \bar{y}_2 + [y_2]^\varepsilon\right\rangle - \frac{1}{2rT_0^2}\left(\sum_{\lambda=1}^{r} [\theta_1^\lambda]^\varepsilon \omega_\lambda(y_0)\right)^2.\end{aligned}$$

We use the expansions derived in Section 5 to analyse the energy $E_\varepsilon^\perp$ on different scales. To this end, we expand $E_\varepsilon^\perp = \sum_{\lambda=1}^r \theta_\varepsilon^\lambda \omega_\lambda(y_\varepsilon)$ and write $E_\varepsilon^\perp = E_0^\perp + \varepsilon[\bar{E}_1^\perp]^\varepsilon + \varepsilon^2[\bar{E}_2^\perp]^\varepsilon + \varepsilon^2 E_3^{\perp\varepsilon}$ with $E_3^{\perp\varepsilon} \to 0$ in $C([0,T])$, where

$$\begin{aligned}E_0^\perp &:= \sum_{\lambda=1}^{r} \theta_*^\lambda \omega_\lambda(y_0), \qquad [\bar{E}_1^\perp]^\varepsilon := \sum_{\lambda=1}^{r} \left[\theta_1^\lambda\right]^\varepsilon \omega_\lambda(y_0), \\ [\bar{E}_2^\perp]^\varepsilon &:= \sum_{\lambda=1}^{r} \left(\bar{\theta}_2^\lambda + [\theta_2^\lambda]^\varepsilon\right)\omega_\lambda(y_0) + \sum_{\lambda=1}^{r} \theta_*^\lambda \left\langle D\omega_\lambda(y_0), \bar{y}_2 + [y_2]^\varepsilon\right\rangle.\end{aligned}$$

6.1. Leading-order thermodynamics

We now analyse the energy $E_\varepsilon^\perp$ in the limit $\varepsilon \to 0$ from a thermodynamic perspective. For $\varepsilon \to 0$, the temperature, entropy and external force are given by the expressions T_0, S_0 and F_0 as in (47).

While the temperature captures the average collective dynamics of the weighted frequencies $\theta_*^\lambda \omega_\lambda(y_0)$, the entropy depends on the dynamics of the weighted frequency ratios $\theta_*^\lambda \omega_\lambda(y_0)/\omega_\mu(y_0)$. In contrast to the simplified model in [11], which can be regarded as the degenerate case of one fast degree of freedom, the entropy S_0 is constant if and only if all weighted frequency ratios $\theta_*^\lambda \omega_\lambda(y_0)/\omega_\mu(y_0)$ $(\lambda, \mu = 1, \dots, r)$ are constant, regardless of the number of fast degrees of freedom. In this case, the motion of the fast degrees of freedom can be described as a quasi-periodic motion. Thus, the entropy can be considered as an indicator of the homogeneity of the frequencies with respect to y_0 and therefore serves as a measure of chaos for the fast subsystem. In the case of one fast degree of freedom as in [11], the weighted frequency ratio is naturally constant and hence the entropy remains constant. Therefore, we can regard – in reference to classical thermodynamic theory – the leading-order dynamics of the fast subsystem in the case of a constant entropy as an adiabatic thermodynamic process and non-constant entropy as a non-adiabatic thermodynamic process. We remark that we make this thermodynamic interpretation despite the fact that the fast subsystem is non-ergodic.

Finally, by expressing the leading-order energy of the fast subsystem $E_0^\perp = \sum_{\lambda=1}^r \theta_*^\lambda \omega_\lambda(y_0)$ as a function of S_0 and y_0, it can be written as

$$E_0^\perp(S_0, y_0) = e^{S_0/r} \prod_{\lambda=1}^{r} \omega_\lambda^{1/r}(y_0).$$

As a consequence, the differential is given by

$$dE_0^\perp = \sum_{j=1}^{n} F_0^j dy_0^j + T_0 dS_0, \tag{48}$$

which coincides with the fundamental thermodynamic energy relation in (44).

6.2. Second-order thermodynamics

In contrast to the ε-independent thermodynamic expressions to leading-order discussed in Section 6.1, the asymptotic expansion terms to higher-order are ε-dependent. In particular, they contain terms that rapidly oscillate around zero, and terms that yield the average motion of the higher-order asymptotic expansions. As the thermodynamic theory aims to describe many-particle systems by their average dynamics, we analyse, similar to [11], the average dynamics of the higher-order asymptotic expansion in $E_\varepsilon^\perp$ and S_ε by studying the weak* limit of $[\bar{E}_1^\perp]^\varepsilon$, $[\bar{E}_2^\perp]^\varepsilon$, $[\bar{S}_1]^\varepsilon$ and $[\bar{S}_2]^\varepsilon$, i.e.,

$$[\bar{E}_1^\perp]^\varepsilon \overset{*}{\rightharpoonup} 0 \quad \text{in} \quad L^\infty([0,T]), \qquad [\bar{E}_2^\perp]^\varepsilon \overset{*}{\rightharpoonup} \bar{E}_2^\perp \quad \text{in} \quad L^\infty([0,T]),$$
$$[\bar{S}_1]^\varepsilon \overset{*}{\rightharpoonup} 0 \quad \text{in} \quad L^\infty([0,T]), \qquad [\bar{S}_2]^\varepsilon \overset{*}{\rightharpoonup} \bar{S}_2 \quad \text{in} \quad L^\infty([0,T]),$$

where

$$\bar{E}_2^{\perp} := \sum_{\lambda=1}^{r} \bar{\theta}_2^{\lambda} \omega_\lambda(y_0) + \sum_{\lambda=1}^{r} \theta_*^{\lambda} \langle D\omega_\lambda(y_0), \bar{y}_2 \rangle$$

and

$$\begin{aligned} \bar{S}_2 &:= \frac{1}{T_0} \sum_{\lambda=1}^{r} \bar{\theta}_2^{\lambda} \omega_\lambda(y_0) + \frac{1}{T_0} \sum_{\lambda=1}^{r} \theta_*^{\lambda} \langle D\omega_\lambda(y_0), \bar{y}_2 \rangle - \sum_{\lambda=1}^{r} \left\langle DL_0^{\lambda}, \bar{y}_2 \right\rangle - \frac{1}{16 r T_0^2} \sum_{\lambda=1}^{r} \left(\theta_*^{\lambda} \cdot D_t L_0^{\lambda}\right)^2 \\ &= \frac{\bar{E}_2^{\perp}}{T_0} - \sum_{\lambda=1}^{r} \left\langle DL_0^{\lambda}, \bar{y}_2 \right\rangle - \frac{1}{16 r T_0^2} \sum_{\lambda=1}^{r} \left(\theta_*^{\lambda} \cdot D_t L_0^{\lambda}\right)^2 . \end{aligned} \tag{49}$$

Note that the expression of the entropy is in this case not constant. Intuitively, this follows from the second-order asymptotic expansion of the slow degrees of freedom y_ε^j $(j = 1, \dots, n)$. These exhibit according to Theorem 5.2 a decomposition into slowly varying components $\bar{y}_2^j$ and rapidly varying components $[y_2^j]^\varepsilon$. The existence of this decomposition gives rise to a non-constant entropy discussed in more detail in [11]. Moreover, we notice that the last term in $\bar{S}_2$ originates from $[\bar{S}_1]^\varepsilon$, the rapidly oscillating first-order correction of S_0.

Finally, after rearranging (49), we derive for $\bar{E}_2^{\perp} = \bar{E}_2^{\perp}(\bar{S}_2, \bar{y}_2; y_0, p_0)$ the expression

$$\bar{E}_2^{\perp} := \langle F_0, \bar{y}_2 \rangle + T_0 \bar{S}_2 + \frac{1}{16 r T_0} \sum_{\lambda=1}^{r} \left(\theta_*^{\lambda} \cdot D_t L_0^{\lambda}\right)^2 .$$

Now, by analysing the differential of $\bar{E}_2^{\perp}$ for fixed (y_0, p_0), which is given by

$$d\bar{E}_2^{\perp} = \sum_{j=1}^{n} F_0^j d\bar{y}_2^j + T_0 d\bar{S}_2,$$

we find, similar to (48), a remarkable resemblance to the fundamental thermodynamic relation as presented in Eq. (44).

6.3. Analysis of the total energy

Finally, we inspect how the thermodynamic energy transfer in form of work and heat is realised in the second-order asymptotic expansion of the total energy E_ε. Recalling the analysis above, we split the total energy E_ε into $E_\varepsilon^{\perp}$ and $E_\varepsilon^{\parallel}$ (compare with (45)), i.e., $E_\varepsilon = E_\varepsilon^{\perp} + E_\varepsilon^{\parallel}$, where

$$E_\varepsilon^{\parallel} := \frac{1}{2}|p_\varepsilon|^2 + V(y_\varepsilon) + \frac{\varepsilon}{2} \sum_{\lambda=1}^{r} \theta_\varepsilon^{\lambda} \left\langle p_\varepsilon, DL_\varepsilon^{\lambda} \right\rangle \sin(2\varepsilon^{-1}\phi_\varepsilon^{\lambda}) + \frac{\varepsilon^2}{8} \sum_{\lambda=1}^{r} \sum_{\mu=1}^{r} \theta_\varepsilon^{\lambda} \theta_\varepsilon^{\mu} \left\langle DL_\varepsilon^{\lambda}, DL_\varepsilon^{\mu} \right\rangle \sin(2\varepsilon^{-1}\phi_\varepsilon^{\lambda}) \sin(2\varepsilon^{-1}\phi_\varepsilon^{\mu}).$$

Similar to before, we use the expressions derived in Theorem 5.2 to expand the energy $E_\varepsilon^{\parallel}$, i.e., $E_\varepsilon^{\parallel} = E_0^{\parallel} + \varepsilon[\bar{E}_1^{\parallel}]^\varepsilon + \varepsilon^2[\bar{E}_2^{\parallel}]^\varepsilon + \varepsilon^2 E_3^{\parallel\varepsilon}$ with $E_3^{\parallel\varepsilon} \to 0$ in $C([0,T])$, where

$$E_0^{\parallel} := \frac{1}{2}|p_0|^2 + V(y_0), \qquad [\bar{E}_1^{\parallel}]^\varepsilon := \frac{1}{2} \sum_{\lambda=1}^{r} \theta_*^{\lambda} D_t L_0^{\lambda} \sin(2\varepsilon^{-1}\phi_0^{\lambda})$$

and

$$\begin{aligned} [\bar{E}_2^{\parallel}]^\varepsilon &:= \langle p_0, \bar{p}_2 + [p_2]^\varepsilon \rangle + \langle DV(y_0), \bar{y}_2 + [y_2]^\varepsilon \rangle + \frac{1}{2} \sum_{\lambda=1}^{r} [\theta_1^{\lambda}]^\varepsilon D_t L_0^{\lambda} \sin(2\varepsilon^{-1}\phi_0^{\lambda}) \\ &\quad + \sum_{\lambda=1}^{r} \theta_\varepsilon^{\lambda} D_t L_0^{\lambda} \cos(2\varepsilon^{-1}\phi_0^{\lambda})(\bar{\phi}_2 + [\phi_2]^\varepsilon) + \frac{1}{8} \sum_{\lambda=1}^{r} \sum_{\mu=1}^{r} \theta_*^{\lambda} \theta_*^{\mu} \left\langle DL_0^{\mu}, DL_0^{\lambda} \right\rangle \sin(2\varepsilon^{-1}\phi_0^{\lambda}) \sin(2\varepsilon^{-1}\phi_0^{\mu}). \end{aligned}$$

To determine the average energy correction at first- and second-order, we take the weak* limit and derive $E_1^{\parallel\varepsilon} \overset{*}{\rightharpoonup} 0$ in $L^\infty([0,T])$ and

$$[\bar{E}_2^{\parallel}]^\varepsilon \overset{*}{\rightharpoonup} \bar{E}_2^{\parallel} := \langle p_0, \bar{p}_2 \rangle + \langle DV(y_0), \bar{y}_2 \rangle - \sum_{\lambda=1}^{r} \frac{\theta_*^{\lambda} (D_t L_0^{\lambda})^2}{4\omega_\lambda(y_0)} + \sum_{\lambda=1}^{r} \frac{(\theta_*^{\lambda})^2 |DL_0^{\lambda}|^2}{16} \quad \text{in} \quad L^\infty([0,T]).$$

The following theorem shows how the Hamiltonian character of the problem and the thermodynamic interpretation materialise for the averaged second-order energy correction $\bar{E}_2 = \bar{E}_2^{\parallel} + \bar{E}_2^{\perp}$.

Theorem 6.1. *Let (y_0, p_0) be as in (14) and $(\bar{y}_2, \bar{p}_2)$ be as in Theorem 5.2. Let $\bar{E}_2$ be the averaged second-order energy correction $\bar{E}_2 = \bar{E}_2^{\parallel} + \bar{E}_2^{\perp}$, where*

$$\bar{E}_2^{\parallel}(\bar{y}_2, \bar{p}_2; y_0, p_0) = \langle p_0, \bar{p}_2 \rangle + \langle DV(y_0), \bar{y}_2 \rangle - \sum_{\lambda=1}^{r} \frac{\theta_*^{\lambda} \langle p_0, D\omega_\lambda(y_0) \rangle^2}{4\omega_\lambda^3(y_0)} + \sum_{\lambda=1}^{r} \frac{(\theta_*^{\lambda})^2 \, |D\omega_\lambda(y_0)|^2}{16\omega_\lambda^2(y_0)}$$

and

$$\bar{E}_2^{\perp}(\bar{y}_2; y_0, p_0) = \sum_{\lambda=1}^{r} \bar{\theta}_2^{\lambda}(y_0, p_0)\omega_\lambda(y_0) + \sum_{\lambda=1}^{r} \theta_*^{\lambda} \langle D\omega_\lambda(y_0), \bar{y}_2\rangle ,$$

with

$$\bar{\theta}_2^{\lambda}(y_0, p_0) = \frac{\theta_*^{\lambda} \langle p_0, D\omega_\lambda(y_0)\rangle^2}{8\omega_\lambda^4(y_0)} + C_{\bar{\theta}_2^{\lambda}}, \qquad C_{\bar{\theta}_2^{\lambda}} = -\frac{\theta_*^{\lambda} \langle p_*, D\omega_\lambda(y_*)\rangle^2}{8\omega_\lambda^4(y_*)} - [\theta_2^{\lambda}]^{\varepsilon}(0).$$

Then the differential equations (23c) and (23d) take the form

$$\frac{d\bar{y}_2^j}{dt} = \frac{\partial \bar{E}_2}{\partial p_0^j}, \qquad \frac{d\bar{p}_2^j}{dt} = -\frac{\partial \bar{E}_2}{\partial y_0^j}, \tag{50}$$

for $j = 1, \dots, n$. Moreover, with the functions T_0, $\bar{S}_2$ and F_0 given in (47) and (49), which can be interpreted as the temperature, entropy and external force in the fast subsystem, the energy $\bar{E}_2^{\perp}$ can be written as

$$\bar{E}_2^{\perp}(\bar{S}_2, \bar{y}_2; y_0, p_0) = \langle F_0(y_0), \bar{y}_2\rangle + T_0(y_0)\bar{S}_2 + \frac{1}{16rT_0(y_0)} \sum_{\lambda=1}^{r} \frac{(\theta_*^{\lambda})^2 \langle p_0, D\omega_\lambda(y_0)\rangle^2}{\omega_\lambda^2(y_0)}.$$

With this notation, the energy $\bar{E}_2^{\perp}$ satisfies the constituent equations

$$T_0 = \frac{\partial \bar{E}_2^{\perp}}{\partial \bar{S}_2}, \qquad F_0^j = \frac{\partial \bar{E}_2^{\perp}}{\partial \bar{y}_2^j}. \tag{51}$$

Proof. The evolution equations (50) follow directly from (23c) and (23d). The constituent equations (51) follow from (44). □

Remark. With $E_\varepsilon = E_* = E_0$, according to (8), the expansion of the energy $E_\varepsilon = E_0 + \varepsilon[\bar{E}_1]^{\varepsilon} + \varepsilon^2[\bar{E}_2]^{\varepsilon} + \varepsilon^2 E_3^{\varepsilon}$ implies that $[\bar{E}_1]^{\varepsilon} = [\bar{E}_2]^{\varepsilon} = E_3^{\varepsilon} \equiv 0$ and thus $\bar{E}_2 \equiv 0$. As a consequence, the averaged energy function $\bar{E}_2$ acts as a constraint on the system and ε^2 can be regarded as a Lagrange multiplier. Note that the evolution equations (50) resemble Hamilton's canonical equations.

7. Simulations

Fast–slow Hamiltonian systems model, for example, the evolution of molecular systems, where the slow degrees of freedom represent the conformal motion of a molecule and the fast degrees of freedom represent the molecular vibrations. A crucial component in the fast–slow Hamiltonian system with the Lagrangian of Section 2 is the scale parameter ε. It often represents a fixed parameter determined by the problem in terms of the ratio of the typical timescales of the fast (here z_ε) and slow (here y_ε) degrees of freedom.

In the analysis of molecular systems, one is often primarily interested in the slow conformal motion of molecules. As such, a small scale parameter ε causes costly overhead in the numerical derivation of y_ε from (6), since the step size has to be chosen sufficiently small to account for the fast, oscillatory motion of z_ε. Theorem 2.1 provides a possible solution to this problem by deriving the homogenised system (6).

The homogenised system describes the evolution of the slow degrees of freedom y_0 only, which can be used to approximate the evolution of y_ε. The approximation of y_ε by y_0 comes, however, with a trade-off. On the one hand, one can choose a larger step size for the computation of y_0 from (9) than for that of y_ε from (6). This significantly reduces the computational cost of the numerical integration. On the other hand, approximating y_ε by y_0 introduces an approximation error which depends on the scale parameter ε, namely $\|y_\varepsilon - y_0\|_{L^\infty([0,T],\mathbb{R}^n)} = \mathcal{O}(\varepsilon^2)$. Therefore, we extend in this article the leading-order asymptotic expansion and derive in Theorem 5.2 the second-order correction $[\bar{y}_2]^{\varepsilon}$ to y_0 such that $\left\|\varepsilon^{-2}(y_\varepsilon - y_0) - [\bar{y}_2]^{\varepsilon}\right\|_{L^\infty([0,T],\mathbb{R}^n)} \to 0$ as $\varepsilon \to 0$. Here, $[\bar{y}_2]^{\varepsilon}$ takes the form $[\bar{y}_2]^{\varepsilon} = \bar{y}_2 + [y_2]^{\varepsilon}$, where $\bar{y}_2$ traces the average motion of the second-order correction and can be derived as the solution to a slow system of differential equations (50) and $[y_2]^{\varepsilon}$ is the explicitly given rapidly oscillating term of the second-order correction.

We compare the global error of approximating y_ε by y_0 and by $y_0 + \varepsilon^2[\bar{y}_2]^{\varepsilon}$ both on a short and a long time interval, and the associated computation times for a specific fast–slow Hamiltonian system described in the next paragraph. The key finding is that the computation of y_0 and $[\bar{y}_2]^{\varepsilon}$, which can be done in parallel, is up to two orders of magnitude faster than the computation of y_ε of similar accuracy. Moreover, the total computation time for y_0 and $y_0 + \varepsilon^2[\bar{y}_2]^{\varepsilon}$ is practically identical, while the global error $\|y_\varepsilon - y_0 - \varepsilon^2[\bar{y}_2]^{\varepsilon}\|_{L^\infty([0,T],\mathbb{R}^n)}$ is significantly smaller than the global error $\|y_\varepsilon - y_0\|_{L^\infty([0,T],\mathbb{R}^n)}$ on short as well as on long time intervals.

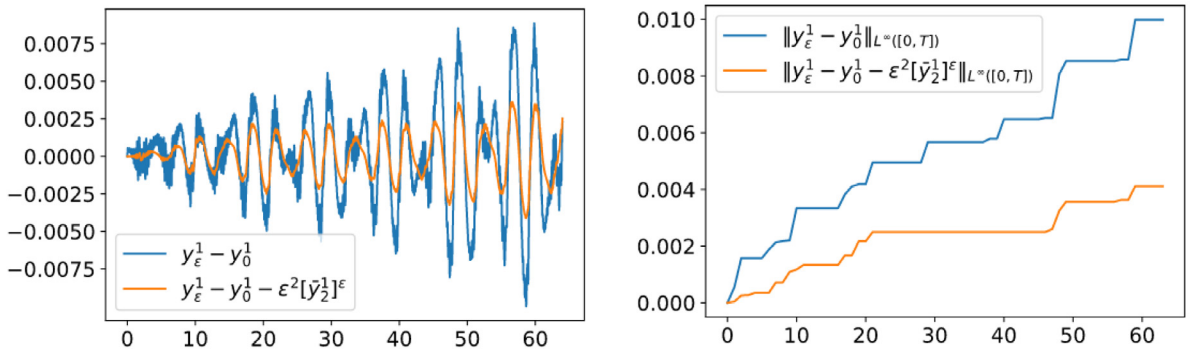
The test model. We consider a fast–slow Hamiltonian system as described in Section 2, defined on the Euclidean configuration space $M = \mathbb{R}^4$. The test model describes the evolution of two fast and two slow degrees of freedom such that $x = (y, z) \in \mathbb{R}^2 \times \mathbb{R}^2 = \mathbb{R}^4$. Their dynamics is governed by the Lagrangian as described in (3), with

$$V(y_\varepsilon) = \tfrac{1}{2}(y_\varepsilon^1)^4 + \tfrac{1}{2}(y_\varepsilon^2)^4, \qquad \omega_1(y_\varepsilon) = 4 + (y_\varepsilon^1 y_\varepsilon^2)^2, \qquad \omega_2(y_\varepsilon) = 2 + \sin(y_\varepsilon^1)$$

and initial values

$$y_\varepsilon(0) = (1, -0.5), \qquad \dot{y}_\varepsilon(0) = (1, 1.2), \qquad z_\varepsilon(0) = (0, 0), \qquad \dot{z}_\varepsilon(0) = (3, 2).$$

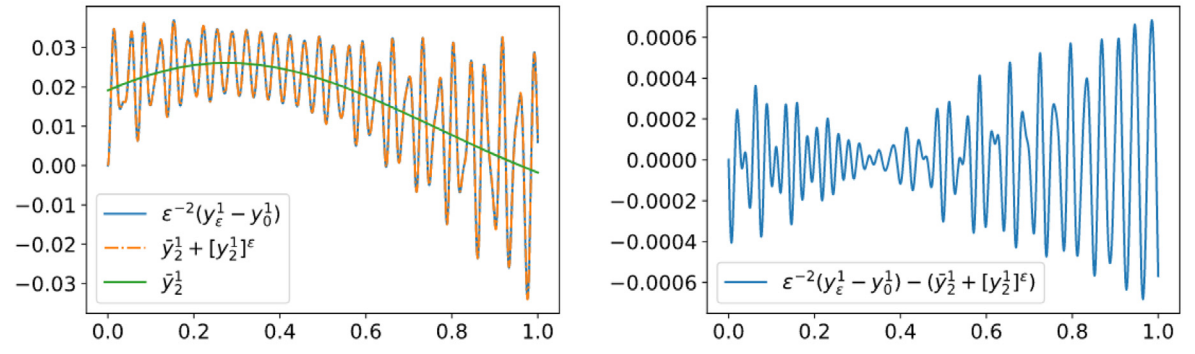
We simulate the full solution $y_\varepsilon = (y_\varepsilon^1, y_\varepsilon^2)$ and homogenised approximations, in particular the second-order approximation based on Theorems 2.1 and 5.2. Specifically, we compare the second-order asymptotic expansion $y_0^1 + \varepsilon^2[\bar{y}_2^1]^{\varepsilon}$ with the full trajectory of y_ε^1 for short and long time intervals. Here, the superscript 1 denotes the index of the first component of y_ε and indicates the first slow degree of freedom in the system. A similar comparison of the second slow degree of freedom, y_ε^2 and $y_0^2 + \varepsilon^2[\bar{y}_2^2]^{\varepsilon}$, is analogous.



(a) $y_\varepsilon^1(t) - y_0^1(t)$ and $y_\varepsilon^1(t) - y_0^1(t) - \varepsilon^2[\bar{y}_2^1]^\varepsilon(t)$ for $t \in [0, \varepsilon^{-2}]$

(b) $\left\|y_\varepsilon^1 - y_0^1\right\|_{L^\infty([0,T])}$ and $\left\|y_\varepsilon^1 - y_0^1 - \varepsilon^2[\bar{y}_2^1]^\varepsilon\right\|_{L^\infty([0,T])}$ for $T \in [0, \varepsilon^{-2}]$

Fig. 1. Comparison of the full dynamics y_ε^1 as the solution to (6), y_0^1 as the solution to the homogenised limit equation (9) and $[\bar{y}_2^1]$ as the second-order approximation, where $\bar{y}_2^1$ is derived from (50). The parameter choice is $\varepsilon = 0.5^3$.



(a) $\varepsilon^{-2}(y_\varepsilon^1(t) - y_0^1(t))$, $\bar{y}_2^1(t) + [y_2^1]^\varepsilon(t)$ and $\bar{y}_2^1(t)$ for $t \in [0, 1]$

(b) $\varepsilon^{-2}(y_\varepsilon^1(t) - y_0^1(t)) - (\bar{y}_2^1(t) + [y_2^1]^\varepsilon(t))$ for $t \in [0, 1]$

Fig. 2. Comparison of $\varepsilon^{-2}(y_\varepsilon^1(t) - y_0^1(t))$ and $\bar{y}_2^1(t) + [y_2^1]^\varepsilon(t)$ on a short time interval $t \in [0, 1]$, with y_ε^1, y_0^1 and $\bar{y}_2^1$ as the solutions of (6), (9) and (50), with $\varepsilon = 0.5^5$. The function $[y_2^1]^\varepsilon$ is given explicitly in Definition 5.1.

Fig. 1(a) displays the trajectory of $y_\varepsilon^1 - y_0^1 - \varepsilon^2[\bar{y}_2^1]^\varepsilon$ superimposed on $y_\varepsilon^1 - y_0^1$, for a long time interval with final time $T = \varepsilon^{-2}$, where $\varepsilon = 0.5^3$. It is evident that the second-order error $y_\varepsilon^1 - y_0^1 - \varepsilon^2[\bar{y}_2^1]^\varepsilon$ is significantly smaller throughout the entire time interval than the leading-order error $y_\varepsilon^1 - y_0^1$. This becomes even clearer in Fig. 1(b). There, we observe that the leading-order error grows faster than the second-order error, illustrating the increased importance of the second-order correction $[\bar{y}_2^1]^\varepsilon$ with time.

The reason why an approximation of y_ε^1 by y_0^1 performs worse than an approximation by $y_0^1 + \varepsilon^2[\bar{y}_2^1]^\varepsilon$ on long time intervals is that y_ε^1 is highly oscillatory at higher-orders, which is not captured by y_0^1. This difference becomes evident only to higher-order. Fig. 2(a) illustrates the oscillatory behaviour of y_ε^1 to second-order. Here, we superimpose the second-order correction $[\bar{y}_2^1]^\varepsilon = \bar{y}_2^1 + [y_2^1]^\varepsilon$ on top of $y_2^{1\varepsilon} = \varepsilon^{-2}(y_\varepsilon^1 - y_0^1)$ to visualise the oscillatory dynamics to higher-order and illustrate the approximation quality of $[\bar{y}_2^1]^\varepsilon$ for short time intervals. For this purpose, we integrate system (6), (9) and (50) for the test model with $\varepsilon = 0.5^5$, on a short time interval $t \in [0, 1]$. The trajectories of $y_2^{1\varepsilon}$ and $[\bar{y}_2^1]^\varepsilon$ are almost indistinguishable; the difference becomes visible only at third-order, as shown in Fig. 2(b).

Although the error $y_\varepsilon^1 - y_0^1 - \varepsilon^2[\bar{y}_2^1]^\varepsilon$ looks very accurate on short time intervals, the accuracy decreases, as Fig. 2(b) suggest, for long time intervals. Fig. 3 reveals how this error increases for long time intervals. Here, we integrated (6), (9) and (50) for $\varepsilon = 0.5^3$, where $T = \varepsilon^{-2}$.

7.1. Comparison of the execution time

As mentioned earlier, the approximation of y_ε^1 by the homogenisation limit y_0^1 comes with a trade-off. The simulation of y_0^1 is faster than that of y_ε^1 but introduces an approximation error of order $\mathcal{O}(\varepsilon^2)$. This error can be reduced by approximating y_ε^1 by $y_0^1 + \varepsilon^2[\bar{y}_2^1]^\varepsilon$, i.e., the second-order asymptotic expansion derived in Theorem 5.2. The leading-order and second-order errors are discussed in the previous section. We now discuss the computational costs of the simulations in this section.

Before comparing the total runtime for deriving y_ε^1, y_0^1 and $y_0^1 + \varepsilon^2[\bar{y}_2^1]^\varepsilon$, we note that in $[\bar{y}_2^1]^\varepsilon = \bar{y}_2^1 + [y_2^1]^\varepsilon$, the function $\bar{y}_2^1$ traces the slow, average motion of the second-order correction term and is given as the solution to (50), while $[y_2^1]^\varepsilon$ is the explicitly given rapidly oscillating term of the second-order correction. Moreover, we point out that the derivation of y_0^1 and $\bar{y}_2^1$ can be carried out in parallel. As such, there is little additional simulational overhead in computing the second-order asymptotic expansion $[\bar{y}_2^1]^\varepsilon$.

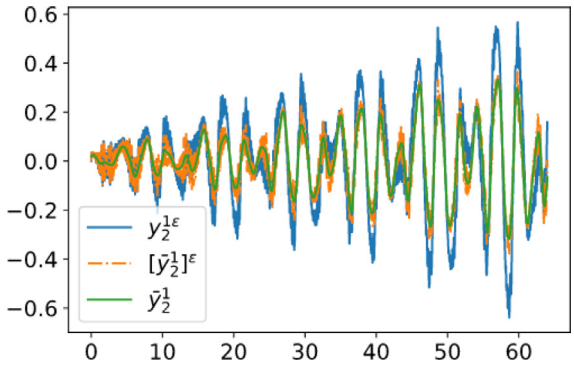


Fig. 3. The second-order correction $[\bar{y}_2^1]^\varepsilon(t) = \bar{y}_2^1(t) + [y_2^1]^\varepsilon(t)$ and its average motion $\bar{y}_2^1(t)$ superimposed on $y_2^{1\varepsilon}(t) = \varepsilon^{-2}(y_\varepsilon^1(t) - y_0^1(t))$ for $t \in [0, \varepsilon^{-2}]$ where $\varepsilon = 0.5^3$.

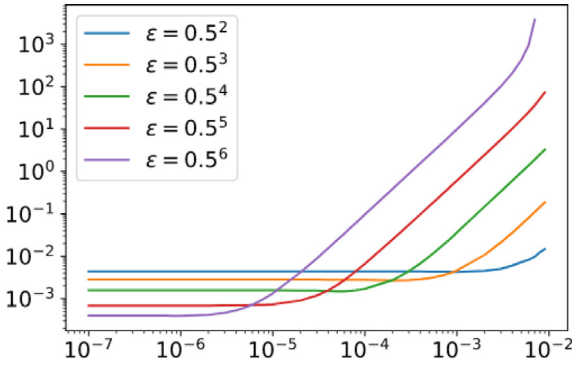


Fig. 4. Graphs of $\|y_2^{1\varepsilon} - [\bar{y}_2^1]^\varepsilon\|_{L^\infty([0,T])}$ versus step size dt_{2,y_ε^1} for different values of ε. The start of an upwards slope indicates the maximal step size.

To analyse the execution time for simulating y_ε^1, y_0^1 and $y_0^1 + \varepsilon^2[\bar{y}_2^1]^\varepsilon$, we determine the maximal step sizes such that certain convergence properties are still satisfied. More precisely, we determine the maximal step size $dt_{0,y_\varepsilon^1}^{\max}$ to compute y_ε^1 and $dt_{0,y_0^1}^{\max}$ to compute y_0^1 such that

$$\|y_\varepsilon^1 - y_0^1\|_{L^\infty([0,T])} = \mathcal{O}(\varepsilon^2) \tag{52}$$

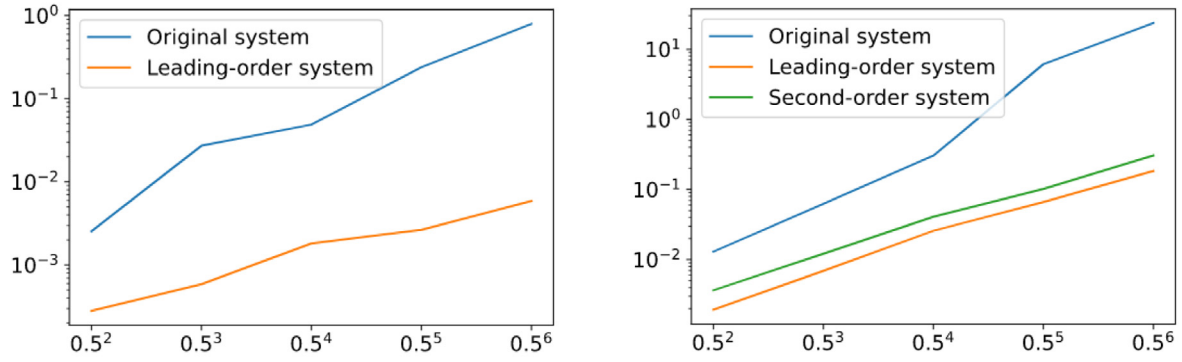
and the maximal step size $dt_{2,y_\varepsilon^1}^{\max}$ to compute y_ε^1 and $dt_{2,y_0^1,\bar{y}_2^1}^{\max}$ to compute y_0^1 and $\bar{y}_2^1$ such that

$$\|y_\varepsilon^1 - y_0^1 - \varepsilon^2[\bar{y}_2^1]^\varepsilon\|_{L^\infty([0,T])} = \mathcal{O}(\varepsilon^3). \tag{53}$$

To determine, for instance, $dt_{2,y_\varepsilon^1}^{\max}$, we fix ε and derive y_0^1 and $\bar{y}_2^1$ with a small but fixed step size $dt_{2,y_0^1,\bar{y}_2^1}$, and solve system (6) for increasingly larger dt_{2,y_ε^1}. This process results in an error plot as shown in Fig. 4. The error is constant for small step sizes dt_{2,y_ε^1} and increases after crossing an ε-dependent threshold value. This value expresses the maximal step size $dt_{2,y_\varepsilon^1}^{\max}$ which still ensures that property (53) holds. A similar procedure was applied to determine the step size $dt_{0,y_\varepsilon^1}^{\max}$ such that property (52) holds, and conversely to determine $dt_{0,y_0^1}^{\max}$ and $dt_{2,y_0^1,\bar{y}_2^1}^{\max}$.

With this procedure, we find for $\varepsilon = 0.5^k$ $(k = 2, \ldots, 7)$ the maximal step size such that the properties (52) and (53) are still satisfied. That is, for the leading-order error (52) we derive $d_{0,y_\varepsilon^1}^{\max} = \mathcal{O}(\varepsilon^2)$ and $d_{0,y_0^1}^{\max} = \mathcal{O}(\varepsilon)$, and for the second-order error (53) we obtain $dt_{2,y_\varepsilon^1}^{\max} = \mathcal{O}(\varepsilon^3)$ and $dt_{2,y_0^1,\bar{y}_2^1}^{\max} = \mathcal{O}(\varepsilon^{3/2})$. The exact maximal step sizes are listed in Tables 3 and 4 in Appendix B.

With these maximal step sizes, we can determine the average runtime for simulating y_ε^1, y_0^1 and $y_0^1 + \varepsilon^2[\bar{y}_2^1]^\varepsilon$ as depicted in Figs. 5(a) and 5(b). Most significantly, we see that the derivation of the leading-order asymptotic expansion in Fig. 5(a) and the second-order asymptotic expansion in Fig. 5(b) are up to two orders of magnitude faster than the simulation of y_ε^1 via (6). This directly reflects the differences in the maximal step sizes as explained above. Moreover, Fig. 5(b) shows that the execution time for simulating y_0^1 as the leading-order approximation and $\bar{y}_2^1$ as the averaged second-order correction are comparable. The minimal difference can be explained by the evolution equation (50), which is more complicated than in (9), resulting in an increase of floating points operations per time-step.



(a) Average runtime in seconds to solve (6) and (9) for different values of ε.

(b) Average runtime in seconds to solve (6), (9) and (50) for different values of ε.

Fig. 5. Total runtime to simulate y_ε^1 from (6), y_0^1 from (9) and $\bar{y}_2^1$ from (50). The exact computation times are listed in Tables 3 and 4 in Appendix B.

7.2. Details of the implementation

To compare the runtime of solving system (6) for y_ε with an accuracy that describes its evolution up to second-order with both the leading-order approximation (Theorem 2.1) and second-order approximation (Theorem 5.2), one needs a numerical integration scheme that allows to solve each of the three systems of differential equations (6), (9) and (50).

We note that system (6) and (9) are given as two autonomous, second-order systems of differential equations. As such, a simple Velocity-Verlet algorithm, which is frequently used in the numerical integration of molecular dynamic systems, can be used to integrate these systems. However, system (50) is non-autonomous. Thus, a numerical integration scheme from the family of Runge–Kutta methods could be used to integrate each of the three systems of differential equations. We notice that system (50) resembles Hamilton's canonical equations. In particular, the system is separable, which allows for the implementation of efficient partitioned Runge–Kutta methods. Furthermore, because of the Hamiltonian structure of systems (6) and (9), it seems natural to apply a symplectic partitioned Runge–Kutta method as an integration scheme for solving the three systems. On that account, the simulations in this article were derived on the basis of a second-order symplectic partitioned Runge–Kutta method which combines the following Lobatto IIIA (Table 1) and Lobatto IIIB (Table 2) tableaux (taken from [27, Chapter IV.5]). Sun [28] proved (also see [29]), that this specific method is symplectic. A detailed description of the implementation can be found in [30, Chapter 8 and 14].

Table 1
Lobatto IIIA.

0	0	0
1	1/2	1/2
	1/2	1/2

Table 2
Lobatto IIIB.

0	1/2	0
1	1/2	0
	1/2	1/2

8. Conclusion

In this article, we studied a class of fast–slow Hamiltonian systems with energy functions given by

$$E_\varepsilon = \frac{1}{2}|\dot{y}_\varepsilon|^2 + \frac{1}{2}|\dot{z}_\varepsilon|^2 + V(y_\varepsilon) + \frac{1}{2}\varepsilon^{-2}\sum_{\lambda=1}^{r}\omega_\lambda^2(y_\varepsilon)(z_\varepsilon^\lambda)^2,$$

where y_ε^j $(j = 1, \dots, n)$ are the slow and z_ε^λ $(\lambda = 1, \dots, r)$ are the non-ergodic fast degrees of freedom and $0 < \varepsilon < \varepsilon_0 < \infty$ is a parameter characterising their typical timescale ratio. A simplified version of one fast and one slow degree of freedom was already studied in [11].

In the first part of this article, we introduced a transformation of the fast degrees of freedom into action–angle variables $(z_\varepsilon, \dot{z}_\varepsilon) \mapsto (\theta_\varepsilon, \phi_\varepsilon)$, which also required a transformation of the momenta $\dot{y}_\varepsilon \mapsto p_\varepsilon$. We derived subsequently the second-order asymptotic expansion of the transformed degrees of freedom. Furthermore, we showed that these expansions can be decomposed into terms that oscillate rapidly around zero and slow terms that trace the average motion of the expansion. While the rapidly oscillating terms are given explicitly, the slow, average terms are given as solutions to an inhomogeneous linear system of differential equations.

In the second part of this article, we studied the fast subsystem characterised by the energy function

$$E_\varepsilon^\perp = \frac{1}{2}|\dot{z}_\varepsilon|^2 + \frac{1}{2}\varepsilon^{-2}\sum_{\lambda=1}^{r}\omega_\lambda^2(y_\varepsilon)(z_\varepsilon^\lambda)^2.$$

Guided by the thermodynamic theory for ergodic Hamiltonian systems described by Hertz, we regard the dynamics of the fast degrees of freedom z_ε^λ ($\lambda = 1, \dots, r$) as a system that is slowly perturbed by the interaction with the slow degrees of freedom y_ε^j ($j = 1, \dots, n$). Because the fast subsystem is not ergodic, we followed along the lines of [14] and replaced the time-average in classical statistical mechanics by an ensemble-average and defined otherwise, following Hertz, the temperature T_ε, the entropy S_ε and the external force F_ε of the fast subsystem.

Together with the second-order asymptotic expansion derived in the first part of this article, we expanded $E_\varepsilon^\perp$, T_ε, S_ε and F_ε. After analysing the leading-order asymptotic expansion of these terms, we found that they obey an energy relation akin to the first and second law of thermodynamics (in the sense of Carathéodory)

$$dE_0^\perp = \sum_{j=1}^{n} F_0^j dy_0^j + T_0 dS_0.$$

In contrast to the case studied in [11], the entropy is not always constant. Indeed, the entropy is constant if and only if all weighted frequency ratios $\theta_*^\lambda \omega_\lambda(y_0)/\omega_\mu(y_0)$ ($\lambda, \mu = 1, \dots, r$) are constant. In this case, the fast subsystem's dynamics is a rigid (quasi-)periodic motion. We infer that, in the case of a constant entropy, the fast subsystem can be regarded as an adiabatic thermodynamic system, while in the case of a non-constant entropy, it can be interpreted as a non-adiabatic thermodynamic system.

Remarkably, for the second-order asymptotic expansion we find, for fixed (y_0, p_0), a thermodynamic energy relation of the form

$$d\bar{E}_2^\perp = \sum_{j=1}^{n} F_0^j d\bar{y}_2^j + T_0 d\bar{S}_2.$$

With a second-order entropy expression $\bar{S}_2$ that is not constant, we can interpret the averaged second-order asymptotic dynamics as a non-adiabatic thermodynamic process.

Finally, in the third part of this article, we analysed the model problem from a numerical point of view. In particular, we compared by means of a specific test model the quality of the short- and long-term approximation of y_ε by the leading-order asymptotic expansion y_0 and by the second-order asymptotic expansion $y_0 + \varepsilon^2[\bar{y}_2]^\varepsilon$. Most importantly we found that the time interval for which $y_0 + \varepsilon^2[\bar{y}_2]^\varepsilon$ ceases to be a viable approximation of y_ε is significantly longer than for an approximation by y_0 alone. Moreover, we analysed in a series of tests how the total runtime of numerically computing y_ε, y_0 and $\bar{y}_2$ depends on the value of the scale parameter ε. We derived experimentally the largest step size so that certain convergence properties are still satisfied. In contrast to system (9) and (50), which only require the integration of slow degrees of freedom and thus allow for choosing a relatively large step size, the integration of system (6) requires the choice of a relatively small step size to accurately replicate small-scale oscillations in the numerical solution. As a consequence, we found that the runtime for simulating y_0 and $\bar{y}_2$, and thus for simulating the second-order asymptotic expansion $y_0 + \varepsilon^2[\bar{y}_2]^\varepsilon$, is up to two orders of magnitude faster than the simulation of y_ε from the original system, for a similar accuracy.

The analysis of this article is restricted to a simple Hamiltonian. A significant limitation of the current analysis is the choice of the interaction potential U in (4). The diagonal structure implies that fast modes interact only indirectly, through slow modes as intermediaries, via multiplicative coupling. Such a coupling appears in the Caldeira–Leggett Hamiltonian [31], with Lagrangian

$$\mathscr{L}(y, z, \dot{y}, \dot{z}) = \frac{1}{2}M\dot{y}^2 - V(y) + \frac{1}{2}\sum_{\lambda=1}^{r} m_\lambda \dot{z}_\lambda^2 - \frac{1}{2}\sum_{\lambda=1}^{r} m_\lambda c_\lambda z_\lambda^2 - y\sum_{\lambda=1}^{r} \omega_\lambda z_\lambda$$

with $c_\lambda > 0$. (Note in the framework of this article, the small parameter would here not be the mass ratio m_λ/M, but the limit of increasing coupling ω_λ.) For direct practical applications such as chemical reactions, for example, the evolution of the butane molecule, an extension of the results presented here to more complex potentials is required. One of the key insights of this paper is the existence of thermodynamic potentials far from equilibrium, albeit in the special situation of diagonal, or diagonalisable, interaction potentials U. If this observation holds in greater generality, then this can lead to a better understanding and better computational approaches away from equilibrium, such as a chain of atoms linked to two reservoirs assigning the outer atoms different temperatures. This is a matter of future investigation.

CRediT authorship contribution statement

Matthias Klar: Conceptualization, Methodology, Formal analysis, Investigation, Software, Validation, Writing – original draft, Writing – review & editing, Visualization. **Karsten Matthies:** Conceptualization, Methodology, Formal analysis, Supervision, Writing – review & editing. **Celia Reina:** Conceptualization, Writing – review & editing. **Johannes Zimmer:** Conceptualization, Methodology, Formal analysis, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank Ben Leimkuhler for stimulating discussions. MK is supported by a scholarship from the EPSRC Centre for Doctoral Training in Statistical Applied Mathematics at Bath (SAMBa), under the project EP/L015684/1. JZ gratefully acknowledges funding by a Royal Society Wolfson Research Merit Award. C. R. acknowledges support from NSF CAREER Award, United States, CMMI-2047506.

Appendix A. Hertz' approach to thermodynamics

As mentioned earlier, the authors in [11] analyse a simplified version of the model problem as presented in Section 2 from a thermodynamic point of view. More precisely, they focus on a system of one fast and one slow degree of freedom, i.e., $n = r = 1$, whose fast subsystem is by construction ergodic. That analysis builds on the thermodynamic theory described by Hertz as presented in [14]. Because of the similarity of the two models, we will focus on the differences in the derivation of the temperature, entropy and external force as given in (46) and refer the interested reader for a detailed discussion to [11].

A.1. Introduction to thermodynamics for non-ergodic systems

To illustrate the difference in the derivation of the thermodynamic quantities in [11] and here, we recall how the temperature is derived for the ergodic system studied in [11] and explain why the same approach fails for non-ergodic systems as studied in this article.

Let us start by analysing the dynamics of a generalised position $z_\varepsilon \in \mathbb{R}^r$ and momentum $\zeta_\varepsilon \in \mathbb{R}^r$ governed by a Hamiltonian of the form

$$H_\varepsilon^\perp(z_\varepsilon, \zeta_\varepsilon; y_\varepsilon) := \sum_{\lambda=1}^{r} H_\varepsilon^\lambda(z_\varepsilon, \zeta_\varepsilon; y_\varepsilon), \qquad \text{where} \qquad H_\varepsilon^\lambda(z_\varepsilon, \zeta_\varepsilon; y_\varepsilon) := \frac{1}{2}(\zeta_\varepsilon^\lambda)^2 + \frac{1}{2}\omega_\lambda^2(y_\varepsilon)(z_\varepsilon^\lambda)^2, \qquad \lambda = 1, \dots, r,$$

and $y_\varepsilon(t) = y(\varepsilon t) \in \mathbb{R}^n$ are slow external parameters with $\dot{y}_\varepsilon = \mathcal{O}(\varepsilon)$. This setting of a Hamiltonian system which is slowly perturbed by an external parameter is fundamental in the thermodynamic formulation derived by Hertz. For $\varepsilon = 0$, the unperturbed Hamiltonian is given by

$$H_0^\perp(z_0, \zeta_0; y_0) = \sum_{\lambda=1}^{r} H_0^\lambda(z_0, \zeta_0; y_0), \qquad \text{where} \qquad H_0^\lambda(z_0, \zeta_0; y_0) = \frac{1}{2}(\zeta_0^\lambda)^2 + \frac{1}{2}\omega_\lambda^2(y_0)(z_0^\lambda)^2, \qquad \lambda = 1, \dots, r,$$

and $y_0(t) = y(0) \equiv y_*$. With initial values of the form $z_0^\lambda(0) = 0$ and $\zeta_0^\lambda(0) = \sqrt{2E_*^\lambda}$ the solutions to the corresponding Hamilton's equations are then given for $\lambda = 1, \dots, r$ by

$$z_0^\lambda(t) = \sqrt{\frac{2E_*^\lambda}{\omega_\lambda^2(y_*)}} \sin\left(\omega_\lambda(y_*)t\right), \qquad \zeta_0^\lambda(t) = \sqrt{2E_*^\lambda} \cos\left(\omega_\lambda(y_*)t\right). \tag{54}$$

Moreover, we define the constant total energy

$$E_*^\perp := \sum_{\lambda=1}^{r} E_*^\lambda, \qquad \text{where} \qquad E_*^\lambda := \frac{1}{2}(\zeta_0^\lambda)^2 + \frac{1}{2}\omega_\lambda^2(y_*)(z_0^\lambda)^2, \qquad \lambda = 1, \dots, r.$$

If $r = 1$, the trajectory of (z_0, ζ_0) covers the entire energy surface $\{(z_0, \zeta_0) \in \mathbb{R}^2 : H_0^\perp(z_0, \zeta_0; y_0) = E_*^\perp\}$. Hence, the system is ergodic. In this case, which corresponds to the model studied in [11], the temperature in thermal equilibrium is defined via the time average, indicated by angle brackets $\langle\cdot\rangle$, of twice the kinetic energy. More precisely, we obtain

$$T_\lambda(E_*^\lambda, y_*) := \left\langle \zeta_0^\lambda \frac{\partial H_0^\lambda}{\partial \zeta_0^\lambda} \right\rangle = \lim_{\theta\to\infty} \frac{1}{\theta} \int_0^\theta 2E_*^\lambda \cos^2(\omega_\lambda(y_*)t)\, \mathrm{d}t = E_*^\lambda,$$

which is unique in the case $r = 1$.

However, if $r > 1$, the energies E_*^λ $(\lambda = 1, \dots, r)$ form distinct integrals of motion. This implies that the system is non-ergodic. A naïve application of the definition of temperature above results in distinct temperature expressions that are unsuitable to describe the thermodynamic state of the whole system, because their values are in general path-dependent, i.e., $T_\lambda = E_*^\lambda \neq E_*^\mu = T_\mu$ for $\lambda \neq \mu$ $(\lambda, \mu = 1, \dots, r)$. Therefore, we define as in [14] the temperature for non-ergodic systems via the ensemble-average. This gives a unique measure for the thermodynamic state of the whole system.

A.1.1. The Birkhoff–Khinchin theorem for non-ergodic systems

A suitable expression for the temperature, which provides a unique measure for the whole system, can be derived if, in addition to averaging with respect to time, one averages with respect to all uniformly distributed initial values on the energy surface, making the temperature path-independent. This ensemble average allows us to define a temperature expression as a measure of the average kinetic motion of the whole system. We follow [14] for the definition of the ensemble average and its application to Hamiltonian systems. Let $x = x(t, x_0)$ be the parametric form of the trajectory in phase-space starting at the point x_0. Then, the average value of some function ϕ with respect to any phase trajectory $x(t)$, i.e.,

$$\langle\phi(x)\rangle = \lim_{\theta\to\infty} \frac{1}{\theta} \int_0^\theta \phi(x(t, x_0))\, \mathrm{d}t,$$

depends, in general, on x_0. An ensemble of systems is given by varying initial data x_0, independent and identically distributed over the phase region $E \le H(x) \le E + \Delta E$. The probability density of x_0 in this region is constant and is equal to $(\Gamma(E + \Delta E) - \Gamma(E))^{-1}$.

The ensemble average, *E.A.* $\langle\phi\rangle$, of the function ϕ is defined by

$$E.A.\ \langle\phi\rangle := \lim_{\Delta E\to 0} \frac{1}{\Gamma(E + \Delta E) - \Gamma(E)} \int_{E \le H(x_0) \le E + \Delta E} \lim_{\theta\to\infty} \frac{1}{\theta} \int_0^\theta \phi(x(t, x_0))\, \mathrm{d}t\, \mathrm{d}x_0.$$

Suppose that the order of calculation of the integral over x_0 and $\lim_{\theta\to\infty}$ can be changed. Then

$$E.A.\ \langle\phi\rangle = \lim_{\Delta E\to 0}\frac{1}{\Gamma(E+\Delta E)-\Gamma(E)}\lim_{\theta\to\infty}\frac{1}{\theta}\int_0^\theta\int_{E\le H(x_0)\le E+\Delta E}\phi(x(t,x_0))\,\mathrm{d}x_0\,\mathrm{d}t. \tag{55}$$

The region $E \le H(x_0) \le E+\Delta E$ is invariant under the action of the phase flow $x(t,x_0)$. Hence, in calculating the integral over x_0 in (55), one can make a change of the variables $x_0 \mapsto x$. Since the determinant of this transformation is 1 by Liouville's theorem, we can write

$$\int_{E\le H(x_0)\le E+\Delta E}\phi(x(t,x_0))\,\mathrm{d}x_0 = \int_{E\le H(x)\le E+\Delta E}\phi(x)\,\mathrm{d}x.$$

Thus, the integral does not depend on time. For small ΔE, this integral is given by

$$\int_{E\le H(x)\le E+\Delta E}\phi(x)\,\mathrm{d}x \approx \Delta E\int\phi(x)\frac{\mathrm{d}\sigma}{|\nabla H|}.$$

Therefore, we arrive at an "analogous" version of the Birkhoff–Khinchin theorem: for any Hamiltonian system

$$E.A.\ \langle\phi\rangle = \frac{\int_\Sigma \phi(x)\frac{\mathrm{d}\sigma}{|\nabla H|}}{\int_\Sigma\frac{\mathrm{d}\sigma}{|\nabla H|}}. \tag{56}$$

This version of the Birkhoff–Khinchin theorem reflects the "average" (with respect to initial data) behaviour of non-ergodic Hamiltonian systems and is thus used to define the temperature in non-ergodic systems.

A.2. Derivation of thermodynamic relations in non-ergodic systems

As we saw in the previous section, the ensemble average of a function can be derived from the equality

$$E.A.\ \langle\phi\rangle = \frac{\int_\Sigma \phi(x)\frac{\mathrm{d}\sigma}{|\nabla H|}}{\int_\Sigma\frac{\mathrm{d}\sigma}{|\nabla H|}},$$

where $\Sigma = \{(z_0,\zeta_0)\in\mathbb{R}^{2r}: H_0^\perp(z_0,\zeta_0;y_*) = E_*^\perp\}$, $\mathrm{d}\sigma$ is a surface element on the energy surface and

$$|\nabla H_0^\perp| = \left[\sum_{\lambda=1}^r\left(\frac{\partial H_0^\perp}{\partial\zeta_0^\lambda}\right)^2 + \left(\frac{\partial H_0^\perp}{\partial z_0^\lambda}\right)^2\right]^{1/2}.$$

The *temperature* for non-ergodic Hamiltonian systems is defined via the ensemble average by

$$T(E_*^\perp,y_*) := E.A.\left\langle\zeta_0^\lambda\frac{\partial H_0^\perp(z_0,\zeta_0;y_*)}{\partial\zeta_0^\lambda}\right\rangle = \frac{\displaystyle\int_\Sigma\zeta_0^\lambda\frac{\partial H_0^\perp}{\partial\zeta_0^\lambda}\frac{\mathrm{d}\sigma}{|\nabla H_0^\perp|}}{\displaystyle\int_\Sigma\frac{\mathrm{d}\sigma}{|\nabla H_0^\perp|}}. \tag{57}$$

The numerator can be evaluated by noting that $\partial H_0^\perp/\partial\zeta_0^\lambda$ is the λth component of the vector $\nabla H_0^\perp$ and hence

$$n_\zeta^\lambda := \frac{\partial H_0^\perp/\partial\zeta_0^\lambda}{|\nabla H_0^\perp|}$$

is the λth component of the outer unit vector $n = \nabla H_0^\perp/|\nabla H_0^\perp|$ on the energy surface. Therefore, we can write the numerator in the form

$$\int_\Sigma\zeta_0^\lambda\frac{\partial H_0^\perp}{\partial\zeta_0^\lambda}\frac{\mathrm{d}\sigma}{|\nabla H_0^\perp|} = \int_\Sigma\zeta_0^\lambda n_\zeta^\lambda\,\mathrm{d}\sigma = \int_{H_0^\perp(z_0,\zeta_0;y_*)\le E_*^\perp}\mathrm{d}^n(z_0,\zeta_0) =: \Gamma(E_*^\perp,y_*), \tag{58}$$

which follows from Gauss' theorem, where $\Gamma(E_*^\perp,y_*)$ is the phase-space volume enclosed by the trajectories of (54). To derive the denominator in (57), we calculate the derivative of $\Gamma(E_*^\perp,y_*)$ with respect to $E_*^\perp$ and find

$$\Gamma(E_*^\perp+\Delta E_*^\perp,y_*)-\Gamma(E_*^\perp,y_*) = \int_{E_*^\perp\le H_0^\perp(z_0,\zeta_0;y_*)\le E_*^\perp+\Delta E_*^\perp}\mathrm{d}^n(z_0,\zeta_0)\approx\int_{H_0^\perp(z_0,\zeta_0;y_*)=E_*^\perp}\Delta n\,\mathrm{d}\sigma,$$

where Δn is the distance between the energy surface $H_0^\perp(z_0+n_z\Delta n,\zeta_0+n_\zeta\Delta n;y_*) = E_*^\perp+\Delta E_*^\perp$ and $H_0^\perp(z_0,\zeta_0;y_*) = E_*^\perp$. A Taylor expansion gives $\Delta n = \Delta E_*^\perp/|\nabla H_0^\perp|$ and hence

$$\frac{\partial\Gamma(E_*^\perp,y_*)}{\partial E_*^\perp} = \int_{H_0^\perp(z_0,\zeta_0;y_*)=E_*^\perp}\frac{\mathrm{d}\sigma}{|\nabla H_0^\perp|}. \tag{59}$$

Combining Eqs. (57)–(59), the temperature T can thus be expressed in terms of the phase-space volume $\Gamma(E_*^\perp,y_*)$:

$$T(E_*^\perp,y_*) = \frac{\Gamma(E_*^\perp,y_*)}{\partial\Gamma(E_*^\perp,y_*)/\partial E_*^\perp}. \tag{60}$$

Similar to [11] we integrate (60) with respect to $E_*^\perp$ and obtain for the entropy

$$S(E_*^\perp, y_*) = \log\left(\Gamma(E_*^\perp, y_*)\right) + f(y_*),$$

where $f(y_*)$ is a constant of integration with respect to $E_*^\perp$. To find the dependence of S on y_* we follow again the derivation presented in [11]. Using (56), we calculate for $j = 1, \dots, n$ the external force

$$F_j(E_*^\perp, y_*) = E.A.\left\langle \frac{\partial H_0^\perp(z_0, \zeta_0; y_*)}{\partial y_*^j} \right\rangle = \frac{\displaystyle\int_\Sigma \frac{\partial H_0^\perp}{\partial y_*^j} \frac{\mathrm{d}\sigma}{|\nabla H_0^\perp|}}{\displaystyle\int_\Sigma \frac{\mathrm{d}\sigma}{|\nabla H_0^\perp|}}. \tag{61}$$

For the numerator, we calculate the derivative of $\Gamma(E_*^\perp, y_*)$ with respect to y_*^j. Similarly to before, we have

$$\begin{aligned}\Gamma(E_*^\perp, y_* + \Delta y_*) - \Gamma(E_*^\perp, y_*) &= \int_{H_0^\perp(z_0,\zeta_0;y_*+\Delta y_*)\le E_*^\perp} \mathrm{d}^n(z_0, \zeta_0) - \int_{H_0^\perp(z_0,\zeta_0;y_*)\le E_*^\perp} \mathrm{d}^n(z_0, \zeta_0) \\ &\approx \int_{H_0^\perp(z_0,\zeta_0;y_*)=E_*^\perp} \Delta n \, \mathrm{d}\sigma,\end{aligned}$$

where Δn indicates the distance between the energy surface $H_0^\perp(z_0 + n_z \Delta n, \zeta_0 + n_\zeta \Delta n; y_*^j + \Delta y_*^j) = E_*^\perp$ and $H_0^\perp(z_0, \zeta_0; y_*) = E_*^\perp$. A Taylor expansion gives

$$\Delta n = -\frac{1}{|\nabla H_0^\perp|} \frac{\partial H_0^\perp}{\partial y_*^j} \Delta y_*^j$$

and we obtain

$$\frac{\partial \Gamma(E_*^\perp, y_*)}{\partial y_*^j} = -\int_{H_0^\perp(z_0,\zeta_0;y_*)=E_*^\perp} \frac{\partial H_0^\perp}{\partial y_*^j} \frac{\mathrm{d}\sigma}{|\nabla H_0^\perp|}. \tag{62}$$

Combining Eqs. (59), (61) and (62) we obtain

$$F_j(E_*^\perp, y_*) = E.A.\left\langle \frac{\partial H_0^\perp}{\partial y_*^j} \right\rangle = -\frac{\partial \Gamma(E_*^\perp, y_*)/\partial y_*^j}{\partial \Gamma(E_*^\perp, y_*)/\partial E_*^\perp}. \tag{63}$$

We thus find

$$S(E_*^\perp, y_*) = \log\left(\Gamma(E_*^\perp, y_*)\right) + C. \tag{64}$$

The constant C is chosen such that the entropy is dimensionless. This is the key result of Hertz' thermodynamic formulation: the explicit derivation of the entropy of a Hamiltonian system under the influence of a slowly varying parameter is (up to a constant) the logarithm of the phase-space volume.

A.3. Application to the model problem

The analysis of the previous section reveals that thermodynamic properties of Hamiltonian systems are intrinsically connected to the phase-space volume.

In general, the set $\{x \in \mathbb{R}^d : x^T \Sigma^{-1} x \le R^2\}$, where $\Sigma = \mathrm{diag}(a_1^2, a_2^2, \dots, a_d^2)$ with $a_1, \dots, a_d \in \mathbb{R}$, describes a hyperellipsoid in $\mathbb{R}^d$. Its d-dimensional volume is given by

$$\Gamma = \Gamma_d |\Sigma|^{1/2} R^d, \tag{65}$$

where Γ_d is the volume of the d-dimensional hypersphere.

To calculate the phase-space volume for the model problem as presented in Section 2 note, that the set $\{(z_0, \zeta_0) \in \mathbb{R}^{2r} : H_0^\perp(z_0, \zeta_0; y_*) = E_*^\perp\}$ with

$$E_*^\perp = \sum_{\lambda=1}^r E_*^\lambda = \sum_{\lambda=1}^r \frac{1}{2}(\zeta_0^\lambda)^2 + \frac{1}{2}\omega_\lambda^2(y_*)(z_0^\lambda)^2, \tag{66}$$

describes a hyperellipsoid in $\mathbb{R}^{2r}$. Eq. (66) can be written in the form $E_*^\perp = x^T \Sigma^{-1} x$ with $x = (z_0^1, z_0^2, \dots, z_0^r, \zeta_0^1, \zeta_0^2, \dots, \zeta_0^r)$ and

$$\Sigma = \mathrm{diag}\left(2, 2, \dots, 2, \frac{2}{\omega_1^2(y_*)}, \frac{2}{\omega_2^2(y_*)}, \dots, \frac{2}{\omega_r^2(y_*)}\right).$$

Therefore, with $d = 2r$ and $R^2 = E_*^\perp$, the volume of the hyperellipsoid (66) is according to (65) given by

$$\Gamma(E_*^\perp, y_*) = \Gamma_{2r} \frac{\left(2E_*^\perp\right)^r}{\prod_{\lambda=1}^r \omega_\lambda(y_*)}.$$

We reason by analogy that the ε-dependent phase-space volume, characterised by the energy of the fast subsystem

$$E_\varepsilon^\perp = \sum_{\lambda=1}^r E_\varepsilon^\lambda = \sum_{\lambda=1}^r \frac{1}{2}(\zeta_\varepsilon^\lambda)^2 + \frac{1}{2}\varepsilon^{-2}\omega_\lambda^2(y_\varepsilon)(z_\varepsilon^\lambda)^2,$$

is given by

$$\Gamma_\varepsilon(E_\varepsilon^\perp, y_\varepsilon) = \varepsilon^r \Gamma_{2r} \frac{\left(2E_\varepsilon^\perp\right)^r}{\prod_{\lambda=1}^r \omega_\lambda(y_\varepsilon)}. \tag{67}$$

We therefore define, provided that $\Gamma_\varepsilon(E_\varepsilon^\perp, y_\varepsilon) \neq 0$, the *temperature*, *normalised entropy* and *external force*, in analogy to (60), (63) and (64), as

$$T_\varepsilon(E_\varepsilon^\perp, y_\varepsilon) := \frac{\Gamma_\varepsilon(E_\varepsilon^\perp, y_\varepsilon)}{\partial\Gamma_\varepsilon(E_\varepsilon^\perp, y_\varepsilon)/\partial E_\varepsilon^\perp}, \quad S_\varepsilon(E_\varepsilon^\perp, y_\varepsilon) := \log\left(\Gamma_\varepsilon(E_\varepsilon^\perp, y_\varepsilon)\right) + C_\varepsilon, \quad F_\varepsilon^j(E_\varepsilon^\perp, y_\varepsilon) := -\frac{\partial\Gamma_\varepsilon(E_\varepsilon^\perp, y_\varepsilon)/\partial y_\varepsilon^j}{\partial\Gamma_\varepsilon(E_\varepsilon^\perp, y_\varepsilon)/\partial E_\varepsilon^\perp},$$

which become with (67) and $C_\varepsilon = -\log((2\varepsilon)^r \Gamma_{2r})$ for $j = 1, \ldots, n$

$$T_\varepsilon = \frac{1}{r}\sum_{\lambda=1}^r \theta_\varepsilon^\lambda \omega_\lambda(y_\varepsilon), \qquad S_\varepsilon = \sum_{\lambda=1}^r \log\left(\frac{E_\varepsilon^\perp}{\omega_\lambda(y_\varepsilon)}\right), \qquad F_\varepsilon^j = T_\varepsilon \sum_{\lambda=1}^r D_j L_\varepsilon^\lambda.$$

Appendix B. Computation times for numerical simulations

For completeness, we present in this section the total computation times corresponding to the maximal step sizes used in the simulations presented in this article. Tables 3 and 4 illustrate in column y_ε^1 the total runtime for simulations of system (6) with respect to distinct values of ε and a corresponding maximal step size as discussed in Section 7.1. Similarly, the columns y_0^1 and $\bar{y}_2^1$ indicate the total runtime for simulating systems (9) and (50). We recall that the maximal step size as discussed in Section 7.1 is given for the leading-order approximation under the theoretical global error (52) by $dt_{0,y_\varepsilon^1}^{\max} = \mathcal{O}(\varepsilon^2)$ and $dt_{0,y_0^1}^{\max} = \mathcal{O}(\varepsilon)$, and for the second-order approximation under the theoretical global error (53) by $dt_{2,y_\varepsilon^1}^{\max} = \mathcal{O}(\varepsilon^3)$ and $dt_{2,y_0^1,\bar{y}_2^1}^{\max} = \mathcal{O}(\varepsilon^{3/2})$. Note that we always chose identical step sizes for the derivation of y_0^1 and $\bar{y}_2^1$.

The source code for the numerical integration scheme was written in Python version 3.8.5. The simulations of the systems (6), (9) and (50) as presented in Tables 3 and 4 were performed on a single core Intel® Core™ i5-8250U CPU.

Table 3
Computation times in seconds for y_ε^1 and its leading-order asymptotic expansion y_0^1 for maximally viable step sizes that satisfy the theoretical global error (52).

	Computation times (s) and maximal step sizes			
ε	y_ε^1	$dt_{0,y_\varepsilon^1}^{\max}$	y_0^1	$dt_{0,y_0^1}^{\max}$
0.5^2	0.0026	1×10^{-2}	0.00028	6×10^{-2}
0.5^3	0.0271	1×10^{-3}	0.00059	3×10^{-2}
0.5^4	0.0488	5×10^{-4}	0.00182	1×10^{-2}
0.5^5	0.2392	1×10^{-4}	0.00266	7×10^{-3}
0.5^6	0.7945	3×10^{-5}	0.00590	3×10^{-3}

Table 4
Computation times in seconds for y_ε^1, its leading-order asymptotic expansion y_0^1 and averaged second-order correction $\bar{y}_2^1$ for maximally viable step sizes that satisfy the theoretical global error (53). The step sizes for deriving y_0^1 and $\bar{y}_2^1$ were chosen identical.

	Computation times (s) and maximal step sizes				
ε	y_ε^1	$dt_{2,y_\varepsilon^1}^{\max}$	y_0^1	$\bar{y}_2^1$	$dt_{2,y_0^1,\bar{y}_2^1}^{\max}$
0.5^2	0.013	2×10^{-3}	0.0019	0.0037	1×10^{-2}
0.5^3	0.062	4×10^{-4}	0.0069	0.0121	3×10^{-3}
0.5^4	0.303	8×10^{-5}	0.0258	0.0408	8×10^{-4}
0.5^5	6.122	4×10^{-6}	0.0663	0.1017	3×10^{-4}
0.5^6	23.670	1×10^{-6}	0.1838	0.3041	1×10^{-4}

References

[1] Folkmar Bornemann, Homogenization in time of singularly perturbed mechanical systems, in: Lecture Notes in Mathematics, vol. 1687, Springer-Verlag, Berlin, 1998, Revised version as of Aug 1998. http://www-m3.ma.tum.de/foswiki/pub/M3/Allgemeines/FolkmarBornemannPublications/Habil.pdf.
[2] Sabyasachi Chatterjee, Amit Acharya, Zvi Artstein, Computing singularly perturbed differential equations, J. Comput. Phys. 354 (2018) 417–446.
[3] Sabyasachi Chatterjee, et al., Plasticity without phenomenology: A first step, J. Mech. Phys. Solids 143 (2020).
[4] L.M. Lerman, E.I. Yakovlev, Geometry of slow–fast Hamiltonian systems and Painlevé equations, Indag. Math. (N.S.) 27 (5) (2016) 1219–1244, Dynamics and Geometry.
[5] Anatoly Neishtadt, Averaging method for systems with separatrix crossing, Nonlinearity 30 (7) (2017) 2871–2917.
[6] Anatoly Neishtadt, On mechanisms of destruction of adiabatic invariance in slow–fast Hamiltonian systems, Nonlinearity 32 (11) (2019) R53–R76.
[7] Kushal Shah, Dmitry Turaev, Vassili Gelfreich, Vered Rom-Kedar, Equilibration of energy in slow–fast systems, Proc. Natl. Acad. Sci. 114 (49) (2017).
[8] J.A. Sanders, F. Verhulst, J. Murdock, Averaging methods in nonlinear dynamical systems, second ed., in: Applied Mathematical Sciences, vol. 59, Springer, New York, 2007, p. xxii+431.
[9] Christian Kuehn, Multiple time scale dynamics, in: Applied Mathematical Sciences, vol. 191, Springer, Cham, 2015.
[10] Grigorios A. Pavliotis, Andrew M. Stuart, Multiscale methods, in: Texts in Applied Mathematics, vol. 53, Springer, New York, 2008.
[11] Matthias Klar, Karsten Matthies, Johannes Zimmer, Emergence of a nonconstant entropy for a fast-slow Hamiltonian system in its second-order asymptotic expansion, 2020, arXiv: 2010.10971 [math-ph].
[12] J.W. Gibbs, Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundations of Thermodynamics, C. Scribner's sons, 1902.
[13] Paul Hertz, Über die mechanischen Grundlagen der Thermodynamik, Ann. Phys. 338 (12) (1910) 225–274, 537–552.

[14] Victor L. Berdichevsky, Thermodynamics of chaos and order, in: Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 90, Longman, Harlow, 1997.
[15] J.H. Weiner, Statistical Mechanics of Elasticity, second ed., Dover Publications, 2002.
[16] Sebastian Reich, Dynamical Systems, Numerical Integration, and Exponentially Small Estimates (Habilitationsschrift), Freie Universitat Berlin, 1998.
[17] Folkmar A. Bornemann, Christof Schütte, Homogenization of Hamiltonian systems with a strong constraining potential, Physica D 102 (1–2) (1997) 57–77.
[18] Xiaoguai Li, Celia Reina, Simultaneous spatial and temporal coarse-graining: From atomistic models to continuum elastodynamics, J. Mech. Phys. Solids 130 (2019) 118–140.
[19] Zhidong Jia, Benedict J. Leimkuhler, A projective thermostatting dynamics technique, Multiscale Model. Simul. 4 (2) (2005) 563–583.
[20] Sebastian Reich, Smoothed Langevin dynamics of highly oscillatory systems, Physica D 138 (3–4) (2000) 210–224.
[21] Stephen Wiggins, Introduction to applied nonlinear dynamical systems and chaos, second ed., in: Texts in Applied Mathematics, vol. 2, Springer-Verlag, New York, 2003.
[22] V.I. Arnol'd, Mathematical methods of classical mechanics, in: Graduate Texts in Mathematics, vol. 60, Springer-Verlag, New York, 1989, Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition.
[23] Gerald Teschl, Ordinary differential equations and dynamical systems, in: Graduate Studies in Mathematics, vol. 140, American Mathematical Society, Providence, RI, 2012.
[24] David Gilbarg, Neil S. Trudinger, Elliptic partial differential equations of second order, in: Classics in Mathematics, Springer-Verlag, Berlin, 1983, Reprint of the 1998 edition.
[25] Elias M. Stein, Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals, in: Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, p. 712, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.
[26] D.S. Mitrinović, J.E. Pečarić, A.M. Fink, Inequalities involving functions and their integrals and derivatives, in: Mathematics and its Applications (East European Series), vol. 53, Kluwer Academic Publishers Group, Dordrecht, 1991.
[27] E. Hairer, G. Wanner, Solving Ordinary Differential Equations. II, in: Springer Series in Computational Mathematics, vol. 14, Springer-Verlag, Berlin, 1991, Stiff and differential-algebraic problems.
[28] Geng Sun, Symplectic partitioned Runge-Kutta methods, J. Comput. Math. 11 (4) (1993) 365–372.
[29] Laurent Jay, Symplectic partitioned Runge-Kutta methods for constrained Hamiltonian systems, SIAM J. Numer. Anal. 33 (1) (1996) 368–387.
[30] J.M. Sanz-Serna, M.P. Calvo, Numerical Hamiltonian Problems, in: Applied Mathematics and Mathematical Computation, vol. 7, Chapman & Hall, London, 1994, p. 224.
[31] A.O. Caldeira, A.J. Leggett, Influence of dissipation on quantum tunneling in macroscopic systems, Phys. Rev. Lett. 46 (1981) 211–214.