Konrad-Zuse-Zentrum für Informationstechnik Berlin
Heilbronner Str. 10, D-10711 Berlin - Wilmersdorf

Folkmar A. Bornemann

Christof Schütte

Homogenization of Highly Oscillatory Hamiltonian Systems

Preprint SC 95–39 (December 1995)

Homogenization of Highly Oscillatory Hamiltonian Systems

Folkmar A. Bornemann and Christof Schütte

Konrad–Zuse–Zentrum Berlin
Heilbronner Str. 10, D-10711 Berlin, Germany
bornemann@zib-berlin.de

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 ϵ of this small time scale tends to zero but the velocity components remain oscillating with an amplitude variation of order $\mathcal{O}(1)$. 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, in plasma physics as the problem of guiding center motion, 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 the 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

AMS classification: 34C15,34C29,34C30,70F20,78A35

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 ϵ 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 $\epsilon \to 0$.

II Only the position is converging pointwise to a limit function as $\epsilon \to 0$. The velocity remains oscillating with an amplitude variation of order $\mathcal{O}(1)$.

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 $\epsilon \to 0$. Thus, case III is ruled out for *bounded energies*. Case I can be handled by standard averaging techniques of perturbation theory, cf. e.g. [16]. 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 [18].

A discussion of this particular homogenization problem is somewhat scattered in the literature. However, it appears at the heart of three 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 extend in the monographs [3][4][9], by means of examples in [15][28]. A mathematically exhaustive investigation of this question is given in [26], which is heavily based on the important early results of [22].

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 the distinction between case I and II. The reader may find quite controversial positions in the literature, cf. [15, p. 8] and [28, p. 104].

B *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 $\mathcal{O}(1)$. Quite early, the results for problems of type A, although unexpected and counterintuitive in that context, have successfully been discovered in the physical literature [2][25][19] for a description of fast Larmor gyration in nonuniform magnetic fields. These results play a key role for the explanation of magnetic traps and magnetic mirrors in plasma physics. They in fact motivated the important mathematical research of [22].

C *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 of *correcting* the weaker potentials as was first noted in [21], where such a correction was suggested on the base of (questionable) additional physical assumptions. However, the right correction can be established on strictly mathematical grounds using the results of [22][26] — at least for certain subsystems as was argued in [7], where the reader will find a detailed discussion of the field.

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(-star) convergence in L^∞ and in the space of distributions $\mathscr{D}'$ 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-seventieth, cf. [18] 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 for problems of type A, B and C. To the best of our knowledge, the collected references are quite complete.

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 in 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 Hamitonian 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 [26], 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 B.

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.,

$$U|_{\mathcal{M}} = 0, \qquad U(x) > 0 \qquad \forall x \in \mathbb{R}^d \setminus \mathcal{M}.$$

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

$$\xi^T D^2U(x)\xi \geq \alpha^2|\xi|^2 \qquad \forall \xi \in N_x\mathcal{M},$$

where $N\mathcal{M}$ denotes the normal bundle of $\mathcal{M}$.

We will denote the potential forces by

$$F(x) = \operatorname{grad} V(x), \qquad G(x) = \operatorname{grad} U(x).$$

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

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

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

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

1 $\mathcal{M}$-Convergence

We start our investigation by a careful study of the convergence properties of the sequence x^ϵ for increasingly strong potential, i.e., $\epsilon \to 0$. We will prove an extraction principle based on energy methods.

Lemma 1.1 *Let a sequence $\epsilon \to 0$ be given, for which the initial position x_0^ϵ as well as the initial energy H^ϵ are bounded. Then, for each choice $T > 0$, there is a subsequence ϵ', such that the unique solutions $x^{\epsilon'} \in C^\infty([0,T],\mathbb{R}^d)$ exist and the following limits hold*

$$x^{\epsilon'} \to x^0 \quad in \quad C([0,T],\mathbb{R}^d), \qquad \dot{x}^{\epsilon'} \overset{*}{\rightharpoonup} \dot{x}^0 \quad in \quad L^\infty([0,T],\mathbb{R}^d), \tag{2}$$

for a Lipschitz continuous function $x^0 \in C^{0,1}([0,T],\mathbb{R}^d)$. This limit function gives a path in $\mathcal{M}$,

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

and we furthermore get uniformly in $[0,T]$

$$\operatorname{dist}(x^{\epsilon'},\mathcal{M}) = \mathcal{O}(\epsilon'). \tag{3}$$

On the other hand, if there is a sequence ϵ', such that (2) and (3) hold, the boundedness of $x_0^{\epsilon'}$ and $H^{\epsilon'}$ follows.

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 $H^\epsilon \le H_*$ and $|x_0^\epsilon| \le K$ for all ϵ implies

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

and therefore by integration

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

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^ϵ and $\dot{x}^\epsilon$ are bounded sequences in $C([0,T],\mathbb{R}^d)$. By the theorems of Arzelà-Ascoli and of Alaoglu (cf. [23, Thm. 11.28/11.29], remember that L^∞ is the dual space of L^1) we now conclude, that there is a subsequence ϵ' 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 the limit relations in (2) hold. Since in consequence $\ddot{x}^{\epsilon'} \overset{\mathscr{D}'}{\rightharpoonup} \ddot{x}^0$ in the sense of distributions and therefore

$$\epsilon'^2 \ddot{x}^{\epsilon'} \overset{\mathscr{D}'}{\rightharpoonup} 0,$$

we get by taking limits in $\mathscr{D}'$ for equation (1) that $G(x^0) = 0$, i.e., x^0 constitutes a path in $\mathcal{M}$. As a consequence we get

$$\operatorname{dist}(x^{\epsilon'}, \mathcal{M}) \to 0 \tag{4}$$

uniformly in $[0, T]$. Thus, for sufficiently small ϵ', the orthogonal projection $x_{\mathcal{M}}^{\epsilon'} \in \mathcal{M}$ of $x^{\epsilon'}$ is a uniquely defined continuous function in t. Taylor expansion of U shows

$$\begin{aligned} \epsilon'^2(H_* - V_*) &\geq U(x^{\epsilon'}) \\ &= \frac{1}{2}(x^{\epsilon'} - x_{\mathcal{M}}^{\epsilon'})^T D^2U(x_{\mathcal{M}}^{\epsilon'})(x^{\epsilon'} - x_{\mathcal{M}}^{\epsilon'}) + \mathcal{O}(\operatorname{dist}(x^{\epsilon'}, \mathcal{M})^3). \end{aligned}$$

Using assumption A3 and the fact that $(x^{\epsilon'} - x_{\mathcal{M}}^{\epsilon'}) \in N_{x_{\mathcal{M}}^{\epsilon'}}\mathcal{M}$, we get the estimate

$$\alpha^2 \operatorname{dist}(x^{\epsilon'}, \mathcal{M})^2 \leq K_1(\epsilon'^2 + \operatorname{dist}(x^{\epsilon'}, \mathcal{M})^3),$$

where K_1 denotes some constant. With $\rho_{\epsilon'} = \operatorname{dist}(x^{\epsilon'}, \mathcal{M})/\epsilon'$ we thus have

$$\rho_{\epsilon'}^2 \leq \frac{K_1}{\alpha^2}\left(1 + \operatorname{dist}(x^{\epsilon'}, \mathcal{M})\rho_{\epsilon'}^2\right).$$

Because of (4), we have for sufficiently small ϵ'

$$\frac{K_1}{\alpha^2} \operatorname{dist}(x^{\epsilon'}, \mathcal{M}) \leq \frac{1}{2},$$

which implies

$$\rho_{\epsilon'} \leq \frac{\sqrt{2K_1}}{\alpha},$$

i.e., assertion (3).

The proof of the converse result is straightforward. □

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

If we do not bound the initial energy H^ϵ, we cannot expect strong convergence of x^ϵ 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

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

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

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

Here, we get merely weak convergence of x^ϵ in L^∞, namely

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

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

In a small neighborhood of a compact path in $\mathcal{M}$ it is possible to introduce uniquely the following decomposition of a point $x \in \mathbb{R}^d$

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

We will 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

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

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

These coordinates can be applied to a sequence x^ϵ, which fulfills the assertions of Lemma 1.1. Using Lagrangian formalism it is straightforward, but tedious, to establish the equations of motion in these new coordinates. The result can be found in the Appendix of this paper. Inserting the details of the convergence results of Lemma 1.1 these equations take the form

$$M(x^\epsilon) \begin{pmatrix} \ddot{x}_{\mathcal{M}}^\epsilon \\ \ddot{x}_N^\epsilon \end{pmatrix} = \begin{pmatrix} -\epsilon^{-2} D_{\mathcal{M}} U(x^\epsilon) + \mathcal{O}(1) \\ -\epsilon^{-2} D_N U(x^\epsilon) + \mathcal{O}(1) \end{pmatrix}, \tag{5}$$

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

$$M(x^\epsilon) = \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(\epsilon) \\ \mathcal{O}(\epsilon) & I \end{pmatrix}, \qquad M(x^\epsilon)^{-1} = \begin{pmatrix} \mathcal{O}(1) & \mathcal{O}(\epsilon) \\ \mathcal{O}(\epsilon) & I + \mathcal{O}(\epsilon) \end{pmatrix}. \tag{6}$$

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

Lemma 1.2 *Under the assumptions of* Lemma 1.1, *there is a subsequence* ϵ', *such that we get in addition to the assertions of* Lemma 1.1

$$x_{\mathcal{M}}^{\epsilon'} \to x^0 \quad in\ C^1, \qquad \ddot{x}_{\mathcal{M}}^{\epsilon'} \stackrel{*}{\rightharpoonup} \ddot{x}^0 \quad in\ L^\infty.$$

Moreover, x^0 *has a Lipschitz continuous first derivative, i.e.,* $x^0 \in C^{1,1}$.

Proof. Without loss of generality we may assume, that the assertions of Lemma 1.1 already hold for the sequence ϵ. Taylor expansion as in the proof of Lemma 1.1 yields

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

and

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

where we have used $DU|_{\mathcal{M}} = 0$ and the estimate $x_N^\epsilon = \mathcal{O}(\epsilon)$. Thus, the equations of motion (5) and the expression (6) for the inverse of the mass matrix give

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

Applying the theorems of Arzelà-Ascoli and Alaoglu once more, we conclude that there is a subsequence ϵ', such that

$$\dot{x}_{\mathcal{M}}^{\epsilon'} \to \dot{x}^0$$

uniformly in $t \in [0, T]$ and $\ddot{x}_{\mathcal{M}}^{\epsilon'} \stackrel{*}{\rightharpoonup} \ddot{x}^0$. □

For later purposes we consider the quantity

$$\eta^{\epsilon'} = x_N^{\epsilon'}/\epsilon' = \mathcal{O}(1).$$

Using the Alaoglu theorem, we get by a further extraction of a subsequence ϵ'' the convergences $H^{\epsilon''} \to H^0$ as well as

$$\eta^{\epsilon''} \stackrel{*}{\rightharpoonup} \eta \quad \text{in } L^\infty([0,T], \mathbb{R}^d), \qquad \eta^{\epsilon''} \otimes \eta^{\epsilon''} \stackrel{*}{\rightharpoonup} \Sigma \quad \text{in } L^\infty([0,T], \mathbb{R}^{d \times d}).$$

Now, we combine all these convergences in a single notion.

Definition 1.3 Given a sequence $\epsilon \to 0$ and a corresponding sequence of functions $x^\epsilon \in C^2([0,T],\mathbb{R}^d)$. The sequence $\mathcal{M}$-*converges* to a function $x^0 \in C^{1,1}([0,T],\mathcal{M})$ if $H^\epsilon \to H^0$ in $\mathbb{R}$ and

$$x^\epsilon_{\mathcal{M}} \to x^0 \quad \text{in} \;\; C^1([0,T],\mathcal{M}), \qquad \ddot{x}^\epsilon_{\mathcal{M}} \overset{*}{\rightharpoonup} \ddot{x}^0 \quad \text{in} \;\; L^\infty([0,T],\mathbb{R}^d),$$

as well as

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

and furthermore if there exist the limits

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

Here, we assume that the sequence ϵ is restricted to sufficiently small values, for which the coordinate decomposition $x^\epsilon = x^\epsilon_{\mathcal{M}} + x^\epsilon_N$ makes sense.

The results of Lemma 1.1 and Lemma 1.2 are summarized by the following theorem.

Theorem 1.4 *Let a sequence $\epsilon \to 0$ be given, for which the initial position x^ϵ_0 as well as the initial energy H^ϵ are bounded. Then, for each choice $T > 0$, there is a subsequence ϵ', such that $x^{\epsilon'}$ $\mathcal{M}$-converges to some $x^0 \in C^{1,1}$. Conversely, if x^ϵ $\mathcal{M}$-converges, the initial position x^ϵ_0 as well as the initial energy H^ϵ are bounded.*

For later purposes we state the following simple consequence for the initial values $x^0_0 = x^0(0)$ and $\dot{x}^0_0 = \dot{x}^0(0)$.

Corollary 1.5 *Suppose x^ϵ $\mathcal{M}$-converges to x^0. Then we have*

$$x^0_0 = \lim_{\epsilon\to 0} x^\epsilon_0 \;\in \mathcal{M}$$

and

$$\dot{x}^0_0 = \lim_{\epsilon\to 0} \dot{x}^\epsilon_{0\mathcal{M}} \;\in T_{x^0_0}\mathcal{M}.$$

2 Abstract Homogenization

The question arises, whether the limit x^0 of a $\mathcal{M}$-converging sequence x^ϵ is itself a solution of an initial value problem on $\mathcal{M}$. In this section we offer an abstract approach for unfolding the structure of such a limiting equation. We call this process "homogenization", since it eliminates the fast oscillation x^ϵ_N normal to the manifold, cf. the analogous situation in [18].

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

$$\ddot{x}^\epsilon \stackrel{\mathscr{D}'}{\rightharpoonup} \ddot{x}^0$$

in $\mathscr{D}'$, in fact, even in the sense of distributions of *first order*, i.e., in $\mathscr{D}'^1$, cf. [12]. Thus, taking limits in (1), we get

$$\ddot{x}^0 + F(x^0) + \mathscr{D}'^1\text{-}\lim_{\epsilon \to 0} \epsilon^{-2} G(x^\epsilon) = 0. \tag{7}$$

This limit expression can be evaluated.

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

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

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

$$\ddot{x}^0 + F(x^0) + DG(x^0) \cdot \lambda^* + \frac{1}{2} D^2 G(x^0) : \Sigma = 0 \tag{8}$$

almost everywhere. The quantities η^ϵ and Σ are from Definition 1.3.

Proof. Taylor expansion of second order yields

$$\begin{aligned} \epsilon^{-2} G(x^\epsilon) &= \epsilon^{-2} \left(G(x^\epsilon) - G(x^\epsilon_{\mathcal{M}}) \right) \\ &= DG(x^\epsilon_{\mathcal{M}}) \cdot \frac{\eta^\epsilon}{\epsilon} + \\ &\qquad \int_0^1 (1-s) D^2 G(x^\epsilon_{\mathcal{M}} + s x^\epsilon_N) : (\eta^\epsilon \otimes \eta^\epsilon)\, ds. \end{aligned} \tag{9}$$

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

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

we get

$$\int_0^1 (1-s) D^2 G(x^\epsilon_{\mathcal{M}} + s x^\epsilon_N) : (\eta^\epsilon \otimes \eta^\epsilon)\, ds \stackrel{*}{\rightharpoonup} \frac{1}{2} D^2 G(x^0) : \Sigma. \tag{10}$$

Equation (7) shows that $\mathscr{D}'^1\text{-}\lim_{\epsilon \to 0} \epsilon^{-2} G(x^\epsilon)$ exists as a function in L^∞. Thus, the relations (9) and (10) yield the existence of the limit

$$\mathscr{D}'^1\text{-}\lim_{\epsilon \to 0} DG(x^\epsilon_{\mathcal{M}}) \cdot \frac{\eta^\epsilon}{\epsilon} = \hat{\lambda}$$

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

$$DG(x^\epsilon_{\mathcal{M}}) \cdot \frac{\eta^\epsilon}{\epsilon} = \begin{pmatrix} 0 \\ D^2_N U(x^\epsilon_{\mathcal{M}}) \eta^\epsilon/\epsilon \end{pmatrix}.$$

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

$$D^2_N U(x^\epsilon_{\mathcal{M}})^{-1} \;\to\; D^2_N U(x^0)^{-1}$$

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

$$C^1 \times \mathscr{D}'^1 \to \mathscr{D}'^1,$$

we get the existence of the limit

$$\lambda^* = \mathscr{D}'^1\text{-}\lim_{\epsilon\to 0} \frac{\eta^\epsilon}{\epsilon} = \mathscr{D}'^1\text{-}\lim_{\epsilon\to 0} D^2_N U(x^\epsilon_{\mathcal{M}})^{-1} \cdot D^2_N U(x^\epsilon_{\mathcal{M}}) \cdot \frac{\eta^\epsilon}{\epsilon} = D^2_N U(x^0)^{-1}\hat{\lambda}$$

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

$$\mathscr{D}'^1\text{-}\lim_{\epsilon\to 0} DG(x^\epsilon_{\mathcal{M}}) \cdot \frac{\eta^\epsilon}{\epsilon} = DG(x^0) \cdot \lambda^*,$$

which finally gives the desired limit equation. □

Remarks.

1. Note, that the existence of $\mathscr{D}'^1\text{-}\lim_{\epsilon\to 0} \eta^\epsilon/\epsilon$ implies that $\eta^\epsilon \overset{*}{\rightharpoonup} 0$. However, in general we will still have

   $$\eta^\epsilon \otimes \eta^\epsilon \overset{*}{\rightharpoonup} \Sigma \neq 0.$$

   For example, $\eta^\epsilon(t) = \sin(t/\epsilon)$ yields $\eta^\epsilon \overset{*}{\rightharpoonup} 0$ but $\eta^\epsilon \otimes \eta^\epsilon = |\eta^\epsilon|^2 \overset{*}{\rightharpoonup} \Sigma = 1/2$. Thus, the product mapping is not weakly continuous, cf. [8].

2. The first order Taylor expansion

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

   instead of (9) cannot be used to evaluate the $\mathscr{D}'^1$-limit of the expression. The reason is, that certainly

   $$\int_0^1 DG(x^\epsilon_{\mathcal{M}} + s x^\epsilon_N)\, ds \;\to\; DG(x^0)$$

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 $C\times\mathscr{D}'^1$ as the following example shows: Take $\phi^\epsilon(t) = \epsilon\cos(t/\epsilon)$ and $\psi^\epsilon = \epsilon^{-1}\cos(t/\epsilon)$. We have $\phi^\epsilon \to 0$ in C, but not in C^1, and $\mathscr{D}'^1\text{-}\lim_{\epsilon\to 0}\psi^\epsilon = 0$, since $\psi^\epsilon(t) = d\sin(t/\epsilon)/dt$. However, the product converges not to zero, $\phi^\epsilon\psi^\epsilon \stackrel{*}{\rightharpoonup} 1/2$.

3. This abstract theorem shows, that in general there exists an additional force for the limit motion on $\mathcal{M}$, which involves *third* derivatives of the strong potential U.

4. For codimension $r = 1$, the limit equation (8) was already stated by Koppe and Jensen [15, eq. (5)] and by van Kampen [28, eq. (8.33)] using suitable averaging operators to express Σ.

The abstract homogenization process of this section does *not* yield an *intrinsic* description of x^0 on $\mathcal{M}$. This is even not possible, since the "shadow" Σ 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 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 (8) vanishes in the tangential direction, i.e., if

$$\frac{1}{2}D^2G(x^0) : \Sigma \;\in N_{x^0}\mathcal{M}, \tag{11}$$

the limit function obeys

$$\ddot{x}^0 + \operatorname{grad} V(x^0) \in N_{x^0}\mathcal{M}, \tag{12}$$

because $DG(x^0)\cdot\lambda^* \in N_{x^0}\mathcal{M}$ holds in any case. By the d'Alembert-Lagrange principle, the relation (12) describes the motion due to the potential V under the *holonomic constraints*

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

Thus, the limit $\epsilon \to 0$ "realizes" holonomic constraints with potential V, if and only if the condition (11) is fulfilled. Standard textbooks on classical mechanics like [1][3][4][17] prove the existence of a *unique* solution $x^0 \in C^2([0,T],\mathcal{M})$ of (12) for given initial values $\dot{x}^0_0 \in T_{x^0_0}\mathcal{M}$. The arguments used in the proof of Theorem 2.1 show that, given a solution of (12), the equation

$$\ddot{x}^0 + \operatorname{grad} V(x^0) = -DG(x^0) \cdot \lambda$$

can always uniquely be solved for the *Lagrange-parameter* $\lambda(t) \in N_{x^0}\mathcal{M}$. Summarizing our discussion yields the following

Theorem 3.1 *Suppose that condition* (11) *holds for a given* $\mathcal{M}$*-converging sequence* x^ϵ*. Then, there is a unique* $\lambda \in L^\infty([0,T], N_{x^0}\mathcal{M})$*, such that*

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

On the other hand, the initial value problem

$$\ddot{x}^0 + F(x^0) + DG(x^0)\lambda = 0, \qquad G(x^0) = 0, \tag{13}$$

with $x^0(0) = x_0^0 \in \mathcal{M}$*,* $\dot{x}^0(0) = \dot{x}_0^0 \in T_{x_0^0}\mathcal{M}$ *has a unique solution* $x^0 \in C^2([0,T], \mathcal{M})$ *and* $\lambda \in C([0,T], N_{x^0}\mathcal{M})$*.*

Remark. For the purposes of numerical integration, one should represent the constraints manifold $\mathcal{M}$ by r independent conditions, i.e.,

$$\mathcal{M} = \{x : \mathbb{R}^d : \psi(x) = 0\}, \; \psi : \mathbb{R}^d \to \mathbb{R}^r, \text{ with } \operatorname{rank} D\psi(x)|_{x\in\mathcal{M}} = r.$$

Then, the equations of motion are described more conveniently as

$$\ddot{x}^0 + F(x^0) + D\psi^T(x^0)\lambda = 0, \qquad \psi(x^0) = 0,$$

with a Lagrange parameter $\lambda \in \mathbb{R}^r$. This is a so called index 3 differential-algebraic system, cf. [11].

There are essentially just two cases, where one can show, that condition (11) holds.

Case I: Vanishing Normal Energy

Theorem 3.2 *Suppose the initial values satisfy*

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

Then, the sequence x^ϵ $\mathcal{M}$*-converges to the unique solution of equation* (13)*.*

Proof. The assumptions on the initial values make standard perturbation theory applicable. With its help LUBICH [16, Theorem 2.2] proves that $x_N^\epsilon = \mathcal{O}(\epsilon^2)$, which implies the strong convergence

$$\eta^\epsilon \to 0,$$

uniformly in $[0,T]$. In this case we conclude that

$$\eta^\epsilon \otimes \eta^\epsilon \to 0$$

strongly, hence $\Sigma = 0$. Theorem 1.4 and Theorem 2.1 yield the existence of a subsequence $x^{\epsilon'}$, which $\mathcal{M}$-converges to the solution x^0 of equation (13). Since this limit is *unique* by Theorem 3.1, we can disregard the extraction of subsequences and have thus proved the convergence of the original sequence. □

Remark. The first mathematical proof of this theorem was given by RUBIN AND UNGAR [22]. It appears in form of an example in the textbook of ARNOLD [3, Chap. 17A]. For codimension $r = 1$ one can find a discussion in [15] and [28]. It is restated as Theorem 9 in [4, Chap. 1,§6.2].

Case II: Constraining Potentials with Constant Gully Width

Theorem 3.3 *Suppose that initial values x_0^ϵ, $\dot{x}_0^\epsilon$ are given with uniformly bounded energy $H^\epsilon = H(x_0^\epsilon, \dot{x}_0^\epsilon; \epsilon)$, such that*

$$x_0^0 = \lim_{\epsilon \to 0} x_0^\epsilon \ \in \mathcal{M}$$

and

$$\dot{x}_0^0 = \lim_{\epsilon \to 0} \dot{x}_{0\mathcal{M}}^\epsilon \ \in T_{x_0^0}\mathcal{M}.$$

If the constraining potential U satisfies

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

the sequence x^ϵ $\mathcal{M}$-converges to the unique solution of equation (13).

Proof. By Theorem 1.4 there is some subsequence $x^{\epsilon'}$, which $\mathcal{M}$-converges to a solution x^0 of the limit equation (8). Since η^ϵ is normal to the manifold, the tensor Σ 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 $\mathcal{M}$

$$
\begin{aligned}
\langle D^2 G(x^0) : \Sigma,\ X\rangle &= \left(D^2 D_X U|_{x=x^0}\right) : \Sigma \\
&= \left(D_N^2 D_X U|_{x=x^0}\right) : \Sigma_{NN} \qquad (14)\\
&= \left(D_X D_N^2 U|_{x=x^0}\right) : \Sigma_{NN},
\end{aligned}
$$

because the tangential derivates D_X commutes with the normal derivative D_N. Since by assumption $D_X D_N^2 U = 0$ for all $X \in T\mathcal{M}$, the condition (11) is satisfied. Theorem 2.1 and Theorem 3.1 show that x^0 is the *unique* solution of the constrained system (13). This uniqueness of the limit x^0 implies, that already the original sequence x^ϵ $\mathcal{M}$-converges to x^0. □

Remarks.

1. In view of Corollary 1.5 the conditions on the initial data are even necessary for the assertions of the theorem. However, they allow a nonvanishing normal velocity component:

$$\dot{x}_N^\epsilon(0) = \mathcal{O}(1).$$

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

3. This theorem has been stated and proved by Gallavotti [9, Chap. 3,§3.8], who calls it "Arnold's theorem" in view of a remark, which was made by Arnold on p. 91f of his textbook [3].

4. Takens [26] offers a proof under somewhat more restrictive conditions on the potential U, cf. his remark on p. 429.

5. This theorem was extended to the case of a time dependent potential V by Schmidt [24]. Our methods extend without difficulty to that case.

6. For codimension $r = 1$ one can find a discussion in [22],[15], and [28].

7. This theorem shows that in the limit the normal and the tangential part of the energy are completely separated,

$$H^0 = \frac{1}{2}|\dot{x}^0|^2 + V(x^0) + E_N^0,$$

where E_N^0 is a constant. Interestingly enough, this energy separation holds for *exponentially large* times in the case $0 < \epsilon \ll 1$. To be precise, for the energy decomposition

$$H^\epsilon = \frac{1}{2}|\dot{x}^\epsilon_{\mathcal{M}}|^2 + V(x^\epsilon_{\mathcal{M}}) + E^\epsilon_N$$

Benettin, Galgani, and Giorgilli [5][6] have proved the following Nekhoroshev-type of result:

$$|E^\epsilon_N(t) - E^\epsilon_N(0)| < \epsilon \qquad \text{for} \quad 0 \le t \le \exp(b\,\epsilon^{-a}),$$

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 $U(x) = \operatorname{dist}(x, \mathcal{M})^2$ yields $a = 1$ in any dimension. These results should be contrasted with the comparatively rather trivial estimate given by Schmidt [24, Prop. 1].

4 The General Case for Codimension 1

In general, the explicit evaluation of the term

$$\frac{1}{2}D^2G(x^0) : \Sigma \tag{15}$$

demands a careful study of the normal oscillations x^ϵ_N. For reasons, which will become clear in the next section, 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 Σ takes the special form

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

where the *scalar* function σ is given by the limit

$$(x^\epsilon_N/\epsilon)^2 \stackrel{*}{\rightharpoonup} \sigma.$$

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

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

in normal direction. Note that $\omega(x) \geq \alpha > 0$ by assumption A3. For a $\mathcal{M}$-converging sequence x^ϵ, we thus introduce the *normal energy*

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

Since $x_N^\epsilon = \mathcal{O}(\epsilon)$, the total energy splits as

$$H^\epsilon = \frac{1}{2}|\dot{x}_{\mathcal{M}}^\epsilon|^2 + V(x_{\mathcal{M}}^\epsilon) + E_N^\epsilon + \mathcal{O}(\epsilon). \tag{16}$$

Lemma 4.1 *For a given $\mathcal{M}$-converging sequence x^ϵ, the normal oscillation x_N^ϵ satisfies the equation*

$$\ddot{x}_N^\epsilon + \epsilon^{-2}\omega^2(x_{\mathcal{M}}^\epsilon)x_N^\epsilon = \mathcal{O}(1). \tag{17}$$

The normal energy converges uniformly to a continuous function $E_N^0 \in C([0,T],\mathbb{R})$ and the constant total energy H^ϵ converges to a constant H^0 which splits as

$$H^0 = \frac{1}{2}|\dot{x}^0|^2 + V(x^0) + E_N^0.$$

Proof. Using the second order equations (5) of motion in $(x_{\mathcal{M}}, x_N)$-coordinates we get

$$\ddot{x}_N^\epsilon + \epsilon^{-2}D_N U(x^\epsilon) = \mathcal{O}(\epsilon)\ddot{x}_{\mathcal{M}}^\epsilon + \mathcal{O}(1).$$

Hence, the Taylor expansion

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

leads us to the desired type of equation for x_N^ϵ. By the Definition 1.3 of $\mathcal{M}$-convergence, we have $H^\epsilon \to H^0$ as constants and therefore by the splitting (16) the uniform convergence

$$E_N^\epsilon \;\to\; H^0 - \left(\frac{1}{2}|\dot{x}^0|^2 + V(x^0)\right)$$

in $C([0,T],\mathbb{R})$. $\square$

This result throws further light on the appearance of the force term $D^2G(x^0):\Sigma/2$ in the limit equation (8). Suppose this term has no influence on the motion of x^0, i.e., the condition (11) of the previous section is fulfilled.

Then, Theorem 3.1 and the discussion preceding it show, that x^0 describes the constrained motion on $\mathcal{M}$ with the *constant* energy

$$H_{\mathcal{M}} = \frac{1}{2}|\dot{x}^0|^2 + V(x^0). \tag{18}$$

Hence, Lemma 4.1 implies that the limit E_N^0 of the normal energy would be necessarily *constant in time*. However, in general it will be a time *dependent* function, therefore enforcing the nontrivial additional force term (15) on the manifold $\mathcal{M}$. We will see, that this additional force term is conservative, i.e., there is an additional potential yielding the equation of motion for the limit x^0.

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 result of Lemma 4.1. The oscillations take place on a time scale $\mathcal{O}(\epsilon)$ and the modulation by $\omega^2(x_{\mathcal{M}}^\epsilon)$ occurs on a time scale $\mathcal{O}(1)$. Thus, the simplified equation describes a slowly modulated harmonic oscillation and the perturbation theory for integrable Hamiltonian systems is applicable. In fact, one can show that the action variable E/ω of a single-frequency system is an *adiabatic invariant*, cf. Theorem 23 in Chapter 5,§4 of [4]. This means that

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

which yields the following expression for the limit normal energy:

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

Thus, the term

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

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

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

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

Rigorous Derivation of the Additional Potential

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

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

and its potential part

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

i.e., $T_N^0 = U_N^0 = E_N^0/2$. This equipartition is a well known fact for the *time averages* of these 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][10][27].

Lemma 4.2 *For a given $\mathcal{M}$-converging sequence x^ϵ the kinetic and the potential part of the normal energy converges weakly in L^∞,*

$$T_N^\epsilon \xrightharpoonup{*} T_N^0, \qquad U_N^\epsilon \xrightharpoonup{*} U_N^0.$$

Moreover, we get

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

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

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

Hence, by Lemma 4.1

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

The next arguments follow closely the proof of the Virial Theorem as given for instance in [10]:

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

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

and therefore

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

This limit can be evaluated in a different way, using the description (17) of the normal oscillations,

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

A comparison of the different evaluations of the limit shows that

$$0 = T_N^0 - U_N^0,$$

which gives the desired result. □

We are now able to show the *adiabatic invariance* of E/ω.

Theorem 4.3 *Suppose x^ϵ $\mathcal{M}$-converges to x^0. Then, there is a constant Θ, such that*

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

Proof. Since we have

$$E_N^0 = H^0 - \frac{1}{2}|\dot{x}^0|^2 - V(x^0),$$

where H^0 is a constant, we get the time derivative

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

If we insert the limit equation (8) of Theorem 2.1 and note that

$$\langle DG(x^0)\lambda^*, \dot{x}^0\rangle = 0$$

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 expression

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

In the last step we have used the identity (14). By Lemma 4.2 we get $\sigma = E_N^0/\omega^2(x^0)$, 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 Θ, such that

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

which finishes the proof if we note the relation between E_N^0 and σ once again. □

Remark. In retrospective, the proof of this theorem was based on the equation of the normal oscillation as given in Lemma 4.1 *and* on the explicit limit equation of Theorem 2.1. Van Kampen [28, p. 103f] argues by formal use of the WKB method, that the invariance of $E_N^0/\omega(x^0)$ follows at once from equation (17). However, the following example shows that his argument is *not* correct since an *arbitrary* $\mathcal{O}(1)$-term can introduce resonances which precludes the adiabatic invariance: Suppose we have

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

with a *constant* frequency ω. 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

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

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

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

Theorem 4.4 *Suppose the initial values x_0^ϵ, $\dot{x}_0^\epsilon$ are given, such that the limits*

$$x_0^0 = \lim_{\epsilon\to 0} x_0^\epsilon \in \mathcal{M}, \qquad \dot{x}_0^0 = \lim_{\epsilon\to 0} \dot{x}_{0\mathcal{M}}^\epsilon \in T_{x_0^0}\mathcal{M}$$

and

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

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

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

with the corrected potential

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

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

Proof. In view of Theorem 2.1 and of Theorem 3.1 as well as the discussion preceding it, we only have to show that

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

for each vector field $X \in T_{x^0}\mathcal{M}$. In fact, using the relation (14), i.e.,

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

we get by Theorem 4.3

$$\begin{aligned} \left\langle \frac{1}{2} D^2 G(x^0) : \Sigma,\, X \right\rangle &= \frac{1}{2}\left\langle \operatorname{grad} \omega^2|_{x=x^0},\, X \right\rangle \cdot \frac{\Theta}{\omega(x^0)} \\ &= \Theta \left\langle \operatorname{grad} \omega|_{x=x^0},\, X \right\rangle. \end{aligned}$$

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

Remarks.

1. This theorem was first proved by Rubin and Ungar [22, p. 82f]. Independently, it can be found by means of an example in the work of Koppe and Jensen [15, eq. (7)]. Keller and Rubinstein [13] generalized the theorem to the wave equation $v_{tt} = \Delta v - \epsilon^{-2} \operatorname{grad} U(v)$.

2. 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 for Codimension $r > 1$

The case of codimension $r > 1$ is considerably more difficult and has been carefully analyzed by Takens [26]. 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 $\mathcal{M}$ *diagonizable*, if there is a field $(e_N^1, \ldots, e_N^r)$ of orthonormal bases of $N\mathcal{M}$, which are eigenvectors of $D_N^2 U$, i.e.,

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

Here, the eigenfrequencies ω_i shall depend *smoothly* on $x \in \mathcal{M}$. Takens proves ([26, Theorem 1]) that Theorem 4.3 extends to each normal component, if one can exclude certain *resonances*, i.e., if for $x \in \mathcal{M}$ we always have

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

and

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

Using this result, we can extend Theorem 4.4 in a straightforward fashion using the same proof. However, in general, there can be situations, where the limit x^0 *cannot* be described intrinsically by a deterministic initial value problem. In fact, Takens [26, 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 μ 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 non-uniquely solvable initial value problems, cf. [20]. Thus, for a fixed parameter μ we cannot describe the limit x^0 by a uniquely solvable initial value problem. Koiller [14] coined the notion "Takens-chaos" for this effect.

6 Two Examples

We will discuss two examples for the general codimension $r = 1$ case, which show the occurrence of the additional potential V_{add}. A further nontrivial example in the context of molecular dynamics including numerical simulations can be found in Bornemann and Schütte [7].

Example I. Illustrative, but Artificial

We take the Hamiltonian

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

describing a motion in $\mathbb{R}^2$. This gives the constraint manifold $\mathcal{M} = \{y = 0\}$ of dimension $m = 1$ and codimension $r = 1$. The equation of motion is given by

$$\begin{aligned} \ddot{x}^\epsilon &= -\partial_x V(x^\epsilon, y^\epsilon) - \epsilon^{-2}\omega(x^\epsilon)\omega'(x^\epsilon)(y^\epsilon)^2, \\ \ddot{y}^\epsilon &= -\partial_y V(x^\epsilon, y^\epsilon) - \epsilon^{-2}\omega^2(x^\epsilon)y^\epsilon. \end{aligned}$$

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 4.4, that

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

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

$$\ddot{x}^0 = -\partial_x V(x^0, 0) - \Theta\,\omega'(x^0), \qquad x^0(0) = x_0,\ \dot{x}^0(0) = v_0.$$

The constant Θ is given by

$$\Theta = \frac{w_0^2}{2\omega(x_0)}.$$

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

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

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 4.3 gives us in the this case a normal energy,

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

which is *constant* in time. Thus, a constant normal energy in the limit is only necessary for the vanishing of the potential correction V_{add}, but *not sufficient*, cf. the discussion behind formula (18).

Remark. This example is discussed at length by Koppe and Jensen [15], who in fact prove Theorem 4.4 only for this example. Gallavotti [9, 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

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

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

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

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

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

which in fact yields V_{wrong}. Interestingly enough, the potential correction suggested by Reich [21], the so called Fixman potential, also turns out to be the *same* V_{wrong}.

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, ϕ the B-field does not depend on the angle ϕ and its ϕ-component vanishes. Hence, there is a vector potential $\mathcal{A}$ with components $\mathcal{A} = (0, 0, A(r,z))$, such that $B = \operatorname{curl}\mathcal{A}$, i.e.,

$$B = \left(-\frac{\partial A}{\partial z}, \frac{\partial A}{\partial r} + \frac{A}{r}, 0\right). \tag{19}$$

The motion takes place according to the Langrangian

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

Since $\mathscr{L}$ does not depend on ϕ we have the conservation law

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

Thus, we can eliminate the ϕ variable. By gauging the vector potential initially to zero and assuming for the sake of simplicity, that the motion initially takes place tangentially to a plane of symmetry, we have

$$A(0) = 0, \qquad \dot{\phi}(0) = 0,$$

which implies

$$\dot{\phi} = -\frac{eA}{mr}$$

for all times. Hence, the Lagrangian reduces in (r, z)-coordinates to

$$\mathscr{L}_{\text{red}} = \frac{m}{2}(\dot{r}^2 + \dot{z}^2) - \frac{e^2A^2}{2m},$$

or equivalently, the motion fits into our framework with the Hamiltonian

$$H = \frac{1}{2}(\dot{r}^2 + \dot{z}^2) + \epsilon^{-2}\underbrace{\frac{A^2}{2}}_{=U(r,z)}, \qquad \epsilon = \frac{m}{e}.$$

For a large specific charge $\epsilon^{-1} = e/m$ we expect that the motion gyrates very rapidly around the line

$$\mathcal{M} = \{(r, z) : A(r, z) = 0\},$$

which is a field line by (19), the so called *guiding center* of the motion. The frequency of gyration is given by $\epsilon^{-1}\omega$ with

$$\begin{aligned} \omega^2 = \frac{1}{2}D^2A^2 : (e_N \otimes e_N)|_{\mathcal{M}} &= (DA \cdot e_N)^2 \\ &= |B|^2, \end{aligned}$$

since $DA \cdot e_N = \pm|B|$ on $\mathcal{M}$ by (19). Thus, just as in the case of uniform magnetic fields the frequency of gyration is given by the *Lamor frequency* $e|B|/m$.

Theorem 4.4 shows, that in the limit $\epsilon \to 0$ the average tangential motion along the guiding center $\mathcal{M}$ is governed by the potential

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

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

$$\ddot{s} = -\Theta\,\frac{\partial}{\partial s}|B|, \tag{20}$$

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 engeneering 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.

Remarks.

1. The first derivation of equation (20) by physical reasoning was given by the Swedish Nobel prize winner ALFVÉN [2, Chapter 2.3], see also [25][19].
2. The first mathematical discussion of the limit $e/m \to \infty$ was given by RUBIN AND UNGAR [22], who also discuss a nice mechanical analogue of the magnetic mirror.

Appendix

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

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

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:

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

We compute the derivatives and simplify them using the asymptotic results of Lemma 1.1:

$$
\begin{aligned}
\left.\frac{\partial \mathscr{L}}{\partial x_{\mathcal{M}}}\right|_{x=x^\epsilon} &= \frac{1}{2} D_{\mathcal{M}} g : (\dot{x}^\epsilon_{\mathcal{M}} \otimes \dot{x}^\epsilon_{\mathcal{M}}) + D_{\mathcal{M}} h : (\dot{x}^\epsilon_{\mathcal{M}} \otimes \dot{x}^\epsilon_{N}) \\
&\qquad - D_{\mathcal{M}} V - \epsilon^{-2} D_{\mathcal{M}} U \\
&= -\epsilon^{-2} D_{\mathcal{M}} U + \mathcal{O}(1),
\end{aligned}
$$

and

$$
\begin{aligned}
\left.\frac{\partial \mathscr{L}}{\partial x_{N}}\right|_{x=x^\epsilon} &= \frac{1}{2} D_{N} g : (\dot{x}^\epsilon_{\mathcal{M}} \otimes \dot{x}^\epsilon_{\mathcal{M}}) + D_{N} h : (\dot{x}^\epsilon_{\mathcal{M}} \otimes \dot{x}^\epsilon_{N}) \\
&\qquad - D_{N} V - \epsilon^{-2} D_{N} U \\
&= -\epsilon^{-2} D_{N} U + \mathcal{O}(1),
\end{aligned}
$$

and

$$
\left.\frac{\partial \mathscr{L}}{\partial \dot{x}_{\mathcal{M}}}\right|_{x=x^\epsilon} = g \cdot \dot{x}^\epsilon_{\mathcal{M}} + h \cdot \dot{x}^\epsilon_{N}, \qquad \left.\frac{\partial \mathscr{L}}{\partial \dot{x}_{N}}\right|_{x=x^\epsilon} = h^T \cdot \dot{x}^\epsilon_{\mathcal{M}} + \dot{x}^\epsilon_{N},
$$

thus,

$$
\left.\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{x}_{\mathcal{M}}}\right|_{x=x^\epsilon} = g \cdot \ddot{x}^\epsilon_{\mathcal{M}} + h \cdot \ddot{x}^\epsilon_{N} + \mathcal{O}(1), \qquad \left.\frac{d}{dt}\frac{\partial \mathscr{L}}{\partial \dot{x}_{N}}\right|_{x=x^\epsilon} = h^T \cdot \ddot{x}^\epsilon_{\mathcal{M}} + \ddot{x}^\epsilon_{N} + \mathcal{O}(1).
$$

Hence, the Euler-Langrange equations take the form (5), where the mass matrix is given by

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

because of $h(x^\epsilon_{\mathcal{M}}, 0) = 0$ and therefore $h(x^\epsilon) = h(x^\epsilon_{\mathcal{M}}, x^\epsilon_{N}) = \mathcal{O}(\epsilon)$. Now,

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

which gives

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

This proves the equations in (6).

References

[1] R. Abraham and J. E. Marsden. "Foundations of Mechanics". Addison-Wesley Publ. Co., Redwood City, New York, Bonn, 1985 printing of the 2nd edition (1985).

[2] H. Alfvén. "Cosmical Electrodynamics". Clarendon Press, Oxford (1950).

[3] V. I. Arnold. "Mathematical Methods of Classical Mechanics". Springer-Verlag, Berlin, Heidelberg, New York (1978).

[4] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. Mathematical Aspects of Classical and Celestial Mechanics. In V. I. Arnold, editor, "Dynamical Systems III". Springer-Verlag, Berlin, Heidelberg, New York (1988).

[5] G. Benettin, L. Galgani, and A. Giorgilli. Realization of holonomic contraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. I. *Commun. Math. Phys.* **113**, 87–103 (1987).

[6] G. Benettin, L. Galgani, and A. Giorgilli. Realization of holonomic contraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory. II. *Commun. Math. Phys.* **121**, 557–601 (1989).

[7] F. A. Bornemann and Ch. Schütte. "A Mathematical Approach to Smoothed Molecular Dynamics: Correcting Potentials for Freezing Bond Angles". Konrad–Zuse–Zentrum, Berlin (1995). Preprint SC 95-30.

[8] B. Dacorogna. "Weak Continuity and Weak Lower Semicontinuity of Non-Linear Functionals". Springer-Verlag, Berlin, Heidelberg, New York (1982).

[9] G. Gallavotti. "The Elements of Mechanics". Springer-Verlag, Berlin, Heidelberg, New York (1983).

[10] H. Goldstein. "Classical Mechanics". Addison-Wesley Publ. Co., Cambridge (1953).

[11] E. Hairer and G. Wanner. "Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems". Springer-Verlag, Berlin, Heidelberg, New York (1991).

[12] L. Hörmander. "The Analysis of Linear Partial Differential Operators I. Distribution Theory and Fourier Analysis". Springer-Verlag, Berlin, Heidelberg, New York, 2nd edition (1990).

[13] J. B. Keller and J. Rubinstein. Nonlinear wave motion in a strong potential. *Wave Motion* **13**, 291–302 (1991).

[14] J. Koiller. A note on classical motions under strong constraints. *J. Phys. A: Math. Gen.* **23**, L521–L527 (1990).

[15] 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).

[16] Ch. Lubich. Integration of stiff mechanical systems by Runge-Kutta methods. *Z. angew. Math. Phys.* **44**, 1022–1053 (1993).

[17] J. E. Marsden and T. S. Ratiu. "Introduction to Mechanics and Symmetry". Springer-Verlag, Berlin, Heidelberg, New York (1994).

[18] F. Murat and L. Tartar. H-convergence. In R. V. Kohn, editor, "Topics in the Mathematical Modelling of Composite Materials". Birkhäuser, Boston, Basel (1994).

[19] T. G. Northrop. "The Adiabadic Motion of Charged Particles". Interscience Publishers, New York, London, Sydney (1963).

[20] C. Pugh. Funnel sections. *J. Diff. Eqs.* **19**, 270–295 (1975).

[21] S. Reich. Smoothed dynamics of highly oscillatory Hamiltonian systems. *Physica D* **to appear** (1996).

[22] H. Rubin and P. Ungar. Motion under a strong constraining force. *Comm. Pure Appl. Math.* **10**, 65–87 (1957).

[23] W. Rudin. "Real and Complex Analysis". McGraw-Hill Publishing Co., New York, London, Sydney, Toronto (1987).

[24] H.-J. Schmidt. Models for constrained motion and d'Alembert's principle. *Z. angew. Math. Mech.* **73**, 155–163 (1993).

[25] L. Spitzer Jr. "Physics of Fully Ionized Gases". Interscience Publishers, New York, London, Sydney (1956).

[26] F. Takens. Motion under the influence of a strong constraining force. In Z. Nitecki and C. Robinson, editors, "Global Theory of Dynamical Systems, Evanston 1979". Springer-Verlag, Berlin, Heidelberg, New York (1980).

[27] M. Toda, R. Kubo, and N. Saitô. "Statistical Physics I. Equilibrium Statistical Mechanics". Springer-Verlag, Berlin, Heidelberg, New York, 2nd edition (1988).

[28] N. G. van Kampen. Elimination of fast variables. *Phys. Rep.* **124**, 69–160 (1985).