



ELSEVIER

Physica D 102 (1997) 57–77

PHYSICA D

## Homogenization of Hamiltonian systems with a strong constraining potential

Folkmar A. Bornemann<sup>a,\*</sup>, Christof Schütte<sup>b,†</sup>

<sup>a</sup> Courant Institute of Mathematical Sciences, 251 Mercer St, New York NY 10012, USA

<sup>b</sup> Konrad-Zuse-Zentrum, Takustr. 7, 14195 Berlin, Germany

Received 20 April 1996; revised 21 October 1996

Communicated by J.M. Ball

---

### Abstract

The paper studies Hamiltonian systems with a strong potential forcing the solutions to oscillate on a very small time scale. In particular, we are interested in the limit situation where the size  $\epsilon$  of this small time scale tends to zero but the velocity components remain oscillating with an amplitude variation of the order  $O(1)$ . The process of establishing an effective initial value problem for the limit positions will be called *homogenization* of the Hamiltonian system. This problem occurs in mechanics as the problem of realization of holonomic constraints, as various singular limits in fluid flow problems, in plasma physics as the problem of guiding center motion and in the simulation of biomolecules as the so-called smoothing problem. We suggest the systematic use of the notion of *weak convergence* in order to approach this problem. This methodology helps to establish unified and short proofs of many known results which throw light on the inherent structure of the problem. Moreover, we give a careful and critical review of the literature.

**Keywords:** Hamiltonian systems; Strong constraining potential; High frequency degrees of freedom; Homogenization; Weak convergence; Virial theorem; Adiabatic invariant; Realization of holonomic constraints; Guiding center; Correcting potential; Smoothing

---

### 0. Introduction

The concern of this paper is the study of Hamiltonian systems with a strong potential forcing the solution to oscillate on a time scale, which is vastly smaller than the time scale of the mean evolution. In particular we are interested in the limit situation where the size  $\epsilon$  of the small time scale is decreased to zero. Depending on the initial values three situations are possible:

- I The position and the velocity are converging pointwise as functions of time to certain limit functions as  $\epsilon \rightarrow 0$ .
- II Only the position is converging pointwise to a limit function as  $\epsilon \rightarrow 0$ . The velocity remains oscillating with an amplitude variation of order  $O(1)$ .

---

\* Corresponding author. E-mail: bornemann@na-net.ornl.gov.

† E-mail: schuette@zib.de.

III Neither position nor velocity are converging pointwise.

We will see, that the positions indeed converge pointwise if the corresponding total energies are bounded in the limit  $\epsilon \rightarrow 0$ . Thus, Case III is ruled out for *bounded energies*. Case I can be handled by standard averaging techniques of perturbation theory, cf. e.g. [21]. The remaining Case II leads to interesting results and deserves special techniques to handle the rapidly oscillating velocities. The specific problem for this Case II is to establish an *effective* initial value problem, which describes the limit solution. We decided to call this problem *homogenization* of the Hamiltonian system in order to have a clear distinction in terminology to the somewhat simpler averaging problem of Case I. This terminology seems to be justified since there is some methodical analogy to the problem of homogenization for elliptic boundary value problems [23].

A discussion of this particular homogenization problem is somewhat scattered in the literature. However, it appears at the heart of several important types of problems:

A *Realization of holonomic constraints*. In some texts on Theoretical Mechanics the question appears whether the *formalism* of the d'Alembert–Lagrange principle for holonomic constraints can be justified by introducing strong, realistic potentials, which – in the limit of infinite stiffness – force the motion to the constraints manifold. This question is discussed to some extent in the monographs [4,5,13], by means of examples in [19,35]. A mathematically exhaustive investigation of this question is given in [33], which is heavily based on the important early results of [27].

It turns out that this intuitive approach to justify the d'Alembert–Lagrange formalism *only* works for either rather special initial data (leading in fact to Case I) or for rather special constraining potentials. The interpretation of the physical meaning of these special potentials is deeply connected to that of distinction between Case I and II. The reader may find quite controversial positions in the literature, cf. [19, p. 8; 35, p. 104].

B *Singular limits in fluid flow problems*. An infinite-dimensional analog to problems of type A is provided by several singular limits of fluid dynamics. So to mention is the incompressible or Zero-Mach-Number limit of compressible fluid flows [8,11,17], and the quasi-geostrophic or Zero-Rossby-Number limit of geophysical fluid flows [12], for which Case II appears to be important in atmosphere–ocean science due to initial data that are neither in geostrophic balance nor in hydrostatic balance.

C *Guiding center of motion of charged particles in nonuniform magnetic fields*. The spiral motion – Larmor gyration – of free charges around magnetic field lines is a well-known phenomenon. The physical importance of Case II is doubtless here, since the velocity of this gyration necessarily remains  $O(1)$ . Quite early, the correct limit description of fast Larmor gyration in nonuniform magnetic fields has successfully been discovered in the physical literature [2,24,32] although unexpected and counterintuitive in the mechanical context of problem type A. These results play a key role in the explanation of magnetic traps and magnetic mirrors in plasma physics. They in fact motivated the important mathematical research of [27].

D *Corrected potentials for introducing constraints in the simulation of biomolecules*. Modeling biomolecules as classical mechanical systems leads to Hamiltonian systems with vastly different time scales. There is a strong need for eliminating the smallest time scales, because they are a severe restriction for numerical simulation. This leads to the idea of just *freezing* the high frequency degrees of freedom. However, the naive way of doing it via holonomic constraints, i.e., via the d'Alembert–Lagrange principle, is bound to produce incorrect results, since there are strong potentials present which do not fit the requirements mentioned for problems of type A. There is a need for *correcting* the weaker potentials as was first noted in [26], where such a correction was suggested on the basis of (questionable) additional physical assumptions. A detailed *mathematical* discussion of this problem can be found in further work of the present authors [31].

In this paper we approach the homogenization problem by making consequent use of the notion of *weak convergence*, which enables us to handle the velocities in a short and lucid way. To be specific, since only *averages* of the velocities are converging, we are led to certain classes of *test functions* in order to have an easy-to-use concept of

convergence. It turns out that the weak\*-convergence in  $L^\infty$  and in the space of distributions  $\mathcal{D}'$  will be appropriate for our purposes. The idea of using weak convergence for homogenization problems was systematically developed by Murat and Tartar in the mid-seventies, cf. [23] and the literature cited therein.

We do not claim to present any new results (except Theorem 2.1), but we hope that the methodical aspects of our presentation help to clarify and unify the whole business. For instance, we will show that the main difficulty of the problem is the *lack of weak continuity* of certain nonlinear functionals like squaring a function. Besides, our aim is to give a critical review of the known literature especially for problems of type A and C. To the best of our knowledge, the collected references are quite complete.

### 0.1. Organization of the paper

In Section 1 an extraction principle is established for solution sequences with bounded energy. The extracted subsequence shows a certain mixture of strong and weak convergences, which is of basic importance for the rest of the paper. We call this mixture  $\mathcal{M}$ -convergence. In Section 2 this concept is used to derive an abstract limit equation, which gives a general answer to the homogenization question. However, this equation is not *intrinsic* and therefore only of minor use. Nevertheless it provides a lot of insight into the structure of the problem and allows to establish short proofs of the more concrete answers for special situations.

Section 3 is devoted to the problems of type A, i.e., realization of holonomic constraints. We give short proofs of the known results. The general case for manifolds  $\mathcal{M}$  of codimension  $r = 1$  is discussed at length in Section 4. We show the connection to the Virial theorem of Statistical Physics and to the theory of adiabatic invariants of Hamiltonian systems.

For the sake of completeness, the general case for codimension  $r > 1$  is shortly reviewed in Section 5. It turns out that resonances and some kind of singularities may cause a nondeterministic behavior of the limit solution. This is the central result of the work of Takens [33], which implies that in general no really satisfactory answer can be given to the homogenization problem. Section 6 presents two examples for the codimension  $r = 1$  case. The first one is academic and completes some aspects of the discussion in Section 4, whereas the second one deals with the problem of type C.

### 0.2. Basic notation

For the sake of simplicity we consider a *model problem* with the following separable Hamiltonian on  $\mathbb{R}^{2d}$ :

$$H(x, \xi; \epsilon) = \frac{1}{2}|\xi|^2 + V(x) + \epsilon^{-2}U(x).$$

Throughout the paper we make the following basic *assumptions*:

- (A1)  $V \in C^\infty$  is bounded from below, i.e.,  $\inf_{x \in \mathbb{R}^d} V(x) \geq V_* > -\infty$ .
- (A2)  $U \in C^\infty$  attains its global minimum 0 on a smooth  $m$ -dimensional manifold  $\mathcal{M}$ , i.e.,  $U|_{\mathcal{M}} = 0$  and  $U(x) > 0$  for all  $x \in \mathbb{R}^d \setminus \mathcal{M}$ . The *codimension* is  $r = d - m$ .
- (A3)  $U$  is uniformly strictly convex in directions *orthogonal* to  $T\mathcal{M}$ , i.e., there is an  $\alpha > 0$  with  $\xi^T D^2 U(x) \xi \geq \alpha^2 |\xi|^2$  for all  $\xi \in N_x \mathcal{M}$ , where  $N\mathcal{M}$  denotes the normal bundle of  $\mathcal{M}$ .

We will denote the potential forces by

$$F(x) = \text{grad } V(x), \quad G(x) = \text{grad } U(x).$$

Thus, the Hamiltonian induces corresponding canonical equations of motion,

$$\epsilon^2 \ddot{x}^\epsilon + \epsilon^2 F(x^\epsilon) + G(x^\epsilon) = 0 \tag{1}$$

with initial values

$$x^\epsilon(0) = x_0^\epsilon, \quad \dot{x}^\epsilon(0) = \dot{x}_0^\epsilon.$$

We denote the energy, which is an invariant of motion, by

$$H^\epsilon = H(x_0^\epsilon, \dot{x}_0^\epsilon, \epsilon).$$

Finally, we will frequently use the following notation for matrices:  $x \otimes y = x y^T \in \mathbb{R}^{d \times d}$ , the *tensor product* of two vectors  $x, y \in \mathbb{R}^d$ , and  $A : B = \text{tr } A^T B$ , the *matrix inner product* of two matrices  $A$  and  $B$  of the same shape. This notation is common, for instance, in mathematical elasticity [9].

## 1. Bounded energy and compactness

We start our investigation with a careful study of the convergence properties of the sequence  $x^\epsilon$  for an increasingly strong potential, i.e.,  $\epsilon \rightarrow 0$ . We prove an *extraction principle* based on energy methods. It turns out that the sequence  $x^\epsilon$  approaches the manifold  $\mathcal{M}$  but with a rather different behavior along the manifold and normal to it. Now, for a small neighborhood of a compact set in  $\mathcal{M}$  it is possible to introduce uniquely the following decomposition of a point  $x \in \mathbb{R}^d$ :

$$x = x_{\mathcal{M}} + x_N, \quad x_{\mathcal{M}} \in \mathcal{M}, \quad x_N \in N_{x_{\mathcal{M}}}\mathcal{M}.$$

We view  $(x_{\mathcal{M}}, x_N)$  as a new coordinate system for this neighborhood. More precisely, the coordinates are given by pulling this decomposition back to a local bundle trivialization

$$\Omega \times (\mathbb{R}^r, \langle \cdot, \cdot \rangle) \rightarrow (N\mathcal{M}, \langle \cdot, \cdot \rangle), \quad \Omega \subset \mathbb{R}^m,$$

which obeys the metric structure. Whenever appropriate, we will – by “abus de langage” – view the coordinates as  $(x_{\mathcal{M}}, x_N) \in \Omega \times \mathbb{R}^r$ .

For the sake of short reference we introduce a special notion for the kind of convergence we are going to establish.

*Definition 1.1.* Given a sequence  $\epsilon \rightarrow 0$  and a corresponding sequence of solutions  $x^\epsilon \in C^2([0, T], \mathbb{R}^d)$  of the Eq. (1). The sequence  $\mathcal{M}$ -converges to a function  $x^0 \in C^{1,1}([0, T], \mathcal{M})$  if  $H^\epsilon \rightarrow H^0$  in  $\mathbb{R}$  and along the manifold

$$x_{\mathcal{M}}^\epsilon \rightarrow x^0 \quad \text{in } C^1([0, T], \mathcal{M}), \quad \ddot{x}_{\mathcal{M}}^\epsilon \overset{*}{\rightharpoonup} \ddot{x}^0 \quad \text{in } L^\infty([0, T], \mathbb{R}^d),$$

as well as normal to it

$$x_N^\epsilon = O(\epsilon) \quad \text{in } C([0, T], \mathbb{R}^d), \quad \dot{x}_N^\epsilon \overset{*}{\rightharpoonup} 0 \quad \text{in } L^\infty([0, T], \mathbb{R}^d)$$

and furthermore if there exist the limits

$$\eta^\epsilon = x_N^\epsilon / \epsilon \overset{*}{\rightharpoonup} \eta \quad \text{in } L^\infty([0, T], \mathbb{R}^d), \quad \eta^\epsilon \otimes \eta^\epsilon \overset{*}{\rightharpoonup} \Sigma \quad \text{in } L^\infty([0, T], \mathbb{R}^{d \times d}).$$

*Remark.* In the following we simplify the notation. All function spaces are understood to denote functions  $[0, T] \rightarrow \mathbb{R}^d$ . Terms like  $O(\epsilon)$ ,  $O(1)$  applied to functions are meant to hold in the space  $C([0, T], \mathbb{R}^d)$ .

We now state and prove our central compactness result.

**Theorem 1.2.** Let a sequence  $\epsilon \rightarrow 0$  be given, for which the initial position  $x_0^\epsilon$  as well as the initial energy  $H^\epsilon$  are bounded. Then, for each choice  $T > 0$ , there is a subsequence  $\epsilon'$ , such that  $x^{\epsilon'}$   $\mathcal{M}$ -converges to some  $x^0 \in C^{1,1}$ . Conversely, if  $x^\epsilon$   $\mathcal{M}$ -converges, the initial position  $x_0^\epsilon$  as well as the initial energy  $H^\epsilon$  are bounded.

*Proof.* Let  $x^\epsilon \in C^\infty([0, T_\epsilon], \mathbb{R}^d)$  be the unique solutions of the equations of motion with the maximal possible choice of  $T_\epsilon > 0$ . The boundedness  $H^\epsilon \leq H_*$  and  $|x_0^\epsilon| \leq K$  for all  $\epsilon$  implies

$$\frac{1}{2}|\dot{x}^\epsilon|^2 \leq H_* - V_*,$$

and therefore by integration

$$|x^\epsilon(t)| \leq |x_0^\epsilon| + t\sqrt{2(H_* - V_*)} \leq K + t\sqrt{2(H_* - V_*)}.$$

Thus, existence and uniqueness theory for ordinary differential equation shows that one can choose  $T_\epsilon = \infty$ . Fixing some *finite*  $T > 0$ , we thus have that  $x^\epsilon$  and  $\dot{x}^\epsilon$  are bounded sequences in  $C([0, T], \mathbb{R}^d)$ . By the theorems of Arzelà–Ascoli and Alaoglu (cf. [28, Theorem. 11.28/11.29], remember that  $L^\infty$  is the dual space of  $L^1$ ) we now conclude, that there is a subsequence of  $\epsilon$ , which we again denote by  $\epsilon$ , and a limit function  $x^0 \in C^{0,1}([0, T], \mathbb{R}^d) = W^{1,\infty}([0, T], \mathbb{R}^d)$  such that

$$x^\epsilon \rightarrow x^0 \text{ in } C, \quad \dot{x}^\epsilon \xrightarrow{*} \dot{x}^0 \text{ in } L^\infty.$$

Since in consequence  $\ddot{x}^\epsilon \xrightarrow{\mathcal{D}'} \ddot{x}^0$  in the sense of distributions and therefore

$$\epsilon^2 \ddot{x}^\epsilon \xrightarrow{\mathcal{D}'} 0,$$

we get by taking limits in  $\mathcal{D}'$  for Eq. (1) that  $G(x^0) = 0$ , i.e.,  $x^0$  constitutes a path in  $\mathcal{M}$ . Hence, for sufficiently small  $\epsilon$  the coordinate split  $x^\epsilon = x_{\mathcal{M}}^\epsilon + x_N^\epsilon$  makes sense and we know that  $x_N^\epsilon \rightarrow 0$  uniformly. Taylor expansion of  $U$  shows

$$\epsilon^2(H_* - V_*) \geq U(x^\epsilon) = \frac{1}{2}D^2U(x_{\mathcal{M}}^\epsilon) : x_N^\epsilon \otimes x_N^\epsilon + O(|x_N^\epsilon|^3).$$

Using assumption (A3) and the uniform convergence of  $x_N^\epsilon$  we get for sufficiently small  $\epsilon$  the estimate

$$|x_N^\epsilon|^2 \leq c(\epsilon^2 + |x_N^\epsilon|^3) \leq c\epsilon^2 + \frac{1}{2}|x_N^\epsilon|^2,$$

where  $c$  denotes some positive constant. This shows  $x_N^\epsilon = O(\epsilon)$ .

Using Lagrangian formalism it is straightforward, but tedious, to establish the equations of motion for coordinates  $(x_{\mathcal{M}}, x_N)$ . The result can be found in Appendix A of this paper. Inserting the details of the convergence established so far these equations take the form

$$M(x^\epsilon) \begin{pmatrix} \ddot{x}_{\mathcal{M}}^\epsilon \\ \ddot{x}_N^\epsilon \end{pmatrix} = \begin{pmatrix} -\epsilon^{-2}D_{\mathcal{M}}U(x^\epsilon) + O(1) \\ -\epsilon^{-2}D_NU(x^\epsilon) + O(1) \end{pmatrix}, \quad (2)$$

where  $D_{\mathcal{M}}$  denotes differentiation with respect to  $x_{\mathcal{M}}$  and  $D_N$  with respect to  $x_N$ . The mass matrix  $M(x^\epsilon)$  (Gramian matrix of the Euclidean metric in the new coordinates) takes the form

$$M(x^\epsilon) = \begin{pmatrix} O(1) & O(\epsilon) \\ O(\epsilon) & I \end{pmatrix}, \quad M(x^\epsilon)^{-1} = \begin{pmatrix} O(1) & O(\epsilon) \\ O(\epsilon) & I + O(\epsilon) \end{pmatrix}. \quad (3)$$

Note, that the metric in the normal coordinates  $x_N$  does not change, because they belong to the Euclidean subspace  $N_{x_M} \mathcal{M}$  of  $\mathbb{R}^d$ . Taylor expansion of the force terms yields

$$D_{\mathcal{M}} U(x^\epsilon) = \frac{1}{2} D_{\mathcal{M}} D^2 U(x_{\mathcal{M}}^\epsilon) : x_N^\epsilon \otimes x_N^\epsilon + O(\epsilon^3) = O(\epsilon^2)$$

and

$$D_N U(x^\epsilon) = D_N^2 U(x_{\mathcal{M}}^\epsilon) x_N^\epsilon + O(\epsilon^2) = O(\epsilon),$$

where we have used  $DU|_{\mathcal{M}} = 0$  and the estimate  $x_N^\epsilon = O(\epsilon)$ . Thus, the equations of motion (2) and the expression (3) for the inverse of the mass matrix give

$$\ddot{x}_{\mathcal{M}}^\epsilon = O(1), \quad \ddot{x}_N^\epsilon = O(\epsilon^{-1}).$$

Moreover  $x_N^\epsilon = O(\epsilon)$  implies the boundedness

$$\eta^\epsilon = x_N^\epsilon / \epsilon = O(1), \quad \eta^\epsilon \otimes \eta^\epsilon = O(1).$$

Applying the theorems of Arzelà–Ascoli and Alaoglu once more, we can – after a further extraction of subsequences – establish the missing parts of what builds up the asserted  $\mathcal{M}$ -convergence. The proof of the converse result is straightforward.  $\square$

If we do not bound the initial energy  $H^\epsilon$ , we cannot expect strong convergence of  $x^\epsilon$  nor can we expect that the limit  $x^0$  has range in  $\mathcal{M}$ .

*Example.* Consider the Hamiltonian

$$H(x, \xi; \epsilon) = \frac{1}{2} \xi^2 + \epsilon^{-2} U(x)$$

with the potential

$$U(x) = \begin{cases} \frac{1}{2} x^2, & x \leq 0, \\ 2x^2, & x \geq 0. \end{cases}$$

For the initial values  $x_0^\epsilon = 1$ ,  $\dot{x}_0^\epsilon = 0$  we get the *unbounded* energy  $H^\epsilon = 2\epsilon^{-2} \rightarrow \infty$ . The solution of the equation of motion is given by the rapidly oscillating function  $x^\epsilon(t) = x(t/\epsilon)$ , where

$$x(t) = \begin{cases} \cos(2t) & 0 \leq t \leq \frac{1}{4}\pi, \\ -2 \sin(t - \frac{1}{4}\pi) & \frac{1}{4}\pi \leq t \leq \frac{5}{4}\pi, \\ \sin(2t - \frac{5}{2}\pi) & \frac{5}{4}\pi \leq t \leq \frac{3}{2}\pi. \end{cases}$$

Here, we get merely weak convergence of  $x^\epsilon$  in  $L^\infty$ , namely

$$x^\epsilon \overset{*}{\rightharpoonup} x^0 \equiv -2/\pi = \frac{1}{3\pi/2} \int_0^{3\pi/2} x(\tau) \, d\tau,$$

which is *not* on the manifold  $\mathcal{M} = \{0\}$  defined by the minimum of  $U$ .

Corresponding to the coordinate splitting  $x^\epsilon = x_{\mathcal{M}}^\epsilon + x_N^\epsilon$  for a given  $\mathcal{M}$ -converging sequence  $x^\epsilon$  there is a splitting of the energy  $H^\epsilon$  into “intrinsic” parts along the manifold and normal to the manifold – at least asymptotically: Using the fact that  $\langle \dot{x}_{\mathcal{M}}^\epsilon, x_N^\epsilon \rangle = 0$  implies

$$\langle \dot{x}_{\mathcal{M}}^\epsilon, \dot{x}_N^\epsilon \rangle = -\langle \ddot{x}_{\mathcal{M}}^\epsilon, x_N^\epsilon \rangle = O(\epsilon),$$

we get by  $|\dot{x}^\epsilon|^2 = |\dot{x}_M^\epsilon|^2 + |\dot{x}_N^\epsilon|^2 + 2\langle \dot{x}_M^\epsilon, \dot{x}_N^\epsilon \rangle$  and Taylor expansion of the potentials  $V$  and  $U$

$$H^\epsilon = E_M^\epsilon + E_N^\epsilon + O(\epsilon)$$

with the tangential and normal energies

$$E_M^\epsilon = \frac{1}{2}|\dot{x}_M^\epsilon|^2 + V(x_M^\epsilon), \quad E_N^\epsilon = \frac{1}{2}|\dot{x}_N^\epsilon|^2 + \frac{1}{2}\epsilon^{-2}D^2U(x_M^\epsilon) : x_N^\epsilon \otimes x_N^\epsilon.$$

The kinetic and the potential part of the tangential energy  $E_M^\epsilon$  are uniformly convergent each, thus  $E_M^\epsilon \rightarrow E_M^0 = |\dot{x}^0|^2/2 + V(x^0)$  uniformly. However, each of the kinetic and potential parts of  $E_N^\epsilon$  converges only *weakly* in general. Nevertheless, since the total energy converges, we surprisingly get the uniform convergence

$$E_N^\epsilon \rightarrow H^0 - E_M^\epsilon,$$

which is of crucial importance later on.

## 2. Abstract homogenization

The question arises, whether the limit  $x^0$  of a  $\mathcal{M}$ -converging sequence  $x^\epsilon$  is itself a solution of an initial value problem constrained to the manifold  $\mathcal{M}$ . In this section we offer an abstract approach for unfolding the structure of such a limiting equation.

The starting point is the observation, that we may take the limit  $\epsilon \rightarrow 0$  in the equations of motion (1) *in the sense of distributions*. The limit  $\dot{x}^\epsilon \xrightarrow{*} \dot{x}^0$  in  $L^\infty$  implies that

$$\ddot{x}^\epsilon \xrightarrow{\mathcal{D}'} \ddot{x}^0$$

in  $\mathcal{D}'$ , in fact, even in the sense of distributions of *first order*, i.e., in  $\mathcal{D}'^1$ , cf. [15]. Thus, taking limits in (1), we get

$$\ddot{x}^0 + F(x^0) + \mathcal{D}'^1\text{-}\lim_{\epsilon \rightarrow 0} \epsilon^{-2}G(x^\epsilon) = 0. \tag{4}$$

This limit expression can be evaluated.

*Theorem 2.1.* Suppose that  $x^\epsilon$   $\mathcal{M}$ -converges to  $x^0$ . Then, the limit

$$\lambda^* = \mathcal{D}'^1\text{-}\lim_{\epsilon \rightarrow 0} \eta^\epsilon / \epsilon$$

exists as a function in  $L^\infty$  and  $x^0 \in C^{1,1}$  fulfills the equation

$$\ddot{x}^0 + F(x^0) + DG(x^0) \cdot \lambda^* + \frac{1}{2}D^2G(x^0) : \Sigma = 0 \tag{5}$$

almost everywhere. The quantities  $\eta^\epsilon$  and  $\Sigma$  are from Definition 1.1.

*Proof.* Taylor expansion of second order yields

$$\begin{aligned} \epsilon^{-2}G(x^\epsilon) &= \epsilon^{-2}(G(x^\epsilon) - G(x_M^\epsilon)) \\ &= DG(x_M^\epsilon) \cdot \frac{\eta^\epsilon}{\epsilon} + \int_0^1 (1-s)D^2G(x_M^\epsilon + s x_N^\epsilon) : (\eta^\epsilon \otimes \eta^\epsilon) ds. \end{aligned} \tag{6}$$

Now we have the convergence  $D^2G(x_{\mathcal{M}}^\epsilon + sx_N^\epsilon) \rightarrow D^2G(x^0)$  uniformly in  $s \in [0, 1]$  and  $t \in [0, T]$ . Since multiplication is continuous as the operator

$$C^0 \times (L^\infty, \text{weak-}^*\text{-topology}) \rightarrow (L^\infty, \text{weak-}^*\text{-topology}),$$

we get

$$\int_0^1 (1-s) D^2G(x_{\mathcal{M}}^\epsilon + sx_N^\epsilon) : (\eta^\epsilon \otimes \eta^\epsilon) \, ds \xrightarrow{*} \frac{1}{2} D^2G(x^0) : \Sigma. \quad (7)$$

Eq. (4) shows that  $\mathcal{D}'\text{-}\lim_{\epsilon \rightarrow 0} \epsilon^{-2}G(x^\epsilon)$  exists as a function in  $L^\infty$ . Thus, relations (6) and (7) yield the existence of the limit

$$\mathcal{D}'\text{-}\lim_{\epsilon \rightarrow 0} DG(x_{\mathcal{M}}^\epsilon) \cdot \frac{\eta^\epsilon}{\epsilon} = \hat{\lambda}$$

as a function in  $L^\infty$ . Using the  $(x_{\mathcal{M}}, x_N)$  coordinates, we have

$$DG(x_{\mathcal{M}}^\epsilon) \cdot \frac{\eta^\epsilon}{\epsilon} = \begin{pmatrix} 0 \\ D_N^2 U(x_{\mathcal{M}}^\epsilon) \eta^\epsilon / \epsilon \end{pmatrix}.$$

Note, that  $\eta^\epsilon / \epsilon \in N_{x_{\mathcal{M}}^\epsilon} \mathcal{M}$ . Our general assumption (A3) implies, that

$$D_N^2 U(x_{\mathcal{M}}^\epsilon)^{-1} \rightarrow D_N^2 U(x^0)^{-1}$$

in  $C^1$ . Thus, using the continuity of the multiplication as an operator

$$C^1 \times \mathcal{D}' \rightarrow \mathcal{D}',$$

we get the existence of the limit

$$\lambda^* = \mathcal{D}'\text{-}\lim_{\epsilon \rightarrow 0} \frac{\eta^\epsilon}{\epsilon} = \mathcal{D}'\text{-}\lim_{\epsilon \rightarrow 0} D_N^2 U(x_{\mathcal{M}}^\epsilon)^{-1} \cdot D_N^2 U(x_{\mathcal{M}}^\epsilon) \cdot \frac{\eta^\epsilon}{\epsilon} = D_N^2 U(x^0)^{-1} \hat{\lambda}$$

as a function in  $L^\infty$ . Turning back to the usual coordinates, we conclude that

$$\mathcal{D}'\text{-}\lim_{\epsilon \rightarrow 0} DG(x_{\mathcal{M}}^\epsilon) \cdot \frac{\eta^\epsilon}{\epsilon} = DG(x^0) \cdot \lambda^*,$$

which finally gives the desired limit equation.  $\square$

*Remark.* For codimension  $r = 1$ , in a somewhat different form the limit equation (5) can be found in [19, Eq. (5); 35, Eq. (8.33)] using suitable *averaging operators* to express  $\Sigma$ .

Note, that the existence of  $\mathcal{D}'\text{-}\lim_{\epsilon \rightarrow 0} \eta^\epsilon / \epsilon$  implies that

$$\eta^\epsilon \xrightarrow{*} 0. \quad (8)$$

However, in general we nevertheless obtain  $\eta^\epsilon \otimes \eta^\epsilon \xrightarrow{*} \Sigma \neq 0$ . For example,  $\eta^\epsilon(t) = \sin(t/\epsilon)$  yields  $\eta^\epsilon \xrightarrow{*} 0$  but  $\eta^\epsilon \otimes \eta^\epsilon = |\eta^\epsilon|^2 \xrightarrow{*} \Sigma = \frac{1}{2}$ . Thus, the product mapping is not weakly continuous, cf. [10]. Likewise, the first order Taylor expansion

$$\epsilon^{-2}G(x^\epsilon) = \int_0^1 DG(x_{\mathcal{M}}^\epsilon + sx_N^\epsilon) \cdot \frac{\eta^\epsilon}{\epsilon} \, ds$$

cannot be used instead of (6) to evaluate the  $\mathcal{D}^1$ -limit of the expression. The reason is, that certainly

$$\int_0^1 DG(x_{\mathcal{M}}^\epsilon + sx_N^\epsilon) ds \rightarrow DG(x^0)$$

in  $C$ , but in general *not* in  $C^1$ . Thus the limit is not simply  $DG(x^0) \cdot \lambda^*$ , since the product is *not* continuous on  $C \times \mathcal{D}^1$  as the following example shows: Take  $\phi^\epsilon(t) = \epsilon \cos(t/\epsilon)$  and  $\psi^\epsilon = \epsilon^{-1} \cos(t/\epsilon)$ . We have  $\phi^\epsilon \rightarrow 0$  in  $C$ , but not in  $C^1$ , and  $\mathcal{D}^1\text{-lim}_{\epsilon \rightarrow 0} \psi^\epsilon = 0$ , since  $\psi^\epsilon(t) = d \sin(t/\epsilon) / dt$ . However, the product converges not to zero,  $\phi^\epsilon \psi^\epsilon \xrightarrow{*} \frac{1}{2}$ .

It is possible to provide further insight into the meaning of the additional force term  $D^2G : \Sigma/2$ . The following lemma shows that the term is directly related to the limit energy dissipation of the normal components. The notion of normal energy was introduced at the end of Section 1.

**Lemma 2.2.** The normal energy  $E_N^\epsilon$  of an  $\mathcal{M}$ -converging sequence  $x^\epsilon$  obeys in the limit

$$E_N^0(t) = E_N^0(0) + \frac{1}{2} \int_0^t \langle \dot{x}^0(\tau), D^2G(x^0(\tau)) : \Sigma(\tau) \rangle d\tau$$

and asymptotically

$$E_N^\epsilon(t) = E_N^\epsilon(0) + \frac{1}{2} \int_0^t \langle \dot{x}^0(\tau), D^2G(x^0(\tau)) : \eta^\epsilon(\tau) \otimes \eta^\epsilon(\tau) \rangle d\tau + o(1).$$

*Proof.* Since  $H^0$  is a constant, differentiating the relation  $E_N^0 = H^0 - \frac{1}{2}|\dot{x}^0|^2 - V(x^0)$  for the limit normal energy with respect to time yields

$$\dot{E}_N^0 = -\langle \dot{x}^0, \ddot{x}^0 + \text{grad } V(x^0) \rangle.$$

If we insert the limit equation (5) of Theorem 2.1 and note that  $\dot{x}^0 \perp DG(x^0)\lambda^*$  because of  $DG(x^0)\lambda^* \in N_{x^0}\mathcal{M}$ ,  $\dot{x}^0 \in T_{x^0}\mathcal{M}$ , we end up with the differential equation

$$\dot{E}_N^0 = \frac{1}{2} \langle \dot{x}^0, D^2G(x^0) : \Sigma \rangle,$$

which proves the limit part of the lemma. The *uniform* asymptotics stated for  $E_N^\epsilon$  can be derived from this limit expression by using the uniform convergence of  $E_N^\epsilon$  and the weak\*-convergence  $\eta^\epsilon \otimes \eta^\epsilon \xrightarrow{*} \Sigma$ , which makes the integral term uniformly convergent to the corresponding integral term of the limit expression by a further application of the Arzelà–Ascoli theorem.  $\square$

The abstract homogenization process of this section does *not* yield an *intrinsic* description of  $x^0$  on  $\mathcal{M}$ . This is not even possible, since the “shadow”  $\Sigma$  of the normal components cannot in general be predicted by its initial value  $\Sigma(0)$  as will be explained in Section 5. However, for certain important situations it is indeed possible to derive a completely intrinsic description of the limit  $x^0$ . This will be the subject of the next two sections.

### 3. Realization of holonomic constraints

If the last force term of the limit equation (5) vanishes in the tangential direction, i.e., if

$$\frac{1}{2} D^2 G(x^0) : \Sigma \in N_{x^0} \mathcal{M}, \quad (9)$$

the limit function obeys

$$\dot{x}^0 + \text{grad } V(x^0) \in N_{x^0} \mathcal{M}, \quad (10)$$

because  $DG(x^0) \cdot \lambda^* \in N_{x^0} \mathcal{M}$  holds in any case. By the d'Alembert–Lagrange principle, relation (10) describes the motion due to the potential  $V$  under the *holonomic constraints*

$$x^0(t) \in \mathcal{M} \quad \forall t \in [0, T].$$

Thus, the limit  $\epsilon \rightarrow 0$  “realizes” holonomic constraints with potential  $V$ , *if and only if* condition (9) is fulfilled. Standard textbooks on classical mechanics like [1,4,5,22] prove the existence of a *unique* solution  $x^0 \in C^2([0, T], \mathcal{M})$  of (10) for given initial values  $\dot{x}_0^0 \in T_{x_0^0} \mathcal{M}$ .

Since the limit tangential energy  $E_{\mathcal{M}}^0 = \frac{1}{2} |\dot{x}^0|^2 + V(x^0)$  is a constant of motion of the constrained system (10), the limit normal energy  $E_N^0 = H^0 - E_{\mathcal{M}}^0$  is *necessarily* constant in time as well. However,  $E_N^0$  being a constant in time is *not sufficient* for condition (9) to hold. We will come back to this point in Section 6.

There are essentially just two cases, where one can show, that condition (9) holds. For the sake of simplicity we state our results for *fixed* initial values only. They can easily be extended to converging initial values.

#### 3.1. Case I: Vanishing normal energy

**Theorem 3.1.** Suppose the initial values satisfy

$$x_0^\epsilon = x_0^0 \in \mathcal{M}, \quad \dot{x}_0^\epsilon = \dot{x}_0^0 \in T_{x_0^0} \mathcal{M}.$$

Then, the sequence  $x^\epsilon$   $\mathcal{M}$ -converges to the unique solution  $x^0$  of the constrained system (10) with initial values  $x_0^0$  and  $\dot{x}_0^0$ .

*Proof.* The assumptions concerning the initial data imply that the initial normal energy is zero,  $E_N^\epsilon(0) = 0$ . Upon introducing the function

$$\phi^\epsilon(t) = \int_0^t |\dot{x}_N^\epsilon(\tau)|^2 + |\eta^\epsilon(\tau)|^2 d\tau,$$

we get by assumption (A3) and Lemma 2.2 the differential inequality

$$\dot{\phi}^\epsilon(t) \leq c_1 E_N^\epsilon(t) \leq \delta(\epsilon) + c_2 \phi^\epsilon(t),$$

where  $c_1$  and  $c_2$  are certain positive constants and  $\delta(\epsilon) \rightarrow 0$  for  $\epsilon \rightarrow 0$  denotes explicitly the  $o(1)$ -term of Lemma 2.2. Gronwall's inequality yields

$$\phi^\epsilon(t) \leq \delta(\epsilon) T e^{c_2 T}, \quad 0 \leq t \leq T.$$

This proves the uniform convergence of  $\phi^\epsilon \rightarrow 0$  which in turn proves the strong  $L^2$ -convergence  $\dot{x}_N^\epsilon \rightarrow 0$  and  $\eta^\epsilon \rightarrow 0$ . This strong convergence allows us to conclude that

$$\eta^\epsilon \otimes \eta^\epsilon \rightarrow 0$$

strongly, cf. [10], hence  $\Sigma = 0$ . Theorems 1.2 and 2.1 yield the existence of a subsequence  $x^{\epsilon'}$ , which  $\mathcal{M}$ -converges to the solution  $x^0$  of the constrained system (10). Since this limit is *unique* we can disregard the extraction of subsequences and have thus proved the convergence of the original sequence.  $\square$

*Remark.* The first mathematical proof of this theorem was given by Rubin and Ungar [27]. It appears in the form of an example in the textbook of Arnold [4, Chap. 17A]. For codimension  $r = 1$  one can find a discussion in [19,35]. It is restated as Theorem 9 in [5, Chap. 1, Section 6.2].

Because of the occurrence of strong convergence in this case it is possible to apply the “slow manifold” technique of Kreiss [20] which is worked out in [21] with no explicit mention, however.

An infinite-dimensional analog of the case is provided by the incompressible limit of fluid dynamics for balanced initial data. A Hamiltonian approach was given by Ebin [11], a careful perturbation analysis can be found in the work of Klainerman and Majda [17], and weak topologies are considered in [8].

The proof of Theorem 3.1 allows us to state a necessary and sufficient condition for the strong convergence of the normal velocities that only involves the given initial data:  $\dot{x}_N^\epsilon \rightarrow 0$  strongly if and only if  $E_N^0(0) = 0$ .

### 3.2. Case II: Constraining potentials with constant gully width

*Theorem 3.2.* Suppose that initial values satisfy

$$x_0^\epsilon = x_0^0 \in \mathcal{M}, \quad \dot{x}_0^\epsilon = v \in \mathbb{R}^d.$$

If the constraining potential  $U$  satisfies

$$D_N^2 U|_{\mathcal{M}} = \text{const.},$$

the sequence  $x^\epsilon$   $\mathcal{M}$ -converges to the unique solution  $x^0$  of the system (10) with initial values  $x_0^0$  and  $\dot{x}_0^0 = v_{\mathcal{M}} \in T_{x_0^0} \mathcal{M}$ , the orthogonal projection of  $v$ .

*Proof.* By Theorem 1.2 there is a subsequence  $x^{\epsilon'}$ , which  $\mathcal{M}$ -converges to a solution  $x^0$  of the limit equation (5). Since  $\eta^\epsilon$  is normal to the manifold, the tensor  $\Sigma$  takes the form

$$\Sigma = \begin{pmatrix} 0 & 0 \\ 0 & \Sigma_{NN} \end{pmatrix},$$

using  $(x_{\mathcal{M}}, x_N)$ -coordinates. For a given vector field  $X \in T_{x^0} \mathcal{M}$  we get in the metric of  $\mathcal{M}$

$$\langle D^2 G(x^0) : \Sigma, X \rangle = \langle D^2 D_X U|_{x=x^0} : \Sigma \rangle = \langle D_N^2 D_X U|_{x=x^0} : \Sigma_{NN} \rangle = \langle D_X D_N^2 U|_{x=x^0} : \Sigma_{NN} \rangle, \quad (11)$$

because the tangential derivatives  $D_X$  commutes with the normal derivative  $D_N$ . Note, that by construction the normal components are provided with an Euclidean structure which allows the use of the matrix inner product ‘:’. Since by assumption  $D_X D_N^2 U = 0$  for all  $X \in T\mathcal{M}$ , condition (9) is satisfied. Theorem 2.1 shows that  $x^0$  is the *unique* solution of the constrained system (10). This uniqueness of the limit  $x^0$  implies that already the original sequence  $x^\epsilon$   $\mathcal{M}$ -converges to  $x^0$ .  $\square$

*Remark.* An example of a potential  $U$  satisfying the condition of this theorem is provided by  $U(x) = \text{dist}(x, \mathcal{M})^2$ .

This theorem has been stated and proved by Gallavotti [13, Chap. 3, Section 3.8], who calls it “Arnold’s theorem” in view of a remark, which was made by Arnold on p. 91f of his textbook [4]. Takens [33] offers a proof under somewhat more restrictive conditions than ours, cf. his remark on p. 429. For codimension  $r = 1$  one can find a discussion in [19,27,35].

As mentioned at the beginning of this section, the limit normal energy  $E_N^0$  is constant in time. Benettin et al. [6,7] have shown that even in the case  $0 < \epsilon \ll 1$  the normal energy is nearly a constant for *exponentially large* times. To be precise they have proved the following Nekhoroshev-type of result:

$$|E_N^\epsilon(t) - E_N^\epsilon(0)| < \epsilon \quad \text{for } 0 \leq t \leq \exp(b\epsilon^{-a}),$$

where  $a$  and  $b$  are positive constants. In general, one has  $a = 1/r$ , where  $r$  denotes the codimension of the constraints manifold  $\mathcal{M}$ , but for instance, the special potential  $U(x) = \text{dist}(x, \mathcal{M})^2$  yields  $a = 1$  in any dimension. These results should be contrasted with the comparatively rather trivial estimate given by Schmidt [29, Proposition 1].

An infinite-dimensional analog is provided by certain fluid flow problems with unbalanced initial data. If the constraining forces appear to be isotropic they fulfill conditions like the one given in our theorem. For the incompressible limit consult [30] and for the quasi-geostrophic limit [12].

#### 4. The general case for codimension 1

In general, the explicit evaluation of the term

$$\frac{1}{2} D^2 G(x^0) : \Sigma \tag{12}$$

demands a careful study of the normal oscillations  $x_N^\epsilon$ . For reasons, which will become clear in Section 5, we restrict this study to the codimension  $r = 1$  case. We are slightly changing notation in this case: Since we are interested in local properties, we may assume without loss of generality that the manifold  $\mathcal{M}$  is orientable. Let  $e_N \in N\mathcal{M}$  be a smooth field of unit normal vectors. Now, the local coordinate system of points  $x$  near to the range of the limit function  $x^0$  is given by

$$x = x_{\mathcal{M}} + x_N \cdot e_N(x_{\mathcal{M}}).$$

In this way, the matrix  $\Sigma$  takes the special form

$$\Sigma = \sigma \cdot e_N(x^0) \otimes e_N(x^0),$$

where the *scalar* function  $\sigma$  is given by the limit

$$(x_N^\epsilon/\epsilon)^2 \xrightarrow{*} \sigma.$$

On the manifold  $\mathcal{M}$  the constraining potential  $U$  shows the “spring constant”

$$\omega^2(x) = D_N^2 U(x) = D^2 U(x) : (e_N(x) \otimes e_N(x)) \quad \forall x \in \mathcal{M},$$

in normal direction. Note that  $\omega(x) \geq \alpha > 0$  by assumption (A3). The normal energy resembles the energy of a harmonic oscillator,

$$E_N^\epsilon = \frac{1}{2} (\dot{x}_N^\epsilon)^2 + \frac{1}{2} \epsilon^{-2} \omega^2(x_{\mathcal{M}}^\epsilon) \cdot (x_N^\epsilon)^2.$$

In fact, the equation of motion in the  $x_N$ -coordinate is nearly that of a slowly modulated harmonic oscillator,

$$\ddot{x}_N^\epsilon + \epsilon^{-2} \omega^2(x_{\mathcal{M}}^\epsilon) x_N^\epsilon = O(1). \quad (13)$$

This equation can be established using the second order equations (2) of motion and a further Taylor expansion.

We will see, that the additional force term (12) is conservative, i.e., there is an *additional potential* yielding the equation of motion for the limit  $x^0$  as a constrained Hamiltonian system.

#### 4.1. Heuristic derivation of the additional potential

The structure of this additional potential can easily be derived, if we assume that the normal oscillation is described by the equation

$$\ddot{x}_N^\epsilon + \epsilon^{-2} \omega^2(x_{\mathcal{M}}^\epsilon) x_N^\epsilon = 0,$$

thus oversimplifying the asymptotic result (13). Now, the perturbation theory for integrable Hamiltonian systems is applicable. In fact, one can show that the action variable  $E/\omega$  of a single-frequency system is an *adiabatic invariant*, cf. [5] Chap. 5, Section 4, Theorem 23. This means that

$$\lim_{\epsilon \rightarrow 0} \frac{E_N^\epsilon(t)}{\omega(x_{\mathcal{M}}^\epsilon(t))} = \Theta = \text{const.},$$

which yields the following expression for the limit normal energy:

$$E_N^0 = \Theta \omega(x^0).$$

Thus, the term

$$H^0 = \frac{1}{2} |x^0|^2 + V(x^0) + \Theta \omega(x^0)$$

would be a first integral of the motion on  $\mathcal{M}$ . This motivates, that  $x^0$  is described by holonomic constrained Hamiltonian mechanics with the potential

$$W(x) = V(x) + V_{\text{add}}(x), \quad V_{\text{add}}(x) = \Theta \omega(x),$$

which is defined for all  $x \in \mathcal{M}$ .

#### 4.2. Rigorous derivation of the additional potential

We will base the rigorous derivation of the above given additional potential  $V_{\text{add}}$  on the fact that in the limit  $\epsilon \rightarrow 0$  the normal energy  $E_N^\epsilon$  is *equipartitioned* into its kinetic

$$T_N^\epsilon = \frac{1}{2} |\dot{x}_N^\epsilon|^2$$

and its potential part

$$U_N^\epsilon = \frac{1}{2} \epsilon^{-2} \omega^2(x_{\mathcal{M}}^\epsilon) \cdot (x_N^\epsilon)^2,$$

i.e.,  $T_N^0 = U_N^0 = \frac{1}{2} E_N^0$ . This equipartition is a well-known fact for the *time averages* of the corresponding energy parts for harmonic oscillations and is connected to the so-called *Virial theorem* of Statistical Mechanics, a *mathematical* result which has the appearance of an ergodic theorem, but no ergodicity is assumed, cf. [1,14,34].

**Lemma 4.1.** For a given  $\mathcal{M}$ -converging sequence  $x^\epsilon$  the kinetic and the potential part of the normal energy converges weakly in  $L^\infty$ ,

$$T_N^\epsilon \xrightarrow{*} T_N^0, \quad U_N^\epsilon \xrightarrow{*} U_N^0.$$

Moreover, we get

$$T_N^0 = U_N^0 = \frac{1}{2} E_N^0 = \frac{1}{2} \omega^2(x^0) \sigma.$$

*Proof.* By definition of  $\sigma$  and the strong convergence  $x_{\mathcal{M}}^\epsilon \rightarrow x^0$  we have

$$U_N^\epsilon = \frac{\omega^2(x_{\mathcal{M}}^\epsilon)}{2} \left( \frac{x_N^\epsilon}{\epsilon} \right)^2 \xrightarrow{*} \frac{\omega^2(x^0)}{2} \sigma.$$

Hence, by the uniform convergence  $E_N^\epsilon \rightarrow E_N^0$

$$T_N^\epsilon = E_N^\epsilon - U_N^\epsilon \xrightarrow{*} E_N^0 - U_N^0.$$

The next arguments follow closely the proof of the Virial theorem as given for instance in [14]:

Since  $\dot{x}_N^\epsilon$  is a bounded sequence and  $x_N^\epsilon = O(\epsilon)$ , we get the uniform convergence

$$\frac{1}{2} x_N^\epsilon \dot{x}_N^\epsilon \rightarrow 0$$

and therefore

$$\frac{d}{dt} \frac{1}{2} x_N^\epsilon \dot{x}_N^\epsilon \xrightarrow{\mathcal{D}'} 0.$$

This limit can be evaluated in a different way, using the description (13) of the normal oscillations,

$$\begin{aligned} \frac{d}{dt} \frac{1}{2} x_N^\epsilon \dot{x}_N^\epsilon &= \frac{1}{2} |\dot{x}_N^\epsilon|^2 + \frac{1}{2} x_N^\epsilon \ddot{x}_N^\epsilon = \frac{1}{2} |\dot{x}_N^\epsilon|^2 - \frac{1}{2} \epsilon^{-2} \omega^2(x_{\mathcal{M}}^\epsilon) (x_N^\epsilon)^2 + O(\epsilon) \\ &= T_N^\epsilon - U_N^\epsilon + O(\epsilon) \xrightarrow{*} T_N^0 - U_N^0, \end{aligned}$$

which gives the desired result.  $\square$

We are now able to show the *adiabatic invariance* of  $E/\omega$ .

**Theorem 4.2.** Suppose  $x^\epsilon$   $\mathcal{M}$ -converges to  $x^0$ . Then, there is a constant  $\Theta$ , such that

$$E_N^0 = \Theta \omega(x^0), \quad \sigma = \frac{\Theta}{\omega(x^0)}.$$

*Proof.* Lemma 2.2 establishes a further relation between the limit normal energy and the unknown quantity  $\sigma$ :

$$\dot{E}_N^0 = \frac{1}{2} \langle D^2 G(x^0) : \Sigma, \dot{x}^0 \rangle = \frac{1}{2} \left( \frac{d}{dt} \omega^2(x^0) \right) \cdot \sigma.$$

In the last step we have used the identity (11). On the other hand, Lemma 4.1 gives  $\sigma = E_N^0 / \omega^2(x^0)$ , which finally yields the differential equation

$$\frac{\dot{E}_N^0}{E_N^0} = \frac{1}{2} \frac{d\omega^2(x^0)/dt}{\omega^2(x^0)} = \frac{d\omega(x^0)/dt}{\omega(x^0)}.$$

Thus, there is a constant  $\Theta$ , such that

$$E_N^0 = \Theta \omega(x^0),$$

finishing the proof if we note the relation between  $E_N^0$  and  $\sigma$  once again.  $\square$

*Remark.* In retrospective, the proof of this theorem was based on Eq. (13) of the normal oscillation *and* on the explicit limit equation of Theorem 2.1. Van Kampen [35, p. 103f] argues by formal use of the WKB method, that the invariance of  $E_N^0/\omega(x^0)$  follows at once from Eq. (13). However, the following example shows that his argument is *not* correct since an *arbitrary*  $O(1)$ -term can introduce resonances which precludes the adiabatic invariance: Suppose we have

$$\ddot{x}_N^\epsilon + \epsilon^{-2}\omega^2 x_N^\epsilon = \cos \frac{\omega t}{\epsilon} = O(1)$$

with a *constant* frequency  $\omega$ . For the initial data  $x_N^\epsilon(0) = \dot{x}_N^\epsilon(0) = 0$  we get the solution

$$x_N^\epsilon(t) = \frac{\epsilon t}{2\omega} \sin \frac{\omega t}{\epsilon}.$$

This gives a limit normal energy

$$E_N^\epsilon(t) = \frac{1}{2}|\dot{x}_N^\epsilon(t)|^2 + \frac{1}{2}\epsilon^{-2}\omega^2|x_N^\epsilon(t)|^2 = \frac{1}{8}t^2 + O(\epsilon) \rightarrow E_N^0(t) = \frac{1}{8}t^2,$$

which is *not* of the form  $\Theta\omega$ .

The limit function  $x^0$  can now be described in a completely intrinsic way.

*Theorem 4.3.* Suppose the initial values  $x_0^\epsilon, \dot{x}_0^\epsilon$  are given, such that the limits

$$x_0^0 = \lim_{\epsilon \rightarrow 0} x_0^\epsilon \in \mathcal{M}, \quad \dot{x}_0^0 = \lim_{\epsilon \rightarrow 0} \dot{x}_0^\epsilon \in T_{x_0^0}\mathcal{M}$$

and

$$\Theta = \lim_{\epsilon \rightarrow 0} \frac{(\dot{x}_{0N}^\epsilon)^2 + \epsilon^{-2}\omega^2(x_0^0)(x_{0N}^\epsilon)^2}{2\omega(x_0^0)}$$

exist. Then, the sequence  $x^\epsilon$   $\mathcal{M}$ -converges to the unique solution  $x^0 \in C^2$  of the constrained system

$$\ddot{x}^0 + \text{grad } W(x^0) \in N_{x^0}\mathcal{M}$$

with the corrected potential

$$W(x) = V(x) + V_{\text{add}}(x), \quad V_{\text{add}}(x) = \Theta\omega(x),$$

which is defined for all  $x \in \mathcal{M}$ .

*Proof.* In view of the discussion at the beginning of Section 3, we only have to show that

$$\langle \frac{1}{2}D^2G(x^0) : \Sigma, X \rangle = \langle \text{grad } V_{\text{add}}|_{x=x^0}, X \rangle$$

for each vector field  $X \in T_{x^0}\mathcal{M}$ . In fact, using relation (11), i.e.,

$$\langle \frac{1}{2}D^2G(x^0) : \Sigma, X \rangle = \frac{1}{2}(D_X D_N^2 U(x^0)) \cdot \sigma,$$

we get by Theorem 4.2

$$\langle \frac{1}{2} D^2 G(x^0) : \Sigma, X \rangle = \frac{1}{2} \langle \text{grad } \omega^2|_{x=x^0}, X \rangle \cdot \frac{\Theta}{\omega(x^0)} = \Theta \langle \text{grad } \omega|_{x=x^0}, X \rangle.$$

Again, the uniqueness of the limit  $x^0$  allows us to disregard the extraction of subsequences.  $\square$

*Remark.* This theorem was first proved by Rubin and Ungar [27, p. 82f], however, the result is somewhat hidden in their paper. Independently, it can be found by means of an example in the work of Koppe and Jensen [19, Eq. (7)].

The additional potential does not have any influence if and only if either  $\Theta = 0$ , i.e., the normal energy vanishes initially in the limit, or  $\omega(x)$  is a constant on the manifold  $\mathcal{M}$ . This shows, that the two cases discussed in Section 3 essentially exhaust all possibilities for the realization of holonomic constrained motions under the potential  $V$ .

## 5. A review of the general case

The case of codimension  $r > 1$  is considerably more difficult and has been carefully analyzed by Takens [33]. We restrict ourselves to a short review of his results. Takens calls the Hessian matrix  $D_N^2 U$  of the strong potential in the normal directions of  $\mathcal{M}$  *smoothly diagonalizable*, if there is a smooth field  $(e_N^1, \dots, e_N^r)$  of orthonormal bases of  $N\mathcal{M}$ , which are eigenvectors of  $D_N^2 U$ , i.e.,

$$D_N^2 U(x) : (e_N^i(x) \otimes e_N^j(x)) = \omega_i^2(x) \cdot \delta_{ij} \quad \forall x \in \mathcal{M}.$$

Here, the eigenfrequencies  $\omega_i$  shall depend *smoothly* on  $x \in \mathcal{M}$ . Takens ([33, Theorem 1]) proves that Theorem 4.2 extends to each normal component, if one can exclude certain *resonances*,<sup>2</sup> i.e., if for  $x \in \mathcal{M}$  we always have

$$\omega_i(x) \neq \omega_j(x), \quad 1 \leq i, j \leq r, \quad i \neq j,$$

and

$$\omega_i(x) \neq \omega_j(x) + \omega_k(x), \quad 1 \leq i, j, k \leq r.$$

Using this result, we can extend Theorem 4.3 in a straightforward fashion using the same proof. However, in general *not* smoothly diagonalizable case, there can be situations, where the limit  $x^0$  *cannot* be described intrinsically by a deterministic initial value problem. In fact, Takens [33, Theorem 3] constructs an example with  $d = 4, r = 2$ , where a one-parameter family of initial data  $x^\epsilon(0; \mu), \dot{x}^\epsilon(0; \mu)$ , depending on  $\mu \in [0, 1]$ , yields a one-parameter family of limit solutions  $x^0(t; \mu)$  having the following property: There is a time  $t_* > 0$  such that

$$x^0(t; \mu) = x^0(t)$$

does not depend on the parameter  $\mu$  for  $0 \leq t \leq t_*$ . However, for fixed  $t > t_*$  the values of  $x^0(t; \mu), \mu \in [0, 1]$ , constitute a *continuum*, i.e., for  $t > t_*$  the family forms a *funnel*. This resembles the properties of nonuniquely solvable initial value problems, cf. [25]. Thus, for a fixed parameter  $\mu$  we cannot describe the limit  $x^0$  by a uniquely solvable initial value problem. Koiller [18] coined the notion “Takens-chaos” for this effect.

*Remark.* Keller and Rubinstein [16] deal with the corresponding general case for the semilinear wave equation  $v_{tt} = \Delta v - \epsilon^{-2} \text{grad } U(v)$ . Using an ingenious multiple scale asymptotics they arrive at the same limit equation as

<sup>2</sup> *Note added in proof:* By extending the methods developed in this paper the first author was recently able to show that these resonances have only to be excluded along the limit  $x^0$  *almost everywhere*, thus allowing for simple eigenvalue crossings.

Takens. However, their analysis being only formal it cannot predict difficulties at resonances or even the appearance of Takens chaos.

## 6. Two examples

We will discuss two examples for the general codimension  $r = 1$  case, which illustrate the occurrence of the additional potential  $V_{\text{add}}$ . A further nontrivial example in the context of molecular dynamics including numerical simulations of the Butan molecule can be found in [31].

### 6.1. Example 1. Illustrative, but artificial

We take the Hamiltonian

$$H^\epsilon = \frac{1}{2}\dot{x}^2 + \frac{1}{2}\dot{y}^2 + V(x, y) + \epsilon^{-2}U(x, y), \quad U(x, y) = \frac{1}{2}\omega^2(x)y^2,$$

describing a motion in  $\mathbb{R}^2$  with the corresponding constraint manifold  $\mathcal{M} = \{y = 0\}$  of dimension  $m = 1$  and codimension  $r = 1$ . The equation of motion reads as

$$\ddot{x}^\epsilon = -V_x(x^\epsilon, y^\epsilon) - \epsilon^{-2}\omega(x^\epsilon)\omega'(x^\epsilon)(y^\epsilon)^2, \quad \ddot{y}^\epsilon = -V_y(x^\epsilon, y^\epsilon) - \epsilon^{-2}\omega^2(x^\epsilon)y^\epsilon.$$

Using the initial values

$$x^\epsilon(0) = x_0, \quad \dot{x}^\epsilon(0) = v_0, \quad y^\epsilon(0) = 0, \quad \dot{y}^\epsilon(0) = w_0,$$

we get as an immediate consequence of Theorem 4.3 that

$$x^\epsilon \rightarrow x^0 \quad \text{in } C^1, \quad y^\epsilon \rightarrow 0 \quad \text{in } C^0,$$

where  $x^0$  is the solution of the initial value problem

$$\ddot{x}^0 = -V_x(x^0, 0) - \Theta \omega'(x^0), \quad x^0(0) = x_0, \quad \dot{x}^0(0) = v_0, \quad \Theta = w_0^2/2\omega(x_0).$$

This result enables us to discuss a further important point. If we consider the special situation that initially

$$-V_x(x_0, 0) = \frac{w_0^2}{2\omega(x_0)}\omega'(x_0) \neq 0 \quad \text{and} \quad v_0 = 0,$$

we get the *stationary* solution  $x^0 \equiv x_0$ . On the one hand, this proves the *necessity* of the additional potential  $V_{\text{add}} = \Theta\omega$ . On the other hand, Theorem 4.2 gives us in this case a normal energy,

$$E_N^0 = \frac{1}{2}w_0^2,$$

which is *constant* in time. Thus, a constant limit normal energy is only necessary for the vanishing of the potential correction  $V_{\text{add}}$ , but *not sufficient*, cf. the discussion in Section 3.

*Remark.* This example is discussed at length by Koppe and Jensen [19], who in fact prove Theorem 4.3 only for this example. Gallavotti [13, p. 172ff] discusses the special case  $\omega^2(x) = 1 + x^2$  “only in a heuristic, nonrigorous way”, as he writes. Instead of

$$V_{\text{add}}(x) = \frac{w_0^2}{2} \sqrt{\frac{1+x^2}{1+x_0^2}}$$

he arrives at the *wrong* potential correction

$$V_{\text{wrong}}(x) = \frac{1}{4} w_0^2 \log(1 + x^2).$$

The reason for this flaw is that he first *correctly* derives for  $x \approx x_0$

$$\text{grad } V_{\text{add}}(x) \approx \frac{w_0^2}{2} \frac{x}{1 + x_0^2},$$

but in turn he argues that  $\text{grad } V_{\text{add}}$  is therefore given by

$$\frac{w_0^2}{2} \frac{x}{1 + x^2},$$

which in fact yields  $V_{\text{wrong}}$ . Interestingly enough, the potential correction suggested by Reich [26], the so-called Fixman potential, also turns out to be the *same*  $V_{\text{wrong}}$ .

## 6.2. Example II. The magnetic mirror

We consider the motion of a charged particle in a nonuniform axially symmetric magnetic  $B$ -field, whose field lines lie in planes passing through the symmetry axis. Thus, in cylindrical coordinates  $r, z, \phi$  the  $B$ -field does not depend on the angle  $\phi$  and its  $\phi$ -component vanishes. Hence, there is a vector potential  $\mathcal{A}$  with components  $\mathcal{A} = (0, 0, A(r, z))$ , such that  $B = \text{curl } \mathcal{A}$ , i.e.,

$$B = \left( -\frac{\partial A}{\partial z}, \frac{\partial A}{\partial r} + \frac{A}{r}, 0 \right). \quad (14)$$

The motion takes place according to the Lagrangian

$$\mathcal{L} = \frac{1}{2} m |\dot{x}|^2 + e \langle \dot{x}, \mathcal{A} \rangle = \frac{1}{2} m (\dot{r}^2 + \dot{z}^2 + r^2 \dot{\phi}^2) + er \dot{\phi} A.$$

Since  $\mathcal{L}$  does not depend on  $\phi$  we obtain conservation of the angular momentum,

$$\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = J = \text{const.}, \quad \text{i.e.,} \quad mr^2 \dot{\phi} + erA = J.$$

We eliminate the cyclic variable  $\phi$  by the classical method of Routh reducing the Lagrangian in  $(r, z)$ -coordinates to (cf. [5, Chap. 3, Section 2.1])

$$\mathcal{L}_{\text{red}} = \mathcal{L} - J \dot{\phi} |_{mr^2 \dot{\phi} + erA = J} = \frac{m}{2} (\dot{r}^2 + \dot{z}^2) - \frac{e^2}{2m} \left( A - \frac{J}{er} \right)^2.$$

Equivalently, the corresponding motion fits into our framework with the Hamiltonian

$$H = \frac{1}{2} (\dot{r}^2 + \dot{z}^2) + \epsilon^{-2} U(r, z), \quad U(r, z) = \frac{1}{2} \left( A - \frac{J}{er} \right)^2, \quad \epsilon = \frac{m}{e}.$$

The theory developed in this paper shows that for a large specific charge  $\epsilon^{-1} = e/m$  the projection of the motion to the  $(r, z)$ -plane oscillates very rapidly in a small neighborhood of the line

$$\mathcal{M}_{\text{red}} = \{(r, z): A(r, z) = J/er\}.$$

In addition, we are able to describe the secular oscillations of the angular variable  $\phi$ . Using the notation and results of Section 4 together with the limit relation (8) we obtain

$$\dot{\phi}^\epsilon = \frac{\epsilon^{-1}}{r^\epsilon} \left( \frac{J}{er^\epsilon} - A(r^\epsilon, z^\epsilon) \right) = \pm \sqrt{2\epsilon^{-2}U/r^\epsilon} = \frac{\omega(x_{\mathcal{M}}^\epsilon)}{r^\epsilon} \eta^\epsilon + O(\epsilon^{1/2}) \xrightarrow{*} 0,$$

implying by the Arzelà–Ascoli theorem the uniform convergence  $\phi^\epsilon \rightarrow \phi_0$  to a fixed initial value  $\phi_0 = \phi^\epsilon(0)$ . Thus, the actual motion in *space* is a small amplitude *gyration* around the line

$$\mathcal{M} = \{(r, z, \phi): A(r, z) = J/er, \phi = \phi_0\},$$

the so-called *guiding center* of the motion, which in fact is a field line of the magnetic field (14). The frequency of gyration is given by  $\epsilon^{-1}\omega$  with

$$\omega^2 = D^2U : (e_N \otimes e_N)|_{\mathcal{M}_{\text{red}}} = (D(A - J/er) \cdot e_N)^2 = |B|^2,$$

since  $D(A - J/er) \cdot e_N = \pm|B|$  on  $\mathcal{M}_{\text{red}}$  by (14). Thus, just as in the case of uniform magnetic fields the particle gyrates with *Larmor frequency*  $e|B|/m$ .

Theorem 4.3 shows, that in the limit  $\epsilon \rightarrow 0$  the average tangential motion along the guiding center  $\mathcal{M}$  is governed by the potential

$$W = \Theta|B|,$$

where the adiabatic invariant  $\Theta$  is the *magnetic moment* of the particle motion. The equation of motion now reads

$$\ddot{s} = -\Theta \frac{\partial}{\partial s} |B|, \quad (15)$$

where  $s$  denotes arc length on the line  $\mathcal{M}$ . As we see, the appearance of the additional potential  $W$  introduces the *only* force term for the limit motion. This force term is of utmost importance in engineering and natural sciences: Charged particles are moderated by an increasingly strong magnetic field – and that the more, the bigger the initial normal velocity was. This is the working principle of magnetic traps and magnetic mirrors in plasma physics, as well as of the Van Allen radiation belt of the earth with all its implications for northern lights and astronautics.

*Remark.* The first derivation of Eq. (15) by physical reasoning was given by the Swedish Nobel prize winner Alfvén [2, Chap. 2.3], see also [24,32].

The first mathematical discussion of the limit  $e/m \rightarrow \infty$  was given by Rubin and Ungar [27], who also discuss a nice mechanical analog of the magnetic mirror. However, they only consider the *reduced* motion in the  $(r, z)$ -plane. The adiabatic invariance of the magnetic moment was also shown by Arnold in his seminal paper [3].

## Appendix A

Here, we derive the equations of motion in the  $(x_{\mathcal{M}}, x_N)$ -coordinates of Section 1. The Lagrangian is given in these coordinates by

$$\mathcal{L} = \frac{1}{2}g(x_{\mathcal{M}}, x_N) : (\dot{x}_{\mathcal{M}} \otimes \dot{x}_{\mathcal{M}}) + h(x_{\mathcal{M}}, x_N) : (\dot{x}_{\mathcal{M}} \otimes \dot{x}_N) + \frac{1}{2}|\dot{x}_N|^2 - V(x_{\mathcal{M}}, x_N) - \epsilon^{-2}U(x_{\mathcal{M}}, x_N),$$

where  $g$  denotes the metric tensor on  $\mathcal{M}$  and we have  $h(x, 0) = 0$  for  $x \in \mathcal{M}$  because of the orthogonality of the coordinate splitting. The equations of motion are given by the Euler–Lagrange equations:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_{\mathcal{M}}} = \frac{\partial \mathcal{L}}{\partial x_{\mathcal{M}}}, \quad \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_N} = \frac{\partial \mathcal{L}}{\partial x_N}.$$

We compute the derivatives and simplify them using the asymptotic results  $\dot{x}^\epsilon = O(1)$  and  $x_N^\epsilon = O(\epsilon)$ :

$$\begin{aligned} \left. \frac{\partial \mathcal{L}}{\partial x_M} \right|_{x=x^\epsilon} &= \frac{1}{2} D_M g : (\dot{x}_M^\epsilon \otimes \dot{x}_M^\epsilon) + D_M h : (\dot{x}_M^\epsilon \otimes \dot{x}_N^\epsilon) - D_M V - \epsilon^{-2} D_M U \\ &= -\epsilon^{-2} D_M U + O(1), \end{aligned}$$

$$\begin{aligned} \left. \frac{\partial \mathcal{L}}{\partial x_N} \right|_{x=x^\epsilon} &= \frac{1}{2} D_N g : (\dot{x}_M^\epsilon \otimes \dot{x}_M^\epsilon) + D_N h : (\dot{x}_M^\epsilon \otimes \dot{x}_N^\epsilon) - D_N V - \epsilon^{-2} D_N U \\ &= -\epsilon^{-2} D_N U + O(1), \end{aligned}$$

and

$$\left. \frac{\partial \mathcal{L}}{\partial \dot{x}_M} \right|_{x=x^\epsilon} = g \cdot \dot{x}_M^\epsilon + h \cdot \dot{x}_N^\epsilon, \quad \left. \frac{\partial \mathcal{L}}{\partial \dot{x}_N} \right|_{x=x^\epsilon} = h^T \cdot \dot{x}_M^\epsilon + \dot{x}_N^\epsilon,$$

thus,

$$\frac{d}{dt} \left. \frac{\partial \mathcal{L}}{\partial \dot{x}_M} \right|_{x=x^\epsilon} = g \cdot \ddot{x}_M^\epsilon + h \cdot \ddot{x}_N^\epsilon + O(1), \quad \frac{d}{dt} \left. \frac{\partial \mathcal{L}}{\partial \dot{x}_N} \right|_{x=x^\epsilon} = h^T \cdot \ddot{x}_M^\epsilon + \ddot{x}_N^\epsilon + O(1).$$

Hence, the Euler–Lagrange equations take the form (2), where the mass matrix is given by

$$M(x^\epsilon) = \begin{pmatrix} g & h \\ h^T & I \end{pmatrix} = \begin{pmatrix} O(1) & O(\epsilon) \\ O(\epsilon) & I \end{pmatrix},$$

because of  $h(x_M^\epsilon, 0) = 0$  and therefore  $h(x^\epsilon) = h(x_M^\epsilon, x_N^\epsilon) = O(\epsilon)$ . Now,

$$M \cdot \begin{pmatrix} I & -h \\ -h^T & g \end{pmatrix} = \begin{pmatrix} g - hh^T & hg - gh \\ 0 & g - h^T h \end{pmatrix} = \begin{pmatrix} g & 0 \\ 0 & g \end{pmatrix} + O(\epsilon),$$

which gives

$$M^{-1} = \begin{pmatrix} g^{-1} & -hg^{-1} \\ -h^T g^{-1} & I \end{pmatrix} + O(\epsilon) = \begin{pmatrix} O(1) & O(\epsilon) \\ O(\epsilon) & I + O(\epsilon) \end{pmatrix}.$$

This proves the equations in (3).

## Acknowledgements

It is our pleasure to thank Peter Deuffhard for his steady support. His interest in smoothed molecular dynamics initiated our research concerning the subject of this paper. We thank Andrew Majda for pointing out the relevance of Case II for geophysical flows with gravity waves present initially [12]. We are indebted to an anonymous referee for improving the presentation of the second example.

## References

- [1] R. Abraham and J.E. Marsden, *Foundations of Mechanics*, 2nd Ed. (Addison-Wesley, Redwood City, 1985).
- [2] H. Alfvén, *Cosmical Electrodynamics* (Clarendon Press, Oxford, 1950).

- [3] V.I. Arnold, Small denominators and problems of stability of motion in classical and celestial mechanics, *Russ. Math. Surv.* 18 (1963) 85–192.
- [4] V.I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer, Berlin, 1978).
- [5] V.I. Arnold, V.V. Kozlov and A.I. Neishtadt, *Mathematical Aspects of Classical and Celestial Mechanics*, in: *Dynamical Systems III*, ed. V.I. Arnold, 2nd Ed. (Springer, Berlin, 1993).
- [6] G. Benettin, L. Galgani and A. Giorgilli, Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. I, *Commun. Math. Phys.* 113 (1987) 87–103.
- [7] G. Benettin, L. Galgani and A. Giorgilli, Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. II, *Commun. Math. Phys.* 121 (1989) 557–601.
- [8] J.Y. Chemin, A propos d'un problème de pénalisation de type anti symétrique, *C.R.A.S. Paris Ser. I* 321 (1995) 861–864.
- [9] P.G. Ciarlet, *Mathematical Elasticity I* (North-Holland, Amsterdam, 1988).
- [10] B. Dacorogna, *Weak Continuity and Weak Lower Semicontinuity of Non-Linear Functionals* (Springer, Berlin, 1982).
- [11] D.G. Ebin, The motion of slightly compressible fluids viewed as a motion with strong constraining force, *Ann. Math.* 105 (1977) 141–200.
- [12] P.F. Embid and A.J. Majda, Averaging over fast gravity waves for geophysical flows with arbitrary potential vorticity, *Comm. Partial Diff. Equat.* 21 (1996) 619–658.
- [13] G. Gallavotti, *The Elements of Mechanics* (Springer, Berlin, 1983).
- [14] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Cambridge, 1953).
- [15] L. Hörmander, *The Analysis of Linear Partial Differential Operators I. Distribution Theory and Fourier Analysis*, 2nd Ed. (Springer, Berlin, 1990).
- [16] J.B. Keller and J. Rubinstein, Nonlinear wave motion in a strong potential, *Wave Motion* 13 (1991) 291–302.
- [17] S. Klainerman and A.J. Majda, Singular limits of quasilinear hyperbolic systems with large parameters and the incompressible limit of compressible fluids, *Commun. Pure Appl. Math.* 43 (1981) 481–524.
- [18] J. Koiller, A note on classical motions under strong constraints, *J. Phys. A* 23 (1990) L521–L527.
- [19] H. Koppe and H. Jensen, Das Prinzip von d'Alembert in der Klassischen Mechanik und in der Quantenmechanik, *Sitz.-Ber. Heidelb. Akad. Wiss. Math.-Naturwiss. Kl. 5. Abh.* (1971).
- [20] H.-O. Kreiss, Problems with different time scales, in: *Acta Numerica 1992*, ed. A. Iserles (Cambridge University Press, Cambridge, 1992).
- [21] C. Lubich, Integration of stiff mechanical systems by Runge–Kutta methods, *Z. Angew. Math. Phys.* 44 (1993) 1022–1053.
- [22] J.E. Marsden and T.S. Ratiu, *Introduction to Mechanics and Symmetry* (Springer, Berlin, 1994).
- [23] F. Murat and L. Tartar, *H-convergence*, *Topics in the Mathematical Modeling of Composite Materials*, ed. R.V. Kohn (Birkhäuser, Basel, 1994).
- [24] T.G. Northrop, *The Adiabatic Motion of Charged Particles* (Interscience, New York, 1963).
- [25] C. Pugh, Funnel sections, *J. Diff. Eqs.* 19 (1975) 270–295.
- [26] S. Reich, Smoothed dynamics of highly oscillatory Hamiltonian systems, *Physica D* 89 (1995) 28–42.
- [27] H. Rubin and P. Ungar, Motion under a strong constraining force, *Commun. Pure Appl. Math.* 10 (1957) 65–87.
- [28] W. Rudin, *Real and Complex Analysis* (McGraw-Hill, New York, 1987).
- [29] H.-J. Schmidt, Models for constrained motion and d'Alembert's principle, *Z. Angew. Math. Mech.* 73 (1993) 155–163.
- [30] S. Schochet, Fast singular limits of hyperbolic partial differential equations, *J. Diff. Eqs.* 114 (1994) 476–512.
- [31] C. Schütte and F. Bornemann, Homogenization approach to smoothed molecular dynamics, in: *World Congress of Nonlinear Analysts '96: Proc. Second Congress (Athens, Greece)*, ed. V. Lakshmikantham (Walter de Gruyter, Berlin, to appear).
- [32] L. Spitzer Jr., *Physics of Fully Ionized Gases* (Interscience, New York, 1956).
- [33] F. Takens, Motion under the influence of a strong constraining force, in: *Global Theory of Dynamical Systems*, Evanston 1979, eds. Z. Nitecki and C. Robinson (Springer, Berlin, 1980).
- [34] M. Toda, R. Kubo and N. Saitô, *Statistical Physics I. Equilibrium Statistical Mechanics*, 2nd Ed. (Springer, Berlin, 1988).
- [35] N. G. van Kampen, Elimination of fast variables, *Phys. Rep.* 124 (1985) 69–160.