Doctoral Thesis

Stable and chaotic behavior in oscillatory and rotational motion with symmetry

Author(s):
Manz, Ueli

Publication Date:
2004

Permanent Link:
https://doi.org/10.3929/ethz-a-004930347

Rights / License:
In Copyright - Non-Commercial Use Permitted
Stable and Chaotic Behavior
in Oscillatory and Rotational Motion
with Symmetry

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
for the degree of
Doctor of Mathematics

presented by
UELI MANZ
Dipl. Math. ETH
born 10.08.1964
citizen of Pfäffikon, ZH

accepted on the recommendation of
Prof. Dr. Urs Kirchgraber, examiner
Prof. Dr. Jürg Marti, co-examiner
PD Dr. Daniel Stoffer, co-examiner

2004
Seite Leer / Blank leaf
Contents

Abstracts 7
Introduction 9

I Stability of Linear Periodic Systems of Differential Equations with Symmetry and an Application to Wind Turbine Motion 13

1 A Simple Model for a Wind Turbine 15
   1.1 Derivation of the System of Differential Equation 16
   1.2 Canonical Reduction to Differential Equations of Order 1 24
   1.3 Stable and Unstable Wind Turbines 26

2 A General Framework 29
   2.1 The Condition for Stability 29
   2.2 The Transition Stability–Instability 36

3 A Normal Form for Families of Matrices 39
   3.1 Block Triangular Form for Families of Matrices 39
   3.2 Properties of the Triangularized Matrices 44

4 Non-Autonomous Systems 45
   4.1 Auxilliary Results 45
   4.2 Dissipative Systems 53
   4.3 Canonical or Reversible Systems 58
   4.4 Four Dimensional Systems 69

5 Autonomous Systems 71
   5.1 Auxilliary Results 71
   5.2 Dissipative Systems 73
   5.3 Canonical or Reversible Systems 77
   5.4 Four Dimensional Systems 86

6 A Comparsion 89
   6.1 The Lyapunov-Schmidt Reduction 89
CONTENTS

6.2 The Comparison ................................................. 92

7 Tracing the Stability Boundary 95
7.1 The Computation of the Monodromy Matrix ................. 96
7.2 The Transformation to Normal Form ......................... 97
7.3 Tracing the Stability Boundary ............................ 99
7.4 The Stability Diagram of a Symplectic Map ................. 110

8 Stability of a Simple Model of a Two Bladed Rotor 115
8.1 Estimating the Moments of Intertia and the Stiffness Coefficients .............. 116
8.2 Computing the Stability Boundary .......................... 118

II Stability and Chaotic Behavior in Non-Linear Oscillations in the Presence of Symmetry with an Application to the Rotational Motion of a Dumbbell Satellite 127

1 The Dumbbell Satellite 129
1.1 The Equations of Lagrange .................................. 130
1.2 The Limiting Case \( t \to 0 \) .................................. 130
1.3 Stationary Solutions ......................................... 131

2 Keplerian Motion 135
2.1 Elliptic Motion .............................................. 135
2.2 Kepler's Laws ............................................... 138
2.3 Dynamics of Elliptic Motion ................................ 138
2.4 Motion Around the Center of Mass .......................... 140
2.5 Basic Properties ............................................ 142

3 A General Framework 143
3.1 The Definition of the General Framework ................. 143
3.2 Discussion of Assumption A5 ............................... 144
3.3 Lyapunov Stability vs. Practical Stability .................. 148

4 Preliminary Coordinate Transformations 151
4.1 Local Coordinates .......................................... 151
4.2 The Theory of Floquet ..................................... 152
4.3 The Structure of the Linear System ......................... 161
4.4 Removing the Linear Term of the Inhomogeneity ........... 162
4.5 The Complete Transformation .............................. 166

5 Formal First Integrals 169
5.1 The Definition of Formal First Integrals ................. 170
5.2 The Formal Perturbation Scheme .......................... 171
5.3 Inverting the Operator \( L \) ................................. 177
5.4 The Symmetry of the Perturbation Equations .............. 185
CONTENTS

5.5 The Solvability of the Perturbation Scheme .................. 198

6 Rigorous Estimates for First Integrals .......................... 203
  6.1 Definition of Norms ............................................. 204
  6.2 Recursive Estimates for $I^k_i$ ................................. 211
  6.3 A Priori Estimates for $I^k_i$ ................................. 217
  6.4 Estimates for $\dot{I}^r_i$ ..................................... 219
  6.5 Estimates for $I^r_i(t)$ and $\dot{I}^r_i(t)$ .................... 222
  6.6 The Optimal Truncation Order ................................. 225

7 First Integrals and Practical Stability ......................... 233
  7.1 Practical Stability ........................................... 234
  7.2 Explicit Computations and Estimates ......................... 240
  7.3 Nekoroshev-Type Results .................................... 242

8 Stability of the Dumbbell Satellite ............................ 249
  8.1 $2\pi$-periodic Solutions .................................... 250
  8.2 Linear Stability .............................................. 254
  8.3 The Stability of the Trivial Solution for $e = 0$ ............ 255
  8.4 Practical Stability for Prescribed Values of $e$ ............ 262
  8.5 Practical Stability for a Range of Values of $e$ ............ 281
  8.6 The Influence of the Coordinate Transformation ............. 305
  8.7 Computational Notes ........................................ 312

9 Chaotic Behavior for Small Eccentricities ..................... 315
  9.1 The Method of Melnikov ..................................... 316
  9.2 Chaotic Behavior for Small Eccentricities ................. 323

10 Verification of Chaos for Large Eccentricities ............... 331
  10.1 Numerical Shadowing ....................................... 332
  10.2 An Algorithm for the Verification of Chaos ................ 347
  10.3 Verification of Chaos for $e = 0.3$ .......................... 358
  10.4 Computational Notes ...................................... 363

Bibliography .......................................................... 365

Curriculum Vitae .................................................... 367
English Abstract

In the first part of this Ph.D. thesis we consider a simple model of a two-bladed rotor, already investigated by Meyer in 1984 (cf. [18]). It leads to a linear, homogeneous system of differential equations with time-dependent coefficients that depends on two parameters. It is of interest to determine the points in the parameter plane for which the system is stable, i.e. all solutions are bounded. For that purpose we investigate the spectrum of the so-called monodromy matrix. Since the dimension of this matrix is large, we transform the matrix in a whole neighborhood of a point in the parameter plane to block triangular form. The spectrum remains invariant under this normal form transformation. In the canonical or reversible case the spectra of the submatrices in the main diagonal show the same symmetries as the spectrum of the original matrix.

Eventually we obtain fairly accessible conditions for stability boundaries and we can prove (under certain conditions) that the boundary between a stable and an unstable region is a curve, locally at least. In contrast to Meyer's our approach covers the canonical and reversible case as well.

Based on these results we develop an efficient algorithm to compute stability boundaries and we apply it to the rotor problem.

In the second part of the Ph.D. thesis we consider a model of a satellite in the shape of a dumbbell. While the center of mass is moving on a Keplerian ellipse, the bar of the dumbbell may revolve around its center of mass. The corresponding differential equation is of second order, periodic in the true anomaly and has the eccentricity of the ellipse as its only parameter. In contrast to the first part of this work the differential equation is nonlinear and it therefore shows a much greater variety of phenomena: For certain initial positions the motion of the bar around its center of mass is stable for "very long" times, while for other initial positions chaotic behavior is established.

The stability results are obtained by means of formal first integrals following ideas of Giorgilli (cf. [7], [8]). On the one hand we construct these integrals explicitly by means of a computer algebra system and on the other hand we construct a priori estimates.

To prove chaotic behavior we make use of a suitable version of the Shadowing Lemma. For small eccentricities we apply the well-known method of Melnikov, while for larger eccentricities we develop a computer assisted method following ideas of Palmer, Stoffer and Kirchgraber (cf. [24], [22]).
Deutsche Zusammenfassung

Im ersten Teil dieser Dissertation betrachten wir ein einfaches Modell eines zweiflügligen Rotors, welches Meyer 1984 untersuchte (cf. [18]). Es führt auf ein lineares, homogenes System von Differentialgleichungen mit zeitperiodischen Koeffizienten, das von zwei Parametern abhängt.

Es ist nun von Interesse herauszufinden, für welche Punkte der Parameterebene, der Rotor stabil ist, d.h. alle Lösungen der Differentialgleichungen beschränkt sind. Dazu untersuchen wir das Spektrum der sog. Monodromiematrix. Da die Dimension dieser Matrix groß ist, transformieren wir sie für eine ganze Umgebung eines Punktes in der Parameterebene in Block-Dreiecksform. Unter dieser Normalformtransformation bleibt das Spektrum invariant. Im kanonischen oder reversiblen Fall zeigen die Spekten der Submatrizen in der Hauptdiagonalen dieselben Symmetrien wie das Spektrum der ursprünglichen Matrix.

Wir gewinnen daraus relativ einfach zugängliche Bedingungen für Stabilitätsgrenzen und können unter gewissen Bedingungen so insbesondere zeigen, dass die Grenze zwischen stabilen und instabilen Gebieten lokal eine Kurve ist. Im Gegensatz zur Arbeit von Meyer funktioniert dieser Ansatz auch im delikateren kanonischen und reversiblen Fall. Basierend auf diesen Resultaten entwickeln wir einen effizienten Algorithmus zur Berechnung von Stabilitätsgrenzen und wenden ihn auf das Modell des Rotors an.


Die Stabilitätsaussagen erhalten wir mit Hilfe formalcr erster Integrale, wobei wir Ideen von Giorgilli folgen (cf. [7], [8]). Auf der einen Seite konstruieren wir diese Integrale explizit mit Hilfe eines Computer-algebrasystems, auf der andere Seite gewinnen wir a priori-Abschätzungen.

Um chaotisches Verhalten nachzuweisen, setzen wir geeignete Versionen des Schattenlemmas ein. Für kleine Exzentrizitäten benutzen wir die bekannte Methode von Melnikov, während wir für grosse Exzentrizitäten eine Computer-unterstützte Methode entwickeln, die auf Ideen von Palmer, Stoffer und Kirchgraber zurückgreift (cf. [24], [22]).
Introduction

Stable and Chaotic Behavior in Oscillatory and Rotational Motion with Symmetry

In this thesis we consider a model for the motion of a wind turbine and a model for the rotational motion of a dumbbell shaped satellite. In both models questions of stability are of interest. In addition the dumbbell satellite may even behave chaotically, as will be shown.

Stability of Linear Periodic Systems of Differential Equations with Symmetry and an Application to Wind Turbine Motion

In Part I of this work we reconsider a simplified model of a wind turbine introduced by Meyer in 1984 (cf. [18]). It consists of a two-bladed rotor attached to a rigid hub/nacelle system at the top of a flexible tower. The motion of the top is described by the horizontal displacements and via the pitching, rolling and yawing angle.

For small displacements and small angles the tower acts approximately like a massless spring. Applying the principle of linear and angular momentum we derive a set of ten first-order non-autonomous linear differential equations with two parameters $\Omega$ and $\varepsilon$, describing the angular velocity and the unbalance of the rotor, respectively. This system is of the form $\dot{z} = A(t,\varepsilon,\Omega)z$, where $z$ is a vector of dimension 10 and $A(t,\varepsilon,\Omega)$ is a $T$-periodic $10 \times 10$-matrix function.

For a particular pair of values of the parameters the system is called stable, if all solutions are bounded and unstable otherwise. Due to the periodicity of $A$ with respect to $t$, the stability of the system depends on the spectrum of the monodromy matrix $M := X(T)$, where $X(t)$ denotes its principal matrix solution of $\dot{z} = A(t,\varepsilon,\Omega)z$.

For dissipative systems the spectrum of the monodromy matrix is but symmetric with respect to the real axis in general. Therefore stability is typically lost, if either one real eigenvalue, or a pair of complex conjugate eigenvalues leaves the unit circle, while the other eigenvalues remain inside the unit circle. For systems with additional symmetries, e.g. if the system is Hamiltonian or reversible, the situation is more subtle. In these cases the spectrum of the monodromy matrix in addition is symmetric with respect to the unit circle. Therefore stability is typically lost if either a pair of simple complex conjugate eigenvalues on the unit circle coalesces at $\pm 1$ and then leaves the unit circle or if two
pairs of simple complex conjugate eigenvalues on the unit circle coalesce and then leave the unit circle.

For fixed values of the parameters the monodromy matrix and its eigenvalues may be computed numerically. Thus one may investigate stability for a certain range of the parameters by inspecting a suitable fine grid of points in the parameter plane with respect to their stability behavior. Yet this "brute force"-method is rather expensive. A more economical approach is to trace stability boundaries, i.e. we look for curves in the $\varepsilon$-$\Omega$-plane with the following property: In every neighborhood of a point $(\varepsilon_0, \Omega_0)$ on such a curve there are points for which the system is stable as well as points for which it is unstable.

For dissipative systems Kirchgraber, Meyer and Schweitzer (cf. [21]) derived a one-dimensional equation for the stability boundary using Lyapunov-Schmidt reduction techniques. Using the Implicit Function Theorem they concluded that the stability boundaries are curves, at least locally, indeed. In the mathematically more interesting canonical (Hamiltonian) case the (standard) Lyapunov-Schmidt reduction led to a two-dimensional system of equations to which the Implicit Function Theorem did not apply, and therefore Kirchgraber, Meyer and Schweitzer were not able to prove that the stability boundaries are curves locally. These difficulties may be avoided by using reduction techniques that respect symmetries. Indeed, Vanderbauwhede developed such an approach for symplectic operators after having learned from our problem (cf. [26]).

Our approach presented here (see also [25]) does not care at all about preserving the symmetry of the operator. Rather it takes advantage of the symmetry properties of its spectrum, which is preserved anyway. As an additional benefit of this approach we can treat the Hamiltonian case as the reversible case simultaneously.

Here is a brief description of our approach. We first introduce a block triangular normal form for the monodromy matrix $M$ for $(\varepsilon, \Omega)$ in a full neighborhood of some given fixed point $(\varepsilon_0, \Omega_0)$. This allows us to investigate the spectra of every submatrix in the main diagonal separately. Since the dimensions of the submatrices are small we obtain fairly accessible conditions to discriminate stable from unstable behavior. This allows us, in particular, to prove that in some neighborhood of a given point $(\varepsilon_0, \Omega_0)$ the stability boundary typically is either a curve or consists of two curves that intersect transversally at $(\varepsilon_0, \Omega_0)$, both in the Hamiltonian and the reversible case.

These results are used to construct an algorithm for tracing stability boundary curves. A detailed description of the algorithm is presented together with an application to the wind turbine model.
Stability and Chaotic Behavior in Non-Linear Oscillations in the Presence of Symmetry with an Application to the Rotational Motion of a Dumbbell Satellite

In Part II of this work we consider a satellite with a shape of a dumbbell. More precisely, the satellite consists of two point masses connected by a rigid massless bar. The motion of such a satellite is conveniently described in terms of polar coordinates for its center of mass and by the angle $\theta$ between the axis of the dumbbell and the radius vector of the center of mass.

Applying the method of Lagrange we obtain a set of three non-linear second-order differential equations. Since the length of the satellite is small compared to the diameter of its orbit we consider the limiting case where the length tends to 0. The first two equations then decouple. They describe the Keplerian Motion of the center of mass. Solving Kepler's problem and substituting the result into the third equation we obtain a single equation for the angle $\theta$.

This equation may be written as a two-dimensional system of periodic first-order differential equations with the eccentricity $e$ of the underlying elliptic motion being the only parameter of the problem. This system is nonlinear and is reversible, a fact that is very important for the subsequent considerations.

If the satellite is moving on a circular orbit, $e = 0$, then the equation for $\theta$ is reduced to the equation of a simple pendulum. This problem has two equilibria, $\theta = 0$ is of elliptic type, $\theta = \frac{\pi}{2}$ is a saddle point. Since $\theta$ is the angle between the axis of the dumbbell and the radius vector of the center of mass we have the following picture. For $\theta = 0$ the dumbbell satellite moves on the circular orbit like a spoke of a wheel, while for $\theta = \frac{\pi}{2}$ the dumbbell is in tangential position to the circular orbit. While the first equilibrium is stable the second is not.

Now assume that $e$ does no longer vanish. Then the following two questions arise: Will the periodic orbits which emanate from the stable equilibrium be stable? Will the saddle type equilibrium generate chaotic behavior?

Consider the first question in the following frame. Assume that we are given a periodic reversible solution of a nonlinear reversible system of differential equations. Starting with local coordinates we perform a series of coordinate changes that eventually lead to a system of periodically perturbed harmonic oscillator with variables $x$ and $y$ corresponding to $\theta$ and $\alpha$.

Then we take up an idea introduced by Giorgilli for the Hamiltonian case in [8]. The crucial observation is that a system of perturbed Hamiltonian oscillators admits formal first integrals of the form $I(x, y) := \frac{1}{2}(x^2 + y^2) + \sum_{k=3}^{\infty} I^k(x, y)$. Since these power series do not converge in general, Giorgilli suggested to work with truncated first integrals. Such truncated series are only almost constant along solutions. Therefore one obtains stability results for finite times only. Given a disk with radius $\rho_{\text{start}}$ for the initial values and a disk with radius $\rho_{\text{end}} > \rho_{\text{start}}$ one can estimate (from below) the maximum time $T_{\text{max}}$ for which solutions will stay within the second disk. This type of stability concept is referred to as practical stability.
There are basically two possibilities to estimate $T_{\text{max}}$. Either one computes the polynomials $I_k$ explicitly using a computer algebra system, or else one can derive a priori estimates for $I_k$. This permits to determine the so-called optimal truncation order, which leads to exponentially long stability time intervals, i.e., $T_{\text{max}} \to \infty$ exponentially as $\rho_{\text{vol}} \to 0$.

In the present work Giorgilli's approach is extended to time-periodic reversible systems and moreover expansions with respect to system's parameters are systematically taken into account. To formally determine the functions $I_k$ of the sought integrals one obtains a recursive scheme of so-called perturbation equations of the form $LI_k = K_k$, where $L$ is a linear partial differential operator, $I_k$ is a polynomial of order $k$ and $K_k$ is a polynomial of order $k$ depending on the solutions of the perturbations up to order $k - 1$.

An important step is to prove that the perturbation equations are solvable under the assumption of reversibility. This is done here for the first time by so to speak explicit computation. To this end we study in detail the properties of the differential operator $L$ and the symmetry properties of the involved polynomials. Careful analyses eventually show that reversibility indeed implies the solvability of the perturbation equations. Then following Giorgilli we derive a priori estimates for the truncated first integrals and for their derivatives. These estimates eventually yield long-term stability results. Even if explicit computations are used one needs a priori estimates for the derivatives, since infinite series are involved.

Finally we present practical stability results for the dumbbell satellite problem. On the one hand we give results for fixed values of the eccentricity, on the other hand we illustrate that our approach is capable to guarantee long-term stability for intervals of $e$, the only parameter present in the problem.

In the last chapters we discuss the question of chaotic behavior. For small eccentricities we apply the well-known Melnikov method to prove the existence of chaotic behavior following a paper of Kirchgraber and Stoffer (cf. [15]). Unfortunately the Melnikov method does not give concrete information on the size of $e$ to which it applies.

For larger eccentricities we use a new computer assisted method due to Palmer, Stoffer and Kirchgraber (cf. [24], [22]). The principal idea is to compute two orbits that are periodic up to discretisation and round-off errors and that are very close at some point. Combining arbitrarily such "periodic" orbits one constructs (uncountably many) pseudo orbits. Applying a suitable version of the Shadowing Lemma guarantees the existence of shadowing orbits and chaotic behavior may be established. We had to slightly generalize the previously available theory due to difficulties caused by the strong hyperbolicity of our problem leading to very stiff equations. The key step is to replace the Poincaré map by a series of maps, which are easier to handle.
Part I

Stability of Linear Periodic Systems of Differential Equations with Symmetry and an Application to Wind Turbine Motion
A Simple Model for a Wind Turbine

We consider a simplified model of a wind turbine (cf. [18] and [21]). It consists of a two-bladed rotor attached to a rigid hub/nacelle system at the top of a flexible tower (cf. Figure 1.1).

Fig. 1.1: A two-bladed rotor attached to a rigid hub/nacelle system at the top of a flexible tower.
1. A Simple Model for a Wind Turbine

1.1 Derivation of the System of Differential Equation

Let \((e_x, e_y, e_z)\) be the coordinate system attached to the top of the tower. For further purposes we define a second coordinate system \((e_1, e_2, e_3)\) attached to the rotating blades (cf. Figure 1.2).

![Diagram of coordinate systems](image)

**Fig. 1.2:** The coordinate systems \((e_x, e_y, e_z)\) attached to the top of the tower and \((e_1, e_2, e_3)\) attached to the rotating blades.

In the \((e_x, e_y, e_z)\)-coordinate system the top of the tower may be described by the displacements \(x\) and \(y\) and the angles \(\alpha\, \beta\) and \(\gamma\) (cf. Figure 1.3).

A Linear Model for the Forces and Angular Momentum

Assume that the cross section of the tower is symmetric and that \(x, y, \alpha, \beta, \gamma\) and their derivative with respect to the time are small.

Let \(F = F_x e_x + F_y e_y\) the force caused by a displacement and let \(M = M_x e_x + M_y e_y + M_z e_z\) be the angular momentum caused by a rotation of the top of the tower.

For small displacements \(x, y\) and small angles \(\alpha, \beta\) and \(\gamma\) the flexible tower acts like a massless spring described by the following set of equations:

\[
\begin{align*}
  y &= \frac{F_y}{c_1}, & \alpha &= -\frac{F_y}{c_2}, & x &= \frac{M_y}{c_2}, & \beta &= \frac{M_y}{c_3}, \\
  x &= \frac{F_x}{c_1}, & \beta &= \frac{F_x}{c_2}, & y &= -\frac{M_y}{c_2}, & \alpha &= \frac{M_x}{c_3},
\end{align*}
\]
1.1. Derivation of the System of Differential Equation

\[ \gamma = \frac{M_z}{c_4}, \]

where \( c_1, c_2, c_3 \) and \( c_4 \) are positive constants.

\[ c_1, c_2, c_3 \text{ and } c_4 \]

\[ \text{yawing angle } \gamma, \text{ pitching angle } \alpha, \text{ axial displacement } y, \text{ radial displacement } x, \text{ rolling angle } \beta \]

\[ \text{Fig. 1.3: The displacements } x \text{ and } y \text{ and the angles } \alpha, \beta \text{ and } \gamma \text{ of the top of the tower.} \]

A superposition of the actions leads to

\[ x = \frac{M_y}{c_2} + \frac{F_x}{c_1}, \quad \alpha = \frac{M_x}{c_3} - \frac{F_y}{c_2}, \]
\[ y = -\frac{M_x}{c_2} + \frac{F_y}{c_1}, \quad \beta = \frac{M_y}{c_3} + \frac{F_x}{c_2}, \]
\[ \gamma = \frac{M_z}{c_4}. \]

Solving this linear system of equations for \( F_x, F_y, M_x, M_y \) and \( M_z \) gives

\[ F_x = k_1 x - k_2 \beta, \quad M_x = k_2 y + k_3 \alpha, \]
\[ F_y = k_1 y + k_2 \alpha, \quad M_y = -k_2 x + k_3 \beta, \]
\[ M_z = k_4 \gamma, \]

where the constants \( k_1, k_2, k_3 \) and \( k_4 \) are defined by

\[ k_1 := \frac{c_1 c_2^2}{c_2^2 - c_1 c_3}, \quad k_2 := \frac{c_1 c_2 c_3}{c_2^2 - c_1 c_3}, \quad k_3 := \frac{c_2^2 c_3}{c_2^2 - c_1 c_3}, \quad k_4 := c_4. \]
The Inertia Tensor

In the rotating coordinate system \((e_1, e_2, e_3)\) the inertia tensor is in diagonal form

\[
\bar{I} = \begin{pmatrix}
A & 0 & 0 \\
0 & B & 0 \\
0 & 0 & C
\end{pmatrix},
\]

where \(A\), \(B\) and \(C\) are the moments of inertia with respect to a rotation around the axis of the coordinate system.

Using the transformation matrix

\[
T = \begin{pmatrix}
\cos(\Omega t) & 0 & -\sin(\Omega t) \\
0 & 1 & 0 \\
\sin(\Omega t) & 0 & \cos(\Omega t)
\end{pmatrix}
\]

we may compute the inertia tensor \(I\) in the \((e_x, e_y, e_z)\)-coordinate system:

\[
T^T \bar{I} T = \begin{pmatrix}
\cos(\Omega t) & 0 & -\sin(\Omega t) \\
0 & 1 & 0 \\
\sin(\Omega t) & 0 & \cos(\Omega t)
\end{pmatrix} \begin{pmatrix}
A & 0 & 0 \\
0 & B & 0 \\
0 & 0 & C
\end{pmatrix} \begin{pmatrix}
\cos(\Omega t) & 0 & \sin(\Omega t) \\
0 & 1 & 0 \\
-\sin(\Omega t) & 0 & \cos(\Omega t)
\end{pmatrix} =
\]

\[
= \begin{pmatrix}
A \cos(\Omega t)^2 + C \sin(\Omega t) & B & (A - C) \cos(\Omega t) \sin(\Omega t)
\end{pmatrix}
\]

\[
= \begin{pmatrix}
A + \frac{A-C}{2} (-1 + \cos(2\Omega t)) & 0 & \frac{A-C}{2} \sin(2\Omega t)
\end{pmatrix}
\]

\[
=: \begin{pmatrix}
I_x & 0 & I_{xz}
\end{pmatrix} =: I.
\]

\(I_x\), \(I_y\) and \(I_z\) are the moments of inertia with respect to the \((e_x, e_y, e_z)\)-coordinate system, while \(I_{xz}\) is the only non-vanishing product of inertia.

The Linear and the Angular Momentum

The linear and the angular momentum of the rotor are defined by

\[
p = mv,
\]

\[
L = I\omega,
\]

where \(m\) is the total mass of the nacelle and the blades and \(I\) is the inertia tensor computed above, \(v\) and \(\omega\) are the linear and angular velocity

\[
v := \dot{x}e_x + \dot{y}e_y \quad \text{and} \quad \omega := \dot{\alpha}e_x + (\beta + \Omega)e_y + \dot{\gamma}e_z,
\]
1.1. Derivation of the System of Differential Equation

respectively. Thus the following holds:

\[ \mathbf{p} = m\dot{x}\mathbf{e}_x + m\dot{y}\mathbf{e}_y, \]

\[ \mathbf{L} = (I_x\dot{\alpha} + I_{xx}\dot{\gamma})\mathbf{e}_x + I_y(\dot{\beta} + \Omega)e_y + (I_{xz}\dot{\alpha} + I_z\dot{\gamma})\mathbf{e}_z. \]

**Principle of Linear and Angular Momentum**

Applying Newton’s second law we obtain the following equations of motion:

\[ \dot{\mathbf{p}} = -\mathbf{F}, \]

\[ \dot{\mathbf{L}} = -\mathbf{M}. \tag{1.1} \]

Note that the forces and angular momentum acting on the top of the tower have signs opposite to the forces and angular momentum resulting from displacement and rotations computed in a previous subsection.

For the derivatives of \( \mathbf{p} \) and \( \mathbf{L} \) with respect to the time \( t \) we have

\[ \dot{\mathbf{p}} = m\ddot{x}\mathbf{e}_x + m\ddot{y}\mathbf{e}_y + m\dot{x}\dot{e}_x + m\dot{y}\dot{e}_y \]

\[ \dot{\mathbf{L}} = \frac{d}{dt} \left( I_x\dot{\alpha} + I_{xx}\dot{\gamma} \right)\mathbf{e}_x + \left( I_x\ddot{\alpha} + I_{xx}\ddot{\gamma} \right)\dot{e}_x + \frac{d}{dt} \left( I_y(\dot{\beta} + \Omega) \right)e_y + \left( I_y(\ddot{\beta} + \Omega) \right)\dot{e}_y + \frac{d}{dt} \left( I_{xz}\dot{\alpha} + I_z\dot{\gamma} \right)e_z + \left( I_{xz}\ddot{\alpha} + I_z\ddot{\gamma} \right)\dot{e}_z. \]

Since \( \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z \) is rotating with the angular velocity

\[ \omega = \dot{\alpha}\mathbf{e}_x + \dot{\beta}\mathbf{e}_y + \dot{\gamma}\mathbf{e}_z \]

we have

\[ \dot{e}_x = \omega \times e_x = \gamma e_y - \beta e_z, \]

\[ \dot{e}_y = \omega \times e_y = -\gamma e_x + \alpha e_z, \]

\[ \dot{e}_z = \omega \times e_z = \beta e_x - \alpha e_y. \]

These formulae allows us to complete the computation of \( \dot{\mathbf{p}} \) and \( \dot{\mathbf{L}} \). Note, since \( \dot{x}, \dot{y}, \dot{\alpha}, \dot{\beta} \) and \( \dot{\gamma} \) are assumed to be small, we may neglect terms of order 2 and higher with respect to \( \dot{\alpha}, \dot{\beta} \) and \( \dot{\gamma} \). This leads to

\[ \dot{\mathbf{p}} = m\ddot{x}\mathbf{e}_x + m\ddot{y}\mathbf{e}_y, \]

\[ \dot{\mathbf{L}} = \left( \frac{d}{dt} \left( I_x\dot{\alpha} + I_{xx}\dot{\gamma} \right) - I_y\Omega \dot{\gamma} \right)\mathbf{e}_x + I_y\ddot{\beta}\mathbf{e}_y + \left( \frac{d}{dt} \left( I_{xz}\dot{\alpha} + I_z\dot{\gamma} \right) + I_y\Omega \dot{\alpha} \right)\mathbf{e}_z. \]
Substituting the expressions for $\mathbf{F}$, $\mathbf{M}$, $\mathbf{p}$ and $\dot{\mathbf{L}}$ obtained above we find

\begin{align*}
m\ddot{x} + k_1 x - k_2 \beta &= 0, \\
m\ddot{y} + k_1 y + k_2 \alpha &= 0,
\end{align*}

\begin{align*}
\frac{d}{dt} (I_x \dot{\alpha} + I_{xz} \dot{\gamma}) - I_y \Omega \dot{\gamma} + k_3 \alpha + k_2 y &= 0, \\
I_y \ddot{\beta} + k_3 \beta - k_2 x &= 0, \\
\frac{d}{dt} (I_{xz} \dot{\alpha} + I_z \dot{\gamma}) + I_y \Omega \dot{\alpha} + k_4 \gamma &= 0.
\end{align*}

Finally we substitute the terms found for moments of inertia $I_x$, $I_y$, $I_z$ and the product of inertia $I_{xz}$:

\begin{align*}
\frac{d}{dt} (m\ddot{x} + k_1 x - k_2 \beta) &= 0, \\
\frac{d}{dt} (m\ddot{y} + k_1 y + k_2 \alpha) &= 0,
\end{align*}

\begin{align*}
\frac{d}{dt} \left( (A + \frac{A-C}{2} (-1 + \cos(2\Omega t))) \dot{\alpha} + \frac{A-C}{2} \sin(2\Omega t) \dot{\gamma} \right) - B \Omega \dot{\gamma} + k_3 \alpha + k_2 y &= 0, \\
\frac{d}{dt} (B \ddot{\beta}) + k_3 \beta - k_2 x &= 0, \\
\frac{d}{dt} \left( \frac{A-C}{2} \sin(2\Omega t) \dot{\alpha} + (A + \frac{A-C}{2} (-1 - \cos(2\Omega t))) \dot{\gamma} \right) + B \Omega \dot{\alpha} + k_4 \gamma &= 0.
\end{align*}

Decomposition of the System of Differential Equations

The system of differential equations (1.3) may be decomposed into two uncoupled subsystems. To this end let

\[
q := \begin{pmatrix} q^I \\ q^{II} \end{pmatrix}, \quad \text{where} \quad q^I := \begin{pmatrix} x \\ \beta \end{pmatrix} \quad \text{and} \quad q^{II} := \begin{pmatrix} y \\ -\alpha \end{pmatrix}.
\]

Further let

\[
M_0^I := \begin{pmatrix} m & 0 \\ 0 & B \end{pmatrix}, \quad M_0^{II} := \begin{pmatrix} m & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & A \end{pmatrix},
\]

\[
M_1(t, \Omega)^I := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_1(t, \Omega)^{II} := \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{A}{2} (-1 + \cos(2\Omega t)) & \frac{A}{2} \sin(2\Omega t) \\ 0 & \frac{A}{2} \sin(2\Omega t) & \frac{A}{2} (-1 - \cos(2\Omega t)) \end{pmatrix},
\]
1.1. Derivation of the System of Differential Equation

\[ G_0^I := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad G_0^{II} := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \Omega B \\ 0 & -\Omega B & 0 \end{pmatrix}, \]

\[ K_0^I := \begin{pmatrix} k_1 & -k_2 \\ -k_2 & k_3 \end{pmatrix}, \quad K_0^{II} := \begin{pmatrix} k_1 & -k_2 & 0 \\ -k_2 & k_3 & 0 \\ 0 & 0 & k_4 \end{pmatrix} \]

and

\[ M_0 := \begin{pmatrix} M_0^I \\ M_0^{II} \end{pmatrix}, \quad M_1(t, \Omega) := \begin{pmatrix} M_1^I(t, \Omega) \\ M_1^{II}(t, \Omega) \end{pmatrix}, \]

\[ G_0 := \begin{pmatrix} G_0^I \\ G_0^{II} \end{pmatrix}, \quad K_0 := \begin{pmatrix} K_0^I \\ K_0^{II} \end{pmatrix}. \]

Using these notations and putting

\[ \varepsilon := \frac{A - C}{A}, \]

Eqs. (1.3) may be written as

\[ \frac{d}{dt} \left( (M_0 + \varepsilon M_1(t, \Omega)) \dot{q} \right) + G_0 \ddot{q} + K_0 q = 0 \] (1.4)

or split into two decoupled subsystems:

\[ \frac{d}{dt} \left( (M_0^I + \varepsilon M_1^I(t, \Omega)) \dot{q}^I \right) + G_0^I \ddot{q}^I + K_0^I q^I = 0, \]

\[ \frac{d}{dt} \left( (M_0^{II} + \varepsilon M_1^{II}(t, \Omega)) \dot{q}^{II} \right) + G_0^{II} \ddot{q}^{II} + K_0^{II} q^{II} = 0. \] (1.5)

The latter equations are refereed as subsystem I and subsystem II, respectively.

Wind Turbines With More Than Two Rotor Blades

Our derivation of the system of differential equations do not depend on the number of rotor blades. The differences lie only in the moments of intertia, as we show in the following.

Consider a rotor with \( n \) rotor blades as shown in Figure 1.4. In a first step we compute the intertia tensors of the individual blades. Due to symmetry reasons the inertia tensor \( I_0 \) has the following simple form

\[ I_0 = \begin{pmatrix} A_0 & 0 & 0 \\ 0 & B_0 & 0 \\ 0 & 0 & C_0 \end{pmatrix}. \]
1. A Simple Model for a Wind Turbine

Fig. 1.4: A wind turbine with $n$ rotor blades.

To compute the inertia tensors of the other blades we introduce the following transformation matrix

$$ T := \begin{pmatrix} \cos\left(\frac{2\pi k}{n}\right) & 0 & -\sin\left(\frac{2\pi k}{n}\right) \\ 0 & 1 & 0 \\ -\sin\left(\frac{2\pi k}{n}\right) & 0 & \cos\left(\frac{2\pi k}{n}\right) \end{pmatrix}. $$

Thus we find

$$ I_k = T^T I_0 T = $$

$$ = \begin{pmatrix} \cos\left(\frac{2\pi k}{n}\right) & 0 & -\sin\left(\frac{2\pi k}{n}\right) \\ 0 & 1 & 0 \\ -\sin\left(\frac{2\pi k}{n}\right) & 0 & \cos\left(\frac{2\pi k}{n}\right) \end{pmatrix} \begin{pmatrix} A_0 & 0 & 0 \\ 0 & B_0 & 0 \\ 0 & 0 & C_0 \end{pmatrix} \begin{pmatrix} \cos\left(\frac{2\pi k}{n}\right) & 0 & -\sin\left(\frac{2\pi k}{n}\right) \\ 0 & 1 & 0 \\ -\sin\left(\frac{2\pi k}{n}\right) & 0 & \cos\left(\frac{2\pi k}{n}\right) \end{pmatrix} = $$

$$ = \begin{pmatrix} A_0 \cos^2\left(\frac{2\pi k}{n}\right) + C_0 \sin^2\left(\frac{2\pi k}{n}\right) & 0 & (A_0 - C_0) \cos\left(\frac{2\pi k}{n}\right) \sin\left(\frac{2\pi k}{n}\right) \\ 0 & B_0 & 0 \\ (A_0 - C_0) \cos\left(\frac{2\pi k}{n}\right) \sin\left(\frac{2\pi k}{n}\right) & 0 & A_0 \cos^2\left(\frac{2\pi k}{n}\right) + C_0 \sin^2\left(\frac{2\pi k}{n}\right) \end{pmatrix}. $$

The inertia tensor of the whole rotor then reads

$$ I := \sum_{k=0}^{n-1} I_k. $$

For $n = 2$ we find

$$ \sum_{k=0}^{n-1} \cos^2\left(\frac{2\pi k}{n}\right) = \cos^2(0) + \cos^2(\pi) = 2, $$

$$ \sum_{k=0}^{n-1} \sin^2\left(\frac{2\pi k}{n}\right) = \sin^2(0) + \sin^2(\pi) = 0. $$
1.1. Derivation of the System of Differential Equation

and

\[ \sum_{k=0}^{n-1} \cos \left( \frac{2\pi k}{n} \right) \sin \left( \frac{2\pi k}{n} \right) = \cos(0) \sin(0) + \cos(\pi) \sin(\pi) = 0. \]

For \( n > 2 \) we find

\[ \sum_{k=0}^{n-1} \cos^2 \left( \frac{2\pi k}{n} \right) = \sum_{k=0}^{n-1} \frac{1}{4} \left( e^{\frac{2\pi ik}{n}} + e^{-\frac{2\pi ik}{n}} \right)^2 = \sum_{k=0}^{n-1} \left( \frac{1}{2} + \frac{1}{4} \left( e^{\frac{4\pi ik}{n}} + e^{-\frac{4\pi ik}{n}} \right) \right) = \]

\[ = \frac{n}{2} + \frac{1}{2} \sum_{k=0}^{n-1} e^{\frac{4\pi ik}{n}} = \frac{n}{2} + \frac{1}{2} \frac{e^{\frac{4\pi in}{n}} - 1}{e^{\frac{4\pi i}{n}} - 1} = \frac{n}{2}, \]

\[ \sum_{k=0}^{n-1} \sin^2 \left( \frac{2\pi k}{n} \right) = \sum_{k=0}^{n-1} \left( 1 - \cos^2 \left( \frac{2\pi k}{n} \right) \right) = n - \sum_{k=0}^{n-1} \cos^2 \left( \frac{2\pi k}{n} \right) = \frac{n}{2} \]

and

\[ \sum_{k=0}^{n-1} \cos \left( \frac{2\pi k}{n} \right) \sin \left( \frac{2\pi k}{n} \right) = \sum_{k=0}^{n-1} \frac{1}{2i} \left( e^{\frac{2\pi ik}{n}} + e^{-\frac{2\pi ik}{n}} \right) \frac{1}{2i} \left( e^{\frac{2\pi ik}{n}} - e^{-\frac{2\pi ik}{n}} \right) = \]

\[ = \frac{1}{4i} \sum_{k=0}^{n-1} \left( e^{\frac{4\pi ik}{n}} - e^{-\frac{4\pi ik}{n}} \right) = \]

\[ = \frac{1}{4i} \sum_{k=0}^{n-1} \left( e^{\frac{4\pi ik}{n}} \right) - \frac{1}{4i} \sum_{k=0}^{n-1} \left( e^{\frac{-4\pi ik}{n}} \right) = \]

\[ = \frac{1}{4i} \frac{e^{\frac{4\pi in}{n}} - 1}{e^{\frac{4\pi i}{n}} - 1} - \frac{1}{4i} \frac{e^{-\frac{4\pi i}{n}} - 1}{e^{-\frac{4\pi i}{n}} - 1} = 0. \]

Thus we have the following situation:

- For two-bladed rotors the tensor of inertia reads

\[ I = \begin{pmatrix} 2A_0 & 0 & 0 \\ 0 & 2B_0 & 0 \\ 0 & 0 & 2C_0 \end{pmatrix} =: \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix} \]

Since \( C_0 \ll A_0 \) holds, the parameter \( \varepsilon = \frac{A - C}{A} \) lies near 1 and the system is non-autonomous.

- For rotors with more than 2 blades the tensor of inertia reads

\[ I = \begin{pmatrix} \frac{n}{2}(A_0 + C_0) & 0 & 0 \\ 0 & nB_0 & 0 \\ 0 & 0 & \frac{n}{2}(A_0 + C_0) \end{pmatrix} =: \begin{pmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{pmatrix}. \]
Here the parameter \( \varepsilon = \frac{\Delta C}{\Delta} \) vanishes and the system is autonomous. Yet a small \( \varepsilon \) may be introduced to describe a small unbalance.

**Structural Damping**

An important property of canonical systems in the absence of damping. While on the one hand the canonical case is mathematically more interesting and challenging, the dissipative case on the other hand is of more practical importance. In order to deal with all aspects we additionally introduce a *structural damping term*

\[ \delta K_0 \dot{q}, \]

where \( \delta \) is a measure for the strength of the damping. This leads to the following *dissipative model* of the wind turbine:

\[ \frac{d}{dt} \left( (M_0 + \varepsilon M_0) \dot{q} \right) + (G_0 + \delta K_0) \dot{q} + K_0 q = 0. \tag{1.6} \]

### 1.2 Canonical Reduction to Differential Equations of Order 1

In this section we show that our system of differential equations (1.4) is canonical. To this end we start with a general lemma (cf. [29], pp. 111).

**Lemma 1.2.1**

Consider a system of differential equations of the following type:

\[ \frac{d}{dt} \left( M(t) \dot{q} + N(t) q \right) - N^T(t) \dot{q} + P(t) q = 0 \tag{1.7} \]

with the following properties

\[ M = M^T, \quad P = P^T \quad \text{and} \quad \det M \neq 0. \]

Then (1.7) may be written as a canonical system of differential equations of first order

\[ J \dot{z} = Hz, \tag{1.8} \]

or

\[ \dot{z} = Az, \tag{1.9} \]

where

\[ z := \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} q \\ Mq + Nq \end{pmatrix}, \]
1.2. Canonical Reduction to Differential Equations of Order 1

\[ H := \begin{pmatrix} -P - NTM^{-1}N & N^T M^{-1} \\ M^{-1}N & -M^{-1} \end{pmatrix} = H^T \]

and

\[ A := -JH = \begin{pmatrix} M^{-1}N & M^{-1} \\ -P - NTM^{-1}N & N^T M^{-1} \end{pmatrix}. \quad (1.10) \]

**Proof:** On the one hand we have

\[ J\ddot{z} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \dot{q} \\ N^T \dot{q} - Pq \end{pmatrix} = \begin{pmatrix} N^T \dot{q} - Pq \\ -\dot{q} \end{pmatrix}, \]

and on the other

\[ H\dot{z} = \begin{pmatrix} -P - NTM^{-1}N & N^T M^{-1} \\ M^{-1}N & -M^{-1} \end{pmatrix} \begin{pmatrix} q \\ M\dot{q} + Nq \end{pmatrix} = \begin{pmatrix} -Pq + N^T \dot{q} \\ -\dot{q} \end{pmatrix}. \]

Comparing the right-hand sides we see that the claim of the lemma holds. \( \square \)

Finally let

\[ M^I := M_0^I + \varepsilon M_1^I = \begin{pmatrix} m & 0 \\ 0 & B \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \]

\[ M^{II} := M_0^{II} + \varepsilon M_1^{II}(t, \Omega) = \]

\[ = \begin{pmatrix} m & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & A \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 & \frac{\Omega}{2}(-1 + \cos(2\Omega t)) \\ 0 & \frac{\Omega}{2}\sin(2\Omega t) & \frac{\Omega}{2}(-1 - \cos(2\Omega t)) \end{pmatrix}, \]

\[ N^I := N^I(t) := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \]

\[ N^{II} := N^{II}(t) := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \Omega B \\ 0 & 0 & 0 \end{pmatrix}, \]

\[ P^I := P^I(t) := K_0 = \begin{pmatrix} k_1 & -k_2 \\ -k_2 & k_3 \end{pmatrix}, \]
1. A Simple Model for a Wind Turbine

\[ P^{II} := P^{II}(t) := K^{II}_0 = \begin{pmatrix} k_1 & -k_2 & 0 \\ -k_2 & k_3 & 0 \\ 0 & 0 & k_4 \end{pmatrix}. \]

Using (1.10) Eqs. (1.5) may be written as

\[ \dot{z}^I = A^I(t, \varepsilon, \Omega)z^I, \]
\[ \dot{z}^I = A^I(t, \varepsilon, \Omega)z^I. \]

(1.11)

If the structural damping term vanishes, then Lemma 1.2.1 implies that system (1.11) is canonical. Moreover a tedious computation shows that it is also reversible with reversibility matrices

\[ R^I := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad R^{II} := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \]

1.3 Stable and Unstable Wind Turbines

In our model the wind turbine is regarded to be stable for certain values of the parameters, if the top of the tower does not move too violently. This can be made mathematically precise in the following way. The wind turbine is said to be stable, if all solutions of (1.9) are bounded and unstable otherwise.

Thus the parameter plane is subdivided into stable and an unstable regions. There arise two questions:

- Are the boundaries between the stable and the unstable regions smooth curves, at least locally?
- If the answer of the first question is positive, how can we compute these curves efficiently?

These questions are the subject of Part I of this dissertation. The organisation of the chapters is as follows:

**Chapter 2**: For the sake of generality of the subsequent considerations a general framework is introduced.

**Chapter 3**: We introduce a certain normal form for families of matrices.

**Chapter 4**: We derive equations for the stability boundaries for non-autonomous systems and prove that the boundaries are curves, locally.
Chapter 5: We derive equations for the stability boundaries for autonomous systems and prove that the boundaries are curves, locally.

Chapter 6: We compare the normal form approach, introduced in the previous chapters, with an older approach based on standard Lyapunov-Schmidt techniques.

Chapter 7: We develop an algorithm for tracing the stability boundary.

Chapter 8: We apply our results to the wind turbine model.
1. A Simple Model for a Wind Turbine
A General Framework

We consider a two parameter family of differential equations

\[ \dot{x} = A(t, \varepsilon, \Omega)x, \quad (2.1) \]

where

\[ A : \mathbb{R} \times \mathbb{R}^2 \rightarrow L(\mathbb{R}^n), \quad (t, \varepsilon, \Omega) \mapsto A(t, \varepsilon, \Omega) \]

is a smooth matrix function, \( T \)-periodic in \( t \).

If for certain values of the parameters all solutions of (2.1) are bounded, we say: (2.1) is stable for the pair \((\varepsilon, \Omega)\). If for other values of the parameters at least one solution is unbounded, then we say: (2.1) is unstable for the pair \((\varepsilon, \Omega)\). A point \((\varepsilon, \Omega)\) of the parameter plane belongs to the stability boundary, if in every neighborhood of \((\varepsilon, \Omega)\) there are points for which (2.1) is stable as well as points for which (2.1) is unstable.

The goal is to describe conditions such that the stability boundary is a smooth curve, at least locally. Moreover we establish an equation which allows us to compute the stability boundary efficiently.

The problem under consideration also comes up in connection with the restricted problem of three bodies (cf. [4]).

The chapter is organized as follows:

Section 2.1: We discuss conditions for stability under several assumptions.

Section 2.2: We present typical situations for the loss of stability.

2.1 The Condition for Stability

In the autonomous case (2.1) is stable, if and only if (cf. [1], p.183)

(i) \( \text{re} \lambda \leq 0 \) for all eigenvalues \( \lambda \) of \( A(\varepsilon, \Omega) \) and

(ii) every eigenvalue of \( A(\varepsilon, \Omega) \) with \( \text{re} \lambda = 0 \) is semisimple.
We may therefore study the eigenvalues of $A(\varepsilon, \Omega)$.

In the non-autonomous case we consider the principal fundamental matrix solution of (2.1), i.e., the matrix solution $X(t, \varepsilon, \Omega)$ of (2.1) satisfying the initial condition $X(0, \varepsilon, \Omega) = I$. Since $X(t + T, \varepsilon, \Omega) = X(t, \varepsilon, \Omega)X(T, \varepsilon, \Omega)$ it suffices to study the eigenvalues of the so-called monodromy matrix $M(\varepsilon, \Omega) := X(T, \varepsilon, \Omega)$: (2.1) is stable, if and only if (cf. [1], p.311)

(i) $|\mu| < 1$ for all eigenvalues $\mu$ of $(M(\varepsilon, \Omega))$,

(ii) every eigenvalue of $M(\varepsilon, \Omega)$ with $|\mu| = 1$ is semisimple.

If we assume Eq. (2.1) to fulfill additional symmetry conditions, it has some consequences on the spectrum of $A(\varepsilon, \Omega)$ in the autonomous case and of $M(\varepsilon, \Omega)$ in the non-autonomous case, respectively.

**Definition 2.1.1 (canonical systems)**

Let $n$ be even. The $n$-dimensional system of differential equations (2.1) is called **canonical**, if the matrix $A(t, \varepsilon, \Omega)$ is **infinitesimally symplectic**, i.e.

$$A^T(t, \varepsilon, \Omega)J + JA(t, \varepsilon, \Omega) = 0,$$

where $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$.

(2.2)

This symmetry condition has implications on the eigenvalues of the matrix $A$.

**Lemma 2.1.1**

Let (2.1) be canonical and autonomous. Then the following holds:

(i) The spectrum of $A(\varepsilon, \Omega)$ is symmetric with respect to the real and the imaginary axis.

(ii) The characteristic polynomial of $A(\varepsilon, \Omega)$ has real coefficients $a_k$ and is even:

$$P_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

with $a_{2k-1} = 0$ for all $1 \leq k \leq \frac{n}{2}$.

**Proof:**

(i) Since $A(\varepsilon, \Omega)$ is real, the spectrum is symmetric with respect to the real axis. From (2.2) it follows that $A = -J^{-1}A^TJ$ and therefore

$$\sigma(A) = \sigma(-J^{-1}A^TJ) = -\sigma(J^{-1}A^TJ) = -\sigma(A^T) = -\sigma(A).$$

Thus the spectrum of $A$ is also symmetric with respect to the origin. The symmetry with respect to the imaginary axis now follows immediately.
The Condition for Stability

(ii) Let \( P_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 \) be the characteristic polynomial of \( A \). If \( \lambda \) is a solution of \( P_A(\lambda) = 0 \) then \( \lambda, -\lambda \) and \(-\lambda\) are solutions too, as we have seen in (i). Thus \( P_A(\lambda) \) is even and the coefficients are real.

The symmetry condition also has implications on the eigenvalues of the monodromy matrix \( M \).

Lemma 2.1.2

Let (2.1) be canonical. Then the following holds:

(i) The monodromy matrix is symplectic, i.e.

\[
M^T(\varepsilon, \Omega)JM(\varepsilon, \Omega) = J.
\]

(ii) The spectrum of the monodromy matrix \( M(\varepsilon, \Omega) \) is symmetric with respect to the unit circle and to the real axis.

(iii) The characteristic polynomial of \( M(\varepsilon, \Omega) \) has real coefficients and is reflexive (cf. [14]), i.e. it has the form

\[
P_M(\mu) = \mu^n + a_{n-1}\mu^{n-1} + \cdots + a_1\mu + 1
\]

with \( a_{n-k} = a_k \) for all \( 1 \leq k \leq n \).

Proof:

(i) We compute the derivative of \( X(t)^TJX(t) \) with respect to \( t \):

\[
\frac{d}{dt} X^T JX = X^T \dot{J}X + X^T J \dot{X} = (AX)^T J X + X^T JAX =
\]

\[
= X^T A^T JX + X^T JAX = X^T (-JA)X + X^T JAX = 0.
\]

In the penultimate step we used equation (2.2). Therefore \( X(t)^TJX(t) \) is constant. From \( X(0)^TJX(0) = I^TJI = J \) it follows that \( X(t, \varepsilon, \Omega) \) and particularly \( M(\varepsilon, \Omega) \) are symplectic.

(ii) From \( M^TJM = J \) it now follows that \( M^{-1} = J^{-1}M^TJ \). Therefore the spectra of \( M^{-1} \) and \( M^T \) are identical. Of course the same is true for \( M^T \) and \( M \). Thus if \( \mu \) is an eigenvalue of \( M \) so \( \mu^{-1} \) is. Now the claim follows immediately from the symmetry with respect to the real axis.

(iii) We first determine the constant coefficient \( a_0 \). To this end we factorize the characteristic polynomial of \( M \):

\[
P_M(\mu) = \mu^n + a_{n-1}\mu^{n-1} + \cdots + a_1\mu + a_0 =
\]

\[
= (\mu - \mu_1) \cdots (\mu - \mu_n),
\]
where $\mu_1, \ldots, \mu_n$ are the eigenvalues of $M$. Thus we have

$$a_0 = \prod_{i=1}^{n} (-\mu_i) = (-1)^n \det M.$$ 

From the proof of (i) it follows that the determinant of $X(t)$ is equal to 1 for all $t$. Thus we have $\det M = \det X(T) = 1$. Finally since in the canonical case $n$ is even we have $a_0 = 1$.

Now let $\mu$ be a solution of $P_M(\mu) = 0$. Then $\mu^{-1}$ is a solution too, as we have seen in (ii). Therefore

$$0 = P_M(\mu^{-1}) = \mu^{-n} + a_{n-1}\mu^{-(n-1)} + \cdots + a_1\mu^{-1} + 1 =$$

$$= \mu^{-n} \left( 1 + a_{n-1}\mu + \cdots + a_1\mu^{n-1} + \mu^n \right) =$$

$$= \mu^{-n} \hat{P}_M(\mu)$$

Since $\mu \neq 0$ the polynomials $P_M(\mu)$ and $\hat{P}_M(\mu)$ have the same zeros. Thus they are identical. Therefore they have the same coefficients, i.e. $a_k = a_{n-k}$ for all $k$. □

We now discuss a further symmetry condition.

**Definition 2.1.2 (reversible systems)**

The system of differential equations (2.1) is called **reversible**, if

$$A(t, \varepsilon, \Omega)R + RA(-t, \varepsilon, \Omega) = 0,$$

(2.3)

where the so-called **reversibility matrix** $R$ is assumed to be invertible.

In the autonomous case this is simplified to

$$A(\varepsilon, \Omega)R + RA(\varepsilon, \Omega) = 0.$$  (2.4)

In the reversible case we find the following implications on the matrix $A$.

**Lemma 2.1.3**

Let (2.1) be reversible and autonomous. Then the following holds:

(i) The spectrum of $A(\varepsilon, \Omega)$ is symmetric with respect to the real and the imaginary axis.

(ii) If the dimension $n$ of (2.1) is even, then the characteristic polynomial of $A(\varepsilon, \Omega)$ has real coefficients and is even

$$P_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$$

with $a_{2k-1} = 0$ for all $1 \leq k \leq \frac{n}{2}$. If the dimension $n$ of (2.1) is odd, then the characteristic polynomial of $M(\varepsilon, \Omega)$ may be written as
2.1. The Condition for Stability

\[ P_A(\lambda) = \lambda \cdot Q_A(\lambda), \]

where \( Q_A(\lambda) \) has real coefficients and is even:

\[ Q_A(\lambda) = \lambda^n + b_{n-1}\lambda^{n-1} + \cdots + b_1 \lambda + b_0 \]

with \( b_{2k-1} = 0 \) for all \( 1 \leq k \leq \frac{n}{2} \).

Proof:

(i) Since \( A \) is a real matrix, its spectrum is symmetric with respect to the real axis. From (2.4) it follows that \( A = -R^{-1}AR \) and therefore

\[ \sigma(A) = \sigma(-R^{-1}AR) = -\sigma(R^{-1}AR) = -\sigma(A). \]

Thus the spectrum of \( A \) is also symmetric with respect to the origin. Combining these two symmetries we obtain the symmetry with respect to the imaginary axis.

(ii) For \( n \) even we may repeat the proof of Lemma 2.1.1 (iii).

For \( n \) odd we note that \( P_A(\lambda) \) is odd, since the degree is odd and because of the symmetry of the spectrum. Thus \( P_A(\lambda) \) may be written as

\[ P_A(\lambda) = \lambda Q_a(\lambda), \]

where \( Q_A(\lambda) \) has real coefficients and is even.

For the monodromy matrix \( M \) we find the following properties.

Lemma 2.1.4

Let (2.1) be reversible. Then the following holds:

(i) The monodromy matrix (as a map) is reversible:

\[ M(\varepsilon, \Omega)RM(\varepsilon, \Omega) = R. \]

(ii) The spectrum of the monodromy matrix \( M(\varepsilon, \Omega) \) is symmetric with respect to the unit circle and to the real axis.

(iii) If the dimension \( n \) of (2.1) is even, then the characteristic polynomial of \( M(\varepsilon, \Omega) \) has real coefficients and is reflexive:

\[ P_M(\mu) = \mu^n + a_{n-1}\mu^{n-1} + \cdots + a_1 \mu + 1 \]

with \( a_{n-k} = a_k \) for all \( k \).

If the dimension \( n \) of (2.1) is odd, then the characteristic polynomial of \( M(\varepsilon, \Omega) \) may be written as

\[ P_M(\mu) = (\mu - 1)Q_M(\mu), \]

where \( Q_M(\mu) \) has real coefficients and is reflexive:
\[ Q_M(\mu) = \mu^n + b_{n-1}\mu^{n-1} + \cdots + b_1\mu + 1 \]

with \( b_{n-k} = -b_k \) for all \( k \).

**Proof:**

(i) We put \( Y(t) := R^{-1}X(-t)R \). Then we obtain

\[
\dot{Y} = \frac{d}{dt} R^{-1}X(-t)R = -R^{-1}\dot{X}(-t)R = -R^{-1}A(-t)X(-t)R = A(t)R^{-1}X(-t)R = A(t)Y(t).
\]

We used equation (2.3). With \( Y(0) = R^{-1}X(0)R = R^{-1}R = I \) it follows from the uniqueness of the solutions that \( R^{-1}X(-t)R = X(t) \). It follows that \( M = X(T) \) is reversible.

(ii) From the reversibility we conclude that \( M \) and \( M^{-1} \) have the same spectrum: if \( \mu \) is an eigenvalue so \( \mu^{-1} \) is an eigenvalue too. Since the spectrum of a real matrix is also symmetric with respect to the real axis the claim follows immediately.

(iii) We first determine the constant coefficient \( a_0 \). To this end we factorize the characteristic polynomial of \( M \):

\[ P_M(\mu) = \mu^n + a_{n-1}\mu^{n-1} + \cdots + a_1\mu + a_0 = (\mu - \mu_1) \cdots (\mu - \mu_n), \]

where \( \mu_1, \ldots, \mu_n \) are the eigenvalues of \( M \). Thus we have

\[ a_0 = \prod_{i=1}^{n} (-\mu_i) = (-1)^n \det M. \]

The Theorem of Liouville (cf. [1], p. 154) implies that

\[ \det M = \det X(T) = \det X(0)e^{\int_0^T \text{tr} A(s) ds} = e^{\int_0^T \text{tr} A(s) ds}. \]

Using the periodicity and the reversibility of \( A \) we obtain for the integral:

\[
\int_0^T \text{tr} A(s) ds = \int_{-T/2}^{T/2} \text{tr} A(s) ds = \int_{-T/2}^0 \text{tr} A(s) ds + \int_0^{T/2} \text{tr} A(s) ds = \int_0^{T/2} \text{tr} A(-s) ds + \int_0^{T/2} \text{tr} A(s) ds = \int_0^{T/2} (\text{tr}(-R^{-1}A(s)R) + \text{tr} A(s)) ds = \int_0^{T/2} (-\text{tr} A(s) + \text{tr} A(s)) ds = 0.
\]
We conclude that $\det M = 1$ and therefore $a_0 = (-1)^n$.

Let $n$ be even. If $\mu$ is a solution of $P_M(\mu) = 0$, then $\mu^{-1}$ is a solution too (cf. (ii)). Therefore we have

$$0 = P_M(\mu^{-1}) = \mu^{-n} + a_{n-1}\mu^{-(n-1)} + \cdots + a_1\mu^{-1} - 1 =
= \mu^{-n} (1 + a_{n-1}\mu + \cdots + a_1\mu^{n-1} + \mu^n) =
= \mu^{-n} \hat{P}_M(\mu).$$

Since $\mu \neq 0$ the polynomials $P_M(\mu)$ and $\hat{P}_M(\mu)$ have the same zeros. Thus they are identical. Therefore they have the same coefficients, i.e.

$$a_k = a_{n-k}, \quad \text{for} \quad 0 \leq k \leq n.$$

Let $n$ be odd. If $\mu$ is a solution of $P_M(\mu) = 0$, then $\mu_1$ is a solution too (cf. (ii)). Therefore we have

$$0 = P_M(\mu^{-1}) = \mu^{-n} + a_{n-1}\mu^{-(n-1)} + \cdots + a_1\mu^{-1} - 1 =
= -\mu^{-n} (-1 - a_{n-1}\mu - \cdots - a_1\mu^{n-1} + \mu^n) =
=: -\mu^{-n} \hat{P}_M(\mu).$$

Since $\mu \neq 0$ the polynomials $P_M(\mu)$ and $\hat{P}_M(\mu)$ have the same zeros. Thus they are identical. Therefore we have

$$a_k = -a_{n-k}, \quad \text{for} \quad 0 \leq k \leq n. \quad (2.5)$$

Let $n = 2m + 1$. Then

$$P_M(\mu) = \mu^{2m+1} + a_{2m}\mu^{2m} + \cdots + a_{m+1}\mu^{m+1} -
- (a_{m+1}\mu^m + \cdots + a_{2m}\mu + 1).$$

Obviously $P_M(1) = 0$ and $P_M(\mu)$ may be written as

$$P_M(\mu) = (\mu - 1)Q_M(\mu).$$

Let

$$Q_M(\mu) := b_{2m}\mu^{2m} + b_{2m-1}\mu^{2m-1} + \cdots + b_1\mu + b_0.$$

Then

$$(\mu - 1)Q_M(\mu) = b_{2m}\mu^{2m+1} + b_{2m-1}\mu^{2m} + \cdots + b_1\mu^2 + b_0\mu -
- b_{2m}\mu^{2m} - \cdots - b_2\mu^2 - b_1\mu - b_0 =
= b_{2m}\mu^{2m+1} + (b_{2m-1} - b_{2m})\mu^{2m} + \cdots + (b_0 - b_1)\mu - b_0.$$
A comparison with $P_M(\mu)$ leads to

$$a_{2m+1} = b_{2m},$$

$$a_k = b_{k-1} - b_k \quad \text{for} \quad 1 \leq k \leq 2m \quad (2.6)$$

$$a_0 = -b_0.$$

It remains to show that $b_{2m-k} = b_k$:

$k = 0 : \quad b_{2m} = a_{2m+1} = -a_0 = b_0.$

$k - 1 \Rightarrow k : \quad$ Suppose that $b_{2m-(k-1)} = b_{k-1}$. Then we have

$$b_{2m-k} = b_{(2m+1)-(k+1)} = a_{(2m+1)-k} + b_{(2m+1)-k} =$$

$$= a_{(2m+1)-k} + b_{2m-(k-1)} = -a_k + b_{k-1} = b_k. \quad \square$$

Thus if (2.1) is either canonical or reversible we have the following conditions for stability:

In the autonomous case

- (2.1) is stable, if and only if all the eigenvalues of $A(\varepsilon, \Omega)$ are on the imaginary axis and are semisimple,

- a simple eigenvalue cannot leave the imaginary axis under a change of $(\varepsilon, \Omega)$, since there had to appear a second symmetric eigenvalue “out of nothing”.

In the non-autonomous case

- (2.1) is stable, if and only if all the eigenvalues of $M(\varepsilon, \Omega)$ are on the unit circle and are semisimple,

- a simple eigenvalue cannot leave the unit circle under a change of $(\varepsilon, \Omega)$, since there had to appear a second symmetric eigenvalue “out of nothing”.

### 2.2 The Transition Stability–Instability

In this section we consider the generic situations for a transition from stability to instability. Table 2.1 gives a survey for the dissipative case.

In the canonical or reversible case the additional symmetry of the spectrum leads to more complicated situations compared to the dissipative case as shown in Table 2.2

To compute the eigenvalues of $A(\varepsilon, \Omega)$ or $M(\varepsilon, \Omega)$, respectively, as functions of $(\varepsilon, \Omega)$, we first transform the matrix to a suitable normal form.
2.2. The Transition Stability–Instability

**Tab. 2.1:** A survey on the typical situations for the loss of stability in the dissipative case

<table>
<thead>
<tr>
<th>autonomous systems</th>
<th>non-autonomous systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>eigenvalues of $A(\varepsilon_0, \Omega_0)$</td>
<td>eigenvalues of $M(\varepsilon_0, \Omega_0)$</td>
</tr>
<tr>
<td><img src="image" alt="Diagram of simple real eigenvalue leaving the left half plane." /></td>
<td><img src="image" alt="Diagram of simple real eigenvalue leaving the unit circle at -1 or +1, respectively." /></td>
</tr>
<tr>
<td>A simple real eigenvalue leaves the left half plane.</td>
<td>A simple real eigenvalue leaves the unit circle at -1 or +1, respectively.</td>
</tr>
<tr>
<td><img src="image" alt="Diagram of complex conjugate eigenvalue leaving the left half plane." /></td>
<td><img src="image" alt="Diagram of complex conjugate eigenvalue leaving the unit circle." /></td>
</tr>
<tr>
<td>A pair of complex conjugate eigenvalues leaves the left half plane.</td>
<td>A pair of complex conjugate eigenvalues leaves the unit circle.</td>
</tr>
</tbody>
</table>

**Tab. 2.2:** A survey on the typical situations for the loss of stability in the canonical or reversible case

<table>
<thead>
<tr>
<th>autonomous systems</th>
<th>non-autonomous systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>eigenvalues of $A(\varepsilon_0, \Omega_0)$</td>
<td>eigenvalues of $M(\varepsilon_0, \Omega_0)$</td>
</tr>
<tr>
<td><img src="image" alt="Diagram of complex conjugate eigenvalue coalescing on the imaginary axis and leaving on the real axis." /></td>
<td><img src="image" alt="Diagram of complex conjugate eigenvalue coalescing on the unit circle at ±1 and leaving on the real axis." /></td>
</tr>
<tr>
<td>A pair of complex conjugate eigenvalues on the imaginary axis coalesce and leaves on the real axis.</td>
<td>A pair of complex conjugate eigenvalues on the unit cycle coalesce at ±1 and leaves on the real axis.</td>
</tr>
<tr>
<td><img src="image" alt="Diagram of complex conjugate eigenvalue coalescing on the imaginary axis and leaving the imaginary axis." /></td>
<td><img src="image" alt="Diagram of complex conjugate eigenvalue coalescing on the unit circle (Krein collision)." /></td>
</tr>
<tr>
<td>Two pairs of complex conjugate eigenvalues on the imaginary axis coalesce and leave the imaginary axis.</td>
<td>Two pairs of complex conjugate eigenvalues on the imaginary axis coalesce and leave the unit circle (Krein collision).</td>
</tr>
</tbody>
</table>
3

A Normal Form for Families of Matrices

In this chapter we derive a normal form for a family of matrices. The chapter is organized as follows:

**Section 3.1:** We introduce a certain similarity transformation that transforms a given family of matrices to block triangular form.

**Section 3.2:** We discuss the properties of the spectra of the transformed matrices.

3.1 Block Triangular Form for Families of Matrices

Consider a family of matrices $A$ depending on a parameter $\sigma$. Suppose that $A(\sigma^0)$ is in block triangular form. We show that it is possible to transform $A(\sigma)$ to block triangular form for $\sigma$ close enough to $\sigma^0$ by a similarity transformation depending on $\sigma$.

**Lemma 3.1.1**

*Let $A : \mathbb{R}^s \to L(\mathbb{C}^n), \sigma \mapsto A(\sigma)$ be a $C^r$-family of matrices ($r > 2$). Assume $A^0 := A(\sigma^0)$ to have the block form

$$
A^0 = \begin{pmatrix}
A_{11} & A_{12}^0 \\
0 & A_{22}^0
\end{pmatrix}
$$

with $A_{11}^0 \in L(\mathbb{C}^k), A_{12}^0 \in L(\mathbb{C}^l, \mathbb{C}^k), A_{22}^0 \in L(\mathbb{C}^l)$ and $0 \in L(\mathbb{C}^k, \mathbb{C}^l)$, where $k + l = n$. Assume further that the spectra of $A_{11}^0$ and $A_{22}^0$ are disjoint.*

*Then there exists a neighborhood $U$ of $\sigma^0$ and a $C^r$-family of matrices

$$
\Phi : U \subset \mathbb{R}^s \to GL(\mathbb{C}^n), \sigma \mapsto \Phi(\sigma),
$$

such that*
3. A Normal Form for Families of Matrices

\( \Phi(\sigma^0) = I_n, \)
\( \tilde{A}(\sigma) := \Phi^{-1}(\sigma) A(\sigma) \Phi(\sigma) = \begin{pmatrix} \tilde{A}_{11}(\sigma) & \tilde{A}_{12}(\sigma) \\ 0 & \tilde{A}_{22}(\sigma) \end{pmatrix} \quad \text{for } \sigma \in U. \)

**Proof:** Put
\[
A(\sigma) = \begin{pmatrix} A_{11}(\sigma) & A_{12}(\sigma) \\ A_{21}(\sigma) & A_{22}(\sigma) \end{pmatrix} \quad \text{and} \quad \Phi(\sigma) = \begin{pmatrix} I_k & 0 \\ \phi(\sigma) & I_l \end{pmatrix}
\]
with \( \phi(\sigma) \in L(\mathbb{C}^k, \mathbb{C}^l). \) Then we have \( \Phi^{-1}(\sigma) = \begin{pmatrix} I_k & 0 \\ -\phi(\sigma) & I_l \end{pmatrix} \) and we obtain
\[
\Phi^{-1} A \Phi = \begin{pmatrix} I_k & 0 \\ -\phi & I_l \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} I_k & 0 \\ \phi & I_l \end{pmatrix} = \begin{pmatrix} A_{11} + A_{12} \phi & A_{12} \\ A_{21} + A_{22} \phi & A_{22} \end{pmatrix} = \begin{pmatrix} A_{11} + A_{12} \phi & A_{12} \\ -\phi A_{11} + \phi A_{12} \phi + A_{21} + A_{22} \phi & -\phi A_{12} + A_{22} \end{pmatrix}.
\]
Therefore \( \phi \) must solve the equation
\[
A_{21} - \phi A_{11} + A_{22} \phi - \phi A_{12} \phi = 0. \quad (3.2)
\]

We define the map
\[
F: L(\mathbb{C}^k, \mathbb{C}^l) \times \mathbb{R}^n \rightarrow L(\mathbb{C}^k, \mathbb{C}^l)
\]
\[
\phi \mapsto F(\phi, \sigma) := A_{21} - \phi A_{11} + A_{22} \phi - \phi A_{12} \phi.
\]
Now we show that \( F(\phi, \sigma) = 0 \) has a solution for every \( \sigma \) sufficiently close to \( \sigma_0. \) To prove this we use the Implicit Function Theorem (e.g. cf. [6, p.253ff]). Clearly \( F(0, \sigma_0) = 0: \)
\[
F(0, \sigma_0) = A_{21}^0 - 0 \cdot A_{11}^0 + A_{22}^0 \cdot 0 - 0 \cdot A_{12}^0 \cdot 0 = A_{21}^0 = 0.
\]
The derivative of \( F \) with respect to \( \phi \) at \( (0, \sigma_0) \) reads \( D_1 F(0, \sigma_0) \psi = -\psi A_{11}^0 + A_{22}^0 \psi: \)
\[
\lim_{\|\Delta \psi\| \rightarrow 0} \frac{\|F(0 + \Delta \psi, \sigma_0) - F(0, \sigma_0) - D_1 F(0, \sigma_0) \Delta \psi\|}{\|\Delta \psi\|} = \lim_{\|\Delta \psi\| \rightarrow 0} \frac{\||-\Delta \psi A_{11}^0 + A_{22}^0 \Delta \psi - \Delta \psi A_{12}^0 \Delta \psi - D_1 F(0, \sigma_0) \Delta \psi\|}{\|\Delta \psi\|} = \leq \lim_{\|\Delta \psi\| \rightarrow 0} \frac{\||-\Delta \psi A_{11}^0 + A_{22}^0 \Delta \psi - D_1 F(0, \sigma_0) \Delta \psi\|}{\|\Delta \psi\|} + \lim_{\|\Delta \psi\| \rightarrow 0} \frac{\||-\Delta \psi A_{12}^0 \Delta \psi\|}{\|\Delta \psi\|} = 0.
\]
This proves the claim.
3.1. Block Triangular Form for Families of Matrices

It remains to show that the map

\[ L^0 : L(\mathbb{C}^k, \mathbb{C}^l) \to L(\mathbb{C}^k, \mathbb{C}^l), \phi \mapsto L^0(\phi) := -\phi A_{11}^0 + A_{22}^0 \phi \]

is invertible. First we have

\[
L^0(\phi) = -\phi A_{11}^0 + A_{22}^0 \phi = \begin{pmatrix}
\phi_{11} & \ldots & \phi_{1k} \\
\vdots & \ddots & \vdots \\
\phi_{l1} & \ldots & \phi_{lk}
\end{pmatrix}
\begin{pmatrix}
A_{11,11}^0 & \ldots & A_{11,1k}^0 \\
\vdots & \ddots & \vdots \\
A_{11,k1}^0 & \ldots & A_{11,kk}^0
\end{pmatrix}
\begin{pmatrix}
\phi_{11} \\
\vdots \\
\phi_{lk}
\end{pmatrix}
\]

\[
+ \begin{pmatrix}
A_{22,11}^0 & \ldots & A_{22,1l}^0 \\
\vdots & \ddots & \vdots \\
A_{22,l1}^0 & \ldots & A_{22,ll}^0
\end{pmatrix}
\begin{pmatrix}
\phi_{11} & \phi_{1k} \\
\vdots & \ddots \\
\phi_{l1} & \phi_{lk}
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\sum_{j=1}^k \phi_{1j} A_{11,1j}^0 & \ldots & \sum_{j=1}^k \phi_{1j} A_{11,jk}^0 \\
\vdots & \ddots & \vdots \\
\sum_{j=1}^k \phi_{lj} A_{11,1j}^0 & \ldots & \sum_{j=1}^k \phi_{lj} A_{11,jk}^0
\end{pmatrix}
\begin{pmatrix}
\phi_{11} \\
\vdots \\
\phi_{lk}
\end{pmatrix}
\]

\[
+ \begin{pmatrix}
\sum_{j=1}^l \phi_{j1} A_{22,1j}^0 & \ldots & \sum_{j=1}^l \phi_{j1} A_{22,jk}^0 \\
\vdots & \ddots & \vdots \\
\sum_{j=1}^l \phi_{lj} A_{22,1j}^0 & \ldots & \sum_{j=1}^l \phi_{lj} A_{22,jk}^0
\end{pmatrix}
\begin{pmatrix}
\phi_{11} & \phi_{1k} \\
\vdots & \ddots \\
\phi_{l1} & \phi_{lk}
\end{pmatrix}
\]

Now we identify \( L(\mathbb{C}^k, \mathbb{C}^l) \) with \( \mathbb{C}^{kl} \). So we are able to represent \( L^0 \) as a \( kl \times kl \)-matrix \( \tilde{L}^0 \):

\[
\tilde{L}^0 \phi = -\begin{pmatrix}
A_{11,11}^0 & \ldots & 0 & A_{11,k1}^0 & \ldots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \ldots & A_{11,11}^0 & 0 & \ldots & A_{11,k1}^0 \\
\vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
0 & \ldots & A_{11,k1}^0 & 0 & \ldots & A_{11,kk}^0
\end{pmatrix}
\begin{pmatrix}
\phi_{11} \\
\vdots \\
\phi_{lk}
\end{pmatrix}
\]

\[
+ \begin{pmatrix}
A_{22,11}^0 & \ldots & A_{22,1l}^0 & 0 & \ldots & 0 \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & \ldots & 0 & A_{22,11}^0 & \ldots & A_{22,1l}^0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & A_{22,l1}^0 & \ldots & A_{22,ll}^0
\end{pmatrix}
\begin{pmatrix}
\phi_{11} \\
\vdots \\
\phi_{lk}
\end{pmatrix}
\]
Without loss of generality we may assume that $A_0$ and therefore $A_{11}$ and $A_{22}$ are in Jordan normal form. Then we have

$$
\begin{pmatrix}
-A_1 + A_2 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & -A_{11,1} + A_{11,2} + I_k \otimes A_{22}
\end{pmatrix}
$$

where $\lambda_{11,i}^0$ is the $i$-th eigenvalue of $A_{11}^0$ and the asterisks stand for the zero matrix $0_t$ or the identity matrix $I_t$. Therefore we can compute the determinant of $L^0$:

$$
det L^0 = \prod_{i=0}^{k} \det(-\lambda_{11,i}^0 I_t + A_{22}^0).
$$

Since the spectra of $A_{11}^0$ and $A_{22}^0$ are disjoint, we have

$$
det(-\lambda_{11,i}^0 I_t + A_{22}^0) \neq 0.
$$

Thus $L^0$ is indeed invertible.

Now the Implicit Function Theorem guarantees the existence of a neighborhood $U$ of $\sigma^0$ and a smooth map $\phi : U \rightarrow L(\mathbb{C}^t, \mathbb{C}^k)$ such that $\phi(\sigma^0) = \sigma$ and

$$
F(\phi(\sigma), \sigma) = 0, \quad \forall \sigma \in U.
$$

This proves the lemma. □

This lemma immediately generalizes to the following result which is basic for the subsequent considerations.

**Theorem 3.1.1**

Let $A : \mathbb{R}^d \rightarrow A(\sigma)$ be a $C^r$-family of matrices. Assume $A^0 := A(\sigma^0)$ to have the block form

$$
A^0 = \begin{pmatrix}
A_{11}^0 & \cdots & A_{1m}^0 \\
0 & \ddots & \vdots \\
\vdots & \ddots & \ddots \\
0 & \cdots & 0 & A_{mm}^0
\end{pmatrix},
$$

where $A_{ii}^0 \in L(\mathbb{C}^{d_i}), A_{ij}^0 \in L(\mathbb{C}^{d_i}, \mathbb{C}^{d_j}), \sum_{i=1}^{m} d_i = n$, and assume that the $A_{ii}^0$ have pairwise disjoint spectra.
Then there exists a neighborhood $U$ of $a^0$ and a $C^r$-family of matrices

$$\Phi : U \subset \mathbb{R}^n \rightarrow GL(\mathbb{C}^n), \sigma \mapsto \Phi(\sigma)$$

such that

(i) $\Phi(a^0) = I_n$,

(ii) $\Phi^{-1}(\sigma) A(\sigma) \Phi(\sigma) = \begin{pmatrix} A_{11}(\sigma) & \ldots & A_{1m}(\sigma) \\ 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots \\ 0 & \ldots & A_{mm}(\sigma) \end{pmatrix}$ for $\sigma \in U$.

**Proof:** First we write $A^0 \in L(\mathbb{C}^n)$ as

$$A^0 = \begin{pmatrix} A_{11}^0 & B_{11}^0 \\ 0 & C_{11}^0 \end{pmatrix}.$$ 

With Lemma 3.1.1 there is a $:\mathbb{R}^6 \rightarrow GL(\mathbb{C}^n)$ with

$$\Phi^{-1}(\sigma) A(\sigma) \Phi(\sigma) = \begin{pmatrix} A_{11}(\sigma) & B_{11}(\sigma) \\ 0 & C_{11}(\sigma) \end{pmatrix}.$$ 

Now we write $C_{11}^0 \in L(\mathbb{C}^{n-d_1})$ as

$$C_{11}^0 = \begin{pmatrix} A_{22}^0 & B_{22}^0 \\ 0 & C_{22}^0 \end{pmatrix}.$$ 

Again with Lemma 3.1.1 there exists a $:\mathbb{R}^6 \rightarrow GL(\mathbb{C}^{n-d_1})$ with

$$\Phi_2^{-1}(\sigma) C_{11}(\sigma) \Phi_2(\sigma) = \begin{pmatrix} A_{22}(\sigma) & B_{22}(\sigma) \\ 0 & C_{22}(\sigma) \end{pmatrix}.$$ 

In this way we go on until we reach the last block. Finally we define

$$\Phi_2 := \begin{pmatrix} I_{d_1} & 0 \\ 0 & \Phi_2 \end{pmatrix}, \quad \Phi_3 := \begin{pmatrix} I_{d_2} & 0 \\ 0 & I_{d_2} & 0 \end{pmatrix}, \quad \ldots.$$ 

Then: $\Phi(\sigma) := \Phi_{m-1}(\sigma) \cdot \ldots \cdot \Phi_1(\sigma)$ is the desired transformation matrix.
3.2 Properties of the Triangularized Matrices

The transformation to block triangular form introduced in the last section does not preserve neither symplecticity nor reversibility with respect to the operator $R$. But it preserves the symmetries of the spectrum of the matrices.

**Lemma 3.2.1**

Let the assumptions of Theorem 3.1.1 hold.

(i) If the spectrum of $A(\sigma)$ for all $\sigma \in U$ is symmetric with respect to the real axis or the imaginary axis or the unit circle then the same holds true for the spectrum of $
abla(\sigma) := \Phi^{-1}(\sigma)A(\sigma)\Phi(\sigma)$.

(ii) Moreover if the spectra of $A_{ii}(\sigma_0)$ are symmetric with respect to the real axis or the imaginary axis or the unit circle, then the same hold true for $\nabla_{ii}(\sigma)$ for $\sigma$ sufficiently close to $\sigma_0$.

**Proof:**

(i) A similarity transformation does not change the spectrum of a matrix. Thus the spectra of $A(\sigma)$ and $\nabla(\sigma)$ are identical.

(ii) Since the spectra of $A_{11}^0, \ldots, A_{nm}^0$ are pairwise disjoint there exists a neighborhood $U(\sigma^0)$ of $\sigma^0$ such that the the spectra of $\nabla_{11}(\sigma), \ldots, \nabla_{mm}(\sigma)$ are also pairwise disjoint for $\sigma \in U(\sigma^0)$.

Let the spectra of $A_{11}^0, \ldots, A_{mm}^0$ be symmetric with respect to the real axis, say. Thus the eigenvalues of $A_{ii}^0$ occur in pairs $\mu^0, \overline{\mu}^0$. Now assume that there exists a $\sigma^1 \in U(\sigma^0)$ such that the spectrum of $\nabla_{ii}(\sigma^1)$ is not symmetric with respect to the real axis.

Let $\mu^1$ be an eigenvalue of $\nabla_{ii}(\sigma^1)$ such that $\overline{\mu}^1$ is not an eigenvalue of $\nabla_{ii}(\sigma^1)$. Since the spectrum of $\nabla(\sigma^1)$ is symmetric with respect to the real axis $\overline{\mu}^1$ is an eigenvalue of $\nabla(\sigma^1)$ and therefore of some matrix $\nabla_{jj}(\sigma) (j \neq i)$. But this is not possible since the eigenvalues depend continuously on the parameter $\sigma$ and the spectra of $\nabla_{ii}(\sigma)$ and $\nabla_{jj}(\sigma)$ are disjoint. By this contradiction we conclude that the spectrum of $\nabla_{ii}(\sigma^1)$ is symmetric with respect to the real axis. The other symmetries are proved in the same way. □
4

Non-Autonomous Systems

In this chapter we derive equations for the stability boundaries of the differential equation

$$\dot{x} = A(t, \varepsilon, \Omega)x.$$ \hspace{1cm} (2.1)

and prove that the stability boundaries are curves, locally. As we already saw in the introduction the symmetry properties of (2.1) influence the way stability is lost. Therefore we organize this chapter according to the generic situations, considered in Section 2.2.

Section 4.1: In the first section we provide some auxilliary results.

Section 4.2: In the second section we discuss the stability of dissipative systems.

Section 4.3: In the third section we discuss the stability of canonical or reversible systems.

4.1 Auxilliary Results

In Section 2.1 we saw that the spectrum of the monodromy matrix $M(\varepsilon, \Omega)$ of (2.1) determines the stability.

Applying the normal form transformation introduced in the previous chapter to the monodromy matrix we obtain a matrix in a block triangular form. That allows us to work with the spectra of the submatrices in the main diagonal. This reduces the degree of the characteristic equations to 1 or 2 in the dissipative case and to 2 or 4 in the canonical or reversible case.

In the latter case the spectrum of the monodromy matrix is symmetric with respect to the real axis and to the unit circle (cf. Lemma 2.1.2 and Lemma 2.1.4). Due to Lemma 3.2.1 the spectra of the submatrices in the main diagonal show the same symmetries. These symmetries may be used for a further reduction of the degree of the characteristic equations using an idea presented by Brouke (cf. [4] and [14]).
We start with the two dimensional case. Let $M$ be a $2 \times 2$-matrix with a reflexive characteristic polynomial

$$P_c(\mu) := \mu^2 - b\mu + 1$$

with real coefficient $b$ (cf. Lemma 2.1.2).

Since the zeroes $\mu$ of $P_c(\mu)$ do not vanish, we may multiply $P_c(\mu)$ by $\mu^{-1}$:

$$\mu^{-1}P_c = \mu^{-1}(\mu^2 - b\mu + 1) = \mu - b + \mu^{-1} = (\mu + \mu^{-1}) - b.$$

If we put $\rho := \mu + \mu^{-1}$, then we obtain the so-called reduced characteristic polynomial:

$$Q_c(\rho) := \rho - b. \quad (4.1)$$

Its zeroes are called reduced eigenvalues of $M$. The next lemma states some properties of reduced eigenvalues.

**Lemma 4.1.1**

Let $M$ be a $2 \times 2$-matrix whose spectrum is symmetric with respect to the real axis and to the unit circle. Then the following statements hold:

(i) $\mu$ is an eigenvalue of $M$ if and only if $\rho$ is a reduced eigenvalue.

(ii) An eigenvalue $\mu$ of $M$ lies on the unit circle if and only if the reduced eigenvalue $\rho$ lies in the interval $[-2, 2]$.

(iii) If $\mu = \pm 1$ is a double eigenvalue of $M$, then $\rho = \pm 2$ is a simple reduced eigenvalue of $M$.

**Proof:**

(i) The first claim follows immediately from the definition of the reduced characteristic polynomial.

(ii) On the one hand if $\mu$ lies on the unit circle, then we have

$$\rho = \mu + \mu^{-1} = \mu + \bar{\mu} = 2 \text{ re } \mu \in [-2, 2].$$

On the other hand solving $\rho = \mu + \mu^{-1}$ for $\mu$ we find

$$\mu^\pm = \frac{\rho}{2} \pm \sqrt{\frac{\rho^2}{4} - 1}.$$

Then $\rho \in [-2, 2]$ implies that $\mu^\pm$ lie on the unit circle.

(iii) Let $\pm 1$ be a double eigenvalue of $M$. Then the characteristic polynomial reads

$$P_c(\mu) = (\mu \mp 1)^2 = \mu^2 \mp 2\mu + 1$$

and the reduced characteristic polynomial reads

$$Q_c(\rho) = \rho \mp 2.$$

Thus $\pm 2$ is a simple reduce eigenvalue of $M$. \qed
4.1. Auxiliary Results

In the next lemma we show how to compute the coefficients of the characteristic equation.

**Lemma 4.1.2**

Let $M$ be a $2 \times 2$-matrix. Then for coefficients $b$ and $c$ of the characteristic polynomial

$$P_c(\beta) := \beta^2 - \beta b + c$$

the following holds:

$$b = \text{tr } M,$$

$$c = \det M = \frac{1}{2} \left( \text{tr}^2 M - \text{tr}(M^2) \right).$$

**Proof:** Let $\mu_1$ and $\mu_2$ be the eigenvalues of $M$. Then we obtain for the characteristic polynomial

$$P_c(\mu) = (\mu - \mu_1)(\mu - \mu_2) = \mu^2 - (\mu_1 + \mu_2)\mu + \mu_1\mu_2.$$  

We immediately see that $b = \text{tr } M$ and $c = \det M$.

Let

$$M := \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}.$$  

Then

$$\text{tr}^2 M - \text{tr}(M^2) = \text{tr}^2 \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} - \text{tr} \begin{pmatrix} m_{11}^2 + m_{12}m_{21} & m_{11}m_{12} + m_{12}m_{22} \\ m_{11}m_{21} + m_{21}m_{22} & m_{12}m_{21} + m_{22}^2 \end{pmatrix} =$$

$$= (m_{11} + m_{22})^2 - (m_{11}^2 + m_{12}m_{21} + m_{12}m_{21} + m_{22}^2) =$$

$$= 2(m_{11}m_{22} - m_{12}m_{21}) = 2\det M.$$

This completes the proof. \qed

We now come to the four dimensional case. Consider a $4 \times 4$-matrix with a reflexive characteristic polynomial

$$P_c(\mu) := \mu^4 - B\mu^3 + C\mu^2 - B\mu + 1$$

with real coefficients $B$ and $C$ (cf. Lemma 2.1.2).

As in the two dimensional case the eigenvalues do not vanish. We may therefore multiply the characteristic polynomial $P_c(\mu)$ by $\mu^{-2}$:

$$\mu^{-2}P_c(\mu) = \mu^{-2}(\mu^4 - B\mu^3 + C\mu^2 - B\mu + 1) =$$

$$= \mu^2 - B\mu + C - B\mu^{-1} + \mu^{-2} = (\mu + \mu^{-1})^2 - B \cdot (\mu + \mu^{-1}) + (C - 2).$$
Again we put $\rho := \mu + \mu^{-1}$ and obtain the so-called reduced characteristic polynomial:

$$Q_c(\rho) := \rho^2 - B \rho + (C - 2). \quad (4.2)$$

The next lemma gives some results for the reduced eigenvalues for the case of a $4 \times 4$-matrix.

**Lemma 4.1.3**

Let $M$ be a $4 \times 4$-matrix whose spectrum is symmetric with respect to the real axis and to the unit circle. Then the following statements hold:

(i) $\mu$ is an eigenvalue of $M$ if and only if $\rho$ is a reduced eigenvalue.

(ii) An eigenvalue $\mu$ of $M$ lies on the unit circle if and only if the reduced eigenvalue $\rho$ lies in the interval $[-2, 2]$.

(iii) If $\mu$ is a double eigenvalue of $M$ different from $\pm 1$, then $\rho$ is a double reduced eigenvalue of $M$.

(iv) If $\mu = \pm 1$ is a double eigenvalue of $M$, then $\rho = \pm 2$ is a simple eigenvalue of $M$.

**Proof:** The proofs of (i) and (ii) are identical to those of Lemma 4.1.1. It remains to prove (iii) and (iv).

(iii) Let $\mu_1 \neq \pm 1$ be a double zero of $P_2$. Then $|\mu_1| = 1$ and $\bar{\mu}_1 = \mu_1^{-1}$ is also a double eigenvalue. Thus we have

$$\mu^{-2} P_2(\mu) = \mu^{-2} (\mu^4 - \mu^3 B + \mu^2 C - \mu B + 1) = \mu^{-2} (\mu - \mu_1)^2 (\mu - \mu_1^{-1})^2 =$$

$$= \mu^{-2} (\mu^2 - \mu (\mu_1 + \mu^{-1}) + 1)^2 = (\mu - (\mu_1 + \mu_1^{-1} + \mu^{-1})^2$$

and therefore

$$Q_c(\rho) = (\rho - (\mu_1 + \mu_1^{-1}))^2.$$ 

We conclude that $\rho_1 := \mu_1 + \mu_1^{-1}$ is a double reduced eigenvalue.

(iv) If $\pm 1$ is a double eigenvalue, then there exists a further pair of eigenvalues $\mu_{1,2}$ with $\mu_{1,2} \neq \pm 1$ and $\mu_1 \mu_2 = 1$. Thus we have

$$\mu^{-2} P_2(\mu) = \mu^{-2} (\mu \mp 1)^2 (\mu^2 - \mu (\mu_1 + \mu_2) + \mu_1 \mu_2) =$$

$$= \mu^{-2} (\mu \mp 1)^2 (\mu^2 - \mu (\mu_1 + \mu_2) + 1) =$$

$$= (\mu \mp 2 + \mu^{-1}) ((\mu - (\mu_1 + \mu_2) + \mu^{-1})$$

and therefore

$$Q_c(\rho) = (\rho \mp 2)(\rho - (\mu_1 + \mu_2)).$$

Since $\mu_{1,2} \neq \pm 1$ it follows that $\rho$ is a simple reduced eigenvalue. \qed
In the next lemma we show how to compute the coefficients of the characteristic equation.

**Lemma 4.1.4**

Let $M$ be a $4 \times 4$-matrix. Then for coefficients $B$ and $C$ of the characteristic polynomial

$$P(\mu) := \mu^4 - \mu^3 B + \mu^2 C - \mu D + E$$

the following holds:

$$B = \text{tr} M,$$

$$C = \frac{1}{2} \left( \text{tr}^2 M - \text{tr} M^2 \right) = \sum_{i=1}^{3} \sum_{j=i+1}^{4} \det \begin{pmatrix} m_{ii} & m_{ij} \\ m_{ji} & m_{jj} \end{pmatrix}.$$

**Proof:** Let $\mu_1, \ldots, \mu_4$ denote the eigenvalues of $M$. Then we obtain for the characteristic polynomial

$$P(\mu) = \prod_{i=1}^{4} (\mu - \mu_i) = \mu^4 - \mu^3 \sum_{i=1}^{4} \mu_i + \mu^2 \sum_{i=1}^{3} \sum_{j=i+1}^{4} \mu_i \mu_j + \ldots + (-1)^4 \prod_{i=1}^{4} \mu_i.$$ 

We see immediately that $B = \text{tr} M$. It remains to prove the claims for the coefficient $C$. First we have

$$\frac{1}{2} \left( \text{tr}^2 M - \text{tr} M^2 \right) = \frac{1}{2} \left( \left( \sum_{i=1}^{4} \mu_i \right)^2 - \sum_{i=1}^{4} \mu_i^2 \right) = \sum_{i=1}^{3} \sum_{j=i+1}^{4} \mu_i \mu_j = C.$$

and further

$$\frac{1}{2} \left( \text{tr}^2 M - \text{tr} M^2 \right) = \frac{1}{2} \left( \left( \sum_{j=1}^{4} m_{jj} \right)^2 - \sum_{i=1}^{4} \sum_{j=1}^{4} m_{ij} m_{ji} \right) =$$

$$= \frac{1}{2} \left( \left( \sum_{j=1}^{4} m_{jj}^2 + 2 \sum_{i=1}^{3} \sum_{j=i+1}^{4} m_{ij} m_{ji} \right) - \sum_{j=1}^{4} m_{jj}^2 - 2 \sum_{i=1}^{3} \sum_{j=i+1}^{4} m_{ij} m_{ji} \right) =$$

$$= \sum_{i=1}^{3} \sum_{j=i+1}^{4} (m_{ii} m_{jj} - m_{ij} m_{ji}) = \sum_{i=1}^{4} \sum_{j=i+1}^{4} \det \begin{pmatrix} m_{ii} & m_{ij} \\ m_{ji} & m_{jj} \end{pmatrix}. \quad \square$$

We close this section with a result on the existence of curves.
Lemma 4.1.5

Consider the equation

$$ F(\varepsilon, \Omega) = 0, \quad (4.3) $$

where $F : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is a smooth function. Assume that there exists a solution $(\varepsilon_0, \Omega_0)$ of (4.3).

Then the following holds:

Case 1: If $D_\varepsilon F(\varepsilon_0, \Omega_0)$ and $D_\Omega F(\varepsilon_0, \Omega_0)$ do not vanish simultaneously, then there exists a neighborhood $U$ of $(\varepsilon_0, \Omega_0)$ such that (4.3) admits a unique smooth solution curve in $U$.

This curve divides $U$ in two regions. In one region one has $F(\varepsilon, \Omega) < 0$ and in the other $F(\varepsilon, \Omega) > 0$ as illustrated in Figure 4.1.

Case 2: If $D_\varepsilon F(\varepsilon_0, \Omega_0)$ and $D_\Omega F(\varepsilon_0, \Omega_0)$ do vanish simultaneously and

$$ D_{\varepsilon \varepsilon} F(\varepsilon_0, \Omega_0) D_{\Omega \Omega} F(\varepsilon_0, \Omega_0) - (D_{\varepsilon \Omega} F(\varepsilon_0, \Omega_0))^2 < 0 $$

holds, then there exists a neighborhood $U$ of $(\varepsilon_0, \Omega_0)$ such that (4.3) admits two smooth curves that intersect transversally at $(\varepsilon_0, \Omega_0)$. These two curves divide $U$ in four regions. In two opposite regions one has $F(\varepsilon, \Omega) < 0$ and in the other $F(\varepsilon, \Omega) > 0$ as illustrated in Figure 4.2.

Fig. 4.1: The existence of a unique smooth solution curve.

Fig. 4.2: The existence of a two smooth solutions curves.
**Proof:** We start with Case 1. Here the existence of a neighborhood of \((\varepsilon_0, \Omega_0)\) such that \(F(\varepsilon, \Omega) = 0\) admits a unique smooth solution curve follows immediately from the Implicit Function Theorem.

The fact that \(F\) has different signs on different sides of the solution curve is a direct consequence of the non-vanishing gradient.

Now we come to Case 2. We first claim that \(F(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega)\) may be written in the following form:

\[
F(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = F(\varepsilon_0, \Omega_0) + F_\varepsilon(\varepsilon_0, \Omega_0) \Delta \varepsilon + F_\Omega(\varepsilon_0, \Omega_0) \Delta \Omega +
\]

\[
+ \int_0^1 (1 - t) \left(F_{\varepsilon\varepsilon}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \varepsilon^2 +
\]

\[
+ 2F_{\varepsilon\Omega}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \varepsilon \Delta \Omega +
\]

\[
+ F_{\Omega\Omega}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \Omega^2 \right) dt.
\]

To prove this, we integrate by parts:

\[
\int_0^1 (1 - t) \left(F_{\varepsilon\varepsilon}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \varepsilon^2 + 2F_{\varepsilon\Omega}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \varepsilon \Delta \Omega +
\]

\[
+ F_{\Omega\Omega}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \Omega^2 \right) dt =
\]

\[
= (1 - t) \left(F_\varepsilon(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \varepsilon + F_\Omega(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \Omega \right)|_0^1 =
\]

\[
- \int_0^1 (-1) \left(F_\varepsilon(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \varepsilon + F_\Omega(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) \Delta \Omega \right) dt =
\]

\[
= - F_\varepsilon(\varepsilon_0, \Omega_0) - F_\Omega(\varepsilon_0, \Omega_0) + \int_0^1 \frac{d}{dt} F(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) dt =
\]

\[
= - F_\varepsilon(\varepsilon_0, \Omega_0) - F_\Omega(\varepsilon_0, \Omega_0) + F(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) - F(\varepsilon_0, \Omega_0).
\]

Now the claim follows immediately.

Taking into account that in Case 2 the constant and the linear parts vanish, we may further write

\[
F(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = a \Delta \varepsilon^2 + 2b \Delta \varepsilon \Delta \Omega + c \Delta \Omega^2,
\]

where

\[
a := a(\varepsilon_0, \Omega_0, \Delta \varepsilon, \Delta \Omega) := \int_0^1 (1 - t) F_{\varepsilon\varepsilon}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) dt,
\]

\[
b := b(\varepsilon_0, \Omega_0, \Delta \varepsilon, \Delta \Omega) := \int_0^1 (1 - t) F_{\varepsilon\Omega}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) dt,
\]

\[
c := c(\varepsilon_0, \Omega_0, \Delta \varepsilon, \Delta \Omega) := \int_0^1 (1 - t) F_{\Omega\Omega}(\varepsilon_0 + t \Delta \varepsilon, \Omega_0 + t \Delta \Omega) dt.
\]
c := c(ε0, Ω0, Δε, ΔΩ) := \int_0^1 (1 - t)F_{\epsilon\Omega}(\epsilon_0 + t\Delta\epsilon, \Omega_0 + t\Delta\Omega)dt.

Also from the assumptions it follows that

\[ b_0^2 - a_0c_0 := b_0^2(\epsilon_0, \Omega_0) - a(\epsilon_0, \Omega_0)c(\epsilon_0, \Omega_0) = \]
\[ = \left(\frac{1}{2}F_{\epsilon\Omega}(\epsilon_0, \Omega_0)\right)^2 - \frac{1}{2}F_{cc}(\epsilon_0, \Omega_0) \cdot \frac{1}{2}F_{\Omega\Omega}(\epsilon_0, \Omega_0) > 0 \]

It therefore exists a neighborhood \( U_1 \) of \((\epsilon_0, \Omega_0)\) such that

\[ b_0^2(\epsilon_0, \Omega_0, \Delta\epsilon, \Delta\Omega) - a(\epsilon_0, \Omega_0, \Delta\epsilon, \Delta\Omega)c(\epsilon_0, \Omega_0, \Delta\epsilon, \Delta\Omega) > 0 \]

for \((\epsilon_0 + \Delta\epsilon, \Omega_0 + \Delta\Omega)\) in \( U_1 \).

Now let

\[ d := \begin{cases} 
    b + \sqrt{b^2 - ac} & \text{if } b_0 \geq 0 \\
    b - \sqrt{b^2 - ac} & \text{if } b_0 < 0.
\end{cases} \]

It follows immediately that \( d_0 \neq 0 \). Thus there exists a neighborhood \( U_2 \subset U_1 \) and a constant \( \rho_0 > 0 \) such that |\( d \)| > \( \rho_0 \).

This definition allows us to factorize \( F(\epsilon_0 + \Delta\epsilon, \Omega_0 + \Delta\Omega) \):

\[ F(\epsilon_0 + \Delta\epsilon, \Omega_0 + \Delta\Omega) = \frac{1}{d} (a\Delta\epsilon + d\Delta\Omega) (d\Delta\epsilon + c\Delta\Omega). \]

The proof is straightforward:

\[ \frac{1}{d}(a\Delta\epsilon + d\Delta\Omega)(d\Delta\epsilon + c\Delta\Omega) = \]
\[ = a\Delta\epsilon^2 + c\Delta\Omega^2 + \frac{ac + d^2}{d} \Delta\epsilon \Delta\Omega = \]
\[ = a\Delta\epsilon^2 + c\Delta\Omega^2 + \frac{ac + b^2 \pm 2b\sqrt{b^2 - ac} + b^2 - ac}{d} \Delta\epsilon \Delta\Omega = \]
\[ = a\Delta\epsilon^2 + c\Delta\Omega^2 + \frac{2b(b \pm \sqrt{b^2 - ac})}{d} \Delta\epsilon \Delta\Omega = \]
\[ = a\Delta\epsilon^2 + c\Delta\Omega^2 + \frac{2bd}{d} \Delta\epsilon \Delta\Omega = \]
\[ = a\Delta\epsilon^2 + c\Delta\Omega^2 + 2b\Delta\epsilon \Delta\Omega. \]

Thus \( F(\epsilon_0 + \Delta\epsilon, \Omega_0 + \Delta\Omega) = 0 \) implies either

\[ g(\epsilon_0 + \Delta\epsilon, \Omega_0 + \Delta\Omega) := a\Delta\epsilon + d\Delta\Omega = 0 \quad (4.4) \]
4.2 Dissipative Systems

or

\[ h(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) := d\Delta \varepsilon + c\Delta \Omega = 0. \]  

(4.5)

Since \( d_0 \neq 0 \) holds, it follows by the Implicit Function Theorem that in \( U_3 \subset U_2 \) equations (4.4) and (4.5) admit each a unique smooth solution curve \( C_g \) and \( C_h \), respectively.

From the definition of \( d \) it follows that \( |d_0| > |b_0| \). Therefore there exists a neighborhood \( U \subset U_3 \) such that \( |d| \geq |b| \). This implies \( ac - d^2 \leq ac - b^2 < 0 \). Thus in \( U \) the curves intersect only at \((\Delta \varepsilon, \Delta \Omega) = (0, 0)\).

Since \( d_0 \neq 0 \) holds, the gradient of \( g \) does not vanish. Thus \( g \) admits different signs on different sides of the curve \( C_g \). For \( h \) we have an analogous result.

Thus in \( U \) the equation \( F(\varepsilon, \Omega) = 0 \) admits two solution curves that intersect transversally in \((\varepsilon_0, \Omega_0)\). These two curves divide \( U \) in four regions. In two opposite regions one has \( F(\varepsilon, \Omega) < 0 \) and in the other \( F(\varepsilon, \Omega) > 0 \).

4.2 Dissipative Systems

We first assume the system to be dissipative. Then there are three typical situations for the loss of stability as discussed in Section 2.2.

The Case of a Simple Real Eigenvalue at ±1

We start with the case, where a simple real eigenvalue leaves the unit circle at ±1.

**Theorem 4.2.1**

Let the differential equation

\[ \dot{z} = A(t, \varepsilon, \Omega)z \]

be dissipative.

Assume that for a certain \((\varepsilon_0, \Omega_0)\) the monodromy matrix \( M(\varepsilon, \Omega) \) has a single real eigenvalue \( \mu_c(\varepsilon_0, \Omega_0) = \pm 1 \) and that the other eigenvalues lie inside the unit circle.

By Theorem 3.1.1 the monodromy matrix \( M(\varepsilon, \Omega) \) may be transformed to normal form

\[ \tilde{M}(\varepsilon, \Omega) = \begin{pmatrix} \mu_c(\varepsilon, \Omega) & * \\ 0 & \tilde{M}_{nc}(\varepsilon, \Omega) \end{pmatrix}, \]  

(4.6)

where

- for the critical eigenvalue \( \mu_c(\varepsilon, \Omega) \) one has \( \mu_c(\varepsilon_0, \Omega_0) = \pm 1 \) and
- for \((\varepsilon, \Omega)\) sufficiently close to \((\varepsilon_0, \Omega_0)\) the eigenvalues of \( \tilde{M}_{nc}(\varepsilon, \Omega) \) are stable, i.e. they lie inside the unit circle.

In some neighborhood of \((\varepsilon_0, \Omega_0)\) the system is stable, if \( |\mu_c(\varepsilon, \Omega)| \leq 1 \) and unstable, if \( |\mu_c(\varepsilon, \Omega)| > 1 \).
Let

\[ \mu_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \pm 1 + \Delta \varepsilon \mu_c^{10} + \Delta \Omega \mu_c^{01} + \Delta \varepsilon^2 \mu_c^{20} + \Delta \varepsilon \Delta \Omega \mu_c^{11} + \Delta \Omega^2 \mu_c^{02} + \ldots \]  

be the expansion of the critical eigenvalue at \((\varepsilon_0, \Omega_0)\). Then the following holds:

Case 1: If \(\mu_c^{10}\) and \(\mu_c^{01}\) do not vanish simultaneously, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) the equation

\[ \mu_c(\varepsilon, \Omega) = \pm 1 \]  

admits a unique smooth solution curve. This curve divides the neighborhood of \((\varepsilon_0, \Omega_0)\) in two regions. In one region the system is stable and in the other region it is unstable as illustrated in Figure 4.3.

\[ ((\mu_c^{10})^2 \pm 2\mu_c^{20}) ((\mu_c^{01})^2 \pm 2\mu_c^{02}) - (2\mu_c^{10}\mu_c^{01} \pm 2\mu_c^{11})^2 \]

holds, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) equation (4.8) admits two smooth solution curves that intersect transversally at \((\varepsilon_0, \Omega_0)\). These two curves divide the neighborhood of \((\varepsilon_0, \Omega_0)\) in four regions. In two opposite regions the system is stable and in the other regions the system is unstable as illustrated in Figure 4.4.

Fig. 4.3: The existence of the stability boundary.

Fig. 4.4: The existence of the stability boundaries that intersect transversally.
4.2. Dissipative Systems

Proof: For $|\mu_c| \leq 1$ the system is stable, while for $|\mu_c| > 1$ the system is unstable. Since $\mu_c$ cannot leave the real axis we apply Lemma 4.1.5 to the equation

$$F(\varepsilon, \Omega) := \mu_c^2(\varepsilon, \Omega) - 1 = 0. \quad (4.9)$$

To this end we substitute the expansion (4.7) into (4.9):

$$0 = \pm 2 \Delta \varepsilon \mu_c^{10} + 2 \Delta \Omega \mu_c^{01} +$$

$$+ \Delta \varepsilon^2 \left( (\mu_c^{10})^2 \pm 2 \mu_c^{20} \right) + \Delta \varepsilon \Delta \Omega \left( 2 \mu_c^{10} \mu_c^{01} \pm 2 \mu_c^{11} \right) + \Delta \Omega^2 \left( (\mu_c^{01})^2 \pm 2 \mu_c^{02} \right).$$

Under the assumptions of Case 1 we conclude that there exists a neighborhood $U$ of $(\varepsilon_0, \Omega_0)$ such that (4.9) defines a unique smooth curve in $U$. This curve divides $U$ in two regions. In one region one has $F(\varepsilon, \Omega) := \mu_c^2(\varepsilon, \Omega) - 1 \leq 0$ and therefore the system is stable. In the other region of $U$ one has $F(\varepsilon, \Omega) := \mu_c^2(\varepsilon, \Omega) - 1 > 0$ and therefore the system is unstable. Thus in $U$ the stability boundary is a unique smooth curve defined by $F(\varepsilon, \Omega) := \mu_c^2(\varepsilon, \Omega) - 1 = 0$.

Under the assumptions of Case 2 we conclude that there exists a neighborhood $U$ of $(\varepsilon_0, \Omega_0)$ such that (4.9) defines two smooth curves in $U$ that intersect transversally at $(\varepsilon_0, \Omega_0)$. These curves divide $U$ in four regions. In two opposite regions one has $F(\varepsilon, \Omega) := \mu_c^2(\varepsilon, \Omega) - 1 \leq 0$ and therefore the system is stable. In the other regions of $U$ one has $F(\varepsilon, \Omega) := \mu_c^2(\varepsilon, \Omega) - 1 > 0$ and therefore the system is unstable. Thus in $U$ the stability boundary consists of two smooth curves defined by $F(\varepsilon, \Omega) := \mu_c^2(\varepsilon, \Omega) - 1 = 0$. □

The Case of a Pair of Simple Eigenvalues on the Unit Circle

Now we come to the case of a pair of simple complex conjugate eigenvalues leaving the unit circle.

**Theorem 4.2.2**

Let the differential equation

$$\dot{z} = A(t, \varepsilon, \Omega)z$$

be dissipative.

Assume that for a certain $(\varepsilon_0, \Omega_0)$ the monodromy matrix $M(\varepsilon, \Omega)$ has a pair of simple complex conjugate eigenvalues $\mu_c^0$ and $\mu_c^\ast$ and that the other eigenvalues lie inside the unit circle.

By Theorem 3.1.1 the monodromy matrix $M(\varepsilon, \Omega)$ may be transformed to normal form

$$\tilde{M}(\varepsilon, \Omega) = \begin{pmatrix} \mu_c(\varepsilon, \Omega) & * & * \\ 0 & \bar{\mu}_c(\varepsilon, \Omega) & * \\ 0 & 0 & \tilde{M}_{nc}(\varepsilon, \Omega) \end{pmatrix}, \quad (4.10)$$

where
for the critical eigenvalues \( \mu_c(\varepsilon, \Omega) \) and \( \bar{\mu}_c(\varepsilon, \Omega) \) one has \( |\mu_c^0| = |\mu_c(\varepsilon_0, \Omega_0)| = |\bar{\mu}_c(\varepsilon_0, \Omega_0)| = 1 \) and

- for \((\varepsilon, \Omega)\) sufficiently close to \((\varepsilon_0, \Omega_0)\) the eigenvalues of \( \bar{M}_{nc}(\varepsilon, \Omega) \) are stable, i.e. they lie inside the unit circle.

In some neighborhood of \((\varepsilon_0, \Omega_0)\) the system is stable, if \( |\mu_c(\varepsilon, \Omega)| \leq 1 \) and unstable, if \( |\mu_c(\varepsilon, \Omega)| > 1 \).

Let

\[
\mu_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \mu_c^0 + \Delta \varepsilon \mu_c^{10} + \Delta \Omega \mu_c^{01} + \\
+ \Delta \varepsilon^2 \mu_c^{20} + \Delta \varepsilon \Delta \Omega \mu_c^{11} + \Delta \Omega^2 \mu_c^{02} + O((|\Delta \varepsilon| + |\Delta \Omega|)^3) \tag{4.11}
\]

be the expansion of the critical eigenvalue at \((\varepsilon_0, \Omega_0)\). Then the following holds:

Case 1: If \( \text{re}(\bar{\mu}_c^0 \mu_c^{10}) \) and \( \text{re}(\bar{\mu}_c^0 \mu_c^{01}) \) do not vanish simultaneously, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) the equation

\[
|\mu_c(\varepsilon, \Omega)| = 1 \tag{4.12}
\]

admits a unique smooth solution curve. This curve divides the neighborhood of \((\varepsilon_0, \Omega_0)\) in two regions. In one region the system is stable and in the other region it is unstable as illustrated in Figure 4.5.

![Fig. 4.5: The existence of the stability boundary.](image)

Case 2: If \( \text{re}(\bar{\mu}_c^0 \mu_c^{10}) \) and \( \text{re}(\bar{\mu}_c^0 \mu_c^{01}) \) do vanish simultaneously and if

\[
(2 \text{re}(\bar{\mu}_c^0 \mu_c^{20}) + |\mu_c^{10}|^2)(2 \text{re}(\bar{\mu}_c^0 \mu_c^{02}) + |\mu_c^{01}|^2) - \\
-(\text{re}(\bar{\mu}_c^0 \mu_c^{11}) + \text{re}(\bar{\mu}_c^{10} \mu_c^{01}))^2 < 0
\]

holds, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) equation (4.12) admits two smooth solution curves that intersect transversally at \((\varepsilon_0, \Omega_0)\). These two curves divide the neighborhood of \((\varepsilon_0, \Omega_0)\) in four regions. In two opposite regions the system is stable and in the other two regions the system is unstable as illustrated in Figure 4.6.
4.2. Dissipative Systems

Fig. 4.6: The existence of two stability boundaries that intersect transversally.

**Proof:** For $|\mu_c| \leq 1$ the system is stable, while for $|\mu_c| > 1$ the system is unstable. Thus we apply Lemma 4.1.5 to the equation

$$F(\varepsilon, \Omega) := |\mu_c(\varepsilon, \Omega)|^2 - 1 = 0. \quad (4.13)$$

To this end we substitute the expansion (4.11) into (4.13):

$$0 = \tilde{\mu}_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega)\mu_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) - 1 =$$

$$= \Delta \varepsilon \left( \tilde{\mu}_c^0 \mu_c^{10} + \mu_c^0 \tilde{\mu}_c^{10} \right) +$$

$$+ \Delta \Omega \left( \tilde{\mu}_c^0\mu_c^{01} + \mu_c^0\tilde{\mu}_c^{01} \right) +$$

$$+ \Delta \varepsilon^2 \left( \tilde{\mu}_c^0 \mu_c^{20} + \tilde{\mu}_c^{20} \mu_c^0 + \mu_c^{20} \tilde{\mu}_c^0 \right) +$$

$$+ \Delta \varepsilon \Delta \Omega \left( \tilde{\mu}_c^0 \mu_c^{11} + \tilde{\mu}_c^{11} \mu_c^0 + \mu_c^{11} \tilde{\mu}_c^0 \right) +$$

$$+ \Delta \Omega^2 \left( \tilde{\mu}_c^0 \mu_c^{02} + \tilde{\mu}_c^{02} \mu_c^0 + \mu_c^{02} \tilde{\mu}_c^0 \right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3) =$$

$$= \Delta \varepsilon \left( 2 \text{re}(\tilde{\mu}_c^0 \mu_c^{10}) \right) +$$

$$+ \Delta \Omega \left( 2 \text{re}(\mu_c^{01}) \right) +$$

$$+ \Delta \varepsilon^2 \left( 2 \text{re}(\tilde{\mu}_c^0 \mu_c^{20}) + |\mu_c^{10}|^2 \right) +$$

$$+ \Delta \varepsilon \Delta \Omega \left( 2 \text{re}(\tilde{\mu}_c^0 \mu_c^{11}) + 2 \text{re}(\mu_c^{10} \mu_c^{01}) \right) +$$

$$+ \Delta \Omega^2 \left( 2 \text{re}(\mu_c^{02}) + |\mu_c^{01}|^2 \right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$
Now the rest of the proof is identical with the proof of Theorem 4.2.1.

4.3 Canonical or Reversible Systems

Now we assume the system to be canonical or reversible. Again there are three typical situations for the loss of stability (cf. Section 2.2). In contrast to dissipative systems only double eigenvalues are involved in the loss of stability.

The Case of a Collision at ±1

For simplicity we start with the case, where two simple complex conjugate eigenvalues coalesce at ±1 and leave the unit circle on the real axis.

Theorem 4.3.1

Let the differential equation
\[ \dot{z} = A(t, \varepsilon, \Omega)z \]
be canonical or reversible.
Assume that for a certain (\varepsilon_0, \Omega_0) the monodromy matrix \( M(\varepsilon, \Omega) \) has a double real eigenvalue \( \mu_c^0 := \pm 1 \) and that the other eigenvalues lie on the unit circle and are simple. By Theorem 3.1.1 the monodromy matrix \( M(\varepsilon, \Omega) \) may be transformed to normal form
\[ \tilde{M}(\varepsilon, \Omega) = \begin{pmatrix} 
\tilde{M}_\varepsilon(\varepsilon, \Omega) & * \\
0 & \tilde{M}_\text{nc}(\varepsilon, \Omega) 
\end{pmatrix}, \quad (4.14) \]
where

- the so-called critical submatrix \( \tilde{M}_\varepsilon(\varepsilon, \Omega) \) has dimension 2,
- the spectrum of \( \tilde{M}_\varepsilon(\varepsilon, \Omega) \) is symmetric with respect to the real axis and the unit circle,
- for \( (\varepsilon, \Omega) = (\varepsilon_0, \Omega_0) \) the matrix \( \tilde{M}_\varepsilon(\varepsilon, \Omega) \) has a double eigenvalue at ±1 and
- for \( (\varepsilon, \Omega) \) sufficiently close to \( (\varepsilon_0, \Omega_0) \) the eigenvalues of \( \tilde{M}_\text{nc}(\varepsilon, \Omega) \) are stable, i.e. they lie on the unit circle and are simple.

In some neighborhood of \( (\varepsilon_0, \Omega_0) \) the following holds:
If \( |\text{tr}(\tilde{M}_\varepsilon(\varepsilon, \Omega))| - 2 \) is negative or if it vanishes and the eigenvalues of \( \tilde{M}_\varepsilon(\varepsilon, \Omega) \) are semisimple, then the system is stable.
If \( |\text{tr}(\tilde{M}_\varepsilon(\varepsilon, \Omega))| - 2 \) is positive or if it vanishes and the eigenvalues of \( \tilde{M}_\varepsilon(\varepsilon, \Omega) \) are not semisimple, then the system is unstable.

Let
4.3. Canonical or Reversible Systems

\[
\tilde{M}_c(e_0 + \Delta e, \Omega_0 + \Delta \Omega) = \tilde{M}_c^0 + \Delta e \tilde{M}_c^{10} + \Delta \Omega \tilde{M}_c^{01} + \\
+ \Delta e^2 \tilde{M}_c^{20} + \Delta e \Delta \Omega \tilde{M}_c^{11} + \Delta \Omega^2 \tilde{M}_c^{02} + \\
+ O((|\Delta e| + |\Delta \Omega|)^3)
\] (4.15)

be the expansion of \(\tilde{M}_c(e, \Omega)\) at \((e_0, \Omega_0)\). Then the following holds:

Case 1: Assumption: The double eigenvalue \(\mu_c^0 = \pm 1\) of \(M_c(e_0, \Omega_0)\) is not semisimple.

Claim: If tr \(\tilde{M}_c^{10}\) and tr \(\tilde{M}_c^{01}\) do not vanish simultaneously, then in a neighborhood of \((e_0, \Omega_0)\) the equation

\[
\text{tr}(\tilde{M}_c(e, \Omega)) \mp 2 = 0
\] (4.16)

admits a unique smooth solution curve. This curve divides the neighborhood of \((e_0, \Omega_0)\) in two regions. In one region the system is stable and in the other region it is unstable as illustrated in Figure 4.7.

![Fig. 4.7: The existence of the stability boundary.](image)

Case 2: Assumption: The double eigenvalue \(\mu_c^0 = \pm 1\) of \(M_c(e_0, \Omega_0)\) is semisimple. Then tr \(\tilde{M}_c^{10}\) and tr \(\tilde{M}_c^{01}\) do vanish simultaneously.

Claim: If

\[
4 \text{tr} \tilde{M}_c^{20} \text{tr} \tilde{M}_c^{02} - \left(\text{tr} \tilde{M}_c^{11}\right)^2 < 0
\]

holds, then in a neighborhood of \((e_0, \Omega_0)\) equation (4.16) admits two smooth solution curves that intersect transversally at \((e_0, \Omega_0)\). These two curves divide the neighborhood of \((e_0, \Omega_0)\) in four regions. In two opposite regions the system is stable and in the other regions the system is unstable as illustrated in Figure 4.8.
60 4. Non-Autonomous Systems

Fig. 4.8: The existence of two stability boundaries that intersect transversally.

**Proof:** From the assumptions on the matrix $\tilde{M}_c(\varepsilon, \Omega)$ it follows that the characteristic polynomial is reflexive:

$$P(\mu) := \mu^2 - \mu b(\varepsilon, \Omega) + 1$$

with real coefficient $b(\varepsilon, \Omega)$ (cf. Lemma 2.1.2). From Lemma 4.1.2 we know that

$$b(\varepsilon, \Omega) = \text{tr} \tilde{M}_c(\varepsilon, \Omega).$$

For the reduced characteristic polynomial we obtain

$$Q_r(\mu) = \mu^{-1}(\mu^2 - \mu b + 1) = \mu - b + \mu^{-1} =: \rho - b.$$  

Now Lemma 4.1.1 implies that the zeroes of $P_c$ lie on the unit circle, if and only if $|b| < 2$ holds. Thus the system is stable, if and only if $|b| - 2$ is negative or if it vanishes and the eigenvalues are semisimple. Thus we apply Lemma 4.1.5 to the equation

$$F(\varepsilon, \Omega) := b(\varepsilon, \Omega) \mp 2 = \text{tr} \tilde{M}_c(\varepsilon, \Omega) \mp 2 = 0. \quad (4.17)$$

We substitute the expansion of $\tilde{M}_c(\varepsilon, \Omega)$ into (4.17):

$$0 = b(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) \mp 2 = \text{tr} \tilde{M}_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) \mp 2 =$$

$$= \text{tr} \tilde{M}_c^0 \mp 2 + \Delta \varepsilon \text{tr} \tilde{M}_c^{10} + \Delta \Omega \text{tr} \tilde{M}_c^{01} +$$

$$+ \Delta \varepsilon^2 \text{tr} \tilde{M}_c^{20} + \Delta \varepsilon \Delta \Omega \text{tr} \tilde{M}_c^{11} + \Delta \Omega^2 \text{tr} \tilde{M}_c^{02} +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3) =$$

$$= \Delta \varepsilon \text{tr} \tilde{M}_c^{10} + \Delta \Omega \text{tr} \tilde{M}_c^{01} +$$

$$+ \Delta \varepsilon^2 \text{tr} \tilde{M}_c^{20} + \Delta \varepsilon \Delta \Omega \text{tr} \tilde{M}_c^{11} + \Delta \Omega^2 \text{tr} \tilde{M}_c^{02} +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$
Now we show that the coefficients of $\Delta \varepsilon$ and $\Delta \Omega$ vanish, if $\tilde{M}_c^0$ is semisimple. From the reflexivity of the characteristic polynomial and Lemma 4.1.2 we know that

$$\det \tilde{M}_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = 1.$$ 

Using Lemma 4.1.2 we obtain

$$2 = 2 \det \tilde{M}_c = \text{tr}^2 \tilde{M}_c - \text{tr}(\tilde{M}_c^2).$$

The expansion of the right-hand side yields

$$2 \det \tilde{M}_c = \left(\text{tr}^2 \tilde{M}_c - \text{tr}(\tilde{M}_c^2)\right) =$$

$$= \left(\text{tr}^2 \tilde{M}_c^0 + 2 \Delta \varepsilon \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{10} + 2 \Delta \Omega \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{01}\right) -$$

$$- \left(\text{tr}((\tilde{M}_c^0)^2) + 2 \Delta \varepsilon \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{10}) + 2 \Delta \Omega \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{01})\right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|^2)) =$$

$$= \left(\text{tr}^2 \tilde{M}_c^0 - \text{tr}((\tilde{M}_c^0)^2)\right) + \Delta \varepsilon \left(2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{10} - 2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{10})\right) +$$

$$+ \Delta \Omega \left(2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{01} - 2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{01})\right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|^2)).$$

The matrix $\tilde{M}_c^0$ may be put into Jordan normal form via a matrix $S$:

$$S^{-1} \tilde{M}_c^0 S = \pm I + \delta N, \quad \text{where} \quad N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and $\delta = 0$ in the semisimple case and $\delta = 1$ in the non-semisimple case. Taking this into account and the fact that the trace is invariant under such transformations, we further obtain

$$\det \tilde{M}_c = 1 + \Delta \varepsilon \left(\pm 2 \text{tr} \tilde{M}_c^{10} - \text{tr}(S^{-1} \tilde{M}_c^0 S S^{-1} \tilde{M}_c^{10} S)\right) +$$

$$+ \Delta \Omega \left(\pm 2 \text{tr} \tilde{M}_c^{01} - \text{tr}(S^{-1} \tilde{M}_c^0 S S^{-1} \tilde{M}_c^{01} S)\right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|^2)) =$$

$$= 1 + \Delta \varepsilon \left(\pm 2 \text{tr} \tilde{M}_c^{10} - \text{tr}((\pm I + \delta N) S^{-1} \tilde{M}_c^{10} S)\right) +$$

$$+ \Delta \Omega \left(\pm 2 \text{tr} \tilde{M}_c^{01} - \text{tr}((\pm I + \delta N) S^{-1} \tilde{M}_c^{01} S)\right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|^2)) =$$
Thus the following identities hold:

\[ \pm \text{tr} \tilde{M}^0 = \delta \text{tr}(NS^{-1}\tilde{M}^{10}S), \quad \text{and} \quad \pm \text{tr} \tilde{M}^{01} = \delta \text{tr}(NS^{-1}\tilde{M}^{01}S). \]

Substituting these identities into the expansion of \( h(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) \equiv 2 = 0 \) given above, we obtain

\[ 0 = \pm \Delta \varepsilon \delta \text{tr}(NS^{-1}\tilde{M}^{10}S) \pm \Delta \Omega \delta \text{tr}(NS^{-1}\tilde{M}^{01}S) + O((|\Delta \varepsilon| + |\Delta \Omega|)^2). \]

This leads to the following conclusions:

\begin{itemize}
  \item In the semisimple case the linear parts of the expansion of the equation of the stability boundary vanish.
  \item In the non-semisimple case it depends on the matrices \( \tilde{M}^{10} \) and \( \tilde{M}^{01} \) whether the linear parts vanish.
\end{itemize}

Now the rest of the proof is identical with the proof of Theorem 4.2.1. \( \square \)

**The Case of a Krein Collision**

Now we come to the case of a Krein collision.

**Theorem 4.3.2**

Let the differential equation

\[ \dot{z} = A(t, \varepsilon, \Omega)z \]

be canonical or reversible.

Assume that for a certain \((\varepsilon_0, \Omega_0)\) the monodromy matrix \( M(\varepsilon, \Omega) \) has a pair of double complex conjugate eigenvalues \( \mu^0_c \) and \( \bar{\mu}^0_c \) on the unit circle and that the other eigenvalues lie on the unit circle and are simple.

By Theorem 3.1.1 the monodromy matrix \( M(\varepsilon, \Omega) \) may be transformed to normal form:

\[ \tilde{M}(\varepsilon, \Omega) = \begin{pmatrix} \tilde{M}_c(\varepsilon, \Omega) & * \\ 0 & \tilde{M}_{nc}(\varepsilon, \Omega) \end{pmatrix} \]  

(4.18)

with

\[ \tilde{M}_c(\varepsilon, \Omega) = \begin{pmatrix} \tilde{M}_{11}(\varepsilon, \Omega) & * \\ 0 & \tilde{M}_{22}(\varepsilon, \Omega) \end{pmatrix}, \]

(4.19)

where
4.3. Canonical or Reversible Systems

- the so-called critical submatrix $\tilde{M}_c(\varepsilon, \Omega)$ has dimension 4,
- the spectrum of $\tilde{M}_c(\varepsilon, \Omega)$ is symmetric with respect to the real axis and the unit circle,
- for $(\varepsilon, \Omega) = (\varepsilon_0, \Omega_0)$ the submatrices $\tilde{M}_{11}(\varepsilon, \Omega)$ and $\tilde{M}_{22}(\varepsilon, \Omega)$ have the double eigenvalues $\mu^c_0$ and $\tilde{\mu}^c_0$, respectively, and
- for $(\varepsilon, \Omega)$ sufficiently close to $(\varepsilon_0, \Omega_0)$ the eigenvalues of $\tilde{M}_{nc}(\varepsilon, \Omega)$ are stable, i.e. they lie on the unit circle and are simple.

In some neighborhood of $(\varepsilon_0, \Omega_0)$ the following holds:
If $2 \text{tr}((\tilde{M}_c(\varepsilon, \Omega))^2) - \text{tr}^2 \tilde{M}_c(\varepsilon, \Omega) + 8$ is positive or if it vanishes and the eigenvalues of $\tilde{M}_c(\varepsilon, \Omega)$ are semisimple, then the system is stable.
If $2 \text{tr}((\tilde{M}_c(\varepsilon, \Omega))^2) - \text{tr}^2 \tilde{M}_c(\varepsilon, \Omega) + 8$ is negative or if it vanishes and the eigenvalues of $\tilde{M}_c(\varepsilon, \Omega)$ are not semisimple, then the system is unstable.

Let
$$\tilde{M}_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \tilde{M}_c^0 + \Delta \varepsilon \tilde{M}_c^{10} + \Delta \Omega \tilde{M}_c^{01} +$$
$$+ \Delta^2 \tilde{M}_c^{20} + \Delta \varepsilon \Delta \Omega \tilde{M}_c^{11} + \Delta \Omega^2 \tilde{M}_c^{02} +$$
$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3) \quad (4.20)$$

be the expansion of $\tilde{M}_c(\varepsilon, \Omega)$ at $(\varepsilon_0, \Omega_0)$. Then the following holds:

Case 1: Assumption: The double eigenvalue $\mu^c_0 \neq \pm 1$ of $\tilde{M}_c(\varepsilon_0, \Omega_0)$ is not semisimple.
Claim: If
$$\text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{10} - 2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{10})$$
and
$$\text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{01} - 2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{01})$$
do not vanish simultaneously, then in a neighborhood of $(\varepsilon_0, \Omega_0)$ the equation
$$2 \text{tr}((\tilde{M}_c(\varepsilon, \Omega))^2) - \text{tr}^2 \tilde{M}_c(\varepsilon, \Omega) + 8 = 0 \quad (4.21)$$
admits a unique smooth solution curve. This curve divides the neighborhood of $(\varepsilon_0, \Omega_0)$ in two regions. In one region the system is stable and in the other region it is unstable as illustrated in Figure 4.9.

![Fig. 4.9: The existence of the stability boundary.](image-url)
Case 2: Assumption: The double eigenvalue $\mu_c^0 \neq \pm 1$ of $M_c(\varepsilon_0, \Omega_0)$ is semisimple. Then

$$\text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{10} - 2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{10})$$

and

$$\text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{01} - 2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{01})$$

do vanish simultaneously.

Claim: If

$$(2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{20} - 4 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{20}) + \text{tr}^2 \tilde{M}_c^{10} - 2 \text{tr}((\tilde{M}_c^{10})^2)).$$

$$(2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{02} - 4 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{02}) + \text{tr}^2 \tilde{M}_c^{02} - 2 \text{tr}((\tilde{M}_c^{02})^2)) -$$

$$(\text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{11} - 2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{11}) + \text{tr} \tilde{M}_c^{10} \text{tr} \tilde{M}_c^{01} - 2 \text{tr}(\tilde{M}_c^{10} \tilde{M}_c^{01}))^2 < 0$$

holds, then in a neighborhood of $(\varepsilon_0, \Omega_0)$ equation (4.21) admits two smooth solution curves that intersect transversally at $(\varepsilon_0, \Omega_0)$. These two curves divide the neighborhood of $(\varepsilon_0, \Omega_0)$ in four regions. In two opposite regions the system is stable and in the other regions the system is unstable as illustrated in Figure 4.10.

![Fig. 4.10: The existence of two stability boundaries that intersect transversally.](image)

**Proof:** From the assumptions on the matrix $\tilde{M}_c(\varepsilon, \Omega)$ it follows that the characteristic polynomial is reflexive:

$$P(\mu) := \mu^4 - \mu^3 B(\varepsilon, \Omega) + \mu^2 C(\varepsilon, \Omega) - \mu B(\varepsilon, \Omega) + 1$$

with real coefficients $B(\varepsilon, \Omega)$ and $C(\varepsilon, \Omega)$ (cf. Lemma 2.1.2). From Lemma 4.1.4 we know that

$$B(\varepsilon, \Omega) = \text{tr} \tilde{M}_c(\varepsilon, \Omega),$$

$$C(\varepsilon, \Omega) = \frac{1}{2} \left( \text{tr}^2 \tilde{M}_c(\varepsilon, \Omega) - \text{tr}((\tilde{M}_c(\varepsilon, \Omega))^2) \right).$$
4.3. Canonical or Reversible Systems

For the reduced characteristic polynomial we find (cf. (4.2))

$$Q_c(p) = p^2 - pB + (C - 2).$$

The zeroes of the reduced characteristic equation read

$$\rho_{1,2} = \frac{B \pm \sqrt{B^2 - 4(C - 2)}}{2}.$$ 

Thus we have a double reduced eigenvalue if and only if $B^2 - 4C + 8 = 0$. For $B^2 - 4C + 8 < 0$ the reduced eigenvalue does not lie in the interval $[-2, 2]$. We conclude by Lemma 4.1.3 that the corresponding eigenvalues do not lie on the unit circle and that the system is therefore unstable. For $B^2 - 4C + 8 > 0$ the reduced eigenvalues are real and for $(\varepsilon, \Omega)$ sufficiently close to $(\varepsilon_0, \Omega_0)$ they lie in the interval $[-2, 2]$. Therefore the system is stable. Thus we apply Lemma 4.1.5 to the equation

$$F(\varepsilon, \Omega) := B^2(\varepsilon, \Omega) - 4C(\varepsilon, \Omega) + 8 = 0. \quad (4.22)$$

Using the definition of $B$ and $C$ together with Lemma 4.1.4 we obtain from (4.22):

$$0 = B^2(\varepsilon, \Omega) - 4C(\varepsilon, \Omega) + 8 =
= \text{tr}^2 \tilde{M}_c(\varepsilon, \Omega) - 4 \cdot \frac{1}{2} \left( \text{tr}^2 \tilde{M}_c(\varepsilon, \Omega) - \text{tr}((\tilde{M}_c(\varepsilon, \Omega))^2) \right) + 8 =
= 2 \text{tr}((\tilde{M}_c(\varepsilon, \Omega))^2) - \text{tr}^2 \tilde{M}_c(\varepsilon, \Omega) + 8. \quad (4.23)$$

In order to compute the expansion of (4.23) we first consider the expansion of $\text{tr}(\tilde{M}_c(\varepsilon_0 + \Delta\varepsilon, \Omega_0 + \Delta\Omega))^2$:

$$\text{tr}(\tilde{M}_c(\varepsilon_0 + \Delta\varepsilon, \Omega_0 + \Delta\Omega))^2 = \text{tr}(\tilde{M}_c^0)^2 + \Delta\varepsilon \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{10} + \tilde{M}_c^{10} \tilde{M}_c^0) +$
$$+ \Delta\Omega \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{01} + \tilde{M}_c^{01} \tilde{M}_c^0) +$
$$+ \Delta\varepsilon^2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{20} + \tilde{M}_c^{20} \tilde{M}_c^0 + (\tilde{M}_c^{10})^2) +$
$$+ \Delta\varepsilon \Delta\Omega \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{11} + \tilde{M}_c^{11} \tilde{M}_c^0 +$
$$+ \tilde{M}_c^{10} \tilde{M}_c^{01} + \tilde{M}_c^{01} \tilde{M}_c^{10}) +$
$$+ \Delta\Omega^2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{02} + \tilde{M}_c^{02} \tilde{M}_c^0 + (\tilde{M}_c^{01})^2) +$
$$+ O((|\Delta\varepsilon| + |\Delta\Omega|)^3)$$

Using $\text{tr}(AB) = \text{tr}(BA)$ we further obtain

$$\text{tr} \left( (\tilde{M}_c(\varepsilon_0 + \Delta\varepsilon, \Omega_0 + \Delta\Omega))^2 \right) = \text{tr}((\tilde{M}_c^0)^2) + \Delta\varepsilon \text{tr}((\tilde{M}_c^0 \tilde{M}_c^{10} + \tilde{M}_c^{10} \tilde{M}_c^0) +$
$$+ \Delta\Omega \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{01} + \tilde{M}_c^{01} \tilde{M}_c^0) +$
$$+ \Delta\varepsilon^2 (2 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{20}) + \text{tr}((\tilde{M}_c^{10})^2)) +$$
Now we consider the expansion \( \text{tr}^2 \tilde{M}_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) \): 
\[
\text{tr}^2 \tilde{M}_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \text{tr}^2 \tilde{M}_c^0 + \Delta \varepsilon \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{01} + \\
+ \Delta \Omega \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{10} + \\
+ \Delta \varepsilon^2 (2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{20} + \text{tr}^2 \tilde{M}_c^{10}) + \\
+ \Delta \varepsilon \Delta \Omega (2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{11} + 2 \text{tr} \tilde{M}_c^{10} \text{tr} \tilde{M}_c^{01}) + \\
+ \Delta \Omega^2 (2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{02} + \text{tr}^2 \tilde{M}_c^{01}) + \\
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).
\]

Inserting these expansions into equation (4.23) we obtain
\[
0 = \Delta \varepsilon \left( 4 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{10}) - 2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{10} \right) + \\
+ \Delta \Omega \left( 4 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{01}) - 2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{01} \right) + \\
+ \Delta \varepsilon^2 \left( 4 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{20}) - 2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{20} + 2 \text{tr}(\tilde{M}_c^{10})^2 - \text{tr}^2 \tilde{M}_c^{10} \right) + \\
+ \Delta \varepsilon \Delta \Omega \left( 4 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{11}) - 2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{11} \right) + \\
+ \Delta \Omega^2 \left( 4 \text{tr}(\tilde{M}_c^0 \tilde{M}_c^{02}) - 2 \text{tr} \tilde{M}_c^0 \text{tr} \tilde{M}_c^{02} + 2 \text{tr}(\tilde{M}_c^{01})^2 - \text{tr}^2 \tilde{M}_c^{01} \right) + \\
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).
\]

Note: From the assumptions of the theorem it follows that (4.23) holds for \((\varepsilon, \Omega) = (\varepsilon_0, \Omega_0)\). Therefore the constant term of (4.24) vanishes.

Now we show that the coefficients of \(\Delta \varepsilon\) and \(\Delta \Omega\) vanish, if \(\tilde{M}_c^0\) is semisimple.

We expand the matrices \(\tilde{M}_{11}(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega)\) and \(\tilde{M}_{22}(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega)\) up to order 1:
\[
\tilde{M}_{11}(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \tilde{M}_{11}^0 + \Delta \varepsilon \tilde{M}_{11}^{10} + \Delta \Omega \tilde{M}_{11}^{01} + O((|\Delta \varepsilon| + |\Delta \Omega|)^2),
\]
\[
\tilde{M}_{22}(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \tilde{M}_{22}^0 + \Delta \varepsilon \tilde{M}_{22}^{10} + \Delta \Omega \tilde{M}_{22}^{01} + O((|\Delta \varepsilon| + |\Delta \Omega|)^2).
\]
4.3. Canonical or Reversible Systems

From the reflexivity of the characteristic polynomial and Lemma 4.1.4 we know that
\[ \det \tilde{M}_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = 1. \]

Using Lemma 4.1.2 we obtain in a first step
\[ 1 = \det \begin{pmatrix} M_{11} & * \\ 0 & M_{22} \end{pmatrix} = \det \tilde{M}_{11} \cdot \det \tilde{M}_{22} = \]
\[ = \frac{1}{2} \left( \text{tr}^2 \tilde{M}_{11} - \text{tr}(\tilde{M}^2_{11}) \right) \cdot \frac{1}{2} \left( \text{tr}^2 \tilde{M}_{22} - \text{tr}(\tilde{M}^2_{22}) \right). \]

The expansion of the factors on the right-hand side was already computed in the proof of Theorem 4.3.1. We find
\[ 2 \det \tilde{M}_{ii} = \left( \text{tr}^2 \tilde{M}_{ii} - \text{tr}((\tilde{M}^0_{ii})^2) \right) + \Delta \varepsilon \left( 2 \text{tr} \tilde{M}^0_{ii} \text{tr} \tilde{M}^{10}_{ii} - 2 \text{tr}(\tilde{M}^0_{ii} \tilde{M}^{10}_{ii}) \right) + \]
\[ + \Delta \Omega \left( 2 \text{tr} \tilde{M}^0_{ii} \text{tr} \tilde{M}^{01}_{ii} - 2 \text{tr}(\tilde{M}^0_{ii} \tilde{M}^{01}_{ii}) \right) + \]
\[ + O((|\Delta \varepsilon| + |\Delta \Omega|)^2). \]

The matrices \( \tilde{M}^0_{ii} \) may be put into Jordan normal form via matrices \( S_i \)
\[ S_1^{-1} \tilde{M}^0_{11} S_1 = \mu^0_c I + N, \quad \text{where} \quad N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \]
\[ S_2^{-1} \tilde{M}^0_{22} S_2 = \mu^0_c I + N, \]
and \( \delta = 0 \) in the semisimple case and \( \delta = 1 \) in the non-semisimple case. Taking this into account and the fact that the trace is invariant under such transformations, we further obtain
\[ \det \tilde{M}_{11} = (\mu^0_c)^2 + \Delta \varepsilon \left( \mu^0_c \text{tr} \tilde{M}^{10}_{11} - \delta \text{tr}(NS_1^{-1} \tilde{M}^{10}_{11} S_1) \right) + \]
\[ + \Delta \Omega \left( \mu^0_c \text{tr} \tilde{M}^{01}_{11} - \delta \text{tr}(NS_1^{-1} \tilde{M}^{01}_{11} S_1) \right) + O((|\Delta \varepsilon| + |\Delta \Omega|)^2), \]
\[ \det \tilde{M}_{22} = (\mu^0_c)^2 + \Delta \varepsilon \left( \mu^0_c \text{tr} \tilde{M}^{10}_{22} - \delta \text{tr}(NS_2^{-1} \tilde{M}^{10}_{22} S_2) \right) + \]
\[ + \Delta \Omega \left( \mu^0_c \text{tr} \tilde{M}^{01}_{22} - \delta \text{tr}(NS_2^{-1} \tilde{M}^{01}_{22} S_2) \right) + O((|\Delta \varepsilon| + |\Delta \Omega|)^2). \]

Combining the expansions for \( \det \tilde{M}_{11} \) and \( \det \tilde{M}_{22} \) we obtain
\[ 1 = 1 + \Delta \varepsilon \left( \mu^0_c \text{tr} \tilde{M}^{10}_{22} - \delta(\mu^0_c)^2 \text{tr}(NS_2^{-1} \tilde{M}^{10}_{22} S_2) + \mu^0_c \text{tr} \tilde{M}^{10}_{11} - \delta(\mu^0_c)^2 \text{tr}(NS_1^{-1} \tilde{M}^{10}_{11} S_1) \right) + \]
\[ + \Delta \Omega \left( \mu^0_c \text{tr} \tilde{M}^{01}_{22} - \delta(\mu^0_c)^2 \text{tr}(NS_2^{-1} \tilde{M}^{01}_{22} S_2) + \mu^0_c \text{tr} \tilde{M}^{01}_{11} - \delta(\mu^0_c)^2 \text{tr}(NS_1^{-1} \tilde{M}^{01}_{11} S_1) \right) + \]
\[ + O((|\Delta \varepsilon| + |\Delta \Omega|)^2). \]
Thus the following identities hold:

\[
\mu_c^0 \text{tr} \tilde{M}_{22}^{10} + \mu_c^0 \text{tr} \tilde{M}_{11}^{10} = \delta \left( (\mu_c^0)^2 \text{tr} (NS_2^{-1} \tilde{M}_{22}^{10} S_2) + (\mu_c^0)^2 \text{tr} (NS_1^{-1} \tilde{M}_{11}^{10} S_1) \right), \quad (4.25)
\]

\[
\mu_c^0 \text{tr} \tilde{M}_{22}^{01} + \mu_c^0 \text{tr} \tilde{M}_{11}^{01} = \delta \left( (\mu_c^0)^2 \text{tr} (NS_2^{-1} \tilde{M}_{22}^{01} S_2) + (\mu_c^0)^2 \text{tr} (NS_1^{-1} \tilde{M}_{11}^{01} S_1) \right). \quad (4.26)
\]

In order to use these identities we first have to express the coefficients of the linear parts of (4.24) in terms of the coefficients of the expansions of \(\tilde{M}_{11}\) and \(\tilde{M}_{22}\), respectively. For convenience we write \(ij\) for 10 or 01, respectively. From (4.24) we obtain

\[
4 \text{tr}(\tilde{M}_c^0 \tilde{M}^{ij}_c) - 2 \text{tr} \tilde{M}^0_c \text{tr} \tilde{M}^{ij}_c =
\]

\[
= + 4 \left( \left( \begin{array}{cc}
\tilde{M}^0_{11} & \ast \\
0 & \tilde{M}^0_{22}
\end{array} \right) \left( \begin{array}{cc}
\tilde{M}^{ij}_{11} & \ast \\
0 & \tilde{M}^{ij}_{22}
\end{array} \right) \right) -
\]

\[
- 2 \left( \text{tr} \tilde{M}^0_{11} + \text{tr} \tilde{M}^0_{22} \right) \left( \text{tr} \tilde{M}^{ij}_{11} + \text{tr} \tilde{M}^{ij}_{22} \right) =
\]

\[
= + 4 \left( \mu_c^0 \text{tr} \tilde{M}^{ij}_{11} + \delta \text{tr} (NS_1^{-1} \tilde{M}^{ij}_{11} S_1) + \mu_c^0 \text{tr} \tilde{M}^{ij}_{22} + \delta \text{tr} (NS_2^{-1} \tilde{M}^{ij}_{22} S_2) \right) -
\]

\[
- 2 \left( 2\mu_c^0 + 2\mu_c^0 \right) \left( \text{tr} \tilde{M}^{ij}_{11} + \text{tr} \tilde{M}^{ij}_{22} \right) =
\]

\[
= - 4\mu_c^0 \text{tr} \tilde{M}^{ij}_{22} - 4\mu_c^0 \tilde{M}^{ij}_{11} + 4\delta \left( \text{tr} (NS_1^{-1} \tilde{M}^{ij}_{11} S_1) + \text{tr} (NS_2^{-1} \tilde{M}^{ij}_{22} S_2) \right)
\]

and finally using the identities (4.25) and (4.26), respectively

\[
= 4\delta \left( (1 - (\mu_c^0)^2) \text{tr} (NS_1^{-1} \tilde{M}^{ij}_{11} S_1) + (1 - (\mu_c^0)^2) \text{tr} (NS_2^{-1} \tilde{M}^{ij}_{22} S_2) \right).
\]

This leads to the following conclusions:

- In the semisimple case the linear parts of the expansion of the equation of the stability boundary vanish.

- In the non-semisimple case it depends on the matrices \(\tilde{M}_{11}^{10}\), \(\tilde{M}_{22}^{10}\) and \(\tilde{M}_{11}^{01}\), \(\tilde{M}_{22}^{01}\), respectively, whether the linear parts vanish.

Now the claims follow from Lemma 4.1.5. \(\square\)
4.4 Four Dimensional Systems

We close this chapter with a short consideration on the four dimensional case. The restriction on four dimensions allows us to combine the results of the previous sections.

Let

$$\tilde{M}_c(\varepsilon, \Omega) = \begin{pmatrix} \tilde{M}_{11}(\varepsilon, \Omega) & \ast \\ 0 & \tilde{M}_{22}(\varepsilon, \Omega) \end{pmatrix}$$

and

$$B(\varepsilon, \Omega) = \text{tr} \tilde{M}_c(\varepsilon, \Omega),$$

$$C(\varepsilon, \Omega) = \frac{1}{2} \left( \text{tr}^2 \tilde{M}_c - \text{tr}((\tilde{M}_c)^2) \right)$$

(ef. Section 4.3).

First we recall the equation (4.22) for the stability boundary in the case of a Krein collision:

$$B^2(\varepsilon, \Omega) - 4C(\varepsilon, \Omega) + 8 = 0. \quad (4.22)$$

On the one hand if we insert $\rho = \pm 2$ into the reduced characteristic equation (4.2)

$$Q(\rho) = \rho^2 - \rho B(\varepsilon, \Omega) + C(\varepsilon, \Omega) - 2 = 0,$$

we obtain

$$\mp 2B(\varepsilon, \Omega) + C(\varepsilon, \Omega) + 2 = 0. \quad (4.27)$$

On the other hand if (4.27) holds, then we may substitute $C = \pm 2B - 2$ into the reduced characteristic equation

$$0 = \rho^2 - \rho B + C - 2 = \rho^2 - \rho B \pm 2B - 4.$$

The solutions are

$$\rho_{1,2} = \frac{B \pm \sqrt{B^2 - 4(\pm 2B - 4)}}{2} = \frac{B \pm (B \mp 4)}{2}.$$

Thus $\rho = \pm 2$ is a reduced eigenvalue. Moreover for $|B| \neq 4$ it is simple and for $|B| > 4$ the second reduced eigenvalue does not lie in the interval $[-2, 2]$.

We conclude that (4.27) describes the stability boundary in the case of a collision at $\pm 1$. Since the coefficients $B(\varepsilon, \Omega)$ and $C(\varepsilon, \Omega)$ are real (ef. Lemma 2.1.2 and Lemma 2.1.4), the left-hand sides of the equations (4.22) and (4.27) are real. This allows us to represent the solutions in a $B$-$C$-coordinate system.

The curves are the solutions of (4.22) and (4.27). The small diagrams show qualitatively the situation of the eigenvalues of $M_c(\varepsilon, \Omega)$. We see that only in the shaded area enclosed by the parabola and the straight lines the eigenvalues of $M_c(\varepsilon, \Omega)$ are on the unit circle.
Fig. 4.11: The solution of Equations (4.22) and (4.27).
If the system of differential equations (2.1) is autonomous, we do not need to compute the monodromy matrix, but we may work with the matrix $A$ itself. We adopt the organization of the last chapter:

Section 5.1: We prove some auxiliary results.
Section 5.2: We discuss the stability of dissipative systems.
Section 5.3: We discuss the stability of canonical or reversible systems.

5.1 Auxiliary Results

Consider the system of differential equations

$$\dot{x} = A(\varepsilon, \Omega)x. \quad (2.1)$$

In Section 2.1 we saw that the spectrum of the matrix $A(\varepsilon, \Omega)$ determines the stability. Applying the normal form transformation introduced in Chapter 3 to the matrix $A$ we obtain a matrix in block triangular form. That allows us to work with the spectra of the submatrices in the main diagonal. This reduces the degree of the characteristic equations to 1 or 2 in the dissipative case and to 2 or 4 in the canonical or reversible case.

In the latter case the spectrum of the matrix $A$ is symmetric with respect to the real and imaginary axis (cf. Lemma 2.1.1 and Lemma 2.1.3). Due to Lemma 3.2.1 the spectra of the submatrices in the main diagonal show the same symmetries. These symmetries may be used for a further reduction of the degree of the characteristic equations.

We start with the two dimensional case. Let $A$ be a $2 \times 2$-matrix whose spectrum is symmetric with respect to the real and imaginary axis. Then the characteristic polynomial has the form

$$p(\lambda) := \lambda^2 + c$$

with real coefficient $c$ (cf. the proof of Lemma 2.1.1 \textit{(ii)}).
If we put \( \rho := \lambda^2 \), then we obtain the so-called reduced characteristic polynomial:

\[
Q_c(\rho) := \rho - c. \tag{5.1}
\]

Its zeroes are called reduced eigenvalues of \( A \). The next lemma states some properties of reduced eigenvalues.

**Lemma 5.1.1**

Let \( A \) be a 2 \( \times \) 2-matrix whose spectrum is symmetric with respect to the real and the imaginary axis. Then the following statements hold:

(i) \( \lambda \) is an eigenvalue of \( A \) if and only if \( \rho \) is a reduced eigenvalue.

(ii) An eigenvalue \( \lambda \) of \( A \) lies on the imaginary axis if and only if the reduced eigenvalue \( \rho \) lies on the non-positive real axis.

(iii) If \( \lambda = 0 \) is an eigenvalue of \( A \), then \( \rho = 0 \) is a simple reduced eigenvalue.

**Proof:** The claims follow immediately from the definition of the reduced eigenvalue. \( \square \)

We now come to the four dimensional case. Consider a 4 \( \times \) 4-matrix whose spectrum is symmetric with respect to the real and imaginary axis. Then the characteristic equation is of the form

\[
P_c(\lambda) := \lambda^4 + B\lambda^2 + C = 0
\]

with real coefficients \( B \) and \( C \) (cf. the proof of Lemma 2.1.1 (ii)).

Again we put \( \rho := \lambda^2 \) and obtain the so-called reduced characteristic polynomial:

\[
Q_c(\rho) := \rho^2 + B\rho + C. \tag{5.2}
\]

The next lemma gives some results for the reduced eigenvalues for the case of a 4 \( \times \) 4-matrix.

**Lemma 5.1.2**

Let \( A \) be a 4 \( \times \) 4-matrix whose spectrum is symmetric with respect to the real and imaginary axis. Then the following statements hold:

(i) \( \lambda \) is an eigenvalue of \( A \) if and only if \( \rho \) is a reduced eigenvalue.

(ii) An eigenvalue \( \lambda \) of \( A \) lies on the imaginary axis if and only if the reduced eigenvalue \( \rho \) lies on the non-positive real axis.

(iii) If \( \lambda \) is a double eigenvalue of \( A \) different from 0, then \( \rho \) is a double reduced eigenvalue of \( A \).

(iv) If \( \lambda = 0 \) is a double eigenvalue of \( A \), then \( \rho = 0 \) is a simple eigenvalue of \( A \).
5.2 Dissipative Systems

Proof: The proofs of (i), (ii) and (iv) follow immediately from the definition of the reduced eigenvalue. It remains to prove (iii).

Let \( \lambda_1 \neq 0 \) be a double zero of \( P_c \). Then \( \lambda_1 \in \mathbb{R} \) or \( \lambda_0 \in i\mathbb{R} \). Thus \( -\lambda_1 \) is also a double eigenvalue. Thus we have

\[
0 = \lambda_1^4 + \lambda_1^2B + C = (\lambda - \lambda_1)^2(\lambda + \lambda_1)^2 = (\lambda^2 - \lambda_1^2)^2 = (\rho - \rho_1)^2.
\]

We conclude that \( \rho_1 \) is a double reduced eigenvalue. \( \square \)

5.2 Dissipative Systems

We first assume the system to be dissipative. Then there are two typical situations for the loss of stability as discussed in Section 2.2.

The Case of a Simple Eigenvalue at 0

We start with the case, where a simple real eigenvalue leaves the left half plane at 0.

Theorem 5.2.1

Let the differential equation

\[
\dot{z} = A(\varepsilon, \Omega)z
\]

be dissipative.

Assume that for a certain \((\varepsilon_0, \Omega_0)\) the matrix \(A(\varepsilon, \Omega)\) has a single real eigenvalue \( \lambda_c^0 := 0 \) and that the other eigenvalues lie in the left half plane.

By Theorem 3.1.1 the matrix \(A(\varepsilon, \Omega)\) may be transformed to normal form

\[
\tilde{A}(\varepsilon, \Omega) = \begin{pmatrix} \lambda_c(\varepsilon, \Omega) & * \\ 0 & \tilde{A}_{nc}(\varepsilon, \Omega) \end{pmatrix},
\]

where

- for the critical eigenvalue \( \lambda_c(\varepsilon, \Omega) \) one has \( \lambda_c^0 = \lambda_c(\varepsilon_0, \Omega_0) = 0 \) and
- for \((\varepsilon, \Omega)\) sufficiently close to \((\varepsilon_0, \Omega_0)\) the eigenvalues of \( \tilde{A}_{nc}(\varepsilon, \Omega) \) are stable, i.e. they lie in the left half plane.

In some neighborhood of \((\varepsilon_0, \Omega_0)\) the system is stable, if \( \lambda_c(\varepsilon, \Omega) \leq 0 \) and unstable, if \( \lambda_c(\varepsilon, \Omega) > 0 \).

Let

\[
\lambda_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \Delta \varepsilon \lambda_c^{10} + \Delta \Omega \lambda_c^{01} + \\
+ \Delta \varepsilon^2 \lambda_c^{20} + \Delta \varepsilon \Delta \Omega \lambda_c^{11} + \Delta \Omega^2 \lambda_c^{02} \\
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3)
\]

be the expansion the critical eigenvalue at \((\varepsilon_0, \Omega_0)\). Then the following holds:
Case 1: If \( \lambda_{c0}^{10} \) and \( \lambda_{c0}^{01} \) do not vanish simultaneously, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) the equation

\[ \lambda_c(\varepsilon, \Omega) = 0 \]  

(5.5)

admits a unique smooth solution curve. This curve divides the neighborhood of \((\varepsilon_0, \Omega_0)\) in two regions. In one region the system is stable and in the other region it is unstable as illustrated in Figure 5.1.

![Fig. 5.1: The existence of the stability boundary.](image)

Case 2: If \( \lambda_{c0}^{10} \) and \( \lambda_{c0}^{01} \) do vanish simultaneously and if

\[ 4\lambda_{c0}^{20} \lambda_{c0}^{02} - (\lambda_{c0}^{11})^2 < 0 \]

holds, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) equation (5.5) admits two smooth solution curves that intersect transversally at \((\varepsilon_0, \Omega_0)\). These two curves divide the neighborhood of \((\varepsilon_0, \Omega_0)\) in four regions. In two opposite regions the system is stable and in the other regions the system is unstable as illustrated in Figure 5.2.

![Fig. 5.2: The existence of the stability boundaries that intersect transversally.](image)

Proof: For \( \lambda_c \leq 0 \) the system is stable, while for \( \lambda_c > 0 \) the system is unstable. Thus we apply Lemma 4.1.5 to the equation

\[ F(\varepsilon, \Omega) := \lambda_c(\varepsilon, \Omega) = 0. \]  

(5.6)
5.2. Dissipative Systems

To this end we substitute the expansion (5.4) into (5.6):

\[ 0 = \lambda_c (\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \lambda_c \lambda_c^{10} + \Delta \Omega \lambda_c^{01} + \]

\[ + \Delta \varepsilon^2 \lambda_c^{20} + \Delta \varepsilon \Delta \Omega \lambda_c^{11} + \Delta \Omega^2 \lambda_c^{02} + O((|\Delta \varepsilon| + |\Delta \Omega|)^3) \]

Now the rest of the proof is identical to the proof of Theorem 4.2.1. \( \square \)

**The Case of a Pair of Simple Eigenvalues on the Imaginary Axis**

Now we come to the case of a pair of simple complex conjugate eigenvalues leaving the left half plane.

**Theorem 5.2.2**

Let the differential equation

\[ \dot{z} = A(\varepsilon, \Omega)z \]

be dissipative. Assume that for a certain \((\varepsilon_0, \Omega_0)\) the matrix \(A(\varepsilon, \Omega)\) has a pair of single complex conjugate eigenvalues \(\lambda_c^0\) and \(\lambda_c^*\) on the imaginary axis and that the other eigenvalues lie in the left half plane. By Theorem 3.1.1 the matrix \(A(\varepsilon, \Omega)\) may be transformed to normal form

\[ \tilde{A}(\varepsilon, \Omega) = \begin{pmatrix} \lambda_c(\varepsilon, \Omega) & * & * \\ 0 & \tilde{\lambda}_c(\varepsilon, \Omega) & * \\ 0 & 0 & \tilde{A}_{nc}(\varepsilon, \Omega) \end{pmatrix} \quad (5.7) \]

where

- for the critical eigenvalues \(\lambda_c(\varepsilon, \Omega)\) and \(\tilde{\lambda}_c(\varepsilon, \Omega)\) one has \(\text{re} \lambda_c(\varepsilon_0, \Omega_0) = \text{re} \tilde{\lambda}_c(\varepsilon, \Omega) = 0\) and

- for \((\varepsilon, \Omega)\) sufficiently close to \((\varepsilon_0, \Omega_0)\) the eigenvalues of \(\tilde{A}_{nc}(\varepsilon, \Omega)\) are stable, i.e. they lie in the left half plane.

In some neighborhood of \((\varepsilon_0, \Omega_0)\) the system is stable, if \(\text{re} \lambda_c(\varepsilon, \Omega) \leq 0\) and unstable, if \(\text{re} \lambda_c(\varepsilon, \Omega) > 0\).

Let

\[ \lambda_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \lambda_c^0 + \Delta \varepsilon \lambda_c^{10} + \Delta \Omega \lambda_c^{01} + \]

\[ + \Delta \varepsilon^2 \lambda_c^{20} + \Delta \varepsilon \Delta \Omega \lambda_c^{11} + \Delta \Omega^2 \lambda_c^{02} + O((|\Delta \varepsilon| + |\Delta \Omega|)^3) \quad (5.8) \]

be the expansion the critical eigenvalue at \((\varepsilon_0, \Omega_0)\). Then the following holds:
Case 1: If \( \text{re} \lambda^0_c \) and \( \text{re} \lambda^0_f \) do not vanish simultaneously, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) the equation

\[
\text{re} \lambda_c(\varepsilon, \Omega) = 0 \tag{5.9}
\]

admits a unique smooth solution curve. This curve divides the neighborhood of \((\varepsilon_0, \Omega_0)\) in two regions. In one region the system is stable and in the other region it is unstable as illustrated in Figure 5.3.

![Fig. 5.3: The existence of the stability boundary.](image)

Case 2: If \( \text{re} \lambda^0_c \) and \( \text{re} \lambda^0_f \) do vanish simultaneously and if

\[
4 \text{re} \lambda_c^0 \text{re} \lambda_f^0 - (\text{re} \lambda_c^1)^2 < 0
\]

holds, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) equation (5.9) admits two smooth solution curves that intersect transversally at \((\varepsilon_0, \Omega_0)\). These two curves divide the neighborhood of \((\varepsilon_0, \Omega_0)\) in four regions. In two opposite regions the system is stable and in the other regions the system is unstable as illustrated in Figure 5.4.

![Fig. 5.4: The existence of two stability boundaries that intersect transversally.](image)

**Proof:** For \( \text{re} \lambda_c \leq 0 \) the system is stable, while for \( \text{re} \lambda_c > 0 \) the system is unstable. Thus we apply Lemma 4.1.5 to the equation

\[
F(\varepsilon, \Omega) := \text{re} \lambda_c(\varepsilon, \Omega) = 0. \tag{5.10}
\]
5.3. Canonical or Reversible Systems

To this end we substitute the expansion (5.8) into (5.10):

$$0 = \text{re} \lambda_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \text{re} \lambda_c^0 + \Delta \varepsilon \text{re} \lambda_c^{01} + \Delta \Omega \text{re} \lambda_c^{02} +$$

$$+ \Delta \varepsilon^2 \text{re} \lambda_c^{20} + \Delta \varepsilon \Delta \Omega \text{re} \lambda_c^{11} + \Delta \Omega^2 \text{re} \lambda_c^{02} + O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$

Now the rest of the proof is identical to the proof of Theorem 4.2.1. □

5.3 Canonical or Reversible Systems

Now we assume the system to be canonical or reversible. Again there are three typical situations for the loss of stability (cf. Section 2.2). In contrast to dissipative systems only double eigenvalues are involved in the loss of stability.

The Case of a Collision at 0

For simplicity we start with the case, where two simple complex conjugate eigenvalues coalesce at 0 and leave the imaginary axis on the real axis.

Theorem 5.3.1

Let the differential equation

$$\dot{z} = A(\varepsilon, \Omega)z$$

be canonical or reversible.

Assume that for a certain \((\varepsilon_0, \Omega_0)\) the matrix \(A(\varepsilon, \Omega)\) has a double eigenvalue \(\lambda_c^0 := 0\) and that the other eigenvalues lie on the imaginary axis and are simple.

By Theorem 3.1.1 the matrix \(A(\varepsilon, \Omega)\) may be transformed to normal form

$$\tilde{A}(\varepsilon, \Omega) = \begin{pmatrix} \tilde{A}_c(\varepsilon, \Omega) & 0 \\ 0 & \tilde{A}_{nc}(\varepsilon, \Omega) \end{pmatrix}^*,$$

(5.11)

where

- the so-called critical submatrix \(\tilde{A}_c(\varepsilon, \Omega)\) has dimension 2,
- the spectrum of \(\tilde{A}_c(\varepsilon, \Omega)\) is symmetric with respect to the real and the imaginary axis,
- for \((\varepsilon, \Omega) = (\varepsilon_0, \Omega_0)\) the matrix \(\tilde{A}_c(\varepsilon, \Omega)\) has a double eigenvalue at 0 and
- for \((\varepsilon, \Omega)\) sufficiently close to \((\varepsilon_0, \Omega_0)\) the eigenvalues of \(\tilde{A}_{nc}(\varepsilon, \Omega)\) are stable, i.e. they lie on the imaginary axis and are simple.

In some neighborhood of \((\varepsilon_0, \Omega_0)\) the following holds:

If \(\text{tr}^2 \tilde{A}_c(\varepsilon, \Omega) - \text{tr}((\tilde{A}_c(\varepsilon, \Omega))^2)\) is negative or if it vanishes and the eigenvalues of \(\tilde{A}_c(\varepsilon, \Omega)\) are semisimple, then the system is stable.

If \(\text{tr}^2 \tilde{A}_c(\varepsilon, \Omega) - \text{tr}((\tilde{A}_c(\varepsilon, \Omega))^2)\) is positive or if it vanishes and the eigenvalues of \(\tilde{A}_c(\varepsilon, \Omega)\) are not semisimple, then the system is unstable.
Let
\[
\tilde{A}_c(\varepsilon + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \tilde{A}^0_c + \Delta \varepsilon \tilde{A}^{10}_c + \Delta \Omega \tilde{A}^{01}_c + \\
+ \Delta \varepsilon^2 \tilde{A}^{20}_c + \Delta \varepsilon \Delta \Omega \tilde{A}^{11}_c + \Delta \Omega^2 \tilde{A}^{02}_c + \\
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3)
\]
be the expansion of \(\tilde{A}_c(\varepsilon, \Omega)\) at \((\varepsilon_0, \Omega_0)\). Then the following holds:

**Case 1:** Assumption: The double eigenvalue \(\lambda^0_c = 0\) of \(A(\varepsilon_0, \Omega_0)\) is not semisimple.

**Claim:** If
\[
\text{tr} \tilde{A}^0_{11} \text{tr} \tilde{A}^{10}_{11} - \text{tr}(\tilde{A}^0_{11} \tilde{A}^{10}_{11})
\]
and
\[
\text{tr} \tilde{A}^0_{11} \text{tr} \tilde{A}^{01}_{11} - \text{tr}(\tilde{A}^0_{11} \tilde{A}^{01}_{11})
\]
do not vanish simultaneously, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) the equation
\[
\text{tr}^2 \tilde{A}_c(\varepsilon, \Omega) - \text{tr}((\tilde{A}_c(\varepsilon, \Omega))^2) = 0
\]
adopts a unique smooth solution curve. This curve divides the neighborhood of \((\varepsilon_0, \Omega_0)\) in two regions. In one region the system is stable and in the other region it is unstable as illustrated in Figure 5.5.

**Case 2:** Assumption: The double eigenvalue \(\lambda^0_c = 0\) of \(A(\varepsilon_0, \Omega_0)\) is semisimple.

Then
\[
\text{tr} \tilde{A}^0_{11} \text{tr} \tilde{A}^{10}_{11} - \text{tr}(\tilde{A}^0_{11} \tilde{A}^{10}_{11})
\]
and
\[
\text{tr} \tilde{A}^0_{11} \text{tr} \tilde{A}^{01}_{11} - \text{tr}(\tilde{A}^0_{11} \tilde{A}^{01}_{11})
\]
do vanish simultaneously.

**Claim:** If
\[
\left(2 \text{tr} \tilde{A}^0_c \text{tr} \tilde{A}^{20}_c - 2 \text{tr}(\tilde{A}^0_c \tilde{A}^{20}_c) + \text{tr}^2 \tilde{A}^{10}_c - \text{tr}((\tilde{A}^{10}_c)^2)\right) \\
- \left(2 \text{tr} \tilde{A}^0_c \text{tr} \tilde{A}^{02}_c - 2 \text{tr}(\tilde{A}^0_c \tilde{A}^{02}_c) + \text{tr}^2 \tilde{A}^{01}_c - \text{tr}((\tilde{A}^{01}_c)^2)\right) \\
+ \left(\text{tr} \tilde{A}^0_c \text{tr} \tilde{A}^c_c - \text{tr}(\tilde{A}^0_c \tilde{A}^c_c) + \text{tr} \tilde{A}^{10}_c \text{tr} \tilde{A}^{01}_c - \text{tr}(\tilde{A}^{10}_c \tilde{A}^{01}_c)\right) < 0
\]
5.3. Canonical or Reversible Systems

holds, then in a neighborhood of \((\varepsilon_0, \Omega_0)\) equation (5.13) admits two smooth solution curves that intersect transversally at \((\varepsilon_0, \Omega_0)\). These two curves divide the neighborhood of \((\varepsilon_0, \Omega_0)\) in four regions. In two opposite regions the system is stable and in the other regions the system is unstable as illustrated in Figure 5.6.

![Diagram showing stability boundaries](image)

**Fig. 5.6:** The existence of two stability boundaries that intersect transversally.

**Proof:** From the assumption on the matrix \(A_c(\varepsilon, \Omega)\) it follows that the characteristic polynomial of \(\tilde{A}_c(\varepsilon, \Omega)\) is of the form

\[
P(\lambda) := \lambda^2 + c(\varepsilon, \Omega) = 0
\]

with real coefficient \(c(\varepsilon, \Omega)\) (cf. the proof of Lemma 2.1.1 (ii)). From Lemma 4.1.4 we know that

\[
c = \det \tilde{A}_c = \frac{1}{2} \left( \text{tr}^2 \tilde{A}_c - \text{tr}(\tilde{A}_c^2) \right).
\]

Now Lemma 5.1.1 implies that the zeroes of \(P_c\) lie on the imaginary axis, if and only if \(c \leq 0\) holds. Thus the system is stable, if and only if \(c\) is negative or if \(c\) vanishes and the eigenvalues are semisimple. Thus we apply Lemma 4.1.5 to the equation

\[
F(\varepsilon, \Omega) := 2c(\varepsilon, \Omega) = \text{tr}^2 \tilde{A}_c(\varepsilon, \Omega) - \text{tr}(\tilde{A}_c(\varepsilon, \Omega))^2 = 0. \quad (5.14)
\]

The expansions of \(\text{tr}^2 \tilde{A}_c\) and \(\text{tr} \tilde{A}^2_c\), respectively, read

\[
\text{tr}^2 \tilde{A}_c = \text{tr}^2 \tilde{A}^0_c + \Delta \varepsilon \left( 2 \text{tr} \tilde{A}^0_c \text{tr} \tilde{A}^{10}_c \right) + \Delta \Omega \left( 2 \text{tr} \tilde{A}^0_c \text{tr} \tilde{A}^{01}_c \right) +
\]

\[
+ \Delta \varepsilon^2 \left( 2 \text{tr} \tilde{A}^0_c \text{tr} \tilde{A}^{20}_c + \text{tr}^2 \tilde{A}^{10}_c \right) +
\]

\[
+ \Delta \varepsilon \Delta \Omega \left( 2 \text{tr} \tilde{A}^0_c \text{tr} \tilde{A}^{11}_c + 2 \text{tr} \tilde{A}^{10}_c \text{tr} \tilde{A}^{01}_c \right) +
\]

\[
+ \Delta \Omega^2 \left( 2 \text{tr} \tilde{A}^0_c \text{tr} \tilde{A}^{02}_c + \text{tr}^2 \tilde{A}^{01}_c \right) +
\]

\[
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3)
\]
and
\[
\text{tr}((\tilde{A}_c)^2) = \text{tr}((\tilde{A}_c^0)^2) + \Delta \varepsilon \left(2 \text{tr} \tilde{A}_c^0 \tilde{A}_c^{10} + \Delta \Omega \left(2 \text{tr} \tilde{A}_c^0 \tilde{A}_c^{01}\right) + \right.
\]
\[
+ \Delta \varepsilon^2 \left(2 \text{tr} \tilde{A}_c^0 \tilde{A}_c^{20} + \text{tr}((\tilde{A}_c^{10})^2)\right) +
\]
\[
+ \Delta \varepsilon \Delta \Omega \left(2 \text{tr} \tilde{A}_c^0 \tilde{A}_c^{11} + 2 \text{tr} \tilde{A}_c^{10} \tilde{A}_c^{01}\right) +
\]
\[
+ \Delta \Omega^2 \left(2 \text{tr} \tilde{A}_c^0 \tilde{A}_c^{02} + \text{tr}((\tilde{A}_c^{01})^2)\right) +
\]
\[
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).
\]

It follows that
\[
c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) =
\]
\[
\Delta \varepsilon \left(\text{tr} \tilde{A}_c^0 \text{tr} \tilde{A}_c^{10} - \text{tr}((\tilde{A}_c^0 \tilde{A}_c^{10})\right) + \Delta \Omega \left(\text{tr} \tilde{A}_c^0 \text{tr} \tilde{A}_c^{01} - \text{tr}(\tilde{A}_c^0 \tilde{A}_c^{01})\right) +
\]
\[
+ \Delta \varepsilon^2 \left(\text{tr} \tilde{A}_c^0 \text{tr} \tilde{A}_c^{20} - \text{tr}((\tilde{A}_c^{10})^2)\right)\right)
\]
\[
+ \Delta \varepsilon \Delta \Omega \left(\text{tr} \tilde{A}_c^0 \text{tr} \tilde{A}_c^{11} - \text{tr}(\tilde{A}_c^0 \tilde{A}_c^{11}) + \text{tr} \tilde{A}_c^{10} \text{tr} \tilde{A}_c^{01} - \text{tr}(\tilde{A}_c^{10} \tilde{A}_c^{01})\right) +
\]
\[
+ \Delta \Omega^2 \left(\text{tr} \tilde{A}_c^0 \text{tr} \tilde{A}_c^{02} - \text{tr}(\tilde{A}_c^0 \tilde{A}_c^{02}) + \frac{1}{2} \left(\text{tr}^2 \tilde{A}_c^{01} - \text{tr}((\tilde{A}_c^{01})^2)\right)\right) +
\]
\[
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).
\]

Now we show that the coefficients of $\Delta \varepsilon$ and $\Delta \Omega$ vanish, if $\tilde{A}_c^0$ is semisimple.
The matrix $\tilde{A}_c^0$ may be put into Jordan normal form via a matrix $S$
\[
S^{-1} \tilde{A}_c^0 S = \delta N, \quad \text{where} \quad N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}
\]
and $\delta = 0$ in the semisimple case and $\delta = 1$ in the non-semisimple case. Taking this into account and the fact that the trace is invariant under such transformations, we further obtain
\[
c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) =
\]
\[
= \Delta \varepsilon \left(\text{tr}(S^{-1} \tilde{A}_c^0 S) \text{tr}(S^{-1} \tilde{A}_c^{10} S) - \text{tr}(S^{-1} \tilde{A}_c^0 S S^{-1} \tilde{A}_c^{10} S)\right) +
\]
\[
+ \Delta \Omega \left(\text{tr}(S^{-1} \tilde{A}_c^0 S) \text{tr}(S^{-1} \tilde{A}_c^{01} S) - \text{tr}(S^{-1} \tilde{A}_c^0 S S^{-1} \tilde{A}_c^{01} S)\right) +
\]
5.3. Canonical or Reversible Systems

\[ + O((|\Delta \varepsilon| + |\Delta \Omega|)^2) = \]
\[ = \Delta \varepsilon \left( \delta \text{ tr } N \text{ tr } (S^{-1} \tilde{A}_c^{10} S) - \delta \text{ tr } (NS^{-1} \tilde{A}_c^{10} S) \right) + \]
\[ + \Delta \Omega \left( \delta \text{ tr } N \text{ tr } (S^{-1} \tilde{A}_c^{01} S) - \delta \text{ tr } (NS^{-1} \tilde{A}_c^{01} S) \right) + \]
\[ + O((|\Delta \varepsilon| + |\Delta \Omega|)^2) = \]
\[ = - \Delta \varepsilon \delta \text{ tr } (NS^{-1} \tilde{A}_c^{10} S) - \Delta \Omega \delta \text{ tr } (NS^{-1} \tilde{A}_c^{01} S) + O((|\Delta \varepsilon| + |\Delta \Omega|)^2). \]

This leads to the following conclusions:

- In the semisimple case the linear parts of the expansion of the equation of the stability boundary vanish.
- In the non-semisimple case it depends on the matrices \( \tilde{A}_c^{10} \) and \( \tilde{A}_c^{01} \) whether the linear parts vanish.

Now the rest of the proof is identical with the proof of Theorem 4.2.1. \( \Box \)

The Case of a Krein Collision

Now we come to the case of a Krein collision.

**Theorem 5.3.2**

Let the differential equation

\[ \dot{z} = A(\varepsilon, \Omega)z \]

be canonical or reversible.

Assume that for a certain \((\varepsilon_0, \Omega_0)\) the matrix \( A(\varepsilon, \Omega) \) has a pair of double complex conjugate eigenvalues \( \lambda_c^0 \) and \( \lambda'_c^0 \) on the imaginary axis and that the other eigenvalues lie on the imaginary axis and are simple.

By Theorem 3.1.1 the matrix \( A(\varepsilon, \Omega) \) may be transformed to normal form:

\[ \tilde{A}(\varepsilon, \Omega) = \begin{pmatrix} \tilde{A}_c(\varepsilon, \Omega) & * \\ 0 & \tilde{A}_{nc}(\varepsilon, \Omega) \end{pmatrix} \]  \( (5.16) \)

with

\[ \tilde{A}_c(\varepsilon, \Omega) = \begin{pmatrix} \tilde{A}_{11}(\varepsilon, \Omega) & * \\ 0 & \tilde{A}_{22}(\varepsilon, \Omega) \end{pmatrix}, \]  \( (5.17) \)

where

- the so-called critical submatrix \( \tilde{A}_c(\varepsilon, \Omega) \) has dimension 4.
• the spectrum of $\tilde{A}_c(\varepsilon, \Omega)$ is symmetric with respect to the real and the imaginary axis,

• for $(\varepsilon, \Omega) = (\varepsilon_0, \Omega_0)$ the submatrices $\tilde{A}_{11}(\varepsilon, \Omega)$ and $\tilde{A}_{22}(\varepsilon, \Omega)$ have the double eigenvalues $\lambda_0^0$ and $\lambda_0^c$, respectively, and

• for $(\varepsilon, \Omega)$ sufficiently close to $(\varepsilon_0, \Omega_0)$ the eigenvalues of $\tilde{A}_{nc}(\varepsilon, \Omega)$ are stable, i.e. they lie on the imaginary axis and are simple.

Then in some neighborhood of $(\varepsilon_0, \Omega_0)$ the following holds:

If $2 \text{tr}((\tilde{A}_{11}(\varepsilon, \Omega))^2) - \text{tr}^2 \tilde{A}_{11}(\varepsilon, \Omega)$ is positive or if it vanishes and the eigenvalues of $\tilde{A}_c(\varepsilon, \Omega)$ are semisimple, then the system is stable.

If $2 \text{tr}((\tilde{A}_{11}(\varepsilon, \Omega))^2) - \text{tr}^2 \tilde{A}_{11}(\varepsilon, \Omega)$ is negative or if it vanishes and the eigenvalues of $\tilde{A}_c(\varepsilon, \Omega)$ are non semisimple, then the system is unstable.

Let

$$\tilde{A}_c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \tilde{A}_c^0 + \Delta \varepsilon \tilde{A}_c^{10} + \Delta \Omega \tilde{A}_c^{01} +$$

$$+ \Delta \varepsilon^2 \tilde{A}_c^{20} + \Delta \varepsilon \Delta \Omega \tilde{A}_c^{11} + \Delta \Omega^2 \tilde{A}_c^{02} +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$

be the expansion of $\tilde{A}_c(\varepsilon, \Omega)$ at $(\varepsilon_0, \Omega_0)$. Then the following holds:

Case 1: Assumption: The double eigenvalue $\lambda_0^0 \neq 0$ of $A(\varepsilon_0, \Omega_0)$ is not semisimple.

Claim: If

$$2 \text{tr}(\tilde{A}_{11}^0 \tilde{A}_{11}^{10}) - \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^{10}$$

and

$$2 \text{tr}(\tilde{A}_{11}^0 \tilde{A}_{11}^{01}) - \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^{01}$$

do not vanish simultaneously, then in a neighborhood of $(\varepsilon_0, \Omega_0)$ the equation

$$2 \text{tr}((\tilde{A}_{11}(\varepsilon, \Omega))^2) - \text{tr}^2 \tilde{A}_{11}(\varepsilon, \Omega) = 0$$

admits a unique smooth solution curve. This curve divides the neighborhood of $(\varepsilon_0, \Omega_0)$ in two regions. In one region the system is stable and in the other region it is unstable as illustrated in Figure 5.7.

Fig. 5.7: The existence of the stability boundary.
Case 2: Assumption: The double eigenvalue $\lambda^0_c \neq 0$ of $A(\varepsilon_0, \Omega_0)$ is semisimple. Then

$$2 \operatorname{tr}(\tilde{A}^0_{11}, \tilde{A}^{10}_{11}) - \operatorname{tr} \tilde{A}^0_{11} \operatorname{tr} \tilde{A}^{10}_{11}$$

and

$$2 \operatorname{tr}(\tilde{A}^0_{11}, \tilde{A}^{01}_{11}) - \operatorname{tr} \tilde{A}^0_{11} \operatorname{tr} \tilde{A}^{01}_{11}$$
do vanish simultaneously.

Claim: If

$$\left(4 \operatorname{tr}(\tilde{A}^0_{11}, \tilde{A}^{20}_{11}) - 2 \operatorname{tr} \tilde{A}^0_{11} \operatorname{tr} \tilde{A}^{20}_{11} + 2 \operatorname{tr}((\tilde{A}^{10}_{11})^2) - \operatorname{tr}^2 \tilde{A}^{10}_{11}\right).$$

$$\left(4 \operatorname{tr}(\tilde{A}^0_{11}, \tilde{A}^{02}_{11}) - 2 \operatorname{tr} \tilde{A}^0_{11} \operatorname{tr} \tilde{A}^{02}_{11} + 2 \operatorname{tr}((\tilde{A}^{01}_{11})^2) - \operatorname{tr}^2 \tilde{A}^{01}_{11}\right) -$$

$$\left(2 \operatorname{tr}(\tilde{A}^0_{11}, \tilde{A}^{11}_{11}) - \operatorname{tr} \tilde{A}^0_{11} \operatorname{tr} \tilde{A}^{11}_{11} + 2 \operatorname{tr} \tilde{A}^{10}_{11} \operatorname{tr} \tilde{A}^{01}_{11} - \operatorname{tr} \tilde{A}^{10}_{11} \operatorname{tr} \tilde{A}^{01}_{11}\right)^2 < 0$$

holds, then in a neighborhood of $(\varepsilon_0, \Omega_0)$ equation (5.19) admits two smooth solution curves that intersect transversally in $(\varepsilon_0, \Omega_0)$. These two curves divide the neighborhood of $(\varepsilon_0, \Omega_0)$ in four regions. In two opposite regions the system is stable and in the other regions the system is unstable as illustrated in Figure 5.8.

**Proof:** From the assumption on the matrix $A_c(\varepsilon, \Omega)$ it follows that the characteristic equation is of the form

$$P_c(\lambda) := \lambda^4 + \lambda^2 B(\varepsilon, \Omega) + C(\varepsilon, \Omega) = 0$$

where $B(\varepsilon, \Omega)$ and $C(\varepsilon, \Omega)$ are real (cf. the proof of Lemma 2.1.1 (ii)). From Lemma 4.1.4 we know that

$$B(\varepsilon, \Omega) = \frac{1}{2} \left(\operatorname{tr}^2 \tilde{A}_c(\varepsilon, \Omega) - \operatorname{tr}((\tilde{A}_c(\varepsilon, \Omega))^2)\right),$$

$$C(\varepsilon, \Omega) = \det \tilde{A}_c(\varepsilon, \Omega).$$
For the reduced characteristic polynomial we find (cf. (5.1))
\[ Q(p) := p^2 + pB(\varepsilon, \Omega) + C(\varepsilon, \Omega) = 0. \]
The zeroes of the reduced characteristic equation read
\[ \rho_{1,2} = \frac{-B \pm \sqrt{B^2 - 4C}}{2}. \]
Thus we have a double reduced eigenvalue if and only if \( B^2 - 4C = 0 \). For \( B^2 - 4C < 0 \) the reduced eigenvalue does not lie on the non-positive real axis. We conclude by Lemma 5.1.2 that the corresponding eigenvalues do not lie on the imaginary axis and that the system is therefore unstable. For \( B^2 - 4C > 0 \) the reduced eigenvalue is real and for \( (\varepsilon, \Omega) \) sufficiently close to \( (\varepsilon_0, \Omega_0) \) it is negative. Therefore the system is stable.

Thus we apply Lemma 4.1.5 to the equation
\[ F(\varepsilon, \Omega) := B^2(\varepsilon, \Omega) - 4C(\varepsilon, \Omega) = 0. \tag{5.20} \]
In contrast to the non-autonomous case the coefficient \( C \) is the determinant of the 4x4-matrix and therefore a non-linear function in \( (\varepsilon, \Omega) \). We avoid this problem using the symmetry of the spectrum of \( A_c(\varepsilon, \Omega) \).
From the symmetries of the spectra of \( \tilde{A}_{11}(\varepsilon, \Omega) \) and \( \tilde{A}_{22}(\varepsilon, \Omega) \) we obtain for their characteristic equations
\[ p_{11}(\lambda) = \lambda^2 - \lambda b(\varepsilon, \Omega) + c(\varepsilon, \Omega) = 0, \]
\[ p_{22}(\lambda) = \lambda^2 + \lambda b(\varepsilon, \Omega) + c(\varepsilon, \Omega) = 0, \]
where \( b(\varepsilon, \Omega) := \text{tr} \tilde{A}_{11}(\varepsilon, \Omega) \) and \( c(\varepsilon, \Omega) := \det \tilde{A}_{11}(\varepsilon, \Omega) \) are real. Thus the stability boundary is described by
\[ b^2(\varepsilon, \Omega) - 4c(\varepsilon, \Omega) = 0. \tag{5.21} \]
The product of \( p_{11}(\lambda) \) and \( p_{22}(\lambda) \) is equal to \( P_c(\lambda) \):
\[ P_c(\lambda) = \lambda^4 + \lambda^2 B + C = (\lambda^2 - \lambda b + c)(\lambda^2 + \lambda b + c) = \lambda^4 + \lambda^2 (2c - b^2) + c^2. \]
Thus we have
\[ B(\varepsilon, \Omega) = 2c(\varepsilon, \Omega) - b^2(\varepsilon, \Omega) \quad \text{and} \quad C(\varepsilon, \Omega) = c^2(\varepsilon, \Omega). \]
Hence we obtain
\[ B^2 - 4C = (2c - b^2)^2 - 4c^2 = b^2(b^2 - 4c). \]
In the case of a collision at \( \lambda \neq 0 \), we have \( b \neq 0 \). Thus the equations \( B^2 - 4C = 0 \) and \( b^2 - 4c = 0 \) are equivalent.
By Lemma 4.1.2 we have
\[ b^2 - 4c = \text{tr}^2 \tilde{A}_{11} - 4 \det \tilde{A}_{11} = \text{tr}^2 \tilde{A}_{11} - 2 \left( \text{tr}^2 \tilde{A}_{11} - \text{tr} \tilde{A}_{11}^2 \right) = 2 \text{tr} \tilde{A}_{11}^2 - \text{tr}^2 \tilde{A}_{11}. \]
5.3. Canonical or Reversible Systems

We have already computed the expansions of $\text{tr}^2 \tilde{A}_{11}$ and $\text{tr} \tilde{A}_{11}^2$ in the previous section. Inserting these expansions into the equation of the stability boundary we obtain

$$b^2(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) - 4c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) =$$

$$= \Delta \varepsilon \left( 4 \text{tr}(\tilde{A}_{11}^0 \tilde{A}_{11}^0) - 2 \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^0 \right) + \Delta \Omega \left( 4 \text{tr}(\tilde{A}_{11}^0 \tilde{A}_{11}^0) - 2 \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^0 \right) +$$

$$+ \Delta \varepsilon \left( 4 \text{tr}(\tilde{A}_{11}^0 \tilde{A}_{11}^0) - 2 \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^0 + 2 \text{tr}((\tilde{A}_{11}^0)^2) - \text{tr}^2 \tilde{A}_{11}^0 \right) +$$

$$+ \Delta \Omega \left( 4 \text{tr}(\tilde{A}_{11}^0 \tilde{A}_{11}^0) - 2 \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^0 + 4 \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^0 - 2 \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^0 \right) +$$

$$+ \Delta \varepsilon \Delta \Omega \left( 4 \text{tr}(\tilde{A}_{11}^0 \tilde{A}_{11}^0) - 2 \text{tr} \tilde{A}_{11}^0 \text{tr} \tilde{A}_{11}^0 + 2 \text{tr}((\tilde{A}_{11}^0)^2) - \text{tr}^2 \tilde{A}_{11}^0 \right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$

Now we show that the coefficients of $\Delta \varepsilon$ and $\Delta \Omega$ vanish, if $\tilde{A}_{11}^0$ is semisimple.

The matrices $\tilde{A}_{11}^0$ may be put into Jordan normal form via matrices $S_i$:

$$S_i^{-1} \tilde{A}_{11}^0 S_i = \lambda_i I + N,$$

where $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

and $\delta = 0$ in the semisimple case and $\delta = 1$ in the non-semisimple case. Taking this into account and the fact that the trace is invariant under such transformations, we further obtain

$$b^2(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) - 4c(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) =$$

$$= \Delta \varepsilon \left( 4 \text{tr}(S_i^{-1} \tilde{A}_{11}^0 S_i S_i^{-1} \tilde{A}_{11}^0 S_i) - 2 \text{tr}(S_i^{-1} \tilde{A}_{11}^0 S_i) \text{tr}(S_i^{-1} \tilde{A}_{11}^0 S_i) \right) +$$

$$+ \Delta \Omega \left( 4 \text{tr}(S_i^{-1} \tilde{A}_{11}^0 S_i S_i^{-1} \tilde{A}_{11}^0 S_i) - 2 \text{tr}(S_i^{-1} \tilde{A}_{11}^0 S_i) \text{tr}(S_i^{-1} \tilde{A}_{11}^0 S_i) \right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^2) =$$

$$= \Delta \varepsilon \left( 4 \text{tr}((\lambda_c I + \delta N)S_i^{-1} \tilde{A}_{11}^0 S_i) - 2 \text{tr}(\lambda_c I + \delta N) \text{tr}(S_i^{-1} \tilde{A}_{11}^0 S_i) \right) +$$

$$+ \Delta \Omega \left( 4 \text{tr}((\lambda_c I + \delta N)S_i^{-1} \tilde{A}_{11}^0 S_i) - 2 \text{tr}(\lambda_c I + \delta N) \text{tr}(S_i^{-1} \tilde{A}_{11}^0 S_i) \right) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^2) =$$
\[
\Delta \varepsilon \left( 4\lambda_c \text{tr}(S^{-1}_1 \tilde{A}^{10}_{11} S_1) + 4\delta \text{tr}(NS^{-1}_1 \tilde{A}^{10}_{11} S_1) - 4\lambda_c \text{tr}(S^{-1}_1 \tilde{A}^{10}_{11} S_1) \right) + \\
\Delta \Omega \left( 4\lambda_c \text{tr}(S^{-1}_1 \tilde{A}^{01}_{11} S_1) + 4\delta \text{tr}(NS^{-1}_1 \tilde{A}^{01}_{11} S_1) - 4\lambda_c \text{tr}(S^{-1}_1 \tilde{A}^{01}_{11} S_1) \right) + \\
O((|\Delta \varepsilon| + |\Delta \Omega|)^2) = \\
4 \Delta \varepsilon \delta \text{tr}(NS^{-1}_1 \tilde{A}^{10}_{11} S_1) + 4 \Delta \Omega \delta \text{tr}(NS^{-1}_1 \tilde{A}^{01}_{11} S_1) + O((|\Delta \varepsilon| + |\Delta \Omega|)^2). 
\]

This leads to the following conclusions:

- In the semisimple case the linear parts of the expansion of the equation of the stability boundary vanish.
- In the non-semisimple case it depends on the matrices \( \tilde{A}^{10}_{11}, \tilde{A}^{01}_{11} \) respectively, whether the linear parts vanish.

Now the rest of the proof is identical with the proof of Theorem 4.2.1. \( \square \)

### 5.4 Four Dimensional Systems

We close this chapter with a short consideration on the four dimensional case. The restriction on four dimensions allows us to combine the results of the previous sections.

Let
\[
\tilde{A}_c(\varepsilon, \Omega) := \begin{pmatrix} \tilde{A}_{11}(\varepsilon, \Omega) & * \\ 0 & \tilde{A}_{22}(\varepsilon, \Omega) \end{pmatrix}
\]
and
\[
B(\varepsilon, \Omega) = \frac{1}{2} \left( \text{tr}^2 \tilde{A}_c - \text{tr}((\tilde{A}_c)^2) \right),
\]
\[
C(\varepsilon, \Omega) = \det \tilde{A}_c.
\]

(cf. Section 5.3).

First we recall the equations (5.20) for the stability boundary in the case of a Krein collision
\[
B^2(\varepsilon, \Omega) - 4C(\varepsilon, \Omega) = 0.
\]

On the one hand if we insert \( \rho = 0 \) into the reduced characteristic equation (5.2)
\[
Q(\rho) = \rho^2 - \rho B(\varepsilon, \Omega) + C(\varepsilon, \Omega) = 0,
\]
we obtain
\[
C(\varepsilon, \Omega) = 0. 
\]

(5.23)

On the other hand if (5.23) holds, then the reduced characteristic equation is simplified to
\[
0 = \rho^2 + \rho B = \rho(\rho - B).
\]
Thus $\rho = 0$ is a reduced eigenvalue. Moreover for $B \neq 0$ it is simple and for $B < 0$ the second reduced eigenvalue does not lie on the non-positive real axis.

We concluded that (5.23) describes the stability boundary in the case of a collision at 0.

Since the coefficients $B(\varepsilon, \Omega)$ and $C(\varepsilon, \Omega)$ are real (cf. Lemma 2.1.2 and Lemma 2.1.4), the left-hand sides of the equations (5.20) and (5.23) are real. This allows us to represent the solutions in a $B$-$C$-coordinate system.

The curves are the solutions of (5.20) and (5.23). The small diagrams show qualitatively the situation of the eigenvalues of $\tilde{A}_c(\varepsilon, \Omega)$. We see that only in the shaded area enclosed by the parabola and the straight lines the eigenvalues of $\tilde{A}_c(\varepsilon, \Omega)$ are on the unit circle.
5. Autonomous Systems
A Comparison between the Method of Lyapunov-Schmidt and the Normal Form Method

Some years ago F. Meyer and U. Kirchgraber investigated the stability problem (cf. [18], [21]). They chose a completely different way to reduce the dimension of the problem: the so-called Lyapunov-Schmidt reduction. They treated the dissipative case completely, but in the case of canonical systems they did not succeed in proving the existence of the stability boundary. We compare the Lyapunov-Schmidt method with the normal form method to show why the normal form method is more successful.

The chapter is organized as follows:

Section 6.1: We present a short sketch of the Lyapunov-Schmidt reduction. The goal is to give an idea of the structure of the resulting equation rather than a complete derivation. We therefore omit the proofs. A general presentation of the method of Lyapunov-Schmidt may be found in [5].

Section 6.2: We compare the method of Lyapunov-Schmidt with the normal form method and discuss why the normal form approach is more successful.

6.1 The Lyapunov-Schmidt Reduction

We treat the case of a canonical or reversible system of differential equations. Then the spectrum or the monodromy matrix is symmetric with respect to the unit circle. We therefore assume that for \((\varepsilon, \Omega) = (\varepsilon_0, \Omega_0)\) the monodromy matrix \(M(\varepsilon, \Omega)\) has a double non-semisimple eigenvalue \(\mu^0\).

Let \(C(\varepsilon, \Omega, \mu) := M(\varepsilon, \Omega) - \mu I\). Every eigenvector of \(M(\varepsilon_0, \Omega_0)\) belonging to the eigenvalue \(\mu^0\) is a nontrivial solution of \(C_0 x := C(\varepsilon_0, \Omega_0, \mu^0) x = 0\). Now we are looking for solutions of \(C(\varepsilon, \Omega, \mu)x = 0\) near \((\varepsilon_0, \Omega_0, \mu^0)\). To this end we rewrite this equation as

\[
0 = C(\varepsilon, \Omega, \mu)x = C_0 x - (C_0 - C(\varepsilon, \Omega, \mu))x =: C_0 x - \Delta C(\varepsilon, \Omega, \mu)x
\]
or
\[ C_0 x = \Delta C(\varepsilon, \Omega, \mu) x, \] (6.1)
respectively. Note that \( \Delta C(\varepsilon_0, \Omega_0, \mu^0) = 0 \).

As stated above the eigenvectors of \( M(\varepsilon_0, \Omega_0) \) belonging to the eigenvalue \( \mu^0 \) are non-trivial solutions of \( C_0 x = 0 \). Therefore \( C_0 \) is not invertible, but we may decompose \( \mathbb{C}^n \) into two subspaces such that the restriction of \( C_0 \) to one of these subspaces is invertible. This allows us to reduce the dimension of the problem from \( n \) to \( 2 \).

### The Decomposition of the Space

Let \( \{a_1, \ldots, a_n\} \) be the set of the generalized right eigenvectors of \( M(\varepsilon_0, \Omega_0) \). Without loss of generality we assume that \( \{a_1, a_2\} \) span the generalized right eigenvectorspace belonging to the eigenvalue \( \mu^0 \). We define
\[
\mathcal{L} := \{a_1, a_2\},
\]
\[
\mathcal{D} := \{a_3, \ldots, a_n\}.
\]
It is well known that \( \mathbb{C}^n = \mathcal{L} \oplus \mathcal{D} \). Obviously \( \mathcal{L} \) and \( \mathcal{D} \) are invariant under \( C_0 \). Moreover the restriction of \( C_0 \) to \( \mathcal{D} \) is invertible. We denote the inverse by \( S_0 \).

We also introduce the set of generalized left eigenvectors \( \{b_1, \ldots, b_n\} \). Again we assume that \( \{b_1, b_2\} \) span the generalized left eigenvectorspace belonging to the eigenvalue \( \mu^0 \). Let
\[
\mathcal{L}^* := \{b_1, b_2\},
\]
\[
\mathcal{D}^* := \{b_3, \ldots, b_n\}.
\]
We may normalize the generalized eigenvectors such that they form a biorthonormal basis of \( \mathbb{C}^n \), i.e. \( b_i^* a_j = \delta_{ij} \), where \( \delta_{ij} \) denotes the well known Kronecker symbol. Thus \( \mathbb{C}^n = \mathcal{L}^* \oplus \mathcal{D}^* \) is an orthonormal decomposition of \( \mathbb{C}^n \).

Since \( \mathcal{L} \) and \( \mathcal{L}^* \) are isomorphic vector spaces we obtain a further decomposition \( \mathbb{C}^n = \mathcal{L}^* \oplus \mathcal{D} \).

### The Decomposition of the Equation

We introduce the projector \( P \) onto \( \mathcal{D} \) along \( \mathcal{L}^* \) by
\[
P : \mathbb{C}^n \rightarrow \mathcal{D}, \quad x := x_{\mathcal{L}^*} + x_{\mathcal{D}} \mapsto x_{\mathcal{D}}.
\]
It follows easily that \( P \) is linear and idempotent, i.e. \( P \) is a projector, indeed. Furthermore \( P \) and \( C_0 \) commute: \( PC_0 = C_0 P \).

Now we decompose equation (6.1) by means of the projector \( P \):
\[
C_0 x = \Delta C(\varepsilon, \Omega, \mu) x \iff P C_0 x = P \Delta C(\varepsilon, \Omega, \mu) x \quad (6.1_\mathcal{D})
\]
\[
(I - P) C_0 x = (I - P) \Delta C(\varepsilon, \Omega, \mu) x \quad (6.1_\mathcal{L})
\]
By equation (6.1-p) we may express $x$ by its restriction to $\mathcal{L}$:

$$PC_0x = P\Delta C(\varepsilon, \Omega, \mu)x$$

$$\Rightarrow \quad C_0Px = P\Delta C(\varepsilon, \Omega, \mu)x$$

$$\Rightarrow \quad Px = S_0P\Delta C(\varepsilon, \Omega, \mu)x \quad \text{with} \quad S_0 := (C_0|_D)^{-1}$$

$$\Rightarrow \quad (I-P)x = (I-S_0P\Delta C(\varepsilon, \Omega, \mu))x$$

$$\Rightarrow \quad x = (I-S_0P\Delta C(\varepsilon, \Omega, \mu))^{-1}(I-P)x. \quad (6.2)$$

In the last step we used that $\Delta C(\varepsilon, \Omega, \mu)$ is small for $(\varepsilon, \Omega, \mu)$ close to $(\varepsilon_0, \Omega_0, \mu_0)$. Now we transform (6.1$_x$) using $(I-P)^2 = I-P$ and $C_0(I-P) = (I-P)C_0$:

$$(I-P)C_0x = (I-P)\Delta C(\varepsilon, \Omega, \mu)x$$

$$\Rightarrow \quad (I-P)^2C_0x = (I-P)\Delta C(\varepsilon, \Omega, \mu)x$$

$$\Rightarrow \quad (I-P)C_0(I-P)x = (I-P)\Delta C(\varepsilon, \Omega, \mu)x \quad (6.3)$$

Finally combining (6.2) and (6.3) we obtain the so-called bifurcation equation

$$(I-P)C_0(I-P)x = (I-P)\Delta C(\varepsilon, \Omega, \mu)(I-S_0P\Delta(\varepsilon, \Omega, \mu))^{-1}(I-P)x$$

or

$$(I-P)(C_0 - \Delta C(\varepsilon, \Omega, \mu)(I-S_0P\Delta(\varepsilon, \Omega, \mu))^{-1}) (I-P)x = 0, \quad (6.4)$$

respectively. Thus we have sketched the proof of the following lemma.

**Lemma 6.1.1**

For $(\varepsilon, \Omega, \mu)$ sufficiently close to $(\varepsilon_0, \Omega_0, \mu_0)$ the following equations are equivalent:

(i) $(I-P)C_0x = (I-P)\Delta C(\varepsilon, \Omega, \mu)x$

(ii) $(I-P)(C_0 - \Delta C(\varepsilon, \Omega, \mu)(I-S_0P\Delta(\varepsilon, \Omega, \mu))^{-1})(I-P)x = 0.$

Since $(I-P)x$ lies in $\mathcal{L}$ we may write $(I-P)x = \xi_1a_1 + \xi_2a_2$ with $\xi_1, \xi_2 \in \mathbb{C}$. For $i = 1, 2$ the generalized lefteigenvectors are orthonormal to $\mathcal{D}$. Therefore we have

$$b_i^*y = b_i^*((Py + (I-P)y)) = b_i^*Py + b_i^*(I-P)y = b_i^*(I-P)y \quad \text{for} \quad i = 1, 2.$$

Using these two arguments we obtain from Lemma 6.1.1 for $i = 1, 2$

$$0 = b_i^*(I-P)(C_0 - \Delta C(\varepsilon, \Omega, \mu)(I-S_0P\Delta(\varepsilon, \Omega, \mu))^{-1})(I-P)x =$$

$$= b_i^*(C_0 - \Delta C(\varepsilon, \Omega, \mu)(I-S_0P\Delta(\varepsilon, \Omega, \mu))^{-1})(\xi_1a_1 + \xi_2a_2) =$$
Thus we have a system of two linear equations for two unknown variables $\xi_1$ and $\xi_2$. This system has a non-trivial solution if and only if the determinant vanishes. This proves one part of the following theorem.

**Theorem 6.1.1**

For $(\epsilon, \Omega, \mu)$ sufficiently close to $(\epsilon_0, \Omega_0, \mu^0)$ the following equations are equivalent:

$$C_0 x = \Delta C(\epsilon, \Omega, \mu)x, \quad x \neq 0$$

$$\det \begin{pmatrix} N_{11}(\epsilon, \Omega, \mu) & N_{12}(\epsilon, \Omega, \mu) \\ N_{21}(\epsilon, \Omega, \mu) & N_{22}(\epsilon, \Omega, \mu) \end{pmatrix} = 0.$$  

We conclude: A point $(\epsilon, \Omega)$ of the parameter plane close to $(\epsilon_0, \Omega_0)$ lies on the stability boundary if and only if equation (6.5) has a double zero $\mu$.

### 6.2 The Comparison

We are now in the position to discuss the major difference between the method of Lyapunov-Schmidt and the normal form method.

**The Lyapunov-Schmidt-Method**

The Lyapunov-Schmidt reduction provides equation (6.5) for the critical eigenvalues of $M(\epsilon, \Omega)$ for $(\epsilon, \Omega)$ close to $(\epsilon_0, \Omega_0)$. This equation is *highly nonlinear* in $\mu$.

Since $\mu^0$ is a double zero for $(\epsilon, \Omega) = (\epsilon_0, \Omega_0)$ we may not apply the theorem of implicit functions to guarantee further zeroes in the neighborhood of $\mu^0$. We are also not in the position to derive a condition for double zeroes to which we may apply the Implicit Function Theorem. Thus we have no proof of the existence of the stability boundary.

**The Normal Form Method**

The normal form reduction provides a *quadratic* equation in $\mu$ with coefficients depending on $(\epsilon, \Omega)$. We may solve this equation easily and give explicitly a simple condition for the existence of double zeroes. Applying the Implicit Function Theorem to this condition we are able to obtain the existence of the stability boundary.
Conclusions

The main difference between the two methods is that the normal form method provides a \textit{quadratic} equation which may be treated easily, while the method of Lyapunov-Schmidt provides a \textit{highly nonlinear} equation which allows no further conclusions.
Seite Leer / Blank leaf
In this chapter we describe the design of an efficient algorithm to compute the stability boundary of the differential equation (2.1).

We develop the algorithm for the non-autonomous case. The autonomous case is treated similarly. In that case the computation of the monodromy matrix may be omitted and a slightly different equation for the stability boundary is used. We also concentrate on the canonical or reversible case. The dissipative case is handled similarly.

Let us assume throughout this chapter that we only have one complex conjugate pair of double eigenvalues, which may cause a change of stability.

Starting points are the Theorems 4.3.1 and 4.3.2 respectively, that not only guarantee the existence of the stability boundary but also provide equations for its computation.

Let $(e_0, \Omega_0)$ be our initial point close to the stability boundary. We perform the following steps:

1. We compute the monodromy matrix $M(e_0 + \Delta e, \Omega_0 + \Delta \Omega)$ up to order two in $\Delta e$ and $\Delta \Omega$.

2. We transform $\hat{M}(e_0 + \Delta e, \Omega_0 + \Delta \Omega)$ to normal form up to order two and we compute the expansion of equation (4.16) or (4.21), respectively.

3. We perform a correction and a prediction step to obtain an approximative point of the stability boundary and a new initial point, respectively.

The steps 1, 2 and 3 as described below can be performed independent of whether the initial point $(e_0, \Omega_0)$ lies on the stability boundary or only near to it. The resulting equation is an approximate condition for a double eigenvalue of the monodromy matrix, anyway, and therefore an approximate condition for the stability boundary. This equation is solved in order to find a point very close to the stability boundary.

The chapter is organized as follows:

**Section 7.1:** We discuss the problem of computing the monodromy matrix of a system of differential equations depending on two parameters.

**Section 7.2:** We consider the transformation of the monodromy matrix to normal form.
7. Tracing the Stability Boundary

Section 7.3: We outline the basic steps of a tracing algorithm for the stability boundary.

7.1 The Computation of the Monodromy Matrix

Since the matrix $A(t, \varepsilon, \Omega)$ in the differential equation (2.1) depends on parameters, we can not apply numerical integration methods directly to obtain the monodromy matrix $M(\varepsilon, \Omega)$. We first expand $A(t, \varepsilon, \Omega)$ at $(\varepsilon_0, \Omega_0)$:

$$A(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = A^0 + \Delta \varepsilon A^{10} + \Delta \Omega A^{01} +$$

$$+ \Delta \varepsilon^2 A^{20} + \Delta \varepsilon \Delta \Omega A^{11} + \Delta \Omega^2 A^{02} +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$

Let $X(t, \varepsilon, \Omega)$ denote the principal matrix solution of (2.1). The expansion of $X(t, \varepsilon, \Omega)$ reads

$$X(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = X^0 + \Delta \varepsilon X^{10} + \Delta \Omega X^{01} +$$

$$+ \Delta \varepsilon^2 X^{20} + \Delta \varepsilon \Delta \Omega X^{11} + \Delta \Omega^2 X^{02} +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$

If we insert this expansion into (2.1), then we obtain for the left-hand side

$$\dot{X}(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = \dot{X}^0 + \Delta \varepsilon \dot{X}^{10} + \Delta \Omega \dot{X}^{01} +$$

$$+ \Delta \varepsilon^2 \dot{X}^{20} + \Delta \varepsilon \Delta \Omega \dot{X}^{11} + \Delta \Omega^2 \dot{X}^{02} +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$

and for the right-hand side

$$A(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega)X(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) =$$

$$= A^0 X^0 + \Delta \varepsilon (A^0 X^{10} + A^{10} X^0) +$$

$$+ \Delta \Omega (A^0 X^{01} + A^{01} X^0) +$$

$$+ \Delta \varepsilon^2 (A^0 X^{20} + A^{10} X^{10} + A^{20} X^0) +$$

$$+ \Delta \varepsilon \Delta \Omega (A^0 X^{11} + A^{10} X^{01} + A^{01} X^{10} + A^{11} X^0) +$$

$$+ \Delta \Omega^2 (A^0 X^{02} + A^{01} X^{01} + A^{02} X^0) +$$

$$+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).$$
Comparing coefficients we obtain the following sequence of systems of differential equations:

\[
\begin{align*}
X^0 &= A^0(t)X^0, & X^0(0) &= I, \\
\dot{X}^{10} &= A^0(t)X^{10} + A^{10}(t)X^0, & X^{10}(0) &= 0, \\
\dot{X}^{01} &= A^0(t)X^{01} + A^{01}(t)X^0, & X^{01}(0) &= 0, \\
\dot{X}^{20} &= A^0(t)X^{20} + A^{10}(t)X^{10} + A^{20}(t)X^0, & X^{20}(0) &= 0, \\
\dot{X}^{11} &= A^0(t)X^{11} + A^{01}(t)X^{10} + A^{10}(t)X^{01} + A^{11}(t)X^0, & X^{11}(0) &= 0, \\
\dot{X}^{02} &= A^0(t)X^{02} + A^{01}(t)X^{01} + A^{02}(t)X^0, & X^{02}(0) &= 0.
\end{align*}
\]

In order to obtain the monodromy matrix, we have to integrate these differential equations numerically over one period $T$. There are a variety of methods to do that, but there is no method which is efficient and reliable in all situations. We refer to the Ph.D.-Thesis of Meyer [18], where he discusses this problem and refers to further literature.

### 7.2 The Transformation to Normal Form

Assume that we have an expansion of the monodromy matrix up to order two with respect to $\varepsilon$, $\Omega$ at some point $(\varepsilon_0, \Omega_0)$. We first transform $M^0$ to upper block triangular form. This can be done efficiently by the so-called Schur decomposition. If necessary we may rearrange the matrix in such a way that the critical eigenvalues appear in the first block on the diagonal.

Assume that this transformation has been performed. For simplicity we denote the resulting matrix again by $M(\varepsilon, \Omega)$. Let

\[
M(\varepsilon, \Omega) = \begin{pmatrix} M_{11}(\varepsilon, \Omega) & M_{12}(\varepsilon, \Omega) \\ M_{21}(\varepsilon, \Omega) & M_{22}(\varepsilon, \Omega) \end{pmatrix}
\]

with

\[
M_{ij}(\varepsilon, \Omega) = M_{ij}^0 + \Delta \varepsilon M_{ij}^{10} + \Delta \Omega M_{ij}^{01} + \\
+ \Delta \varepsilon^2 M_{ij}^{20} + \Delta \varepsilon \Delta \Omega M_{ij}^{11} + \Delta \Omega^2 M_{ij}^{02} + \\
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3) \quad \text{for} \quad i, j = 1, 2,
\]

and $M_{21}^0 = 0$. Let

\[
\Phi(\varepsilon, \Omega) = \begin{pmatrix} I & 0 \\ \phi(\varepsilon, \Omega) & I \end{pmatrix}
\]
with
\[
\phi(\varepsilon, \Omega) = \Delta \varepsilon \phi^{10} + \Delta \Omega \phi^{01} + \Delta \varepsilon^2 \phi^{20} + \Delta \varepsilon \Delta \Omega \phi^{11} + \Delta \Omega^2 \phi^{02} + O((|\Delta \varepsilon| + |\Delta \Omega|)^3).
\]
Inserting this into equation (3.2) yields
\[
0 = M_{21} - \phi M_{11} + M_{22} \phi - \Omega M_{12} \phi =
- \Delta \varepsilon (\phi^{10} M_{11}^0 - M_{22}^0 \phi^{10}) - M_{21}^{10}
- \Delta \Omega (\phi^{01} M_{11}^0 - M_{22}^0 \phi^{01}) - M_{21}^{01}
- \Delta \varepsilon^2 (\phi^{20} M_{11}^0 - M_{22}^0 \phi^{20}) +
+ (\phi^{10} M_{11}^0 - M_{22}^0 \phi^{10}) + \phi^{10} M_{12}^0 \phi^{10} - M_{21}^{20}
- \Delta \varepsilon \Delta \Omega ((\phi^{11} M_{11}^0 - M_{22}^0 \phi^{11}) + (\phi^{10} M_{11}^0 - M_{22}^0 \phi^{01}) +
+ (\phi^{01} M_{11}^0 - M_{22}^0 \phi^{01}) + (\phi^{10} M_{12}^0 \phi^{01} + \phi^{01} M_{12}^0 \phi^{10}) - M_{21}^{11}) -
- \Delta \Omega^2 ((\phi^{02} M_{11}^0 - M_{22}^0 \phi^{02}) +
+ (\phi^{01} M_{11}^0 - M_{22}^0 \phi^{01}) + \phi^{01} M_{12}^0 \phi^{01} - M_{21}^{02}) -
+ O((|\Delta \varepsilon| + |\Delta \Omega|)^3).
\]
Thus we have to solve the following equations
\[
\phi^{ij} M_{11}^0 - M_{22}^0 \phi^{ij} = C^{ij}
\]
with
\[
C^{10} := M_{21}^{10},
C^{01} := M_{21}^{01},
C^{20} := M_{21}^{20} - (\phi^{10} M_{11}^0 - M_{22}^0 \phi^{10}) - \phi^{10} M_{12}^0 \phi^{10},
C^{11} := M_{21}^{11} - (\phi^{10} M_{11}^0 - M_{22}^0 \phi^{10}) - (\phi^{01} M_{11}^0 + M_{22}^0 \phi^{01}) -
- (\phi^{10} M_{12}^0 \phi^{01} + \phi^{01} M_{12}^0 \phi^{10}),
C^{02} := M_{21}^{02} - (\phi^{01} M_{11}^0 - M_{22}^0 \phi^{01}) - \phi^{01} M_{12}^0 \phi^{01}.
\]
Note that $C^{20}$, $C^{11}$ and $C^{02}$ depend on $\phi^{10}$ and $\phi^{01}$.

Let
\[
\Phi^{-1}(\varepsilon, \Omega) M(\varepsilon, \Omega) \Phi(\varepsilon, \Omega) = \tilde{M}(\varepsilon, \Omega) = \begin{pmatrix}
\tilde{M}_{11}(\varepsilon, \Omega) & * \\
0 & \tilde{M}_{22}(\varepsilon, \Omega)
\end{pmatrix}
\]
7.3. Tracing the Stability Boundary

Let \((\varepsilon_0, \Omega_0)\) be a point near the stability boundary, i.e. in some neighborhood \(U_0\) of \((\varepsilon_0, \Omega_0)\) there exists a point \((\varepsilon_1, \Omega_1)\) for which the assumptions of Theorem 4.3.1 or Theorem 4.3.2, respectively, hold. Then in some neighborhood \(U_1\) of \((\varepsilon_1, \Omega_1)\) the stability boundary is described by

\[
F(\varepsilon, \Omega) := 2 \text{tr}((\bar{M}_c(\varepsilon, \Omega))^2) - \text{tr}^2 \bar{M}_c(\varepsilon, \Omega) + 8 = 0
\]

or

\[
F(\varepsilon, \Omega) := \text{tr}((\bar{M}_{11}(\varepsilon, \Omega))^2) \mp 2 = 0,
\]

respectively.

Now let \(U_2\) be a neighborhood of \((\varepsilon_0, \Omega_0)\) in which the expansion of the monodromy matrix exists and the normal form transformation may be performed. Then in \(U := U_1 \cap U_2\) the stability boundary is described by the expansion

\[
2F(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) := F^0 + \Delta \varepsilon F^{10} + \Delta \Omega F^{01} + \Delta \varepsilon^2 F^{00} + \Delta \varepsilon \Delta \Omega F^{11} + \Delta \Omega^2 F^{02} + O((|\Delta \varepsilon| + |\Delta \Omega|)^3).
\]
Let
\[ F_1(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) := F_0^0 + \Delta \varepsilon F_1^0 + \Delta \Omega F_0^0, \]
\[ F_2(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) := F_0^0 + \Delta \varepsilon F_1^0 + \Delta \Omega F_0^0 + \Delta \varepsilon^2 F_2^0 + \Delta \varepsilon \Delta \Omega F_1^1 + \Delta \Omega^2 F_0^0 \]
be its truncations up to order one and two, respectively.

We now develop a predictor-corrector algorithm for tracing the stability boundary. Let us assume for the moment that \( F_1^0 \) and \( F_0^1 \) do not vanish simultaneously. Then the solutions of
\[ F_1(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = 0 \]
lie on a straight line \( C_{F_1} \), while the solutions of
\[ F_2(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = 0 \]
lie on a conic section \( C_{F_2} \) as illustrated in Figure 7.1.

**The Corrector Step**

For the corrector step we draw a straight line perpendicular to \( C_{F_1} \). It intersects the curves \( C_{F_1} \) and \( C_{F_2} \) in \((\varepsilon_1, \Omega_1)\) and \((\varepsilon_{2,3}, \Omega_{2,3})\), respectively. The situation is illustrated in Figure 7.2.

\((\varepsilon_1, \Omega_1)\) is the first approximation to the point on the stability boundary. Since we consider the second approximation as a correction to the first, we choose that point on \( C_{F_2} \) that lies next to \((\varepsilon_1, \Omega_1)\). To be more precise: We compute the distances \( d_{2,3} \) between \((\varepsilon_1, \Omega_1)\) and \((\varepsilon_{2,3}, \Omega_{2,3})\), respectively. Then we distinguish three cases:
7.3. Tracing the Stability Boundary

![Diagram showing the situation of the corrector step.](image)

Fig. 7.2: The situation of the corrector step.

- If \( \frac{d_2}{d_3} \) is bigger than some positive constant \( c_1 > 1 \), then we choose the point with the coordinates \((\varepsilon_3, \Omega_3)\).
- If \( \frac{d_2}{d_3} \) is smaller than \( 1/c_1 \), then we choose \((\varepsilon_2, \Omega_2)\).
- If \( \frac{d_2}{d_3} \) is between \( 1/c_1 \) and \( c_1 \), then the quality of the approximation is not good enough. This is a hint that we may be close to an intersection point of two branches of the stability boundary. We treat this case below.

The algorithmic realization is illustrated by a flow chart in Figure 7.10 at the end of this chapter.

The Predictior Step

For the predictor step we proceed similarly as for the corrector step, but we start with a step tangential to the stability boundary: We compute a point \((\varepsilon'_1, \Omega'_1)\) on \(C_{F_1}\) at some distance \(\Delta s\) away from \((\varepsilon_1, \Omega_1)\), as illustrated in Figure 7.3. Then we draw the straight line perpendicular to \(C_{F_1}\) through this point and compute the intersection points with \(C_{F_3}\) and go on as before.

- If \( \frac{d'_2}{d'_3} \) is bigger than some positive constant \( c_2 > 1 \) or smaller than \( 1/c_2 \) then we may double the step size \(\Delta s\) and repeat the computations.
- If \( \frac{d'_2}{d'_3} \) lies between \( c_3 \) and \( c_2 \) or between \( 1/c_2 \) and \( 1/c_3 \) for some positive constant \( 1 < c_3 < c_2 \) then choose \((\varepsilon'_3, \Omega'_3)\) or \((\varepsilon'_2, \Omega'_2)\), respectively.
- If \( \frac{d'_2}{d'_3} \) is smaller than \( c_3 \) or bigger than \( 1/c_3 \) then we halve the step size.

For \( c_2 < c_1 \) this process certainly terminates, since the corrector step was successful, but \(\Delta s\) may become very small. This is again a hint that we may be close to the intersection point of two branches of the stability boundary.

The flow chart of the predictor step is illustrated in Figure 7.11 at the end of this chapter.
Intersection Points

If the corrector step fails or if in the prediction step the step size becomes very small, then we typically have one of the following situations

**Situation A**

Two branches of the stability boundary intersect in a point \((\tilde{\varepsilon}, \tilde{\Omega})\) (cf. Figure 7.4). Such a point is called intersection or bifurcation point. If we expand \(F(\varepsilon, \Omega)\) at this point, then the constant and the linear parts vanish (cf. Theorems 4.3.1 and 4.3.2). That means, that the following equations hold:

\[
F(\varepsilon, \Omega) = 0, \quad (7.1)
\]

\[
F_\varepsilon(\varepsilon, \tilde{\Omega}) = 0,
\]

\[
F_\Omega(\varepsilon, \tilde{\Omega}) = 0. \quad (7.2)
\]

**Situation B**

We have no intersection point (cf. Figure 7.5). There still may be some point \((\varepsilon', \Omega')\), for
7.3. Tracing the Stability Boundary

Fig. 7.5: Situation B: There is no intersection point.

which Eqs. (7.2) hold, but for this point Eq. (7.1) is no longer fulfilled.

It is impossible to predict which situation we will find. This is immediately clear, if we consider the graph of $F(\varepsilon, \Omega)$ shown in Figure 7.6.

In both situations we have a saddle point. In situation A the height of the saddle is 0. Thus the level curves of the height 0 intersect in the saddle point. But under a small perturbation of the function $F$ the height of the saddle changes and the level curves of height 0 no longer intersect, as in situation B.

Since we do not know in advance, which situation occurs, we treat both in the same way. First we solve (7.2) and then check whether (7.1) is fulfilled or not.

Let us consider Eqs. (7.2). Since we know the expansion of equation (7.1) at $(\varepsilon_0, \Omega_0)$ up to order two, we may compute the expansions of (7.2) up to order one:

$$F_\varepsilon(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = F^{10} + 2\Delta \varepsilon F^{20} + \Delta \Omega F^{11} + O((|\Delta \varepsilon| + |\Delta \Omega|)^2) = 0,$$

$$F_\Omega(\varepsilon_0 + \Delta \varepsilon, \Omega_0 + \Delta \Omega) = F^{01} + \Delta \varepsilon F^{11} + 2\Delta \Omega F^{02} + O((|\Delta \varepsilon| + |\Delta \Omega|)^2) = 0.$$
The solution of the truncated system is

\[ \Delta \varepsilon = -\frac{2F^{10}F^{02} - F^{01}F^{11}}{4F^{20}F^{02} - (F^{11})^2}, \]

\[ \Delta \Omega = -\frac{2F^{01}F^{20} - F^{10}F^{11}}{4F^{20}F^{02} - (F^{11})^2}, \]

provided \( 4F^{20}F^{02} - (F^{11})^2 \) does not vanish. This allows us to compute the solution of (7.2) iteratively. Then we check, whether (7.1) is fulfilled or not. The further procedure depends on the situation encountered:

**Situation A:** If we have a bifurcation point, then we may leave this point tangential to one of the branches. The tangential directions are easily computed using the coefficients \( F^{ij} \).

**Situation B:** If we have no bifurcation point, then we have to proceed with very small steps.

If the maximum step size is chosen too large the algorithm may not detect such critical situations. It then may \"overlook\" intersection points and possibly even switch to another branch.

In addition to the situations just described intersection points may also occur in a genuinely a different case. We treat this case in the following subsection.

**Situation C**

Let \( \tilde{M}(\varepsilon_0, \Omega_0) \) denote the block triangular form of the monodromy matrix. So far we assumed that only one single block in the diagonal has critical eigenvalues. Let us now assume that two blocks in the diagonal have critical eigenvalues. Then for both these blocks there exists locally a smooth curve corresponding to critical eigenvalues. These curves intersect at \((\varepsilon_0, \Omega_0)\). This situation is sketched in Figure 7.7. The stable region is shaded. The small icons indicate qualitatively the situation of the interesting eigenvalues.

**Fig. 7.7:** Situation C: Several eigenvalues are critical.

While in Situation A the intersection point arises from the degeneration of the equation that determines the stability boundary here the intersection point arises from critical eigenvalues of different block matrices.
To keep track of stability we therefore compute the characteristic equations for every block in the diagonal of the normal form of the monodromy matrix. After every predictor step we check whether stability is lost. If this is the case we determine the block which causes the loss of stability. This block provides the equation of the stability boundary for the next corrector and predictor step. Before we compute the intersection point. This point serves as a new initial point for tracing the new branch of the stability boundary. The flow chart of the whole algorithm that detects and computes intersection points is shown in Figure 7.12 at the end of this chapter.

The Tracing Algorithm

The flow chart in Figure 7.9 gives an overview on the subroutine GetNextPoint that computes the next point on the stability boundary.

The subroutine GetEquation provides the current equation of the stability boundary. To this end it calls subroutines for computing the monodromy matrix and for transforming into normal form, as discussed in the first two sections of this chapter.

GetNextPoint calls the subroutines Corrector, Predictor and IntersectionPoint discussed above in detail.

It controls the step size and appends continuously the next points of the stability boundary to a list called StabPointList. If intersection points are detected, they are appended together with their tangential directions to a list called IntersectionPointList.

The main program then may be kept very short. After some initialization the subroutine GetNextPoint is called iteratively a certain number of times determined by the user (cf. Figure 7.8).
Fig. 7.9: The flow chart of GetNextPoint.
7.3. Tracing the Stability Boundary

Fig. 7.10: The flow chart of the corrector step.
7. Tracing the Stability Boundary

$(e_0, \Omega_0)$, tangential direction

compute the coefficients of the quadratic approximation of the stability boundary

the quadratic approximation is degenerated

no

compute $(e'_1, \Omega'_1)$

compute $(e'_2, \Omega'_2)$

compute $(e'_3, \Omega'_3)$

compute $d'_1, d'_2$

$\frac{1}{c_2} \leq \frac{d'_1}{d'_2} \leq c_2$

$\Delta s < \frac{1}{2} \Delta s_{\text{max}}$

yes

double step size $\Delta s$

no

$\frac{1}{c_3} \leq \frac{d'_2}{d'_3} \leq c_3$

no

halve step size $\Delta s$

yes

$\frac{c_3}{d'_3} < \frac{d'_2}{d'_3} < c_2$

no

$i = 2$

$i = 3$

$(e_1, \Omega_1)$, tangential direction

Fig. 7.11: The flow chart of the predictor step.
7.3. Tracing the Stability Boundary

\((\varepsilon_0, \Omega_0)\), tangential direction

- yes: stability is lost
- no: compute the coefficients of the quadratic approximation of the stability boundary

- determine which block causes instability
- compute corresponding stability boundaries and their intersection point

- yes: compute the linear part of
  \[ F_1(\varepsilon_0 + \Delta \varepsilon, \Omega_n + \Delta \Omega) = 0 \]
  \[ F_0(\varepsilon_0 + \Delta \varepsilon, \Omega_n + \Delta \Omega) = 0 \]
  vanishes
- no: compute next approximation \((\varepsilon_0 + \Delta \varepsilon, \Omega_n + \Delta \Omega)\)

- yes: \( F(\varepsilon_0, \Omega_0) = 0 \)
- no: return empty list

- compute the tangential directions
- compute a point on one tangent at the distance \(\Delta s\) from the intersection point
- return intersection point, new initial point, and tangential directions

**Fig. 7.12:** The flow chart for the detection and computation of an intersection point.
7.4 The Stability Diagram of a Symplectic Map

In this section we illustrate the tracing of the stability boundary. We compare the result with the exact stability boundary obtained by scanning the parameter plane with a small step size.

Since we concentrate on the tracing, we do not start from a differential equation, but from a symplectic map. This also allows us to compute the exact stability boundary with moderate costs. Moreover we may choose an example which shows several interesting phenomena.

The Map

We define a map $M$ by

$$M : \mathbb{R}^4 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^4, (x; (\varepsilon, \Omega)) \longmapsto M(\varepsilon, \Omega)x,$$

where

$$M(\varepsilon, \Omega) = \begin{pmatrix}
1 & 0 & b_1 & b_2 \\
0 & 1 & b_2 & b_4 \\
c_1 & c_3 & 1 + b_1c_1 + b_2c_3 & b_2c_1 + b_4c_3 \\
c_3 & c_4 & b_1c_3 + b_2c_4 & 1 + b_2c_3 + b_4c_4
\end{pmatrix}.$$

It is easy to show that this map is symplectic, indeed. We choose for the coefficients

$$b_1(\varepsilon, \Omega) := 1 - \varepsilon - \Omega - \frac{1}{2}\varepsilon\Omega,$$

$$c_1(\varepsilon, \Omega) := -2\varepsilon + \Omega + \varepsilon\Omega - \Omega^2,$$

$$b_2(\varepsilon, \Omega) := \varepsilon\Omega,$$

$$c_3(\varepsilon, \Omega) := 1 + \varepsilon - \Omega^2,$$

$$b_4(\varepsilon, \Omega) := -1 + \Omega^2,$$

$$c_4(\varepsilon, \Omega) := 2 + \varepsilon\Omega.$$

The Stability Diagram

We investigate the stability of the map $M$ in a quadratic section of the parameter plane given by

$$-4 \leq \varepsilon, \Omega \leq 4.$$

In Figure 7.13 we show the curves of solution of the equations

$$B^2(\varepsilon, \Omega) - 4C(\varepsilon, \Omega) + 8 = 0 : \text{black},$$

$$-2B(\varepsilon, \Omega) + C(\varepsilon, \Omega) + 2 = 0 : \text{dark-gray},$$

$$+2B(\varepsilon, \Omega) + C(\varepsilon, \Omega) + 2 = 0 : \text{light-gray},$$

(7.3)
obtained by scanning the parameter plane with a small step size ($\Delta h = 0.02$) in direction of the $\varepsilon$-axis and the $\Omega$-axis. The small icons show qualitatively the situation of the eigenvalues. The region of stability is shaded. The boxed areas are magnified in Figures 7.14, 7.15, 7.16.

In Figure 7.17 finally we show the tracing of the stability boundary.

![Diagram of stability diagram](image)

**Fig. 7.13**: The solutions of the equations (7.3).

We found a maximum step size of $\Delta s_{\text{max}} = 0.2$ and constants $c_1 = 16$, $c_2 = 32$ and $c_3 = 8$ suitable for this problem. With larger maximum step sizes some of the intersection points may not be detected. Bigger values for the constants $c_i$ require more steps, smaller values may cause some problems finding the intersection points.
Fig. 7.14: Magnification of the near miss of the stability boundaries.

Fig. 7.15: Magnification of the intersection point of the stability boundary. At the intersection point the double eigenvalues are semisimple.
7.4. The Stability Diagram of a Symplectic Map

Fig. 7.16: Magnification of the intersection point. All eigenvalues are equal to 1.

Fig. 7.17: Tracing the stability boundary. The small dots mark the points computed by the tracing algorithm.
7. Tracing the Stability Boundary
Stability of a Simple Model of a Two Bladed Rotor

We return to the question of stability of the wind turbine. To this end we recall the equations of motion developed in Chapter 1:

\[ \dot{z}^I = A^I(t, \varepsilon, \Omega)z^I, \]
\[ \dot{z}^{II} = A^{II}(t, \varepsilon, \Omega)z^{II}, \]

where

\[ M^I := M_0^I + \varepsilon M_1^I = \begin{pmatrix} m & 0 \\ 0 & B \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \]
\[ M^{II} := M_0^{II} + \varepsilon M_1^{II}(t, \Omega) = \]
\[ = \begin{pmatrix} m & 0 & 0 \\ 0 & A & 0 \\ 0 & 0 & A \end{pmatrix} + \varepsilon \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{d}{2}(-1 + \cos(2\Omega t)) & \frac{d}{2}(-1 + \cos(2\Omega t)) \\ 0 & \frac{d}{2}(-1 + \cos(2\Omega t)) & \frac{d}{2}(-1 + \cos(2\Omega t)) \end{pmatrix}, \]
\[ N^I := N^I(t) := \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \]
\[ N^{II} := N^{II}(t) := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \Omega B \\ 0 & 0 & 0 \end{pmatrix}, \]
\[ P^I := P^I(t) := K_0^I = \begin{pmatrix} k_1 & -k_2 \\ -k_2 & k_3 \end{pmatrix}, \]
\[ P^{II} := P^{II}(t) := K_0^{II} = \begin{pmatrix} k_1 & -k_2 & 0 \\ -k_2 & k_3 & 0 \\ 0 & 0 & k_4 \end{pmatrix}. \]

The chapter is organized as follows:
Section 8.1: In the first section we give rough values for the moments of inertia and the stiffness coefficients.

Section 8.2: In the second section we apply the general methods developed in the previous chapters to prove that the stability boundary is locally a curve and to compute an approximation to this curve.

8.1 Estimating the Moments of Inertia and the Stiffness Coefficients

It turned out to be very difficult to obtain specific information about existing wind turbines. Thus we had to make several assumptions and rough estimates.

Figure 8.1 shows a model of a wind turbine composed of simple geometrical elements. In Table 8.1 we present a plausible example.

![Diagram of a wind turbine](image)

Fig. 8.1: Typical dimensions of a large wind turbine.
8.1. Estimating the Moments of Intertia and the Stiffness Coefficients

Tab. 8.1: Typical dimensions of a large wind turbine

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Height of the tower</td>
<td>100m</td>
</tr>
<tr>
<td>Radius of the tower</td>
<td>3m</td>
</tr>
<tr>
<td>Length of the nacelle</td>
<td>12m</td>
</tr>
<tr>
<td>Radius of the nacelle</td>
<td>1.5m</td>
</tr>
<tr>
<td>Mass of the nacelle</td>
<td>$10^5$kg</td>
</tr>
<tr>
<td>Length of the blades</td>
<td>70m</td>
</tr>
<tr>
<td>Mass of a blade</td>
<td>$2.6 \cdot 10^4$kg</td>
</tr>
</tbody>
</table>

The basic constants and the computations that follow may be found in every text book on mechanical engineering, e.g. [2].

The Stiffness Coefficients

In order to compute the stiffness coefficients of the steel tower we need an assumption on the wall thickness $d$. It turns out that $d = 0.03$m is a good choice. This choice leads to the following moment of inertia of area $I$ and the polar moment of inertia $I_p$:

$$ I = \frac{\pi}{4} \left( 3^4 - 2.97^4 \right) = 1.49m^4, $$

$$ I_p = 2I = 2.98m^4. $$

Let $E \approx 2.1 \cdot 10^{11} \frac{N}{m^2}$ be the elastic modulus and let $G \approx 0.81 \cdot 10^{11} \frac{N}{m^2}$ be the shear modulus of steel. Then flexual stiffness coefficients are given by

$$ k_1 = \frac{12EI}{h^3} = \frac{12 \cdot 2.1 \cdot 10^{11} \cdot 1.49}{100^3} \approx 37.55 \cdot 10^5 \frac{N}{m}, $$

$$ k_2 = \frac{6EI}{h^2} = \frac{6 \cdot 2.1 \cdot 10^{11} \cdot 1.49}{100^2} \approx 18.77 \cdot 10^7 \frac{N}{rad}, $$

$$ k_3 = \frac{4EI}{h} = \frac{4 \cdot 2.1 \cdot 10^{11} \cdot 1.49}{100} \approx 12.52 \cdot 10^9 \frac{Nm}{rad} $$

and the torsional stiffness coefficient by

$$ k_4 = \frac{GI_p}{h} = \frac{0.81 \cdot 10^{11} \cdot 2.89}{100} \approx 2.414 \cdot 10^9 \frac{Nm}{rad}. $$
The Moments of Inertia

To compute the moments of inertia we assume the nacelle to be a cylinder and the blades to be thin rods. This leads to the following results:

\[
A = \frac{m_{\text{nac}} l_{\text{nac}}^2}{12} + 2 \left( \frac{m_{\text{bl}} l_{\text{bl}}^2}{12} + d_1^2 m_{\text{bl}} \right) = \\
= \frac{10^5 \cdot 12^2}{12} + 2 \left( \frac{2.6 \cdot 10^4 \cdot 70^2}{12} + (8^2 + 37.5^2) \cdot 2.6 \cdot 10^4 \right) \approx \\
\approx 989 \cdot 10^5 \text{kgm}^2,
\]

\[
B = m_{\text{nac}} R_{\text{nac}}^2 + 2 \left( \frac{m_{\text{bl}} l_{\text{bl}}^2}{12} + d_2^2 m_{\text{bl}} \right) = \\
= 10^5 \cdot 1.5^2 + 2 \left( \frac{2.6 \cdot 10^4 \cdot 70^2}{12} + 37.5^2 \cdot 2.6 \cdot 10^4 \right) \approx \\
\approx 945 \cdot 10^5 \text{kgm}^2,
\]

\[
C = \frac{m_{\text{nac}} l_{\text{nac}}^2}{12} + 2d_3^2 m_{\text{bl}} = \\
= \frac{10^5 \cdot 12^2}{12} + 2 \cdot 2.6 \cdot 10^4 \cdot 8^2 \approx \\
\approx 45 \cdot 10^5 \text{kgm}^2.
\]

Dimensionless Constants

Now we introduce dimensionless constants. This allows us to obtain results that are independent of the actual dimensions of the wind turbine under consideration.

To this end we put

\[10^5 \text{kg} \rightarrow 1, \quad 10^2 \text{m} \rightarrow 1, \quad 1 \text{s} \rightarrow 1.\]

Table 8.2 gives a survey on the dimensionless constants used in our model of the wind turbine.

8.2 Computing the Stability Boundary

After these preparations we are in the position to compute the stability boundary.
8.2. Computing the Stability Boundary

Tab. 8.2: Dimensionless constants

<table>
<thead>
<tr>
<th>$m_{\text{tot}}$</th>
<th>1.52</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>0.0989</td>
</tr>
<tr>
<td>$B$</td>
<td>0.0945</td>
</tr>
<tr>
<td>$C$</td>
<td>0.0045</td>
</tr>
<tr>
<td>$k_1$</td>
<td>37.55</td>
</tr>
<tr>
<td>$k_2$</td>
<td>18.77</td>
</tr>
<tr>
<td>$k_3$</td>
<td>12.52</td>
</tr>
<tr>
<td>$k_4$</td>
<td>2.414</td>
</tr>
</tbody>
</table>

Subsystem I

Subsystem I is not affected by the structural damping term. Thus the results for the canonical case hold for the dissipative case as well.

The corresponding differential equation reads

$$\dot{q}' = A'q',$$

where

$$A' := -JH' = -\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} -k_0' & 0 \\ 0 & -(M_0')^{-1} \end{pmatrix} = \begin{pmatrix} 0 & (M_0')^{-1} \\ -P' & 0 \end{pmatrix} =$$

$$\begin{pmatrix} 0 & 0 & \frac{1}{m} & 0 \\ 0 & 0 & 0 & \frac{1}{p} \\ -k_1 & k_2 & 0 & 0 \\ k_2 & -k_3 & 0 & 0 \end{pmatrix}. $$

Thus the system is autonomous. Therefore it suffices to study the spectrum of $A'$. The characteristic equation reads

$$\lambda^4 + \lambda^2 \left( \frac{k_1}{m} + \frac{k_3}{B} \right) - \frac{k_2^2 - k_1 k_3}{Bm} = 0.$$ 

With $\rho := \lambda^2$ we obtain the reduced characteristic equation

$$\rho^2 + \rho \left( \frac{k_1}{m} + \frac{k_3}{B} \right) - \frac{k_2^2 - k_1 k_3}{Bm} = 0.$$ 

It has the following solutions:

$$\rho_{1,2} = -\left( \frac{k_1}{m} + \frac{k_3}{B} \right) \pm \frac{1}{2} \sqrt{\left( \frac{k_1}{m} + \frac{k_3}{B} \right)^2 + 4 \frac{k_2^2 - k_1 k_3}{Bm}} =$$

$$= -\left( \frac{k_1}{m} + \frac{k_3}{B} \right) \pm \frac{1}{2} \sqrt{\left( \frac{k_1}{m} - \frac{k_3}{B} \right)^2 + 4 \frac{k_2^2}{Bm}}.$$
From
\[
\left( \frac{k_1}{m} - \frac{k_3}{B} \right)^2 + 4 \frac{k_2^2}{Bm} > 0
\]
we conclude that the reduced eigenvalues \( \rho_{1,2} \) are real. Thus the eigenvalues \( \lambda_{1,2,3,4} \) are either real or purely imaginary. Obviously one of the reduced eigenvalues is negative. The other reduced eigenvalue has opposite sign if and only if
\[
\frac{k_2^2 - k_1 k_3}{Bm} > 0.
\]
Since \( B > 0 \) and \( m > 0 \) we have
\[
\lambda_{1,2,3,4} \in i \mathbb{R} \iff \rho_{1,2} < 0 \iff k_2^2 - k_1 k_3 < 0.
\]
Thus \( k_2^2 - k_1 k_3 < 0 \) is a necessary condition for the stability of the whole system (1.7) independent of the values of the parameters \( \varepsilon \) and \( \Omega \).

For our example we find
\[
k_2^2 - k_1 k_3 = 18.77^2 - 37.55 \cdot 12.52 \approx -117.81 < 0.
\]
Thus the necessary condition for stability holds.

**Subsystem II**

In contrast to subsystem I subsystem II is non-autonomous and depends on the parameters \( \varepsilon \) and \( \Omega \). Here we have a typical situation to apply our methods developed in Chapters 5 and 7. We do this for the damped rotor as well as for the undamped.

In order to compare the approximation of the stability boundaries obtained by the methods developed in Chapters 5 and 7 with the actual stability boundary we scanned sections of the parameter plane with a small step size. This leads to the shaded regions shown in the figures below. Stable regions are shaded in light-gray while unstable regions are shaded in dark-gray.

Figure 8.2 gives a survey on the regions of stability and instability for the damped rotor with a damping coefficient \( \delta = 0.0003 \). The small rectangles are magnified in Figures 8.3, 8.4, 8.5 and 8.6, where we compare linear and quadratic approximations of the stability boundary obtained at single points \( (\varepsilon_0, \Omega_0) \) with the actual stability boundary. Note that a good correspondence is achieved only for small neighborhoods. The small icons show qualitatively the eigenvalues of the monodromy matrix.

Figure 8.7 gives a survey on the regions of stability and instability for the undamped rotor. Again the small rectangles are magnified in the following figures, where we compare linear and quadratic approximations of the stability boundary obtained at single points \( (\varepsilon_0, \Omega_0) \) with the actual stability boundary (cf. Figures 8.8, 8.9, 8.10 and 8.11).

Note that for our wind turbine model the stability boundary may be computed locally without problems, but tracing the complete stability boundary would be very challenging.
Fig. 8.2: Survey on the regions of stability for the damped system with $\delta = 0.0003$. 
8. Stability of a Simple Model of a Two Bladed Rotor

Fig. 8.3: A comparison between the actual stability boundary and the linear and quadratic approximations for the damped system with $\delta = 0.0003$.

Fig. 8.4: A comparison between the actual stability boundary and the linear and quadratic approximations for the damped system with $\delta = 0.0003$. 
8.2. Computing the Stability Boundary

Fig. 8.5: A comparison between the actual stability boundary and the linear and quadratic approximations for the damped system with $\delta = 0.0003$.

Fig. 8.6: A comparison between the actual stability boundary and the linear and quadratic approximations for the damped system with $\delta = 0.0003$. 
Fig. 8.7: Survey on the regions of stability for the undamped system ($\delta = 0.0$).
8.2. Computing the Stability Boundary

Fig. 8.8: A comparison between the actual stability boundary and the linear and quadratic approximations for the undamped system.

Fig. 8.9: A comparison between the actual stability boundary and the linear and quadratic approximations for the undamped system.
Fig. 8.10: A comparison between the actual stability boundary and the linear and quadratic approximations for the undamped system.

Fig. 8.11: A comparison between the actual stability boundary and the linear and quadratic approximations for the undamped system.
Part II

Stability and Chaotic Behavior in Non-Linear Oscillations in the Presence of Symmetry with an Application to the Rotational Motion of a Dumbbell Satellite
We consider a satellite in the shape of a dumbbell: It is composed of two point masses \( m_1 \) and \( m_2 \) connected by a rigid massless bar of length \( l \) (cf. Figure 1.1).

It is assumed that the orbital motion of the satellite takes place in a plane which is taken to be the \( x-y \)-plane. The motion of the satellite may be described by polar coordinates \( r \) and \( \phi \). The angle \( \theta \) between the axis of the dumbbell and the vector of the center of mass is used to describe the relative motion of the bar.

To derive the equations of motion we use the method of Lagrange.
1. The Equations of Lagrange

The kinetic and the potential energy of the dumbbell are, respectively,

\[ T = \frac{1}{2}(m_1 + m_2)v^2 + \frac{1}{2}(m_1l_1^2 + m_2l_2^2)(\dot{\theta} + \dot{\phi})^2 = \]
\[ = \frac{1}{2}(m_1 + m_2)r^2 + \frac{1}{2}(m_1 + m_2)r^2\dot{\varphi}^2 + \frac{1}{2}(m_1l_1^2 + m_2l_2^2)(\dot{\theta} + \dot{\phi})^2, \]

\[ U = -\left( \frac{m_1}{\sqrt{r^2 + l_1^2 + 2rl_1 \cos \theta}} + \frac{m_2}{\sqrt{r^2 + l_2^2 - 2rl_2 \cos \theta}} \right), \]

where \( l_1 \) and \( l_2 \) are determined by \( l_1 + l_2 = l \) and \( l_1m_1 = l_2m_2 \).

The Lagrangian function \( L \) is given by

\[ L := T - U = \frac{1}{2}(m_1 + m_2)r^2 + \frac{1}{2}(m_1 + m_2)r^2\dot{\varphi}^2 + \frac{1}{2}(m_1l_1^2 + m_2l_2^2)(\dot{\theta} + \dot{\phi})^2 + \]
\[ + \frac{m_1}{\sqrt{r^2 + l_1^2 + 2rl_1 \cos \theta}} + \frac{m_2}{\sqrt{r^2 + l_2^2 - 2rl_2 \cos \theta}}. \]

The Lagrangian equation read

\[ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0 \]

with \( q = r, \varphi, \theta \), respectively. The equations of motion are found to be

\[ \ddot{r} = r\dot{\varphi}^2 - \frac{1}{m_1 + m_2} \left( \frac{m_1(r + l_1 \cos \theta)}{(r^2 + l_1^2 + 2rl_1 \cos \theta)^{3/2}} + \frac{m_2(r - l_2 \cos \theta)}{(r^2 + l_2^2 - 2rl_2 \cos \theta)^{3/2}} \right), \]  

(1.11)

\[ 0 = 2r\ddot{r} + r^2\ddot{\varphi} - \frac{m_1}{m_1 + m_2}l_1^2 + \frac{m_2}{m_1 + m_2}l_2^2 \left( \ddot{\theta} + \ddot{\phi} \right), \]  

(1.12)

\[ \ddot{\theta} + \ddot{\phi} = \frac{r}{m_1l_1^2 + m_2l_2^2} \left( \frac{m_1l_1}{(r^2 + l_1^2 + 2rl_1 \cos \theta)^{3/2}} - \frac{m_2l_2}{(r^2 + l_2^2 - 2rl_2 \cos \theta)^{3/2}} \right) \sin \theta. \]  

(1.13)

Since the diameter \( l \) of the satellite is extremely small compared with the radius \( r \) of the orbit, we consider the limiting case \( l \to 0 \).

1.2 The Limiting Case \( l \to 0 \)

In equations (1.11) and (1.12) we may simply put \( l_1 := 0 \) and \( l_2 := 0 \):

\[ \ddot{r} = r\dot{\varphi}^2 = -\frac{1}{r^2} \]  

(1.21)

\[ \frac{d}{dt} \left( r^2\dot{\varphi} \right) = 0. \]  

(1.22)
1.3. Stationary Solutions

These equations show that the motion of the center of mass of the satellite is the Keplerian motion in this limiting case.

Equation (1.13) must be treated with more care. Let \( r_1^* := r^2 + l_1^2 + 2r_1 l_1 \cos \theta \) and \( r_2^* := r^2 + l_2^2 - 2r_2 l_2 \cos \theta \). Then we have

\[
\frac{r}{m_1 l_1^2 + m_2 l_2^2} \left( \frac{m_1 l_1}{r_1^3} - \frac{m_2 l_2}{r_2^3} \right) \sin \theta = \frac{r}{m_1 l_1^2 + m_2 l_2^2} \left( \frac{m_1 l_1}{r_1^3} - \frac{m_2 m_3 l_1}{r_2^3} \right) \sin \theta =
\]

\[
= \frac{m_2}{m_1 + m_2} \cdot \frac{r}{l_1} \left( \frac{1}{r_1^3} - \frac{1}{r_2^3} \right) \sin \theta = \frac{m_2}{m_1 + m_2} \cdot \frac{r}{l_1} \frac{r_2^3 - r_1^3}{r_1^3 r_2^3} \sin \theta =
\]

\[
= \frac{m_2}{m_1 + m_2} \cdot \frac{r}{l_1} \frac{(r_2 - r_1)(r_2^3 + r_1 r_2 + r_1^3)}{r_1^3 r_2^3} \sin \theta =
\]

\[
= \frac{m_2}{m_1 + m_2} \cdot \frac{r}{l_1} \frac{r_2^2 + r_1 r_2 + r_1^2}{r_1^3 r_2^3 (r_2 + r_1)} \left( r^2 + \frac{m_2^2 l_1^2}{m_2} - 2r \frac{l_1}{m_2} \cos \theta - r_1^2 - 2r l_1 \cos \theta \right) \sin \theta =
\]

\[
= \frac{m_2}{m_1 + m_2} \cdot \frac{r}{l_1} \frac{r_2^2 + r_1 r_2 + r_1^2}{r_1^3 r_2^3 (r_2 + r_1)} \left( \frac{m_1 - m_2}{m_2} r_1^2 l_1 - 2r l_1 \frac{m_1 + m_2}{m_2} \cos \theta \right) \sin \theta =
\]

Thus for \( l \to 0 \) equation (1.13) is simplified to

\[
\ddot{\theta} + \frac{3 \sin(2\theta)}{\sqrt{r^3}} = -\dot{\phi}
\]

This equation describes the motion of the bar around the center of mass.

It was essential to first compute the Lagrangian equations and only then to take the limit \( l \to 0 \).

1.3 Stationary Solutions

In this section we consider two stationary solutions of (1.2) and discuss their stability.

Let \( r = r_0 \) be fixed, i.e. the center of mass moves uniformly on a circle. To fulfill equation (1.21) we choose

\[
\dot{\phi} = \dot{\phi}_0 := \pm r_0^{-3/2}.
\]

Then equation (1.22) holds identically and equation (1.23) reduces to

\[
\ddot{\theta} + \frac{3 \sin(2\theta)}{2 \sqrt{r_0^3}} = 0.
\]
This is the equation of motion of a pendulum. Hence we have two equilibrium solutions $\theta_0 = 0$ and $\theta_0 = \frac{\pi}{2}$. They correspond to radial and tangential orientation of the dumbbell, respectively, as shown in Figure 1.2.

![Radial and tangential orientation of the dumbbell satellite.](image)

**Fig. 1.2:** Radial and tangential orientation of the dumbbell satellite.

It is well-known that the solution $\theta_0 = 0$ is stable, while the solution $\theta_0 = \frac{\pi}{2}$ is unstable. Suppose now that the orbit of the center of motion is no longer a circle, but an ellipse. Then two questions arise:

- Will the first solution still be stable?
- Will the second solution generate chaotic behavior?

These questions are the subject of Part II of this dissertation. The first eight chapters are dedicated to the stability problem. The remaining two chapters are devoted to chaotic behavior.

The organisation of the chapters is as follows:

**Chapter 2:** We simplify the problem at hand by introducing the solution to Kepler's problem. The result is a non-autonomous second order equation.

**Chapter 3:** For the sake of generality of the subsequent considerations a general framework is introduced.

**Chapter 4:** We perform some coordinate transformations to obtain perturbed harmonic oscillators.

**Chapter 5:** We recursively construct formal first integrals and prove the existence of solutions of the recursion scheme.

**Chapter 6:** We provide both recursive as well as a priori estimates for truncated first integrals.

**Chapter 7:** We offer results on the stability of periodic solutions based on approximate first integrals.
Chapter 8: We apply the general results of the previous chapters to the dumbbell satellite problem. For the explicit construction of the approximate first integrals we make use of a computer algebra package.

Chapter 9: We prove chaotic behavior for small eccentricities using the theory of Melnikov.

Chapter 10: We prove chaotic behavior for large eccentricities using numerical shadowing.
Keplerian Motion

The chapter is organized as follows:

Section 2.1: We solve Eqs. (1.21) and show that the center of mass moves on an ellipse with the earth in one of the foci.

Section 2.2: We derive Kepler’s Laws.

Section 2.3: We describe the dynamics of the elliptic motion. An exhaustive treatment of Keplerian motion is found in every textbook on celestial mechanics, e.g. [20].

Section 2.4: We substitute the results obtained above into equation (1.23). This resulting non-autonomous equation is the starting point of our subsequent analysis.

Section 2.5: We summarize the main properties of the resulting equation.

2.1 Elliptic Motion

Let

\[ x := r \cos \varphi, \]
\[ y := r \sin \varphi. \]

Taking derivatives we obtain

\[ \dot{x} = \dot{r} \cos \varphi - r \sin \varphi \dot{\varphi}, \]
\[ \dot{y} = \dot{r} \sin \varphi + r \cos \varphi \dot{\varphi} \]

and

\[ \ddot{x} = \cos \varphi (\ddot{r} - r \dot{\varphi}^2) - \sin \varphi (2\dot{\varphi} \dot{\varphi} + r \ddot{\varphi}), \]
\[ \ddot{y} = \sin \varphi (\ddot{r} - r \dot{\varphi}^2) + \cos \varphi (2\dot{\varphi} \dot{\varphi} + r \ddot{\varphi}). \]
We start our computations with equation (1.22). It may be written as
\[ r^2 \dot{\varphi} = \sigma \]  
(1.2a)
with some fixed constant \( \sigma \) or as
\[ 2r \ddot{r} \dot{\varphi} + r^2 \dot{\varphi}^2 = 0. \]  
(1.2b)
From (1.2a) we obtain
\[ \frac{1}{r} \frac{\partial}{\partial r} (r^2 \varphi) = \frac{1}{r} \frac{\partial}{\partial \varphi} (r^2 \varphi) = \frac{1}{r^2} (r \cos \varphi + r \cos \varphi \dot{\varphi}) - r \sin \varphi (r \cos \varphi - r \sin \varphi \dot{\varphi}) = \]  
\[ = r^2 \dot{\varphi} = \sigma. \]  
(1.2c)
From (1.2a) and (1.2b) we obtain
\[ \dot{x} = -\frac{1}{r^2} \cos \varphi, \]  
\[ \dot{y} = -\frac{1}{r^2} \sin \varphi \]  
and further
\[ \frac{d}{dt} \left( \frac{x}{r} \right) = \frac{d}{dt} \cos \varphi = -\sin \varphi \dot{\varphi} = \sigma \dot{y}, \]  
\[ \frac{d}{dt} \left( \frac{y}{r} \right) = \frac{d}{dt} \sin \varphi = \cos \varphi \dot{\varphi} = -\sigma \dot{x}. \]
These equations are referred to as Lagrangian integrals. They may be written as
\[ \sigma \dot{y} = \frac{x}{r} + e, \]  
\[ \sigma \dot{x} = -\frac{y}{r} + f, \]  
(2.1)
where \( e \) and \( f \) are constants of integration.
Now we chose our coordinate system such that the \( x \)-axis intersects the solution curve of (2.1) perpendicularly. Then in the intersection point we have \( y = 0 \) and \( \dot{x} = 0 \). This implies that \( f \) vanishes for this choice of the coordinate system.
Multiplying the first equation of (2.1) by \( x \), the second (with \( f = 0 \)) by \( -y \) and adding the resulting equations, we obtain
\[ \sigma (x \dot{y} - \dot{x} y) = r + ex. \]
Using equation (1.2c) we find
\[ \sigma^2 = r + er \cos \varphi. \]
Thus we have
\[ r = \frac{\sigma^2}{1 + e \cos \varphi}. \]
2.1. Elliptic Motion

This equation describes a conic section with one focus at the origin. For $0 < e < 1$ it is an ellipse with eccentricity $e$.

Let $a$ denote the semi-major axis. Then we have

$$\frac{\sigma^2}{1+e} + \frac{\sigma^2}{1-e} = 2a$$

and therefore

$$\sigma^2 = a(1-e^2). \tag{2.2}$$

Thus the equation of the conic section now reads

$$r = \frac{a(1-e^2)}{1+e \cos \varphi}. \tag{2.3}$$

For the distance of the foci from the center we obtain

$$a - \frac{a(1-e^2)}{1+e} = a - a(1-e) = ae.$$  

Thus for the semi-minor axis $b$ we eventually obtain

$$b = a\sqrt{1-e^2}.$$
2.2 Kepler’s Laws

The results deduced above from Eqs. (1.2) and (1.22) give rise to Kepler’s famous laws:

1. Kepler’s first law
   The satellite moves on a conic section with semi-major axis \( a \) and eccentricity \( e \) with the earth in one of the foci. In polar coordinates the equation of the ellipse reads
   \[
r = \frac{a(1 - e^2)}{1 + e \cos \varphi}.
   \]

2. Kepler’s second law
   The area \( A(t) \) swept out by the radius vector is proportional to the time elapsed. Indeed from (1.22a) we obtain
   \[
   A(t) = \int_{\varphi_1}^{\varphi_2} \frac{1}{2} r^2 d\varphi = \int_{t_1}^{t_2} \frac{1}{2} r^2 \dot{\varphi} dt = \int_{t_1}^{t_2} \frac{1}{2} \sigma dt = \frac{1}{2} \sigma(t_2 - t_1).
   \]

3. Kepler’s third law
   The cubic power of the semi-major axis is proportional to the square of the satellite’s period.
   On the one hand the area of the ellipse is
   \[
   A = \pi ab = \pi a^2 \sqrt{1 - e^2}.
   \]
   On the other hand from the derivation of Kepler’s second law we have
   \[
   A = \frac{1}{2} \sigma T.
   \]
   Combing these results we indeed find:
   \[
   \frac{T^2}{a^3} = 4\pi^2.
   \]

2.3 Dynamics of Elliptic Motion

To describe the dynamics of the elliptic motion we introduce the so called eccentric anomaly \( E \). By elementary geometric arguments we express \( r \) and \( \varphi \) in terms of \( E \).

- We start with the computation of \( r \):
  \[
  r^2 = a^2 (\cos E - e)^2 + b^2 \sin^2 E =
  = a^2 (\cos^2 E - 2e \cos E + e^2 + \sin^2 E - e^2 \sin^2 E) =
  = a^2 (1 - e \cos E)^2.
  \]

Therefore we have
\[
r = a(1 - e \cos E). \quad (2.4)
\]
2.3. Dynamics of Elliptic Motion

Next we consider $\varphi$:

$$a \cos E = ae + r \cos \varphi = ae + \frac{a(1 - e^2)}{1 + e \cos \varphi} \cos \varphi =$$

$$= \frac{ae + ae^2 \cos \varphi + a \cos \varphi - ae^2 \cos \varphi}{1 + e \cos \varphi} = a \cdot \frac{e + \cos \varphi}{1 + e \cos \varphi}.$$  

Hence it follows that

$$1 - \cos E = 1 - \frac{e + \cos \varphi}{1 + e \cos \varphi} = \frac{(1 - e)(1 - \cos \varphi)}{1 + e \cos \varphi},$$

$$1 + \cos E = 1 + \frac{e + \cos \varphi}{1 + e \cos \varphi} = \frac{(1 + e)(1 + \cos \varphi)}{1 + e \cos \varphi}$$

and finally

$$\tan^2 \frac{E}{2} = 1 - \cos E = 1 - \frac{e}{1 + e} \tan^2 \frac{\varphi}{2}.$$  \hspace{1cm} (2.5)

This last formula allows us to compute the integral in Kepler's second law:

$$\sqrt{a(1 - e^2)} \cdot t = \sigma \cdot t = \int_0^t \sigma ds = \int_0^t r^2 d\varphi =$$

$$= \int_0^\varphi r^2 d\varphi = \int_0^\varphi \frac{a^2(1 - e^2)^2}{(1 + e \cos \varphi)^2} d\varphi.$$

By (2.5) and the identity

$$\cos \varphi = \frac{1 - \tan^2 \left(\frac{\varphi}{2}\right)}{1 + \tan^2 \left(\frac{\varphi}{2}\right)}$$
we obtain

$$\sqrt{a(1-e^2)} \cdot t =$$

$$= a^2(1-e^2)^2 \int_0^E \left(1+e \frac{1-\frac{1+e}{1-e} \tan^2 \frac{E}{2}}{1+\frac{1+e}{1-e} \tan^2 \frac{E}{2}}\right)^{-2} \sqrt{\frac{1+e}{1-e} \frac{1+\tan^2 \frac{E}{2}}{1-e} \frac{1+\tan^2 \frac{E}{2}}{1+\tan^2 \frac{E}{2}}} dE =$$

$$= a^2(1-e^2)^2 \sqrt{\frac{1+e}{1-e} \int_0^E \frac{(1+\tan^2 \frac{E}{2}) (1+\frac{1+e}{1-e} \tan^2 \frac{E}{2})}{(1+\frac{1+e}{1-e} \tan^2 \frac{E}{2} + e - \frac{1+e}{1-e} \tan^2 \frac{E}{2})^2} dE =$$

$$= a^2(1-e^2)^2 \sqrt{\frac{1+e}{1-e} \int_0^E \frac{1+\tan^2 \frac{E}{2}}{1-e} \frac{1}{(1+e)^2 (1+\tan^2 \frac{E}{2})^2} \frac{1}{1-e} (1+\tan^2 \frac{E}{2}) dE =$$

$$= a^2 \sqrt{1-e^2} \int_0^E \left(1-e \frac{1-\frac{1+e}{1-e} \tan^2 \frac{E}{2}}{1+\tan^2 \frac{E}{2}}\right) dE =$$

$$= a^2 \sqrt{1-e^2} \int_0^E (1-e \cos E) dE =$$

$$= a^2 \sqrt{1-e^2} (E - e \sin E).$$

Thus we have

$$t = a^{3/2} (E - e \sin E) \tag{2.6}$$

This equation is referred to as Kepler's equation.

### 2.4 Motion Around the Center of Mass

We are now in a position to compute $r$ and $\dot{\varphi}$ in terms of $E$ and we substitute the result into (1.23). From (1.22b) it follows that

$$\dot{r} = -\frac{2 \dot{E}}{r}. \tag{2.7}$$

Therefore we need $\dot{r}$ and $\dot{\varphi}$ in terms of $E$:

- $\dot{r}$: From Eq. (2.4) we obtain

$$\dot{r} = ae \sin E \dot{E}$$

and from Eq. (2.6)

$$1 = a^{3/2} (1 - e \cos E) \dot{E}.$$
Combining these results we have
\[ \dot{r} = ae \sin E \cdot \frac{1}{a^{3/2}(1 - e \cos E)} = \frac{e \sin E}{\sqrt{a(1 - e \cos E)}}. \]

- \( \dot{\psi} \): From Eqs. (1.2)\(_2\), (2.2) and (2.4) we obtain
\[ \dot{\psi} = \frac{\sqrt{a(1 - e^2)}}{r^2} = \frac{\sqrt{1 - e^2}}{a^{3/2}(1 - e \cos E)^2}. \]

Substituting these findings into Eq. (2.7) we have
\[ \ddot{\psi} = -\frac{2e \sqrt{1 - e^2} \sin E}{a^3(1 - e \cos E)^4}. \]

Therefore Eq. (1.23) may be written as
\[ \ddot{\theta} + \frac{3}{2} \frac{\sin(2\theta)}{a^3(1 - e \cos E)^3} = \frac{2e \sqrt{1 - e^2} \sin E}{a^3(1 - e \cos E)^4}. \]

Finally we eliminate \( t \) by replacing the derivative with respect to \( t \) by the derivative with respect with the eccentric anomaly \( E \). From (2.6) it follows that
\[ \frac{dE}{dt} = \frac{1}{a^{3/2}(1 - e \cos E)}. \]

Therefore we have
\[ \frac{d\theta}{dt} = \frac{d\theta}{dE} \frac{dE}{dt} = \frac{d\theta}{dE} \frac{1}{a^{3/2}(1 - e \cos E)}, \]

and
\[ \frac{d^2\theta}{dt^2} = \frac{d}{dt} \left( \frac{d\theta}{dt} \right) = \frac{d}{dE} \left( \frac{d\theta}{dE} \frac{1}{a^{3/2}(1 - e \cos E)} \right) \frac{1}{a^{3/2}(1 - e \cos E)} = \]
\[ = \frac{d^2\theta}{dE^2} \frac{1}{a^3(1 - e \cos E)^2} - \frac{d\theta}{dE} \frac{e \sin E}{a^3(1 - e \cos E)^3}. \]

This eventually leads to
\[ \frac{d^2\theta}{dE^2} - \frac{d\theta}{dE} \frac{e \sin E}{1 - e \cos E} + \frac{3}{2} \frac{\sin(2\theta)}{1 - e \cos E} = \frac{2e \sqrt{1 - e^2} \sin E}{(1 - e \cos E)^2} \]

This is the equation of the dumbbell satellite problem.
2.5 Basic Properties

For further treatment let us introduce the following notation:

\[ t := E, \quad x := \theta, \quad y := \dot{\theta}, \quad z = (x, y)^T. \]

Then equation (2.8) is of the form

\[ \dot{z} = F(t, z, e), \quad (2.9) \]

where the function \( F \) defined by

\[ F(t, z, e) := F(t, x, y, e) = \left( \frac{3}{2} \sin(2\theta) + \frac{y}{1 - e \cos t} + \frac{2ye \sqrt{1 - e^2} \sin t}{(1 - e \cos t)^2} \right) \]

has the following properties:

**P₁:** \( F \) is in \( C^\infty(\mathbb{R} \times \mathbb{R}^2 \times ]-1,1[, \mathbb{R}^2) \).

**P₂:** \( F \) is 2\( \pi \)-periodic in \( t \):

\[ F(t + 2\pi, z, e) = F(t, z, e). \]

**P₃:** \( F \) is reversible:

\[ F(-t, Rz, e) = -RF(t, z, e), \]

with the reversibility matrix

\[ R := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \]
3

A General Framework

In the last chapter we derived a differential equation for the movement of the dumbbell satellite around its center of mass and we proved some basic properties.

Now for the sake of generality of the subsequent considerations we introduce a general framework. Furthermore we introduce the concept of practical stability.

The chapter is organized as follows:

Section 3.1: We define the general framework.

Section 3.2: We discuss one of the assumptions made in the general framework.

Section 3.3: We introduce the concept of practical stability.

3.1 The Definition of the General Framework

We consider a system of differential equations

\[ \dot{z} = F(t, z, \sigma), \] (3.1)

where

\[ F : \mathbb{R} \times (U \subset \mathbb{R}^{2n}) \times (U_\sigma \subset \mathbb{R}^m) \rightarrow \mathbb{R}^{2n}, \ (t, z, \sigma) \mapsto F(t, z, \sigma). \]

Let the following assumptions hold.

General Assumptions

A_1: F is sufficiently regular in t, z and \( \sigma \).

A_2: F is T-periodic in t:

\[ F(t + T, z, \sigma) = F(t, z, \sigma), \]

for all t, z, \( \sigma \) in their domains of definition.
A3: $F$ fulfills the reversibility condition

$$F(-t, Rz, \sigma) = -RF(t, z, \sigma),$$

where the so-called reversibility matrix $R$ is assumed to be invertible.

A4: Assume further that for some fixed value $\sigma = \sigma^0$ of the parameter the system of differential equations (3.1) admits a $T$-periodic solution $z^0_{\text{per}}$ such that $z^0_{\text{per}}$ fulfills the symmetry condition

$$z^0_{\text{per}}(-t) = Rz^0_{\text{per}}(t).$$

A5: Let $Z_1$ denote the principal fundamental matrix solution of

$$\dot{z}_1 = D_2F(t, z^0_{\text{per}}(t), \sigma^0)z_1,$$

i.e.,

$$\dot{Z}_1 = D_2F(t, z^0_{\text{per}}(t), \sigma^0)Z_1 \text{ and } Z_1(0) = I.$$

Assume that the monodromy matrix $Z_1(T)$ has the following properties:

A5a: The eigenvalues lie on the unit circle and are simple:

$$\mu_1^{\pm} = e^{\pm i\omega_1 T}, \mu_2^{\pm} = e^{\pm i\omega_2 T}, \ldots \mu_n^{\pm} = e^{\pm i\omega_n T}.$$

The $\mu_i$ are called characteristic multipliers.

A5b: For some $N \in \mathbb{N}$ or $N = \infty$ let there exist constants $\gamma > 0, \tau \geq 0$ such that

$$|l_0 \omega_0 + l \cdot \omega| \geq \frac{1}{|l|_1}$$

(NR$_N$)

for $l_0 \in \mathbb{Z}$, $l \in \mathbb{N}^n$, $|l|_1 \leq N$, $|l_0| + |l|_1 \neq 0$,

where $\omega_0 := \frac{2\pi}{T}$, $\omega := (\omega_1, \ldots, \omega_n)$ and $|l|_1 := \sum_{i=1}^{n} |l_i|$.

Our goal is to study the stability of the periodic solution $z^0_{\text{per}}$.

### 3.2 Discussion of Assumption A5

In Assumption A5a it is required that the eigenvalues of the monodromy matrix $M := Z_1(T)$

- lie on the unit circle and
- are simple.

We show that the first part is a necessary condition for the stability of reversible systems and that the second part follows from Assumption A5b.
Furthermore we give equivalent formulations of the non-resonance condition in Assumption \( A_{5b} \) for finite \( N \).

\[ 3.2. \text{ Discussion of Assumption } A_5 \]

A Necessary Condition for Stability

Let \( A(t) := D_2 F(t, z_{\text{per}}^0(t, \sigma^0), \sigma^0) \). Then the linearization of (3.1) along the periodic solution \( z_{\text{per}}^0 \) may be written as

\[ \dot{z}_1 = A(t)z_1. \quad (3.2) \]

Let \( Z_1(t) \) denote the principal fundamental matrix solution of (3.2) (cf. \( A_5 \)). It is well-known (cf. [1]) that the spectrum of the monodromy matrix \( M := Z_1(T) \) provides a criterion for the stability of the equilibrium solution \( z_1 = 0 \) of Eq. (3.2).

Lemma 3.2.1

The equilibrium solution \( z_1 = 0 \) of (3.2) is stable if and only if the following conditions on the eigenvalues \( \mu \) of the monodromy matrix \( M \) hold:

(i) \( |\mu| \leq 1 \),

(ii) if \( |\mu| = 1 \), then \( \mu \) is semi-simple.

A proof of this lemma may be found in every textbook on differential equation (e.g. [1], p. 311).

Assumption \( A_3 \) implies that the linear system is reversible too.

\[ A(-t)R = -RA(t). \]

The reversibility of the linear system has further implications on the spectrum of the monodromy matrix \( M \).

Lemma 3.2.2

Let the linear system be reversible. Then the following holds:

(i) \( Z_1(-t) = RZ_1(t)R^{-1} \).

(ii) \( M^{-1} = RMR^{-1} \).

(iii) The spectrum of \( M \) is symmetric with respect to the real axis and to the unit circle.

Proof:

(i) Let \( Y(t) := RZ_1(-t)R^{-1} \). Then we have
\[ \dot{Y}(t) = -RZ_1(-t)R^{-1} = -RA(-t)Z_1(-t)R^{-1} = A(t)RZ_1(-t)R^{-1} = A(t)Y(t) \]

and

\[ Y(0) = RZ_1(0)R^{-1} = RIR^{-1} = I. \]

Uniqueness of solutions implies that \( Y(t) = Z_1(t) \).

(ii) By (i) we have

\[ M^{-1} = Z_1(T)^{-1} = Z_1(-T) = RZ_1(T)R^{-1} = RMR^{-1}. \]

(iii) Since the matrix \( M \) is real, its spectrum is symmetric with respect to the real axis. Part (ii) implies that the matrices \( M \) and \( M^{-1} \) are similar. Thus their spectra are identical. Combining these results we find the following. If \( \mu \) is an eigenvalue of \( M \), then \( \mu^{-1} \) is an eigenvalue of \( M^{-1} \) and therefore also of \( M \). Moreover, since \( M \) is real \( \bar{\mu}^{-1} \) is an eigenvalue of \( M \), too. Thus the spectrum of \( M \) is also symmetric with respect to the unit circle.

We conclude that an eigenvalue of the monodromy matrix off the unit circle implies instability for reversible systems. Or in other words: It is necessary for the stability of a reversible system that the eigenvalues of the monodromy matrix lie on the unit circle.

A Consequence of Assumption A5b

Now we consider the second part of Assumption A5a: the simplicity of the eigenvalues. We show that it follows from the non-resonance condition in Assumption A5b.

To this end let \( \mu_1 := e^{i\omega_1 T} \) and \( \mu_2 := e^{i\omega_2 T} \) be two eigenvalues of \( M \). From \( \mu_1 = \mu_2 \) it follows that \( \omega_2 T = \omega_1 T + 2\pi l \), for some \( l \in \mathbb{Z} \). Substituting \( T \) by \( \frac{2\pi}{\omega_0} \) we find that

\[ l \cdot \omega_0 + 1 \cdot \omega_1 + (-1) \cdot \omega_2 = 0 \]

in contradiction to the non-resonance condition in Assumption A5b.

Several Formulations of the Non-Resonance Condition

For finite numbers \( N \) there are several equivalent formulations of the non-resonance condition \( \text{NR}_N \).

Lemma 3.2.3

Let \( n \in \mathbb{N} \) or \( N = \infty \) and consider the following non-resonance conditions.

\[ \text{NR}_N : \text{ There exist constants } \gamma >, \tau \geq 0 \text{ such that} \]

\[ |l_0 \omega_0 + l \cdot \omega| \geq \gamma \frac{1}{|l|} \]

for \( l_0 \in \mathbb{Z}, l \in \mathbb{Z}^n, |l|_1 \leq N, |l_0| + |l|_1 \neq 0. \)
3.2. Discussion of Assumption A5

NR′_N : There exists a non-increasing sequence (α_k)_{1 \leq k \leq N} such that
\[ |l_0 \omega_0 + l \cdot \omega| \geq \alpha_k > 0 \]
for \( l_0 \in \mathbb{Z}, l \in \mathbb{Z}^n, |l|_1 = k \leq N, |l_0| + |l|_1 \neq 0. \)

NR''_N : The following inequality holds:
\[ l_0 \omega_0 + l \cdot \omega \neq 0 \]
for \( l_0 \in \mathbb{Z}, l \in \mathbb{Z}^n, |l|_1 \leq N, |l_0| + |l|_1 \neq 0. \)

Then the following holds:
(i) For \( N = \infty \):
\[ NR_N \Rightarrow NR'_N \Rightarrow NR''_N. \]
(ii) For finite \( N \):
\[ NR_N \iff NR'_N \iff NR''_N. \]

Proof:
(i) This part of the lemma is obvious.
(ii) We show that \( NR''_N \) implies \( NR_N \). To this end we first prove that \( |l_0 \omega_0 + l \cdot \omega| \)
cannot take its minimum if \( l_0 \) is large. Let
\[ l_i := (0, \ldots, 0, 1, 0, \ldots, 0). \]

It follows immediately that
\[ \min_{0 \leq |l|_1 \leq k} |l \cdot \omega| \leq \min_{1 \leq i \leq n} |l_i \cdot \omega| = \min_{1 \leq i \leq n} |\omega_i| < \max_{1 \leq i \leq n} |\omega_i| =: |\omega|_\infty. \]

For
\[ l_0 \geq k_0 := 2k \cdot \left\lfloor \frac{|\omega|_\infty}{|\omega_0|} \right\rfloor \]
we obtain using \( 1 \leq |l|_1 \leq k \)
\[ |(l_0; l) \cdot (\omega_0; \omega)| = |l_0 \omega_0 + l \cdot \omega| \]
\[ \geq |l_0 \omega_0| - |l \cdot \omega| \geq 2|\omega|_\infty \cdot k - |l|_1 \cdot |\omega|_\infty \geq 2k|\omega|_\infty - k|\omega|_\infty = k|\omega|_\infty \geq |\omega|_\infty \geq \min_{0 \leq |l|_1 \leq k} |l \cdot \omega| \geq \min_{0 \leq |l|_1 \leq k, |l_0| + |l|_1 \neq 0} |l_0 \omega_0 + l \cdot \omega|. \]
This implies
\[
\min_{0 \leq |l| \leq k} |l_0 \omega_0 + l \cdot \omega| = \min_{0 \leq |l_1| \leq k} |l_0 \omega_0 + l \cdot \omega| = \min_{0 \leq |l_1| \leq k} |(l_0; l) \cdot (\omega_0; \omega)|. \tag{3.3}
\]

Now we define the sets
\[
M_k := \{(l_0; l) \in \mathbb{Z} \times \mathbb{Z}^n \mid |l_0| \leq k_0, |l_1| \leq k, |l_0| + |l_1| \neq 0\}, \quad \text{for} \quad k \leq N.
\]

Since \((l_0; l) \cdot (\omega_0; \omega) \neq 0\) for \((l_0; l) \in M_k\) and since the set \(M_k\) is finite the minimum of \(|(l_0; l) \cdot (\omega_0; \omega)|\) exists and is positive
\[
\beta_k := \min_{(l_0; l) \in M_k} |(l_0; l) \cdot (\omega_0; \omega)| > 0.
\]

Since the sequence \(\beta_1, \ldots, \beta_N\) is finite there exist constants \(\gamma > 0\) and \(\tau \geq 0\) such that \(\beta_k \geq \gamma k^{-\tau}\). Thus condition \(\text{NR}_N\) is satisfied. \(\square\)

The proof of this lemma allows us to check the non-resonance condition \(\text{NR}_N\) for a given finite number \(N\) even if the frequencies \((\omega_0; \omega_1, \ldots, \omega_n)\) are known only with finite precision.

3.3 Lyapunov Stability vs. Practical Stability

Stability in the Sense of Lyapunov

Let \(z_{\text{per}}^0\) denote a \(T\)-periodic solution of (3.1). By
\[
z = z_{\text{per}}^0(t) + z_1
\]
equation (3.1) is transformed to
\[
\dot{z}_1 = F(t, z_{\text{per}}^0(t) + z_1, \sigma^0) - F(t, z_{\text{per}}^0(t), \sigma^0), \tag{3.4}
\]
and \(z_{\text{per}}^0\) is transformed to \(z_1 = 0\). The question of stability of periodic solutions may therefore be reduced to the question of stability of equilibrium solutions.

**Definition 3.3.1 (Stability in the Sense of Lyapunov)**

An equilibrium solution \(z_0\) of a periodic system of differential equations is called **stable in the sense of Lyapunov**, if for every neighborhood \(U_{\text{tol}}\) of \(z_0\) there exists a neighborhood \(U_{\text{start}}\) of \(z_0\) such that every solution of (3.4) with initial values in \(U_{\text{start}}\) will stay in \(U_{\text{tol}}\) for all positive times.
3.3. Lyapunov Stability vs. Practical Stability

This definition together with the coordinate transformation allows us to define stability of periodic solutions.

**Definition 3.3.2 (Stability of a Periodic Solution)**

A periodic solution $z_{per}$ of (3.1) is called **stable in the sense of Lyapunov**, if the corresponding equilibrium solution $z_1 = 0$ of (3.4) is stable in the sense of Lyapunov.

**Practical Stability**

The definition of stability given above has two disadvantages:

- The notion of stability in the sense of Lyapunov is purely qualitative i.e. there is no information on the size of the starting neighborhood $U_{start}$ in terms of the size of the tolerance neighborhood $U_{tol}$.
  But for practical purposes it is important that the size of the starting neighborhood may be determined quantitatively. Moreover the size of $U_{start}$ should not be too small compared with the size of the tolerance neighborhood $U_{tol}$.

- The requirement “for all positive times” is too restrictive for many purposes. In the theory of perturbed integrable Hamiltonian systems stability in the sense of Lyapunov may not hold, if the number of degrees of freedom is greater than two, due to the possibility of Arnold diffusion.

Both disadvantages may be avoided if we replace the requirement “for all positive times” by “for long positive times”. Of course it depends on the situation which times we call...
"long". In celestial mechanics for example a time is called "long" if it is comparable with the age of the universe. This kind of stability is called \textit{practical} or \textit{effective} stability.
4

Preliminary Coordinate Transformations

We start with system (3.1) subject to Assumption $A_1$–$A_5$ and perform a series of preliminary coordinate transformations to achieve a set of periodically perturbed harmonic oscillators leading to $0$ as equilibrium solution.

The chapter is organized as follows:

Section 4.1: We introduce local coordinates.

Section 4.2: We apply the so-called Floquet transformation to remove the time dependence in the linearized homogenous differential equation.

Section 4.3: We transform the system such that the linear part consists of coupled harmonic oscillators and we describe the structure of the transformed reversibility matrix.

Section 4.4: We remove the linear term of the inhomogeneity.

Section 4.5: We present an overview of the transformations of the previous sections.

4.1 Local Coordinates

Let $\sigma^0$ be some fixed value of the parameter $\sigma$ and let $z_{\text{per}}^0$ denote the periodic solution of (3.1) for $\sigma = \sigma^0$ which fulfills the symmetry condition $z_{\text{per}}^0(-t) = R z_{\text{per}}^0(t)$ (cf. Assumption $A_4$):

$$z_{\text{per}}^0(t) := z_{\text{per}}(t, \sigma^0).$$

We introduce local coordinates by

$$\sigma =: \sigma^0 + \Delta \sigma,$$
$$z =: z_{\text{per}}^0(t) + z_1.$$
Substituting this into (3.1) we obtain
\[
\ddot{z}_\text{per} + \dot{z}_1 = = F(t, z^0_{\text{per}}(t) + z_1, \sigma^0 + \Delta \sigma) = \\
= F(t, z^0_{\text{per}}(t), \sigma^0) + (F(t, z^0_{\text{per}}(t) + z_1, \sigma^0 + \Delta \sigma) - F(t, z^0_{\text{per}}(t), \sigma^0)).
\]

Thus we have the following differential equation for \(z_1\):
\[
\ddot{z}_1 = F(t, z^0_{\text{per}}(t) + z_1, \sigma^0 + \Delta \sigma) - F(t, z^0_{\text{per}}(t), \sigma^0). \tag{4.1}
\]

Due to the reversibility of \(F\) and the symmetry condition of \(z^0_{\text{per}}\), the differential equation (4.1) is still reversible:
\[
F(-t, z^0_{\text{per}}(-t) + Rz_1, \sigma^0 + \Delta \sigma) - F(-t, z^0_{\text{per}}(-t), \sigma^0) = \\
= F(-t, R(z^0_{\text{per}}(t) + z_1), \sigma^0 + \Delta \sigma) - F(-t, Rz^0_{\text{per}}(t), \sigma^0) = \\
= -R(F(t, z^0_{\text{per}} + z_1, \sigma^0 + \Delta \sigma) - F(t, z^0_{\text{per}}, \sigma^0)).
\]

4.2 The Theory of Floquet

The right-hand side of (4.1) may be written in the following form
\[
\ddot{z}_1 = F^{10}(t)z_1 + F^{01}(t)\Delta \sigma + F^{22}(t, z_1, \Delta \sigma), \tag{4.2}
\]
where \(F^{10}\), \(F^{01} \Delta \sigma\) and \(F^{22}\) are \(T\)-periodic in \(t\).

Let \(Z_1\) denote the principal fundamental matrix solution of
\[
\dot{z}_1 = F^{10}(t)z_1, \tag{4.3}
\]
i.e.: \(\dot{Z}_1 = F^{10}(t)Z_1\) and \(Z_1(0) = I\). In Lemma 3.2.2 we showed that \(Z_1(T)\) is reversible:
\[
Z_1^{-1}(T) = RZ_1(T)R^{-1}.
\]

We start this section with two auxiliary lemmas.

**Lemma 4.2.1**

*By Assumption \(A_5\) there exists a matrix \(S\) that transforms the monodromy matrix \(Z_1(T)\) into block diagonal form:

\[
Z_2(T) := S^{-1}Z_1(T)S = \begin{pmatrix} Z_{2,1}(T) & & \\
& \ddots & \\
& & Z_{2,n}(T) \end{pmatrix}
\]

with

\[
Z_{2,i}(T) := \begin{pmatrix} \cos(\omega_i T) & \sin(\omega_i T) \\
-\sin(\omega_i T) & \cos(\omega_i T) \end{pmatrix}.
\]

Then the following holds:
4.2. The Theory of Floquet

(i) $Z_2(T)$ is reversible with a reversibility matrix $R_D := S^{-1}RS$:

$$Z_2^{-1}(T) = R_D Z_2(T) R_D^{-1}.$$  

(ii) $R_D$ is in the following block diagonal form:

$$R_D := \begin{pmatrix} R_1 & & \\ & \ddots & \\ & & R_n \end{pmatrix}, \quad \text{where} \quad R_i := r_i \begin{pmatrix} \sin \beta_i & \cos \beta_i \\ \cos \beta_i & -\sin \beta_i \end{pmatrix},$$

for suitably chosen $r_i > 0$ and $\beta_i$.

Proof:

(i) Using the definition of $Z_2(T)$ and $R_D$ we find

$$Z_2^{-1}(T) = (S^{-1}Z_1(T)S)^{-1} = S^{-1}Z_1^{-1}(T)S = S^{-1}(RZ_1(T)R^{-1})S =$$

$$= R_D S^{-1}Z_1(T)SR_D^{-1} = R_D Z_2(T) R_D^{-1}.$$  

(ii) For convenience we rewrite the reversibility condition $Z_2^{-1}(T) = R_D Z_2(T) R_D^{-1}$ as follows:

$$Z_2^{-1}(T) R_D = R_D Z_2(T). \quad (4.4)$$

Let

$$Z_2(T) := \begin{pmatrix} Z_{2,1}(T) \\ \vdots \\ Z_{2,n}(T) \end{pmatrix} \quad \text{with} \quad Z_{2,i}(T) := \begin{pmatrix} \cos(\omega_i T) & \sin(\omega_i T) \\ -\sin(\omega_i T) & \cos(\omega_i T) \end{pmatrix}$$

and

$$R_D := \begin{pmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{n1} & \cdots & R_{nn} \end{pmatrix} \quad \text{with} \quad R_{ij} := \begin{pmatrix} r_{ij}^{11} & r_{ij}^{12} \\ r_{ij}^{21} & r_{ij}^{22} \end{pmatrix}.$$  

Substituting these definitions into (4.4) we obtain on the left-hand side:

$$Z_2^{-1}(T) R_D = \begin{pmatrix} Z_{2,1}^{-1}(T) \\ \vdots \\ Z_{2,n}^{-1}(T) \end{pmatrix} \begin{pmatrix} R_{11} & \cdots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{n1} & \cdots & R_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} Z_{2,1}^{-1}(T) R_{11} & \cdots & Z_{2,1}^{-1}(T)^{-1} R_{1n} \\ \vdots & \ddots & \vdots \\ Z_{2,n}^{-1}(T) R_{n1} & \cdots & Z_{2,n}^{-1}(T) R_{nn} \end{pmatrix}.$$
and on the right-hand side

\[
R_{ij} Z_{2,j}(T) = \begin{pmatrix}
R_{11} & \ldots & R_{1n} \\
\vdots & & \vdots \\
R_{n1} & \ldots & R_{nn}
\end{pmatrix}
\begin{pmatrix}
Z_{2,1}(T) \\
\vdots \\
Z_{2,n}(T)
\end{pmatrix} =
\begin{pmatrix}
R_{11}Z_{2,1}(T) & \ldots & R_{1n}Z_{2,n}(T) \\
\vdots & & \vdots \\
R_{n1}Z_{2,1}(T) & \ldots & R_{nn}Z_{2,n}(T)
\end{pmatrix}.
\]

We conclude that the blocks satisfy the following equation:

\[
Z_{2,i}(T)^{-1} R_{ij} = R_{ij} Z_{2,j}(T).
\] (4.5)

Using the definitions of \(Z_{2,i}(T)\) and \(R_{ij}\) we obtain for the left-hand side of (4.5)

\[
Z_{2,i}^{-1}(T) R_{ij} = \begin{pmatrix}
\cos(\omega_i T) & -\sin(\omega_i T) \\
\sin(\omega_i T) & \cos(\omega_i T)
\end{pmatrix}
\begin{pmatrix}
r_{1i}^{11} & r_{1i}^{12} \\
r_{1i}^{21} & r_{1i}^{22}
\end{pmatrix}

= \begin{pmatrix}
r_{1i}^{11} \cos(\omega_i T) - r_{1i}^{21} \sin(\omega_i T) & \quad r_{1i}^{12} \cos(\omega_i T) - r_{1i}^{22} \sin(\omega_i T) \\
r_{1i}^{21} \sin(\omega_i T) + r_{1i}^{22} \cos(\omega_i T) & \quad r_{1i}^{21} \sin(\omega_i T) + r_{1i}^{22} \cos(\omega_i T)
\end{pmatrix}
\]

and for the right-hand side

\[
R_{ij} Z_{2,j}(T) = \begin{pmatrix}
r_{1j}^{11} & r_{1j}^{12} \\
r_{1j}^{21} & r_{1j}^{22}
\end{pmatrix}
\begin{pmatrix}
\cos(\omega_j T) & \sin(\omega_j T) \\
-\sin(\omega_j T) & \cos(\omega_j T)
\end{pmatrix}

= \begin{pmatrix}
r_{1j}^{11} \cos(\omega_j T) - r_{1j}^{12} \sin(\omega_j T) & \quad r_{1j}^{11} \sin(\omega_j T) + r_{1j}^{12} \cos(\omega_j T) \\
r_{1j}^{21} \sin(\omega_j T) - r_{1j}^{22} \cos(\omega_j T) & \quad r_{1j}^{21} \sin(\omega_j T) + r_{1j}^{22} \cos(\omega_j T)
\end{pmatrix}.
\]

Comparing both sides leads to the following system of equations:

\[
\begin{pmatrix}
c & b & -a & 0 \\
-b & c & 0 & -a \\
a & 0 & c & b \\
0 & a & -b & c
\end{pmatrix}
\begin{pmatrix}
r_{1j}^{11} \\
r_{1j}^{12} \\
r_{1j}^{21} \\
r_{1j}^{22}
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
0 \\
0
\end{pmatrix}
\] (4.6)

with

\[a := \sin \omega_i T, \quad b := \sin \omega_j T, \quad c := \cos \omega_i T - \cos \omega_j T.\]

We distinguish between two cases:

\* \(i \neq j\):

We compute the determinant of (4.6) to see whether the trivial solution is unique.

\[
det
\begin{pmatrix}
c & b & -a & 0 \\
-b & c & 0 & -a \\
a & 0 & c & b \\
0 & a & -b & c
\end{pmatrix}
\]
4.2. The Theory of Floquet

\[\begin{align*}
&= (-1)^{1+1} \cdot c \cdot \det \begin{pmatrix}
  c & 0 & -a \\
  0 & c & b \\
  a & -b & c 
\end{pmatrix}
+ (-1)^{1+2} \cdot b \cdot \det \begin{pmatrix}
  -b & 0 & -a \\
  a & c & b \\
  0 & -b & c 
\end{pmatrix}
+ (-1)^{1+3} \cdot (-a) \cdot \det \begin{pmatrix}
  -b & c & -a \\
  a & 0 & b \\
  0 & a & c 
\end{pmatrix}
+ (-1)^{1+4} \cdot 0 = \\
&= c \cdot (c^3 + a^2c + b^2c) - b \cdot (-bc^2 + a^2b - b^3) - a \cdot (-a^3 - ac^2 + ab^2) = \\
&= a^4 + b^4 + c^4 - 2a^2b^2 + 2a^2c^2 + 2b^2c^2 = ((a + b)^2 + c^2)((a - b)^2 + c^2).
\end{align*}\]

Thus the determinant vanishes if and only if \(c = 0\) and \(a \pm b = 0\) or expressed in the original variables

\[
\cos \omega_i T = \cos \omega_j T \\
\sin \omega_i T = \pm \sin \omega_j T.
\]

Equations (4.7) hold if and only if \(\omega_i T = \pm \omega_j T + k_0 \cdot 2\pi\) for some \(k_0 \in \mathbb{Z}\). Substituting \(T\) by \(\frac{2\pi}{\omega_0}\) we obtain

\[
2\pi \frac{\omega_i}{\omega_0} = \pm 2\pi \frac{\omega_j}{\omega_0} + k_0 \cdot 2\pi
\]

and further

\[
0 = \omega_i \mp \omega_j - k_0 \omega_0
\]

in contradiction to the non-resonance condition in Assumption A_{5b}. Thus the determinant does not vanish and equation (4.6) has only the trivial solution \(r_{ij} = r_{ij} = r_{ij} = r_{ij} = 0\), i.e.

\[
R_{ij} = 0 \quad \text{for} \quad i \neq j.
\]

* \(i = j\):

In this case equation (4.6) is simplified to

\[
\begin{pmatrix}
  0 & a & -a \\
  -a & 0 & 0 \\
  a & 0 & 0 \\
  0 & a & -a 
\end{pmatrix}
\begin{pmatrix}
  r_{11}^{11} \\
  r_{12}^{12} \\
  r_{21}^{21} \\
  r_{22}^{22} 
\end{pmatrix}
= \begin{pmatrix}
  0 \\
  0 \\
  0 \\
  0 
\end{pmatrix}.
\]

The parameter \(a\) vanishes if and only if \(\omega_i T = k_0 \pi\), for some \(k_0 \in \mathbb{Z}\). Substituting \(T\) by \(\frac{2\pi}{\omega_0}\) we obtain

\[
2\pi \frac{\omega_i}{\omega_0} = k_0 \cdot \pi
\]

and further

\[
0 = 2\omega_i - k_0 \omega_0
\]
in contradiction to the non-resonance condition. Thus $a \neq 0$ and it follows that $r_{ii}^{22} = -r_{ii}^{11}$ and $r_{ii}^{21} = r_{ii}^{12}$, i.e.

$$R_{ii} = \begin{pmatrix} r_{ii}^{11} & r_{ii}^{12} \\ r_{ii}^{21} & r_{ii}^{22} \end{pmatrix}.$$ 

Since $R$ is invertible we have

$$0 \neq \det R = \det R_D = \prod_{i=1}^{n} \det R_{ii} = (-1)^n \prod_{i=1}^{n} ((r_{ii}^{11})^2 + (r_{ii}^{12})^2).$$

We may therefore put:

$$R_i := R_{ii} = r_i \begin{pmatrix} \sin \beta_i & \cos \beta_i \\ \cos \beta_i & -\sin \beta_i \end{pmatrix} \quad \text{with} \quad r_i := \sqrt{(r_{ii}^{11})^2 + (r_{ii}^{12})^2}.$$ 

Thus $R_D$ has the form stated in the claim of the lemma.

\[ \square \]

**Lemma 4.2.2**

Let $Z_1(T)$ be the monodromy matrix of (4.3) and let

$$H_{m}^{10} := \begin{pmatrix} H_{m,1}^{10} \\ \vdots \\ H_{m,n}^{10} \end{pmatrix}, \quad m \in \mathbb{Z}$$

with

$$H_{m,i}^{10} := \begin{pmatrix} 0 & \omega_i + m \frac{2\pi}{T} \\ -\omega_i - m \frac{2\pi}{T} & 0 \end{pmatrix}.$$ 

Then the following holds:

(i) The transformed monodromy matrix $Z_2(T)$ may be written as

$$Z_2(T) = e^{H_{m}^{10}T}.$$ 

(ii) The matrices $H_{m}^{10}$ are reversible with respect to the reversibility matrix $R_D$ introduced in Lemma 4.2.1:

$$H_{m}^{10}R_D + R_DH_{m}^{10} = 0.$$ 

(iii) The original monodromy matrix $Z_1(T)$ may be written as

$$Z_1(T) = e^{G_{m}^{10}T},$$

where

$$G_{m}^{10} := SH_{m}^{10}S^{-1}.$$ 

(iv) The matrices $G_{m}^{10}$ are reversible with respect to the original reversibility matrix $R$

$$G_{m}^{10}R + RG_{m}^{10} = 0.$$
4.2. The Theory of Floquet

Proof:

(i) For a single block of $H_{m}^{10}$ we find

$$e^{H_{m,i}^{10}T} = \sum_{k=0}^{\infty} \frac{1}{k!} \left( \begin{array}{cc} 0 & \omega_i T + 2\pi m \\ -\omega_i T - 2\pi m & 0 \end{array} \right)^k =$$

$$= \sum_{k=0}^{\infty} \frac{1}{(2k)!} \left( \begin{array}{cc} 0 & \omega_i T + 2\pi m \\ -\omega_i T - 2\pi m & 0 \end{array} \right)^{2k} +$$

$$+ \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \left( \begin{array}{cc} 0 & \omega_i T + 2\pi m \\ -\omega_i T - 2\pi m & 0 \end{array} \right)^{2k+1} =$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \left( \begin{array}{cc} \omega_i T + 2\pi m & 0 \\ 0 & \omega_i T + 2\pi m \end{array} \right)^{2k} +$$

$$+ \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} \left( \begin{array}{cc} \omega_i T + 2\pi m & 0 \\ 0 & \omega_i T + 2\pi m \end{array} \right)^{2k+1} =$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \sum_{\sigma=0}^{\infty} \frac{1}{(2k+1)!} \left( \begin{array}{cc} \omega_i T + 2\pi m & 0 \\ 0 & \omega_i T + 2\pi m \end{array} \right)^{2k+1} =$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} \sum_{\sigma=0}^{\infty} \frac{1}{(2k+1)!} \left( \begin{array}{cc} \omega_i T + 2\pi m & 0 \\ 0 & \omega_i T + 2\pi m \end{array} \right)^{2k+1} =$$

$$= \left( \begin{array}{cc} \cos(\omega_i T + 2\pi m) & \sin(\omega_i T + 2\pi m) \\ -\sin(\omega_i T + 2\pi m) & \cos(\omega_i T + 2\pi m) \end{array} \right) =$$

$$= Z_{2,i}(T).$$

Thus the full matrices $H_{m}^{10}$ satisfy

$$e^{H_{m}^{10}T} = Z_{2}(T).$$

(ii) Lemma 4.2.1 implies in that

$$H_{m}^{10}R_{D} + R_{D}H_{m}^{10} = \left( \begin{array}{cc} H_{m,1}^{10} + R_{1}H_{m,1}^{10} & \cdots \\ \cdots & \cdots \\ H_{m,n}^{10} + R_{n}H_{m,n}^{10} \end{array} \right).$$

By the definitions of $H_{m,i}^{0}$ and $R_{i}$ we find for a single block

$$H_{m,i}^{10}R_{i} + R_{i}H_{m,i}^{10} = \left( \begin{array}{cc} 0 & \omega_{i} + m_{i} \frac{2\pi}{T} \\ -\omega_{i} - m_{i} \frac{2\pi}{T} & 0 \end{array} \right) r_{i} \left( \begin{array}{cc} \sin \beta_{i} & \cos \beta_{i} \\ \cos \beta_{i} & -\sin \beta_{i} \end{array} \right) +$$

$$+ r_{i} \left( \begin{array}{cc} \sin \beta_{i} & \cos \beta_{i} \\ \cos \beta_{i} & -\sin \beta_{i} \end{array} \right) \left( \begin{array}{cc} 0 & \omega_{i} + m_{i} \frac{2\pi}{T} \\ -\omega_{i} - m_{i} \frac{2\pi}{T} & 0 \end{array} \right) =$$
4. Preliminary Coordinate Transformations

\[ r_i \left( (\omega_i + m \frac{2\pi}{T}) \cos \beta_i - (\omega_i + m \frac{2\pi}{T}) \sin \beta_i \right) + \]

\[ + r_i \left( (\omega_i + m \frac{2\pi}{T}) \sin \beta_i - (\omega_i + m \frac{2\pi}{T}) \cos \beta_i \right) = 0. \]

(iii) Taken into account the definition of \( Z_2(T) \) (cf. Lemma 4.2.1) it follows from (i) that
\[ e^{G_m^{10}T} = e^{SH_{m}^{10}S^{-1}T} = S \cdot Z_2(T) \cdot S^{-1} = S \cdot S^{-1} \cdot Z_1(T) \cdot S \cdot S^{-1} = Z_1(T). \]

(iv) Using the definitions of \( R_D \) and \( H_{m}^{10} \) and (ii) we find
\[ G_{m}^{10} R + R G_{m}^{10} = (SH_{m}^{10}S^{-1})(SR_{D}S^{-1}) + (SR_{D}S^{-1})(SH_{m}^{10}S^{-1}) = \]
\[ = S^{-1} \cdot 0 = 0. \]

This completes the proof of this lemma. \( \square \)

Now we are in a position to state the central theorem of this section.

**Theorem 4.2.1 (Reversible Floquet Transformation)**

*Let*
\[ \dot{z}_1 = F^{10}(t)z_1 + F^{01}(t)\Delta \sigma + F^{22}(t, z_1, \Delta \sigma), \quad (4.2) \]

*where* \( F^{10}, F^{01} \) \( \text{and} F^{22} \) *are* \( T \)-*periodic in* \( t \).

*Then there exists a function* \( Q \) *such that*
\[ z_1 =: Q(t)z_2 \]

*transforms (4.2) to*
\[ \dot{z}_2 = G^{10}z_2 + G^{01}(t)\Delta \sigma + G^{22}(t, z_2, \Delta \sigma), \quad (4.8) \]

*where* \( e^{G_{m}^{10}T} = Z_1(T) \) \( \text{(cf. Lemma 4.2.2)} \) *and*
\[ G^{01}(t) := Q^{-1}(t)F^{01}(t), \]
\[ G^{22}(t, z_2, \Delta \sigma) := Q^{-1}(t)F^{22}(t, Q(t)z_2, \Delta \sigma). \]

*Moreover*

(i) \( Q(0) = I \),

(ii) \( Q \) is \( T \)-*periodic in* \( t \): \( Q(T + t) = Q(t) \),

(iii) \( Q \) is time-reversible: \( Q(-t) = RQ(t)R^{-1} \),
(iv) the right-hand side of (4.8) is $T$-periodic:

$$G^{01}(t + T)\Delta \sigma = G^{01}(t)\Delta \sigma,$$

$$G^{22}(t + T, z_2, \Delta \sigma) = G^{22}(t, z_2, \Delta \sigma),$$

(v) (4.8) is time-reversible:

$$G^{10}R = -RG^{10},$$

$$G^{01}(-t)\Delta \sigma = -RG^{01}(t)\Delta \sigma,$$

$$G^{22}(-t, Rz_2, \Delta \sigma) = -RG^{22}(t, z_2, \Delta \sigma).$$

Proof: From Lemma 4.2.2 (iii) we know that there exists a matrix $G^{10}$ such that $Z_1(T) = e^{G^{10}T}$. Let

$$Q(t) := Z_1(t)e^{-G^{10}t}.$$

In a first step we show that $Q(t)$ has the properties claimed in the theorem.

(i) $Q(0) = I$:

$$Q(0) = Z_1(0)e^{-G^{10}0} = I \cdot I = I.$$

(ii) Periodicity:

By the definition of $Q$ and $G^{10}$ we obtain

$$Q(t + T) = Z_1(t + T)e^{-G^{10}(t+T)} = Z_1(t)Z_1(T)e^{-G^{10}T}e^{-G^{10}t} = Z_1(t)e^{-G^{10}t} = Q(t).$$

(iii) Reversibility:

The reversibility of $G^{10}$ is proved in Lemma 4.2.2 (iv). For $Q(t)$ we obtain

$$Q(-t) = Z_1(-t)e^{-G^{10}(-t)} = RZ_1(t)R^{-1}e^{-RG^{10}R^{-1}t} = RZ_1(t)R^{-1}Re^{-G^{10}t}R^{-1} = RZ_1(t)e^{-G^{10}t}R^{-1} = RQ(t)R^{-1}.$$

In a second step we compute the transformed differential equation. To this end we substitute $z_1 = Q(t)z_2$ into (4.2):

$$\dot{Q}(t)z_2 + Q(t)\dot{z}_2 = F^{10}(t)Q(t)z_2 + F^{01}(t)\Delta \sigma + F^{22}(t, Q(t)z_2, \Delta \sigma).$$

(4.9)

By the definition of $Q$ we have

$$\dot{Q}(t) = \dot{Z}_1(t)e^{-G^{10}t} - Z_1(t)e^{-G^{10}t}G^{10} = F^{10}(t)Q(t) - Q(t)G^{10}.$$

Substituting this into (4.9) we obtain

$$-Q(t)G^{10}z_2 + Q(t)\dot{z}_2 = F^{01}(t)\Delta \sigma + F^{22}(t, Q(t)z_2, \Delta \sigma).$$
and furthermore

\[ z_2 = G^{10}z_2 + Q^{-1}(t)F^{01}(t)\Delta \sigma + Q^{-1}(t)F^{22}(t, Q(t)z_2, \Delta \sigma) = \]

\[ = G^{10}z_2 + G^{01}(t)\Delta \sigma + G^{22}(t, z_2, \Delta \sigma). \]

In a final step we prove that the transformed differential equation has the properties claimed in the theorem.

(iv) \(T\)-periodicity:

It follows immediately from the definition of \(F^{01}\) and \(F^{22}\) and the periodicity of \(Q\) that

\[ G^{01}(t + T) = Q^{-1}(t + T)F^{01}(t + T) = \]

\[ = Q^{-1}(t)F^{01}(t) = \]

\[ = G^{01}(t), \]

\[ G^{22}(t + T, z_2, \Delta \sigma) = Q^{-1}(t + T)F^{22}(t + T, Q(t + T)z_2, \Delta \sigma) = \]

\[ = Q^{-1}(t)F^{22}(t, Q(t)z_2, \Delta \sigma) = \]

\[ = G^{22}(t, z_2, \Delta \sigma). \]

(v) Reversibility:

We have already proved in Lemma 4.2.2 (iv) that \(G^{10}\) satisfies the reversibility condition. It remains to prove that \(G^{01}(t)\) and \(G^{22}(t, z_2, \Delta \sigma)\) fulfill the reversibility condition too.

By the definition of \(G^{01}(t)\Delta \sigma\) and \(G^{22}(t, z_2, \Delta \sigma)\) and the reversibility of \(Q(t, \Delta \sigma)\), \(F^{01}(t)\) and \(F^{22}(t, z_1, \Delta \sigma)\) we obtain

\[ G^{01}(-t) = Q^{-1}(-t)F^{01}(-t) = -(RQ(t)R^{-1})^{-1}RF^{01}(t) = \]

\[ = -RQ^{-1}(t)R^{-1}RF^{01}(t) = \]

\[ = -RQ^{-1}(t)F^{01}(t) = -RG^{01}(t), \]

\[ G^{22}(-t, Rz_2, \Delta \sigma) = Q^{-1}(-t)F^{22}(-t, Q(-t)Rz_2, \Delta \sigma) = \]

\[ = (RQ(t)R^{-1})^{-1}F^{22}(-t, Q(-t)Rz_2, \Delta \sigma) = \]

\[ = RQ^{-1}(t)R^{-1}F^{22}(-t, RQ(t)R^{-1}Rz_2, \Delta \sigma) = \]

\[ = -RQ^{-1}(t)R^{-1}RF^{22}(t, Q(t)z_2, \Delta \sigma) = \]

\[ = -RQ^{-1}(t)F^{22}(t, Q(t)z_2, \Delta \sigma) = \]

\[ = -RG^{22}(t, z_2, \Delta \sigma). \]
4.3. The Structure of the Linear System

Recall the transformation matrix $S$ introduced in Lemma 4.2.1. By

$$z_2 := S z_3$$

Eq. (4.8) is transformed to

$$\dot{z}_3 = H^{10} z_3 + H^{01}(t) \Delta \sigma + H^{22}(t, z_3, \Delta \sigma), \quad (4.10)$$

where

$$H^{10} := S^{-1} G^{10} S = \begin{pmatrix} H_1^{10} & \cdots & \vdots \\ \vdots & \ddots & \vdots \\ H_n^{10} \end{pmatrix} \quad \text{with} \quad H_j^{10} := \begin{pmatrix} 0 & \omega_j \\ -\omega_j & 0 \end{pmatrix},$$

$$H^{01}(t) := S^{-1} G^{01}(t),$$

$$H^{22}(t, z_3, \Delta \sigma) := S^{-1} G^{22}(t, S z_3, \Delta \sigma).$$

The transformed system (4.10) is still reversible.

Lemma 4.3.1

Let (4.8) be reversible. Then the following holds:

(i) (4.10) is reversible with the reversibility matrix $R_D := S^{-1} R S$:

$$H^{10} R_D = -R_D H^{10},$$

$$H^{01}(-t) = -R_D H^{01}(t),$$

$$H^{22}(-t, R_D z_3, \Delta \sigma) = -R_D H^{22}(t, z_3, \Delta \sigma).$$

(ii) $R_D$ is in the following block diagonal form

$$R_D = \begin{pmatrix} R_1 \\ \vdots \\ R_n \end{pmatrix}, \quad \text{where} \quad R_j := r_j \begin{pmatrix} \sin \beta_j & \cos \beta_j \\ \cos \beta_j & -\sin \beta_j \end{pmatrix}. $$
Proof: The reversibility of $H^{10}$ is proved in Lemma 4.2.2 (ii). By the definition of $H^{01}$ and $R_D$ and the reversibility condition of $G^{01}$ we find

$$H^{01}(-t) = S^{-1}G^{01}(-t) = -S^{-1}RG^{01}(t) = -R_D S^{-1}G^{01}(t) = -R_D H^{01}(t).$$

By Lemma 4.2.1 and Theorem 4.2.1 we obtain

$$H^{22}(-t, R_D z_3, \Delta \sigma) = S^{-1}G^{22}(-t, S R_D z_3, \Delta \sigma) =
S^{-1}G^{22}(-t, R S z_3, \Delta \sigma) = -S^{-1}RG^{22}(t, S z_3, \Delta \sigma) =
-R_D S^{-1}G^{22}(t, S z_3, \Delta \sigma) = -R_D H^{22}(t, z_3, \Delta \sigma).$$

The second part of the lemma is proved in Lemma 4.2.1.

4.4 Removing the Linear Term of the Inhomogeneity

In general the linear term in $\Delta \sigma$ of (4.10) does not vanish. In this case we may apply an additional transformation to make $H^{01}$ vanish. Let

$$z_3 := z_3^0(t) \Delta \sigma + z_4, \quad (4.11)$$

where

$$z_3^0 : \mathbb{R} \rightarrow L(\mathbb{R}^m, \mathbb{R}^{2n})$$

is a $T$-periodic matrix function to be determined.

If we substitute (4.11) into (4.10), then we obtain

$$z_3^0(t) \Delta \sigma + z_4 = H^{10}(z_3^0(t) \Delta \sigma + z_4) + H^{01}(t) \Delta \sigma + H^{22}(t, z_3^0(t) \Delta \sigma + z_4, \Delta \sigma) =
(H^{10} z_3^0(t) + H^{01}(t)) \Delta \sigma + H^{10} z_4 + H^{22}(t, z_3^0(t) \Delta \sigma + z_4, \Delta \sigma).$$

Thus to remove the linear term in $\Delta \sigma$ we have to look for a $T$-periodic solution of the following differential equation:

$$\dot{z}_3^0 = H^{10} z_3^0 + H^{01}(t). \quad (4.12)$$

The existence of a $T$-periodic solution is the content of the next lemma.

Lemma 4.4.1

Let $H^{01}(t) =: \sum_{k=0}^{\infty} H^{01,k} e^{i k \omega_0 t}$ be the Fourier expansion of $H^{01}$, where $\omega_0 := \frac{2 \pi}{T}$.

The general solution of (4.12) is given by

$$z_3^0(t) := e^{H^{10} t} z_3^0(0) + \int_0^t e^{H^{10}(t-s)} H^{01}(s) ds. \quad (4.13)$$
4.4. Removing the Linear Term of the Inhomogeneity

For the initial value \( z^0_3(0) \) defined by

\[
z^0_3(0) := \sum_{k_0=-\infty}^{\infty} (-H^{10} + ik_0\omega_0I)^{-1} H^{01,k_0}
\]  

(4.14)

the function \( z^0_3 \) is \( T \)-periodic and has the following symmetry property:

\[
z^0_3(-t) = R_D z^0_3(t).
\]

**Proof:** The Variation-of-Constant-Formula leads immediately to (4.13). It remains to show that \( z^0_3 \) is \( T \)-periodic and satisfies the symmetry condition \( R_D z^0_3(-t) = z^0_3(t) \).

Substituting the Fourier expansion of \( H^{01} \) into (4.13) we have

\[
z^0_3(t) = e^{H^{10}t} z^0_3(0) + \int_0^t e^{H^{10}(t-s)} H^{01}(s) ds = \\
e^{H^{10}t} z^0_3(0) + \int_0^t e^{H^{10}(t-s)} \left( \sum_{k_0=-\infty}^{\infty} e^{ik_0\omega_0s} H^{01,k_0} \right) ds.
\]

For sufficiently regular \( H^{01} \), e.g. \( H^{01} \in C^m, m \geq 3 \) the Fourier coefficients \( H^{01,k_0} \) are of order \( k_0^{-m} \) (cf. [3], p. 344). Thus we may interchange integral and sum:

\[
z^0_3(t) = e^{H^{10}t} z^0_3(0) + \sum_{k_0=-\infty}^{\infty} e^{H^{10}t} \cdot \int_0^t e^{(-H^{10} + ik_0\omega_0I)s} ds \cdot H^{01,k_0}.
\]

The exponent is in block diagonal form (cf. Lemma 4.2.2). This allows us to consider a single block:

\[
-H^{10}_j + ik_0\omega_0I_2 = \begin{pmatrix} 0 & \omega_j \\ -\omega_j & 0 \end{pmatrix} + ik_0\omega_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} ik_0\omega_0 & -\omega_j \\ \omega_j & ik_0\omega_0 \end{pmatrix}.
\]

For the determinant of a single block we obtain

\[
\det \begin{pmatrix} ik_0\omega_0 & -\omega_j \\ \omega_j & ik_0\omega_0 \end{pmatrix} = -k_0^2\omega_0^2 + \omega_j^2 = (\omega_j + k_0\omega_0)(\omega_j - k_0\omega_0).
\]

Due to the non-resonance condition the determinant of a single block does not vanish. Thus we have

\[
(-H_j^{10} + ik_0\omega_0I_2)^{-1} = \frac{1}{-k_0^2\omega_0^2 + \omega_j^2} \begin{pmatrix} ik_0\omega_0 & \omega_j \\ \omega_j & ik_0\omega_0 \end{pmatrix}.
\]  

(4.15)

It follows that the full matrix is invertible. This allows us to compute the integral:

\[
\int_0^t e^{(-H^{10} + ik_0\omega_0I)s} ds = (-H^{10} + ik_0\omega_0I)^{-1} \left( e^{(-H^{10} + ik_0\omega_0I)t} - I \right).
\]
We therefore have
\[ z_3^0(t) = e^{H^{10} t} z_3^0(0) + \sum_{k_0=-\infty}^{\infty} e^{H^{10} t} \cdot (-H^{10} + i k_0 \omega_0 I)^{-1} \left( e^{(-H^{10} + i k_0 \omega_0 I) t} - I \right) H^{01, k_0} = \]
\[ = e^{H^{10} t} z_3^0(0) + \sum_{k_0=-\infty}^{\infty} (-H^{10} + i k_0 \omega_0 I)^{-1} \left( e^{i k_0 \omega_0 t} - e^{H^{10} t} \right) H^{01, k_0} = \]
\[ = e^{H^{10} t} z_3^0(0) + \sum_{k_0=-\infty}^{\infty} \left(-(-H^{10} + i k_0 \omega_0 I)^{-1} e^{H^{10} t} H^{01, k_0} + \right. \]
\[ \left. + (-H^{10} + i k_0 \omega_0 I)^{-1} e^{i k_0 \omega_0 t} H^{01, k_0} \right). \]

From the order of the Fourier coefficients \( H^{01, k_0} \) and from (4.15) it follows that \((-H^{10} + i k_0 \omega_0 I)^{-1} H^{01, k_0} \) is of order \( k_0^{-(m+1)} \). Thus we find that the series
\[ \sum_{k_0=-\infty}^{\infty} (-H^{10} + i k_0 \omega_0 I)^{-1} e^{i k_0 \omega_0 t} H^{01, k_0} \]

and
\[ \sum_{k_0=-\infty}^{\infty} (-H^{10} + i k_0 \omega_0 I)^{-1} e^{H^{10} t} H^{01, k_0} \]
converge. Thus we have
\[ z_3^0(t) = e^{H^{10} t} \left( z_3^0(0) - \sum_{k_0=-\infty}^{\infty} (-H^{10} + i k_0 \omega_0 I)^{-1} H^{01, k_0} \right) + \]
\[ + \sum_{k_0=-\infty}^{\infty} (-H^{10} + i k_0 \omega_0 I)^{-1} e^{i k_0 \omega_0 t} H^{01, k_0} \]

and (4.14) implies that \( z_3^0 \) is \( T \)-periodic.

Now we show that \( z_3^0(0) \) given in (4.14) satisfies the symmetry condition \( R_D z_3^0(0) = z_3^0(0) \).

From the symmetry condition \( H^{01}(-t) = -R_D H^{01}(t) \) we derive a symmetry condition for the Fourier coefficients of \( H^{01}(t) \). On the left-hand side we have
\[ H^{01}(-t) = \sum_{k_0=-\infty}^{\infty} H^{01, k_0} e^{-i k_0 \omega_0 t} = \sum_{k_0=-\infty}^{\infty} H^{01, -k_0} e^{i k_0 \omega_0 t} \]

and on the right-hand side
\[ -R_D H^{01}(t) = -R_D \sum_{k_0=-\infty}^{\infty} H^{01, k_0} e^{i k_0 \omega_0 t} = \sum_{k_0=-\infty}^{\infty} (-R_D H^{01, k_0}) e^{i k_0 \omega_0 t}. \]
Comparing both sides we find that
\[ H^{01,-k_0} = -R_D H^{01,k_0}. \]

Using this result and Lemma 4.3.1 (i) we obtain
\[
R_D z_3^0(0) = \sum_{k_0=-\infty}^{\infty} R_D \left(-H^{10} + i k_0 \omega_0 I\right)^{-1} H^{01,k_0} = \sum_{k_0=-\infty}^{\infty} \left(H^{10} + i k_0 \omega_0 I\right)^{-1} R_D H^{01,k_0} =
\]
\[
= - \sum_{k_0=-\infty}^{\infty} \left(H^{10} + i k_0 \omega_0 I\right)^{-1} H^{01,-k_0} = - \sum_{k_0=-\infty}^{\infty} \left(H^{10} - i k_0 \omega_0 I\right)^{-1} H^{01,k_0} =
\]
\[
= \sum_{k_0=-\infty}^{\infty} \left(-H^{10} + i k_0 \omega_0 I\right)^{-1} H^{01,k_0} = z_3^0(0).
\]

Finally we show that the $T$-periodic solution $z_3^0$ fulfills the symmetry condition $z_3^0(-t) = R_D z_3^0(t)$:
\[
R_D z_3^0(t) = R_D e^{H^{10} t} z_3^0(0) + \int_0^t R_D e^{H^{10} (t-s)} H^{01}(s) ds =
\]
\[
= e^{H^{10} t} R_D z_3^0(0) + \int_0^t e^{H^{10} (t-s)} R_D H^{01}(s) ds =
\]
\[
= e^{H^{10} t} z_3^0(0) + \int_0^t e^{H^{10} (t-s)} H^{01}(-s) ds =
\]
\[
= e^{H^{10} t} z_3^0(0) + \int_0^{-t} e^{H^{10} (t-s)} H^{01}(s) ds =
\]
\[
= z_3^0(-t).
\]

This completes the proof. \( \square \)

From this lemma we conclude that it is possible to remove the linear term in $\Delta \sigma$ by the transformation
\[ z_3 = z_3^0(t) \Delta \sigma + z_4. \]

We obtain
\[ \dot{z}_4 = H^{10} z_4 + H^{2}(-t, z_3^0(t) \Delta \sigma + z_4, \Delta \sigma). \] (4.16)

It remains to show that (4.16) is still reversible:
\[ H^{10} R_D z_4 + H^{2}(-t, z_3^0(-t) \Delta \sigma + R_D z_4, \Delta \sigma) =
\]
\[ = H^{10} R_D z_4 + H^{2}(-t, R_D (z_3^0(t) \Delta \sigma + z_4), \Delta \sigma) =
\]
\[ = - R_D H^{10} z_4 - R_D H^{2}(-t, z_3^0(t) \Delta \sigma + z_4, \Delta \sigma). \]
4. Preliminary Coordinate Transformations

4.5 The Complete Transformation

The results obtained in this chapter are summarized in the following theorem.

**Theorem 4.5.1**

Consider the system of differential equations (3.1):

\[ \dot{z} = F(t, z, \sigma) \tag{3.1} \]

and let Assumptions A1–A5 hold (cf. Section 3.1). Then the following holds:

(i) The system of differential equations (3.1) may be transformed to a system of perturbed harmonic oscillators:

\[ \begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t, x, y, \Delta \sigma) \\
\dot{y}_i &= -\omega_i x_i + h_{2,i}(t, x, y, \Delta \sigma)
\end{align*} \tag{4.17} \]

(ii) The functions \( h_{1,i}(t, x, y, \Delta \sigma) \) and \( h_{2,i}(t, x, y, \Delta \sigma) \)

- are \( T \)-periodic in \( t \):
  \[ h_{m,i}(t + T, x, y, \Delta \sigma) = h_{m,i}(t, x, y, \Delta \sigma), \quad m = 1, 2, \]

- are of order 2 or higher in \( x \), \( y \) and \( \Delta \sigma \),

- are time-reversible with reversibility matrix \( R_D \) (cf. Lemma 4.2.1):
  \[ \begin{pmatrix}
  h_{1,i}(-t, R_D z_4, \Delta \sigma) \\
  h_{2,i}(-t, R_D z_4, \Delta \sigma)
  \end{pmatrix} = -R_i \begin{pmatrix}
  h_{1,i}(t, z_4, \Delta \sigma) \\
  h_{2,i}(t, z_4, \Delta \sigma)
  \end{pmatrix} \]

where \( z_4 =: (x_1, y_1, \ldots, x_n, y_n)^T \).

(iii) The transformation is given by

\[
\begin{pmatrix} z \\ \sigma \end{pmatrix} \mapsto \begin{pmatrix} z_4 \\ \Delta \sigma \end{pmatrix} := \begin{pmatrix} S^{-1}Q(t)(z(t, \sigma) - z_{\text{per}}(t)) - z_3^0(t)(\sigma - \sigma^0) \end{pmatrix}, \tag{4.18}
\]

where

- \( \sigma^0 \) is a fixed value of the parameter \( \sigma \) (cf. Assumption A4),

- \( z_{\text{per}}^0 \) is a periodic solution of (3.1) for \( \sigma = \sigma^0 \),

- \( Q(t) \) is the matrix of the Floquet transformation defined in Theorem 4.2.1,

- \( z_3^0 \) is a matrix function given in Lemma 4.4.1,

- \( S \) is the matrix of the transformation that takes the monodromy matrix of the linearization of (3.1) to block diagonal form (cf. Lemma 4.3.1).

(iv) The back transformation is given by

\[
\begin{pmatrix} z_4 \\ \Delta \sigma \end{pmatrix} \mapsto \begin{pmatrix} z \\ \sigma \end{pmatrix} = \begin{pmatrix} z(t, \sigma) = z_{\text{per}}^0(t) + Q(t)S(z_4(t, \Delta \sigma) + z_3^0(t)(\sigma - \sigma^0)) \end{pmatrix}, \tag{4.19}
\]
4.5. The Complete Transformation

(v) Let \( z(t, \sigma) \) and \( z_4(t, \Delta \sigma) \), respectively, be corresponding solutions of equations 3.1 and 4.17, respectively, then for integer multiples of the period \( T \) we have

\[
    z(nT, \sigma) = z^0_{\text{per}}(0) + S \left( z_4(nT, \sigma - \sigma^0) + z_3^0(0)(\sigma - \sigma^0) \right)
\]

\[
    z_4(nT, \Delta \sigma) = S^{-1} \left( z(nT, \sigma^0 + \Delta \sigma) - z^0_{\text{per}}(0) \right) - z_3^0(0) \Delta \sigma
\]

for \( n \in \mathbb{N} \).

(vi) Let

\[
    F(t, z^0_{\text{per}}(t) + z_1, \sigma^0 + \Delta \sigma) := F^{10} z_1 + F^{01} \Delta \sigma + F^{22}(t, z_1, \Delta \sigma),
\]

then

\[
    \begin{pmatrix}
        h_{1,1}(t, z_4, \Delta \sigma) \\
        h_{1,2}(t, z_4, \Delta \sigma) \\
        \vdots \\
        h_{n,1}(t, z_4, \Delta \sigma) \\
        h_{n,1}(t, z_4, \Delta \sigma)
    \end{pmatrix}
    := H^{22}(t, z_4, \Delta \sigma) :=
    S^{-1}Q^{-1}(t)F^{22}(t, Q(t)z^0_4 \Delta \sigma + Q(t)z_4, \Delta \sigma).
Formal First Integrals

Remember system (3.1) subject to the basic Assumptions $A_1 - A_5$. In view of the work of Chapter 4 (cf. Theorem 4.5.1) system (3.1) can be transformed to the following system of periodically perturbed harmonic oscillators:

$$
\begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t, x, y, \Delta \sigma) \\
\dot{y}_i &= -\omega_i x_i + h_{2,i}(t, x, y, \Delta \sigma) \\
\end{align*}
$$

where $1 \leq i \leq n$, (5.1)

where the functions $h_{i,j}(t, x, y, \Delta \sigma)$ are

- $T$-periodic in $t$: $h_{i,j}(t + T, x, y, \Delta \sigma) = h_{i,j}(t, x, y, \Delta \sigma)$,
- of order 2 or higher in $x, y$ and $\Delta \sigma$,
- time-reversible with reversibility matrix $R_D$ (cf. Lemma 4.2.1):

$$
(h_{1,i}(-t, R_D z_4, \Delta \sigma)) = -R_i (h_{2,i}(-t, R_D z_4, \Delta \sigma)) \\
$$

where $z_4 := (x_1, y_1, \ldots, x_n, y_n)^T$.

Note that system (5.1) admits 0 as equilibrium solution reflecting the fact that (3.1) admits the periodic solution $z_{\text{per}}^0$ for $\sigma = \sigma^0$. To study the stability behavior of $z_{\text{per}}^0$ we may study the stability properties of 0 of system (5.1). The basic strategy following Giorgilli (cf. [7],[8]) is to use suitable truncations of formal first integrals of system (5.1).

The purpose of this chapter is to construct these formal integrals and to study their algebraic properties.

The chapter is organized as follows:

**Section 5.1:** We define $n$ formal first integrals of (5.1)

$$
I_i(t, x, y, \Delta \sigma) := \sum_{k=2}^{\infty} I^k_i(t, x, y, \Delta \sigma).
$$

**Section 5.2:** The $I^k_i$ happen to satisfy first order partial differential equations of the type

$$
L I^k_i = K^k_i \quad \text{for} \quad 1 \leq i \leq n,
$$

derived in this section.
Section 5.3: We study the properties of the operator $L$ and its inverse.

Section 5.4: We discuss the implication of reversibility.

Section 5.5: We establish the existence of formal first integrals.

5.1 The Definition of Formal First Integrals

We start with the definition of first integrals.

Definition 5.1.1

A function

$$I : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times (U_{\Delta \sigma} \subset \mathbb{R}^n) \rightarrow \mathbb{R}, \quad (t, x, y, \Delta \sigma) \mapsto I(t, x, y, \Delta \sigma)$$

is called a first integral of the system of differential equations (5.1), if it is constant along solutions of (5.1):

$$\frac{d}{dt} I(t, x(t), y(t), \Delta \sigma) = 0 \text{ for every solution } t \mapsto (x(t), y(t)) \text{ of } (5.1). \quad (5.2)$$

First Integrals of Unperturbed Harmonic Oscillators

For a system of $n$ unperturbed harmonic oscillators

$$\dot{x}_i = \omega_i y_i, \quad \dot{y}_i = -\omega_i x_i \quad 1 \leq i \leq n,$$

we have $n$ first integrals

$$I_i^2(x, y) = \frac{1}{2}(x_i^2 + y_i^2) \quad 1 \leq i \leq n.$$

First Integrals of Perturbed Harmonic Oscillators

Keeping the above result in mind we show that (5.1) is formally integrable, i.e. it has $n$ formal first integrals of the form

$$I_i(t, x, y, \Delta \sigma) := \frac{1}{2}(x_i^2 + y_i^2) + \sum_{k=3}^{\infty} I_i^k(t, x, y, \Delta \sigma), \quad 1 \leq i \leq n, \quad (5.3)$$

where $I_i^k(t, x, y, \Delta \sigma)$ are polynomials of order $k$ in $x$, $y$ and $\Delta \sigma$ with $T$-periodic coefficients.
Definition 5.1.2
Let \( \mathcal{F} \) denote the space of smooth \( \frac{2\pi}{w_0} \)-periodic functions:

\[
\mathcal{F} := \left\{ f \in C^\infty(\mathbb{R}, \mathbb{C}) \mid f(t + \frac{2\pi}{w_0}) = f(t) \right\}
\]

and let

\[
\Pi_k := \left\{ f \mid f(t, x, y, \Delta \sigma) = \sum_{|p|_1 + |q|_1 + |s|_1 = k} f^{p,q,s}(t) x^p y^q \Delta \sigma^s \text{ with } f^{p,q,s} \in \mathcal{F} \right\}
\]

be the spaces of polynomials of order \( k \) in \( x, y \) and \( \Delta \sigma \) with \( T \)-periodic coefficients.

It is well known that the power series in (5.3) generally do not converge, even if the right-hand side of the differential equation (5.1) is analytic. Nevertheless they are very helpful: We use truncations of these formal power series as approximate first integrals.

5.2 The Formal Perturbation Scheme

We substitute the formal power series (5.3) into the definition (5.2) of a first integral.

The Perturbation Equations

The power expansion of \( h_{i,i}(x, y, \Delta \sigma) \) reads

\[
h_{i,i}(t, x, y, \Delta \sigma) := \sum_{k=2}^{\infty} h_{i,i}^k(t, x, y, \Delta \sigma),
\]

where the polynomials

\[
h_{i,i}^k(t, x, y, \Delta \sigma) := \sum_{|p|_1 + |q|_1 + |s|_1 = k} h_{i,i}^{p,q,s}(t) x^p y^q \Delta \sigma^s = \sum_{|p|_1 + |q|_1 + |s|_1 = k} \sum_{k_0 = -\infty}^{\infty} h_{i,i}^{p,q,s,k_0} e^{i k_0 \omega_0 t} x^p y^q \Delta \sigma^s
\]

are in \( \Pi_k \).

Lemma 5.2.1

Let the polynomials \( l_i^k \in \Pi_k \) be recursively defined by the so called perturbation equations

\[
Ll_i^k = k_i^k \quad \text{for} \quad 3 \leq k,
\]

where the operator \( L \) is given by
$L : \Pi_k \rightarrow \Pi_k, f \mapsto Lf := \sum_{j=1}^{n} \left( \frac{\partial f}{\partial x_j} y_j - \frac{\partial f}{\partial y_j} x_j \right) \omega_j + \frac{\partial f}{\partial t}$,

and the polynomials $K_k^i$ by

$$K_1^{3} := - x_i h_{1,i}^2 - y_i h_{2,i}^2,$$

$$K_k^k := - x_i h_{1,i}^{k-1} - y_i h_{2,i}^{k-1} - \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial I_i^l}{\partial x_j} h_{1,j}^{k-l+1} + \frac{\partial I_i^l}{\partial y_j} h_{2,j}^{k-l+1} \right).$$

Then

$$I_i(t, x, y, \Delta \sigma) := I_i^2(x, y) + \sum_{k=3}^{\infty} I_i^k(t, x, y, \Delta \sigma) \quad \text{with} \quad I_i^2(x, y) := \frac{1}{2} (x_i^2 + y_i^2)$$

is a formal first integral of (5.1).

**Proof:** We insert the expansion of $I_i(t, x, y, \Delta \sigma)$ into (5.2):

$$0 = \frac{d}{dt} \left( \frac{1}{2} (x_i^2 + y_i^2) + \sum_{k=3}^{\infty} I_i^k \right) = \left( x_i \dot{x}_i + y_i \dot{y}_i \right) + \sum_{k=3}^{\infty} \left( \sum_{j=1}^{n} \left( \frac{\partial I_i^k}{\partial x_j} \dot{x}_j + \frac{\partial I_i^k}{\partial y_j} \dot{y}_j + \frac{\partial I_i^k}{\partial t} \right) \right).$$

Substituting $\dot{x}_i$ and $\dot{y}_i$ by the right-hand sides of (5.1) we find

$$0 = x_i \sum_{k=2}^{\infty} h_{1,i}^k + y_i \sum_{k=2}^{\infty} h_{2,i}^k +$$

$$+ \sum_{k=3}^{\infty} \left( \sum_{j=1}^{n} \left( \frac{\partial I_i^k}{\partial x_j} \omega_j - \frac{\partial I_i^k}{\partial y_j} \omega_j x_j \right) \right) + \frac{\partial I_i^k}{\partial t} +$$

$$+ \sum_{k=3}^{\infty} \sum_{l=2}^{\infty} \sum_{j=1}^{n} \left( \frac{\partial I_i^k}{\partial x_j} h_{1,j}^{l-1} + \frac{\partial I_i^k}{\partial y_j} h_{2,j}^{l-1} \right) =$$

$$= \left( \sum_{j=1}^{n} \left( \frac{\partial I_i^3}{\partial x_j} y_j - \frac{\partial I_i^3}{\partial y_j} x_j \right) \omega_j + \frac{\partial I_i^3}{\partial t} + (x_i h_{1,i}^2 + y_i h_{2,i}^2) \right) +$$

$$+ \sum_{k=4}^{\infty} \sum_{j=1}^{n} \left( \frac{\partial I_i^k}{\partial x_j} y_j - \frac{\partial I_i^k}{\partial y_j} x_j \right) \omega_j + \frac{\partial I_i^k}{\partial t} +$$

$$+ (x_i h_{1,i}^{k-1} + y_i h_{2,i}^{k-1}) + \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial I_i^l}{\partial x_j} h_{1,j}^{k-l+1} + \frac{\partial I_i^l}{\partial y_j} h_{2,j}^{k-l+1} \right).$$
5.2. The Formal Perturbation Scheme

Using the definitions of $L$ and $K_i^k$ we obtain

$$\frac{d}{dt}I_i = \sum_{k=3}^{\infty} (L I_i^k - K_i^k) = 0.$$ 

This completes the proof. \halmos

In order to solve (5.4) for $I_i^k$ we must invert $L$. To achieve this goal we introduce new coordinates to obtain a more suitable form of $L$.

A Transformation of the Perturbation Equations

Consider the following change of coordinates:

$$\tau : \mathbb{C} \rightarrow \mathbb{C}, \quad (\xi, \eta) \mapsto (\xi, \eta) := \left(\frac{1}{\sqrt{2}}(\xi + i\eta), \frac{1}{\sqrt{2}}(\xi - i\eta)\right).$$

The inverse map is given by

$$\tau^{-1} : \mathbb{C} \rightarrow \mathbb{C}, \quad (x, y) \mapsto (x, y) := \left(\frac{1}{\sqrt{2}}(x - iy), -\frac{1}{\sqrt{2}}(x + iy)\right).$$

This transformation $\tau$ is well-known from the theory of Hamiltonian systems.

Since the solvability of the perturbation equations (5.4) is somewhat delicate it is decisive to distinguish strictly between polynomials in $x$-$y$-coordinates and polynomials in $\xi$-$\eta$-coordinates. Therefore we denote the space of polynomials in $\xi$-$\eta$-coordinates by $\Pi_k$ rather than by $\Pi_k^\sigma$.

**Definition 5.2.1**

Let $\Pi_k$ denote the space of polynomials of order $k$ in $\xi$, $\eta$ and $\Delta \sigma$:

$$\Pi_k := \left\{ \tilde{f} \left| \tilde{f}(t, \xi, \eta, \Delta \sigma) = \sum_{|p| + |q| + |r| = k} \tilde{f}^{p,q,r}(t)\xi^p \eta^q \Delta \sigma^r \text{ with } \tilde{f}^{p,q,r} \in \mathcal{F} \right. \right\}$$

and let $T$ be the homeomorphism between $\Pi_k$ and $\Pi_k$ induced by $\tau$:

$$T : \Pi_k \rightarrow \Pi_k, \; f \mapsto \tilde{f} := T f \quad \text{with} \quad \tilde{f}(t, \xi, \eta, \Delta \sigma) := f(t, \tau(\xi, \eta), \Delta \sigma)$$

and

$$T^{-1} : \Pi_k \rightarrow \Pi_k, \; \tilde{f} \mapsto f = T^{-1} \tilde{f} \quad \text{with} \quad f(t, x, y, \Delta \sigma) := \tilde{f}(t, \tau^{-1}(x, y), \Delta \sigma).$$
We apply the change of coordinates $\tau$ to the system of perturbed harmonic oscillators (5.1):

\[
\frac{1}{\sqrt{2}}(\xi_i + i\eta_i) = \omega_i \frac{i}{\sqrt{2}}(\xi_i - i\eta_i) + h_{1,i}(t, \tau(\xi, \eta), \Delta \sigma),
\]

\[
\frac{i}{\sqrt{2}}(\xi_i - i\eta_i) = -\omega_i \frac{1}{\sqrt{2}}(\xi_i + i\eta_i) + h_{2,i}(t, \tau(\xi, \eta), \Delta \sigma)
\]

and obtain the transformed system

\[
\begin{align*}
\dot{\xi}_i &= \omega_i \xi_i + \frac{1}{\sqrt{2}}(\bar{h}_{1,i}(t, \xi, \eta, \Delta \sigma) - i\bar{h}_{2,i}(t, \xi, \eta, \Delta \sigma)) \\
\dot{\eta}_i &= -\omega_i \eta_i - \frac{1}{\sqrt{2}}(\bar{h}_{1,i}(t, \xi, \eta, \Delta \sigma) + i\bar{h}_{2,i}(t, \xi, \eta, \Delta \sigma)) \\
\end{align*}
\]

(5.5)

**Lemma 5.2.2**

Let the polynomials $\bar{I}^k_1 \in \bar{\Pi}_k$ be recursively defined by the perturbation equations

\[
\bar{L}\bar{I}^k_1 = \bar{K}^k_1 \quad \text{for} \quad 3 \leq k,
\]

where the operator $\bar{L}$ and the polynomials $\bar{K}^k_1$ are given by

\[
\bar{L} : \quad \bar{\Pi}_k \rightarrow \bar{\Pi}_k, \quad \bar{f} \mapsto \bar{L}\bar{f} := i \sum_{j=1}^{n} \left( \frac{\partial \bar{f}}{\partial \xi_j} \xi_j - \frac{\partial \bar{f}}{\partial \eta_j} \eta_j \right) \omega_j + \frac{\partial \bar{f}}{\partial t},
\]

\[
\bar{K}^3_1 := -\frac{1}{\sqrt{2}} \left( (\xi_i + i\eta_i) \bar{h}_{1,i}^2 + i(\xi_i - i\eta_i) \bar{h}_{2,i}^2 \right),
\]

\[
\bar{K}^k_1 := -\frac{1}{\sqrt{2}} \left( (\xi_i + i\eta_i) \bar{h}_{1,i}^{k-1} + i(\xi_i - i\eta_i) \bar{h}_{2,i}^{k-1} \right) - \frac{1}{\sqrt{2}} \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \left( \frac{\partial \bar{I}^l_1}{\partial \xi_j} \xi_j - \frac{\partial \bar{I}^l_1}{\partial \eta_j} \eta_j \right) \bar{h}_{1,j}^{k-l-1} - i \left( \frac{\partial \bar{I}^l_1}{\partial \xi_j} \xi_j + \frac{\partial \bar{I}^l_1}{\partial \eta_j} \eta_j \right) \bar{h}_{2,j}^{k-l-1} \right).
\]

Then

\[
\bar{I}_1(t, \xi, \eta, \Delta \sigma) := \bar{I}^2_1(\xi, \eta) + \sum_{k=3}^{\infty} \bar{I}^k_1(t, x, y, \Delta \sigma) \quad \text{with} \quad \bar{I}^k_1(\xi, \eta) := i\xi_1\eta_1
\]

is a formal first integral of (5.5).

**Proof:** We insert the expansion of $I_1(t, \xi, \eta, \Delta \sigma)$ into (5.2):

\[
0 = \frac{d}{dt} \left( i\xi_1\eta_1 + \sum_{k=3}^{\infty} \bar{I}^k_1 \right) = i \left( \xi_1\eta_1 + \xi_1\eta_1 \right) + \sum_{k=3}^{\infty} \left( \sum_{j=1}^{n} \left( \frac{\partial \bar{I}^k_1}{\partial \xi_j} \xi_j + \frac{\partial \bar{I}^k_1}{\partial \eta_j} \eta_j \right) + \frac{\partial \bar{I}^k_1}{\partial t} \right).
\]
Substituting $\dot{\xi}_i$ and $\dot{\eta}_i$ by the right-hand sides of (5.5) we find
\[ 0 = \frac{i\eta}{2} \sum_{k=2}^{\infty} (\tilde{h}_{1,i}^k - i\tilde{h}_{2,i}^k) - i\xi_1 \frac{i}{2} \sum_{k=2}^{\infty} (\tilde{h}_{1,1}^k + i\tilde{h}_{2,1}^k) + \]
\[ + \sum_{k=3}^{\infty} \left( \sum_{j=1}^{n} \left( \frac{\partial \tilde{I}_1^k}{\partial \xi_j} \omega_j \xi_j - \frac{\partial \tilde{I}_1^k}{\partial \eta_j} \omega_j \eta_j \right) \omega_j + \frac{\partial \tilde{I}_1^k}{\partial t} \right) + \]
\[ + \sum_{k=3}^{\infty} \sum_{l=2}^{\infty} \sum_{j=1}^{n} \left( \frac{\partial \tilde{I}_1^k}{\partial \xi_j} \frac{1}{\sqrt{2}} (\tilde{h}_{1,j}^l - i\tilde{h}_{2,j}^l) - \frac{\partial \tilde{I}_1^k}{\partial \eta_j} \frac{i}{\sqrt{2}} (\tilde{h}_{1,j}^l + i\tilde{h}_{2,j}^l) \right) = \]
\[ = \left( \sum_{j=1}^{n} \left( \frac{\partial \tilde{I}_1^3}{\partial \xi_j} - \frac{\partial \tilde{I}_1^3}{\partial \eta_j} \right) \omega_j + \frac{\partial \tilde{I}_1^3}{\partial t} + \left( \frac{1}{\sqrt{2}} (\xi_i + i\eta_i) \tilde{h}_{1,i}^2 + i \frac{1}{\sqrt{2}} (\xi_i - i\eta_i) \tilde{h}_{2,i}^2 \right) \right) + \]
\[ + \sum_{k=4}^{\infty} \left( \sum_{j=1}^{n} \left( \frac{\partial \tilde{I}_1^k}{\partial \xi_j} - \frac{\partial \tilde{I}_1^k}{\partial \eta_j} \right) \omega_j + \frac{\partial \tilde{I}_1^k}{\partial t} + \right. \]
\[ + \left( \frac{1}{\sqrt{2}} (\xi_i + i\eta_i) \tilde{h}_{1,i}^{k-1} + i \frac{1}{\sqrt{2}} (\xi_i - i\eta_i) \tilde{h}_{2,i}^{k-1} \right) + \]
\[ + \frac{1}{\sqrt{2}} \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial \tilde{I}_1^l}{\partial \xi_j} - i \frac{\partial \tilde{I}_1^l}{\partial \eta_j} \right) \tilde{h}_{1,j}^{k-1} - i \left( \frac{\partial \tilde{I}_1^l}{\partial \xi_j} + i \frac{\partial \tilde{I}_1^l}{\partial \eta_j} \right) \tilde{h}_{2,j}^{k-1} \right). \]

Using the definitions of $\tilde{L}$ and $\tilde{K}_i^k$ we obtain
\[ 0 = \frac{d}{dt} \tilde{I}_i = \sum_{k=3}^{\infty} (\tilde{L} \tilde{I}_i^k - \tilde{K}_i^k). \]

This completes the proof. \qed

**Lemma 5.2.3**
Consider the operators $L$ and $\tilde{L}$ defined in (5.2.1) and (5.2.2), respectively. Then the following holds:

(i) The following diagram commutes:
\[ \begin{array}{ccc}
\Pi_k & \xrightarrow{L} & \Pi_k \\
\, \downarrow T & & \downarrow T \\
\Pi_k & \xrightarrow{\tilde{L}} & \Pi_k \\
\end{array} \]

(ii) If the polynomials $I_i^k$ are solutions of (5.4) then the transformed polynomials $\tilde{I}_i^k := T \tilde{I}_i^k$ are solutions of (5.6) with $\tilde{K}_i^k = TK_i^k$. 


(iii) If the polynomials \( \tilde{I}^k_i \) are solutions of (5.6) then the polynomials \( I^k_i := T^{-1}\tilde{I}^k_i \) are solutions of (5.4) with \( K^k_i = T^{-1}\tilde{K}^k_i \).

**Proof:**

(i) Let \( f \in \Pi_k \). By the chain rule we obtain

\[
T(Lf) = T \left( \sum_{j=1}^{n} \left( \frac{\partial f}{\partial x_j} y_j - \frac{\partial f}{\partial y_j} x_j \right) \omega_j + \frac{\partial f}{\partial t} \right) =
\]

\[
= \sum_{j=1}^{n} \left( \frac{\partial \tilde{f}}{\partial \xi_j} \frac{1}{\sqrt{2}} + \frac{\partial \tilde{f}}{\partial \eta_j} \frac{i}{\sqrt{2}} \right) \frac{1}{\sqrt{2}} (\xi_j - i\eta_j) - \left( \frac{\partial \tilde{f}}{\partial \xi_j} \frac{1}{\sqrt{2}} + \frac{\partial \tilde{f}}{\partial \eta_j} \frac{1}{\sqrt{2}} \right) \frac{1}{\sqrt{2}} (\xi_j + i\eta_j) \omega_j + \frac{\partial \tilde{f}}{\partial t} =
\]

\[
= i \sum_{j=1}^{n} \left( \frac{\partial \tilde{f}}{\partial \xi_j} \xi_j - \frac{\partial \tilde{f}}{\partial \eta_j} \eta_j \right) \omega_j + \frac{\partial \tilde{f}}{\partial t} = \tilde{L}\tilde{f}.
\]

Thus we have

\[
TL = \tilde{L}T.
\]

\( L(\Pi_k) \subset \Pi_k \) and \( \tilde{L}(\Pi_k) \subset \Pi_k \) are obvious.

(ii) We prove this part by induction.

For \( k = 3 \) we obtain

\[
TK^3_i = T \left( x_i h^2_{1,i} + y_i h^2_{2,i} \right) = \left( \frac{1}{\sqrt{2}} (\xi_i + i\eta_i) \tilde{h}^2_{1,i} + \frac{i}{\sqrt{2}} (\xi_i - i\eta_i) \tilde{h}^2_{2,i} \right) = \tilde{K}^3_i
\]

and by the definition of \( \tilde{I}^3_i \) and (i)

\[
\tilde{L}I^3_i = \tilde{L}I^3_i = TL I^3_i = TK^3_i = \tilde{K}^3_i.
\]

Now assume that the polynomials \( \tilde{I}^l_i := TI^l_i \) are solutions of (5.6) for \( 3 \leq l \leq k-1 \). We obtain

\[
TK^k_i = T \left( - (x_i h^{k-1}_{1,i} + y_i h^{k-1}_{2,i}) - \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial I^l_i}{\partial x_j} h^{k-l+1}_{1,j} + \frac{\partial I^l_i}{\partial y_j} h^{k-l+1}_{2,j} \right) \right) =
\]

\[
= - \left( \frac{1}{\sqrt{2}} (\xi_i + i\eta_i) \tilde{h}^{k-1}_{1,i} + \frac{i}{\sqrt{2}} (\xi_i - i\eta_i) \tilde{h}^{k-1}_{2,i} \right) -
\]

\[
- \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial I^l_i}{\partial \xi_j} \frac{1}{\sqrt{2}} + \frac{\partial I^l_i}{\partial \eta_j} \frac{i}{\sqrt{2}} \right) \tilde{h}^{k-l+1}_{1,j} + \left( \frac{\partial I^l_i}{\partial \xi_j} \frac{1}{\sqrt{2}} + \frac{\partial I^l_i}{\partial \eta_j} \frac{i}{\sqrt{2}} \right) \tilde{h}^{k-l+1}_{2,j} =
\]
5.3. Inverting the Operator $L$

In this section we show that the equation $Lf = g$ is solvable under certain assumptions on $g$.

**Conditions for the Solvability of $Lf = g$**

We consider the equation $Lf = g$ for monomials $f, g$ in $\Pi_k$ with $k \leq N$ (for the definition of $N$ cf. Assumption $A_{5b}$):

$$f(t, \xi, \eta, \Delta \sigma) := \tilde{f}^{p,q,s}(t)\xi^p\eta^q\Delta \sigma^s$$

$$g(t, \xi, \eta, \Delta \sigma) := \tilde{g}^{p,q,s}(t)\xi^p\eta^q\Delta \sigma^s.$$  

From

$$\tilde{L}f(t, \xi, \eta, \Delta \sigma) = i(p-q)\omega \tilde{f}^{p,q,s}(t)\xi^p\eta^q\Delta \sigma^s + \tilde{f}^{p,q,s}(t)\xi^p\eta^q\Delta \sigma^s$$

it follows that the function $\tilde{f}^{p,q,s}(t)$ must be a $\frac{2\pi}{\omega_0}$-periodic solution of the differential equation

$$\tilde{f}^{p,q,s} = -i(p-q)\omega \tilde{f}^{p,q,s} + \tilde{g}^{p,q,s}.$$  

(5.7)

The existence of a solution is guaranteed by the next lemma.

**Lemma 5.3.1**

Consider Eq. (5.7) for $\tilde{g}^{p,q,s} \in F$, where $|p|_1 + |q|_1 + |s|_1 = k$ for some fixed $k \leq N$, $N$ defined in Assumption $A_{5b}$, and either $p \neq q$ or $p = q$ and $\tilde{g}^{p,p,s,0} = 0$. Then the following assertions hold:

(i) There exists a solution $\tilde{f}^{p,q,s} \in F$.

(ii) Let

$$\tilde{g}^{p,q,s}(t) := \sum_{k_0 = -\infty}^{\infty} \tilde{g}^{p,q,s,k_0} e^{ik_0\omega_0 t}$$

\[= -\frac{1}{\sqrt{2}} \left( (\xi_1 + i\eta_1)\tilde{h}_{1,i}^{k-1} + i(\xi_1 - i\eta_1)\tilde{h}_{2,i}^{k-1} \right) -

- \frac{1}{\sqrt{2}} \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \left( \frac{\partial \tilde{t}_i}{\partial \xi_j} - i \frac{\partial \tilde{t}_i}{\partial \eta_j} \right) \tilde{h}_{1,j}^{k-l+1} - i \left( \frac{\partial \tilde{t}_i}{\partial \xi_j} + i \frac{\partial \tilde{t}_i}{\partial \eta_j} \right) \tilde{h}_{2,j}^{k-l+1} \right) = K_i^k

and by the definition of $\tilde{T}_i^k$ and (i)

$$\tilde{L}I_i^k = \tilde{L}T I_i^k = T \tilde{L}I_i^k = T K_i^k = \tilde{K}_i.$$  

Thus we may conclude that the polynomials $\tilde{I}_i^k := T \tilde{I}_i^k$ solve (5.6).

(iii) The proof of this part is similar to the proof of part (ii). It is therefore omitted. □
be the Fourier expansion of \( \tilde{g}^{p,q,s} \), then the Fourier expansion of \( \tilde{f}^{p,q,s} \) is given by

\[
\tilde{f}^{p,q,s}(t) = \sum_{k_0=-\infty}^{\infty} \frac{\tilde{g}^{p,q,s,k_0}}{i ((p-q)\omega + k_0\omega_0)} e^{i(k_0\omega_0)t}.
\]

**Proof:** We first treat the case \( p \neq q \).

Using the Variation-of-Constant-Formula we obtain

\[
\tilde{f}^{p,q,s}(t) = e^{-i(p-q)\omega t} \tilde{f}^{p,q,s}(0) + \int_0^t e^{-i(p-q)\omega (t-\tau)} \tilde{g}^{p,q,s}(\tau) d\tau.
\]

Obviously \( \tilde{f}^{p,q,s} \) has the same degree of regularity as \( \tilde{g}^{p,q,s} \). Substituting the Fourier expansion of \( \tilde{g}^{p,q,s} \) into the above formula we obtain

\[
\tilde{f}^{p,q,s}(t) = e^{-i(p-q)\omega t} \tilde{f}^{p,q,s}(0) + e^{-i(p-q)\omega t} \int_0^t e^{i(p-q)\omega \tau} \sum_{k_0=-\infty}^{\infty} \tilde{g}^{p,q,s,k_0} e^{i k_0 \omega_0 \tau} d\tau.
\]

Since \( \tilde{g}^{p,q,s} \) is in \( C^\infty \) (cf. the definition of \( \mathcal{F} \)) the Fourier expansion is absolutely convergent. Thus we may exchange the sum and the integral:

\[
\tilde{f}^{p,q,s}(t) = e^{-i(p-q)\omega t} \tilde{f}^{p,q,s}(0) + e^{-i(p-q)\omega t} \sum_{k_0=-\infty}^{\infty} \tilde{g}^{p,q,s,k_0} \int_0^t e^{i(p-q)\omega \tau + i k_0 \omega_0 \tau} d\tau =
\]

\[
eq e^{-i(p-q)\omega t} \tilde{f}^{p,q,s}(0) + e^{-i(p-q)\omega t} \sum_{k_0=-\infty}^{\infty} \tilde{g}^{p,q,s,k_0} \left( \frac{e^{i((p-q)\omega + k_0 \omega_0) t} - 1}{i ((p-q)\omega + k_0 \omega_0)} \right).
\]

In the last step we used that the denominators do not vanish due to assumptions of this lemma. Moreover since \( k \leq N \) is fixed the denominators may be estimated from below by some positive numbers \( \alpha_k \) introduced in Lemma 3.2.3:

\[
\left| (p-q)\omega + k_0 \omega_0 \right| \geq \alpha_k \quad \text{for} \quad |p|_1 + |q|_1 + |s|_1 \leq k, k_0 \in \mathbb{Z}.
\]

Thus the series

\[
\sum_{k_0=-\infty}^{\infty} \frac{\tilde{g}^{p,q,s,k_0}}{i ((p-q)\omega + k_0 \omega_0)}
\]

and

\[
\sum_{k_0=-\infty}^{\infty} \frac{\tilde{g}^{p,q,s,k_0}}{i ((p-q)\omega + k_0 \omega_0)} e^{i k_0 \omega_0 t}
\]
are absolutely convergent and we find
\[
\tilde{f}^{p,q,s}(t) = e^{-i(p-q)\omega t} \left( \sum_{k_0}^{\infty} \frac{\tilde{g}^{p,q,s,k_0}}{i((p-q)\omega + k_0\omega_0)} + \sum_{k_0}^{\infty} \frac{\tilde{g}^{p,q,s,k_0}}{i((p-q)\omega + k_0\omega_0)} e^{i(k_0\omega_0)t} \right).
\]
Thus \(\tilde{f}^{p,q,s}\) is \(2\pi\omega_0\)-periodic if and only if
\[
\tilde{f}^{p,q,s}(0) = \sum_{k_0}^{\infty} \frac{\tilde{g}^{p,q,s,k_0}}{i((p-q)\omega + k_0\omega_0)}.
\]
It follows that a periodic solution exists and that it admits the following Fourier expansion:
\[
\tilde{f}^{p,q,s}(t) = \sum_{k_0}^{\infty} \frac{\tilde{g}^{p,q,s,k_0}}{i((p-q)\omega + k_0\omega_0)} e^{i(k_0\omega_0)t}.
\]
This completes the first part of the proof. We now treat the case \(p = q\).
For \(p = q\) the differential equation (5.7) is simplified to
\[
\tilde{f}^{p,p,s} = \tilde{g}^{p,p,s},
\]
Thus we have
\[
\tilde{f}^{p,p,s}(t) = \int_{0}^{t} \tilde{g}^{p,p,s}(\tau) d\tau.
\]
Obviously \(\tilde{f}^{p,p,s}\) has the same degree of regularity as \(\tilde{g}^{p,p,s}\) and it is \(2\pi\omega_0\)-periodic if and only if \(\tilde{g}^{p,p,s,0} = 0\). This completes the second part of the proof.

**Complex Solutions of \(Lf = g\)**

Lemma 5.3.1 suggests the following decomposition of \(\Pi_k\) into two subspaces.

**Lemma 5.3.2**

Let \(\Pi_k^{nc}\) and \(\Pi_k^{c}\) denote the spaces of the so-called non-critical and critical polynomials, respectively:

\[
\Pi_k^{nc} := \left\{ \tilde{f} \in \Pi_k \mid \tilde{f}(t, \xi, \eta, \Delta \sigma) = \sum_{|t| + |\xi| + |\eta| = k} \sum_{k_0 = -\infty}^{\infty} \tilde{f}^{p,q,s,k_0} e^{ik_0\omega_0 t} \xi_p \eta^q \Delta \sigma^s \right\},
\]

with \(\tilde{f}^{p,p,s,0} = 0\), for all \(p\),

\[
\Pi_k^{c} := \left\{ \tilde{f} \in \Pi_k \mid \tilde{f}(t, \xi, \eta, \Delta \sigma) = \sum_{2|\xi| + |\eta| = k} \tilde{f}^{p,p,s,0} \xi_p \eta^q \Delta \sigma^s \right\}.
\]
Then for \( k \leq N \) (cf. Assumption A5b) the following holds:

(i) \( \Pi_k = \Pi_k^{nc} \oplus \Pi_k^{c} \).
(ii) \( \ker(\bar{L}) = \Pi_k^{c} \).
(iii) \( \bar{L}(\Pi_k) = \Pi_k^{nc} \).
(iv) \( \bar{L} : \Pi_k^{nc} \to \Pi_k^{nc} \) is bijective. Hence its inverse exists is given by

\[
\bar{L}^{-1} : \Pi_k^{nc} 
\to \Pi_k^{nc} , \, \tilde{g} \mapsto \bar{f}, \text{ where } \bar{L}\bar{f} = \tilde{g}.
\]

(v) Let \( \bar{f} := \sum_{|p|+|q|+|s_1|+|k_0|=k} \bar{g}_{p,q,s,k_0} e^{ik_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s \in \Pi_k^{nc} \),

then

\[
\bar{L}^{-1} \bar{g} = \sum_{|p|+|q|+|s_1|+|k_0|=k} \sum_{k_0=-\infty}^{\infty} \frac{\bar{g}_{p,q,s,k_0}}{i ((p-q)\omega + k_0 \omega_0)} e^{ik_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s.
\]

Proof:

(i) \( \Pi_k = \Pi_k^{nc} \oplus \Pi_k^{c} \) follows from the definition of \( \Pi_k^{nc} \) and \( \Pi_k^{c} \).

(ii) Let \( \bar{f} := \sum_{|p|+|q|+|s_1|+|k_0|=k} \bar{g}_{p,q,s,k_0} e^{ik_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s \) be a monomial in \( \Pi_k \) with \( \bar{g}_{p,q,s,k_0} \neq 0 \). Then we have

\[
\bar{L}\bar{f} = i ((p-q)\omega + k_0 \omega_0) \bar{f}_{p,q,s,k_0} e^{ik_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s.
\]

On the one hand assume that \( \bar{f} \in \Pi_k^{c} \). It follows from the definition of \( \Pi_k^{c} \) that \( p = q \) and \( k_0 = 0 \). Therefore we have \( \bar{L}\bar{f} = 0 \), i.e. \( \bar{f} \in \ker(\bar{L}) \).

On the other hand assume that \( \bar{f} \in \ker(\bar{L}) \). It follows that \( i ((p-q)\omega + k_0 \omega_0) = 0 \). The non-resonance condition implies that \( p = q \) and \( k_0 = 0 \). Thus we \( \bar{f} \in \Pi_k^{c} \).

(iii) We first show that \( \bar{L}(\Pi_k^{nc}) = \bar{L}(\Pi_k) \). Obviously we have \( \bar{L}(\Pi_k^{nc}) \subset \bar{L}(\Pi_k) \). Thus assume that \( \bar{f} \in \Pi_k^{nc} \). According to (i) we may write \( \bar{f} = \bar{f}_{nc} + \bar{f}_{c} \), where \( \bar{f}_{nc} \in \Pi_k^{nc} \) and \( \bar{f}_{c} \in \Pi_k^{c} \). Using (ii) we find

\[
\bar{L}(\bar{f}) = \bar{L}(\bar{f}_{nc} + \bar{f}_{c}) = \bar{L}\bar{f}_{nc} + \bar{L}\bar{f}_{c} = \bar{L}\bar{f}_{nc}
\]

and therefore \( \bar{L}(\Pi_k) \subset \bar{L}(\Pi_k^{nc}) \).

Let \( \tilde{g} := \tilde{g}_{p,q,s,k_0} e^{ki_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s \) be a monomial in \( \Pi_k \). On the one hand assume that \( \tilde{g} \in \Pi_k^{nc} \). It follows from the non-resonance condition that \( i ((p-q)\omega + k_0 \omega_0) \neq 0 \).

This allows us to define a monomial \( \bar{f} := f_{p,q,s,k_0} e^{ki_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s \) by

\[
\bar{f}_{p,q,s,k_0} := \frac{\tilde{g}_{p,q,s,k_0}}{i ((p-q)\omega + k_0 \omega_0)}.
\]

Obvious we have \( \bar{f} \in \Pi_k^{nc} \) and \( \bar{L}\bar{f} = \tilde{g} \). It follows that and \( \tilde{g} \in \bar{L}(\Pi_k^{nc}) \).

On the other hand assume that \( \tilde{g} \in \bar{L}(\Pi_k^{nc}) \). It follows that there exists a monomial
5.3. Inverting the Operator $L$

Let $f$ with $Lf = g$. This implies that

$$i((p-q)\omega + k_0\omega_0)\tilde{f}^p,q,\omega,k_0 e^{i\omega_0k_0t}\zeta^p\eta^q \Delta \sigma = \tilde{g}^p,q,\omega,k_0 e^{i\omega_0k_0t}\zeta^p\eta^q \Delta \sigma.$$

It follows that $i((p-q)\omega + k_0\omega_0) \neq 0$. By the non-resonance condition we conclude that either $p \neq q$ or $p = q$ and $k_0 \neq 0$. Thus we have $\tilde{g} \in \Pi_k^{nc}$.

(iv) From (ii) and (iii) it follows that $\bar{L}$ restricted to $\Pi_k^{nc}$ is surjective. Thus it remains to show that $\bar{L}$ is also injective. To this end let $\bar{f}$ and $\bar{f}'$ be two functions in $\Pi_k^{nc}$ with $\bar{L}\bar{f} = \bar{L}\bar{f}'$. From the linearity of $\bar{L}$ it follows that

$$0 = \bar{L}\bar{f} - \bar{L}\bar{f}' = \bar{L}(\bar{f} - \bar{f}).$$

Thus $\bar{f} - \bar{f}' \in \ker \bar{L}$. From (ii) it follows further that $\bar{f} - \bar{f}' \in \Pi_k^{nc}$. Now (i) implies that $\bar{f} - \bar{f}' = 0$. Thus $\bar{L}$ is injective.

(v) The claim follows immediately from the definition of the operator $\bar{L}$ and from Lemma 5.3.1.

Using the homeomorphism $T$ we may define non-critical and critical polynomials for the original $x$-$y$-coordinates.

Lemma 5.3.3

For $k \leq N$ (cf. Assumption $A_{5b}$) let

$$\Pi_k^{nc} := \{ f \in \Pi_k \mid (Tf)(t,\xi,\eta,\Delta \sigma) \in \Pi_k^{nc} \} ,$$

$$\Pi_k^c := \{ f \in \Pi_k \mid (Tf)(t,\xi,\eta,\Delta \sigma) \in \Pi_k^{c} \} ,$$

then the following holds:

(i) $\Pi_k = \Pi_k^{nc} \oplus \Pi_k^c$.

(ii) $\ker(L) = \Pi_k^c$.

(iii) $L(\Pi_k) = \Pi_k^{nc}$.

(iv) $L : \Pi_k^{nc} \longrightarrow \Pi_k^{nc}$ is bijective and hence its inverse exists:

$$L^{-1} : \Pi_k^{nc} \longrightarrow \Pi_k^{nc} : g \longrightarrow f, \text{ where } f = T^{-1}(L^{-1}(Tg)).$$

Proof:

(i) We first show that $\Pi_k \subset \Pi_k^{nc} \oplus \Pi_k^c$. Let $f \in \Pi_k$. Then $\bar{f} := Tf$ lies in $\Pi_k$. Lemma 5.3.2 implies that there exist $\bar{f}^{nc} \in \Pi_k^{nc}$ and $\bar{f}^c \in \Pi_k^c$ such that $\bar{f} = \bar{f}^{nc} + \bar{f}^c$. Therefore we have $f = f^{nc} + f^c$ with $f^{nc} := T^{-1}\bar{f}^{nc} \in \Pi_k^{nc}$ and $f^c := T^{-1}\bar{f}^c \in \Pi_k^c$. We conclude that $f$ lies in $\Pi_k^{nc} \oplus \Pi_k^c$.

Now we show that $\Pi_k^{nc} + \Pi_k^c \subset \Pi_k$. Since $\Pi_k^{nc}$ and $\Pi_k^c$ are subspaces of $\Pi_k$, it follows
immediately that $\Pi_k^c \oplus \Pi_k^c \subset \Pi_k$.

It remains to show that $\Pi_k^nc \cap \Pi_k^c = \{0\}$. Let $f \in \Pi_k^nc \cap \Pi_k^c$. It follows that $\tilde{f} := T \tilde{f}$ lies in $\Pi_k^nc \cap \Pi_k^c$. Lemma 5.3.2 implies that $\tilde{f} = 0$. Therefore we have $f = T^{-1} \tilde{f} = T^{-1}0 = 0$. This completes the proof of the first part of this lemma.

(iv) We first show that the operator $L$ is injective in $\Pi_k^nc$. Let $f, f' \in \Pi_k^nc$ with $Lf = Lf'$. Since $T$ is a homeomorphism we have $\tilde{L}f = \tilde{L}f'$ with $\tilde{f} := T \tilde{f} \in \Pi_k^nc$, $\tilde{f}' := T \tilde{f}' \in \Pi_k^nc$. Since the operator $\tilde{L}$ is injective on $\Pi_k^nc$ we find $\tilde{f} = \tilde{f}'$ and therefore $f = f'$. Thus $L$ is injective in $\Pi_k^nc$. Now we show that $L$ is surjective in $\Pi_k^nc$. To this end let $g \in \Pi_k^nc$. It follows from (iii) that there exists a polynomial $f \in \Pi_k$ such that $Lf = g$. Since $f = f^nc + f^c$ (cf. (i)) we find $g = Lf = Lf^nc + Lf^c = Lf'^nc$ and we conclude that $g \in L(\Pi_k^nc).$

Finally let $g \in \Pi_k^nc$. There exists a polynomial $f \in \Pi_k^nc$ such that $Lf = g$. Since $T$ is a homeomorphism we have $\tilde{L}f = \tilde{T}g$ with $T \tilde{g} \in \Pi_k^nc$. Lemma 5.2.3 implies that $\tilde{L}Tf = \tilde{T}g$ with $T \tilde{f}, T \tilde{g} \in \Pi_k^nc$. Due to Lemma 5.3.2 we may invert the operator $\tilde{L}$ and obtain $T \tilde{f} = \tilde{T}^-1T \tilde{g}$ with $T \tilde{f}, T \tilde{g} \in \Pi_k$. Thus we have $f = T^-1(\tilde{T}^-1(T \tilde{g}))$.

\[ \square \]

**Real Solutions of $Lf = g$**

At the moment we know that for any $g \in \Pi_k^nc$ there exists a complex polynomial $f$ such that $Lf = g$. It remains to show that for a real polynomial $g$ the solution $f$ is real too. To this end we decompose $\Pi_k$ and $\Pi_k^c$.

**Lemma 5.3.4**

For $k \leq N$ (cf. Assumption $A_{5b}$) let

\[ \Pi_{R,k} := \{ f \in \Pi_k \mid f^{p,q,s}(t) \in \mathbb{R} \}, \quad \Pi_{iR,k} := \{ f \in \Pi_k \mid f^{p,q,s}(t) \in i\mathbb{R} \} \]

and

\[ f_R \in \Pi_{R,k} := T(\Pi_{R,k}) \quad f_{iR} \in \Pi_{iR,k} := T(\Pi_{iR,k}) \]
Finally let
\[ L_R := L|_{\Pi_R,k} \quad \text{and} \quad \bar{L}_R := L|_{\bar{\Pi}_R,k} \]

Then the following holds:
(i) \( \Pi_k = \Pi_R,k \oplus \Pi_I,k \) with \( f = f_R + f_I \), where \( f_R \in \Pi_R,k \), \( f_I \in \Pi_I,k \).
(ii) \( \bar{\Pi}_k = \bar{\Pi}_R,k \oplus \bar{\Pi}_I,k \) with \( \bar{f} = \bar{f}_R + \bar{f}_I \), where \( \bar{f}_R \in \bar{\Pi}_R,k \), \( \bar{f}_I \in \bar{\Pi}_I,k \).
(iii) The following diagram commutes:

\[
\begin{array}{ccc}
\Pi_R,k & \xrightarrow{L_R} & \Pi_R,k \\
\downarrow T & & \downarrow T \\
\bar{\Pi}_R,k & \xrightarrow{L_R} & \bar{\Pi}_R,k
\end{array}
\]

Proof: The commutation property \( TL_R = \bar{L}_R T \) follows from Lemma 5.2.3. Obviously we have \( L_R(\Pi_R,k) \subset \Pi_R,k \) and therefore \( L_R(\bar{\Pi}_R,k) = \bar{L}_R T(\Pi_R,k) = TL_R(\Pi_R,k) \subset T(\Pi_I,k) = \bar{\Pi}_R,k \).

Lemma 5.3.5
For \( k \leq N \) (cf. Assumption A5b) let
\[
\begin{align*}
\Pi_R,n^c, & := \Pi_R,k \cap \Pi_k^c, \\
\Pi_I,n^c, & := \Pi_I,k \cap \Pi_k^c, \\
\Pi_R,k \cap \Pi_k^c, & := \Pi_R,k \cap \Pi_k^c \\
\Pi_I,k \cap \Pi_k^c, & := \Pi_I,k \cap \Pi_k^c 
\end{align*}
\]
then the following holds:
(i) \( \Pi_R,k = \Pi_R,n^c, \oplus \Pi_I,n^c, \).
(ii) \( \ker(\bar{L}_R) = \Pi_I,n^c, \).
(iii) \( \bar{L}_R(\Pi_R,k,n^c) = \Pi_R,n^c, \).
(iv) \( \bar{L}_R : \Pi_{R,n}^c, \longrightarrow \Pi_{R,k,n}^c, \) is bijective. Thus the inverse of \( \bar{L}_R \) on \( \Pi_{R,n}^c, \) is uniquely defined by:
\[
\bar{L}_R^{-1} : \Pi_R,n^c, \longrightarrow \Pi_{R,k,n}^c, \quad \bar{f} \mapsto f, \quad \text{where} \quad \bar{L}_R f = \bar{g}.
\]

Proof:
(i) We first prove that the intersection contains only the element 0.
\[
\Pi_{R,k} \cap \Pi_{k}^c = (\Pi_{R,k} \cap \Pi_k^c) \cap (\Pi_{R,k} \cap \Pi_k^c) = \\
= \Pi_{R,k} \cap (\Pi_k^c \cap \Pi_k^c) = \Pi_{R,k} \cap \{0\} = \{0\}.
\]
Now let \( f \in \Pi_{R,k} \). Since \( f \) also lies in \( \Pi_k \) we may decompose it into \( f = f^nc + f^c \), where \( f^nc \in \Pi^nc_k \) and \( f^c \in \Pi^c_k \). Using Lemma 5.3.4 (ii) we may further write \( f^nc = f^nc_R + f^nc_k \) and \( f^c = f^c_R + f^c_k \) and we find
\[
\tilde{f} = f^nc_R + f^c_R = (f^nc_R + f^nc_k) + (f^c_R + f^c_k) = (\tilde{f}^nc_R + \tilde{f}^c_R) + (\tilde{f}^nc_k + \tilde{f}^c_k).
\]
Since the left-hand side lies in \( \Pi_R \) we conclude with Lemma 5.3.4 (iii) that \( \tilde{f}^nc_R + \tilde{f}^c_R = 0 \). Moreover since the decomposition of 0 is unique we find \( \tilde{f}^nc_R = 0 \) and \( \tilde{f}^c_R = 0 \). This implies \( \tilde{f}^nc_k \in \Pi^nc_{R,k} \) and therefore \( \Pi^nc_{R,k} \subset \Pi^nc_{R,k} \). Finally let \( \tilde{f} \in \Pi^nc_{R,K} \). Since \( \Pi^nc_{R,k} \subset \Pi^nc_{R,k} \) and \( \Pi^c_{R,k} \subset \Pi^c_{R,k} \) we have \( \tilde{f} \in \Pi^nc_{R,k} \) and therefore \( \Pi^nc_{R,k} \subset \Pi^nc_{R,k} \). This completes the proof of part (i).

(ii) \( \tilde{f} \in \ker(L_R) \iff \tilde{f} \in \Pi^nc_{R,k} \). \( \tilde{f} \in \ker(L_R) \iff \tilde{f} \in \Pi^nc_{R,k} \). \( \tilde{f} \in \ker(L_R) \iff \tilde{f} \in \Pi^nc_{R,k} \).

(iii) In a first step we show that \( L_R(\Pi^nc_{R,k}) \subset \Pi^nc_{R,k} \). On the one hand it follows from Lemma 5.3.3 that
\[
L_R(\Pi^nc_{R,k}) \subset L(\Pi_k) = \Pi^nc_k.
\]
On the other hand it follows from Lemma 5.2.3 that
\[
L_R(\Pi^nc_{R,k}) \subset L(\Pi_{R,k}) = \Pi_{R,k} \subset \Pi^nc_{R,k}.
\]
Combining these findings we have \( L_R(\Pi^nc_{R,k}) \subset \Pi^nc_{R,k} \). It remains to show that even the equality sign holds.

Let \( \tilde{g} \in \Pi^nc_{R,k} \). From Lemma 5.3.2 it follows that there exists a function \( \tilde{f} \in \Pi^nc_k \) with \( L\tilde{f} = \tilde{g} \). We show that one can even choose \( \tilde{f} \in \Pi^nc_{R,k} \). To this end we decompose \( \tilde{f} \) into \( \tilde{f} = \tilde{f}_R + \tilde{f}_R \) (cf. Lemma 5.3.4 (ii)). The homeomorphism \( T \) induces polynomials \( f \in \Pi_k \), \( f_R \in \Pi^nc_{R,k} \) and \( f_R \in \Pi^c_{R,k} \) such that \( f = f_R + f_R \). It follows that
\[
\tilde{g} = L\tilde{f} = L(f_R + f_R) = \tilde{f}_R + \tilde{f}_R = TLTf_R + TLTf_R = TLf_R + TLf_R.
\]
On the one hand the definition of \( L \) implies that \( TLf_R \in \Pi^nc_{R,k} \) but on the other hand we have \( \tilde{g} \in \Pi^nc_{R,k} \). It follows that \( TLf_R = 0 \) and therefore \( \tilde{f}_R = 0 \). This implies that \( L_R(\Pi^nc_{R,k}) = \Pi^nc_{R,k} \).

(iv) The proof is analogous to the Proof of Lemma 5.3.2 (iv).

Lemma 5.3.6

For \( k \leq N \) (cf. Assumption A_{5b}) let
\[
\Pi^nc_{R,k} := \Pi_{R,k} \cap \Pi^nc_k, \\
\Pi^c_{R,k} := \Pi_{R,k} \cap \Pi^c_k.
\]
Then the following holds:

Using the homeomorphism \( T \) we may prove an analogous result for \( \Pi_k \).
5.4. The Symmetry of the Perturbation Equations

(i) \( \Pi_{R,k} = \Pi_{R,k}^{nc} \oplus \Pi_{R,k}^c \).

(ii) \( \ker(L_R) = \Pi_{R,k}^c \).

(iii) \( L_R(\Pi_{R,k}) = \Pi_{R,k}^{nc} \).

(iv) \( L_R : \Pi_{R,k}^{nc} \rightarrow \Pi_{R,k}^{nc} \) is bijective and its inverse exists and is unique:

\[
L_R^{-1} : \Pi_{R,k}^{nc} \rightarrow \Pi_{R,k}^{nc}, \quad g \mapsto f, \quad \text{where } f = T^{-1}(\tilde{L}_R^{-1}(Tg)).
\]

Proof:

(i) We first show that \( \Pi_{R,k}^{nc} \cap \Pi_{R,k}^c = \{0\} \). Let \( f \in \Pi_{R,k}^{nc} \cap \Pi_{R,k}^c \). It follows that \( \tilde{f} := Tf \) lies in \( \Pi_{R,k}^{nc} \cap \Pi_{R,k}^c \). Lemma 5.3.5 implies that \( \tilde{f} = 0 \). Therefore we have \( f = T^{-1} \tilde{f} = T^{-1}0 = 0 \).

Now we show that \( \Pi_{R,k} \subset \Pi_{R,k}^{nc} \oplus \Pi_{R,k}^c \). Let \( f \in \Pi_{R,k} \). Then \( \tilde{f} := Tf \) lies in \( \Pi_{R,k} \). Lemma 5.3.5 implies that there exist \( \tilde{f}^{nc} \in \Pi_{R,k}^{nc} \) and \( \tilde{f}^c \in \Pi_{R,k}^c \) such that \( \tilde{f} = \tilde{f}^{nc} + \tilde{f}^c \). Therefore we have \( f = f^{nc} + f^c \) with \( f^{nc} := T^{-1} \tilde{f}^{nc} \in \Pi_{R,k}^{nc} \) and \( f^c := T^{-1} \tilde{f}^c \in \Pi_{R,k}^c \). We conclude that \( f \) lies in \( \Pi_{R,k}^{nc} \oplus \Pi_{R,k}^c \).

Finally we show that \( \Pi_{R,k}^{nc} \oplus \Pi_{R,k}^c \subset \Pi_{R,k} \). Since \( \Pi_{R,k}^{nc} \) and \( \Pi_{R,k}^c \) are subspaces of \( \Pi_{R,k} \), it follows immediately that \( \Pi_{R,k}^{nc} \oplus \Pi_{R,k}^c \subset \Pi_{R,k} \). This completes the proof of the first part of this lemma.

(ii) \( f \in \ker L_R \iff L_R f = 0 \iff TL_R f = 0 \iff \tilde{L}_R T f = 0 \iff T f \in \ker(\tilde{L}_R) \iff T f \in \Pi_{R,k}^c \iff f \in \Pi_{R,k}^c \).

(iii) \( L_R(\Pi_{R,k}) = \Pi_{R,k}^{nc} \iff TL_R(\Pi_{R,k}) = T(\Pi_{R,k}^{nc}) \iff \tilde{L}_R T(\Pi_{R,k}) = T(\Pi_{R,k}^{nc}) \iff \tilde{L}_R(\Pi_{R,k}) = \Pi_{R,k}^{nc} \).

The claim follows from Lemma 5.3.5.

(iv) \( L_R f = g \iff TL_R f = Tg \iff \tilde{L}_R T f = Tg \iff T f = \tilde{L}_R^{-1} T g \iff f = T^{-1}(\tilde{L}_R^{-1}(Tg)) \).

5.4 The Symmetry of the Perturbation Equations

In Section 4.3 we showed that the matrix \( H^{10} \) is reversible with a certain reversibility matrix \( R_D \). In Section 5.2 we introduced a last change of coordinates \( \tau \) to diagonalize the operator \( L \).
In this section we first compute the reversibility matrix of the transformed matrix $\bar{H}^{10}$. Then we derive the symmetry property of the perturbation equations (5.4).

**Lemma 5.4.1**

Let $H^{10}$ be the matrix defined in Section 4.3 and let $\tau$ denote the transformation between $x$-$y$-coordinates and $\xi$-$\eta$-coordinates defined in section 5.2.

For $\bar{H}^{10} := (D\tau)^{-1}H^{10}D\tau$ the following holds:

(i) $\bar{H}^{10}$ has block diagonal form:

$$
\bar{H}^{10} := \begin{pmatrix} 
\bar{H}_1^{10} & & \\
& \ddots & \\
& & \bar{H}_n^{10}
\end{pmatrix}, \quad \text{where} \quad \bar{H}_j^{10} := \begin{pmatrix} 
i \omega_j & 0 \\
0 & -i \omega_j
\end{pmatrix}.
$$

(ii) $\bar{H}^{10}$ is reversible with the reversibility matrix $\bar{R} := (D\tau)^{-1}R_DD\tau$:

$$
\bar{R}\bar{H}^{10} + \bar{H}^{10}\bar{R} = 0.
$$

(iii) $\bar{R}$ has block diagonal form:

$$
\bar{R} := \begin{pmatrix} 
\bar{R}_1 & & \\
& \ddots & \\
& & \bar{R}_n
\end{pmatrix}, \quad \text{where} \quad \bar{R}_j := r_j \begin{pmatrix} 
0 & e^{i\delta_j} \\
e^{-i\delta_j} & 0
\end{pmatrix}.
$$

**Proof:**

(i) $D\tau$ has the following form:

$$
D\tau := \begin{pmatrix} 
D_{\tau_1} & & \\
& \ddots & \\
& & D_{\tau_n}
\end{pmatrix}, \quad \text{where} \quad D_{\tau_j} := \frac{1}{\sqrt{2}} \begin{pmatrix} 
1 & i \\
i & 1
\end{pmatrix}.
$$

Therefore we have

$$
\bar{H}^{10} := (D\tau)^{-1}H^{10}D\tau =
$$

$$
= \begin{pmatrix} 
(D\tau_1)^{-1} & & \\
& \ddots & \\
& & (D\tau_n)^{-1}
\end{pmatrix} \begin{pmatrix} 
H_1^{10} & & \\
& \ddots & \\
& & H_n^{10}
\end{pmatrix} \begin{pmatrix} 
D_{\tau_1} & & \\
& \ddots & \\
& & D_{\tau_n}
\end{pmatrix} =
$$

$$
= \begin{pmatrix} 
(D\tau_1)^{-1}H_1^{10}D_{\tau_1} & & \\
& \ddots & \\
& & (D\tau_n)^{-1}H_n^{10}D_{\tau_n}
\end{pmatrix}.
$$
5.4. The Symmetry of the Perturbation Equations

and

\[ \tilde{H}_j^{10} = (D\tau_j)^{-1} H_j^{10} D\tau_j = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} 0 & \omega_j \\ -\omega_j & 0 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \]

\[ \frac{1}{2} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} i\omega_j & \omega_j \\ -\omega_j & -i\omega_j \end{pmatrix} = \begin{pmatrix} i\omega_j & 0 \\ 0 & -i\omega_j \end{pmatrix}. \]

(ii) We check the reversibility condition.

\[ \tilde{R}\tilde{H}^{10} + \tilde{H}^{10}\tilde{R} = ((D\tau)^{-1} R_D D\tau)((D\tau)^{-1} H^{10} D\tau) + \]

\[ + ((D\tau)^{-1} H^{10} D\tau)((D\tau)^{-1} R_D D\tau) = \]

\[ = (D\tau)^{-1} (R_D H^{10} + H^{10} R_D) D\tau = 0. \]

In the last step we used that \( H^{10} \) is reversible (cf. Lemma 4.3.1).

(iii) For \( \tilde{R} \) we find

\[ \tilde{R} := (D\tau)^{-1} R_D D\tau = \]

\[ = \left( (D\tau_1)^{-1} \ldots (D\tau_n)^{-1} \right) \begin{pmatrix} R_1 & \cdots & R_n \end{pmatrix} \begin{pmatrix} D\tau_1 & \cdots & D\tau_n \end{pmatrix} = \]

\[ = \begin{pmatrix} (D\tau_1)^{-1} R_1 D\tau_1 \\ \cdots \\ (D\tau_n)^{-1} R_n D\tau_n \end{pmatrix} \]

with

\[ \tilde{R}_j = (D\tau_j)^{-1} R_j D\tau_j = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} r_j \begin{pmatrix} \sin \beta_j & \cos \beta_j \\ \cos \beta_j & -\sin \beta_j \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} = \]

\[ = \frac{1}{2} r_j \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix} \begin{pmatrix} \sin \beta_j + i \cos \beta_j & i \sin \beta_j + \cos \beta_j \\ \cos \beta_j - i \sin \beta_j & i \cos \beta_j - \sin \beta_j \end{pmatrix} = \]

\[ = r_j \begin{pmatrix} \cos \beta_j - i \sin \beta_j & 0 \\ 0 & \cos \beta_j + i \sin \beta_j \end{pmatrix} = r_j \begin{pmatrix} 0 & e^{i\beta_j} \\ e^{-i\beta_j} & 0 \end{pmatrix}. \]

This completes the proof. \[ \square \]

In the next Lemma we establish the symmetry property of the perturbation terms on the right-hand side of system (5.1).
Lemma 5.4.2

Let

\[ \Pi^i_{R,k} := \left\{ (\hat{h}_{1,i}, \hat{h}_{2,i}) \in \Pi_{R,k} \times \Pi_{R,k} \mid \begin{pmatrix} \hat{h}_{1,i}(-t, \bar{R}_\zeta, \Delta \sigma) \\ \hat{h}_{2,i}(-t, \bar{R}_\zeta, \Delta \sigma) \end{pmatrix} = -R_i \begin{pmatrix} \bar{h}_{1,i}(t, \zeta, \Delta \sigma) \\ \bar{h}_{2,i}(t, \zeta, \Delta \sigma) \end{pmatrix} \right\} . \]

The upper index \( i \) indicates that the spaces depend not only on the full matrices \( R_D \) and \( \bar{R} \), respectively, but in particular on the single block \( R_i \).

Let \( (\hat{h}^k_{1,i}, \hat{h}^k_{2,i}) \in \Pi^i_{R,k} \). Then

\[
\begin{align*}
\hat{h}^1_{1,i}(t, \zeta, \Delta \sigma) &= \sum_{|p|+|q|+|s| = k} h^{p,q,s}_{1,i}(t) \xi^p \eta^q \Delta \sigma^s \\
\hat{h}^1_{2,i}(t, \zeta, \Delta \sigma) &= \sum_{|p|+|q|+|s| = k} h^{p,q,s}_{2,i}(t) \xi^p \eta^q \Delta \sigma^s
\end{align*}
\]

have the following properties:

\[
\begin{align*}
\hat{h}_{1,i}^{p,q,s}(-t) &= -e^{-i \beta (p-q)} r^{-(p+q) + e_i} \left( \sin \beta \hat{h}_{1,i}^{q,p,s}(t) + \cos \beta \hat{h}_{2,i}^{q,p,s}(t) \right) , \\
\hat{h}_{2,i}^{p,q,s}(-t) &= -e^{-i \beta (p-q)} r^{-(p+q) + e_i} \left( \cos \beta \hat{h}_{1,i}^{q,p,s}(t) - \sin \beta \hat{h}_{2,i}^{q,p,s}(t) \right) ,
\end{align*}
\]

where

\[ e_i := (0, \ldots, 0, \frac{1}{i^{th \text{ position}}}, 0, \ldots, 0). \]

Proof: We substitute the expressions of \( \hat{h}^k_{1,i}(t, \zeta, \Delta \sigma) \) and \( \hat{h}^k_{2,i}(t, \zeta, \Delta \sigma) \) into the reversibility condition. This leads to

\[
0 = \begin{pmatrix} \hat{h}_{1,i}(-t, \bar{R}_\zeta, \Delta \sigma) \\ \hat{h}_{2,i}(-t, \bar{R}_\zeta, \Delta \sigma) \end{pmatrix} + R_i \begin{pmatrix} \bar{h}_{1,i}(t, \zeta, \Delta \sigma) \\ \bar{h}_{2,i}(t, \zeta, \Delta \sigma) \end{pmatrix} =
\]

\[
\begin{align*}
&= \sum_{|p|+|q|+|s| = k} \left( \hat{h}_{1,i}^{p,q,s}(-t)(re^{i \beta} \eta^p)(re^{-i \beta} \xi^q) \Delta \sigma^s \right) + \\
&+ \sum_{|p|+|q|+|s| = k} r_i \left( \sin \beta \hat{h}_{1,i}^{p,q,s}(t)(\xi^p \eta^q \Delta \sigma^s) + \cos \beta \hat{h}_{2,i}^{p,q,s}(t)(\xi^p \eta^q \Delta \sigma^s) \right) =
\end{align*}
\]
5.4. The Symmetry of the Perturbation Equations

\[ Y^{\beta}p(f(-t)r^p+q)e^{i\beta(p-q)} + \sum_{|p|+|q|+|s|=k} r^p \left( \sin \beta r^p_{2,1}(t) + \cos \beta r^p_{2,1}(t) \right) \Delta \sigma^s. \]

Now the claim follows immediately. \qed

In the next lemma we introduce two spaces of polynomials with some symmetry properties and investigate the symmetries of the coefficients.

**Lemma 5.4.3**

For \( k \leq N \) (cf. Assumption A5b) let

\[ \Pi_{\mathbb{R},k}^\rightarrow := \{ K \in \Pi_{\mathbb{R},k} \mid K(-t, R\zeta, \Delta \sigma) = -r^2 K(t, \zeta, \Delta \sigma) \}, \]

\[ \Pi_{\mathbb{R},k}^\leftarrow := \{ \bar{K} \in \Pi_{\mathbb{R},k} \mid \bar{K}(-t, R\zeta, \Delta \sigma) = r^2 \bar{K}(t, \zeta, \Delta \sigma) \}. \]

For

\[ \bar{K}^k_i(t, \zeta, \Delta \sigma) := \sum_{|p|+|q|+|s|=k} K^{p,q,s}_i(t) \xi^p \eta^q \Delta \sigma^s \]

\[ \bar{I}^k_i(t, \zeta, \Delta \sigma) := \sum_{|p|+|q|+|s|=k} I^{p,q,s}_i(t) \xi^p \eta^q \Delta \sigma^s \]

the following holds:

(i) \( \bar{K}^k_i \in \Pi_{\mathbb{R},k}^\rightarrow \) if and only if \( \bar{K}^{p,q,s}_i(-t) = -r^{-(p+q)+2} e^{-i\beta(p-q)} K^{p,q,s}_i(t) \) for all coefficients.

(ii) \( \bar{I}^k_i \in \Pi_{\mathbb{R},k}^\leftarrow \) if and only if \( \bar{I}^{p,q,s}_i(-t) = r^{-(p+q)+2} e^{-i\beta(p-q)} I^{p,q,s}_i(t) \) for all coefficients.

**Proof:**

(i) On the one hand we have

\[ -r^2 \bar{K}^1_i(t, \zeta, \Delta \sigma) = -r^2 \sum_{|p|+|q|+|s|=l} K^{p,q,s}_i(t) \xi^p \eta^q \Delta \sigma^s \]

and on the other hand

\[ \bar{K}^1_i(-t, R\zeta, \Delta \sigma) = \sum_{|p|+|q|+|s|=l} \bar{K}^{p,q,s}_i(-t) (r e^{i\beta})^p (r e^{-i\beta})^q \Delta \sigma^s \]
We conclude that $K_i^t \in \Pi_{R,l}^i$ if and only if

$$K_i^{p,q,s}(-t) = -r^{-(p+q)+2\nu\tau}e^{-i\beta(p-q)}K_i^{p,q,s}(t).$$

(ii) The proof of the second part is almost identical. $\square$

**Lemma 5.4.4**

For $k \leq N$ (cf. Assumption $A_{5b}$) let

$$(h_{1,i}, h_{2,i}) \in \Pi_{R,l}^i \quad \text{for} \quad 2 \leq l \leq k-1,$$

$$\bar{h}_l^i \in \Pi_{R,l}^{i,+} \quad \text{for} \quad 2 \leq l \leq k-1$$

and let the right-hand side of the perturbation equation (5.6) be defined by

$$K_i^k := \frac{1}{\sqrt{2}} \left( (\xi_i + i\eta_i)h_{1,i}^{k-1} + i(\xi_i - i\eta_i)h_{2,i}^{k-1} \right) - \frac{1}{\sqrt{2}} \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial h_{1,i}^{l+1}}{\partial \eta_j} - i \frac{\partial h_{2,i}^{l+1}}{\partial \xi_j} \right) \bar{h}_{1,i}^{l+1} - i \left( \frac{\partial h_{1,i}^{l+1}}{\partial \xi_j} + i \frac{\partial h_{2,i}^{l+1}}{\partial \eta_j} \right) \bar{h}_{2,i}^{l+1}.$$ 

Then the following holds:

(i) The coefficients of

$$K_i^k(t, \zeta, \Delta \sigma) =: \sum_{|p|+|q|+|s|=k} K_i^{p,q,s}(t) \xi^p \eta^q \Delta \sigma^s$$

have the following symmetry property:

$$K_i^{p,q,s}(-t) = -r^{-(p+q)+2\nu\tau}e^{-i\beta(p-q)}K_i^{q,p,s}(t).$$

(ii) $K_i^k \in \Pi_{R,k}^i$.

**Proof:**

(i) Let

$$\bar{h}_{1,i}^{k-1} = \sum_{|p|+|q|+|s|=k-1} \bar{h}_{1,i}^{p,q,s}(t) \xi^p \eta^q \Delta \sigma^s,$$

$$\bar{h}_{2,i}^{k-1} = \sum_{|p|+|q|+|s|=k-1} \bar{h}_{2,i}^{p,q,s}(t) \xi^p \eta^q \Delta \sigma^s.$$
5.4. The Symmetry of the Perturbation Equations

We split $-\sqrt{2}\hat{K}^+_{\pm}$ into two parts.
In a first step we consider $(\xi_i + i\eta_i)\hat{h}_{1,i}^{k-1} + i(\xi_i - i\eta_i)\hat{h}_{2,i}^{k-1}$. For this term we obtain

$$(\xi_i + i\eta_i)\hat{h}_{1,i}^{k-1} + i(\xi_i - i\eta_i)\hat{h}_{2,i}^{k-1} =$$

$$= \sum_{|p|+|q|+|s|=k-1} \hat{h}^{p,q,s}_{1,i}(t)(\xi_i + i\eta_i)\xi^q\eta^s \Delta \sigma^s +$$

$$+ i \sum_{|p|+|q|+|s|=k-1} \hat{h}^{p,q,s}_{2,i}(t)(\xi_i - i\eta_i)\xi^q\eta^s \Delta \sigma^s =$$

$$= \sum_{|p|+|q|+|s|=k-1} (\hat{h}^{p,q,s}_{1,i}(t)\xi^{p+e_i\eta^q} \Delta \sigma^s + i\hat{h}^{p,q,s}_{2,i}(t)\xi^{p+e_i\eta^q} \Delta \sigma^s) +$$

$$+ i \sum_{|p|+|q|+|s|=k-1} (\hat{h}^{p,q,s}_{1,i}(t)\xi^{p+e_i\eta^q} \Delta \sigma^s - i\hat{h}^{p,q,s}_{2,i}(t)\xi^{p+e_i\eta^q} \Delta \sigma^s) =$$

$$= \sum_{|p|+|q|+|s|=k-1} (\hat{h}^{p,q,s}_{1,i}(t) + i\hat{h}^{p,q,s}_{2,i}(t))\xi^{p+e_i\eta^q} \Delta \sigma^s +$$

$$+ i \sum_{|p|+|q|+|s|=k-1} (\hat{h}^{p,q,s}_{1,i}(t) - i\hat{h}^{p,q,s}_{2,i}(t))\xi^{p+e_i\eta^q} \Delta \sigma^s.$$

Thus the coefficients of $\xi^{p\eta^q} \Delta \sigma^s$ read

$$c^{p,q,s}_i(t) := (\hat{h}^{p-q,-e_i,s}_{1,i}(t) + i\hat{h}^{p-q,-e_i,s}_{2,i}(t)) + i (\hat{h}^{p-q,-e_i,s}_{1,i}(t) - i\hat{h}^{p-q,-e_i,s}_{2,i}(t)).$$

With Lemma 5.4.2 we find

$$c^{p,q,s}_i(-t) = (\hat{h}^{p-q,-e_i,s}_{1,i}(-t) + i\hat{h}^{p-q,-e_i,s}_{2,i}(-t)) + i (\hat{h}^{p-q,-e_i,s}_{1,i}(-t) - i\hat{h}^{p-q,-e_i,s}_{2,i}(-t)) =$$

$$= -r^{-(p+q+e_i+e_i)} e^{-i\beta(p+q-e_i-q)} (\sin \beta_i \hat{h}^{q,-e_i,s}_{1,i}(t) + \cos \beta_i \hat{h}^{q,-e_i,s}_{2,i}(t)) -$$

$$- r^{-(p+q+e_i+e_i)} e^{-i\beta(p+q-e_i-q)} (\cos \beta_i \hat{h}^{q,-e_i,s}_{1,i}(t) - \sin \beta_i \hat{h}^{q,-e_i,s}_{2,i}(t)) -$$

$$- r^{-(p+q+e_i+e_i)} e^{-i\beta(p+q+e_i)} (\sin \beta_i \hat{h}^{q,-e_i,s}_{1,i}(t) + \cos \beta_i \hat{h}^{q,-e_i,s}_{2,i}(t)) -$$

$$- r^{-(p+q+e_i+e_i)} e^{-i\beta(p+q+e_i)} (\cos \beta_i \hat{h}^{q,-e_i,s}_{1,i}(t) - \sin \beta_i \hat{h}^{q,-e_i,s}_{2,i}(t)) =$$

$$= -r^{-(p+q)} e^{-i\beta(p-q)} e^{i\beta_i} (\cos \beta_i - i \sin \beta_i \hat{h}^{q,-e_i,s}_{1,i}(t)) -$$

$$- r^{-(p+q)} e^{-i\beta(p-q)} e^{i\beta_i} (\cos \beta_i - i \sin \beta_i \hat{h}^{q,-e_i,s}_{2,i}(t)) -$$

$$- r^{-(p+q)} e^{-i\beta(p-q)} e^{i\beta_i} (\cos \beta_i + i \sin \beta_i \hat{h}^{q,-e_i,s}_{1,i}(t)) -$$

$$- r^{-(p+q)} e^{-i\beta(p-q)} e^{i\beta_i} (\cos \beta_i + i \sin \beta_i \hat{h}^{q,-e_i,s}_{2,i}(t)) =$$
In a second step we consider the terms

\[
\left( \frac{\partial I_i^l}{\partial \xi_j} - i \frac{\partial I_i^l}{\partial \eta_j} \right) \tilde{h}_{1,j}^{k-l} - i \left( \frac{\partial I_i^l}{\partial \xi_j} + i \frac{\partial I_i^l}{\partial \eta_j} \right) \tilde{h}_{2,j}^{k-l} =
\]

Let

\[
I_i^l(t, \zeta, \Delta \sigma) := \sum_{|p|_1+|q|_1+|s|_1=l} I_i^{p,q,s}(t) \xi^p \eta^q \Delta \sigma^s.
\]

From the assumptions and Lemma 5.4.3 (ii) it follows that the coefficients \( I_i^{p,q,s}(t) \) have the following symmetry property:

\[
I_i^{p,q,s}(-t) = r^{-(p+q)+2e_i e^{-i\beta(p-q)}} I_i^{q,p,s}(t).
\]

With this in mind we are able to go on:

\[
\left( \frac{\partial I_i^l}{\partial \xi_j} - i \frac{\partial I_i^l}{\partial \eta_j} \right) \tilde{h}_{1,j}^{k-l} - i \left( \frac{\partial I_i^l}{\partial \xi_j} + i \frac{\partial I_i^l}{\partial \eta_j} \right) \tilde{h}_{2,j}^{k-l} =
\]

\[
= \sum_{|p'|_1+|q'|_1+|s'|_1=l} p_j^{p',q',s'}(t) \xi^{p'-e_j} \eta^{q'} \Delta \sigma^{s'}.
\]

\[
- \sum_{|p''|_1+|q''|_1+|s''|_1=k-l+1} \left( \tilde{h}_{1,j}^{p''q''s''}(t) - i \tilde{h}_{2,j}^{p''q''s''}(t) \right) \xi^{p''} \eta^{q''} \Delta \sigma^{s''} -
\]

\[
- \sum_{|p'|_1+|q'|_1+|s'|_1=k-l+1} q_j^{p',q',s'}(t) \xi^{p'-e_j} \eta^{q'} \Delta \sigma^{s'}.
\]

\[
+ \sum_{|p''|_1+|q''|_1+|s''|_1=k-l+1} \left( \tilde{h}_{1,j}^{p''q''s''}(t) + i \tilde{h}_{2,j}^{p''q''s''}(t) \right) \xi^{p''} \eta^{q''} \Delta \sigma^{s''} =
\]
5.4. The Symmetry of the Perturbation Equations

\[
\sum_{|p'|_1+|q'|_1+|s'|_1=1} \sum_{|p''|_1+|q''|_1+|s''|=k-l+1} p'_{j,i} p''_{j,i} \left( \tilde{h}^{p'q's'}_{1,j} (t) - i \tilde{h}^{p''q''s''}_{2,j} (t) \right) \cdot \xi^{q''} e^{q'' \Delta \sigma^{s''}} - \\
- i \sum_{|p'|_1+|q'|_1+|s'|_1=1} \sum_{|p''|_1+|q''|_1+|s''|=k-l+1} q'_{j,i} p''_{j,i} \left( \tilde{h}^{p'q's'}_{1,j} (t) + i \tilde{h}^{p''q''s''}_{2,j} (t) \right) \cdot \xi^{q''} e^{q'' \Delta \sigma^{s''}}.
\]

Thus the coefficient of \( \xi^{q''} e^{q'' \Delta \sigma^{s''}} \) consists of terms of the form

\[
d_{i,j}^{p,q,s,s'} (t) := \left( (p'_{j,i} + 1) \tilde{h}^{p'q's'}_{1,j} (t) \left( \tilde{h}^{p-p'q'-s'-s'}_{1,j} (t) - i \tilde{h}^{p-p'q'-s'-s'}_{2,j} (t) \right) - \\
- i (q'_{j,i} + 1) \tilde{h}^{p'q's'}_{1,j} (t) \left( \tilde{h}^{p-p'q'-s'-s'}_{1,j} (t) + i \tilde{h}^{p-p'q'-s'-s'}_{2,j} (t) \right) \right).
\]

From the symmetry properties of \( \tilde{h}^{p,q,s}_{1,j} \), \( \tilde{h}^{p,q,s}_{2,j} \) (cf. Lemma 5.4.2) and \( I^p_{i,q,s} \) (cf. above) it follows that

\[
d_{i,j}^{p,q,s,s'} (-t) = \\
= (p'_{j,i} + 1) \tilde{h}^{p'q's'}_{1,j} (-t) \left( \tilde{h}^{p-p'q'-s'-s'}_{1,j} (-t) - i \tilde{h}^{p-p'q'-s'-s'}_{2,j} (-t) \right) - \\
- i (q'_{j,i} + 1) \tilde{h}^{p'q's'}_{1,j} (-t) \left( \tilde{h}^{p-p'q'-s'-s'}_{1,j} (-t) + i \tilde{h}^{p-p'q'-s'-s'}_{2,j} (-t) \right) = \\
= (p'_{j,i} + 1) r^{-(p'+q') + 2e_{j} e^{-i\beta(p'-q') + e_{j}}} \tilde{h}^{p'q's'}_{1,j} (t) \cdot (-1) \cdot r^{-(p-p' + q') + e_{j} e^{-i\beta((p-p') - (q-q'))}} \cdot \\
\cdot \left( \sin \beta_{j} \tilde{h}^{q-p'-p'q'-s'-s'}_{1,j} (t) \cos \beta_{j} \tilde{h}^{q-p'-p'q'-s'-s'}_{2,j} (t) - \\
- i \left( \cos \beta_{j} \tilde{h}^{q-p'-p'q'-s'-s'}_{1,j} (t) - \sin \beta_{j} \tilde{h}^{q-p'-p'q'-s'-s'}_{2,j} (t) \right) \right) - \\
- i (q'_{j,i} + 1) r^{-(p'+q') + 2e_{j} e^{-i\beta(p'-q') + e_{j}}} \tilde{h}^{p'q's'}_{1,j} (t) \cdot (-1) \cdot r^{-(p-p' + q') + e_{j} e^{-i\beta((p-p') - (q-q'))}} \cdot \\
\cdot \left( \sin \beta_{j} \tilde{h}^{q-p'-p'q'-s'-s'}_{1,j} (t) + \cos \beta_{j} \tilde{h}^{q-p'-p'q'-s'-s'}_{2,j} (t) + \\
+ i \left( \cos \beta_{j} \tilde{h}^{q-p'-p'q'-s'-s'}_{1,j} (t) - \sin \beta_{j} \tilde{h}^{q-p'-p'q'-s'-s'}_{2,j} (t) \right) \right).
\begin{align*}
= -(p' + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( -i \cos \beta_j - i \sin \beta_j \right) \tilde{h}'_{q'+p'+j,s'}(t) + (\cos \beta_j - i \sin \beta_j \right) \tilde{h}'_{q'+p'+j,s'}(t) + \\
+ i(q'_j + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( i \cos \beta_j - i \sin \beta_j \right) \tilde{h}'_{q'+p'+j,s'}(t) + (\cos \beta_j - i \sin \beta_j \right) \tilde{h}'_{q'+p'+j,s'}(t) = \\
= -(p' + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( -i \tilde{h}'_{q'+p'+j,s'}(t) + \tilde{h}'_{q'+p'+j,s'}(t) \right) + \\
+ i(q'_j + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( i \tilde{h}'_{q'+p'+j,s'}(t) + \tilde{h}'_{q'+p'+j,s'}(t) \right) = \\
= -(q'_j + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( \tilde{h}'_{q'+p'+j,s'}(t) - \tilde{h}'_{q'+p'+j,s'}(t) \right) - \\
+ i(p'_j + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( i \tilde{h}'_{q'+p'+j,s'}(t) + i \tilde{h}'_{q'+p'+j,s'}(t) \right) = \\
= -r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{i(p'_j + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( \tilde{h}'_{q'+p'+j,s'}(t) - \tilde{h}'_{q'+p'+j,s'}(t) \right) - \\
+ i(p'_j + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( i \tilde{h}'_{q'+p'+j,s'}(t) + i \tilde{h}'_{q'+p'+j,s'}(t) \right). \\
\end{align*}

The symmetry properties of \( e^{\tilde{p}'q's'(t)} \) and \( d^{p'}q's'(t) \) imply

\[ K^{p,q,s}(-t) = -r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{i(p'_j + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( \tilde{h}'_{q'+p'+j,s'}(t) - \tilde{h}'_{q'+p'+j,s'}(t) \right) - \\
+ i(p'_j + 1)r^{-(p+q)+2e_1e^{-i\beta(p-q)} \cdot e^{-i\beta_j \tilde{h}'_{q'+p'+j,s'}(t)} \cdot \\
\quad \left( i \tilde{h}'_{q'+p'+j,s'}(t) + i \tilde{h}'_{q'+p'+j,s'}(t) \right). \]

This completes the proof of the first part of the lemma.

(ii) The claim follows from the first part of Lemma 5.4.3

The next lemmas transform the results obtained above back to \( x-y \)-coordinates.

**Lemma 5.4.5**

For \( k \leq N \) (cf. Assumption A5b), let

\[ \Pi_{E,k} := \left\{(h_1, h_2) \in \Pi_{E,k} \times \Pi_{E,k} \mid \left(h_1(-t, R_{Qz}, \Delta \sigma), h_2(-t, R_{Qz}, \Delta \sigma) = -R_i \left(h_1(t, z, \Delta \sigma), h_2(t, z, \Delta \sigma) \right) \right\} \]

Then \( (h_{1,a}(t, \zeta, \Delta \sigma), h_{2,a}(t, \zeta, \Delta \sigma))^T := (h_{1,a}(t, \tau(\zeta), \Delta \sigma), h_{2,a}(t, \tau(\zeta), \Delta \sigma))^T \in \Pi_{E,k}^i. \)
5.4. The Symmetry of the Perturbation Equations

Proof: Lemma 5.4.1 (ii) implies that \( R_D \tau(\zeta) = R_D D \tau \zeta = D \tau \tilde{R} \zeta = \tau(\tilde{R} \zeta) \). Thus from the assumptions we find

\[
0 = \left( h_{1,t}(-t, R_D \tau(\zeta), \Delta \sigma) \right) + R_t \left( h_{1,t}(t, \tau(\zeta), \Delta \sigma) \right) =
\]

\[
\left( h_{1,t}(-t, \tau(\tilde{R} \zeta), \Delta \sigma) \right) + R_t \left( h_{1,t}(t, \tau(\zeta), \Delta \sigma) \right) =
\]

\[
\left( h_{1,t}(-t, \tilde{R} \zeta, \Delta \sigma) \right) + R_t \left( h_{1,t}(t, \zeta, \Delta \sigma) \right).
\]

In the last step we used the definition of \( \bar{h}_{1,t}^k \) and \( \bar{h}_{2,t}^k \).

Thus we have \( (\bar{h}_{1,t}^k, \bar{h}_{2,t}^k) \in \Pi_{s,k} \).

Lemma 5.4.6

For \( k \leq N \) (cf. Assumption A5b) let

\[
\Pi_{s,k}^{-} := T^{-1} \Pi_{s,k}^{-}, \quad \text{and} \quad \Pi_{s,k}^{+} := T^{-1} \Pi_{s,k}^{+}.
\]

For \( 2 \leq l \leq k - 1 \)

\[
(\bar{h}_{1,t}^l, \bar{h}_{2,t}^l) \in \Pi_{s,l} \quad \text{and} \quad I_i^l \in \Pi_{s,l}^{+}.
\]

let the right-hand side of the perturbation equation (5.4) be defined by

\[
K_i^k := -x_i h_{1,t}^{k-1} - y_i h_{2,t}^{k-1} - \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial I_i^l}{\partial x_j} h_{1,j}^{k-1-l} + \frac{\partial I_i^l}{\partial y_j} h_{2,j}^{k-1-l} \right). \]

Then \( K_i^k \in \Pi_{s,k}^{-} \).

Proof: Let

\[
(\bar{h}_{1,t}^l, \bar{h}_{2,t}^l) := (T h_{1,t}^l, T h_{2,t}^l) \quad \text{for} \quad 2 \leq l \leq k - 1
\]

\[
\bar{I}_i^l := T I_i^l \quad \text{for} \quad 2 \leq l \leq k - 1.
\]

Lemma 5.4.5, the definition of \( \Pi_{s,k}^{+} \) and Lemma 5.3.4 imply that the assumptions of Lemma 5.4.4 hold for \( K_i^k := T K_i^l \). Thus \( K_i^k \in \Pi_{s,k}^{-} \), and therefore \( K_i^k \in \Pi_{s,k}^{-} \).
Lemma 5.4.7
For \( k \leq N \) (cf. Assumption \( A_{5b} \)) let \( \bar{K}_i^k \in \Pi_{\mathbb{R}, k}^{i,-} \). Then the following holds:

(i) \( \bar{K}_i^k \in \Pi_{\mathbb{R}, k}^{n,k} \).

(ii) \( I_i^k := I_{\mathbb{R}}^{-1} \bar{K}_i^k \in \Pi_{\mathbb{R}, k}^{i,+} \).

Proof:

(i) Let

\[
\bar{K}_i^k(t, \zeta, \Delta \sigma) := \sum_{|p|+|q|+|s|=k} \bar{K}_i^{p,q,s}(t) \xi^p \eta^q \Delta \sigma^s := \sum_{|p|+|q|+|s|=k} \sum_{k_0=-\infty}^{\infty} \bar{K}_i^{p,q,s,k_0} e^{ik_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s.
\]

We first derive the symmetry property of the Fourier coefficients of \( \bar{K}_i^{p,q,s}(t) \).
From Lemma 5.4.3 we have

\[
\bar{K}_i^{p,q,s}(-t) = -r^{-(p+q)+2c_i} e^{-i \beta(p-q)} \bar{K}_i^{q,p,s}(t).
\]

On the one hand we find

\[
\bar{K}_i^{p,q,s}(-t) = \sum_{k_0=-\infty}^{\infty} \bar{K}_i^{p,q,s,k_0} e^{ik_0 \omega_0 t} = \sum_{k_0=-\infty}^{\infty} \bar{K}_i^{p,q,s,-k_0} e^{ik_0 \omega_0 t}
\]

and on the other hand

\[
- r^{-(p+q)+2c_i} e^{-i \beta(p-q)} \bar{K}_i^{q,p,s}(t) = - r^{-(p+q)+2c_i} e^{-i \beta(p-q)} \sum_{k_0=-\infty}^{\infty} \bar{K}_i^{q,p,s,k_0} e^{ik_0 \omega_0 t} = \sum_{k_0=-\infty}^{\infty} \left( - r^{-(p+q)+2c_i} e^{-i \beta(p-q)} \bar{K}_i^{q,p,s,k_0} \right) e^{ik_0 \omega_0 t}.
\]

Thus it follows that

\[
\bar{K}_i^{p,q,s,-k_0} = - r^{-(p+q)+2c_i} e^{-i \beta(p-q)} \bar{K}_i^{q,p,s,k_0}.
\] (5.8)

A single monomial \( \sum_{k_0=-\infty}^{\infty} \bar{K}_i^{p,q,s,k_0} e^{(k_0 \omega_0) t} \xi^p \eta^q \Delta \sigma^s \) is critical if and only if \( p = q \) and \( \bar{K}_i^{p,q,s,0} \neq 0 \) (cf. the definition of \( \Pi_k^{i} \)). But for \( p = q \) and \( k_0 = 0 \) it follows from (5.8) that

\[
\bar{K}_i^{p,p,s,0} = - r^{-(2p-c_i)} \bar{K}_i^{p,p,s,0}.
\]

Since the \( r_j \) are positive we conclude that \( \bar{K}_i^{p,p,s,0} = 0 \) and therefore \( \bar{K}_i^k \in \Pi_{\mathbb{R}, k}^{nc} \).
5.4. The Symmetry of the Perturbation Equations

(ii) We first determine the Fourier coefficients of \( \bar{I}_k^{p,q,s}(t) \). To this end let

\[
\bar{K}_k(t, \zeta, \Delta \sigma) := \sum_{|p|+|q|+|s| = k} \bar{K}_k^{p,q,s}(t) \xi^p \eta^q \Delta \sigma^s := \sum_{|p|+|q|+|s| = k} \sum_{k_0 = -\infty}^{\infty} \bar{K}_k^{p,q,s,k_0} e^{ik_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s.
\]

Then Lemma 5.3.2 (iv) implies that

\[
\bar{I}_k = L^{-1} \bar{K}_k(t) = \sum_{|p|+|q|+|s| = k} \sum_{k_0 = -\infty}^{\infty} \frac{\bar{K}_k^{p,q,s,k_0}}{i((p - q) \omega + k_0 \omega_0)} e^{ik_0 \omega_0 t} \xi^p \eta^q \Delta \sigma^s.
\]

Thus

\[
\bar{I}_k^{p,q,s,k_0} = \frac{1}{i((p - q) \omega + k_0 \omega_0)} \bar{K}_k^{p,q,s,k_0}.
\]

Now we determine the symmetry property of the Fourier coefficients \( \bar{I}_k^{p,q,s,k_0} \).

\[
\bar{I}_k^{p,q,s,-k_0} = \frac{1}{i((p - q) \omega - k_0 \omega_0)} \bar{K}_k^{p,q,s,-k_0} = \frac{1}{i((p - q) \omega - k_0 \omega_0)} (-r^{-(p+q)+2e_i} e^{-i\beta(p-q)}) \bar{K}_k^{q,p,s,k_0} = r^{-(p+q)+2e_i} e^{-i\beta(p-q)} \frac{1}{i((q - p) \omega + k_0 \omega_0)} \bar{K}_k^{q,p,s,k_0} = r^{-(p+q)+2e_i} e^{-i\beta(p-q)} \bar{I}_k^{q,p,s,k_0}.
\]

This result allows us to determine the symmetry property of \( \bar{I}_k^{p,q,s}(t) \).

\[
\bar{I}_k^{p,q,s}(-t) = \sum_{k_0 = -\infty}^{\infty} \bar{I}_k^{p,q,s,k_0} e^{ik_0 \omega_0 (-t)} = \sum_{k_0 = -\infty}^{\infty} \bar{I}_k^{p,q,s,-k_0} e^{ik_0 \omega_0 t} = \sum_{k_0 = -\infty}^{\infty} r^{-(p+q)+2e_i} e^{-i\beta(p-q)} \bar{I}_k^{q,p,s,k_0} e^{ik_0 \omega_0 t} = r^{-(p+q)+2e_i} e^{-i\beta(p-q)} \sum_{k_0 = -\infty}^{\infty} \bar{I}_k^{q,p,s,k_0} e^{ik_0 \omega_0 t} = r^{-(p+q)+2e_i} e^{-i\beta(p-q)} \bar{I}_k^{q,p,s}(t).
\]

Thus we have

\[
\bar{I}_k^{p,q,s}(-t) = r^{-(p+q)+2e_i} e^{-i\beta(p-q)} \bar{I}_k^{q,p,s}(t)
\]

and Lemma 5.4.3 implies \( \bar{I}_k^i \in \bar{I}_k^{i+} \). \( \square \)
Lemma 5.4.8
For $k \leq N$ (cf. Assumption $\mathbf{A}_{5b}$) let $K_i^k \in \Pi^i_{\mathbb{R},k}$. Then the following holds:

(i) $K_i^k \in \Pi^c_{\mathbb{R},k}$.

(ii) $I_i^k := \mathbb{L}_\mathbb{R}^{-1} K_i^k \in \Pi^c_{\mathbb{R},k}$.

Proof:

(i) Let $K_i^k \in \Pi^i_{\mathbb{R},k}$. By the definition of $\Pi^i_{\mathbb{R},k}$ it follows that $K_i^k := T K_i^k \in \Pi^i_{\mathbb{R},k}$.

Lemma 5.4.7 implies that $\overline{K}_i^k \in \Pi^c_{\mathbb{R},k}$ and the definition of $\Pi^c_{\mathbb{R},k}$ (cf. Section 5.3) leads to $K_i^k \in \Pi^c_{\mathbb{R},k}$.

(ii) From the assumption of this lemma and the definition of $\Pi^i_{\mathbb{R},k}$ we conclude that $K_i^k := T K_i^k \in \Pi^i_{\mathbb{R},k}$. By Lemma 5.3.6 we conclude that $I_i^k := T^{-1}(\mathbb{L}_\mathbb{R}^{-1} T K_i^k) = T^{-1}(\mathbb{L}_\mathbb{R}^{-1} K_i^k)$ exists and is unique in $\Pi^c_{\mathbb{R},k}$. Lemma 5.4.7 implies that $\overline{I}_i^k \in \Pi^{c+i}_{\mathbb{R},k}$. By the definition of $\Pi^c_{\mathbb{R},k}$ we conclude that $I_i^k \in \Pi^{c+i}_{\mathbb{R},k}$.

5.5 The Solvability of the Perturbation Scheme

We now are able to prove the solvability of the perturbation equations and therefore the existence of first integrals.

Theorem 5.5.1
Let Assumptions $\mathbf{A}_1 - \mathbf{A}_5$ hold. Then the system of differential equations (3.1) may be transformed to the system of perturbed harmonic oscillators (5.6), where the functions $\bar{h}_{i,j}^k, \bar{h}_{2,i}^k$ on the right-hand side lie in $\Pi^c_{\mathbb{R},k}$ for $2 \leq k \leq N$.

Consider the perturbation scheme

\[ \mathbb{L}_\mathbb{R} \bar{I}_i^k = \bar{K}_i^k, \quad 3 \leq k \leq N, \tag{5.6} \]

where

\[ \mathbb{L}_\mathbb{R} \bar{I}_i^k := i \sum_{j=1}^{n} \left( \frac{\partial \bar{I}_i^k}{\partial \xi_j} \xi_j - \frac{\partial \bar{I}_i^k}{\partial \eta_j} \eta_j \right) \omega_j + \frac{\partial \bar{I}_i^k}{\partial t} \]

and

\[ \bar{K}_i^3 := \frac{1}{\sqrt{2}} \left( (\xi_i + i \eta_i) \bar{h}_{i,i}^2 + i (\xi_i - i \eta_i) \bar{h}_{2,i}^2 \right), \]

\[ \bar{K}_i^k := \frac{1}{\sqrt{2}} \left( (\xi_i + i \eta_i) \bar{h}_{i,i}^{k-1} + i (\xi_i - i \eta_i) \bar{h}_{2,i}^{k-1} \right) - \frac{1}{\sqrt{2}} \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \left( \frac{\partial \bar{I}_i^l}{\partial \xi_j} - i \frac{\partial \bar{I}_i^l}{\partial \eta_j} \right) \bar{h}_{i,j}^{l-1} - i \left( \frac{\partial \bar{I}_i^l}{\partial \xi_j} + i \frac{\partial \bar{I}_i^l}{\partial \eta_j} \right) \bar{h}_{2,j}^{l-1} \right) \quad 4 \leq k \leq N. \]
5.5. The Solvability of the Perturbation Scheme

Then the following holds:

(i) The perturbation scheme is solvable for \( k \leq N \) (cf. Assumption \( A^b \)).

(ii) The polynomials \( \tilde{h}_{i,i}^k, \tilde{h}_{2,i}^k, \tilde{K}_i^k \) and \( \tilde{I}_i^k \) have the following properties:

\[
\tilde{h}_{i,i}^k, \tilde{h}_{2,i}^k \in \Pi_{R,k}^i
\]  (5.9)

\[
\tilde{K}_i^k \in \Pi_{R,k}^i
\]  (5.10)

\[
\tilde{I}_i^k \in \Pi_{R,k}^{i+}
\]  (5.11)

where

\[
\Pi_{R,k} := T(\Pi_{R,k}) = \left\{ T f \in \Pi_k \mid f^{p,q,s}(t) \in \mathbb{R} \right\},
\]

\[
\Pi_{R,k}^i := \left\{ (\tilde{h}_1, \tilde{h}_2) \in \Pi_{R,k} \times \Pi_{R,k} \mid \left( \begin{array}{c}
\tilde{h}_1(-t, R\zeta, \Delta \sigma) \\
\tilde{h}_2(-t, R\zeta, \Delta \sigma)
\end{array} \right) = -R_i \left( \begin{array}{c}
\tilde{h}_1(t, \zeta, \Delta \sigma) \\
\tilde{h}_2(t, \zeta, \Delta \sigma)
\end{array} \right) \right\},
\]

\[
\Pi_{R,k}^{i-} := \left\{ \tilde{K} \in \Pi_{R,k} \mid \tilde{K}(-t, R\zeta, \Delta \sigma) = -r_i^2 \tilde{K}(t, \zeta, \Delta \sigma) \right\},
\]

\[
\Pi_{R,k}^{i+} := \left\{ \tilde{I} \in \Pi_{R,k} \mid \tilde{I}(-t, R\zeta, \Delta \sigma) = r_i^2 \tilde{I}(t, \zeta, \Delta \sigma) \right\}.
\]

**Proof:** The symmetry property (5.9) of the polynomials \( \tilde{h}_{i,i}^k \) follows from Assumption \( A_3 \) and Lemma 5.4.2. Let \( \bar{I}(m) \): denote the following assertion:

\( \bar{I}(m) \): For \( 3 \leq l \leq m \) the following holds:

- \( \tilde{K}_i^l \in \Pi_{R,l}^{i-} \),
- The solutions \( \tilde{I}_i^l \) of the perturbation equations \( \tilde{L}_R \tilde{I}_i^l = \tilde{K}_i^l \) exist and are unique,
- \( \tilde{I}_i^l \in \Pi_{R,l}^{i+} \).

We prove by induction that \( \bar{I}(m) \) holds for every number \( m \geq 3 \).

- For \( m = 3 \) the right-hand side of the perturbation equations (5.6) is reduced to \( \tilde{K}_3^3 = \frac{1}{\sqrt{2}} \left( (\xi + i\eta_1) \tilde{h}_{2,i}^3 + i(\xi - i\eta_1) \tilde{h}_{2,i}^3 \right) \). This implies that the assumptions of Lemma 5.4.4 are fulfilled. It follows that \( \tilde{K}_3^3 \in \Pi_{R,3}^{i-} \). Then Lemma 5.4.7 implies that \( \tilde{I}_3^3 \) uniquely exists and lies in \( \Pi_{R,3}^{i+} \). Thus \( \bar{I}(3) \) holds true.

- \( \bar{I}(m - 1) \Rightarrow \bar{I}(m) \): Let \( \bar{I}(m - 1) \) hold true. This implies that the assumptions of Lemma 5.4.4 are fulfilled. It follows that \( \tilde{K}_i^m \in \Pi_{R,m}^{i-} \). Then Lemma 5.4.7 implies that \( \tilde{I}_i^m \) uniquely exists and lies in \( \Pi_{R,m}^{i+} \). Thus \( \bar{I}(m) \) holds true.
By induction we conclude that $I(m)$ holds for every number $3 \leq m \leq N$.

This theorem implies that the system of perturbed harmonic oscillator (5.6) has approximate first integrals

$$\tilde{I}_i(t, \xi, \eta, \Delta \sigma) := I_i^2(\xi, \eta) + \sum_{k=3}^r \tilde{I}_i^k(t, \xi, \eta, \Delta \sigma),$$

where $\tilde{I}_i^k := i\xi_i \eta_i$.

The theorem above proves the solvability of the perturbation equations in $\xi$-$\eta$-coordinates. Analogously we may state and prove a theorem on the solvability in $x$-$y$-coordinates.

**Theorem 5.5.2**

Let Assumptions $A_1 - A_5$ hold. Then the system of differential equations (3.1) may be transformed to the system of perturbed harmonic oscillators (5.4), where the functions $h_{1,j}^k, h_{2,j}^k$ on the right-hand side lie in $\Pi_{k,k}$ for $2 \leq k$.

Consider the perturbation scheme

$$L_{\mathbb{R}} I_i^k = K_i^k, \quad 3 \leq k \leq N. \quad (5.4)$$

where

$$L_{\mathbb{R}} I_i^k := \sum_{j=1}^n \left( \frac{\partial I_i^k}{\partial x_j} y_j - \frac{\partial I_i^k}{\partial y_j} x_j \right) \omega_j + \frac{\partial I_i^k}{\partial t}$$

and

$$K_i^3 := -x_i h_{1,i}^2 - y_i h_{2,i}^2,$$

$$K_i^k := -x_i h_{1,i}^{k-1} - y_i h_{2,i}^{k-1} - \sum_{l=3}^{k-1} \sum_{j=1}^n \left( \frac{\partial I_i^l}{\partial x_j} h_{1,i}^{k-l} + \frac{\partial I_i^l}{\partial y_j} h_{2,i}^{k-l} \right), \quad 4 \leq k \leq N.$$

Then the following holds:

(i) The perturbation scheme is solvable for $k \leq N$ (cf. Assumption $A_5$).

(ii) The polynomials $h_{1,i}^k, h_{2,i}^k, K_i^k$ and $I_i^k$ have the following symmetry properties:

$$h_{1,i}^k, h_{2,i}^k \in \Pi_{\mathbb{R},k}, \quad (5.12)$$

$$K_i^k \in \Pi_{\mathbb{R},k}^-, \quad (5.13)$$

$$I_i^k \in \Pi_{\mathbb{R},k}^+, \quad (5.14)$$

where

$$\Pi_{\mathbb{R},k} := \Pi_{\mathbb{R},k} = \left\{ f \in \Pi_k \mid f^{p,q,a}(t) \in \mathbb{R} \right\},$$
\[ \Pi_{i,k}^{i} := \left\{ (h_1, h_2) \in \Pi_{i,k} \times \Pi_{i,k} \ \middle| \ \begin{pmatrix} h_1(-t, R_Dz, \Delta \sigma) \\ h_2(-t, R_Dz, \Delta \sigma) \end{pmatrix} = -R_z \begin{pmatrix} h_1(t, z, \Delta \sigma) \\ h_2(t, z, \Delta \sigma) \end{pmatrix} \right\}, \]

\[ \Pi_{i,k}^{-} := T^{-1} \Pi_{i,k}^{i}, \]

\[ \Pi_{i,k}^{+} := T^{-1} \Pi_{i,k}^{i}. \]

**Proof:** The symmetry property (5.12) of the polynomials \( h_{i,j}^k \) follows from Assumption A3 and Theorem 4.5.1. Let \( I(m) \) be the following assertion:

\[ I(m) : \text{For } 3 \leq l \leq m \text{ the following holds:} \]

- \( K_i^l \in \Pi_{i,k}^{-} \),
- The solutions \( I_i^l \) of the perturbation equations \( L_{i,l} = K_i^l \) exist and are unique,
- \( I_i^l \in \Pi_{i,k}^{+} \).

We prove by induction that \( I(m) \) holds for every number \( m \geq 3 \).

- For \( m = 3 \) the right-hand side of the perturbation equations (5.4) is reduced to \( K_i^3 = -x_i h_{1,j}^2 - y_i h_{2,j}^2 \). This implies that the assumptions of Lemma 5.4.6 are fulfilled. It follows that \( K_i^3 \in \Pi_{i,k}^{-} \). Then Lemma 5.4.8 implies that \( I_i^3 \) uniquely exists and lies in \( \Pi_{i,k}^{+} \). Thus \( I(3) \) holds true.

- \( I(m - 1) \Rightarrow I(m) \): Let \( I(m - 1) \) hold true. This implies that the assumptions of Lemma 5.4.6 are fulfilled. It follows that \( K_i^m \in \Pi_{i,k}^{-} \). Then Lemma 5.4.8 implies that \( I_i^m \) uniquely exists and lies in \( \Pi_{i,k}^{+} \). Thus \( I(m) \) holds true.

By induction we conclude that \( I(m) \) holds for every number \( 3 \leq m \leq N \).

This theorem implies that the system of perturbed harmonic oscillators (5.4) has approximate first integrals

\[ I_i(t, x, y, \Delta \sigma) := I_i^3(x, y) + \sum_{k=3}^{\infty} I_i^k(t, x, y, \Delta \sigma), \]

where

\[ I_i^3 := \frac{1}{2} (x_i^2 + y_i^2). \]

We close this chapter with a theorem on the close relationship between the perturbation schemes in \( x,y \)- and \( \xi,\eta \)-coordinates.
Theorem 5.5.3
Let Assumptions $A_1$ – $A_5$ hold with the non-resonance condition NR$_N$.
Let $T$ denote the homeomorphism between $\Pi_{R,k}$ and $\Pi_{R,k}$:

$$T : \Pi_{R,k} \rightarrow \Pi_{R,k}, \ f \mapsto \tilde{f} := Tf \quad \text{with} \quad \tilde{f}(t, \xi, \eta, \Delta \sigma) := f(t, \tau(\xi, \eta), \Delta \sigma),$$

where

$$\tau : \mathbb{C} \rightarrow \mathbb{C}, \ (\xi, \eta) \mapsto \tau(\xi, \eta) := \left(\frac{1}{\sqrt{2}}(\xi + i\eta), \frac{1}{\sqrt{2}}(\xi - i\eta)\right).$$

Consider the perturbation schemes

$$L_R I_i^k = K_i^k \quad 3 \leq k \leq N \quad (5.4)$$

(cf. Theorem 5.5.1) and

$$L_R \bar{I}_i^k = \bar{K}_i^k \quad 3 \leq k \leq N \quad (5.6)$$

(cf. Theorem 5.5.2).
Then the following holds:

(i) The following diagram commutes:

$\begin{array}{ccc}
\Pi_{R,k} & \xrightarrow{L} & \Pi_{\mathbb{C},k} \\
T \downarrow & & \downarrow T \\
\Pi_{R,k} & \xrightarrow{L} & \Pi_{\mathbb{C},k}
\end{array}$

(ii) Let the polynomials $I_i^k$ be solutions of (5.4).
Then the polynomials $\bar{I}_i^k := TI_i^k$ solve the perturbation scheme (5.6).
Moreover for the right-hand sides one has

$$\bar{K}_i^k = TK_i^k.$$

(iii) Let the polynomials $\bar{I}_i^k$ be solutions of (5.6).
Then the polynomials $I_i^k := T^{-1}\bar{I}_i^k$ solve the perturbation scheme (5.4).
Moreover for the right-hand sides one has

$$K_i^k = T^{-1}\bar{K}_i^k.$$

Proof: The proof follows immediately from Lemma 5.2.3 and Lemma 5.3.6. \qed
Rigorous Estimates for First Integrals

Let Assumptions A₁–A₅ hold. Then Theorem 5.5.2 proves the existence of approximate first integrals in $x,y$-coordinates

$$I_i^{(r)}(t,x,y) := I_i^2(x,y) + \sum_{k=3}^{r} I_i^k(t,x,y,\Delta \sigma)$$

for Eqs. (5.4). In terms of the complex variables $\xi, \eta$ the approximate integrals are given by

$$\tilde{I}_i^{(r)}(t,\xi,\eta) := \tilde{I}_i^2(\xi,\eta) + \sum_{k=3}^{r} \tilde{I}_i^k(t,\xi,\eta,\Delta \sigma)$$

and are associated with Eqs. (5.6). Theorem 5.5.3 proves that the polynomials $I_i^k$ and $\tilde{I}_i^k$ are related via the homeomorphism $T$.

The goal of this chapter is to give estimates for the functions $I_i^{(r)}(t,x,y)$ and $\tilde{I}_i^{(r)}(t,x,y)$. Through this chapter we follow the ideas of Giorgilli (cf. [8]). In particular we adopt the norms introduced by Giorgilli for the spaces of polynomials.

The chapter is organized as follows:

Section 6.1: We introduce suitable norms.

Section 6.2: We give recursive estimates for the polynomials $I_i^k$.

Section 6.3: We derive a priori estimates for $I_i^k$.

Section 6.4: We derive a priori estimates for $\tilde{I}_i^k$.

Section 6.5: Combining the results obtained we derive estimates for the functions $I_i^{(r)}(t)$ and $\tilde{I}_i^{(r)}(t)$.

Section 6.6: Eventually we compute the optimal truncation order $r_{opt}$ for a priori estimates.
6.1 Definition of Norms

Consider the $2n$-dimensional system of differential equations

$$z = F(t,z,a), \quad z \in \mathbb{R}^{2n} \quad (3.1)$$

and assume that the vector space $\mathbb{R}^{2n}$ is provided with a certain norm denoted by $|.|$.

The Norm for the Transformed System

The complete transformation of coordinates is given by

$$(z, \sigma) \mapsto (z_4, \Delta \sigma) := \left( S^{-1}Q^{-1}(t) \left( z(t, \sigma^0 + \Delta \sigma) - z_0^0(t) \right) - z_0^0(t)(\sigma - \sigma^0) \right) \quad (4.18)$$

(cf. Theorem 4.5.1). It transforms (3.1) to

$$\begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t,x,y,\Delta \sigma) \\
\dot{y}_i &= - \omega_i x_i + h_{2,i}(t,x,y,\Delta \sigma)
\end{align*} \quad 1 \leq i \leq n, \quad (4.17)$$

where $z_4 =: (x_1, y_1, \ldots, x_n, y_n)^T$.

It will turn out that the following norms are suitable for our problem.

**Definition 6.1.1**

Let $R = (R_z; R_{\Delta \sigma}) := (R_1, \ldots, R_n; R_{n+1}, \ldots, R_{n+m})^T \in \mathbb{R}^n \times \mathbb{R}^m$. We define the following norms:

$$|z_4|_{R_z} := \max_{1 \leq i \leq n} \frac{1}{R_i} \left( x_i^2 + y_i^2 \right)^{1/2}, \quad \text{and} \quad |\Delta \sigma|_{R_{\Delta \sigma}} := \max_{1 \leq j \leq m} \frac{1}{R_{n+j}} |\Delta \sigma_j|.$$

The Norms of the Transformation Matrices $S$ and $S^{-1}$

Consider $z = Sz'$. The norms $||.||_{R_z^{-1}}$ and $||.||_{R_z}$ of the matrices $S$ and $S^{-1}$ has to be choosen compatible with the norms $|.|$ and $||.|_{R_z}$, i.e. they must fulfill the following inequalities:

$$|z| \leq ||S||_{R_z^{-1}} \cdot |z'|_{R_z}, \quad |z'|_{R_z} \leq ||S^{-1}||_{R_z} \cdot |z|.$$
Lemma 6.1.1
Assume that in the original coordinates the norm is given by

\[ |z| := |(z_1, \ldots, z_n)| := \max_{1 \leq i \leq n} |z_i|_2, \quad \text{where} \quad |z_i|_2 := |(x_i, y_i)|_2 := (x_i^2 + y_i^2)^{\frac{1}{2}}. \]

Then the following norms for \( S \) and \( S^{-1} \) are compatible:

\[
\|S\|_{R_e^{-1}} := \left\| \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix} \right\| := \max_{1 \leq i \leq n} \sum_{j=1}^{n} R_j \|S_{ij}\|_2,
\]

and

\[
\|S^{-1}\|_{R_e} := \left\| \begin{pmatrix} (S^{-1})_{11} & \cdots & (S^{-1})_{1n} \\ \vdots & \ddots & \vdots \\ (S^{-1})_{n1} & \cdots & (S^{-1})_{nn} \end{pmatrix} \right\| := \max_{1 \leq i \leq n} \sum_{j=1}^{n} \|(S^{-1})_{ij}\|_2,
\]

where

\[
\|S_{ij}\|_2 := \left\| \begin{pmatrix} s_{ij}^{11} & s_{ij}^{12} \\ s_{ij}^{21} & s_{ij}^{22} \end{pmatrix} \right\|_2 := \left( (s_{ij}^{11})^2 + (s_{ij}^{12})^2 + (s_{ij}^{21})^2 + (s_{ij}^{22})^2 \right)^{\frac{1}{2}}.
\]

Proof: We first show that

\[ |S_{ij}z_j|_2 \leq \|S_{ij}\|_2 \cdot |z_j|_2. \]

Let

\[ S_{ij} := \begin{pmatrix} s_{ij}^{11} & s_{ij}^{12} \\ s_{ij}^{21} & s_{ij}^{22} \end{pmatrix} \quad \text{and} \quad z_j := \begin{pmatrix} x \\ y \end{pmatrix}. \]

Then we have

\[
\|S_{ij}\|_2^2 \cdot |z_j|_2^2 - |S_{ij}z_j|_2^2 = (s_{ij}^{11} + s_{ij}^{12} + s_{ij}^{21} + s_{ij}^{22})(x^2 + y^2) - \\
- ((s_{ij}^{11}x + s_{ij}^{12}y)^2 + (s_{ij}^{21}x + s_{ij}^{22}y)^2) = \\
s_{ij}^{11}x^2 + s_{ij}^{12}y^2 + s_{ij}^{21}x^2 + s_{ij}^{22}y^2 + s_{ij}^{21}x^2 + s_{ij}^{22}y^2 + s_{ij}^{22}x^2 + s_{ij}^{22}y^2 - \\
- (s_{ij}^{12}x^2 + 2s_{ij}^{12}xy + s_{ij}^{22}y^2 + s_{ij}^{21}x^2 + 2s_{ij}^{21}xy + s_{ij}^{22}x^2) = \\
= (s_{ij}^{12}y - s_{ij}^{12}x)^2 + (s_{ij}^{22}y - s_{ij}^{22}x)^2 \geq 0.
\]

We conclude that \( |S_{ij}z_j|_2 \leq \|S_{ij}\|_2 \cdot |z_j|_2. \)
Now we prove the compatibility of the norm for the matrix $S$:

\[
|Sz'| = \left| \begin{pmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & \ddots & \vdots \\ S_{n1} & \cdots & S_{nn} \end{pmatrix} \begin{pmatrix} z'_1 \\ \vdots \\ z'_n \end{pmatrix} \right| = \left| \begin{pmatrix} \sum_{j=1}^{n} S_{ij} z'_j \\ \vdots \\ \sum_{j=1}^{n} S_{nj} z'_j \end{pmatrix} \right| \\
= \max_{1 \leq i \leq n} \left| \sum_{j=1}^{n} S_{ij} z'_j \right|_2 \leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} \|S_{ij}\|_2 \cdot |z'_j|_2 \\
\leq \max_{1 \leq i \leq n} \sum_{j=1}^{n} R_j \|S_{ij}\|_2 \cdot \max_{1 \leq j \leq n} \frac{1}{R_j} |z'_j|_2 = \\
\|S\|_{R_i^{-1}} |z'|_{R_i}.
\]

Finally we prove the compatibility of the norm for the matrix $S^{-1}$:

\[
|S^{-1}z|_{R_i} = \left| \begin{pmatrix} (S^{-1})_{11} & \cdots & (S^{-1})_{1n} \\ \vdots & \ddots & \vdots \\ (S^{-1})_{n1} & \cdots & (S^{-1})_{nn} \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \right| = \left| \begin{pmatrix} \sum_{j=1}^{n} (S^{-1})_{ij} z_j \\ \vdots \\ \sum_{j=1}^{n} (S^{-1})_{nj} z_j \end{pmatrix} \right|_{R_i} \\
\leq \max_{1 \leq i \leq n} \frac{1}{R_i} \sum_{j=1}^{n} \| (S^{-1})_{ij} z_j \|_2 \leq \max_{1 \leq i \leq n} \frac{1}{R_i} \sum_{j=1}^{n} \| (S^{-1})_{ij} \|_2 \cdot |z_j|_2 \\
\leq \max_{1 \leq i \leq n} \frac{1}{R_i} \sum_{j=1}^{n} \| (S^{-1})_{ij} \|_2 \cdot \max_{1 \leq j \leq n} |z_i|_2 = \\
\|S^{-1}\|_{R_i} \cdot |z|.
\]

This completes the proof. \(\square\)

The Norm of the Matrix $z_3^0(0)$

The norm of $z_3^0(0) \in L(\mathbb{R}^m, \mathbb{R}^{2n})$ must be chosen compatible with the norms $|.|_{R_i}$ and $|.|_{R_{\Delta^*}}$, i.e.

\[
|z_3^0(0)\Delta^*|_{R_i} \leq \|z_3^0(0)\|_{\mathbb{R}_i / \mathbb{R}_{\Delta^*}} \cdot |\Delta^*|_{\mathbb{R}_{\Delta^*}}.
\]

The next lemma gives an example.

**Lemma 6.1.2**

Consider the example for the norm $|.|$ introduced in the previous lemma. Then the following norm for $B := z_3^0(0)$ is compatible:
6.1. Definition of Norms

\[
\|B\|_{R_z/R_{\Delta \sigma}} := \left\| \begin{pmatrix} B_{11} & \ldots & B_{1m} \\ \vdots & \ddots & \vdots \\ B_{n1} & \ldots & B_{nm} \end{pmatrix} \right\|_{R_z/R_{\Delta \sigma}} := \max_{1 \leq i \leq n} \sum_{j=1}^{m} \frac{R_{n+j}}{R_i} |B_{ij}|_2, 
\]

where

\[
|B_{ij}|_2 := \left\| \begin{pmatrix} b_{ij}^1 \\ b_{ij}^2 \end{pmatrix} \right\|_2 := \sqrt{(b_{ij}^1)^2 + (b_{ij}^2)^2}.
\]

**Proof:** The proof is straight forward:

\[
B \Delta \sigma = \begin{pmatrix} B_{11} & \ldots & B_{1m} \\ \vdots & \ddots & \vdots \\ B_{n1} & \ldots & B_{nm} \end{pmatrix} \begin{pmatrix} \Delta \sigma_1 \\ \vdots \\ \Delta \sigma_m \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{m} B_{ij} \Delta \sigma_j \\ \vdots \\ \sum_{j=1}^{m} B_{nj} \Delta \sigma_j \end{pmatrix},
\]

where

\[
B_{ij} \Delta \sigma_j = \begin{pmatrix} b_{ij}^1 \\ b_{ij}^2 \end{pmatrix} \Delta \sigma_j = \begin{pmatrix} b_{ij}^1 \Delta \sigma_j \\ b_{ij}^2 \Delta \sigma_j \end{pmatrix}.
\]

Thus we have

\[
|B \Delta \sigma|_{R_z} = \max_{1 \leq i \leq n} \frac{1}{R_i} \left| \sum_{j=1}^{m} B_{ij} \Delta \sigma_j \right|_2 \leq \max_{1 \leq i \leq n} \frac{1}{R_i} \sum_{j=1}^{m} |B_{ij} \Delta \sigma_j|_2 \leq \max_{1 \leq i \leq n} \sum_{j=1}^{m} \frac{R_{n+j}}{R_i} |B_{ij}|_2 \cdot \max_{1 \leq j \leq m} \frac{1}{R_{n+j}} |\Delta \sigma_j| =
\]

\[
= \|B\|_{R_z/R_{\Delta \sigma}} \cdot |\Delta \sigma|_{R_{\Delta \sigma}}.
\]

This completes the proof.

**The Norm for the T-Periodic Coefficients**

Before we define a norm for the spaces of polynomials we must define an norm for the T-periodic coefficients.

**Definition 6.1.2**

On the space

\[
\mathcal{F} := \left\{ f \in C^\infty(\mathbb{R}, \mathbb{C}) \mid f(t + T) = f(t), T := \frac{2\pi}{\omega_0} \right\}
\]

of smooth T-periodic function we define the following norm:

\[
\|f\|_{\mathcal{F}} := \left\| \sum_{k_0 = -\infty}^{\infty} f_{k_0} e^{ik_0 \omega_0 t} \right\|_{\mathcal{F}} = \sum_{k_0 = -\infty}^{\infty} |f_{k_0}|.
\]
The Norm for the Polynomials

Now we are in the position to define a suitable norm for the spaces of polynomials of order $k$.

**Lemma 6.1.3**

On the spaces $\Pi_k$ and $\overline{\Pi}_k$ of polynomial of order $k$ let

$$
\|f\|_R := \sum_{|p|_1+|q|_1+|s|_1=k} \|f_{p,q,s}\|_F R^{(p+q,s)} = \sum_{|p|_1+|q|_1+|s|_1=k} \sum_{k_0=-\infty}^{\infty} |f_{p,q,s,k_0}| R^{(p+q,s)},
$$

where

$$
R^{(p+q,s)} := \prod_{i=1}^{n} R^{p_{i}+q_{i}}_{i} \cdot \prod_{j=1}^{m} R^{s_{j}}_{n+j}.
$$

Then the following holds:

(i) $(F, \|\cdot\|_F)$ is a normed algebra.

(ii) $(\Pi_k, \|\cdot\|_R)$ and $(\overline{\Pi}_k, \|\cdot\|_R)$ are normed spaces.

(iii) Let $f \in \Pi_k$, $g \in \Pi_k$. Then $f \cdot g \in \Pi_{k+k'}$ and $\|f \cdot g\|_R \leq \|f\|_R \cdot \|g\|_R$.

Let $\tilde{f} \in \overline{\Pi}_k$, $\tilde{g} \in \overline{\Pi}_{k'}$. Then $\tilde{f} \cdot \tilde{g} \in \overline{\Pi}_{k+k'}$ and $\|\tilde{f} \cdot \tilde{g}\|_R \leq \|\tilde{f}\|_R \cdot \|\tilde{g}\|_R$.

**Proof:**

(i) Let $f$ be $m$-times continuously differentiable ($m \geq 2$). Then the coefficients $f^{k_0}$ of the Fourier expansion are of order $O(k_0^{-m})$ (cf. [3]). Thus for $m \geq 2$ the series $\sum_{k_0=-\infty}^{\infty} |f^{k_0}|$ on the right-hand side of the definition converges.

The properties of the norm are checked easily. It remains to show that $\|f \cdot g\|_F \leq \|f\|_F \cdot \|g\|_F$:

$$
\|f \cdot g\|_F = \left\| \sum_{k_0=-\infty}^{\infty} f^{k_0} e^{i k_0 \omega_0 t} \cdot \sum_{l_0=-\infty}^{\infty} g^{l_0} e^{i l_0 \omega_0 t} \right\|_F =
$$

$$
= \left\| \sum_{k_0=-\infty}^{\infty} \sum_{l_0=-\infty}^{\infty} f^{k_0} g^{l_0} e^{i (k_0+l_0) \omega_0 t} \right\|_F = \left\| \sum_{k_0=-\infty}^{\infty} \sum_{l_0=-\infty}^{\infty} f^{k_0-l_0} g^{l_0} e^{i k_0 \omega_0 t} \right\|_F =
$$

$$
= \sum_{k_0=-\infty}^{\infty} \sum_{l_0=-\infty}^{\infty} |f^{k_0-l_0} g^{l_0}| \leq \sum_{k_0=-\infty}^{\infty} \sum_{l_0=-\infty}^{\infty} |f^{k_0} | \cdot |g^{l_0}| =
$$

$$
= \sum_{k_0=-\infty}^{\infty} \sum_{l_0=-\infty}^{\infty} |f^{k_0}| \cdot |g^{l_0}| = \sum_{k_0=-\infty}^{\infty} |f^{k_0}| \cdot \sum_{l_0=-\infty}^{\infty} |g^{l_0}| = \|f\|_F \cdot \|g\|_F.
$$
(ii) Let \( f(t, x, y, \Delta \sigma) := \sum_{|p|_1 + |q|_1 + |s|_1 = k} f^{p,q,s}(t) x^p y^q \Delta \sigma^s \in \Pi_k \). Since \( R_i > 0 \) we have \( \|f\|_R \geq 0 \) and

\[
\|f\|_R = 0 \iff \|f^{p,q,s}\|_\mathcal{F} = 0 \text{ for all } p, q, s \text{ with } |p|_1 + |q|_1 + |s|_1 = k
\]

\(
\iff f(t, x, y, \Delta \sigma) = 0 \text{ for all } t, x, y, \Delta \sigma.
\)

Let \( \alpha \in \mathbb{R} \). Then

\[
\|\alpha \cdot f\|_R = \sum_{|p|_1 + |q|_1 + |s|_1 = k} ||\alpha \cdot f^{p,q,s}||_\mathcal{F} R^{(p+q,s)}(p+q,s) = ||\alpha|| \cdot \sum_{|p|_1 + |q|_1 + |s|_1 = k} ||f^{p,q,s}||_\mathcal{F} R^{(p+q,s)}(p+q,s) = ||\alpha|| \cdot \|f\|_R.
\]

Let \( g(x, y, \Delta \sigma) := \sum_{|p|_1 + |q|_1 + |s|_1 = k} g^{p,q,s}(t) x^p y^q \Delta \sigma^s \in \Pi_k \). Then we have

\[
\|f + g\|_R = \sum_{|p|_1 + |q|_1 + |s|_1 = k} ||f^{p,q,s} + g^{p,q,s}||_\mathcal{F} R^{(p+q,s)}(p+q,s) \leq \sum_{|p|_1 + |q|_1 + |s|_1 = k} (||f^{p,q,s}||_\mathcal{F} + ||g^{p,q,s}||_\mathcal{F}) R^{(p+q,s)}(p+q,s) = \sum_{|p|_1 + |q|_1 + |s|_1 = k} ||f^{p,q,s}||_\mathcal{F} R^{(p+q,s)} + \sum_{|p|_1 + |q|_1 + |s|_1 = k} ||g^{p,q,s}||_\mathcal{F} R^{(p+q,s)} = \|f\|_R + \|g\|_R.
\]

Thus \( \|\cdot\|_R \) is a norm indeed.

(iii) Let \( f \in \Pi_k, g \in \Pi_{k'} \). Obviously we have \( f \cdot g \in \Pi_{k+k'} \). It remains to show that

\[
\|f \cdot g\|_R \leq \|f\|_R \cdot \|g\|_R:
\]

\[
\|f \cdot g\|_R = \left\| \sum_{|p|_1 + |q|_1 + |s|_1 = k} f^{p,q,s}(t) x^p y^q \Delta \sigma^s \cdot \sum_{|p'|_1 + |q'|_1 + |s'|_1 = k'} g^{p,q,s}(t) x^{p'} y^{q'} \Delta \sigma^{s'} \right\|_R \leq \sum_{|p|_1 + |q|_1 + |s|_1 = k} \sum_{|p'|_1 + |q'|_1 + |s'|_1 = k'} ||f^{p,q,s}(t) g^{p',q,s'}(t)||_\mathcal{F} R^{(p+q+q'+s+s')}(p+q,s+s') \leq \sum_{|p|_1 + |q|_1 + |s|_1 = k} \sum_{|p'|_1 + |q'|_1 + |s'|_1 = k'} ||f^{p,q,s}||_\mathcal{F} R^{(p+q,s)}(p+q,s) \cdot \sum_{|p'|_1 + |q'|_1 + |s'|_1 = k'} ||g^{p,q,s}||_\mathcal{F} R^{(p'+q',s')}(p'+q',s') = \|f\|_R \cdot \|g\|_R \]

The next lemma provides an estimate for \( f(t, x, y, \Delta \sigma) \).
Lemma 6.1.4
(i) For $f \in F$ we have $|f(t)| \leq \|f\|_{\mathcal{F}}$ for all $t$.
(ii) For $|(x, y)|_{R_{\sigma}} \leq \rho$, $|\Delta \sigma|_{R_{\Delta \sigma}} \leq \rho$ and $f(t, x, y, \Delta \sigma) \in \Pi_k$ we have
$$|f(t, x, y, \Delta \sigma)| \leq \|f\|_{R} \rho^k \text{ for all } t \in \mathbb{R}.$$

Proof:
(i) Let $f(t) := \sum_{k_0=-\infty}^{\infty} f^{k_0} e^{ik_0 \omega_0 t}$. Then
$$|f(t)| \leq \sum_{k_0=-\infty}^{\infty} |f^{k_0} e^{ik_0 \omega_0 t}| \leq \sum_{k_0=-\infty}^{\infty} |f^{k_0}| = \|f\|_{\mathcal{F}}.$$

(ii) Let $f(t, x, y, \Delta \sigma) := \sum_{|p|+|q|+|s|} f^{p,q,s}(t) x^p y^q \Delta \sigma^s$. Then we have
$$|f(t, x, y, \Delta \sigma)| = \left| \sum_{|p|+|q|+|s|} f^{p,q,s}(t) x^p y^q \Delta \sigma^s \right| \leq \sum_{|p|+|q|+|s|} |f^{p,q,s}(t)| x^p y^q \Delta \sigma^s \leq \sum_{|p|+|q|+|s|} \|f^{p,q,s}\|_{\mathcal{F}} |x^p| |y^q| |\Delta \sigma^s| \leq \sum_{|p|+|q|+|s|} \|f^{p,q,s}\|_{\mathcal{F}} (\rho R)^{|p|+|q|+|s|} = \rho^k \cdot \sum_{|p|+|q|+|s|} \|f^{p,q,s}\|_{\mathcal{F}} = \rho^k \|f\|_{R}.$$

We complete this section with a lemma on the the change of coordinates.

Lemma 6.1.5
Let $f \in \Pi_k$ and let $\bar{f} \equiv Tf \in \overline{\Pi}_k$. The following estimates hold:
$$\|\bar{f}\|_{R} \leq 2^\frac{k}{2} \|f\|_{R} \text{ and } \|f\|_{R} \leq 2^\frac{k}{2} \|\bar{f}\|_{R}.$$
6.2 Recursive Estimates for $I^k_i$

**Proof:** Let \( f(t,x,y,\Delta \sigma) = \sum_{|p|+|q|+|s|} f^{p,q,s}(t) x^p y^q \Delta \sigma^s \). Then we have

\[
\tilde{f}(t,\xi,\eta,\Delta \sigma) = \sum_{|p|+|q|+|s|} f^{p,q,s}(t) \left( \frac{1}{\sqrt{2}} (\xi + i\eta) \right)^p \left( \frac{i}{\sqrt{2}} (\xi - i\eta) \right)^q \Delta \sigma^s =
\]

\[
= \sum_{|p|+|q|+|s|} f^{p,q,s}(t) 2^{-\frac{|q|}{2}} (-i^{q-m}) (\xi + i\eta)^p (\xi - i\eta)^q \Delta \sigma^s =
\]

\[
= \sum_{|p|+|q|+|s|} f^{p,q,s}(t) 2^{-\frac{k-|s|}{2}} (-i^{q-m}) \Delta \sigma^s.
\]

And we obtain for the norm

\[
\|\tilde{f}\|_R \leq \sum_{|p|+|q|+|s|} \|f^{p,q,s}\|_R 2^{-\frac{k-|s|}{2}} \prod_{j=1}^n \left( \sum_{l=0}^{p_j} \sum_{m=0}^{q_j} \left( \frac{p_j}{l} \frac{q_j}{m} \right) R^{p_j+q_j} \right) R^{0,s} =
\]

\[
= \sum_{|p|+|q|+|s|} \|f^{p,q,s}\|_R 2^{-\frac{k-|s|}{2}} R^{p+q,s}
\]

\[
\leq 2^{\frac{1}{2}} \|f\|_R.
\]

The second inequality is proved in the same way. \(\square\)

6.2 Recursive Estimates for $I^k_i$

In this section we give recursive estimates for the integrals in $\xi\eta$-coordinates and $x\eta$-coordinates as well. We start with estimates for the differential operators on the right-hand side of (5.4) and (5.6), respectively.
Lemma 6.2.1
For $1 \leq i \leq n$ and $4 \leq k \leq N$ let

$$\tilde{I}^i \in \Pi^{i+}_{R,l},$$

$$\tilde{h}^{k-l+1}_{1,i}, \tilde{h}^{k-l+1}_{2,i} \in \Pi_{R,k-l+1}.$$

Then for $3 \leq l \leq k - 1$ we have

$$\left\| \sum_{j=1}^{n} \left( \left( \frac{\partial \tilde{I}^i}{\partial \xi_j} - i \frac{\partial \tilde{I}^i}{\partial \eta_j} \right) \tilde{h}^{k-l+1}_{1,j} - i \left( \frac{\partial \tilde{I}^i}{\partial \xi_j} + i \frac{\partial \tilde{I}^i}{\partial \eta_j} \right) \tilde{h}^{k-l+1}_{2,j} \right) \right\|_R \leq$$

$$\leq l |R_z^{-1}|_\infty \| \tilde{I}^i \|_{R,1 \leq j \leq n} \max \left( \| \tilde{h}^{k-l+1}_{1,j} \|_R + \| \tilde{h}^{k-l+1}_{2,j} \|_R \right),$$

where $|R_z^{-1}|_\infty := \max_{1 \leq i \leq n} R_i^{-1}$.

Proof: For $I^i(t, \xi, \eta, \Delta \sigma) = \sum_{|p| + |q| + |s| = l} \tilde{I}^{p,q,s}_i(t) \xi^p \eta^q \Delta \sigma^s$ we obtain

$$\left\| \sum_{j=1}^{n} \left( \left( \frac{\partial \tilde{I}^i}{\partial \xi_j} - i \frac{\partial \tilde{I}^i}{\partial \eta_j} \right) \tilde{h}^{k-l+1}_{1,j} - i \left( \frac{\partial \tilde{I}^i}{\partial \xi_j} + i \frac{\partial \tilde{I}^i}{\partial \eta_j} \right) \tilde{h}^{k-l+1}_{2,j} \right) \right\|_R =$$

$$= \left\| \sum_{j=1}^{n} \frac{\partial \tilde{I}^i}{\partial \xi_j} \left( \tilde{h}^{k-l+1}_{1,j} - i \tilde{h}^{k-l+1}_{2,j} \right) - \frac{i \partial \tilde{I}^i}{\partial \eta_j} \left( \tilde{h}^{k-l+1}_{1,j} + i \tilde{h}^{k-l+1}_{2,j} \right) \right\|_R =$$

$$= \left\| \sum_{|p| + |q| + |s| = l} \tilde{I}^{p,q,s}_i(t) \sum_{j=1}^{n} \left( \frac{\xi^p \eta^q}{\xi_j} \Delta \sigma^s \left( \tilde{h}^{k-l+1}_{1,j} - i \tilde{h}^{k-l+1}_{2,j} \right) - i q_j \frac{\xi^p \eta^q}{\xi_j} \Delta \sigma^s \left( \tilde{h}^{k-l+1}_{1,j} + i \tilde{h}^{k-l+1}_{2,j} \right) \right) \right\|_R \leq$$

$$\leq \sum_{|p| + |q| + |s| = l} \| \tilde{I}^{p,q,s}_i \|_F \sum_{j=1}^{n} (p_j + q_j) \frac{R^{p+q,s}}{R_j} \left( \| \tilde{h}^{k-l+1}_{1,j} \|_R + \| \tilde{h}^{k-l+1}_{2,j} \|_R \right) \leq$$

$$\leq l |R_z^{-1}|_\infty \sum_{|p| + |q| + |s| = l} \| \tilde{I}^{p,q,s}_i \|_F R^{p+q,s} \max_{1 \leq j \leq n} \left( \| \tilde{h}^{k-l+1}_{1,j} \|_R + \| \tilde{h}^{k-l+1}_{2,j} \|_R \right) \leq$$

$$\leq l |R_z^{-1}|_\infty \| \tilde{I}^i \|_R \max_{1 \leq j \leq n} \left( \| \tilde{h}^{k-l+1}_{1,j} \|_R + \| \tilde{h}^{k-l+1}_{2,j} \|_R \right).$$ \(\square\)
Lemma 6.2.2
For $1 \leq i \leq n$ and $4 \leq k \leq N$ let
\[ I_i^k \in \Pi_{R,j}^{k+}, \]
\[ h_1^{k-1+1}, h_2^{k-1+1} \in \Pi_{R,k-l+1}. \]
Then for $3 \leq l \leq k-1$ we have
\[
\left\| \sum_{j=1}^{n} \left( \frac{\partial I_i^k}{\partial x_j} h_1^{k-1+1} + \frac{\partial I_i^k}{\partial y_j} h_2^{k-1+1} \right) \right\|_R \leq \frac{l|R_{z}^{-1}\infty \|I_i^k\|_R \max_{1 \leq j \leq n} (\|h_1^{k-1+1}\|_R + \|h_2^{k-1+1}\|_R)}. \]

Proof: For $I_i^k(t,x,y,\Delta \sigma) = \sum_{|p|+|q|+|s|=l} I_i^{p,q,s}(t)x^p y^q \Delta \sigma^s$ we obtain
\[
\left\| \sum_{j=1}^{n} \left( \frac{\partial I_i^k}{\partial x_j} h_1^{k-1+1} + \frac{\partial I_i^k}{\partial y_j} h_2^{k-1+1} \right) \right\|_R \leq \frac{l|R_{z}^{-1}\infty \|I_i^k\|_R \max_{1 \leq j \leq n} (\|h_1^{k-1+1}\|_R + \|h_2^{k-1+1}\|_R)}. \]

Now we give an estimate for the inverse of the operator $L_R$ (cf. Lemma 5.3.4).

Lemma 6.2.3
Assume that the non-resonance condition $\text{NR}_N$ holds. For $1 \leq i \leq n$ and $3 \leq k \leq N$ let $K_i^k \in \bar{\Pi}_{R,k}^{+} \subset \bar{\Pi}_{R,k}^{NC}$. Then $\bar{I}_i^k := L_R^{-1} K_i^k$ exists and lies in $\bar{\Pi}_{R,k}^{+} \subset \bar{\Pi}_{R,k}^{NC}$ and the following estimate holds:
\[
\|\bar{I}_i^k\|_R \leq \frac{1}{\alpha_k} \|K_i^k\|_R, \]
where the sequence $(\alpha_k)_{3 \leq k \leq n}$ is introduced in Lemma 3.2.3.
Proof: Let
\[ I_i^k(t, \xi, \eta, \Delta \sigma) := \sum_{|p| + |q| + |s| = k} \sum_{k_0 = -\infty}^{\infty} \tilde{f}^{p,q,s,k_0} e^{i k \omega_0 t} \xi^p \eta^q \Delta \sigma^s, \]
\[ K_i^k(t, \xi, \eta, \Delta \sigma) := \sum_{|p| + |q| + |s| = k} \sum_{k_0 = -\infty}^{\infty} \tilde{K}^{p,q,s,k_0} e^{i k \omega_0 t} \xi^p \eta^q \Delta \sigma^s. \]
From \( \tilde{I}_i \tilde{I}_i = \tilde{K}_i \) and the definition of the operator \( \tilde{L}_i \) we obtain (cf. Lemma 5.3.2 (iv))
\[ \sum_{|p| + |q| + |s| = k} \sum_{k_0 = -\infty}^{\infty} \tilde{K}^{p,q,s,k_0} e^{i k \omega_0 t} \xi^p \eta^q \Delta \sigma^s = \]
\[ = \sum_{|p| + |q| + |s| = k} \sum_{k_0 = -\infty}^{\infty} i ((p - q) \cdot \omega + k_0 \omega_0) \tilde{I}_i^{p,q,s,k_0} e^{i k \omega_0 t} \xi^p \eta^q \Delta \sigma^s. \]
Since \( \tilde{K}_i \) lies in the space \( \Pi_k^{nc} \) of non-critical polynomials, we have \( (p - q) \cdot \omega + k_0 \omega_0 \neq 0 \) and Lemma 3.2.3 may be applied:
\[ \| \tilde{K}_i \|_R = \sum_{|p| + |q| + |s| = k} \sum_{k_0 = -\infty}^{\infty} |\tilde{K}^{p,q,s,k_0}| R^{(p + q, s)} \]
\[ = \sum_{|p| + |q| + |s| = k} \sum_{k_0 = -\infty}^{\infty} |((p - q) \cdot \omega + k_0 \omega_0) \tilde{I}_i^{p,q,s,k_0}| R^{(p + q, s)} \geq \alpha_k \sum_{|p| + |q| + |s| = k} \sum_{k_0 = -\infty}^{\infty} |\tilde{I}_i^{p,q,s,k_0}| R^{(p + q, s)} = \alpha_k \| \tilde{I}_i \|_R. \]
We are now in the position to give estimates for the right-hand sides of (5.6):
\[ \tilde{K}_i^3 := \frac{1}{\sqrt{2}} \left( (\xi_i + i \eta_i) \tilde{h}_{1,i}^2 + i (\xi_i - i \eta_i) \tilde{h}_{2,i}^2 \right), \]
\[ \tilde{K}_i^k := \frac{1}{\sqrt{2}} \left( (\xi_i + i \eta_i) \tilde{h}_{1,i}^{k-1} + i (\xi_i - i \eta_i) \tilde{h}_{2,i}^{k-1} \right) - \frac{1}{\sqrt{2}} \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial \tilde{F}_j}{\partial \xi_j} - i \frac{\partial \tilde{R}_j}{\partial \eta_j} \right) \tilde{h}_{1,j}^{k-l+1} - i \left( \frac{\partial \tilde{F}_j}{\partial \xi_j} + i \frac{\partial \tilde{R}_j}{\partial \eta_j} \right) \tilde{h}_{2,j}^{k-l+1}. \]

Lemma 6.2.4
Assume that the non-resonance condition \( \text{NR}_N \) holds. For \( 1 \leq i \leq n \) and \( k \leq N \) let
\[ \tilde{I}_i \in \Pi_{R,i}^{nc} \subset \Pi_{R,i}^{nc}, \]
\[ (\tilde{h}_{1,j}^{k-l+1}, \tilde{h}_{2,j}^{k-l+1}) \in \Pi_{R,k-l+1}. \]
6.2. Recursive Estimates for $I^k_i$

Then $\tilde{K}^k_i \in \tilde{\Pi}^k_{\mathcal{I},k} \subset \tilde{\Pi}^n_{\mathcal{I},k}$ and the following estimates holds:

$$\|\tilde{K}^3_i\|_R \leq \sqrt{2}R_i \left(\|\tilde{h}_{1,i}\|_R + \|\tilde{h}_{2,i}\|_R\right),$$

$$\|\tilde{K}^k_i\|_R \leq \sqrt{2}R_i \left(\|\tilde{h}_{1,i}^{k-1}\|_R + \|\tilde{h}_{2,i}^{k-1}\|_R\right) +$$

$$+ \frac{|R_i|_1}{\sqrt{2}} \sum_{l=3}^{k-1} \left\|I^l_i\right\|_R \max_{1 \leq j \leq n} \left(\|\tilde{h}_{1,j}^{k-l+1}\|_R + \|\tilde{h}_{2,j}^{k-l+1}\|_R\right), \quad k > 3.$$

Proof: Using Lemma 6.1.3 we obtain

$$\|\tilde{K}^3_i\|_R \leq \frac{1}{\sqrt{2}} \|\left(\xi_i + \eta_k\right)\tilde{h}_{1,i}\|_R + \frac{1}{\sqrt{2}} \|\left(\xi_i - \eta_k\right)\tilde{h}_{2,i}\|_R \leq$$

$$\leq \frac{1}{\sqrt{2}} \|\left(\xi_i + \eta_k\right)\|_R \|\tilde{h}_{1,i}\|_R + \frac{1}{\sqrt{2}} \|\left(\xi_i - \eta_k\right)\|_R \|\tilde{h}_{2,i}\|_R \leq$$

$$\leq 2 \frac{1}{\sqrt{2}} R_i \left(\|\tilde{h}_{1,i}\|_R + \|\tilde{h}_{2,i}\|_R\right).$$

Using Lemma 6.2.1 and Lemma 6.2.3 we obtain

$$\|\tilde{K}^k_i\|_R \leq \frac{1}{\sqrt{2}} \|\left(\xi_i + \eta_k\right)\tilde{h}_{1,i}^{k-1}\|_R + \frac{1}{\sqrt{2}} \|\left(\xi_i - \eta_k\right)\tilde{h}_{2,i}^{k-1}\|_R +$$

$$+ \frac{1}{\sqrt{2}} \sum_{l=3}^{k-1} |R_i|_1 \left\|I^l_i\right\|_R \max_{1 \leq j \leq n} \left(\|\tilde{h}_{1,j}^{k-l+1}\|_R + \|\tilde{h}_{2,j}^{k-l+1}\|_R\right) \leq$$

$$\leq \sqrt{2}R_i \left(\|\tilde{h}_{1,i}^{k-1}\|_R + \|\tilde{h}_{2,i}^{k-1}\|_R\right) +$$

$$+ \frac{|R_i|_1}{\sqrt{2}} \sum_{l=3}^{k-1} \left\|I^l_i\right\|_R \max_{1 \leq j \leq n} \left(\|\tilde{h}_{1,j}^{k-l+1}\|_R + \|\tilde{h}_{2,j}^{k-l+1}\|_R\right).$$

From the estimates for $\|\tilde{K}^k_i\|_R$ we easily derive estimates for $\|\tilde{I}^k_i\|_R$.

Lemma 6.2.5

Assume that the non-resonance condition $\text{NR}_N$ holds. For $1 \leq i \leq n$ we define

$$\tilde{B}_{i,3} := \sqrt{2}R_i \left(\|\tilde{h}_{1,i}^2\|_R + \|\tilde{h}_{2,i}^2\|_R\right),$$

$$\tilde{B}_{i,k} := \sqrt{2}R_i \left(\|\tilde{h}_{1,i}^{k-1}\|_R + \|\tilde{h}_{2,i}^{k-1}\|_R\right) +$$
where the sequence \((\alpha_k)_{1 \leq k \leq N}\) is introduced in Lemma 3.2.3.

Then we have

\[
\|K_i^k\|_R < B_{i,k}, \quad \text{for } 3 \leq k \leq N
\]

and

\[
\|I_i^k\|_R < \frac{1}{\alpha_k} B_{i,k}, \quad \text{for } 3 \leq k \leq N.
\]

**Proof:** These bounds follow by induction with respect to \(k\) from Lemma 6.2.3 and Lemma 6.2.4.

Now we return to \(x-y\)-coordinates, i.e. we consider equation (5.4):

\[
L_R I_i^k = K_i^k,
\]

with

\[
K_i^3 := -x_i h_{1,i}^2 - y_i h_{2,i}^2,
\]

\[
K_i^k := -x_i h_{1,i}^{k-1} - y_i h_{2,i}^{k-1} - \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial h_{i,j}^l}{\partial x_j} h_{1,i}^{k-l+1} + \frac{\partial h_{2,j}^l}{\partial y_j} h_{2,i}^{k-l+1} \right).
\]

**Lemma 6.2.6**

Assume that the non-resonance condition \(\text{NR}_N\) holds. For \(3 \leq i \leq n\) we define

\[
B_{i,3} := 2^3 R_i \left( \|h_{1,i}^2\|_R + \|h_{2,i}^2\|_R \right),
\]

\[
B_{i,k} := 2^k R_i \left( \|h_{1,i}^{k-1}\|_R + \|h_{2,i}^{k-1}\|_R \right) +
\]

\[
|R_i^{-1}| \sum_{l=3}^{k-1} 2^{k-l-1} \frac{1}{\alpha_l} B_{i,l} \max_{1 \leq j \leq n} \left( \|h_{1,i}^{k-l+1}\|_R + \|h_{2,j}^{k-l+1}\|_R \right) \quad 3 \leq k \leq N,
\]

where the sequence \((\alpha_k)_{1 \leq k \leq N}\) is introduced in Lemma 3.2.3.

Then we have for \(3 \leq k \leq N\)

(i) \(2^k B_{i,k} \leq B_{i,k}\),

(ii) \(\|K_i^k\|_R \leq B_{i,k}\),

(iii) \(\|I_i^k\|_R \leq \frac{1}{\alpha_k} B_{i,k}\).
6.3 A Priori Estimates for $I_i^k$

Proof:

(i) We prove the claim by induction.

Using Theorem 5.5.3 and Lemma 6.1.5 we obtain for $k = 3$

$$2^{3/2} B_{i,3} = 2^{3/2} \sqrt{2} R_i (\|\tilde{h}_{1,i}^2\|_R + \|\tilde{h}_{2,i}^2\|_R) \leq 2^2 R_i (2\|h_{1,i}^2\|_R + 2\|h_{2,i}^2\|_R) =$$

$$= 2^{3/2} R_i (\|\tilde{h}_{1,i}^2\|_R + \|\tilde{h}_{2,i}^2\|_R) =: B_{i,3}.$$

Now assume that $2^{3/2} B_{i,l} \leq B_{i,l}$ for $3 \leq l \leq k-1$ holds true. Using again Lemma 6.1.5 we find

$$2^{5/2} B_{i,k} = 2^{5/2} \sqrt{2} R_i (\|\tilde{h}_{1,i}^{k-1}\|_R + \|\tilde{h}_{2,i}^{k-1}\|_R) +$$

$$+ 2^{5/2} \frac{|R_2|}{\sqrt{2}} \sum_{l=3}^{k-1} \frac{1}{\alpha_l} B_{i,l} \max_{1 \leq j \leq n} (\|\tilde{h}_{1,j}\|_R + \|\tilde{h}_{2,j}\|_R) \leq$$

$$\leq 2^{5/2} R_i \left( 2^{5/2} \|h_{1,i}^{k-1}\|_R + 2^{5/2} \|h_{2,i}^{k-1}\|_R \right) +$$

$$+ 2^{5/2} |R_2| \sum_{l=3}^{k-1} \frac{1}{\alpha_l} \cdot 2^{-l/2} B_{i,l} \max_{1 \leq j \leq n} \left( 2^{k-l-1/2} \|h_{1,j}^{k-l-1}\|_R + 2^{k-l-1/2} \|h_{2,j}^{k-l-1}\|_R \right) =$$

$$= 2^k R_i (\|h_{1,i}^{k-1}\|_R + \|h_{2,i}^{k-1}\|_R) +$$

$$+ |R_2| \sum_{l=3}^{k-1} 2^{k-l-1} \frac{1}{\alpha_l} B_{i,l} \max_{1 \leq j \leq n} (\|h_{1,j}^{k-l-1}\|_R + \|h_{2,j}^{k-l-1}\|_R) =$$

$$=: B_{i,k}.$$

(ii) From Lemma 6.1.5, Lemma 6.2.5 and part (i) of this lemma it follows that

$$\|K_i^k\|_R \leq 2^{5/2} \|\tilde{K}_i^k\|_R \leq 2^{5/2} B_{i,k} \leq B_{i,k}.$$

(iii) From Lemma 6.1.5, Lemma 6.2.5 and part (i) of this lemma it follows that

$$\|I_i^k\|_R \leq 2^{5/2} \|\tilde{I}_i^k\|_R \leq 2^{5/2} \frac{1}{\alpha_k} B_{i,k} \leq \frac{1}{\alpha_k} B_{i,k}. \quad \square$$

6.3 A Priori Estimates for $I_i^k$

Under the assumption that we have simple estimates for the polynomials $h_{1,i}^k$ and $h_{2,i}^k$, we may give a priori estimates for $I_i^k$. 
Lemma 6.3.1

Suppose that the non-resonance condition NR_N holds and that there are constants \( E > 0 \) and \( h \geq 0 \) such that \( \| h_{1,i}^{k} \|_{R} + \| h_{2,i}^{k} \|_{R} \leq Eh^{k} \) for \( i \leq n \). Let the sequence \((b_{i,k})_{3 \leq k \leq N}\) be defined by

\[
b_{i,k} := 16R_{i}E \left( \alpha_{3} + 4|R_{z}^{-1}|_{\infty}Eh \right)^{k-3} h^{k-1} \frac{(k-1)!}{2 \prod_{l=3}^{k} \alpha_{l}} \quad \text{for } 3 \leq k \leq N,
\]

where the sequence \((\alpha_{k})_{3 \leq k \leq N}\) is introduced in Lemma 3.2.3.

Then the recursively defined sequences \((B_{i,k})_{3 \leq k \leq N}\) of Lemma 6.2.6 may be estimated by the explicitly defined sequences \((b_{i,k})_{3 \leq k \leq N}\):

\[
\frac{1}{\alpha_{k}} B_{i,k} \leq b_{i,k} \quad 3 \leq k \leq N.
\]

Proof: Consider the definition of the sequence \((B_{i,k})_{3 \leq k \leq N}\) given in Lemma 6.2.6. If we estimate \( \max_{1 \leq i \leq n} (\| h_{1,i}^{k-1} \|_{R} + \| h_{2,i}^{k-1} \|_{R}) \) by \( 2Eh^{k-1} \) we obtain auxiliary sequences \((\beta_{i,k})_{3 \leq k \leq N}\) with \( B_{i,k} \leq \beta_{i,k}: \)

\[
\beta_{i,3} := 4R_{i}E \cdot (2h)^2,
\]

\[
\beta_{i,k} := 4R_{i}E \cdot (2h)^{k-1} + |R_{z}^{-1}|_{\infty}E \sum_{l=3}^{k-1} l \frac{1}{\alpha_{l}} \beta_{i,l}(2h)^{k-l+1} \quad \text{for } 4 \leq k \leq N.
\]

In a first step we estimate \( \beta_{i,k} \) recursively. For \( k \geq 4 \) we find

\[
\beta_{i,k} = 4R_{i}E \cdot (2h)^{k-1} + |R_{z}^{-1}|_{\infty}E \sum_{l=3}^{k-1} l \frac{1}{\alpha_{l}} \beta_{i,l}(2h)^{k-l+1} =
\]

\[
= 2h \left( 4R_{i}E \cdot (2h)^{k-2} + |R_{z}^{-1}|_{\infty}E \sum_{l=3}^{k-2} l \frac{1}{\alpha_{l}} \beta_{i,l}(2h)(k-l+1) \right) +
\]

\[
+ |R_{z}^{-1}|_{\infty}E \cdot (k-1) \frac{1}{\alpha_{k-1}} \beta_{i,k-1}(2h)^2 =
\]

\[
= 2h \beta_{i,k-1} + |R_{z}^{-1}|_{\infty}E \cdot (k-1) \frac{1}{\alpha_{k-1}} \beta_{i,k-1}(2h)^2 \leq
\]

\[
\leq \frac{k-1}{\alpha_{k-1}} \left( \frac{\alpha_{3}}{2} + 2|R_{z}^{-1}|_{\infty}Eh \right) 2h \beta_{i,k-1}.
\]

For the last estimate we used that the sequence \((\alpha_{k})_{3 \leq k \leq N}\) is non-increasing and that \( 2 \leq k - 1 \) for \( k > 3 \).
6.4. Estimates for $I_i^{(r)}$

In a second step we give an explicit estimate for $\beta_{i,k}$:

$$\beta_{i,k} \leq \frac{(k-1)!}{2 \prod_{l=3}^{k} \alpha_l} \left( \frac{\alpha_3}{2} + 2|R_z^{-1}|_\infty E h \right)^{k-3} (2h)^{k-3} \cdot 4R_iE \cdot (2h)^2 =$$

$$= 16R_iE \left( \alpha_3 + 4|R_z^{-1}|_\infty E h \right)^{k-3} h^{k-1} \cdot \frac{(k-1)!}{2 \prod_{l=3}^{k} \alpha_l} .$$

Finally we find

$$\frac{1}{\alpha_k} B_{i,k} \leq \frac{1}{\alpha_k} \beta_{i,k} = 16R_iE \left( \alpha_3 + 4|R_z^{-1}|_\infty E h \right)^{k-3} h^{k-1} \cdot \frac{(k-1)!}{2 \prod_{l=3}^{k} \alpha_l} =: b_{i,k} .$$

Lemma 6.3.2

Suppose that the non-resonance condition $\text{NR}_N$ holds and that there are constants $E > 0$ and $h \geq 0$ such that $\|h^k_{1,i}\|_R, \|h^k_{2,i}\|_R \leq Eh^k$ for $i \leq n$. Then we have for $3 \leq k \leq N$

(i) $\|K^k_i\|_R \leq B_{i,k} \leq \alpha_k b_{i,k}$,

(ii) $\|I^k_i\|_R \leq \frac{1}{\alpha_k} B_{i,k} \leq b_{i,k}$,

where the sequences $(b_{i,k})_{3 \leq k \leq N}$ are defined in Lemma 6.3.1.

Proof: The proof follows immediately from Lemma 6.2.6 and Lemma 6.3.1. □

6.4 Estimates for $I_i^{(r)}$

In Section 5.2 we found that for a formal first integral $I_i = I_i^2 + \sum_{k=3}^{\infty} I_i^k$ the following equation holds:

$$0 = \dot{I}_i = \frac{d}{dt} I_i = \sum_{k=3}^{\infty} (LI_i^k - K_i^k) .$$

Obviously this is no longer true for the truncation $I_i^{(r)} := I_i^2 + \sum_{k=3}^{r} I_i^k$. The goal of this section is to give estimates for $I_i^{(r)}$.

Lemma 6.4.1

Suppose that the assumptions of Lemma 6.3.1 hold and let $I_i^{(r)}$ denote the truncation of the formal first integral $I_i$ to order $r < N$ introduced above. Then the following holds:

(i) $I_i^{(r)}$ contains only polynomials of order $r + 1$ and higher

$$\dot{I}_i^{(r)} = \sum_{k=r+1}^{\infty} \dot{i}_i^k .$$
where
\[ \dot{I}_i^k := x_i h_{1,i}^{k-1} + y_i h_{2,i}^{k-1} + \sum_{l=3}^r \sum_{j=1}^n \left( \frac{\partial I_i^k}{\partial x_j} h_{1,j}^{k-l+1} + \frac{\partial I_i^k}{\partial y_j} h_{2,j}^{k-l+1} \right). \]

(ii) The polynomials \( \dot{I}_i^k \) are in \( \Pi_k \) and have the following bounds:
\[ \| \dot{I}_i^k \|_R \leq \alpha_{r+1} (2h)^{k-(r+1)} b_{i,r+1}, \]
where \( b_{i,r+1} \) is defined in Lemma 6.3.1.

Proof:
(i) We first compute the derivative of \( I_i^{(r)}(t, x, y, \Delta \sigma) \):
\[
\frac{d}{dt} I_i^{(r)}(t, x, y, \Delta \sigma) = \frac{d}{dt} \left( \frac{1}{2} (x_i^2 + y_i^2) + \sum_{k=3}^r I_i^k \right) =
\]
\[
= (x_i \dot{x}_i + y_i \dot{y}_i) + \sum_{k=3}^r \left( \sum_{j=1}^n \left( \frac{\partial I_i^k}{\partial x_j} \dot{x}_j + \frac{\partial I_i^k}{\partial y_j} \dot{y}_j \right) + \frac{\partial I_i^k}{\partial t} \right). \]

Substituting \( \dot{x}_i \) and \( \dot{y}_i \) by the right-hand sides of the differential equation (5.1) we find
\[
\frac{d}{dt} I_i^{(r)}(t, x, y, \Delta \sigma) =
\]
\[
= x_i \sum_{k=2}^\infty h_{1,i}^k + y_i \sum_{k=2}^\infty h_{2,i}^k +
\]
\[
+ \sum_{k=3}^r \left( \sum_{j=1}^n \left( \frac{\partial I_i^k}{\partial x_j} \omega_j \dot{y}_j - \frac{\partial I_i^k}{\partial y_j} \omega_j \dot{x}_j \right) + \frac{\partial I_i^k}{\partial t} \right) +
\]
\[
+ \sum_{k=3}^r \sum_{l=2}^\infty \sum_{j=1}^n \left( \frac{\partial I_i^k}{\partial x_j} h_{1,j}^l + \frac{\partial I_i^k}{\partial y_j} h_{2,j}^l \right) =
\]
\[
= \left( \sum_{j=1}^n \left( \frac{\partial I_i^3}{\partial x_j} \dot{y}_j - \frac{\partial I_i^3}{\partial y_j} \dot{x}_j \right) \omega_j + \frac{\partial I_i^3}{\partial t} \right) + (x_i h_{1,i}^2 + y_i h_{2,i}^2) +
\]
\[
+ \sum_{k=4}^r \left( \sum_{j=1}^n \left( \frac{\partial I_i^k}{\partial x_j} \dot{y}_j - \frac{\partial I_i^k}{\partial y_j} \dot{x}_j \right) \omega_j + \frac{\partial I_i^k}{\partial t} +
\]
\[
+ (x_i h_{1,i}^{k-1} + y_i h_{2,i}^{k-1}) + \sum_{l=3}^k \sum_{j=1}^n \left( \frac{\partial I_i^l}{\partial x_j} h_{1,j}^{k-l+1} + \frac{\partial I_i^l}{\partial y_j} h_{2,j}^{k-l+1} \right) \right). \]
6.4. Estimates for $I_i^{(r)}$

\[ + \sum_{k=r+1}^{\infty} \left( x_i h_{1,i}^{k-1} + y_i h_{2,i}^{k-1} + \sum_{l=3}^{r} \sum_{j=1}^{n} \left( \frac{\partial I_i^l}{\partial x_j} h_{1,j}^{k-l+1} + \frac{\partial I_i^l}{\partial y_j} h_{2,j}^{k-l+1} \right) \right) = \]

\[ = (L_i^3 - K_i^3) + \sum_{k=4}^{r} (L_i^k - K_i^k) + \sum_{k=r+1}^{\infty} \hat{I}_i^k = \]

\[ = \sum_{k=r+1}^{\infty} \hat{I}_i^k. \]

In the last step we used the definition of the polynomials $\hat{I}_i^k$ and that the polynomials $I_i^k$ solve the perturbation equations for $3 \leq k \leq r$.

(ii) The polynomials $K_i^k$ and $\hat{I}_i^k$, $k \geq r + 1$ differ only in the sign and in the upper limit of the sum. Thus the estimate of $\hat{I}_i^k$ may be established similarly to the estimate of $K_i^k$. Therefore in analogy to the sequences $(B_{i,k})_{k \geq 3}$ and $(\beta_{i,k})_{k \geq 3}$ (cf. Lemma 6.2.6 and Lemma 6.3.1) we define sequences

\[ B'_{i,k} := B_{i,k}, \quad 3 \leq k \leq r, \]

\[ B'_{i,k} := 2^k R_i \left( \|h_{1,i}^{k-1}\|_R + \|h_{2,i}^{k-1}\|_R \right) + \]

\[ + \left| R_z^{-1} \right| \sum_{l=3}^{r} 2^{k-l} \frac{1}{\alpha_l} B_{i,l} \max_{1 \leq j \leq n} \left( \|h_{1,j}^{k-l+1}\|_R + \|h_{2,j}^{k-l+1}\|_R \right), \quad r < k \]

and

\[ \beta'_{i,k} := \beta_{i,k}, \quad 3 \leq k \leq r, \]

\[ \beta'_{i,k} := 4 R_i E \cdot (2h)^{k-1} + \left| R_z^{-1} \right| \sum_{l=3}^{r} \frac{1}{\alpha_l} \beta'_{i,l} (2h)^{k-l+1}, \quad r < k. \]

Taking into account that $\|h_{1,i}^{k}\|_R, \|h_{2,i}^{k}\|_R \leq Eh^k$ we see that $B'_{i,k} \leq \beta'_{i,k}$. Analogously to the estimates in Lemma 6.2.6 we find

\[ \|\hat{I}_i^k\|_R \leq B'_{i,k} \leq \beta'_{i,k}. \]

For $k > r + 1$ we may estimate $\beta'_{i,k}$ in terms of $\beta'_{i,r+1}$:

\[ \beta'_{i,k} = (2h)^{k-(r+1)} \left( 4 R_i E (2h)^{(r+1)-1} + \left| R_z^{-1} \right| \sum_{l=3}^{r} \frac{1}{\alpha_l} \beta'_{i,l} (2h)^{(r+1)-l+1} \right) = \]

\[ = (2h)^{k-(r+1)} \beta'_{i,r+1}. \]

Using the definition of $\beta'_{i,r+1}$ and the last formula of the proof of Lemma 6.3.1 we obtain

\[ \|\hat{I}_i^k\|_R \leq \beta'_{i,k} = (2h)^{k-(r+1)} \beta'_{i,r+1} = (2h)^{k-(r+1)} \beta_{i,r+1} \leq \alpha_{r+1} (2h)^{k-(r+1)} \beta_{i,r+1}. \]
6. Rigorous Estimates for First Integrals

6.5 Estimates for $I_i^{(r)}(t)$ and $\dot{I}_i^{(r)}(t)$

Now we are in the position to present the main result of this chapter:

**Theorem 6.5.1**

Let Assumptions $A_1 - A_5$ hold.

The non-resonance condition $NR_N$ in $A_{5b}$ implies the existence of a non-increasing sequence $(\alpha_k)_{3 \leq k \leq N}$ such that for $k \leq N$ the following holds (cf. Lemma 3.2.3):

$$|l_0 \omega_0 + l \cdot \omega| \geq \alpha_k > 0$$

for $l \in \mathbb{Z}^n$, $0 \leq ||l|| \leq k$, $l_0 \in \mathbb{Z}$ and $||l|| + |l_0| \neq 0$. Assume further that there are constants $E > 0$ and $h > 0$ such that

$$||h_{1,i}(t,x,y,\Delta \sigma)||_{\mathbb{R}}, ||h_{2,i}(t,x,y,\Delta \sigma)||_{\mathbb{R}} \leq Eh^k$$

for $2 \leq k \leq N$, $1 \leq i \leq n$.

Then for $2 < r < N$ the system of coupled oscillators

$$\begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t,x,y,\Delta \sigma) \\
\dot{y}_i &= -\omega_i x_i + h_{2,i}(t,x,y,\Delta \sigma)
\end{align*}$$

has approximate first integrals $I_i^{(r)} := I_i^2 + \sum_{k=3}^{r} I_i^k$, $1 \leq i \leq n$ with $I_i^2 := \frac{1}{2}(x_i^2 + y_i^2)$, such that $I_i^{(r)}$ is a power series starting with terms of degree $r + 1$ or higher.

Let

$$\sigma_0 := 1,$$

$$\sigma_k := (\alpha_3 + 4|R_k^{-1}|_{\infty} Eh)h \left( \frac{(k-1)!}{2 \prod_{l=4}^{k} \alpha_l} \right)^{1-3}$$

for $3 < k \leq N$.

For $\rho < \frac{1}{2h}$ and $|(x,y)|_{R_k} \leq \rho$ and $|\Delta \sigma|_{R_{2\rho}} \leq \rho$ the following bounds hold:

$$|(I_i^{(r)} - I_i^2)(t,x,y,\Delta \sigma)| \leq \frac{16R_i Eh^{2}\rho^3}{\alpha_3} \sum_{k=3}^{r} \left( \sigma_k \rho \right)^{k-3},$$

$$|\dot{I}_i^{(r)}(t,x,y,\Delta \sigma)| \leq 16R_i Eh^{2}\rho^3 \frac{\alpha_{r+1}}{\alpha_3} \frac{(\sigma_{r-2}\rho)^{r-2}}{1 - 2h\rho}.$$
Lemma 6.5.1
For the constants $\sigma_k$ defined in Theorem 6.5.1 the following inequality holds:

$$(\sigma_{k-3})^{k-3} \leq (\sigma_{r-3})^{k-3} \quad \text{for} \quad 3 \leq k \leq r \leq N.$$ 

Proof: Let $3 \leq k \leq r \leq N$. Then we have

$$\frac{\sigma_{k-3}^{k-3}}{\sigma_{r-3}^{k-3}} = \frac{(k-1)!}{(r-1)!} \frac{\prod_{i=4}^{k-3} \alpha_i}{\prod_{i=4}^{r-3} \alpha_i} \leq \frac{(k-1)!}{(r-1)!} \frac{\prod_{i=4}^{k-3} \alpha_i}{\prod_{i=4}^{r-3} \alpha_i}$$

We now give estimates for $C_1$ and $C_2$:

$$C_1 = 2^{-\frac{r-k}{r-3}} \frac{\alpha_{k+1} \cdots \alpha_r}{(\alpha_4 \cdots \alpha_k \cdot \alpha_{k+1} \cdots \alpha_r)^{\frac{r-k}{r-3}}} \leq$$

$$\leq 2^{-\frac{r-k}{r-3}} \frac{\alpha_{k+1} \cdots \alpha_r}{(\alpha_{k+1} \cdots \alpha_r)^{\frac{r-k}{r-3}}} =$$

$$= 2^{-\frac{r-k}{r-3}} \frac{\alpha_{k+1} \cdots \alpha_r}{(\alpha_{k+1} \cdots \alpha_r)^{\frac{r-k}{r-3}}} =$$

$$= 2^{-\frac{r-k}{r-3}} \frac{\alpha_{k+1} \cdots \alpha_r}{(\alpha_{k+1} \cdots \alpha_r)^{\frac{r-k}{r-3}}} \leq$$

$$\leq 2^{-\frac{r-k}{r-3}} \frac{\alpha_{k+1} \cdots \alpha_r}{(\alpha_{k+1} \cdots \alpha_r)^{\frac{r-k}{r-3}}} =$$

$$= 2^{-\frac{r-k}{r-3}} \frac{\alpha_{k+1} \cdots \alpha_r}{\alpha_{k+1} \cdots \alpha_r} = 2^{-\frac{r-k}{r-3}},$$

$$C_2 = \frac{(2 \cdot 3 \cdots (k-1) \cdot k \cdots (r-1))^{\frac{r-k}{r-3}}} {k \cdots (r-1)} \leq$$

$$\leq 2^{-\frac{r-k}{r-3}} \frac{(k-1)^{k-3} \cdot k \cdots (r-1)}{k \cdots (r-1)} =$$

$$= 2^{-\frac{r-k}{r-3}} \frac{(k-1)^{(k-3)(r-k) \cdot k^{r-k} \cdots (r-1)^{r-k}}}{k \cdots (r-1)} =$$

$$= 2^{-\frac{r-k}{r-3}} \frac{(k-1)^{(k-3)(r-k) \cdots (r-1)^{r-k}}}{k \cdots (r-1)} \leq$$
Thus we have
\[
\frac{\sigma_{k-3}^{-3}}{\sigma_{r-3}^{-3}} \leq 1.
\]
This completes the proof.

**Proof of Theorem 6.5.1:** From \(|(x,y)|_{R^2} \leq \rho\) and \(|\Delta \sigma|_{R^\Delta} \leq \rho\) and Lemma 6.1.4 it follows that
\[
|I_i (r) - I_i^2(t, x, y, \Delta \sigma)| \leq \left| \sum_{k=3}^{\infty} I_i^k(t, x, y, \Delta \sigma) \right| \leq \sum_{k=3}^{r} ||I_i^k||_R \rho^k.
\]
By Lemma 6.3.1 and Lemma 6.3.2 we obtain
\[
|I_i (r) - I_i^2(t, x, y, \Delta \sigma)| \leq 16R_iE \sum_{k=3}^{r} (\alpha_3 + 4|R_z|_{\infty} E h)^{k-3} \frac{(k-1)!}{2 \prod_{l=3}^{k} \alpha_l} \rho^k = \\
= 16R_iE \rho^3 h^2 \frac{\alpha_3}{\alpha_3} \cdot \left( 1 + \sum_{k=4}^{r} (\alpha_3 + 4|R_z|_{\infty} E h)^{k-3} h^{k-3} \frac{(k-1)!}{2 \prod_{l=4}^{k} \alpha_l} \rho^{k-3} \right) = \\
= 16R_iE \rho^3 h^2 \frac{\alpha_3}{\alpha_3} \cdot \left( 1 + \sum_{k=4}^{r} \left( (\alpha_3 + 4|R_z|_{\infty} E h)h \cdot \frac{(k-1)!}{2 \prod_{l=4}^{k} \alpha_l} \rho \right)^{k-3} \right) = \\
= 16R_iE \rho^3 h^2 \frac{\alpha_3}{\alpha_3} \cdot \sum_{k=3}^{r} \sigma_{k-3} \rho^{k-3}.
\]
Using Lemma 6.5.1 we may complete the proof of the first statement of Theorem 6.5.1:
\[
|I_i (r) - I_i^2(t, x, y, \Delta \sigma)| \leq \frac{16R_iE \rho^3 h^2}{\alpha_3} \cdot \sum_{k=3}^{r} (\sigma_{r-3} \rho)^{k-3}.
\]
Now we prove the second statement of Theorem 6.5.1. With the definition of the constants \(b_k\) and \(\sigma_k\) and by Lemma 6.4.1 we find
\[
|I_i (r) (t, x, y, \Delta \sigma)| \leq \\
\leq \sum_{k=r+1}^{\infty} ||I_i^k||_R \rho^k \leq \alpha_{r+1} \cdot b_{i,r+1} \cdot \sum_{k=r+1}^{\infty} (2h)^{k-(r+1)} \rho^k =
\]
6.6 The Optimal Truncation Order

The size of \( I_i(t,x,y) \) depends on the truncation order \( r \). Usually we have the following picture: For small values of \( r \) the derivative \( I_i(t,x,y) \) decreases as \( r \) increases. On the other hand for large values of \( r \) the derivative \( I_i(t,x,y) \) increases. Therefore there exists an optimal truncation order \( r_{\text{opt}} \) such that \( |I_i^{(r)}(t,x,y)|_{R_\varepsilon} \) is close to its minimum.

In the case where we work with a priori estimates we are able to determine explicitly the order with smallest error estimate.

Finding the Optimal Truncation Order

In the following lemma we give the optimal truncation order \( r_{\text{opt}} \) and the optimal estimate for \( I_i^{(r)}(t,x,y) \).

**Lemma 6.6.1**

Let the assumptions of Theorem 6.5.1 hold. Moreover assume that the sequence \((\alpha_k)_{3 \leq k \leq N}\) is of the form \( \alpha_k := \gamma k^{-\tau}, \ 3 \leq k \leq N \) with constants \( \gamma > 0, \tau \geq 0 \) (cf. Lemma 3.2.3). Let \( r_{\text{opt}} \) be such that

\[
\left( \frac{\rho_*}{\rho} \right)^{1+\tau} - 1 \leq r_{\text{opt}} < \left( \frac{\rho_*}{\rho} \right)^{1+\tau},
\]

where

\[
\rho_* := \left( \frac{h}{3^\tau} + \frac{4|R_{\varepsilon}^{-1}|_{\infty} Eh^2}{\gamma} \right)^{-1}.
\]

Then for \( N^{-(1+\tau)} \rho_* < \rho < 3^{-(1+\tau)} \rho_* \) the following holds:
(i) \(2 < r_{\text{opt}} < N\).

(ii) The right-hand side of

\[ |\dot{i}_t^{(r)}(t, x, y, \Delta \sigma)| \leq 16R_4Eh^2 \rho^{3 \frac{\alpha_{r+1}}{\alpha_3}} \frac{(\sigma_{r-2}\rho)^{r-2}}{1 - 2h\rho}. \]

as a function of \(r\) is minimal for \(r = r_{\text{opt}}\).

**Proof:**

(i) From the definition of \(r_{\text{opt}}\) and \(\rho < 3^{-(1+r)\rho_*}\) it follows that

\[ r_{\text{opt}} \geq \left( \frac{\rho_*}{\rho} \right)^{\frac{1}{1+r}} - 1 > \left( \frac{\rho_*}{3^{-(1+r)\rho_*}} \right)^{\frac{1}{1+r}} - 1 = 3 - 1 = 2. \]

Similarly it follows from \(\rho > N^{-(1+r)\rho_*}\) that

\[ r_{\text{opt}} < \left( \frac{\rho_*}{\rho} \right)^{\frac{1}{1+r}} \leq \left( \frac{\rho_*}{N^{-(1+r)\rho_*}} \right)^{\frac{1}{1+r}} = N. \]

(ii) We first show that the assumptions of Theorem 6.5.1 are fulfilled. For \(\rho \leq 3^{-(1+r)\rho_*}\) we have

\[ \rho \leq \frac{1}{3^{1+r}h^{\frac{1}{3^r}} + 4|R_z^{-1}|_{\infty} Eh^2} = \frac{1}{3h + 123^r|R_z^{-1}|_{\infty} Eh^2} < \frac{1}{3h} < \frac{1}{2h}. \]

Now consider the upper bound of \(|\dot{i}_t^{(r)}(t, x, y, \Delta \sigma)|:

\[ 16R_4Eh^2 \rho^{3 \frac{\alpha_{r+1}}{\alpha_3}} \frac{(\sigma_{r-2}\rho)^{r-2}}{1 - 2h\rho}. \]

We show that this expression takes its minimum for \(r = r_{\text{opt}}\).

Using the definition of \(\sigma_{r-2}\) we obtain

\[
\frac{\alpha_{r+1}}{\alpha_3} (\sigma_{r-2}\rho)^{r-2} = \left( \alpha_3 + 4|R_z|_{\infty} Eh \right)^{r-2} h^{r-2} \frac{r!}{2 \prod_{l=3}^{r} \alpha_l} \rho^{r-2} = \left( \gamma 3^{-r} + 4|R_z|_{\infty} Eh \right)^{r-2} h^{r-2} \frac{r!}{2 \prod_{l=3}^{r} \gamma l^{-r} \rho^{r-2}} = \left( \frac{h}{3^r} + 4|R_z|_{\infty} Eh^{2} \right)^{r-2} 2^{-(1+r)} (r!)^{1+r} \rho^{r-2} = 2^{-(1+r)} \left( \frac{\rho}{\rho_*} \right)^{r-2} (r!)^{1+r}. \]

It therefore suffices to compute the minimum of the function

\[ f : \{3, \ldots, N-1\} \to \mathbb{R}, \ r \mapsto f(r) := \left( \frac{\rho}{\rho_*} \right)^{r-2} (r!)^{1+r}. \]
To this end we compare $f(r)$ with $f(r-1)$ and $f(r+1)$. First we find

$$f(r-1) = \left(\frac{\rho}{\rho_*}\right)^{-3} ((r-1)!)^{1+\tau} = f(r) \left(\frac{\rho}{\rho_*}\right)^{1-r(r+1)^{1+\tau}}$$
$$f(r+1) = \left(\frac{\rho}{\rho_*}\right)^{-1} ((r+1)!)^{1+\tau} = f(r) \left(\frac{\rho}{\rho_*}\right)^{(r+1)^{1+\tau}}.$$

This implies

$$f(r-1) > f(r) \iff \left(\frac{\rho}{\rho_*}\right)^{-r(r+1)^{1+\tau}} > 1 \iff \left(\frac{\rho_*}{\rho}\right)^{1+\tau} > r$$
and

$$f(r+1) > f(r) \iff \left(\frac{\rho}{\rho_*}\right)^{(r+1)^{1+\tau}} > 1 \iff r \geq \left(\frac{\rho_*}{\rho}\right)^{1+\tau} - 1.$$
has $n$ approximate first integrals $I_i^{\text{opt}}$ such that for $|\langle x, y \rangle|_{\mathcal{R}_*} \leq \rho$ the following estimates hold:

$$|(I_i^{\text{opt}} - I_i^2)(t, x, y, \Delta \sigma)| \leq \frac{20 R_\rho 3^\tau}{|R_i^{-1}|_{\infty} \rho^3},$$

$$|I_i^{\text{opt}}(t, x, y, \Delta \sigma)| \leq \frac{12 R_\rho^2 \rho^2}{|R_i^{-1}|_{\infty}} \left( \frac{\rho}{\rho_*} \right)^{\frac{1}{2}} \left( \frac{e^2}{2} \right)^{1+\tau} \exp \left( -(1+\tau) \left( \frac{\rho}{\rho} \right)^{\frac{1}{2+\tau}} \right).$$

**Proof:** We start this proof with three auxiliary inequalities.

- As an easy consequence of the well-known Stirling Formula
  $$n! = \sqrt{2\pi n} \left( \frac{n}{e} \right)^n e^{\frac{\theta_n}{12n}}$$
  for $0 < \theta_n < 1$,
  we obtain
  $$n! = \sqrt{2\pi n} n^{n+\frac{1}{2}} e^{-n+\frac{6n}{12n}} = n^{n+\frac{1}{2}} e^{-n+\frac{1}{2} \log(2\pi)+\frac{6n}{12n}} < n^{n+\frac{1}{2}} e^{-n+1}$$
  for $1 < n$. (6.4)
  We make use of this inequality twice.

- From the definition of $\rho_*$ we find
  $$1 = \rho_* \left( \frac{h}{3^\tau} + \frac{4|R_i^{-1}|_{\infty} Eh^2}{\gamma} \right) = \rho_* \frac{h}{3^\tau} + \frac{4|R_i^{-1}|_{\infty} Eh^2 \rho_*}{\gamma}$$
  and we conclude that
  $$\frac{4Eh^2}{\gamma} = \frac{1}{|R_i^{-1}|_{\infty} \rho_*} - \frac{h}{|R_i^{-1}|_{\infty} 3^\tau} < \frac{1}{|R_i^{-1}|_{\infty} \rho_*}. (6.5)$$

- In the proof of Lemma 6.6.1 we showed that $\rho < 3^{-(1+\tau)} \rho_*$ implies $\rho < \frac{1}{3^h}$. Thus we have
  $$(1 - 2h\rho)^{-1} < \left( 1 - 2h \frac{1}{3^h} \right)^{-1} = 3. (6.6)$$

After these preparations we are able to prove the first estimate of the theorem. From Theorem 6.5.1 we know that

$$|(I_i^{(\tau)} - I_i^2)(t, x, y, \Delta \sigma)| < \frac{16 R_\rho E h^2 \rho^3}{\alpha_3} \cdot \sum_{k=3}^{r} (\sigma_{r-3}\rho)^{k-3}.$$**
for \( r := r_{\text{opt}} \). For \( r_{\text{opt}} = 3 \) the sum is equal to 1. For \( r_{\text{opt}} > 3 \) we show that it is smaller than 5.

Similarly as in the second part of the proof of Lemma 6.6.1 we obtain

\[
\sigma_{r-3}\rho = \left( \alpha_{3} + 4|R_{z}^{-1}|_{\infty} \right) h \left( \frac{(r-1)!}{2 \prod_{t=4} \alpha_{t}} \right) \frac{1}{r-3} \rho = \\
= \left( \gamma^{3-\tau} + 4|R_{z}^{-1}|_{\infty} h \right) h \left( \frac{(r-1)!}{2 \gamma^{r-3} \left( \frac{4}{6} \right)^{r-7}} \right) \frac{1}{r-3} \rho = \\
= \left( \frac{h}{3^{\tau}} + 4|R_{z}^{-1}|_{\infty} h^{2} \right) \rho \left( \frac{(r)!^{1+\tau}}{2 \cdot 6^{r} \cdot r} \right) \frac{1}{r-3} \\
= \frac{\rho}{\rho_{*}} \left( \frac{(r)!^{1+\tau}}{2 \cdot 6^{r} \cdot r} \right) \frac{1}{r-3}.
\]

From the definition of \( r_{\text{opt}} \) in Lemma 6.6.1 we obtain for \( r := r_{\text{opt}} \)

\[
\sigma_{r-3}\rho < r^{-(1+\tau)} \left( \frac{(r)!^{1+\tau}}{2 \cdot 6^{r} \cdot r} \right) \frac{1}{r-3} \leq \frac{\rho}{\rho_{*}} \left( \frac{(r)!^{1+\tau}}{2 \cdot 6^{r} \cdot r} \right) \frac{1}{r-3}.
\]

In the last step we used that \( r! < 6 \cdot r^{r-3} \) for \( 4 \leq r \) and that \( \tau \) is non-negative.

Now we show that \( \sigma_{r-3}\rho \leq \frac{4}{5} \). To simplify the computations we consider the logarithm of \( \sigma_{r-3}\rho \). Using (6.4) we obtain

\[
\log(\sigma_{r-3}\rho) \leq \frac{1}{r-3} \left( \log((r-1)!)-\log 2 \right) - \log r \leq \\
\leq \frac{1}{r-3} \left( \left( r - \frac{1}{2} \right) \log(r-1) - (r-2) - \log 2 - (r-3) \log r \right).
\]

It remains to prove that

\[
\frac{1}{r-3} \left( \left( r - \frac{1}{2} \right) \log(r-1) - (r-2) - \log 2 - (r-3) \log r \right) \leq \log \frac{4}{5}
\]

or

\[
f(r) := \left( r - \frac{1}{2} \right) \log(r-1) - (r-2) - \log 2 - (r-3) \left( \log r + \log \frac{4}{5} \right) \leq 0,
\]

respectively.

For \( r = 4 \) and \( r = 5 \) we find

\[
f(4) \approx -0.0112 \quad \text{and} \quad f(5) \approx -0.227
\]
and for $r \geq 5$ the function $f$ is decreasing as the following computation shows:

$$f'(r) = \log(r - 1) + \frac{r - \frac{1}{2}}{r - 1} - 1 - \log r - \frac{r - 3}{r} - \log \frac{4}{5} =$$

$$= \log \frac{r - 1}{r} + \frac{r - \frac{1}{2}}{r - 1} - \frac{r - 3}{r} - 1 - \log \frac{4}{5} <$$

$$< 0 + \frac{9}{8} - \frac{2}{5} - 1 - \log \frac{4}{5} \approx -0.05 < 0.$$  

Thus we have

$$\sigma_{r-3\rho} < \frac{4}{5}.$$  

Using this estimate we obtain

$$\sum_{k=3}^{r} (\sigma_{r-3\rho})^{k-3} < \sum_{k=0}^{r-3} \left(\frac{4}{5}\right)^{k} < \frac{1 - \left(\frac{4}{5}\right)^{r-2}}{1 - \frac{4}{5}} < \frac{1}{1 - \frac{4}{5}} = 5$$  

and with Theorem 6.5.1 and (6.5)

$$|I_{i_{\text{opt}}}^2(t, x, y, \Delta \sigma)| < \frac{16R_{i}E\rho^3h^2}{\alpha_3} \sum_{k=3}^{r} (\sigma_{r-3\rho})^{k-3} < \frac{80R_{i}E\rho^3h^2}{\gamma^3_{r-\tau}} <$$

$$< \frac{20R_{i}3^r}{|R^{-1}_{z}|^{\infty}_\rho^3}.$$  

Let us now come to the second statement. By Theorem 6.5.1 and (6.3) we obtain

$$|\hat{I}_{i_{\text{opt}}}(t, x, y, \Delta \sigma)| \leq 16R_{i}Eh^2\rho^3 \frac{\alpha_{r+1}}{\alpha_3} (\sigma_{r-2\rho})^{r-2} \frac{1}{1 - 2h\rho} =$$

$$= 16R_{i}E\rho^3h^2 2^{-(1+\gamma)} \left(\frac{\rho}{\rho}\right)^{r-2} (r!)^{1+\gamma} \frac{1}{1 - 2h\rho}.$$  

Using (6.4) and (6.6) we further have

$$|\hat{I}_{i_{\text{opt}}}(t, x, y, \Delta \sigma)| \leq 16R_{i}E\rho^3h^2 2^{-(1+\gamma)} \left(\frac{\rho}{\rho}\right)^{r-2} \left(r^{1+\gamma} e^{-r+1}\right)^{1+\gamma} \cdot 3.$$  

From the definition of $r_{\text{opt}}$ it follows that

$$|\hat{I}_{i_{\text{opt}}}(t, x, y, \Delta \sigma)| \leq$$

$$\leq 48R_{i}E\rho^3h^2 2^{-(1+\gamma)} \left(\frac{\rho}{\rho}\right)^{r-2} \left(\frac{\rho}{\rho}\right)^{r+1} \cdot \exp \left(-(1 + \gamma) \left(\frac{\rho}{\rho}\right)^{1+\gamma} - 2\right) =$$
6.6. The Optimal Truncation Order

\[ = 48 R_i E \rho^3 h^2 2^{-(1+\gamma)} \left( \frac{\rho_*}{\rho} \right)^{\frac{3}{2}} e^{2(1+\gamma)} \exp \left( -(1 + \tau) \left( \left( \frac{\rho_*}{\rho} \right)^{\frac{1}{1+\tau}} \right) \right) = \]

\[ = 48 R_i E \rho^3 h^2 \left( \frac{\rho}{\rho_*} \right)^{\frac{1}{2}} \left( \frac{e^2}{2} \right)^{1+\tau} \exp \left( -(1 + \tau) \left( \left( \frac{\rho_*}{\rho} \right)^{\frac{1}{1+\tau}} \right) \right) . \]

Finally using (6.5) we find

\[ |j^\text{opt}_i(t, x, y, \Delta \sigma)| \leq \frac{12 R_i^2 \rho_*^2}{|R_e|} \left( \frac{\rho}{\rho_*} \right)^{\frac{1}{2}} \left( \frac{e^2}{2} \right)^{1+\tau} \exp \left( -(1 + \tau) \left( \left( \frac{\rho_*}{\rho} \right)^{\frac{1}{1+\tau}} \right) \right) . \]

This completes the proof of the theorem. \( \square \)
6. Rigorous Estimates for First Integrals
First Integrals and Practical Stability

In this chapter we return to the question of practical stability of a $T$-periodic solution $z_{\text{per}}^0$ of

$$\dot{z} = F(t, z, \sigma)$$

(3.1)

for $\sigma = \sigma^0$.

Let Assumptions $A_1-A_5$ hold. Theorem 4.5.1 shows that the system (3.1) may be transformed to a system of perturbed harmonic oscillators

$$\begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t, x, y, \Delta \sigma) \\
\dot{y}_i &= -\omega_i x_i + h_{2,i}(t, x, y, \Delta \sigma)
\end{align*}$$

(5.1)

which has approximate first integrals

$$I_{i}^{(\tau)}(t, x, y, \Delta \sigma) := \frac{1}{2}(x_i^2 + y_i^2) + \sum_{k=3}^{r} I_{i}^{k}(t, x, y, \Delta \sigma), \quad 1 \leq i \leq n,$$

(5.3)

where the derivative $\dot{I}_{i}^{(\tau)}(t, x, y)$ contains only polynomials in $x$, $y$ and $\Delta \sigma$ of order $r + 1$ and higher

$$\dot{I}_{i}^{(\tau)}(t, x, y) = \sum_{k=r+1}^{\infty} \dot{I}_{i}^{k}(t, x, y, \Delta \sigma),$$

(7.1)

(cf. Theorem 5.5.2 and Lemma 6.4.1). If $\dot{I}_{i}^{(\tau)}(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma) \equiv 0$ for every solution $(x(t, \Delta \sigma), y(t, \Delta \sigma))$ of (3.1) then (5.3) is a first integral of (3.1). Throughout this chapter we assume $\dot{I}_{i}^{(\tau)}(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma) \neq 0$.

This chapter shows how to make use of these approximate first integrals to obtain results on the practical stability of $z_{\text{per}}^0$. There are three possible ways:

- **Explicit Computations**
  We compute the polynomials $I_{i}^{k}$ explicitly using a computer algebra system. For the polynomials $\dot{I}_{i}^{k}$ we may use a mix of explicit computations and a priori estimates (cf. Lemma 6.4.1).
• A priori estimates

We use a priori estimates for the polynomials $I_i^k$ and $I_i^{k^*}$. In this case we may use the optimal truncation order. This leads to a result of Nekoroshev-type.

The chapter is organized as follows:

Section 7.1: We study the stability of the equilibrium solution of the system of perturbed harmonic oscillators (5.1) and the stability of a periodic solution $z_0^{\text{per}}$ of the original system (3.1).

Section 7.2: We discuss the computations of the estimates used in the first section of the chapter.

Section 7.3: We present a Nekoroshev-type result.

7.1 Practical Stability

There is a simple connection between the distance of a point $(x, y)$ from the origin and the first integrals $I_i^2$ of the unperturbed harmonic oscillators.

Lemma 7.1.1

Let $(x(\cdot, \Delta \sigma), y(\cdot, \Delta \sigma))$ be a solution of (5.1) and let $\rho > 0$. Then the following holds:

$$|(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_i} \leq \tilde{\rho} \iff I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma)) \leq \frac{1}{2} R_i^2 \tilde{\rho}^2 \text{ for } 1 \leq i \leq n.$$  

Proof: Let $|(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_i} \leq \tilde{\rho}$. Then we have for $1 \leq i \leq n$

$$I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma)) = \frac{1}{2} (x_i^2(t, \Delta \sigma) + y_i^2(t, \Delta \sigma)) = \frac{1}{2} R_i^2 \cdot \frac{1}{R_i^2} (x_i^2(t, \Delta \sigma) + y_i^2(t, \Delta \sigma)) \leq \frac{1}{2} R_i^2 \cdot \max_{1 \leq i \leq n} \frac{1}{R_i^2} (x_i^2(t, \Delta \sigma) + y_i^2(t, \Delta \sigma)) = \frac{1}{2} R_i^2 \cdot |(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_i}^2 \leq \frac{1}{2} R_i^2 \tilde{\rho}^2.$$  

Now let $I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma)) \leq \frac{1}{2} R_i^2 \tilde{\rho}^2$ hold for $1 \leq i \leq n$. Then it follows that

$$|(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_i}^2 = \max_{1 \leq i \leq n} \frac{1}{R_i^2} (x_i^2(t, \Delta \sigma) + y_i^2(t, \Delta \sigma)) = \max_{1 \leq i \leq n} \frac{2}{R_i^2} I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma)) \leq \max_{1 \leq i \leq n} \frac{2}{R_i^2} \cdot \frac{1}{2} R_i^2 \tilde{\rho}^2 = \tilde{\rho}^2. \quad \square$$
The following lemma allows us to discuss the stability of the equilibrium solution of the system of perturbed harmonic oscillators (5.1).

**Theorem 7.1.1**

Let Assumptions $A_1 - A_5$ hold. For $r < N$ let $I_i^{(r)} := I_i + \sum_{k=3}^{r} I_{ik}$, $1 \leq i \leq n$ be the approximate first integrals of the system of perturbed harmonic oscillators

\[
\begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t, x, y, \Delta \sigma) \\
\dot{y}_i &= -\omega_i x_i + h_{2,i}(t, x, y, \Delta \sigma)
\end{align*}
\]

(5.1)

Let $\bar{\rho} := \bar{\rho}^{(r)}$ be such that

\[
\frac{1}{2} R_i^2 \bar{\rho}^2 > 2 \sum_{k=3}^{r} \| I_i^k \|_R \bar{\rho}^k, \quad 1 \leq i \leq n
\]

holds for $\bar{\rho} \in ]0, \bar{\rho}[$.

For positive numbers $\bar{\rho}_{tol}$ and $\bar{\rho}_{start}$ with

\[
\bar{\rho}_{tol} < \bar{\rho}*,
\]

\[
\bar{\rho}_{start} < \min_{1 \leq i \leq n} \left( \bar{\rho}_{tol}^2 - \frac{4}{R_i^2} \sum_{k=3}^{r} \| I_i^k \|_R \bar{\rho}_{tol}^k \right)
\]

let

\[
T_{\text{max}} := \min_{1 \leq i \leq n} \frac{1}{2} R_i^2 (\bar{\rho}_{tol}^2 - \bar{\rho}_{start}^2) - 2 \sum_{k=3}^{r} \| I_i^k \|_R \bar{\rho}_{tol}^k
\]

(7.2)

Then the following statement holds:

For any solution $(x(., \Delta \sigma), y(., \Delta \sigma))$ of (5.1) and any value of the parameter $\Delta \sigma$ with

\[
|(x(0, \Delta \sigma), y(0, \Delta \sigma))|_{R_z} \leq \bar{\rho}_{start} \quad \text{and} \quad |\Delta \sigma|_{R_{\Delta \sigma}} \leq \bar{\rho}_{tol}
\]

(7.3)

one has

\[
|(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_z} < \bar{\rho}_{tol} \quad \text{for} \quad t < T_{\text{max}}.
\]

(7.4)

**Fig. 7.1:** Practical Stability of the equilibrium.
Proof: Let \((x(t, \Delta \sigma), y(t, \Delta \sigma))\) be a solution of (5.1) fulfilling (7.3). Then either we have
\[ |(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_2} \leq \tilde{\rho}_{\text{tol}} \] for all positive times or there exists a first time \(T_{\text{ex}}\) such that
\[ |(x(T_{\text{ex}}, \Delta \sigma), y(T_{\text{ex}}, \Delta \sigma))|_{R_2} = \tilde{\rho}_{\text{tol}} \text{ and } |(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_2} < \tilde{\rho}_{\text{tol}} \text{ for } 0 \leq T < T_{\text{ex}}. \]

In the first case, (7.4) trivially holds. Thus we concentrate on the second case and show that \(T_{\text{max}} < T_{\text{ex}}\).

Let \(t \in [0, T_{\text{ex}}]\). In a first step we obtain
\[
I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma)) \leq \]
\[
\leq I_i^2(x(0, \Delta \sigma), y(0, \Delta \sigma)) + |I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma)) - I_i^2(x(0, \Delta \sigma), y(0, \Delta \sigma))| \leq \]
\[
\leq I_i^2(x(0, \Delta \sigma), y(0, \Delta \sigma)) + \]
\[
+ |I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma)) - I_i^{(r)}(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma)| + \]
\[
+ |I_i^{(r)}(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma) - I_i^{(r)}(0, x(0, \Delta \sigma), y(0, \Delta \sigma), \Delta \sigma)| + \]
\[
+ |I_i^{(r)}(0, x(0, \Delta \sigma), y(0, \Delta \sigma), \Delta \sigma) - I_i^2(x(0, \Delta \sigma), y(0, \Delta \sigma))| \leq \]
\[
\leq I_i^2(x(0, \Delta \sigma), y(0, \Delta \sigma)) + \]
\[
+ |I_i^2(t, x(t, \Delta \sigma), y(t, \Delta \sigma)) - I_i^{(r)}(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma)| + \]
\[
+ |I_i^{(r)}(0, x(0, \Delta \sigma), y(0, \Delta \sigma), \Delta \sigma) - I_i^2(x(0, \Delta \sigma), y(0, \Delta \sigma))| + \]
\[
+ \sup_{0 \leq t < T_{\text{ex}}} |I_i^{(r)}(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma)| \cdot T_{\text{ex}}. \]

The first term in the last expression may be estimated with the help of Lemma 7.1.1 and with initial conditions (7.3):
\[
I_i^2(x(0, \Delta \sigma), y(0, \Delta \sigma)) \leq \frac{1}{2} R_i^2 \rho_{\text{start}}^2. \]

The second and third term may be estimated with the help of Lemma 6.1.4:
\[
|I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma)) - I_i^{(r)}(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma)| = \]
\[
= \left| \sum_{k=3}^{r} I_i^k(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma) \right| \leq \]
\[
\leq \sum_{k=3}^{r} |I_i^k(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma)| \leq \sum_{k=3}^{r} \|I_i^k\|_{R_i} \rho_{\text{tol}}^k. \]
The last term may be estimated with the help of (7.1) and Lemma 6.1.4:
\[
|\tilde{I}_i(r)(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma)| = \left| \sum_{k=r+1}^{\infty} \tilde{I}_i^k(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma) \right| \leq \\
\sum_{k=r+1}^{\infty} |\tilde{I}_i^k(t, x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma)| \leq \sum_{k=r+1}^{\infty} \|\tilde{I}_i^k\|_R \tilde{\rho}_{tol}^k.
\]

Combining these estimates we obtain
\[
|I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma))| \leq \frac{1}{2} R_i^2 \tilde{\rho}_{start}^2 + 2 \sum_{k=3}^{r} \|I_i^k\|_R \tilde{\rho}_{tol}^k + \sum_{k=r+1}^{\infty} \|I_i^k\|_R \tilde{\rho}_{tol}^k \cdot T_{ex}.
\]

Assume that \( T_{ex} < T_{max} \). Then we have
\[
|I_i^2(x(t, \Delta \sigma), y(t, \Delta \sigma), \Delta \sigma)| \leq \frac{1}{2} R_i^2 \tilde{\rho}_{start}^2 + 2 \sum_{k=3}^{r} \|I_i^k\|_R \tilde{\rho}_{tol}^k + \sum_{k=r+1}^{\infty} \|I_i^k\|_R \tilde{\rho}_{tol}^k \cdot T_{ex} < \\
< \frac{1}{2} R_i^2 \tilde{\rho}_{start}^2 + 2 \sum_{k=3}^{r} \|I_i^k\|_R \tilde{\rho}_{tol}^k + \sum_{k=r+1}^{\infty} \|I_i^k\|_R \tilde{\rho}_{tol}^k \cdot T_{max} = \\
= \frac{1}{2} R_i^2 \tilde{\rho}_{tol}^2.
\]

Now Lemma 7.1.1 implies that \(|(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_x} < \tilde{\rho}_{tol} \) for \( t < T_{max} \) in contradiction to the definition of \( T_{ex} \). Thus the solution \((x(t, \Delta \sigma), y(t, \Delta \sigma))\) of (5.1) is defined for \( t \in [0, T_{max}] \) and is bounded by \( \tilde{\rho}_{tol} \). \( \square \)

Now we return to the original system of differential equations (3.1) and study the stability of periodic solutions \( z_{0, per}^* \). It is sufficient to consider the distance between a solution \( z(., \sigma) \) of (3.1) and \( z_{0, per}^* \) for integer multiples of the period \( T \) only.

The idea is to reduce the problem to the stability of the equilibrium solution of a system of perturbed harmonic oscillators. The link between (3.1) and (5.1) is given by the sequence of coordinate transformations introduced in Chapter 4. They may be summarized as
\[
\left( \begin{array}{c} z_4 \\ \Delta \sigma \end{array} \right) := \left( S^{-1} Q^{-1}(t) \left( z(t, \sigma) - z_{0, per}(t) \right) - z_3(t)(\sigma - \sigma^0) \right).
\]

For details see Theorem 4.5.1.

Now let \( \rho_{start} \) and \( \rho_{tol} \) be given. It remains to determine constants \( \tilde{\rho}_{start} \) and \( \tilde{\rho}_{tol} \) such that the following situation holds:
\[
|z(0, \sigma) - z_{0, per}(0)| \leq \rho_{start} \quad |z(nT, \sigma) - z_{n, per}(nT)| \leq \rho_{tol} \quad \text{for} \quad nT \leq T_{max}
\]

\( \Downarrow \)
\( \Uparrow \)
\[
|z_4(0, \Delta \sigma)|_{R_x} < \tilde{\rho}_{start} \quad \text{Theorem 7.1.1} \quad |z_4(nT, \Delta \sigma)|_{R_x} < \tilde{\rho}_{tol} \quad \text{for} \quad nT \leq T_{max}
\]
Theorem 7.1.2
Let Assumptions $A_1 - A_5$ hold for

$$\dot{z} = F(t, z, \sigma)$$

and let $z_{\text{per}}^{0}$ denote the $T$-periodic solution of (3.1) for $\sigma = \sigma^0$.

The coordinate transformation

$$
\begin{pmatrix}
    z_4 \\
    \Delta \sigma
\end{pmatrix}
:=
(S^{-1}Q^{-1}(t) \left( z(t, \sigma) - z_{\text{per}}^{0}(t) \right) - z_3^{0}(\sigma - \sigma^0))
$$

transforms (3.1) to the system of perturbed harmonic oscillators

$$
\begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t, x, y, \Delta \sigma) \\
\dot{y}_i &= -\omega_i x_i + h_{2,i}(t, x, y, \Delta \sigma)
\end{align*}
\quad 1 \leq i \leq n
$$

(cf. Theorem 4.5.1) with approximate first integrals

$$I_i^{(r)} := I_i^2 + \sum_{k=3}^{r} I_i^k, \quad 1 \leq i \leq n,$$

where $r < N$ (cf. Lemma 6.4.1).

Let $\tilde{\rho}_* := \tilde{\rho}_*(r)$ be chosen such that

$$\frac{1}{2} R_i^2 \tilde{\rho}_*^2 > 2 \sum_{k=3}^{r} \|I_i^k\|_R \tilde{\rho}_*^k, \quad 1 \leq i \leq n$$

holds for $\tilde{\rho} \in [0, \tilde{\rho}_*]$. For positive numbers $\rho_{\text{start}}$ and $\rho_{\text{col}}$ define the quantities

$$\tilde{\rho}_{\text{start}} := \|S^{-1}\|_{R_e} \cdot \rho_{\text{start}} + \|z_3^{0}(0)\|_{R_e/R_{\Delta \sigma}} \cdot \rho_{\text{col}},$$

$$\tilde{\rho}_{\text{col}} := \left( \|S^{-1}\|_{R_e} - \|z_3^{0}(0)\|_{R_e/R_{\Delta \sigma}} \right) \cdot \rho_{\text{col}}$$

and assume that they satisfy the conditions

$$0 < \tilde{\rho}_{\text{col}} < \tilde{\rho}_*,$$

$$\tilde{\rho}_{\text{start}}^2 < \min_{1 \leq i \leq n} \left( \tilde{\rho}_{\text{col}}^2 - \frac{4}{R_i^2} \sum_{k=3}^{r} \|I_i^k\|_R \tilde{\rho}_{\text{col}}^k \right).$$

Finally let $T_{\text{max}}$ be defined by

$$T_{\text{max}} := \min_{1 \leq i \leq n} \frac{\frac{1}{2} R_i^2 (\tilde{\rho}_{\text{col}}^2 - \tilde{\rho}_{\text{start}}^2) - 2 \sum_{k=3}^{r} \|I_i^k\|_R \tilde{\rho}_{\text{col}}^k}{\sum_{k=r+1}^{\infty} \|I_i^k\|_R \tilde{\rho}_{\text{col}}^k}.$$
Then the following holds:
For any solution \( z(., \sigma) \) of (3.1) and any value of the parameter \( \sigma \) with
\[
|z(0, \sigma) - z^0_{\text{per}}(0)| \leq \rho_{\text{start}} \quad \text{and} \quad |\sigma - \sigma^0|_{R_{\Delta \sigma}} \leq \rho_{\text{tol}}
\]
one has
\[
|z(nT, \sigma) - z^0_{\text{per}}(nT)| < \rho_{\text{tol}} \quad \text{for} \quad 0 \leq nT < T_{\text{max}}.
\]

**Fig. 7.2:** Practical stability of periodic solutions.

**Proof:** Let the assumptions of the theorem hold. We show that for the transformed system the assumptions of Theorem 7.1.1 are fulfilled.

- \(|z(0, \sigma^0 + \Delta \sigma) - z^0_{\text{per}}(0)| < \rho_{\text{start}} \implies |z_4(0, \Delta \sigma)|_{R_z} < \tilde{\rho}_{\text{start}}|
With the help of Theorem 4.5.1 (v) we find
\[
|z_4(0, \Delta \sigma)|_{R_z} \leq \|S^{-1}\|_{R_z} |z(0, \sigma^0 + \Delta \sigma) - z^0_{\text{per}}(0)| + \|z^0_{\text{per}}(0)\|_{R_z/R_{\Delta \sigma}} |\Delta \sigma|_{R_{\Delta \sigma}} \leq
\]
\[
\leq \|S^{-1}\|_{R_z} \rho_{\text{start}} + \|z^0_{\text{per}}(0)\|_{R_z/R_{\Delta \sigma}} \rho_{\text{tol}} =
\]
\[
\tilde{\rho}_{\text{start}}.
\]
- \(|z_4(nT, \sigma^0 + \Delta \sigma) < \rho_{\text{tol}} \implies |z(nT, \sigma^0 + \Delta \sigma) - z^0_{\text{per}}(nT)| < \rho_{\text{tol}}|
Using again Theorem 4.5.1 (v) we have
\[
|z(nT, \sigma^0 + \Delta \sigma) - z^0_{\text{per}}(nT)| \leq
\]
\[
\leq \|S\|_{R_z^{-1}} \cdot (|z_4(nT, \Delta \sigma)|_{R_z} + \|z^0(0)\|_{R_z/R_{\Delta \sigma}} |\Delta \sigma|_{R_{\Delta \sigma}} ) \leq
\]
\[
\leq \|S\|_{R_z^{-1}} \cdot (\rho_{\text{tol}} + \|z^0(0)\|_{R_z/R_{\Delta \sigma}} \rho_{\text{tol}}) =
\]
First Integrals and Practical Stability

\[ = \|S\|_{R_\ell} \cdot \left( (\|S\|_{R_\ell}^{-1} - \|z_0^0(0)\|_{R_\ell/\Delta \rho}) \rho_{\text{tol}} + \|z_3^0(0)\|_{R_\ell/\Delta \rho} \rho_{\text{tol}} \right) = \]

\[ =: \rho_{\text{tol}}. \]

Now the claim follows from Theorem 7.1.1. □

7.2 Explicit Computations and Estimates

In the previous section we derived two theorems on practical stability with the help of approximate first integrals of order \( r \). We obtained the lower bound for the maximum time of existence for solutions of (5.1):

\[ T_{\text{max}} := \min_{1 \leq i \leq n} \frac{1}{2} R_i^2 (\rho_{\text{tol}}^2 - \rho_{\text{start}}^2) - 2 \sum_{k=3}^{r} \|I_k\|_{R} \rho_{\text{tol}}. \]

In this section we discuss the possibilities to compute the norms \( \|I_k\|_{R} \) and \( \|I_k\|_{R} \).

The Computation of \( \|h_{1,i}^k\|_{R} \) and \( \|h_{2,i}^k\|_{R} \)

In principle the polynomials \( h_{1,i}^k \) and \( h_{2,i}^k \) of the perturbation may be computed up to any finite order \( M \). But we must be aware that the costs of computation grow rapidly with the order \( M \). So they may become prohibitively high.

In addition we need simple explicit estimates for \( \|h_{1,i}^k\|_{R} \) and \( \|h_{2,i}^k\|_{R} \) to compute the polynomials \( \|I_k\|_{R} \).

In the sequel we assume that the polynomials \( h_{1,i}^k \) and \( h_{2,i}^k \) are computed explicitly up to some order \( M \) bigger than the truncation order \( r \) of the first integrals (e.g. \( M = 2r \)). For higher order terms we use estimates of the form

\[ \|h_{m,i}^k(t,x,y,\Delta \sigma)\|_{R} \leq Eh^k \quad \text{for} \quad m = 1, 2, \quad 1 \leq i \leq n, \quad M \leq k \]

with constants \( E > 0 \) and \( h \geq 0 \).

The Computation of \( \|I_k^k\|_{R} \)

The polynomials \( I_k^k \) are obtained from the recursive scheme (cf. Lemma 5.2.1):

\[ LI_1^k = -x_i h_{1,i}^2 - y_i h_{2,i}^2, \]

\[ LI_k^k = -x_i h_{1,i}^{k-1} - y_i h_{2,i}^{k-1} - \sum_{l=3}^{k-1} \sum_{j=1}^{n} \left( \frac{\partial I_i^{k-1}}{\partial x_j} h_{1,j}^2 + \frac{\partial I_i^{k-1}}{\partial y_j} h_{2,j}^2 \right), \quad 4 \leq k \leq r. \]
These formulae allow us to compute the polynomials $I_k^i$ up to the given order $r$. But again we must be aware that the costs of computation grow rapidly. Thus for large $k$ we may be forced to use the estimate

$$\|I_k^i\|_R \leq \frac{1}{\alpha_k} b_{i,k},$$

where the sequences $(B_{i,k})_{3 \leq k \leq N}$ are recursively defined by

$$B_{i,3} := R_t 2^3 (\|h_{1,i}^3\|_R + \|h_{2,i}^3\|_R),$$

$$B_{i,k} := R_t 2^k (\|h_{1,i}^{k-1}\|_R + \|h_{2,i}^{k-1}\|_R) +$$

$$+ |R_z^{-1}| \sum_{l=3}^{k-1} \frac{1}{\alpha_l} B_{i,l} 2^{k-l-1} \max_{1 \leq j \leq n} (\|h_{1,j}^{k-l-1}\|_R + \|h_{2,j}^{k-l-1}\|_R), \quad 4 \leq k \leq N;$$

(cf. Lemma 6.2.6) or we must even use the estimate

$$\|I_k^i\|_R \leq b_{i,k},$$

where the sequences $(b_{i,k})_{3 \leq k \leq N}$ are explicitly defined by

$$b_{i,k} := 16 R_t E \left( \alpha_3 + 4 |R_z^{-1}|_{\infty} E h \right)^{k-3} h^{k-1} (k-1)! \frac{2 \prod_{l=3}^{k} \alpha_l}{2 \prod_{l=3}^{k} \alpha_l}$$

for $3 \leq k \leq N$

(cf. Lemma 6.3.1). We will see in the following chapter that these estimates are much worse than the estimates obtained by explicit computations of the polynomials.

**The Computation of $\|I_k^i\|_R$**

The polynomials $I_k^i$ are given by

$$I_k^i = x_i h_{1,i}^{k-1} + y_i h_{2,i}^{k-1} + \sum_{l=3}^{r} \sum_{j=1}^{n} \left( \frac{\partial I_l^j}{\partial x_j} h_{1,j}^{k-l+1} + \frac{\partial I_l^j}{\partial y_j} h_{2,j}^{k-l+1} \right), \quad r < k$$

(cf. Lemma 6.4.1). Since the number of polynomials $I_k^i$ is infinite, it is not possible to compute them all explicitly.

In order to keep the sum $\sum_{k=r+1}^{\infty} \|I_k^i\|_R \rho_{\text{tol}}^k$ as small as possible we may use a mix of explicit computations and a priori estimates.

Let $M$ be the order up to which we compute the polynomials $I_k^i$ explicitly. Using the definition of $I_k^i$ in Lemma 6.4.1 and Lemma 6.2.2 we find

$$\sum_{k=r+1}^{\infty} \|I_k^i\|_R \rho_{\text{tol}}^k \leq$$

$$\sum_{k=r+1}^{M} \|I_k^i\|_R \rho_{\text{tol}}^k + \sum_{k=M+1}^{\infty} \left( 2|R_z|_{\infty} E h^{k-1} + 2|R_z^{-1}|_{\infty} E \sum_{l=3}^{r} \|I_l^i\|_R h^{k-l+1} \right) \rho_{\text{tol}}^k \leq$$
The Non-Resonance Condition

For the finite truncation order \( r \) the non-resonance condition \( \text{NR}_N \) must at least hold for \( N = r + 1 \). Moreover due to Lemma 3.2.3 it is equivalent to \( \text{NR}_{r+1}^t \):

\[
|l_0 \omega_0 + l \cdot \omega| \geq \alpha_k > 0
\]

for \( l_0 \in \mathbb{Z}, l \in \mathbb{Z}^n, |l|_1 = k \leq r + 1, |l_0| + |l|_1 \neq 0 \), where \( (\alpha_k)_{1 \leq k \leq r+1} \) is a non-increasing sequence.

The proof of Lemma 3.2.3 shows that \( \text{NR}_{r+1} \) must be checked only for a finite range of numbers \( l_0 \). This allows us to compute the sequence \( (\alpha_k)_{1 \leq k \leq r+1} \).

7.3 Nekoroshev-Type Results

In this section we show that the time \( T_{\text{max}} \) increases exponentially with the smallness of \( \rho_{\text{tol}} \). Again we start with the system of perturbed harmonic oscillators.

**Theorem 7.3.1**

Let Assumptions \( A_1 - A_5 \) hold and consider the system of perturbed harmonic oscillators

\[
\begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t, x, y) \\
\dot{y}_i &= -\omega_i x_i + h_{2,i}(t, x, y)
\end{align*}
\]

1 \leq i \leq n.  

(5.1)

In accordance with the non-resonance condition \( \text{NR}_N \) in \( A_{5b} \) there exists a non-increasing sequence \( (\alpha_k)_{3 \leq k \leq N} \) of the form \( \alpha_k := \gamma k^{-\tau} \) with constants \( \gamma > 0, \tau \geq 0 \) such that for \( 3 \leq k \leq N \) the following holds:

\[
|l_0 \omega_0 + l \cdot \omega| \geq \gamma \frac{1}{|l|_1}
\]

for \( l \in \mathbb{Z}^n, 0 \leq |l|_1 \leq k, l_0 \in \mathbb{Z} \) and \( |l|_1 + |l_0| \neq 0 \).
Assume further that there are constants \( E > 0 \) and \( h > 0 \) such that
\[
\| h_{m,i}(t, x, y, \Delta \sigma) \|_R \leq E h^k \quad \text{for} \quad m = 1, 2, \quad 1 \leq i \leq n, \quad 2 \leq k.
\]
Let
\[
\tilde{\rho}_* := \left( \frac{h}{3^\tau} + \frac{4|R_x^{-1}|_\infty E h^2}{\gamma} \right)^{-1}.
\]
For positive numbers \( \tilde{\rho}_{\text{col}} \) and \( \tilde{\rho}_{\text{start}} \) with
\[
\frac{\tilde{\rho}_*}{N^{1+\tau}} < \tilde{\rho}_{\text{col}} \leq \frac{\tilde{\rho}_*}{40 \cdot 3^{1+\tau}},
\]
\[
\tilde{\rho}_{\text{start}} < \left( 1 - 37 \cdot \tilde{\rho}_{\text{col}} \right) \tilde{\rho}_{\text{col}}
\]
let
\[
T_{\text{max}} := \frac{\tilde{\rho}_{\text{col}}^2 - \tilde{\rho}_{\text{start}}^2 - 80 \cdot 3^\tau \tilde{\rho}_{\text{col}}^3}{24 \gamma \tilde{\rho}_*^3} \left( \frac{2}{e^2} \right)^{1+\tau} \exp \left( (1 + \tau) \left( \frac{\tilde{\rho}_*}{\tilde{\rho}_{\text{col}}} \right)^{1+\tau} \right).
\]
Then the following statement holds:
For any solution \((x(., \Delta \sigma), y(., \Delta \sigma))\) of (5.1) and any value of the parameter \( \Delta \sigma \) with
\[
\|(x(0, \Delta \sigma), y(0, \Delta \sigma))\|_{R_z} \leq \tilde{\rho}_{\text{start}} \quad \text{and} \quad |\Delta \sigma|_{R_{\Delta \sigma}} \leq \tilde{\rho}_{\text{col}} \quad (7.6)
\]
one has
\[
\|(x(t, \Delta \sigma), y(t, \Delta \sigma))\|_{R_z} < \tilde{\rho}_{\text{col}} \quad \text{for} \quad 0 < t < T_{\text{max}}.
\]
\[
\begin{align*}
\tilde{\rho}_* & \quad \tilde{\rho}_{\text{col}} \quad \tilde{\rho}_{\text{start}} \\
(x(t, \Delta \sigma), y(t, \Delta \sigma)) & \quad (x(0, \Delta \sigma), y(0, \Delta \sigma)) \\
0 & \quad 0
\end{align*}
\]
\textbf{Fig. 7.3: Practical Stability of the equilibrium.}

\textbf{Proof:} Let \((x(t, \Delta \sigma), y(t, \Delta \sigma))\) be a solution of (5.1) fulfilling (7.6). Then either we have
\[
\|(x(t, \Delta \sigma), y(t, \Delta \sigma))\|_{R_z} \leq \tilde{\rho}_{\text{col}} \quad \text{for} \quad \text{all positive times or there exists a first time} \ T_{\text{ex}} \text{ such that} \quad \|(x(T_{\text{ex}}, \Delta \sigma), y(T_{\text{ex}}, \Delta \sigma))\|_{R_z} = \tilde{\rho}_{\text{col}} \text{ and} \quad \|(x(t, \Delta \sigma), y(t, \Delta \sigma))\|_{R_z} < \tilde{\rho}_{\text{col}} \quad \text{for} \quad 0 \leq T < T_{\text{ex}}.
\]
In the first case, (7.7) trivially holds. Thus we concentrate on the second case and show that $T_{\max} < T_{\text{ex}}$.

For $\rho_* := \tilde{\rho}_* \text{ and } \rho_{\text{col}} := \tilde{\rho}_{\text{col}}$ the assumptions of Lemma 6.6.1 and Theorem 6.6.1 are fulfilled. Lemma 6.6.1 implies that the optimal truncation order $r_{\text{opt}}$ defined by

$$\left( \frac{\tilde{\rho}_*}{\rho} \right)^{1+\tau} - 1 \leq r_{\text{opt}} < \left( \frac{\tilde{\rho}_*}{\rho} \right)^{1+\tau}$$

lies between 2 and $N - 1$. Thus the approximate first integral

$$I^\text{opt}_i(t, x, y, \Delta \sigma) := \frac{1}{2} \left( x^2_i + y^2_i \right) + \sum_{k=3}^{r_{\text{opt}}} I^k_i(t, x, y, \Delta \sigma)$$

exists and the following estimates hold (cf. Theorem 6.6.1):

$$|I^\text{opt}_i(t, x, y, \Delta \sigma)| \leq \frac{20 R_i 3^\tau}{|R^{-1}_{z_i}|_{\infty} \rho_*} \rho^3,$$

$$|I^\text{opt}_i(t, x, y, \Delta \sigma)| \leq \frac{12 R_i \gamma \rho_*^2}{|R^{-1}_{z_i}|_{\infty} \rho_*} \left( \frac{\rho}{\rho_*} \right)^{1+\tau} \left( \frac{e^2}{2} \right)^{1+\tau} \exp \left( -(1+\tau) \left( \frac{\tilde{\rho}_*}{\rho} \right)^{1+\tau} \right).$$

Now let $t \in [0, T_{\text{ex}}]$. Using the estimates given above we find

$$|I^2_i(x(t, \Delta \sigma), y(t, \Delta \sigma)) - I^\text{opt}_i(t, x(t, \Delta \sigma), y(t, \Delta \sigma))| \leq$$

$$\leq \frac{20 R_i 3^\tau}{|R^{-1}_{z_i}|_{\infty} \rho_*} \rho_{\text{col}} \leq \frac{40 \cdot 3^\tau \tilde{\rho}_{\text{col}}}{|R^{-1}_{z_i}|_{\infty} R_i} \cdot \frac{1}{2} R_i \rho_{\text{col}} \leq$$

$$\leq 40 \cdot 3^\tau \tilde{\rho}_{\text{col}} \cdot \frac{1}{2} R_i \rho_{\text{col}}$$

and

$$\sup_{0 \leq t < T_{\text{ex}}} |\dot{I}^\text{opt}_i(t, x(t, \Delta \sigma), y(t, \Delta \sigma))| \cdot T_{\text{ex}} \leq$$

$$\leq \frac{12 R_i \gamma \rho_*^2}{|R^{-1}_{z_i}|_{\infty} \rho_*} \left( \frac{\tilde{\rho}_{\text{col}}}{\tilde{\rho}_*} \right)^{1+\tau} \left( \frac{e^2}{2} \right)^{1+\tau} \exp \left( -(1+\tau) \left( \frac{\tilde{\rho}_*}{\rho} \right)^{1+\tau} \right) \cdot T_{\text{ex}} \leq$$

$$\leq 24 \gamma \left( \frac{\rho_*}{\rho_{\text{col}}} \right)^{1+\tau} \left( \frac{e^2}{2} \right)^{1+\tau} \exp \left( -(1+\tau) \left( \frac{\tilde{\rho}_*}{\rho} \right)^{1+\tau} \right) \cdot T_{\text{ex}} \cdot \frac{1}{2} R_i \rho_{\text{col}}.$$
7.3. Nekoroshev-Type Results

Combining these estimates we obtain similarly as in the proof of Theorem 7.1.1
\[ |I_1^t(x(t, \Delta \sigma), y(t, \Delta \sigma))| \leq \]
\[ \leq \frac{1}{2} R_1^2 \rho_{\text{start}} + 2 \cdot 40 \cdot 3^{1+\gamma} R_{\text{tol}} + \frac{1}{2} R_1^2 R_{\text{tol}}^2 + \]
\[ + 24 \gamma \left( \frac{\rho_0}{R_{\text{tol}}} \right)^{3/2} \left( e^2 \right)^{1+\gamma} \cdot \exp \left( -\left( 1 + \gamma \right) \left( \frac{\rho_0}{\rho} \right)^{1+\gamma} \right) \cdot T_{\text{ex}} \cdot \frac{1}{2} R_1^2 \rho_{\text{tol}}. \]

Now assume that \( T_{\text{ex}} < T_{\max} \). Then we have
\[ |I_1^t(x(t, \Delta \sigma), y(t, \Delta \sigma))| \leq \]
\[ \leq \frac{1}{2} R_1^2 \rho_{\text{start}} + 2 \cdot 40 \cdot 3^{1+\gamma} R_{\text{tol}} + \frac{1}{2} R_1^2 R_{\text{tol}}^2 + \]
\[ + 24 \gamma \left( \frac{\rho_0}{R_{\text{tol}}} \right)^{3/2} \left( e^2 \right)^{1+\gamma} \cdot \exp \left( -\left( 1 + \gamma \right) \left( \frac{\rho_0}{\rho} \right)^{1+\gamma} \right) \cdot T_{\text{ex}} \cdot \frac{1}{2} R_1^2 \rho_{\text{tol}} < \]
\[ < \frac{1}{2} R_1^2 \rho_{\text{start}} + 2 \cdot 40 \cdot 3^{1+\gamma} R_{\text{tol}} + \frac{1}{2} R_1^2 R_{\text{tol}}^2 + \]
\[ + 24 \gamma \left( \frac{\rho_0}{R_{\text{tol}}} \right)^{3/2} \left( e^2 \right)^{1+\gamma} \cdot \exp \left( -\left( 1 + \gamma \right) \left( \frac{\rho_0}{\rho} \right)^{1+\gamma} \right) \cdot T_{\max} \cdot \frac{1}{2} R_1^2 \rho_{\text{tol}} = \]
\[ = \frac{1}{2} R_1^2 \rho_{\text{tol}}. \]

Now Lemma 7.1.1 implies that \( |(x(t, \Delta \sigma), y(t, \Delta \sigma))|_{R_*} < \rho_{\text{tol}} \) for \( t < T_{\max} \) in contradiction to the definition of \( T_{\text{ex}} \). Thus the solution \((x(t, \Delta \sigma), y(t, \Delta \sigma))\) of (5.1) is defined for \( t \in [0, T_{\max}] \) and is bounded by \( \rho_{\text{tol}} \). \( \square \)

Now we return to the original system (3.1) and derive a Nekoroshev-type result as well.

**Theorem 7.3.2**

Let Assumptions \( A_1 - A_5 \) hold for

\[ \dot{z} = F(t, z, \sigma) \quad (3.1) \]

and let \( z^0_{\text{per}} \) denote the \( T \)-periodic solution of (3.1) for \( \sigma = \sigma^0 \).

The coordinate transformation

\[ \begin{pmatrix} z_4 \\ \Delta \sigma \end{pmatrix} := \begin{pmatrix} S^{-1} Q^{-1}(t) & (z(t, \sigma) - z^0_{\text{per}}(t)) - z^0_3(t)(\sigma - \sigma^0) \\ \sigma - \sigma^0 \end{pmatrix} \]
transforms (3.1) to the system of perturbed harmonic oscillators

\[ \begin{align*}
\dot{x}_i &= \omega_i y_i + h_{1,i}(t, x, y, \Delta \sigma) \\
\dot{y}_i &= -\omega_i x_i + h_{2,i}(t, x, y, \Delta \sigma)
\end{align*} \]

(1 ≤ i ≤ n) \hfill (5.1)

(cf. Theorem 4.5.1).

In accordance with the non-resonance condition \( \text{NR}_N \) in \( A_{5b} \), there exists a non-increasing sequence \((\alpha_k)_{3 < k < N} \) of the form \( \alpha_k := \gamma k^{-\tau} \) with constants \( \gamma > 0 \), \( \tau \geq 0 \) such that for \( 3 < k < N \) the following holds:

\[ |l_0 \omega_0 + l \cdot \omega| \geq \gamma \frac{1}{|l||l|} \]

for \( l \in \mathbb{Z}^n \), \( 0 \leq |l|_1 \leq k \), \( l_0 \in \mathbb{Z} \) and \( |l|_1 + |l_0| \neq 0 \).

Assume further that there are constants \( E > 0 \) and \( h \geq 0 \) such that

\[ \|h^k_{m,i}(t, x, y, \Delta \sigma)\|_R \leq Eh^k \quad \text{for} \quad m = 1, 2, \quad 1 \leq i \leq n, \quad 2 < k. \]

Let

\[ \bar{\rho}_* := \left( \frac{h}{3^r} + \frac{4c^r |R^{-1}_{z,1}|_{\infty} Eh^2}{\gamma} \right)^{-1}. \]

For positive numbers \( \rho_{\text{start}} \) and \( \rho_{\text{tol}} \) define the quantities

\[ \bar{\rho}_{\text{start}} := \|S^{-1}\|_{R_z} \cdot \rho_{\text{start}} + \|z_3^0(0)\|_{R_z/R_{\Delta \sigma}} \cdot \rho_{\text{tol}} \]

\[ \bar{\rho}_{\text{tol}} := \left( \|S^{-1}\|_{R_z} - \|z_3^0(0)\|_{R_z/R_{\Delta \sigma}} \right) \cdot \rho_{\text{tol}} \]

and assume that they satisfy the conditions

\[ \frac{\bar{\rho}_*}{N^{1+\tau}} \leq \bar{\rho}_{\text{tol}} < \frac{\bar{\rho}_*}{40 \cdot 3^{1+\tau}}, \]

\[ \bar{\rho}_{\text{start}} := \left( 1 - 80 \cdot 3^r \frac{\bar{\rho}_{\text{tol}}}{\bar{\rho}_*} \right)^{\frac{1}{2}} \bar{\rho}_{\text{tol}}. \]

Finally let

\[ T_{\text{max}} := \frac{\bar{\rho}_*^2 - \bar{\rho}_{\text{start}}^2 - 80 \cdot 3^r \frac{\bar{\rho}_{\text{tol}}^3}{\rho_*} \left( \frac{2}{e^2} \right)^{1+\tau} \exp \left( (1 + \tau) \left( \frac{\bar{\rho}_*}{\rho_{\text{tol}}} \right)^{\frac{1}{1+\tau}} \right)}{24 \gamma \bar{\rho}_*^{\frac{3}{2}} \rho_{\text{tol}}^{\frac{1}{2}}}. \]

(7.8)

Then the following holds:

For any solution \( z(., \sigma) \) of (3.1) and any value of the parameter \( \sigma \) with

\[ |z(0, \sigma) - z^0_{\text{per}}(0)| \leq \rho_{\text{start}} \quad \text{and} \quad |\sigma - \sigma^0|_{R_{\Delta \sigma}} \leq \rho_{\text{tol}} \]

one has

\[ |z(nT, \sigma) - z^0_{\text{per}}(nT)| < \rho_{\text{tol}} \quad \text{for} \quad 0 \leq nT < T_{\text{max}}. \]
7.3. Nekoroshev-Type Results

Fig. 7.4: Practical stability of periodic solutions.

**Proof:** The proof is done as the proof of Theorem 7.1.2.
Stability of the Dumbbell Satellite

We return to the rotational motion of the dumbbell satellite around its center of mass. The equation of motion is given by

\[ \dot{z} = F(t, z, e), \quad (2.9) \]

where the function \( F \) is defined by

\[
F(t, z, e) := F(t, x, y, e) = \left( \begin{array}{c} 
\frac{3}{2} \sin(2x) 
\frac{y}{2(1 - c \cos t)} 
+ y \frac{e \sin t}{1 - e \cos t} 
+ \frac{2e \sqrt{1 - e^2 \sin t}}{(1 - e \cos t)^2} 
\end{array} \right).
\]

The basic properties are

- \( P_1: \) \( F \) is in \( C^\infty(\mathbb{R} \times \mathbb{R}^2 \times [0, 1[ \times \mathbb{R}^2) \),
- \( P_2: \) \( F \) is 2\( \pi \)-periodic in \( t \):
  \[ F(t + 2\pi, z, e) = F(t, z, e), \]
- \( P_3: \) \( F \) is reversible:
  \[ F(-t, Rz, e) = -RF(t, z, e), \]

where

\[
R := \begin{pmatrix} -1 & 0 \\
0 & 1 \end{pmatrix}.
\]

The chapter is organized as follows:

**Section 8.1:** We show that there exist 2\( \pi \)-periodic solutions for a wide range of eccentricities \( e \).

**Section 8.2:** We discuss the linear stability of the 2\( \pi \)-periodic solutions.

**Section 8.3:** We show that for \( e = 0 \) the trivial solution is stable in the sense of Lyapunov and derive results on practical stability.

**Section 8.4:** We present results on the stability of 2\( \pi \)-periodic solutions for prescribed values of the eccentricity \( e \) using the general results obtained in the last chapter.
Section 8.5: We present results on the stability of $2\pi$-periodic solutions for a whole range of values of the eccentricity $e$.

Section 8.6: We study the influence of the particular choice of the coordinate transformation on our results.

Section 8.7: We close the chapter with some remarks on the computations carried out on the computer.

8.1 $2\pi$-periodic Solutions

In this section we compute $2\pi$-periodic solutions of (2.9). To this end we start with two auxiliary results.

Lemma 8.1.1
Consider a differential equation of the form (2.9) and let Properties $P_1 - P_3$ hold. Let $z(t) = (x(t), y(t))$ denote a solution of (2.9). Then the following holds:

(i) $z'(t) := Rz(-t)$ is also a solution of (2.9).

(ii) If $z(0) = Rz(0)$ then $z(t) = Rz(-t)$.

(iii) For the special form of the reversibility operator $R$ given in Property $P_3$ condition $z(0) = Rz(0)$ is equivalent to $x(0) = 0$.

Proof:

(i) The claim follows from Property $P_3$:

$$z'(t) = -Rz(-t) = -RF(-t, z(-t), e) = F(t, Rz(-t), e) = F(t, z'(t), e).$$

(ii) If $z(0) = Rz(0)$ then $z(t)$ and $z'(t)$ fulfill the same initial value problem. Since the solution of the initial value problem is unique, we conclude that $z(t) = Rz(-t)$.

(iii) On one hand we have

$$Rz(0) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} -x(0) \\ y(0) \end{pmatrix}$$

and on the other hand

$$z(0) = \begin{pmatrix} x(0) \\ y(0) \end{pmatrix}.$$  

It follows that $Rz(0) = z(0)$ is equivalent to $x(0) = 0$. 

\qed
Lemma 8.1.2

Consider a differential equation of the form (2.9) and let Properties $P_1 - P_3$ hold. Let $z(t) = (x(t), y(t))$ denote a solution of (2.9) with the following symmetry property: $z(-t) = Rz(t)$. Then $z(t)$ is $2\pi$-periodic if and only if $x(\pi) = 0$.

Proof: Assuming $z$ to be $2\pi$-periodic we find

$$z(\pi + t) = z(-\pi + t) = Rz(\pi - t)$$

and therefore

$$x(\pi + t) = x(-\pi + t) = -x(\pi - t).$$

For $t := 0$ we obtain $x(\pi) = -x(\pi)$ and therefore $x(\pi) = 0$.

Now let $x(\pi) = 0$ hold. We define functions $x'$ and $y'$ by

$$x'(t) := x(t - 2\pi)$$
$$y'(t) := y(t - 2\pi).$$

Then we have

$$\dot{x}'(t) = \dot{x}(t - 2\pi) = F_1(t - 2\pi, x(t - 2\pi), y(t - 2\pi), e) = F_1(t, x'(t), y'(t), e),$$
$$\dot{y}'(t) = \dot{y}(t - 2\pi) = F_2(t - 2\pi, x(t - 2\pi), y(t - 2\pi), e) = F_2(t, x'(t), y'(t), e),$$

where we used the $2\pi$-periodicity of $F(t, x, y, e)$ in $t$.

Thus $(x', y')$ is a solution of (2.9). Moreover

$$x'(\pi) = x(\pi - 2\pi) = x(-\pi) = -x(\pi) = 0,$$
$$y'(\pi) = y(\pi - 2\pi) = y(-\pi) = y(\pi).$$

Since the solution of the initial value problem is unique, we conclude that $(x', y') = (x, y)$. Thus $(x, y)$ is $2\pi$-periodic.

We may summarize these two lemmas as follows:

$$z(t, z_0, e) = (x(t, (x_0, y_0), e), y(t, (x_0, y_0), e))$$

is a $2\pi$-periodic solution of (2.9) with the desired symmetry property if and only if $x(\pi, (0, y_0), e) = 0$.

Thus we are looking for zeros of $x(\pi, (0, y_0), e)$ as a function of $y_0$ and $e$. Figure 8.1 gives a survey.

We see that for $e < e^{l}_{\text{crit}} \approx 0.45$ we have at least three $2\pi$-periodic solutions. For $e = e^{l}_{\text{crit}}$ the Branches I and II merge in a saddle node bifurcation. For $e > e^{l}_{\text{crit}}$ only one $2\pi$-periodic solution (Branch III) persists.

Figures 8.2, 8.3 and 8.4 show the $2\pi$-periodic motion of the satellite corresponding to the Branches I, II and III.
8. Stability of the Dumbbell Satellite

Fig. 8.1: A plot of $x(\pi,(0,y_0),\epsilon)$. 

- $\epsilon = 0, y_0 = 0$
- $\epsilon = 0.1, y_0 \approx 0.08$
- $\epsilon = 0.2, y_0 = 0.14$
- $\epsilon = 0.3, y_0 = 0.21$
- $\epsilon = 0.4, y_0 = 0.39$
- $\epsilon = 0.44, y_0 = 0.60$

Fig. 8.2: The $2\pi$-periodic motion of the dumbbell satellite corresponding to Branch I.
8.1. $2\pi$-periodic Solutions

Fig. 8.3: The $2\pi$-periodic motion of the dumbbell satellite corresponding to Branch II.

Fig. 8.4: The $2\pi$-periodic motion of the dumbbell satellite corresponding to Branch III.
8.2 Linear Stability

Let $z^0_{per}$ denote a $2\pi$-periodic solution of (2.9) for some $e = e_0$. A necessary condition for the stability of $z^0_{per}$ is the stability of the equilibrium of the linearized system

$$\dot{z}(t) = A(t)z(t),$$

where

$$A(t) := D_z F(t, z^0_{per}(t), e_0).$$

---

**Fig. 8.5:** $k \cdot 2\pi$-periodic solutions ($k \in \mathbb{N}$) and linear stability.
In Figure 8.5 we present an overview on the initial values of $2\pi$-periodic solutions depending on the eccentricity $e$. The thick outlined curves indicate linear stability while the dashed curves indicate instability.

To understand the loss of stability on Branch I and III we show the initial values that lead to solutions of higher periods ($T = 2\pi \cdot k$, $k \in \mathbb{Z}$). Already for $16\pi$-periodic solutions it turned out to be impossible to present a complete survey. Due to this extremely complex situation the picture is inevitably incomplete.

8.3 The Stability of the Trivial Solution for $e = 0$

For $e = 0$ the system of differential equations (2.9) is reduced to

$$\begin{align*}
\dot{x} &= y \\
\dot{y} &= -\frac{3}{2} \sin(2x).
\end{align*}$$

Stability in the Sense of Lyapunov

In this section we show that the trivial periodic solution $(x_0, y_0) \equiv (0, 0)$ is stable in the sense of Lyapunov.

Lemma 8.3.1

Consider the system of differential equations (8.1). The following holds:

(i) The function

$$J(x, y) := \frac{1}{2}(3x^2 + y^2) - \frac{3}{4}(-1 + 2x^2 + \cos(2x)).$$

is a first integral of (8.1). Thus the orbits of (8.1) lie on the level curves of $J$ as shown in Figure 8.6.

(ii) Let $\rho_{\text{tol}}$ and $\rho_{\text{start}}$ be numbers such that

$$0 < \rho_{\text{tol}} \leq \sqrt{3},$$

$$\rho_{\text{start}} = \begin{cases} 
\frac{1}{2} \arccos(1 - \frac{2}{3} \rho_{\text{tol}}^2) & \text{for } \rho_{\text{tol}} \leq \sqrt{\frac{3}{2}(1 - \cos(2x_0))}, \\
\sqrt{x_0^2 + \rho_{\text{tol}}^2 + \frac{3}{2}(\cos(2x_0) - 1)} & \text{else,}
\end{cases}$$

where $x_0 \approx 1.3943$ is the smallest positive solution of $2x - 3\sin(2x) = 0$. Then the following holds:

For all solutions of (8.1) with

$$|(x(0), y(0))| \leq \rho_{\text{start}}$$
one has
\[ |(x(t), y(t))| \leq \rho_{\text{tol}}, \quad \text{for all } t > 0. \]

Fig. 8.6: The first integral \( J(x, y) \) with its level lines.

Proof:

(i) The derivative of \( J(x, y) \) reads
\[
\frac{d}{dt} J(x, y) = 3xx' + yy' - \frac{3}{4}(4x - 2\sin(2x))x'.
\]
Substituting the right-hand sides of the differential equations we find
\[
\frac{d}{dt} J(x, y) = 3xy + y \left( -\frac{\sqrt{3}}{2} \sin(2x) \right) - \frac{3}{4}(4x - 2\sin(2x))y = 0.
\]
Thus we have a first integral of (8.1).

(ii) Consider a circle around the origin with radius \( \rho_{\text{tol}} \). In a first step we determine the largest level curve \( J(x, y) = J_0 \) within this circle. In a second step we determine the largest circle around the origin within the level curve (cf. Figure 8.7). The radius of this circle is \( \rho_{\text{start}} \). Let \( 0 \leq J_0 \leq \frac{3}{2} \). Then the level curves \( J(x, y) = J_0 \) are closed curves that intersect the axes at \( x = \pm \frac{1}{2} \arccos \left( 1 - \frac{4}{3}J_0 \right) \) and \( y = \pm \sqrt{2J_0} \), respectively. For the square of the distance of a point \((x, y)\) on a level curve from the origin we obtain
\[
d^2(x) = x^2 + y^2 = x^2 + 2J_0 + \frac{3}{2}(\cos(2x) - 1).
\]
8.3. The Stability of the Trivial Solution for $c = 0$

\[
\frac{1}{2} (3x^2 + y^2) - \frac{3}{4} (-1 + 2x^2 + \cos(2x)) = J_0
\]

Fig. 8.7: The level curve $J(x, y) = J_0$ and the circles with radius $\rho_{\text{start}}$ and $\rho_{\text{tol}}$.

The derivatives read

\[
d^2(x)' = 2x - 3\sin(2x),
\]
\[
d^2(x)'' = 2 - 6\cos(2x).
\]

It follows that $d^2$ takes its maximum for $x = 0$: $d^2_{\text{max}} = 2J_0$. We conclude that $J(x, y) = J_0$ with $J_0^* := \frac{1}{2}\rho_{\text{tol}}^2$ is the largest level curve within the circle around the origin with radius $\rho_{\text{tol}}$.

For this level curve $d^2$ takes its minimum at

\[
x_{\text{min}} = \begin{cases} 
\frac{1}{2} \arccos \left( 1 - \frac{2}{3}\rho_{\text{tol}}^2 \right) & \text{for } \rho_{\text{tol}} \leq \sqrt{\frac{3}{2} (1 - \cos(2x_0))}, \\
\frac{\pi}{2} & \text{else,}
\end{cases}
\]

where $x_0$ is defined in the assumptions. Thus we have

\[
d^2(x_{\text{min}}) = \begin{cases} 
\frac{1}{4} \arccos^2 \left( 1 - \frac{2}{3}\rho_{\text{tol}}^2 \right) & \text{for } \rho_{\text{tol}} \leq \sqrt{\frac{3}{2} (1 - \cos(2x_0))}, \\
\frac{x_0^2 + \rho_{\text{tol}}^2 + \frac{3}{2} (\cos(2x_0) - 1)} & \text{else.}
\end{cases}
\]

Thus $d(x_{\text{min}})$ is equal to $\rho_{\text{start}}$ as defined in the assumptions. We conclude that the circle around the origin with radius $\rho_{\text{start}}$ is the largest circle within the level curve $J(x, y) = J_0^*$. This completes the proof. \(\square\)

The dependence of $\rho_{\text{start}}$ on $\rho_{\text{tol}}$ is illustrated in Table 8.1.


8. Stability of the Dumbbell Satellite

Tab. 8.1: $\rho_{\text{start}}$ for several $\rho_{\text{tol}}$.

<table>
<thead>
<tr>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{tol}}$</th>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{tol}}$</th>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{tol}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$5.8 \cdot 10^{-10}$</td>
<td>$10^{-9}$</td>
<td>$5.8 \cdot 10^{-9}$</td>
<td>$10^{-8}$</td>
<td>$0.0578$</td>
<td>$0.1$</td>
</tr>
<tr>
<td>$5.8 \cdot 10^{-8}$</td>
<td>$10^{-7}$</td>
<td>$5.8 \cdot 10^{-7}$</td>
<td>$10^{-6}$</td>
<td>$0.1157$</td>
<td>$0.2$</td>
</tr>
<tr>
<td>$5.8 \cdot 10^{-6}$</td>
<td>$10^{-5}$</td>
<td>$5.8 \cdot 10^{-5}$</td>
<td>$10^{-4}$</td>
<td>$0.1741$</td>
<td>$0.3$</td>
</tr>
<tr>
<td>$5.8 \cdot 10^{-4}$</td>
<td>$10^{-3}$</td>
<td>$5.8 \cdot 10^{-4}$</td>
<td>$10^{-3}$</td>
<td>$0.2330$</td>
<td>$0.4$</td>
</tr>
<tr>
<td>$5.8 \cdot 10^{-3}$</td>
<td>$10^{-2}$</td>
<td>$5.8 \cdot 10^{-3}$</td>
<td>$10^{-2}$</td>
<td>$0.3537$</td>
<td>$0.6$</td>
</tr>
<tr>
<td>$5.8 \cdot 10^{-2}$</td>
<td>$10^{-1}$</td>
<td>$5.8 \cdot 10^{-2}$</td>
<td>$10^{-1}$</td>
<td>$0.4160$</td>
<td>$0.8$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$5.8 \cdot 10^{-2}$</td>
<td>$10^{-1}$</td>
<td>$0.4801$</td>
<td>$1.0$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$5.8 \cdot 10^{-2}$</td>
<td>$10^{-1}$</td>
<td>$0.5456$</td>
<td>$1.2$</td>
</tr>
</tbody>
</table>

(a) $\rho_{\text{start}}$ for small $\rho_{\text{tol}}$.

(b) $\rho_{\text{start}}$ for large $\rho_{\text{tol}}$.

Practical Stability

In this section we derive results on the practical stability of the trivial solution of (8.1). The special case $e_0 = 0$ allows us to verify the results obtained by computer assistance.

For $e_0 = 0$ the series of coordinate transformations introduced in Chapter 4 is reduced to a linear transformation. The transformation matrix is determined only up to two parameters:

$$ S = \begin{pmatrix} s_{11} & s_{12} \\ -\sqrt{3}s_{12} & \sqrt{3}s_{11} \end{pmatrix}. $$

For the moment we choose $s_{11} = 1$ and $s_{12} = 0$. The influence of the particular choice of the transformation will be discussed in Section 8.6.

For the transformed system we obtain

$$ \begin{align*}
\dot{x} &= \sqrt{3}y \\
\dot{y} &= -\sqrt{3}x - \frac{\sqrt{3}}{2} (\sin(2x) - 2x).
\end{align*} \quad (8.2) $$

Lemma 8.3.2

Let

$$ \check{J}(x, y) := \frac{1}{3} J(S(x, y)) = \frac{1}{2} (x^2 + y^2) - \frac{1}{4} (-1 + 2x^2 + \cos(2x)) $$

Then $\check{J}(x, y)$ is a first integral of (8.2).

Proof: For the derivative of $\check{J}(x, y)$ we obtain

$$ \frac{d}{dt} \check{J}(x, y) = x\dot{x} + y\dot{y} - \frac{1}{2} (2x - \sin(2x)) \dot{x}. $$
8.3. The Stability of the Trivial Solution for $e = 0$

Substituting $\dot{x}$ and $\dot{y}$ by the right-hand sides of (8.2) we further find

$$
\frac{d}{dt} \dot{J}(x, y) = -\frac{\sqrt{3}}{2} y (\sin(2x) - 2x) + \frac{1}{2} (\sin(2x) - 2x) \sqrt{3} y = 0.
$$

Thus $\dot{J}(x, y)$ is a first integral, indeed.

In order to obtain results on practical stability we truncate the expansion of $\dot{J}(x, y)$ with respect to $x$ and $y$

$$
\dot{J}^{(r)}(x, y) = \frac{1}{2}(x^2 + y^2) + \frac{1}{4} \sum_{k=2}^{\lfloor \frac{r}{2} \rfloor} (-1)^k (2x)^{2k} (2k)!.
$$

This implies that the derivative no longer vanishes:

$$
\frac{d}{dt} \dot{J}^{(r)}(x, y) = \sum_{k=\lfloor \frac{r}{2} \rfloor + 1}^{\infty} j^k \neq 0.
$$

The polynomials $j^k$ are given by

$$
\dot{J}^k(x, y) = xh_1^{k-1} + yh_2^{k-1} + \sum_{l=3}^{\lfloor \frac{k}{2} \rfloor} \left( \frac{\partial \dot{J}^l}{\partial x} h_1^{k-l+1} + \frac{\partial \dot{J}^l}{\partial y} h_2^{k-l+1} \right)
$$

(cf. Lemma 6.4.1), where

$$
h_1^{2k+1}(x, y) = 0,
$$

$$
h_1^{2k}(x, y) = 0,
$$

$$
h_2^{2k+1}(x, y) = -\frac{\sqrt{3}}{2} (-1)^k (2x)^{2k+1} (2k + 1)!,
$$

$$
h_2^{2k}(x, y) = 0
$$

are the terms of the power series expansion of the right-hand side of (8.2).

Since $h_2^{2k} = 0$ and $\frac{\partial h_2^l}{\partial x} = \frac{\partial h_2^l}{\partial y} = 0$ if $l$ is odd and $h_1^{(2k+1)-l+1} = 0$ if $l$ is even, we have $j^{2k+1} = 0$. For $j^{2k}$ we obtain

$$
\dot{J}^{2k} = xh_1^{2k-1} + yh_2^{2k-1} + \sum_{l=2}^{\lfloor \frac{k}{2} \rfloor} \left( \frac{\partial \dot{J}^{2l}}{\partial x} h_1^{2k-2l+1} + \frac{\partial \dot{J}^{2l}}{\partial y} h_2^{2k-2l+1} \right) =
$$

$$
= -y \frac{\sqrt{3}}{2} (-1)^{k-1} (2x)^{2k-1} (2k - 1)! =
$$

$$
= \frac{\sqrt{3}}{2} (-1)^k y (2x)^{2k-1} (2k - 1)!.
$$
We compute the norms of $S$, $h^k$, $j^k$ and $\hat{j}^k$ using the definitions and lemmas given in Section 6.1. In our case the vector $R$ is reduced to a positive number. For the rest of this section we put $R = 1$. By this choice we obtain

$$ \|S\|_{R_z^{-1}} = 2, \quad \|S^{-1}\|_{R_x} = \frac{2}{\sqrt{3}},$$

$$ \|h^2_k\|_R = 0, \quad \|h^2_{k+1}\|_R = 0,$$

$$ \|j^2_k\|_R = \frac{1}{4} \frac{2^{2k}}{(2k)!} R^{2k}, \quad \|j^2_{k+1}\|_R = 0,$$

$$ \|j^2_k\|_R = \frac{\sqrt{2}}{2} \frac{2^{2k-1}}{(2k-1)!} R^{2k} \quad \|j^2_{k+1}\|_R = 0. \quad (8.3)$$

Let $\rho_{\text{start}} := 0.4\rho_{\text{tol}}$. In Table 8.2 we present a survey on the results obtained by Theorem 7.1.2 using the truncation of the first integral $\hat{j}^{(r)}$ together with the estimates given above.

**Tab. 8.2:** $T_{\text{max}}/2\pi$ for various values of $\rho_{\text{tol}}$ and the truncation order $r$ using the estimates given above.

<table>
<thead>
<tr>
<th>$\rho_{\text{tol}}$</th>
<th>$T_{\text{max}}^{(20)}/2\pi$</th>
<th>$T_{\text{max}}^{(40)}/2\pi$</th>
<th>$T_{\text{max}}^{(60)}/2\pi$</th>
<th>$T_{\text{max}}^{(80)}/2\pi$</th>
<th>$T_{\text{max}}^{(100)}/2\pi$</th>
<th>$T_{\text{max}}^{(120)}/2\pi$</th>
<th>$T_{\text{max}}^{(140)}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-9}$</td>
<td>$2.10^{198}$</td>
<td>$1.10^{408}$</td>
<td>$2.10^{622}$</td>
<td>$2.10^{839}$</td>
<td>$4.10^{1058}$</td>
<td>$3.10^{1279}$</td>
<td>$8.10^{1501}$</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>$2.10^{178}$</td>
<td>$1.10^{368}$</td>
<td>$2.10^{562}$</td>
<td>$2.10^{759}$</td>
<td>$4.10^{958}$</td>
<td>$3.10^{1159}$</td>
<td>$8.10^{1361}$</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>$2.10^{138}$</td>
<td>$1.10^{228}$</td>
<td>$2.10^{302}$</td>
<td>$2.10^{679}$</td>
<td>$4.10^{858}$</td>
<td>$3.10^{1039}$</td>
<td>$8.10^{1221}$</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>$2.10^{138}$</td>
<td>$1.10^{288}$</td>
<td>$2.10^{442}$</td>
<td>$2.10^{959}$</td>
<td>$4.10^{758}$</td>
<td>$3.10^{919}$</td>
<td>$8.10^{1081}$</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>$2.10^{118}$</td>
<td>$1.10^{248}$</td>
<td>$2.10^{382}$</td>
<td>$2.10^{519}$</td>
<td>$4.10^{658}$</td>
<td>$3.10^{799}$</td>
<td>$8.10^{941}$</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>$2.10^{108}$</td>
<td>$1.10^{208}$</td>
<td>$2.10^{322}$</td>
<td>$2.10^{439}$</td>
<td>$4.10^{558}$</td>
<td>$3.10^{679}$</td>
<td>$8.10^{801}$</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>$2.10^{78}$</td>
<td>$1.10^{168}$</td>
<td>$2.10^{262}$</td>
<td>$2.10^{359}$</td>
<td>$4.10^{458}$</td>
<td>$3.10^{559}$</td>
<td>$8.10^{661}$</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>$2.10^{58}$</td>
<td>$1.10^{128}$</td>
<td>$2.10^{202}$</td>
<td>$2.10^{279}$</td>
<td>$4.10^{358}$</td>
<td>$3.10^{439}$</td>
<td>$8.10^{521}$</td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>$2.10^{38}$</td>
<td>$1.10^{88}$</td>
<td>$2.10^{142}$</td>
<td>$2.10^{199}$</td>
<td>$4.10^{258}$</td>
<td>$3.10^{319}$</td>
<td>$8.10^{281}$</td>
</tr>
</tbody>
</table>

In Chapter 5 we introduced a method to construct an approximate first integral $I^{(r)}$ recursively. In Table 8.3 we provide results obtained by this method and compare them with the results obtained above.

The difference is due to the fact that the polynomials $I^{2k}$ of the approximate first integral obtained by explicit construction lie in the non-critical space $\Pi^{1_{\text{tol}}}_{2k}$ (cf. Lemma 5.3.6) while the polynomials $\hat{j}^{2k}$ do not, since from

$$ \hat{j}^{2k} = T(\hat{j}^{2k}) = T \left( \frac{1}{4} \frac{(2x)^{2k}}{(2k)!} \right) = \frac{1}{4} (-1)^k \frac{2^{1/2} (\xi + i\eta)^{2k}}{(2k)!} \quad (8.3)$$
8.3. The Stability of the Trivial Solution for $\epsilon = 0$

$$= \frac{1}{4} (-1)^k \frac{2^k}{(2k)!} \sum_{i=0}^{2k} \binom{2k}{i} (i\eta)^{2k-i}$$

it follows that $J^{2k}$ contains the critical polynomial:

$$\frac{1}{4} (-1)^k \frac{(2i)^k}{(2k)!} \binom{2k}{k} \xi^k \eta^k.$$

Since $J^2 = I^2 = i\xi \eta$ and $J^3 = I^3 = 0$ the polynomials $J^4$ and $I^4$ differ only in the critical term:

$$\tilde{J}^4 = T(J^4) = \frac{1}{24} \xi^4 + \frac{i}{6} \xi^3 \eta - \frac{1}{4} \xi^2 \eta^2 - \frac{i}{6} \xi \eta^3 + \frac{1}{24} \eta^4$$

$$\tilde{I}^4 = T(I^4) = \frac{1}{24} \xi^4 + \frac{i}{6} \xi^3 \eta - \frac{i}{6} \xi \eta^3 + \frac{1}{24} \eta^4.$$

If we transform both expressions back to $x$-$y$ coordinates we find

$$\tilde{J}^4 = T^{-1}(\tilde{J}^4) = \frac{1}{6} x^4,$$

$$\tilde{I}^4 = T^{-1}(\tilde{I}^4) = \frac{5}{48} x^4 - \frac{1}{8} x^2 y^2 - \frac{1}{24} y^4.$$

Thus the norms read

$$\|\tilde{J}^4\|_R = \frac{1}{6} \quad \text{and} \quad \|\tilde{I}^4\|_R = \frac{7}{24}.$$  

This suggests that the estimates obtained by $\tilde{J}^{(r)}$ are better than the estimates obtained by $I^{(r)}$.

**Tab. 8.3:** Comparison of $T_{\text{max}}/2\pi$ obtained by $I^{(r)}$ and $J^{(r)}$ for various values of $\rho_{\text{tol}}$.

<table>
<thead>
<tr>
<th>$\rho_{\text{tol}}$</th>
<th>$I^{(r)}$</th>
<th>$J^{(r)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-9}$</td>
<td>5. · 10^{185}</td>
<td>2. · 10^{198}</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>5. · 10^{165}</td>
<td>2. · 10^{178}</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>5. · 10^{145}</td>
<td>2. · 10^{158}</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>5. · 10^{125}</td>
<td>2. · 10^{138}</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>5. · 10^{105}</td>
<td>2. · 10^{118}</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>5. · 10^{05}</td>
<td>2. · 10^{08}</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>5. · 10^{01}</td>
<td>2. · 10^{07}</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>5. · 10^{0}</td>
<td>2. · 10^{05}</td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>5. · 10^{-25}</td>
<td>2. · 10^{38}</td>
</tr>
</tbody>
</table>
8.4 Practical Stability for Prescribed Values of $e$

In this section we present some results on the practical stability for prescribed values of the eccentricity $e$.

**Estimating the Perturbation**

By a series of transformations introduced in Chapter 4 the system of differential equations (2.9) is transformed to a system of perturbed harmonic oscillators

\[
\begin{align*}
\dot{x} &= \omega_1 y + h_1(t, x, y) \\
\dot{y} &= -\omega_1 x + h_2(t, x, y),
\end{align*}
\]

where the perturbations $h_1$ and $h_2$ on the right-hand side are at least of order 2 with respect to $x$ and $y$. The right-hand side of the transformed system of differential equations is given by

\[
\begin{pmatrix}
h_1(t, x, y) \\
h_2(t, x, y)
\end{pmatrix} = S^{-1} Q^{-1}(t) \begin{pmatrix} F_1(t, z^0_{\text{per}}(t) + Q(t) S z, e_0) \\ F_2(t, z^0_{\text{per}}(t) + Q(t) S z, e_0) \end{pmatrix},
\]

where

- $e_0$ is a fixed value of the eccentricity $e$,
- $z^0_{\text{per}} = (x_0, y_0)$ is a periodic solution of (2.9) for $e = e_0$.

Let $\tilde{Q}$ denote the product of the matrix of the Floquet transformation $Q$ with the matrix of the linear transformation $S$ that takes the monodromy matrix of the linearization of (2.9) to diagonal form and let us denote the elements of $\tilde{Q}(t)$ by $\tilde{q}_{ij}(t)$. With this notation we have

\[
\begin{pmatrix}
h_1(t, x, y) \\
h_2(t, x, y)
\end{pmatrix} = \tilde{Q}^{-1}(t) \left( \begin{pmatrix} y_0(t) + \tilde{q}_{21}(t)x + \tilde{q}_{22}(t)y & 0 \\ \frac{3 \sin(2x_0(t) + 2\tilde{q}_{11}(t)x + 2\tilde{q}_{12}(t)y)}{2} & \frac{e_0 \sin t}{1 - e_0 \cos t} \\ 0 & 0 \\ \frac{2e_0 \sqrt{1 - e_0^2 \sin t}}{(1 - e_0 \cos t)^2} & \frac{e_0 \sin t}{1 - e_0 \cos t} \end{pmatrix} \right). \tag{8.5}
\]

The power series expansions of $F_1(t, z^0_{\text{per}} + \tilde{Q} z, e_0)$ and $F_2(t, z^0_{\text{per}} + \tilde{Q} z, e_0)$ read

\[
F_1(t, z^0_{\text{per}} + \tilde{Q} z, e_0) = y_0 + \tilde{q}_{21} x + \tilde{q}_{22} y,
\]

\[
F_2(t, z^0_{\text{per}} + \tilde{Q} z, e_0) = \frac{e_0 \sin t}{1 - e_0 \cos t}.
\]
8.4. Practical Stability for Prescribed Values of \( \epsilon \)

\[
F_2(t, z_0 + \tilde{Q} z, e_0) = -\frac{3}{2} \frac{\sin(2x_0 + 2\tilde{q}_{11} x + 2\tilde{q}_{12} y)}{1 - \epsilon_0 \cos t} + \\
+ \left( y_0 + \tilde{q}_{21} x + \tilde{q}_{22} y \right) \frac{\epsilon_0 \sin t}{1 - \epsilon_0 \cos t} + \frac{2\epsilon_0 \sqrt{1 - \epsilon_0^2} \sin t}{(1 - \epsilon_0 \cos t)^2} = \\
= -\frac{3}{2} \left( \frac{\sin(2x_0)}{1 - \epsilon_0 \cos t} \cos(2\tilde{q}_{11} x + 2\tilde{q}_{12} y) + \\
\cdot \frac{\cos(2x_0)}{1 - \epsilon_0 \cos t} \sin(2\tilde{q}_{11} x + 2\tilde{q}_{12} y) \right) + \\
+ \left( y_0 + \tilde{q}_{21} x + \tilde{q}_{22} y \right) \frac{\epsilon_0 \sin t}{1 - \epsilon_0 \cos t} + \frac{2\epsilon_0 \sqrt{1 - \epsilon_0^2} \sin t}{(1 - \epsilon_0 \cos t)^2} = \\
= -\frac{3}{2} \frac{\sin(2x_0)}{1 - \epsilon_0 \cos t} \sum_{k=0}^{\infty} (-1)^k \frac{(2\tilde{q}_{11} x + 2\tilde{q}_{12} y)^{2k}}{(2k)!} - \\
\cdot \frac{\cos(2x_0)}{1 - \epsilon_0 \cos t} \sum_{k=0}^{\infty} (-1)^k \frac{(2\tilde{q}_{11} x + 2\tilde{q}_{12} y)^{2k+1}}{(2k+1)!} + \\
+ \left( y_0 + \tilde{q}_{21} x + \tilde{q}_{22} y \right) \frac{\epsilon_0 \sin t}{1 - \epsilon_0 \cos t} + \frac{2\epsilon_0 \sqrt{1 - \epsilon_0^2} \sin t}{(1 - \epsilon_0 \cos t)^2} = \\
= \left( -\frac{3}{2} \frac{\sin(2x_0)}{1 - \epsilon_0 \cos t} + y_0 \frac{\epsilon_0 \sin t}{1 - \epsilon_0 \cos t} + \frac{2\epsilon_0 \sqrt{1 - \epsilon_0^2} \sin t}{(1 - \epsilon_0 \cos t)^2} \right) + \\
+ \left( -\frac{3}{2} \frac{\cos(2x_0)}{1 - \epsilon_0 \cos t} (2\tilde{q}_{11} x + 2\tilde{q}_{12} y) + \frac{\epsilon_0 \sin t}{1 - \epsilon_0 \cos t} (\tilde{q}_{21} x + \tilde{q}_{22} y) \right) + \\
+ \left( -\frac{3}{2} \frac{\sin(2x_0)}{1 - \epsilon_0 \cos t} \sum_{k=1}^{\infty} (-1)^k \frac{(2\tilde{q}_{11} x + 2\tilde{q}_{12} y)^{2k}}{(2k)!} \right) + \\
+ \left( -\frac{3}{2} \frac{\cos(2x_0)}{1 - \epsilon_0 \cos t} \sum_{k=1}^{\infty} (-1)^k \frac{(2\tilde{q}_{11} x + 2\tilde{q}_{12} y)^{2k+1}}{(2k+1)!} \right).
\]

Since \((x_0, y_0)\) is a periodic solution of (2.9) the term of order 0 in \( x \) and \( y \) vanish. Similarly the term of order 1 is reduced to the linear part of (8.4) (cf. the proof of Theorem 4.2.1). Thus for \( k \geq 2 \) we have

\[
\left( \begin{array}{c} h_{2k}^{1,2}(t, x, y) \\
(h_{2k}^{2,2}(t, x, y)\end{array} \right) = \tilde{Q}^{-1}(t) \left( \begin{array}{c} 0 \\
\frac{3 \sin(2x_0(t))}{2} (-1)^k (2\tilde{q}_{11}(t)x + 2\tilde{q}_{12}(t)y)^{2k} \end{array} \right),
\]

\[
\left( h_{1}^{2k+1}(t, x, y), h_{2}^{2k+1}(t, x, y) \right) = Q^{-1}(t) \left( \begin{array}{c}
0 \\
-\frac{3 \cos(2\varphi_1(t))}{2} \sin(t) \left( -1 \right)^{k+1} \frac{(2\varphi_1(t)x + 2\varphi_2(t)y)^{2k+1}}{(2k+1)!}
\end{array} \right).
\]

In order to estimate the polynomials \( h_{1}^{k} \) and \( h_{2}^{k} \) we need a lemma.

**Lemma 8.4.1**

(i) Let \(|x| < 1\). Then the following holds:

\[
\frac{1}{(1-x)^m} = \sum_{k=0}^{\infty} \binom{m-1+k}{k} x^k.
\]

(ii) Let \(0 < e_0 < 1\). Then the following estimate holds:

\[
\left\| \frac{1}{(1-e_0 \cos t)^m} \right\|_\mathcal{F} \leq \frac{1}{(1-e_0)^m}.
\]

**Proof:**

(i) For \(|x| < 1\) we have

\[
\frac{1}{(1-x)^m} = \frac{1}{(m-1)!} \frac{d^{m-1}}{dx^{m-1}} \left( \frac{1}{1-x} \right) = \frac{1}{(m-1)!} \frac{d^{m-1}}{dx^{m-1}} \sum_{k=0}^{\infty} x^k =
\]

\[
= \frac{1}{(m-1)!} \sum_{k=m-1}^{\infty} \left( \prod_{j=m-k+2}^{k} j \right) x^{k-(m-1)} =
\]

\[
= \frac{1}{(m-1)!} \sum_{k=0}^{\infty} \left( \prod_{j=0}^{m-2} (k+1+j) \right) x^{k} =
\]

\[
= \sum_{k=0}^{\infty} \binom{m-1+k}{k} x^k.
\]

(ii) Using (i) we obtain

\[
\left\| \frac{1}{(1-e_0 \cos t)^m} \right\|_\mathcal{F} = \left\| \sum_{k=0}^{\infty} \binom{m-1+k}{k} e_0^k \cos^k t \right\|_\mathcal{F} =
\]

\[
= \left\| \sum_{k=0}^{\infty} \binom{m-1+k}{k} \left( e_0 \right)^k \sum_{l=0}^{k} \binom{k}{l} e^{i(2l-k)t} \right\|_\mathcal{F} \leq
\]

\[
\leq \sum_{k=0}^{\infty} \binom{m-1+k}{k} \left( e_0 \right)^k \sum_{l=0}^{k} \binom{k}{l} =
\]
= \sum_{k=0}^{\infty} \left( m - 1 + k \right) \frac{e^k}{k!}.

Now the claim follows immediately using (i). □

Using this lemma we find

\[ \|h^k_1\|_R \leq \frac{3}{2} \max_{1 \leq i, j \leq 2} \left\| (\phi^{-1})_{ij} \right\| \frac{\max(\|\sin(2x_0)\|_F, \|\cos(2x_0)\|_F) (2(\|q_{11}\|_F + \|q_{12}\|_F)R)}{1 - e_0} \frac{m^m}{m!}, \]

(8.6)

In a final step we determine constants \( E \) and \( h \) such that

\[ |h^k_{1,2}(t, x, y)| \leq Eh^k. \]

(8.7)

Let \( m \) be a positive integer. It is easy to see that \( 1/k! \) may be estimated by

\[ \frac{1}{k!} \leq \frac{m^{m-k}}{m!}, \]

where the equality-sign only holds for \( k = m - 1 \) and for \( k = m \). Thus an exponential estimate of the form (8.7) may be achieved by putting

\[ E := \frac{3}{2} \max_{1 \leq i, j \leq 2} \left\| (\phi^{-1})_{ij} \right\| \frac{\max(\|\sin(2x_0)\|_F, \|\cos(2x_0)\|_F)}{1 - e_0} \frac{m^m}{m!}, \]

(8.8)

\[ h := 2(\|q_{11}\|_F + \|q_{12}\|_F)\frac{R}{m}, \]

where \( m \) is a suitable positive integer.

**Practical Stability for Small Periodic Solutions**

In Section 8.2 we showed that there exist small linearly stable 2\( \pi \)-periodic solutions of (2.9) corresponding to Branch 1 in Figure 8.8.

In this section we investigate the stability of these periodic solutions by the method of explicitly constructed approximate first integrals (cf. Theorem 7.1.2).

**The Frequency \( \omega_1 \) as a Function of the Eccentricity \( e \)**

Figure 8.9 shows the frequency \( \omega_1 \) as a function of the eccentricity \( e \). The small icons depict the Floquet multipliers.

**On the Choice of the Norms**

In Section 6.1 we defined several norms. They depend on a vector \( R \) of positive numbers. In our case this vector is reduced to a number. For the subsequent computations we put \( R = 1 \).

For \( S \) and \( S^{-1} \) we use the norms given in Lemma 6.1.1.
Fig. 8.8: Linear Stability of $2\pi$-Periodic Solutions.

Fig. 8.9: The frequency as a function of the eccentricity $e$. 

$e_0 = e_{\text{crit}} \approx 0.35$
8.4. Practical Stability for Prescribed Values of $e$

On the Ratio $\rho_{\text{start}}/\rho_{\text{tol}}$

Our method for practical stability breaks down, if the numerator in the formula for the maximum time

$$T_{\text{max}} := \min_{1 \leq i \leq n} \left\{ \frac{\frac{1}{2} R^2 (\hat{\rho}_{\text{tol}}^2 - \hat{\rho}_{\text{start}}^2) - 2 \sum_{k=3}^{r} \| I_k \|_{R_{k}} \hat{\rho}_{\text{tol}}^{k} }{\sum_{k=r+1}^{\infty} \| I_k \|_{R_{k}} \hat{\rho}_{\text{tol}}^{k} } \right\}$$

(cf. Theorem 7.1.2) becomes negative. Let $q_{\text{break}}^{(r)}$ denote the ratio of $\rho_{\text{start}}$ and $\rho_{\text{tol}}$ where this break-down occurs.

For small values of $\hat{\rho}_{\text{tol}}$ we may neglect the sum $2 \sum_{k=3}^{r} \| I_k \|_{R_{k}} \hat{\rho}_{\text{tol}}^{k}$. Using the definition of $\hat{\rho}_{\text{start}}$ and $\hat{\rho}_{\text{tol}}$

$$\hat{\rho}_{\text{start}} := \| S^{-1} \|_{R_{5}} \rho_{\text{start}},$$

$$\hat{\rho}_{\text{tol}} := \| S^{-1} \|_{R_{5}} \rho_{\text{tol}}$$

(cf. Theorem 7.1.2) we find that the break-down occurs approximately for

$$\frac{\rho_{\text{start}}}{\rho_{\text{tol}}} \approx \frac{1}{\| S^{-1} \|_{R_{5}} \| S \|_{R_{5}}^{-1}} =: q_{\text{break}}^{\approx}.$$

For the definition of the norms see Section 6.1.

In Table 8.4 we compare the values of $q_{\text{break}}^{\approx}$ and $q_{\text{break}}^{(r)}$ for $R = 1$, $\rho_{\text{tol}} = 10^{-3}$ and $r = 20$.

Tab. 8.4: Comparison of $q_{\text{break}}^{\approx}$ and $q_{\text{break}}^{(20)}$ for $R = 1$ and several values of $e_0$.

<table>
<thead>
<tr>
<th>$e_0$</th>
<th>$q_{\text{break}}^{\approx}$</th>
<th>$q_{\text{break}}^{(20)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.433013</td>
<td>0.433013</td>
</tr>
<tr>
<td>0.05</td>
<td>0.425855</td>
<td>0.425794</td>
</tr>
<tr>
<td>0.1</td>
<td>0.418722</td>
<td>0.418561</td>
</tr>
<tr>
<td>0.15</td>
<td>0.412819</td>
<td>0.412494</td>
</tr>
<tr>
<td>0.2</td>
<td>0.410196</td>
<td>0.409546</td>
</tr>
<tr>
<td>0.25</td>
<td>0.415058</td>
<td>0.411974</td>
</tr>
<tr>
<td>0.3</td>
<td>0.439918</td>
<td>0.436215</td>
</tr>
<tr>
<td>0.35</td>
<td>0.430093</td>
<td>0.410662</td>
</tr>
</tbody>
</table>

Estimates of the Perturbation

For our computations we need the norms of the functions $h_{1,2}^k$. For small order $k$ we use the norms obtained by explicit computations. For larger $k$ we make use of the estimates (8.6) given above, where for Nekoroshev-type results we need exponential estimates (8.7). Figures 8.10 and 8.11 give a comparison of the explicitly computed norms with the estimates (8.6) and (8.7), respectively.
8. Stability of the Dumbbell Satellite

Computer Assisted Results

We are now in the position to present some results based on explicit computations performed on a computer algebra system.

According to Theorem 7.1.2 the maximum time $T_{\text{max}}$ is defined by

$$T_{\text{max}} := \min_{1 \leq i \leq n} \frac{1}{2} R_i^2 (\bar{\rho}_{\text{tol}}^2 - \bar{\rho}_{\text{start}}^2) - 2 \sum_{k=3}^{r} \| f_i^k \|_{R} \bar{\rho}_{\text{tol}}^k.$$  (7.2)

While we may compute the numerator explicitely we must use a mix of explicit computations and exponential estimates for the denominator (cf. Section 7.2). In this situation it turns out that the exponential estimates (8.7) provide the best results for $m := r$, where $r$ is the order of the approximate first integral.

Let $\rho_{\text{start}} := 0.4 \rho_{\text{tol}}$ and let the truncation order be $r = 20$. Table 8.5 and Figures 8.12 and 8.13 gives a survey of $T_{\text{max}}/2\pi$ as a function of $\rho_{\text{tol}}$ for various values of the eccentricity $\epsilon_0$. The figures suggest that $T_{\text{max}}/2\pi$ depends exponentially on $\rho_{\text{tol}}$. 

Fig. 8.10: Comparing the norms of $h_{1,2}^k$ with the estimates (8.6) for several values of $\epsilon_0$. 

![Graph showing the norms of $h_{1,2}^k$ and the estimates with different values of $\epsilon_0$.]
8.4. Practical Stability for Prescribed Values of $e$

Fig. 8.11: Comparing the norms of $h_{1,2}^k$ with the estimates (8.7) using (8.8) with $m = r$ for several values of $e_0$.

Tab. 8.5: Dependence of $T_{\text{max}}/2\pi$ on $\rho_{\text{tol}}$ for various values of $e_0$.

<table>
<thead>
<tr>
<th>$\rho_{\text{tol}}$</th>
<th>0</th>
<th>0.05</th>
<th>0.1</th>
<th>0.15</th>
<th>0.2</th>
<th>0.25</th>
<th>0.3</th>
<th>0.35</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_0$</td>
<td>$10^{-9}$</td>
<td>$10^{-8}$</td>
<td>$10^{-7}$</td>
<td>$10^{-6}$</td>
<td>$10^{-5}$</td>
<td>$10^{-4}$</td>
<td>$10^{-3}$</td>
<td>$10^{-2}$</td>
</tr>
<tr>
<td>0.00</td>
<td>$8 \cdot 10^{184}$</td>
<td>$3 \cdot 10^{166}$</td>
<td>$4 \cdot 10^{156}$</td>
<td>$4 \cdot 10^{156}$</td>
<td>$4 \cdot 10^{152}$</td>
<td>$4 \cdot 10^{153}$</td>
<td>$4 \cdot 10^{144}$</td>
<td>$4 \cdot 10^{144}$</td>
</tr>
<tr>
<td>0.05</td>
<td>$8 \cdot 10^{164}$</td>
<td>$3 \cdot 10^{147}$</td>
<td>$4 \cdot 10^{137}$</td>
<td>$4 \cdot 10^{141}$</td>
<td>$4 \cdot 10^{133}$</td>
<td>$4 \cdot 10^{134}$</td>
<td>$4 \cdot 10^{125}$</td>
<td>$4 \cdot 10^{125}$</td>
</tr>
<tr>
<td>0.1</td>
<td>$8 \cdot 10^{144}$</td>
<td>$3 \cdot 10^{128}$</td>
<td>$4 \cdot 10^{118}$</td>
<td>$4 \cdot 10^{122}$</td>
<td>$4 \cdot 10^{114}$</td>
<td>$4 \cdot 10^{115}$</td>
<td>$4 \cdot 10^{106}$</td>
<td>$4 \cdot 10^{106}$</td>
</tr>
<tr>
<td>0.15</td>
<td>$8 \cdot 10^{124}$</td>
<td>$3 \cdot 10^{109}$</td>
<td>$4 \cdot 10^{100}$</td>
<td>$4 \cdot 10^{104}$</td>
<td>$4 \cdot 10^{95}$</td>
<td>$4 \cdot 10^{96}$</td>
<td>$4 \cdot 10^{87}$</td>
<td>$4 \cdot 10^{87}$</td>
</tr>
<tr>
<td>0.2</td>
<td>$8 \cdot 10^{104}$</td>
<td>$3 \cdot 10^{090}$</td>
<td>$4 \cdot 10^{080}$</td>
<td>$4 \cdot 10^{090}$</td>
<td>$4 \cdot 10^{076}$</td>
<td>$4 \cdot 10^{084}$</td>
<td>$4 \cdot 10^{067}$</td>
<td>$4 \cdot 10^{067}$</td>
</tr>
<tr>
<td>0.25</td>
<td>$8 \cdot 10^{084}$</td>
<td>$3 \cdot 10^{071}$</td>
<td>$4 \cdot 10^{061}$</td>
<td>$4 \cdot 10^{065}$</td>
<td>$4 \cdot 10^{057}$</td>
<td>$4 \cdot 10^{058}$</td>
<td>$4 \cdot 10^{049}$</td>
<td>$4 \cdot 10^{049}$</td>
</tr>
<tr>
<td>0.3</td>
<td>$8 \cdot 10^{064}$</td>
<td>$3 \cdot 10^{052}$</td>
<td>$4 \cdot 10^{042}$</td>
<td>$4 \cdot 10^{046}$</td>
<td>$4 \cdot 10^{038}$</td>
<td>$4 \cdot 10^{039}$</td>
<td>$4 \cdot 10^{031}$</td>
<td>$4 \cdot 10^{031}$</td>
</tr>
<tr>
<td>0.35</td>
<td>$8 \cdot 10^{044}$</td>
<td>$3 \cdot 10^{033}$</td>
<td>$4 \cdot 10^{023}$</td>
<td>$4 \cdot 10^{027}$</td>
<td>$4 \cdot 10^{027}$</td>
<td>$4 \cdot 10^{027}$</td>
<td>$4 \cdot 10^{027}$</td>
<td>$4 \cdot 10^{027}$</td>
</tr>
</tbody>
</table>
8. Stability of the Dumbbell Satellite

Fig. 8.12: Dependence of $T_{\text{max}}/2\pi$ on $\rho_{\text{tol}}$ for various values of $e_0$.

Fig. 8.13: A magnification of Fig. 8.12 for large values of $\rho_{\text{tol}}$. 
To obtain an idea of the values presented in Table 8.5, we consider a satellite on an orbit of eccentricity $e = 0.1$ around the earth. The values $\rho_{\text{start}} = 0.04$ and $\rho_{\text{col}} = 0.1$ correspond to angles of about $\theta_{\text{start}} \approx \pm 2.3^\circ$ and $\theta_{\text{max}} \approx \pm 5.7^\circ$, respectively. Assuming the satellite to need one and a half hour for one revolution around the earth, $T_{\text{max}}/2\pi = 3 \cdot 10^{11}$ corresponds to roughly 50 million years. The situation is illustrated in Figure 8.14.

![Figure 8.14: The motion of the satellite around its center of mass.](image)

**Tab. 8.6:** Dependence of $T_{\text{max}}/2\pi$ on the truncation order $r$ for $e_0 = 0$. 

<table>
<thead>
<tr>
<th>$\rho_{\text{col}}$</th>
<th>Truncation order $r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10(^{-9})</td>
<td>4 6 8 10 12 14 16 18 20</td>
</tr>
<tr>
<td>5 \cdot 10^{35}</td>
<td>4 \cdot 10^{73} 1 \cdot 10^{92} 9 \cdot 10^{110} 2 \cdot 10^{129}</td>
</tr>
<tr>
<td>4 \cdot 10^{48}</td>
<td>4 \cdot 10^{65} 9 \cdot 10^{98} 2 \cdot 10^{115}</td>
</tr>
<tr>
<td>10(^{-8})</td>
<td>5 \cdot 10^{31} 4 \cdot 10^{42} 9 \cdot 10^{57} 2 \cdot 10^{101}</td>
</tr>
<tr>
<td>10(^{-7})</td>
<td>5 \cdot 10^{27} 4 \cdot 10^{49} 9 \cdot 10^{72} 2 \cdot 10^{107}</td>
</tr>
<tr>
<td>10(^{-6})</td>
<td>5 \cdot 10^{23} 4 \cdot 10^{36} 9 \cdot 10^{62} 2 \cdot 10^{113}</td>
</tr>
<tr>
<td>10(^{-5})</td>
<td>5 \cdot 10^{19} 4 \cdot 10^{30} 9 \cdot 10^{52} 2 \cdot 10^{119}</td>
</tr>
<tr>
<td>10(^{-4})</td>
<td>5 \cdot 10^{15} 4 \cdot 10^{24} 9 \cdot 10^{42} 2 \cdot 10^{125}</td>
</tr>
<tr>
<td>10(^{-3})</td>
<td>5 \cdot 10^{11} 4 \cdot 10^{18} 9 \cdot 10^{32} 2 \cdot 10^{131}</td>
</tr>
<tr>
<td>10(^{-2})</td>
<td>5 \cdot 10^{7} 4 \cdot 10^{12} 9 \cdot 10^{22} 2 \cdot 10^{137}</td>
</tr>
<tr>
<td>10(^{-1})</td>
<td>5 \cdot 10^{1} 4 \cdot 10^{6} 9 \cdot 10^{12} 2 \cdot 10^{143}</td>
</tr>
</tbody>
</table>
Let still $\rho_{\text{start}} := 0.4 \rho_{\text{tol}}$ hold. Tables 8.6, 8.7 and 8.8 give a survey on $T_{\text{max}}/2\pi$ as a function of $\rho_{\text{tol}}$ for various values of the truncation order $r$. In Figures 8.15, 8.16 and 8.17 these results are presented graphically.

**Tab. 8.7:** Dependence of $T_{\text{max}}/2\pi$ on the truncation order $r$ for $\epsilon_0 = 0.2$.

<table>
<thead>
<tr>
<th>$\rho_{\text{tol}}$</th>
<th>$r$</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>14</th>
<th>16</th>
<th>18</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-9}$</td>
<td></td>
<td>$5 \cdot 10^{24}$</td>
<td>$1 \cdot 10^{49}$</td>
<td>$9 \cdot 10^{57}$</td>
<td>$8 \cdot 10^{62}$</td>
<td>$6 \cdot 10^{89}$</td>
<td>$5 \cdot 10^{126}$</td>
<td>$4 \cdot 10^{143}$</td>
<td>$4 \cdot 10^{160}$</td>
<td></td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td></td>
<td>$5 \cdot 10^{21}$</td>
<td>$1 \cdot 10^{37}$</td>
<td>$9 \cdot 10^{56}$</td>
<td>$8 \cdot 10^{61}$</td>
<td>$6 \cdot 10^{86}$</td>
<td>$5 \cdot 10^{111}$</td>
<td>$4 \cdot 10^{126}$</td>
<td>$4 \cdot 10^{141}$</td>
<td></td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td></td>
<td>$5 \cdot 10^{18}$</td>
<td>$1 \cdot 10^{45}$</td>
<td>$9 \cdot 10^{55}$</td>
<td>$8 \cdot 10^{60}$</td>
<td>$6 \cdot 10^{83}$</td>
<td>$5 \cdot 10^{106}$</td>
<td>$4 \cdot 10^{122}$</td>
<td>$4 \cdot 10^{137}$</td>
<td></td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td></td>
<td>$5 \cdot 10^{15}$</td>
<td>$1 \cdot 10^{38}$</td>
<td>$9 \cdot 10^{48}$</td>
<td>$8 \cdot 10^{59}$</td>
<td>$6 \cdot 10^{79}$</td>
<td>$5 \cdot 10^{91}$</td>
<td>$4 \cdot 10^{102}$</td>
<td>$4 \cdot 10^{113}$</td>
<td></td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td></td>
<td>$5 \cdot 10^{12}$</td>
<td>$1 \cdot 10^{31}$</td>
<td>$9 \cdot 10^{40}$</td>
<td>$8 \cdot 10^{57}$</td>
<td>$6 \cdot 10^{70}$</td>
<td>$5 \cdot 10^{81}$</td>
<td>$4 \cdot 10^{92}$</td>
<td>$4 \cdot 10^{103}$</td>
<td></td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td></td>
<td>$5 \cdot 10^{9}$</td>
<td>$1 \cdot 10^{24}$</td>
<td>$9 \cdot 10^{30}$</td>
<td>$8 \cdot 10^{55}$</td>
<td>$6 \cdot 10^{64}$</td>
<td>$5 \cdot 10^{84}$</td>
<td>$4 \cdot 10^{95}$</td>
<td>$4 \cdot 10^{104}$</td>
<td></td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td></td>
<td>$5 \cdot 10^{6}$</td>
<td>$1 \cdot 10^{17}$</td>
<td>$9 \cdot 10^{21}$</td>
<td>$7 \cdot 10^{56}$</td>
<td>$6 \cdot 10^{64}$</td>
<td>$5 \cdot 10^{86}$</td>
<td>$4 \cdot 10^{101}$</td>
<td>$4 \cdot 10^{111}$</td>
<td></td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td></td>
<td>$2 \cdot 10^{3}$</td>
<td>$4 \cdot 10^{9}$</td>
<td>$4 \cdot 10^{12}$</td>
<td>$3 \cdot 10^{15}$</td>
<td>$2 \cdot 10^{18}$</td>
<td>$2 \cdot 10^{21}$</td>
<td>$2 \cdot 10^{24}$</td>
<td>$1 \cdot 10^{27}$</td>
<td></td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tab. 8.8:** Dependence of $T_{\text{max}}/2\pi$ on the truncation order $r$ for $\epsilon_0 = 0.35$.

<table>
<thead>
<tr>
<th>$\rho_{\text{tol}}$</th>
<th>$r$</th>
<th>4</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>14</th>
<th>16</th>
<th>18</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-9}$</td>
<td></td>
<td>$2 \cdot 10^{21}$</td>
<td>$1 \cdot 10^{37}$</td>
<td>$1 \cdot 10^{53}$</td>
<td>$6 \cdot 10^{68}$</td>
<td>$3 \cdot 10^{84}$</td>
<td>$2 \cdot 10^{99}$</td>
<td>$2 \cdot 10^{114}$</td>
<td>$1 \cdot 10^{129}$</td>
<td>$7 \cdot 10^{143}$</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td></td>
<td>$2 \cdot 10^{18}$</td>
<td>$1 \cdot 10^{32}$</td>
<td>$1 \cdot 10^{46}$</td>
<td>$6 \cdot 10^{59}$</td>
<td>$3 \cdot 10^{73}$</td>
<td>$2 \cdot 10^{86}$</td>
<td>$2 \cdot 10^{99}$</td>
<td>$1 \cdot 10^{112}$</td>
<td>$7 \cdot 10^{124}$</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td></td>
<td>$2 \cdot 10^{15}$</td>
<td>$1 \cdot 10^{27}$</td>
<td>$1 \cdot 10^{39}$</td>
<td>$6 \cdot 10^{50}$</td>
<td>$3 \cdot 10^{62}$</td>
<td>$2 \cdot 10^{73}$</td>
<td>$2 \cdot 10^{84}$</td>
<td>$1 \cdot 10^{95}$</td>
<td>$7 \cdot 10^{105}$</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td></td>
<td>$2 \cdot 10^{12}$</td>
<td>$1 \cdot 10^{22}$</td>
<td>$1 \cdot 10^{32}$</td>
<td>$6 \cdot 10^{41}$</td>
<td>$3 \cdot 10^{51}$</td>
<td>$2 \cdot 10^{60}$</td>
<td>$2 \cdot 10^{60}$</td>
<td>$1 \cdot 10^{78}$</td>
<td>$7 \cdot 10^{86}$</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td></td>
<td>$2 \cdot 10^{9}$</td>
<td>$1 \cdot 10^{17}$</td>
<td>$1 \cdot 10^{25}$</td>
<td>$6 \cdot 10^{32}$</td>
<td>$3 \cdot 10^{10}$</td>
<td>$2 \cdot 10^{47}$</td>
<td>$2 \cdot 10^{54}$</td>
<td>$1 \cdot 10^{61}$</td>
<td>$7 \cdot 10^{67}$</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td></td>
<td>$2 \cdot 10^{6}$</td>
<td>$1 \cdot 10^{12}$</td>
<td>$1 \cdot 10^{18}$</td>
<td>$5 \cdot 10^{23}$</td>
<td>$3 \cdot 10^{29}$</td>
<td>$1 \cdot 10^{34}$</td>
<td>$2 \cdot 10^{39}$</td>
<td>$1 \cdot 10^{44}$</td>
<td>$6 \cdot 10^{48}$</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
8.4. Practical Stability for Prescribed Values of $e$

Fig. 8.15: Dependence of $T_{\text{max}}/2\pi$ on $\rho_{\text{tol}}$ for $e_0 = 0$ and for various values of the truncation order $r$.

Fig. 8.16: Dependence of $T_{\text{max}}/2\pi$ on $\rho_{\text{tol}}$ for $e_0 = 0.2$ and for various values of the truncation order $r$. 
Fig. 8.17: Dependence of $T_{\text{max}}/2\pi$ on $\rho_{\text{tol}}$ for $e_0 = 0.35$ and for various values of the truncation order $r$.

Figures 8.18, 8.19, 8.20 give a survey on $T_{\text{max}}/2\pi$ as a function of the truncation order $r$ for various values of $\rho_{\text{tol}}$.

Fig. 8.18: Dependence of $T_{\text{max}}/2\pi$ on the truncation order $r$ for $e_0 = 0$ and various values of $\rho_{\text{tol}}$. 
8.4. Practical Stability for Prescribed Values of \( \epsilon \)

---

**Fig. 8.19:** Dependence of \( T_{\text{max}}/2\pi \) on the truncation order \( r \) for \( \epsilon_0 = 0.2 \) and various values of \( \rho_{\text{tol}} \).

---

**Fig. 8.20:** Dependence of \( T_{\text{max}}/2\pi \) on the truncation order \( r \) for \( \epsilon_0 = 0.35 \) and various values of \( \rho_{\text{tol}} \).
Practical Stability for Large Periodic Solutions

In this section we consider the $2\pi$-periodic solutions corresponding to Branch III in Figure 8.8. They are linearly stable for $0 < \epsilon_0 < \epsilon_{\text{crit}}^{I} \approx 0.045$. Figure 8.21 shows the frequency $\omega_1$ as a function of the eccentricity $\epsilon$. The small icons depict the Floquet multipliers.

Fig. 8.21: The frequency as a function of the eccentricity.

Compared to the situation on Branch I the break-down of our method occurs for a significantly smaller ratio $q_{\text{break}}^{(r)} := \rho_{\text{start}} / \rho_{\text{col}}$. We have to choose $q_{\text{break}}^{(r)} := 0.03$ instead of $q_{\text{break}}^{(r)} := 0.4$.

In Figure 8.22 we compare the norms of the functions $h_k^0$ with their exponential estimates.

<table>
<thead>
<tr>
<th>$\rho_{\text{col}}$</th>
<th>$\epsilon_0 = 0.01$</th>
<th>$\epsilon_0 = 0.02$</th>
<th>$\epsilon_0 = 0.03$</th>
<th>$\epsilon_0 = 0.04$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-9}$</td>
<td>$7 \cdot 10^{38}$</td>
<td>$4 \cdot 10^{35}$</td>
<td>$6 \cdot 10^{37}$</td>
<td>$9 \cdot 10^{29}$</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>$7 \cdot 10^{19}$</td>
<td>$4 \cdot 10^{20}$</td>
<td>$6 \cdot 10^{18}$</td>
<td>$9 \cdot 10^{10}$</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>$7 \cdot 10^{100}$</td>
<td>$4 \cdot 10^{101}$</td>
<td>$6 \cdot 10^{99}$</td>
<td>$9 \cdot 10^{91}$</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>$7 \cdot 10^{81}$</td>
<td>$4 \cdot 10^{82}$</td>
<td>$6 \cdot 10^{80}$</td>
<td>$9 \cdot 10^{72}$</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>$7 \cdot 10^{62}$</td>
<td>$4 \cdot 10^{63}$</td>
<td>$6 \cdot 10^{61}$</td>
<td>$9 \cdot 10^{53}$</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>$6 \cdot 10^{43}$</td>
<td>$4 \cdot 10^{44}$</td>
<td>$6 \cdot 10^{42}$</td>
<td>$8 \cdot 10^{34}$</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>$4 \cdot 10^{24}$</td>
<td>$3 \cdot 10^{25}$</td>
<td>$5 \cdot 10^{23}$</td>
<td>$3 \cdot 10^{15}$</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
<td>$-$</td>
</tr>
</tbody>
</table>

In Table 8.9 and Figure 8.23 we present results obtained by explicitly constructed approximate first integrals (cf. (7.1.1)). The truncation order is $r = 20$. Note that the times $T_{\text{max}} / 2\pi$ are significantly smaller compared to those obtained for Branch I.
8.4. Practical Stability for Prescribed Values of $c$

Nekoroshev-type Results

We close the section on practical stability for prescribed values of $c_0$ with some Nekoroshev-type results.

On the Ratio $\rho_{\text{start}}/\rho_{\text{tol}}$

Our method for practical stability breaks down, if the numerator in the formula for the maximum time

$$T_{\text{max}} := \frac{\tilde{\rho}_{\text{tol}}^2 - \tilde{\rho}_{\text{start}}^2 - 80 \cdot 3^\tau \hat{\rho}_{\text{tol}}}{24 \gamma \hat{\rho} \frac{3}{2} \hat{\rho}_{\text{tol}}} \left( \frac{2}{e^2} \right)^{1+\frac{1}{1+\tau}} \exp \left( 1 + \tau \left( \frac{\tilde{\rho}_*}{\tilde{\rho}_{\text{tol}}} \right)^{1+\frac{1}{1+\tau}} \right)$$

(cf. Theorem 7.3.2) becomes negative. Let $q^{(r)}_{\text{break}}$ denote the ratio of $\rho_{\text{start}}$ and $\rho_{\text{tol}}$ where this break-down occurs.

For $\tilde{\rho}_{\text{tol}}$ small compared with $\tilde{\rho}_*$ we may neglect the term $80 \cdot 3^\tau \hat{\rho}_{\text{tol}}$. Using the definition of $\tilde{\rho}_{\text{start}}$ and $\tilde{\rho}_{\text{tol}}$

$$\tilde{\rho}_{\text{start}} := \|S^{-1}\|_{R_c} \cdot \rho_{\text{start}},$$
$$\tilde{\rho}_{\text{tol}} := \|S^\tau\|^{-1}_{R_c} \cdot \rho_{\text{tol}}$$

Fig. 8.22: Comparing the norms of $h_{1,2}^k$ with the estimates (8.7) using (8.8) for several values of $c_0$. 
(cf. Theorem 7.3.2) we find that the break-down occurs approximately for

\[ \frac{\rho_{\text{start}}}{\rho_{\text{tol}}} \approx \frac{1}{\| S^{-1} \| R_e \| S \| R_e^{-1}} =: q_{\text{break}}^{\text{approx}} \]

In Table 8.10 we compare the values of \( q_{\text{break}}^{\text{approx}} \) and \( q_{\text{break}}^{(r)} \).

**Estimates of the Perturbation**

The exponential estimates (8.7) allow us also to state some Nekoroshey-type results using Theorem 7.3.2. In this situation it turns out that the exponential estimates (8.7) provide the best results for \( m := 2 \).

**On the Ratio \( \omega_1/\omega_0 \)**

Since the quotient \( \omega_1/\omega_0 \) is known only with finite precision, we treat \( \omega_1/\omega_0 \) as a rational number:

\[ \frac{\omega_1}{\omega_0} = \frac{p_N}{q_N} \]

In Table 8.11 we give a survey on \( \omega_0 \), \( \omega_1 \) and \( p_N/q_N \) for several values of \( \varepsilon_0 \).
Tab. 8.10: Comparison of $q_{\text{break}}^{\text{approx}}$ for several values of $\rho_{\text{tol}}$.

<table>
<thead>
<tr>
<th>$\rho_{\text{tol}}$</th>
<th>$q_{\text{break}}^{(r)}$</th>
<th>$q_{\text{break}}^{\text{approx}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-18}$</td>
<td>0.425855</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-17}$</td>
<td>0.425855</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-16}$</td>
<td>0.425855</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-15}$</td>
<td>0.425855</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-14}$</td>
<td>0.425855</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-13}$</td>
<td>0.425855</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-12}$</td>
<td>0.425855</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-11}$</td>
<td>0.425850</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-10}$</td>
<td>0.425802</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-9}$</td>
<td>0.425319</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>0.420496</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>0.372265</td>
<td>0.425855</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>0.037040</td>
<td>0.425855</td>
</tr>
</tbody>
</table>

(a) Comparison for Branch I and $e_0 = 0.05$.

(b) Comparison for Branch III and $e_0 = 0.01$.

Tab. 8.11: Survey on $\omega_0$, $\omega_1$ and $p_N/q_N$ for several values of $e_0$.

<table>
<thead>
<tr>
<th>$e_0$</th>
<th>$\omega_0$</th>
<th>$\omega_1$</th>
<th>$p_N$</th>
<th>$q_N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1.730449633063415</td>
<td>0.153137241617951</td>
<td>1.7309937</td>
<td>22560674</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>1.725498196373355</td>
<td>0.227624814353213</td>
<td>22560574</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>1.715615417640824</td>
<td>0.297630905260292</td>
<td>22183591</td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td>1.699351302779641</td>
<td>0.380987683333924</td>
<td>22433264</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>1.673000071838066</td>
<td>0.38023566</td>
<td>70540359</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>1.62934581764356</td>
<td>0.3996677</td>
<td>5796686</td>
</tr>
<tr>
<td></td>
<td>0.35</td>
<td>1.532312690754799</td>
<td>0.4770014</td>
<td>19476089</td>
</tr>
</tbody>
</table>

(a) Branch I.

(b) Branch III.

From the theory of continued fraction (cf. [11]) we know that for $\omega_1/\omega_0$ there is a finite series of convergents

$$\frac{p_0}{q_0}, \frac{p_1}{q_1}, \ldots, \frac{p_N}{q_N} = \frac{\omega_1}{\omega_0}$$

for which the following holds:

$$|l_0\omega_0 + l_1\omega_1| = \omega_0 \left| l_0 + l_1 \frac{\omega_1}{\omega_0} \right| \geq \omega_0 \left| p_n + q_n \frac{\omega_1}{\omega_0} \right|,$$

for $l_0 \in \mathbb{N}$, $0 < l_1 \leq q_n$. 

Thus we put
\[ \alpha_k := \omega_0 \left| p_n + q_n \frac{\omega_1}{\omega_0} \right| \quad \text{for} \quad q_{n-1} < k \leq q_n. \]

It remains to determine the constants \( \gamma \) and \( \tau \) such that
\[ \alpha_k \geq \gamma k^{-\tau}. \]

From the theory of finite continued fraction we know that we may choose \( \tau := 1 \). For \( \gamma \) put
\[ \gamma := \min_{1 \leq n \leq N} \alpha_{q_n} q_n^\tau. \]

Table 8.12: Comparison of \( \alpha_k \) and \( \gamma k^{-\tau} \) for \( e_0 = 0.05 \) on Branch I and \( k = q_n \leq q_N \).

<table>
<thead>
<tr>
<th>( k )</th>
<th>( \alpha_k )</th>
<th>( k\alpha_k )</th>
<th>( \gamma k^{-\tau} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3.49460 \cdot 10^{-2}</td>
<td>0.1398</td>
<td>3.91124 \cdot 10^{-3}</td>
</tr>
<tr>
<td>11</td>
<td>8.30954 \cdot 10^{-3}</td>
<td>0.0914</td>
<td>1.42227 \cdot 10^{-3}</td>
</tr>
<tr>
<td>26</td>
<td>1.70780 \cdot 10^{-3}</td>
<td>0.0444</td>
<td>6.01729 \cdot 10^{-4}</td>
</tr>
<tr>
<td>115</td>
<td>1.47833 \cdot 10^{-3}</td>
<td>0.1700</td>
<td>1.36043 \cdot 10^{-4}</td>
</tr>
<tr>
<td>486</td>
<td>2.29471 \cdot 10^{-4}</td>
<td>0.1115</td>
<td>3.21913 \cdot 10^{-5}</td>
</tr>
<tr>
<td>601</td>
<td>1.01505 \cdot 10^{-4}</td>
<td>0.0610</td>
<td>2.60316 \cdot 10^{-5}</td>
</tr>
<tr>
<td>4092</td>
<td>2.64621 \cdot 10^{-5}</td>
<td>0.1083</td>
<td>3.82330 \cdot 10^{-6}</td>
</tr>
<tr>
<td>8785</td>
<td>2.21182 \cdot 10^{-5}</td>
<td>0.1943</td>
<td>1.78087 \cdot 10^{-6}</td>
</tr>
<tr>
<td>30447</td>
<td>4.34389 \cdot 10^{-6}</td>
<td>0.1323</td>
<td>5.13843 \cdot 10^{-7}</td>
</tr>
<tr>
<td>39232</td>
<td>3.98781 \cdot 10^{-7}</td>
<td>0.0156</td>
<td>3.98781 \cdot 10^{-7}</td>
</tr>
<tr>
<td>226607</td>
<td>3.56231 \cdot 10^{-7}</td>
<td>0.0807</td>
<td>6.90401 \cdot 10^{-8}</td>
</tr>
<tr>
<td>2305302</td>
<td>4.28408 \cdot 10^{-8}</td>
<td>0.0988</td>
<td>6.78651 \cdot 10^{-9}</td>
</tr>
<tr>
<td>2531909</td>
<td>1.49012 \cdot 10^{-8}</td>
<td>0.0377</td>
<td>6.17912 \cdot 10^{-9}</td>
</tr>
</tbody>
</table>

In Tables 8.12 and 8.13 the values of \( \alpha_k \) and \( \gamma k^{-\tau} \) are shown for \( e_0 = 0.05 \) on Branch I and \( e_0 = 0.01 \) on Branch III, respectively, and \( k = q_n, 3 \leq n < N \).

Nekoroshev-type Results

We present results of Nekoroshev-type for \( 2\pi \)-periodic solutions corresponding to Branch I. In Tables 8.14, 8.15 and 8.16 we compare the results obtained by explicit computations and by a priori estimates. The optimal truncation orders for the a priori estimates are given in the tables while the truncation order of the explicit computations is kept at \( r = 20 \). Note that for very small \( \rho_{\text{vol}} \) the optimal truncation order \( r_{\text{opt}} \) may become larger.
8.5 Practical Stability for a Range of Values of $e$

Tab. 8.13: Comparison of $\alpha_k$ and $\gamma k^{-\tau}$ for $e_0 = 0.01$ on Branch III and $k = q_n \leq q_N$.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$\alpha_k$</th>
<th>$k\alpha_k$</th>
<th>$\gamma k^{-\tau}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>7.44968 · $10^{-3}$</td>
<td>0.0968</td>
<td>3.28694 · $10^{-3}$</td>
</tr>
<tr>
<td>98</td>
<td>1.76618 · $10^{-3}$</td>
<td>0.1731</td>
<td>4.36023 · $10^{-4}$</td>
</tr>
<tr>
<td>111</td>
<td>3.84957 · $10^{-4}$</td>
<td>0.0427</td>
<td>3.84957 · $10^{-4}$</td>
</tr>
<tr>
<td>542</td>
<td>2.26353 · $10^{-4}$</td>
<td>0.1227</td>
<td>7.88380 · $10^{-5}$</td>
</tr>
<tr>
<td>2279</td>
<td>2.29471 · $10^{-4}$</td>
<td>0.3615</td>
<td>1.87495 · $10^{-5}$</td>
</tr>
<tr>
<td>2821</td>
<td>1.58604 · $10^{-4}$</td>
<td>0.1911</td>
<td>1.51472 · $10^{-5}$</td>
</tr>
<tr>
<td>5100</td>
<td>6.77485 · $10^{-5}$</td>
<td>0.1178</td>
<td>8.37847 · $10^{-6}$</td>
</tr>
<tr>
<td>13021</td>
<td>2.31073 · $10^{-5}$</td>
<td>0.2804</td>
<td>3.28164 · $10^{-6}$</td>
</tr>
<tr>
<td>31142</td>
<td>2.15338 · $10^{-5}$</td>
<td>0.0490</td>
<td>1.37211 · $10^{-6}$</td>
</tr>
<tr>
<td>44163</td>
<td>1.57356 · $10^{-6}$</td>
<td>0.0476</td>
<td>9.67557 · $10^{-7}$</td>
</tr>
<tr>
<td>605261</td>
<td>4.96133 · $10^{-7}$</td>
<td>0.3003</td>
<td>6.57971 · $10^{-8}$</td>
</tr>
<tr>
<td>649424</td>
<td>8.51578 · $10^{-8}$</td>
<td>0.0553</td>
<td>2.24411 · $10^{-8}$</td>
</tr>
<tr>
<td>1904109</td>
<td>7.03149 · $10^{-8}$</td>
<td>0.1340</td>
<td>4.20161 · $10^{-9}$</td>
</tr>
<tr>
<td>10169969</td>
<td>1.46683 · $10^{-8}$</td>
<td>0.1492</td>
<td>6.78651 · $10^{-9}$</td>
</tr>
<tr>
<td>12074078</td>
<td>1.11759 · $10^{-8}$</td>
<td>0.1349</td>
<td>3.53900 · $10^{-9}$</td>
</tr>
<tr>
<td>58466281</td>
<td>3.72529 · $10^{-9}$</td>
<td>0.2178</td>
<td>7.30852 · $10^{-10}$</td>
</tr>
</tbody>
</table>

than the denominator $q_N$ of the best convergent of $\omega_1/\omega_0$. Thus the a priori method breaks down.

The section is completed with similar results for Branch III. In Tables 8.17, 8.18, 8.19 and 8.20 we compare results obtained by explicit computations and by a priori estimates. The optimal truncation orders for the a priori estimates are given in the tables while the truncation order of the explicit computations is kept at $r = 20$.

8.5 Practical Stability for a Range of Values of $e$

In the last section we presented results on practical stability for prescribed values of the eccentricity $e$. There arises the question whether $T_{\text{max}}/2\pi$ may be computed for all values of the eccentricity $e$ in a certain interval. In this section we answer this question in the affirmative.
Tab. 8.14: Comparison of $T_{\text{max}}^{(20)}/2\pi$ and $T_{\text{max}}^{\text{opt}}/2\pi$ for $e_0 = 0.05$ on Branch I.

<table>
<thead>
<tr>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{sol}}$</th>
<th>$r_{\text{opt}}$</th>
<th>$T_{\text{max}}^{(20)}/2\pi$</th>
<th>$T_{\text{max}}^{\text{opt}}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4 \times 10^{-18}$</td>
<td>$10^{-17}$</td>
<td>3050235</td>
<td>$3 \times 10^{318}$</td>
<td>$6 \times 10^{3793182}$</td>
</tr>
<tr>
<td>$4 \times 10^{-17}$</td>
<td>$10^{-16}$</td>
<td>964569</td>
<td>$3 \times 10^{299}$</td>
<td>$8 \times 10^{1199495}$</td>
</tr>
<tr>
<td>$4 \times 10^{-16}$</td>
<td>$10^{-15}$</td>
<td>305023</td>
<td>$3 \times 10^{280}$</td>
<td>$1 \times 10^{4793182}$</td>
</tr>
<tr>
<td>$4 \times 10^{-15}$</td>
<td>$10^{-14}$</td>
<td>96456</td>
<td>$3 \times 10^{261}$</td>
<td>$7 \times 10^{119933}$</td>
</tr>
<tr>
<td>$4 \times 10^{-14}$</td>
<td>$10^{-13}$</td>
<td>30502</td>
<td>$3 \times 10^{242}$</td>
<td>$5 \times 10^{47915}$</td>
</tr>
<tr>
<td>$4 \times 10^{-13}$</td>
<td>$10^{-12}$</td>
<td>9645</td>
<td>$3 \times 10^{223}$</td>
<td>$2 \times 10^{11980}$</td>
</tr>
<tr>
<td>$4 \times 10^{-12}$</td>
<td>$10^{-11}$</td>
<td>3050</td>
<td>$3 \times 10^{204}$</td>
<td>$7 \times 10^{4779}$</td>
</tr>
<tr>
<td>$4 \times 10^{-11}$</td>
<td>$10^{-10}$</td>
<td>964</td>
<td>$3 \times 10^{185}$</td>
<td>$5 \times 10^{1187}$</td>
</tr>
<tr>
<td>$4 \times 10^{-10}$</td>
<td>$10^{-9}$</td>
<td>305</td>
<td>$3 \times 10^{166}$</td>
<td>$9 \times 10^{368}$</td>
</tr>
<tr>
<td>$4 \times 10^{-9}$</td>
<td>$10^{-8}$</td>
<td>96</td>
<td>$3 \times 10^{147}$</td>
<td>$1 \times 10^{111}$</td>
</tr>
<tr>
<td>$4 \times 10^{-8}$</td>
<td>$10^{-7}$</td>
<td>30</td>
<td>$3 \times 10^{128}$</td>
<td>$-$</td>
</tr>
<tr>
<td>$4 \times 10^{-7}$</td>
<td>$10^{-6}$</td>
<td>9</td>
<td>$3 \times 10^{109}$</td>
<td>$-$</td>
</tr>
</tbody>
</table>

Tab. 8.15: Comparison of $T_{\text{max}}^{(20)}/2\pi$ and $T_{\text{max}}^{\text{opt}}/2\pi$ for $e_0 = 0.2$ on Branch I.

<table>
<thead>
<tr>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{sol}}$</th>
<th>$r_{\text{opt}}$</th>
<th>$T_{\text{max}}^{(20)}/2\pi$</th>
<th>$T_{\text{max}}^{\text{opt}}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4 \times 10^{-18}$</td>
<td>$10^{-17}$</td>
<td>968942</td>
<td>$3 \times 10^{312}$</td>
<td>$-$</td>
</tr>
<tr>
<td>$4 \times 10^{-17}$</td>
<td>$10^{-16}$</td>
<td>306406</td>
<td>$3 \times 10^{293}$</td>
<td>$3 \times 10^{391235}$</td>
</tr>
<tr>
<td>$4 \times 10^{-16}$</td>
<td>$10^{-15}$</td>
<td>96894</td>
<td>$3 \times 10^{274}$</td>
<td>$3 \times 10^{123708}$</td>
</tr>
<tr>
<td>$4 \times 10^{-15}$</td>
<td>$10^{-14}$</td>
<td>30640</td>
<td>$3 \times 10^{255}$</td>
<td>$1 \times 10^{39110}$</td>
</tr>
<tr>
<td>$4 \times 10^{-14}$</td>
<td>$10^{-13}$</td>
<td>9689</td>
<td>$3 \times 10^{236}$</td>
<td>$6 \times 10^{12358}$</td>
</tr>
<tr>
<td>$4 \times 10^{-13}$</td>
<td>$10^{-12}$</td>
<td>3064</td>
<td>$3 \times 10^{217}$</td>
<td>$2 \times 10^{3900}$</td>
</tr>
<tr>
<td>$4 \times 10^{-12}$</td>
<td>$10^{-11}$</td>
<td>968</td>
<td>$3 \times 10^{198}$</td>
<td>$3 \times 10^{1226}$</td>
</tr>
<tr>
<td>$4 \times 10^{-11}$</td>
<td>$10^{-10}$</td>
<td>306</td>
<td>$3 \times 10^{179}$</td>
<td>$1 \times 10^{382}$</td>
</tr>
<tr>
<td>$4 \times 10^{-10}$</td>
<td>$10^{-9}$</td>
<td>96</td>
<td>$3 \times 10^{160}$</td>
<td>$7 \times 10^{1115}$</td>
</tr>
<tr>
<td>$4 \times 10^{-9}$</td>
<td>$10^{-8}$</td>
<td>30</td>
<td>$3 \times 10^{141}$</td>
<td>$-$</td>
</tr>
<tr>
<td>$4 \times 10^{-8}$</td>
<td>$10^{-7}$</td>
<td>9</td>
<td>$3 \times 10^{122}$</td>
<td>$-$</td>
</tr>
</tbody>
</table>
8.5. Practical Stability for a Range of Values of $\epsilon$

Tab. 8.16: Comparison of $T_{\text{max}}^{(20)}/2\pi$ and $T_{\text{opt}}^{\text{max}}/2\pi$ for $\epsilon_0 = 0.35$ on Branch I.

<table>
<thead>
<tr>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{tol}}$</th>
<th>$r_{\text{opt}}$</th>
<th>$T_{\text{max}}^{(20)}/2\pi$</th>
<th>$T_{\text{opt}}^{\text{max}}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4 \cdot 10^{-18}$</td>
<td>$10^{-17}$</td>
<td>295188</td>
<td>$7 \cdot 10^{295}$</td>
<td>$-$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-17}$</td>
<td>$10^{-16}$</td>
<td>93346</td>
<td>$7 \cdot 10^{276}$</td>
<td>$1 \cdot 10^{96965}$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-16}$</td>
<td>$10^{-15}$</td>
<td>29518</td>
<td>$7 \cdot 10^{257}$</td>
<td>$1 \cdot 10^{27491}$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-15}$</td>
<td>$10^{-14}$</td>
<td>9334</td>
<td>$7 \cdot 10^{238}$</td>
<td>$4 \cdot 10^{8684}$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-14}$</td>
<td>$10^{-13}$</td>
<td>2951</td>
<td>$7 \cdot 10^{219}$</td>
<td>$4 \cdot 10^{2738}$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-13}$</td>
<td>$10^{-12}$</td>
<td>933</td>
<td>$7 \cdot 10^{200}$</td>
<td>$2 \cdot 10^{859}$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-12}$</td>
<td>$10^{-11}$</td>
<td>295</td>
<td>$7 \cdot 10^{181}$</td>
<td>$1 \cdot 10^{266}$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-11}$</td>
<td>$10^{-10}$</td>
<td>93</td>
<td>$7 \cdot 10^{162}$</td>
<td>$2 \cdot 10^{79}$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-10}$</td>
<td>$10^{-9}$</td>
<td>29</td>
<td>$7 \cdot 10^{143}$</td>
<td>$-$</td>
</tr>
<tr>
<td>$4 \cdot 10^{-9}$</td>
<td>$10^{-8}$</td>
<td>9</td>
<td>$7 \cdot 10^{124}$</td>
<td>$-$</td>
</tr>
</tbody>
</table>

Tab. 8.17: Comparison of $T_{\text{max}}^{(20)}/2\pi$ and $T_{\text{opt}}^{\text{max}}/2\pi$ for $\epsilon_0 = 0.01$ on Branch III.

<table>
<thead>
<tr>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{tol}}$</th>
<th>$r_{\text{opt}}$</th>
<th>$T_{\text{max}}^{(20)}/2\pi$</th>
<th>$T_{\text{opt}}^{\text{max}}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3 \cdot 10^{-10}$</td>
<td>$1 \cdot 10^{-17}$</td>
<td>333517</td>
<td>$7 \cdot 10^{299}$</td>
<td>$1 \cdot 10^{289771}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-18}$</td>
<td>$1 \cdot 10^{-16}$</td>
<td>105467</td>
<td>$7 \cdot 10^{271}$</td>
<td>$5 \cdot 10^{81622}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-17}$</td>
<td>$1 \cdot 10^{-15}$</td>
<td>33351</td>
<td>$7 \cdot 10^{252}$</td>
<td>$5 \cdot 10^{28963}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-16}$</td>
<td>$1 \cdot 10^{-14}$</td>
<td>10546</td>
<td>$7 \cdot 10^{233}$</td>
<td>$1 \cdot 10^{9150}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-15}$</td>
<td>$1 \cdot 10^{-13}$</td>
<td>3335</td>
<td>$7 \cdot 10^{214}$</td>
<td>$4 \cdot 10^{2885}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-14}$</td>
<td>$1 \cdot 10^{-12}$</td>
<td>1054</td>
<td>$7 \cdot 10^{195}$</td>
<td>$4 \cdot 10^{905}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-13}$</td>
<td>$1 \cdot 10^{-11}$</td>
<td>333</td>
<td>$7 \cdot 10^{176}$</td>
<td>$3 \cdot 10^{280}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-12}$</td>
<td>$1 \cdot 10^{-10}$</td>
<td>105</td>
<td>$7 \cdot 10^{157}$</td>
<td>$7 \cdot 10^{83}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-11}$</td>
<td>$1 \cdot 10^{-9}$</td>
<td>33</td>
<td>$7 \cdot 10^{138}$</td>
<td>$2 \cdot 10^{22}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-10}$</td>
<td>$1 \cdot 10^{-8}$</td>
<td>10</td>
<td>$7 \cdot 10^{119}$</td>
<td>$-$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-9}$</td>
<td>$1 \cdot 10^{-7}$</td>
<td>3</td>
<td>$7 \cdot 10^{100}$</td>
<td>$-$</td>
</tr>
</tbody>
</table>
8. Stability of the Dumbbell Satellite

Tab. 8.18: Comparison of $T^{(20)}_{\text{max}}/2\pi$ and $T^{\text{opt}}_{\text{max}}/2\pi$ for $e_0 = 0.02$ on Branch III.

<table>
<thead>
<tr>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{tol}}$</th>
<th>$r_{\text{opt}}$</th>
<th>$T^{(20)}_{\text{max}}/2\pi$</th>
<th>$T^{\text{opt}}_{\text{max}}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3 \cdot 10^{-19}$</td>
<td>$1 \cdot 10^{-17}$</td>
<td>503988</td>
<td>$3 \cdot 10^{291}$</td>
<td>$5 \cdot 10^{438137}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-18}$</td>
<td>$1 \cdot 10^{-16}$</td>
<td>159375</td>
<td>$3 \cdot 10^{272}$</td>
<td>$9 \cdot 10^{438539}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-17}$</td>
<td>$1 \cdot 10^{-15}$</td>
<td>50398</td>
<td>$3 \cdot 10^{253}$</td>
<td>$7 \cdot 10^{43799}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-16}$</td>
<td>$1 \cdot 10^{-14}$</td>
<td>15937</td>
<td>$3 \cdot 10^{234}$</td>
<td>$3 \cdot 10^{13841}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-15}$</td>
<td>$1 \cdot 10^{-13}$</td>
<td>5039</td>
<td>$3 \cdot 10^{215}$</td>
<td>$6 \cdot 10^{4368}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-14}$</td>
<td>$1 \cdot 10^{-12}$</td>
<td>1593</td>
<td>$3 \cdot 10^{196}$</td>
<td>$2 \cdot 10^{1374}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-13}$</td>
<td>$1 \cdot 10^{-11}$</td>
<td>503</td>
<td>$3 \cdot 10^{177}$</td>
<td>$2 \cdot 10^{428}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-12}$</td>
<td>$1 \cdot 10^{-10}$</td>
<td>159</td>
<td>$3 \cdot 10^{130}$</td>
<td>$2 \cdot 10^{131}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-11}$</td>
<td>$1 \cdot 10^{-9}$</td>
<td>50</td>
<td>$3 \cdot 10^{139}$</td>
<td>$9 \cdot 10^{36}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-10}$</td>
<td>$1 \cdot 10^{-8}$</td>
<td>15</td>
<td>$3 \cdot 10^{120}$</td>
<td>-</td>
</tr>
<tr>
<td>$3 \cdot 10^{-9}$</td>
<td>$1 \cdot 10^{-7}$</td>
<td>5</td>
<td>$3 \cdot 10^{101}$</td>
<td>-</td>
</tr>
</tbody>
</table>

Tab. 8.19: Comparison of $T^{(20)}_{\text{max}}/2\pi$ and $T^{\text{opt}}_{\text{max}}/2\pi$ for $e_0 = 0.03$ on Branch III.

<table>
<thead>
<tr>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{tol}}$</th>
<th>$r_{\text{opt}}$</th>
<th>$T^{(20)}_{\text{max}}/2\pi$</th>
<th>$T^{\text{opt}}_{\text{max}}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3 \cdot 10^{-19}$</td>
<td>$1 \cdot 10^{-17}$</td>
<td>300816</td>
<td>$5 \cdot 10^{289}$</td>
<td>$2 \cdot 10^{261820}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-18}$</td>
<td>$1 \cdot 10^{-16}$</td>
<td>95126</td>
<td>$5 \cdot 10^{270}$</td>
<td>$2 \cdot 10^{82784}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-17}$</td>
<td>$1 \cdot 10^{-15}$</td>
<td>30081</td>
<td>$5 \cdot 10^{251}$</td>
<td>$1 \cdot 10^{26169}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-16}$</td>
<td>$1 \cdot 10^{-14}$</td>
<td>9512</td>
<td>$5 \cdot 10^{232}$</td>
<td>$7 \cdot 10^{8266}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-15}$</td>
<td>$1 \cdot 10^{-13}$</td>
<td>3008</td>
<td>$5 \cdot 10^{213}$</td>
<td>$5 \cdot 10^{2606}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-14}$</td>
<td>$1 \cdot 10^{-12}$</td>
<td>951</td>
<td>$5 \cdot 10^{194}$</td>
<td>$6 \cdot 10^{817}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-13}$</td>
<td>$1 \cdot 10^{-11}$</td>
<td>300</td>
<td>$5 \cdot 10^{175}$</td>
<td>$1 \cdot 10^{253}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-12}$</td>
<td>$1 \cdot 10^{-10}$</td>
<td>95</td>
<td>$5 \cdot 10^{156}$</td>
<td>$4 \cdot 10^{75}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-11}$</td>
<td>$1 \cdot 10^{-9}$</td>
<td>30</td>
<td>$5 \cdot 10^{137}$</td>
<td>$2 \cdot 10^{20}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-10}$</td>
<td>$1 \cdot 10^{-8}$</td>
<td>9</td>
<td>$5 \cdot 10^{118}$</td>
<td>-</td>
</tr>
</tbody>
</table>
8.5. Practical Stability for a Range of Values of e

Tab. 8.20: Comparison of $T_{\text{max}}^{(20)}/2\pi$ and $T_{\text{opt}}^{(20)}/2\pi$ for $e_0 = 0.04$ on Branch III.

<table>
<thead>
<tr>
<th>$\rho_{\text{start}}$</th>
<th>$\rho_{\text{tol}}$</th>
<th>$r_{\text{opt}}$</th>
<th>$T_{\text{max}}^{(20)}/2\pi$</th>
<th>$T_{\text{opt}}^{(20)}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3 \cdot 10^{-19}$</td>
<td>$1 \cdot 10^{-17}$</td>
<td>227898</td>
<td>$9 \cdot 10^{281}$</td>
<td>$2 \cdot 10^{199330}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-18}$</td>
<td>$1 \cdot 10^{-16}$</td>
<td>72067</td>
<td>$9 \cdot 10^{262}$</td>
<td>$2 \cdot 10^{63026}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-17}$</td>
<td>$1 \cdot 10^{-15}$</td>
<td>22789</td>
<td>$9 \cdot 10^{243}$</td>
<td>$2 \cdot 10^{19921}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-16}$</td>
<td>$1 \cdot 10^{-14}$</td>
<td>7206</td>
<td>$9 \cdot 10^{224}$</td>
<td>$2 \cdot 10^{6291}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-15}$</td>
<td>$1 \cdot 10^{-13}$</td>
<td>227</td>
<td>$9 \cdot 10^{205}$</td>
<td>$1 \cdot 10^{1982}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-14}$</td>
<td>$1 \cdot 10^{-12}$</td>
<td>72</td>
<td>$9 \cdot 10^{186}$</td>
<td>$3 \cdot 10^{620}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-13}$</td>
<td>$1 \cdot 10^{-11}$</td>
<td>22</td>
<td>$9 \cdot 10^{167}$</td>
<td>$9 \cdot 10^{190}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-12}$</td>
<td>$1 \cdot 10^{-10}$</td>
<td>72</td>
<td>$9 \cdot 10^{148}$</td>
<td>$1 \cdot 10^{56}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-11}$</td>
<td>$1 \cdot 10^{-9}$</td>
<td>22</td>
<td>$9 \cdot 10^{129}$</td>
<td>$2 \cdot 10^{14}$</td>
</tr>
<tr>
<td>$3 \cdot 10^{-10}$</td>
<td>$1 \cdot 10^{-8}$</td>
<td>7</td>
<td>$9 \cdot 10^{110}$</td>
<td>-</td>
</tr>
</tbody>
</table>

Estimating the Perturbation

The right hand side of the original differential equation is given by

$$
\begin{pmatrix}
F_1(t, x, y, e) \\
F_2(t, x, y, e)
\end{pmatrix} := \begin{pmatrix}
-\frac{3}{2} \sin(2x) + \frac{y}{1 - e \cos t} + \frac{y e \sin t}{(1 - e \cos t)^2} \\
1 - \frac{1}{1 - e \cos t}
\end{pmatrix}.
$$

Unlike in Section 8.4 the eccentricity $e$ is now no longer regarded to be constant.

Let

$$
\begin{pmatrix}
x_0(t) \\
y_0(t)
\end{pmatrix} := z_{\text{per}}^0(t), \quad \begin{pmatrix}
\tilde{q}_{11}(t) \\
\tilde{q}_{21}(t) \\
\tilde{q}_{22}(t)
\end{pmatrix} := Q(t)S, \quad \text{and} \quad \begin{pmatrix}
x_4(t) \\
y_4(t)
\end{pmatrix} := Q(t)S z_3^0(t),
$$

where $z_{\text{per}}^0(t)$, $Q(t)$, $S$ and $z_3^0(t)$ are introduced in Chapter 4. Then Theorem 4.5.1 (vi) implies that

$$
\begin{pmatrix}
h_1(t, x, y, \Delta e) \\
h_2(t, x, y, \Delta e)
\end{pmatrix} := S^{-1}Q^{-1} \begin{pmatrix}
F_1^{\geq 2}(t, x_4, y_4, \Delta e) \\
F_2^{\geq 2}(t, x_4, y_4, \Delta e)
\end{pmatrix},
$$

where

$$
\begin{pmatrix}
x_4 \\
y_4
\end{pmatrix} := \begin{pmatrix}
x_0(t) + \tilde{q}_{11}(t) + \tilde{q}_{12}(t)y + x_1(t)\Delta e \\
y_0(t) + \tilde{q}_{21}(t) + \tilde{q}_{22}(t)y + y_1(t)\Delta e
\end{pmatrix}
$$

and $F_i^{\geq 2}(t, x_4, y_4, \Delta e)$ are the power series of $F_i(t, x_4, y_4, \Delta e)$ with terms of order 2 and higher in $x_4, y_4$ and $\Delta e$.

The goal of this section is to give a priori estimates of the form $\|h_m^k\|_R \leq Eh^k$, $m = 1, 2$ for the terms of the expansions

$$
h_m(t, x, y, \Delta e) = \sum_{k=2}^{\infty} h_k^m(t, x, y, \Delta e), \quad m = 1, 2.
$$
The Expansion of the First Expression

We start with the expansion of

\[
\frac{3 \sin(2x_0 + 2\tilde{q}_{11}x + 2\tilde{q}_{12}y + x_1 \Delta e)}{2 1 - (e_0 + \Delta e) \cos t}.
\]

For the numerator we have

\[
\sin(2x_0 + 2\tilde{q}_{11}x + 2\tilde{q}_{12}y + x_1 \Delta e) = \\
\sin(2x_0) \cos(2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1 \Delta e) + \cos(2x_0) \sin(2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1 \Delta e) = \\
\sin(2x_0) \sum_{l=0}^{\infty} (-1)^l \frac{(2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1 \Delta e)^{2l}}{(2l)!} + \\
\cos(2x_0) \sum_{l=0}^{\infty} (-1)^l \frac{(2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1 \Delta e)^{2l+1}}{(2l + 1)!}
\]

and for the denominator

\[
\frac{1}{1 - (e_0 + \Delta e) \cos t} = \\
\frac{1}{1 - e_0 \cos t} - \Delta e \frac{\cos t}{1 - e_0 \cos t} = \\
\frac{1}{1 - e_0 \cos t} \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^k \Delta e^k.
\]

Combining these results we find

\[
\frac{3 \sin(2x_0 + 2\tilde{q}_{11}x + 2\tilde{q}_{12}y + x_1 \Delta e)}{2 1 - (e_0 + \Delta e) \cos t} = \\
\frac{3}{2} \frac{1}{1 - e_0 \cos t} \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^k \Delta e^k.
\]

\[
\cdot \sin(2x_0) \sum_{l=0}^{\infty} (-1)^l \frac{(2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1 \Delta e)^{2l}}{(2l)!} - \\
\frac{3}{2} \frac{1}{1 - e_0 \cos t} \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^k \Delta e^k.
\]

\[
\cdot \cos(2x_0) \sum_{l=0}^{\infty} (-1)^l \frac{(2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1 \Delta e)^{2l+1}}{(2l + 1)!} =
\]
8.5. Practical Stability for a Range of Values of $e$

\[
- \frac{3 \sin(2x_0)}{2 \bar{1} - e_0 \cos t} \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k} \Delta e^{2k} \cdot \sum_{l=0}^{\infty} (-1)^l \frac{(2q_{11}x + 2q_{12}y + 2x_1 \Delta e)^{2l}}{(2l)!} \\
- \frac{3 \cos(2x_0)}{2 \bar{1} - e_0 \cos t} \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k+1} \Delta e^{2k+1} \cdot \sum_{l=0}^{\infty} (-1)^l \frac{(2q_{11}x + 2q_{12}y + 2x_1 \Delta e)^{2l+1}}{(2l + 1)!} \\
- \frac{3 \sin(2x_0)}{2 \bar{1} - e_0 \cos t} \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k+1} \Delta e^{2k+1} \cdot \sum_{l=0}^{\infty} (-1)^l \frac{(2q_{11}x + 2q_{12}y + 2x_1 \Delta e)^{2l}}{(2l)!} \\
- \frac{3 \cos(2x_0)}{2 \bar{1} - e_0 \cos t} \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k} \Delta e^{2k} \cdot \sum_{l=0}^{\infty} (-1)^l \frac{(2q_{11}x + 2q_{12}y + 2x_1 \Delta e)^{2l+1}}{(2l + 1)!} \\
= - \frac{3 \sin(2x_0)}{2 \bar{1} - e_0 \cos t} \sum_{k=0}^{\infty} \sum_{l=0}^{k} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2l} \frac{(-1)^{k-l}}{(2k - 2l)!} \Delta e^{2l} (2q_{11}x + 2q_{12}y + 2x_1 \Delta e)^{2(k-l)} - \\
- \frac{3 \cos(2x_0)}{2 \bar{1} - e_0 \cos t} \sum_{k=0}^{\infty} \sum_{l=0}^{k} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2l+1} \frac{(-1)^{k-l}}{(2k - 2l + 1)!} \Delta e^{2l+1} (2q_{11}x + 2q_{12}y + 2x_1 \Delta e)^{2(k-l)+1} - \\
- \frac{3 \sin(2x_0)}{2 \bar{1} - e_0 \cos t} \sum_{k=0}^{\infty} \sum_{l=0}^{k} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2l+1} \frac{(-1)^{k-l}}{(2(k - l))!} \Delta e^{2l+1} (2q_{11}x + 2q_{12}y + 2x_1 \Delta e)^{2(k-l)} - \\
- \frac{3 \cos(2x_0)}{2 \bar{1} - e_0 \cos t} \sum_{k=0}^{\infty} \sum_{l=0}^{k} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2l+1} \frac{(-1)^{k-l}}{(2(k - l) + 1)!} \Delta e^{2l} (2q_{11}x + 2q_{12}y + 2x_1 \Delta e)^{2(k-l)+1} -
\[ = - \sum_{k=0}^{\infty} \left( \sum_{l=0}^{k} \frac{3}{2} \frac{\sin(2x_0)}{1 - e_0 \cos t} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2l} \frac{(-1)^{k-l}}{(2(k-l))!} \Delta e^{2l} \cdot \left( 2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1\Delta e \right)^{2(k-l)+1} + \right. \]
\[ + \sum_{l=0}^{k-1} \frac{3}{2} \frac{\cos(2x_0)}{1 - e_0 \cos t} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2l+1} \frac{(-1)^{k-l}}{(2(k-l)+1)!} \Delta e^{2l+1} \cdot \left( 2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1\Delta e \right)^{2(k-l)+1} \right) - \]
\[ - \sum_{k=0}^{\infty} \left( \sum_{l=0}^{k} \frac{3}{2} \frac{\sin(2x_0)}{1 - e_0 \cos t} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2l+1} \frac{(-1)^{k-l}}{(2(k-l))!} \Delta e^{2l+1} \cdot \left( 2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1\Delta e \right)^{2(k-l)+1} + \right. \]
\[ + \sum_{l=0}^{k-1} \frac{3}{2} \frac{\cos(2x_0)}{1 - e_0 \cos t} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2l} \frac{(-1)^{k-l}}{(2(k-l)+1)!} \Delta e^{2l} \cdot \left( 2\tilde{q}_{11}x + 2\tilde{q}_{12}y + 2x_1\Delta e \right)^{2(k-l)+1} \right) \]

The Expansion of the Second Expression

Now we consider the expansion of
\[ (y_0 + \tilde{q}_{21}x + \tilde{q}_{22}y + y_1\Delta e) \cdot \frac{(e_0 + \Delta e) \sin t}{1 - (e_0 + \Delta e) \cos t} \]

We obtain
\[ (y_0 + \tilde{q}_{21}x + \tilde{q}_{22}y + y_1\Delta e) \cdot \frac{(e_0 + \Delta e) \sin t}{1 - (e_0 + \Delta e) \cos t} = \]
\[ = \frac{\sin t}{1 - e_0 \cos t} \left( y_0 e_0 + (e_0 \tilde{q}_{21}x + e_0 \tilde{q}_{22}y + y_0 \Delta e + y_1 e_0 \Delta e) + (\tilde{q}_{21}x \Delta e + \tilde{q}_{22}y \Delta e + y_1 \Delta e^2) \right) \]
\[ \cdot \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^k \Delta e^k \]
5. Practical Stability for a Range of Values of $e$

\[
= \frac{\sin t}{1 - e_0 \cos t} \left( y_0 e_0 \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k} \Delta e^{2k} + \right.
\]
\[
+ (e_0 \tilde{q}_{21} x + e_0 \tilde{q}_{22} y + y_0 \Delta e + y_1 e_0 \Delta e) \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k+1} \Delta e^{2k+1} + \right.
\]
\[
+ (\tilde{q}_{21} x \Delta e + \tilde{q}_{22} y \Delta e + y_1 \Delta e^2) \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k} \Delta e^{2k} + \right.
\]
\[
+ \frac{\sin t}{1 - e_0 \cos t} \left( y_0 e_0 \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k+1} \Delta e^{2k+1} + \right.
\]
\[
+ (e_0 \tilde{q}_{21} x + e_0 \tilde{q}_{22} y + y_0 \Delta e + y_1 e_0 \Delta e) \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k} \Delta e^{2k} + \right.
\]
\[
+ (\tilde{q}_{21} x \Delta e + \tilde{q}_{22} y \Delta e + y_1 \Delta e^2) \sum_{k=0}^{\infty} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k+1} \Delta e^{2k+1} + \right.
\]
\[
= \frac{y_0 e_0 \sin t}{1 - e_0 \cos t} + \frac{\sin t}{1 - e_0 \cos t} \left( y_0 \frac{\cos t}{1 - e_0 \cos t} \Delta e + (e_0 \tilde{q}_{21} x + e_0 \tilde{q}_{22} y + y_0 \Delta e + y_1 e_0 \Delta e) \right) + \right.
\]
\[
+ \frac{\sin t}{1 - e_0 \cos t} \sum_{k=0}^{\infty} \left( y_0 e_0 \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k+2} \Delta e^{2k+2} + \right.
\]
\[
+ (e_0 \tilde{q}_{21} x + e_0 \tilde{q}_{22} y + y_0 \Delta e + y_1 e_0 \Delta e) \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k+1} \Delta e^{2k+1} + \right.
\]
\[
+ (\tilde{q}_{21} x \Delta e + \tilde{q}_{22} y \Delta e + y_1 \Delta e^2) \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k} \Delta e^{2k} + \right.
\]
\[
+ \frac{\sin t}{1 - e_0 \cos t} \sum_{k=1}^{\infty} \left( y_0 e_0 \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k+1} \Delta e^{2k+1} + \right.
\]
\[
+ (e_0 \tilde{q}_{21} x + e_0 \tilde{q}_{22} y + y_0 \Delta e + y_1 e_0 \Delta e) \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k} \Delta e^{2k} + \right.
\]
\[
+ (\tilde{q}_{21} x \Delta e + \tilde{q}_{22} y \Delta e + y_1 \Delta e^2) \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{2k-1} \Delta e^{2k-1} .
\]
The Expansion of the Third Expression

Finally we consider the expansion of

$$\frac{2(e_0 + \Delta e)\sqrt{1 - (e_0 + \Delta e)^2}\sin t}{(1 - (e_0 + \Delta e)\cos t)^2}.$$ 

We start with the square root $\sqrt{1 - (e_0 + \Delta e)^2}$. Let

$$\sqrt{1 - (e_0 + \Delta e)^2} =: \sum_{k=0}^{\infty} a_k \Delta e^k.$$ 

Squaring both sides gives

$$(1 - e_0^2) - 2e_0\Delta e - \Delta e^2 = \sum_{k=0}^{\infty} \sum_{l=0}^{k} a_l a_{k-l} \Delta e^k.$$ 

Thus we find for $a_k$:

$$a_0 = \sqrt{1 - e_0^2}, \quad a_1 = \frac{-e_0}{\sqrt{1 - e_0^2}}, \quad a_2 = \frac{-1}{2(1 - e_0^2)^{3/2}}, \quad a_k = \frac{-1}{2a_0} \sum_{l=1}^{k-1} a_l a_{k-l}, \quad k > 2.$$ 

Now let $b_k := (1 - e_0^2)^{k-\frac{3}{2}} a_k$ Then we have

$$b_0 = 1, \quad b_1 = -e_0, \quad b_2 = -\frac{1}{2},$$

$$b_k = (1 - e_0^2)^{k-\frac{3}{2}} a_k = (1 - e_0^2)^{k-\frac{3}{2}} \left( \frac{-1}{2a_0} \right) \sum_{l=1}^{k-1} (1 - e_0^2)^{(l-\frac{3}{2})} b_l (1 - e_0^2)^{-(k-l-\frac{3}{2})} b_{k-l} =$$

$$= \frac{1}{2} \sum_{l=1}^{k-1} b_l b_{k-l}, \quad k > 2.$$ 

The $b_k$ are polynomials in $e_0$ of degree $k$. For $k \geq 1$ all coefficients of $b_k(e_0)$ are negative as one easily sees. Therefore we have

$$\max_{0 \leq e_0 \leq 1} |b_k(e_0)| = |b_k(1)|.$$ 

It remains to give an explicit formula for $|b_k(1)|$. To this end we consider the series of $\sqrt{1 - 2x}$:

$$(1 - 2x)^{1/2} =: \sum_{k=0}^{\infty} c_k x^k.$$ 

Squaring leads to

$$1 - 2x = \sum_{k=0}^{\infty} \sum_{l=0}^{k} c_l c_{k-l} x^k.$$
and further to
\[ c_0 = 1, \quad c_1 = -1, \quad c_k = -\frac{1}{2} \sum_{l=1}^{k-1} c_l c_{k-l} \quad k > 1. \]

Thus we conclude that \(|b_k(1)| = c_k\) for all \(k \in \mathbb{N}\). On the other hand we may express \(\sqrt{1-2x}\) by
\[ (1 - 2x)^{1/2} = \sum_{k=0}^{\infty} \prod_{j=1}^{k} \frac{2j-3}{2j} (2x)^k = \sum_{k=0}^{\infty} \prod_{j=1}^{k} \frac{2j-3}{j} (x)^k. \]

This leads to the desired explicit formula for \(|b_k(1)|\):
\[ |b_k(1)| = \prod_{j=2}^{k} \frac{2j-3}{j}. \]

We may summarize these results as follows:
\[ \sqrt{1 - (e_0 + \Delta e)^2} = \sum_{k=0}^{\infty} (1 - e_0^2)^{1/2 - k} b_k(e_0) \Delta e^k, \]
with
\[ \max_{0 \leq e_0 \leq 1} |b_k(e_0)| = |b_k(1)| = \prod_{j=2}^{k} \frac{2j-3}{j}. \quad (8.10) \]

Now we consider the expression
\[ \frac{(e_0 + \Delta e) \sin t}{(1 - (e_0 + \Delta e) \cos t)^2}. \]

We obtain
\[ \frac{(e_0 + \Delta e) \sin t}{(1 - (e_0 + \Delta e) \cos t)^2} = \]
\[ = \frac{(e_0 + \Delta e) \sin t}{(1 - e_0 \cos t)^2} \cdot \frac{1}{\left(1 - \frac{\cos t}{1 - e_0 \cos t} \Delta e\right)^2} = \]
\[ = \frac{(e_0 + \Delta e) \sin t}{(1 - e_0 \cos t)^2} \sum_{l=0}^{\infty} (l + 1) \left(\frac{\cos t}{1 - e_0 \cos t}\right)^l \Delta e^l = \]
\[ = \frac{\sin t}{(1 - e_0 \cos t)^2} \sum_{l=0}^{\infty} e_0 (l + 1) \left(\frac{\cos t}{1 - e_0 \cos t}\right)^l \Delta e^l + \]
\[ + \frac{\sin t}{(1 - e_0 \cos t)^2} \sum_{l=0}^{\infty} (l + 1) \left(\frac{\cos t}{1 - e_0 \cos t}\right)^l \Delta e^{l+1} = \]
Stability of the Dumbbell Satellite

\[
\frac{e_0 \sin t}{(1 - e_0 \cos t)^2} + \\
\frac{\sin t}{(1 - e_0 \cos t)^2} \sum_{l=1}^{\infty} \left( e_0 (l + 1) \frac{\cos t}{1 - e_0 \cos t} + l \right) \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{l-1} \Delta e^l =
\]

\[
= \frac{e_0 \sin t}{(1 - e_0 \cos t)^2} + \\
\frac{\sin t}{(1 - e_0 \cos t)^2} \sum_{l=1}^{\infty} e_0^l \frac{\cos t + e_0 \cos t + l - e_0 l \cos t}{1 - e_0 \cos t} \left( \frac{\cos t}{1 - e_0 \cos t} \right)^{l-1} \Delta e^l =
\]

\[
= \frac{e_0 \sin t}{(1 - e_0 \cos t)^2} + \sum_{l=1}^{\infty} \frac{(e_0 \cos t + l) \cos^{l-1} t \sin t}{(1 - e_0 \cos t)^{l+2}} \Delta e^l
\]

Combining these results we obtain

\[
\frac{(e_0 + \Delta e) \sqrt{1 - (e_0 + \Delta e)^2} \sin t}{(1 - (e_0 + \Delta e) \cos t)^2} =
\]

\[
= \sum_{k=0}^{\infty} (1 - e_0^2)^{1/2-k} b_k(e_0) \Delta e^k \cdot \sum_{l=0}^{\infty} \frac{(e_0 \cos t + l) \sin t \cos^{l-1} t}{(1 - e_0 \cos t)^{l+2}} \Delta e^l
\]

\[
= \sum_{k=0}^{\infty} \sum_{l=0}^{k} (1 - e_0^2)^{1/2-l} b_l(e_0) \frac{(e_0 \cos t + k - l) \sin t \cos^{k-l-1} t}{(1 - e_0 \cos t)^{k-l+2}} \Delta e^k.
\]

The Estimates of \( h_m^k \)

We are now in the position the give the estimates for the right hand side of (8.9) of the transformed differential equation. Combining the expansions given above and using Lemma 8.4.1 and (8.10) we obtain

\[
\| h_m^2 \|_\mathcal{F} \leq \left\| \max_{1 \leq i, j \leq 2} (Q^{-1})_{ij} \right\|_\mathcal{F} \cdot \\
\cdot \left( \frac{3}{2} \left\| \sin(2x_0) \right\|_\mathcal{F} \sum_{l=0}^{k} \frac{1}{(2k - 2l)!} \left( \frac{R_2}{1 - e_0} \right)^{2l} \cdot \\
\cdot (2\| \tilde{q}_{11} \|_\mathcal{F} R_1 + 2\| \tilde{q}_{12} \|_\mathcal{F} R_1 + 2\| x_1 \|_\mathcal{F} R_2)^{2k-2l} + \\
\right.
\]
8.5. Practical Stability for a Range of Values of e

\[ + \frac{3}{2} \left| \frac{\cos(2x_0)}{1 - e_0} \right| \sum_{l=0}^{k} \frac{1}{(2k - 2l + 1)!} \left( \frac{R_2}{1 - e_0} \right)^{2l+1} \]

\[ \cdot \left( 2||\tilde{q}_{11}||xR_1 + 2||\tilde{q}_{12}||xR_1 + 2||x_1||xR_2 \right)^{2k-2l+1} + \]

\[ + \frac{1}{1 - e_0} \left( ||y_0||x e_0 \left( \frac{R_2}{1 - e_0} \right)^{2k} + \right. \]

\[ + \left( e_0 ||\tilde{q}_{21}||xR_1 + e_0 ||\tilde{q}_{22}||xR_2 + ||y_0||xR_2 + e_0 ||y_1||xR_2 \right) \cdot \left( \frac{1}{1 - e_0} \right)^{2k-1} \]

\[ + \left( ||\tilde{q}_{21}||xR_1R_2 + ||\tilde{q}_{22}||xR_1R_2 + ||y_1||xR_2^2 \right) \cdot \left( \frac{1}{1 - e_0} \right)^{2k-2} \]

\[ + \sum_{l=0}^{2k} (1 - e_0^2)^{1/2-l} \prod_{j=2}^{l} \frac{2j - 3 (e_0 + 2k - l)}{j} (1 - e_0)^{2k-l+2} R_2^{2k} \]

and

\[ ||h_{m}^{2k+1}||_R \leq \left\| \max_{1 \leq i, j \leq 2} (Q^{-1})_{ij} \right\|_x \]

\[ \cdot \left( 3 \left| \frac{\sin(2x_0)}{1 - e_0} \right| \sum_{l=0}^{k} \frac{1}{(2k - 2l)!} \left( \frac{R_2}{1 - e_0} \right)^{2l+1} \right. \]

\[ \cdot \left( 2||\tilde{q}_{11}||xR_1 + 2||\tilde{q}_{12}||xR_1 + 2||x_1||xR_2 \right)^{2k-2l+1} + \]

\[ + \frac{3}{2} \left| \frac{\cos(2x_0)}{1 - e_0} \right| \sum_{l=0}^{k} \frac{1}{(2k - 2l + 1)!} \left( \frac{R_2}{1 - e_0} \right)^{2l} \]

\[ \cdot \left( 2||\tilde{q}_{11}||xR_1 + 2||\tilde{q}_{12}||xR_1 + 2||x_1||xR_2 \right)^{2k-2l+1} + \]

\[ + \frac{1}{1 - e_0} \left( ||y_0||x e_0 \left( \frac{R_2}{1 - e_0} \right)^{2k+1} + \right. \]

\[ + \left( e_0 ||\tilde{q}_{21}||xR_1 + e_0 ||\tilde{q}_{22}||xR_2 + ||y_0||xR_2 + e_0 ||y_1||xR_2 \right) \cdot \left( \frac{1}{1 - e_0} \right)^{2k-1} \]

\[ + \left( ||\tilde{q}_{21}||xR_1R_2 + ||\tilde{q}_{22}||xR_1R_2 + ||y_1||xR_2^2 \right) \cdot \left( \frac{1}{1 - e_0} \right)^{2k-2} \]

\[ + \sum_{l=0}^{2k} (1 - e_0^2)^{1/2-l} \prod_{j=2}^{l} \frac{2j - 3 (e_0 + 2k - l)}{j} (1 - e_0)^{2k-l+2} R_2^{2k} \]
+ (e_0 \| \hat{q}_{21} \| \mathcal{R}_1 + e_0 \| \hat{q}_{22} \| \mathcal{R}_2 + \| y_0 \| \mathcal{R}_2 + e_0 \| y_1 \| \mathcal{R}_2) \cdot \\
\left( \frac{1}{1 - e_0} \right)^{2k} + \\
+ (\| \hat{q}_{21} \| \mathcal{R}_1 \mathcal{R}_2 + \| \hat{q}_{22} \| \mathcal{R}_1 \mathcal{R}_2 + \| y_1 \| \mathcal{R}_2^2) \cdot \\
\left( \frac{1}{1 - e_0} \right)^{2k-1} + \\
\sum_{l=0}^{2k+1} (1 - e_0^2)^{1/2-l} \prod_{j=2}^{l} \frac{2j - 3 (e_0 + 2k + 1 - l)}{(1 - e_0)^{2k-l+3}} \mathcal{R}_2^2, \\
\text{respectively.}
\]

Considering the estimates above we put

\[ h := \max \left\{ 2 \| \hat{q}_{11} \| \mathcal{R}_1 + 2 \| \hat{q}_{12} \| \mathcal{R}_1 + 2 \| x_1 \| \mathcal{R}_2, \right. \\
\left. \frac{4 \mathcal{R}_2}{1 - e_0}, \right. \\
e_0 \| \hat{q}_{21} \| \mathcal{R}_1 + e_0 \| \hat{q}_{22} \| \mathcal{R}_1 + \| y_0 \| \mathcal{R}_2 + e_0 \| y_1 \| \mathcal{R}_2, \\
\| \hat{q}_{21} \| \mathcal{R}_1 \mathcal{R}_2 + \| \hat{q}_{22} \| \mathcal{R}_1 \mathcal{R}_2 + \| y_1 \| \mathcal{R}_2^2 \right\}. \\
\]

Then the term of order \( k \) in the expansion of the first expression may be estimated by

\[ 3 \cdot \frac{\| \sin(2x_0) \| \mathcal{R} + \| \cos(2x_0) \| \mathcal{R}}{1 - e_0} \cdot h^k, \]

where we used that

\[ \sum_{l=0}^{k} \frac{1}{(2k - 2l)!} \leq 2 \quad \text{and} \quad \sum_{l=0}^{k-1} \frac{1}{(2k - 2l - 1)!} \leq 2. \]

For the term of order \( k \) in the expansion of the second expression we obtain the estimate

\[ \frac{1}{1 - e_0} (\| y_0 \| \mathcal{R}_0 + 2) \cdot h^k. \]

For the term of order \( k \) in the expansion of the third expression we find

\[ \frac{\sqrt{1 - e_0^2}}{(1 - e_0)^2} \sum_{l=0}^{k} \frac{c_0 + k - l}{(1 + e_0)^l} \prod_{j=2}^{l} \frac{2j - 3}{j} \left( \frac{\mathcal{R}_2}{1 - e_0} \right)^k \leq \]
8.5. Practical Stability for a Range of Values of $e$

$$\leq (k + 1) \cdot \frac{\sqrt{1 - e_0^2}}{(1 - e_0)^2} \sum_{l=0}^{k} \frac{1}{(1 + e_0)^l} \cdot 2^{l-1} \left( \frac{R_2}{1 - e_0} \right)^k \leq$$

$$\leq \frac{k + 1}{2} \frac{\sqrt{1 - e_0^2}}{(1 - e_0)^2} \sum_{l=0}^{k} \left( \frac{2}{1 + e_0} \right)^l \left( \frac{R_2}{1 - e_0} \right)^k \leq$$

$$\leq \frac{k + 1}{2} \frac{\sqrt{1 - e_0^2}}{(1 - e_0)^2} \sum_{l=0}^{k} 2^l \left( \frac{R_2}{1 - e_0} \right)^k \leq$$

$$\leq \frac{(k + 1) \cdot 2^k}{(1 - e_0)^2} \left( \frac{R_2}{1 - e_0} \right)^k \leq$$

$$\leq \frac{1}{(1 - e_0)^2} \cdot h^k.$$ 

Combining these estimates we obtain

$$\|h_m^k\| \leq \left\| \max_{1 \leq i,j \leq 2} (Q^{-1})_{ij} \right\| \cdot \left( 3 \cdot \frac{\| \sin(2x_0) \|_\mathcal{F} + \| \cos(2x_0) \|_\mathcal{F}}{1 - e_0} + \frac{1}{1 - e_0} (\| y_0 \|_\mathcal{F} e_0 + 2) + \frac{1}{(1 - e_0)^2} \right) \cdot h^k$$

Thus for

$$E := \left\| \max_{1 \leq i,j \leq 2} (Q^{-1})_{ij} \right\| \cdot \left( 3 \cdot \frac{\| \sin(2x_0) \|_\mathcal{F} + \| \cos(2x_0) \|_\mathcal{F}}{1 - e_0} + \frac{1}{1 - e_0} (\| y_0 \|_\mathcal{F} e_0 + 2) + \frac{1}{(1 - e_0)^2} \right),$$

$$h := \max \left\{ 2\| \tilde{q}_{11} \|_{\mathcal{F} R_1} + 2\| \tilde{q}_{12} \|_{\mathcal{F} R_1} + 2\| x_1 \|_{\mathcal{F} R_2}, \frac{4R_2}{1 - e_0}, \right\}$$

$$\varepsilon_0 \| \tilde{q}_{21} \|_{\mathcal{F} R_1} + e_0 \| \tilde{q}_{22} \|_{\mathcal{F} R_1} + \| y_0 \|_{\mathcal{F} R_2} + e_0 \| y_1 \|_{\mathcal{F} R_2},$$

$$\| \tilde{q}_{21} \|_{\mathcal{F} R_1 R_2} + \| \tilde{q}_{22} \|_{\mathcal{F} R_1 R_2} + \| y_1 \|_{\mathcal{F} R_2^2} \right\}$$

we have

$$\|h_m^k\| \leq E h^k$$

(8.12)
Practical Stability for Small Periodic Solutions

In Section 8.2 we showed that there exist small linearly stable $2\pi$-periodic solutions of (2.9) corresponding to Branch I in Figure 8.24.

![Figure 8.24: Linear Stability of $2\pi$-Periodic Solutions.](image)

In Section 8.4 we verified practical stability for prescribed values of the eccentricity $e$. Now we are in the position to verify practical stability for a whole range of $e$.

On the Choice of the Norms

In Section 6.1 we defined several norms. They depend on a vector $R$ of positive numbers. In our case this vector has dimension 2:

$$R = (R_2, R_{\Delta e}) =: (R_1, R_2) \quad \text{with} \quad R_3 > 0, R_2 > 0.$$  

The norms of $S$ and $z_3^0(0)$ defined in Lemma 6.1.1 and Lemma 6.1.2 are simplified to

$$\|S\|_{R_1^{-1}} = \frac{1}{R_1} \left( s_{11}^2 + s_{12}^2 + s_{21}^2 + s_{22}^2 \right)^{\frac{1}{2}},$$

$$\|z_3^0(0)\|_{R_{\Delta e}} = \frac{R_2}{R_1} \left( (z_3^0(0))^2_1 + (z_3^0(0))^2_2 \right)^{\frac{1}{2}}.$$

In Theorem 7.1.2 we defined the number

$$\bar{\rho}_{\text{tol}} := \left( \|S\|_{R_1^{-1}}^{-1} - \|z_3^0(0)\|_{R_1/R_{\Delta e}} \right) \cdot \rho_{\text{tol}} =$$

$$= \left( R_1 \left( s_{11}^2 + s_{12}^2 + s_{21}^2 + s_{22}^2 \right)^{\frac{1}{2}} - \frac{R_2}{R_1} \left( (z_3^0(0))^2_1 + (z_3^0(0))^2_2 \right)^{\frac{1}{2}} \right).$$
8.5. Practical Stability for a Range of Values of $e$

Obviously our method for practical stability breaks down for

$$
(s_i^2 + s_{1i}^2 + s_{2i}^2 + s_{22i})^{ \frac{3}{2}} - \frac{R_2}{R_1} \left((z_3^0(0))_1 + (z_3^0(0))_2\right)^{ \frac{3}{2}} = 0.
$$

For given $R_1$ this determines an upper bound for $R_2$. In Figure 8.25 we show $\hat{\rho}_{tol}$ as a function of $R_2$ for $R_1 = 1$ and for several values $e_i$ of the eccentricity $e$.

To keep the loss small we have to choose $R_2$ small, but then

$$\frac{1}{R_2} |e - e_i| =: |e - e_i|_{R_0,e} \leq \rho_{tol}
$$

(cf. Theorem 7.1.2) implies that $|e - e_i|$ is small too. Hence we must compromise. The following choice turns out to be suitable:

$$R_1 := 1 \quad \text{and} \quad R_2 := 0.4.
$$

**On the Ratio $\rho_{start} / \rho_{tol}$**

Our method for practical stability also breaks down, if the numerator in the formula for the maximum time

$$T_{max} := \min_{1 \leq i \leq n} \frac{\frac{1}{2} R_i^2 (\hat{\rho}_{tol} - \hat{\rho}_{start}) - 2 \sum_{k=r+1}^r \|I_k\|_{R_{\hat{R}_{tol}}}^k}{\sum_{k=r+1}^\infty \|I_k\|_{R_{\hat{R}_{tol}}}^k}
$$

(cf. Theorem 7.1.2) becomes negative. Let $q^{(r)}_{break}$ denote the ratio of $\rho_{start}$ and $\rho_{tol}$ where this break-down occurs.
Using the definition of $\tilde{\rho}_{\text{start}}$ and $\tilde{\rho}_{\text{tol}}$ (cf. Theorem 7.1.2)

$$
\tilde{\rho}_{\text{start}} := \|S^{-1}\|_{R_1} \cdot \rho_{\text{start}} + \|z_3^0(0)\|_{R_1/R_2} \cdot \rho_{\text{tol}},
$$

$$
\tilde{\rho}_{\text{tol}} := \left(\|S\|^{-1}_{R_1^{-1}} - \|z_3^0(0)\|_{R_1/R_2}\right) \cdot \rho_{\text{tol}}
$$

we obtain

$$
q_{\text{break}}^{(r)} = \frac{\rho_{\text{start}}}{\rho_{\text{tol}}} = \frac{\|S\|^{-1}_{R_1} \tilde{\rho}_{\text{start}}}{\rho_{\text{tol}}} = \frac{\left(\|S\|^{-1}_{R_1} - \|z_3^0(0)\|_{R_1/R_2}\right)^{-1}}{\|S\|^{-1}_{R_1} \rho_{\text{start}} - \|z_3^0(0)\|_{R_1/R_2} \rho_{\text{tol}}} = \|S\|^{-1}_{R_1} \left(\|S\|^{-1}_{R_1^{-1}} - \|z_3^0(0)\|_{R_1/R_2}\right)
$$

$$
\leq \|S\|^{-1}_{R_1} \left(\|S\|^{-1}_{R_1^{-1}} - 2\|z_3^0(0)\|_{R_1/R_2}\right) =: \approx q_{\text{break}}^{(r)}.
$$

In Table 8.21 we present the values of $q_{\text{break}}^{\approx}$ and $q_{\text{break}}^{(r)}$ for $\rho_{\text{tol}} = 10^{-2}$.

**Tab. 8.21:** Comparison of $q_{\text{break}}^{\approx}$ and $q_{\text{break}}^{(r)}$ for $\rho_{\text{tol}} = 10^{-2}$ and several values of $e_0$.

<table>
<thead>
<tr>
<th>$e_0$</th>
<th>$q_{\text{break}}^{\approx}$</th>
<th>$q_{\text{break}}^{(r)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.03301270189</td>
<td>0.03155892929</td>
</tr>
<tr>
<td>0.0125</td>
<td>0.04369997308</td>
<td>0.04213755706</td>
</tr>
<tr>
<td>0.025</td>
<td>0.05300124025</td>
<td>0.05132626763</td>
</tr>
<tr>
<td>0.0375</td>
<td>0.06081710359</td>
<td>0.0590263612</td>
</tr>
<tr>
<td>0.05</td>
<td>0.06704775442</td>
<td>0.06513899093</td>
</tr>
<tr>
<td>0.0625</td>
<td>0.07159261787</td>
<td>0.06956466903</td>
</tr>
<tr>
<td>0.075</td>
<td>0.07435016972</td>
<td>0.0722038797</td>
</tr>
<tr>
<td>0.0875</td>
<td>0.07521779375</td>
<td>0.07211370023</td>
</tr>
<tr>
<td>0.1</td>
<td>0.07409274819</td>
<td>0.0717161609</td>
</tr>
</tbody>
</table>

**Estimates of the Perturbation**

For our computations we need the norms of the functions $h_{1,2}^k$. For small order $k$ we use the norms obtained by explicit computations. For larger $k$ we make use of the estimates (8.12) given above.
8.5. Practical Stability for a Range of Values of $e$

Figure 8.26 gives a comparison of the explicitly computed norms with the exponential estimates (8.12).

**On the Range of the Eccentricity $e$**

In Theorem 7.1.2 we proved that for any solution $(x(., e), y(., e))$ and any value of the parameter $e$ with

\[
\frac{1}{R_1} \sqrt{(x(0, e) - x_0(0))^2 + (y(0, e) - y_0(0))^2} \leq \rho_{\text{start}} \quad \text{and} \quad \frac{1}{R_2} |e - e_0| = \frac{1}{R_2} |\Delta e| \leq \rho_{\text{col}}
\]

one has

\[
\frac{1}{R_1} \sqrt{(x(nT, e) - x_0(nT))^2 + (y(nT, e) - y_0(nT))^2} < \rho_{\text{col}} \quad \text{for} \quad 0 \leq nT < T_{\text{max}},
\]

where the maximum time $T_{\text{max}}$ is defined by

\[
T_{\text{max}} := \min_{1 \leq r \leq n} \frac{\frac{1}{2} R_1^2 (\tilde{\rho}_{\text{col}}^2 - \tilde{\rho}_{\text{start}}^2) - 2 \sum_{k=3}^r \|I_k^2\| R_1 \tilde{\rho}_{\text{col}}^k}{\sum_{k=r+1}^\infty \|I_k\| R_1 \tilde{\rho}_{\text{col}}^k}.
\]

This result allows us to verify practical stability for a certain neighborhood $U_{\rho_{\text{col}}}(e_0)$ of a prescribed value of the eccentricity $e_0$ (cf. Figure 8.27).
300 8. Stability of the Dumbbell Satellite

Now let \([e_{\text{min}}, e_{\text{max}}]\) be some interval of the parameter range and let
\[e_1 := e_{\text{min}}, e_2, \ldots, e_{n-1}, e_n := e_{\text{max}}\]
be some subdivision of \([e_{\text{min}}, e_{\text{max}}]\). Then we may apply Theorem 7.1.2 for \(e_0 := e_i, \leq i \leq n\). Thus we may verify practical stability for a series of subintervals \(U_{\rho_{\text{tol}}}(e_i)\) of \([e_{\text{min}}, e_{\text{max}}]\).

For \(\rho_{\text{tol}}\) too small these subintervals \(U_{\rho_{\text{tol}}}(e_i)\) do not cover \([e_{\text{min}}, e_{\text{max}}]\) as illustrated in Figure 8.28.

But for
\[\rho_{\text{tol}} \geq \frac{1}{2R_2} \max_{1 \leq i \leq n-1} |e_{i+1} - e_i|\]
the subintervals \(U_{\rho_{\text{tol}}}(e_i)\) cover the whole interval \([e_{\text{min}}, e_{\text{max}}]\) as shown in Figure 8.29.
8.5. Practical Stability for a Range of Values of $\epsilon$

Computer Assisted Results

We are now in the position to present some results based on explicit computations performed on a computer algebra system.

As in Section 8.4 we use a mix of explicit computations and exponential estimates for the denominator (cf. Section 7.2).

Let $\rho_{\text{start}} := 0.03\rho_{\text{tol}}$ and let the truncation order be $r = 15$.

Figure 8.30 and Table 8.22 give a survey of $T_{\text{max}}/2\pi$ as a function of $\rho_{\text{tol}}$ for

\begin{align*}
\epsilon_1 &= 0, \quad \epsilon_2 = 0.0125, \quad \epsilon_3 = 0.025, \quad \epsilon_4 = 0.0375, \quad \epsilon_5 = 0.05, \\
\epsilon_6 &= 0.0625, \quad \epsilon_7 = 0.075, \quad \epsilon_8 = 0.0875, \quad \epsilon_9 = 0.1.
\end{align*}

The second column indicates the width of the neighborhood $U_{\rho_{\text{tol}}}(\epsilon_i)$. Note that the neighborhoods do not cover the whole interval $[0, 0.1]$.

![Graph showing $T_{\text{max}}/2\pi$ vs. $\rho_{\text{tol}}$](image)

**Fig. 8.30:** Dependence of $T_{\text{max}}/2\pi$ on $\rho_{\text{tol}}$ for several subintervals of width $2R_2\rho_{\text{tol}}$ centered at $\epsilon_i$.

For

$$
\rho_{\text{tol}} \geq \frac{1}{R_2} \max_{1 \leq i \leq n-1} |e_{i+1} - e_i| = \frac{1}{2 \cdot 0.4} \cdot 0.0125 = 0.015625
$$
8. Stability of the Dumbbell Satellite

### Tab. 8.22: Dependence of $T_{\text{max}}/2\pi$ on $\rho_{\text{tol}}$ for various values of $e$.

<table>
<thead>
<tr>
<th>$\rho_{\text{tol}}$</th>
<th>$2R_2\rho_{\text{tol}}$</th>
<th>$e_1 = 0.$</th>
<th>$e_2 = 0.0125$</th>
<th>$e_3 = 0.025$</th>
<th>$e_4 = 0.0375$</th>
<th>$e_5 = 0.05$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-9}$</td>
<td>$0.8 \cdot 10^{-9}$</td>
<td>$1. \cdot 10^{-26}$</td>
<td>$2. \cdot 10^{-26}$</td>
<td>$1. \cdot 10^{-26}$</td>
<td>$6. \cdot 10^{-25}$</td>
<td>$3. \cdot 10^{-25}$</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>$0.8 \cdot 10^{-8}$</td>
<td>$1. \cdot 10^{-14}$</td>
<td>$2. \cdot 10^{-14}$</td>
<td>$1. \cdot 10^{-14}$</td>
<td>$6. \cdot 10^{-14}$</td>
<td>$3. \cdot 10^{-14}$</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>$0.8 \cdot 10^{-7}$</td>
<td>$1. \cdot 10^{-28}$</td>
<td>$2. \cdot 10^{-28}$</td>
<td>$1. \cdot 10^{-28}$</td>
<td>$6. \cdot 10^{-28}$</td>
<td>$3. \cdot 10^{-28}$</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>$0.8 \cdot 10^{-6}$</td>
<td>$1. \cdot 10^{-30}$</td>
<td>$2. \cdot 10^{-30}$</td>
<td>$1. \cdot 10^{-30}$</td>
<td>$6. \cdot 10^{-30}$</td>
<td>$3. \cdot 10^{-30}$</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>$0.8 \cdot 10^{-5}$</td>
<td>$1. \cdot 10^{-31}$</td>
<td>$2. \cdot 10^{-31}$</td>
<td>$1. \cdot 10^{-31}$</td>
<td>$6. \cdot 10^{-31}$</td>
<td>$3. \cdot 10^{-31}$</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>$0.8 \cdot 10^{-4}$</td>
<td>$1. \cdot 10^{-32}$</td>
<td>$2. \cdot 10^{-32}$</td>
<td>$1. \cdot 10^{-32}$</td>
<td>$6. \cdot 10^{-32}$</td>
<td>$3. \cdot 10^{-32}$</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>$0.8 \cdot 10^{-3}$</td>
<td>$7. \cdot 10^{-41}$</td>
<td>$8. \cdot 10^{-41}$</td>
<td>$8. \cdot 10^{-41}$</td>
<td>$4. \cdot 10^{-41}$</td>
<td>$2. \cdot 10^{-41}$</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>$0.8 \cdot 10^{-2}$</td>
<td>$1. \cdot 10^{-27}$</td>
<td>$2. \cdot 10^{-27}$</td>
<td>$2. \cdot 10^{-27}$</td>
<td>$1. \cdot 10^{-27}$</td>
<td>$6. \cdot 10^{-26}$</td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>$0.8 \cdot 10^{-1}$</td>
<td>$3. \cdot 10^{-11}$</td>
<td>$6. \cdot 10^{-11}$</td>
<td>$6. \cdot 10^{-11}$</td>
<td>$3. \cdot 10^{-11}$</td>
<td>$6. \cdot 10^{-11}$</td>
</tr>
</tbody>
</table>

The subintervals $U_{\rho_{\text{tol}}}(e_i)$ cover the whole interval $[0, 0.1]$. Tables 8.23 and Figure 8.31 give a survey of $T_{\text{max}}/2\pi$ as a function of $\rho_{\text{tol}}$. Note that for $\rho_{\text{tol}} < 0.01$ the coverage of the interval $[0, 0.1]$ is complete. We summarize the last results in the Table 8.24.

We complete this section on computer assisted results with some computations concerning the interval $[0.2, 0.25]$. For this interval we use the grid

$$e_1 = 0.2, \quad e_2 = 0.21, \quad e_3 = 0.22, \quad e_4 = 0.23, \quad e_5 = 0.24, \quad e_6 = 0.25$$

and put $\rho_{\text{start}} := 0.08\rho_{\text{tol}}$ and $R_1 = 1$ and $R_2 = 0.2$. The smallest $\rho_{\text{tol}}$ that leads to a complete coverage of the interval is

$$\rho_{\text{tol}} = \frac{1}{2R_2} \max_{1 \leq i \leq 4} |e_{i+1} - e_i| = \frac{1}{2 \cdot 0.2} \cdot 0.01 = 0.025.$$ 

For this $\rho_{\text{tol}}$ we obtain the following result.

For $\rho_{\text{tol}} = 0.025$ and $\rho_{\text{start}} = 0.08\rho_{\text{tol}}$ the maximum time $T_{\text{max}}/2\pi = 7 \cdot 10^7$ is established for the whole interval $[0.195, 0.255]$.

For values of the eccentricity $e$ near $e_7 = 0.26$ the upper bound of $\rho_{\text{tol}}$ becomes very small. This prohibits the coverage of a larger interval as $[0.2, 0.26]$ or even $[0.2, 0.3]$ with reasonable effort.
8.5. Practical Stability for a Range of Values of \( \epsilon \)

### Tab. 8.23: Dependence of \( T_{\text{max}}/2\pi \) on \( \rho_{\text{tol}} \) for various values of \( \epsilon \).

<table>
<thead>
<tr>
<th>( \rho_{\text{tol}} )</th>
<th>( 2R_{2}\rho_{\text{tol}} )</th>
<th>( \epsilon_1 = 0.0 )</th>
<th>( \epsilon_2 = 0.0125 )</th>
<th>( \epsilon_3 = 0.025 )</th>
<th>( \epsilon_4 = 0.0375 )</th>
<th>( \epsilon_5 = 0.05 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.008</td>
<td>1. ( \cdot 10^{27} )</td>
<td>3. ( \cdot 10^{27} )</td>
<td>2. ( \cdot 10^{27} )</td>
<td>1. ( \cdot 10^{27} )</td>
<td>6. ( \cdot 10^{26} )</td>
</tr>
<tr>
<td>0.02</td>
<td>0.016</td>
<td>2. ( \cdot 10^{21} )</td>
<td>1. ( \cdot 10^{23} )</td>
<td>7. ( \cdot 10^{22} )</td>
<td>4. ( \cdot 10^{22} )</td>
<td>2. ( \cdot 10^{22} )</td>
</tr>
<tr>
<td>0.03</td>
<td>0.024</td>
<td>-</td>
<td>2. ( \cdot 10^{20} )</td>
<td>2. ( \cdot 10^{20} )</td>
<td>9. ( \cdot 10^{19} )</td>
<td>4. ( \cdot 10^{19} )</td>
</tr>
<tr>
<td>0.04</td>
<td>0.032</td>
<td>-</td>
<td>2. ( \cdot 10^{18} )</td>
<td>2. ( \cdot 10^{18} )</td>
<td>1. ( \cdot 10^{18} )</td>
<td>6. ( \cdot 10^{17} )</td>
</tr>
<tr>
<td>0.05</td>
<td>0.04</td>
<td>-</td>
<td>6. ( \cdot 10^{16} )</td>
<td>6. ( \cdot 10^{16} )</td>
<td>4. ( \cdot 10^{16} )</td>
<td>2. ( \cdot 10^{16} )</td>
</tr>
<tr>
<td>0.06</td>
<td>0.048</td>
<td>-</td>
<td>3. ( \cdot 10^{15} )</td>
<td>3. ( \cdot 10^{15} )</td>
<td>2. ( \cdot 10^{15} )</td>
<td>1. ( \cdot 10^{15} )</td>
</tr>
<tr>
<td>0.07</td>
<td>0.056</td>
<td>-</td>
<td>2. ( \cdot 10^{14} )</td>
<td>3. ( \cdot 10^{14} )</td>
<td>2. ( \cdot 10^{14} )</td>
<td>1. ( \cdot 10^{14} )</td>
</tr>
<tr>
<td>0.08</td>
<td>0.064</td>
<td>-</td>
<td>5. ( \cdot 10^{13} )</td>
<td>3. ( \cdot 10^{13} )</td>
<td>2. ( \cdot 10^{13} )</td>
<td>1. ( \cdot 10^{13} )</td>
</tr>
<tr>
<td>0.09</td>
<td>0.072</td>
<td>-</td>
<td>-</td>
<td>4. ( \cdot 10^{12} )</td>
<td>3. ( \cdot 10^{12} )</td>
<td>2. ( \cdot 10^{12} )</td>
</tr>
<tr>
<td>0.1</td>
<td>0.08</td>
<td>-</td>
<td>-</td>
<td>6. ( \cdot 10^{11} )</td>
<td>6. ( \cdot 10^{11} )</td>
<td>3. ( \cdot 10^{11} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \rho_{\text{col}} )</th>
<th>( 2R_{2}\rho_{\text{col}} )</th>
<th>( \epsilon_5 = 0.05 )</th>
<th>( \epsilon_6 = 0.0625 )</th>
<th>( \epsilon_7 = 0.075 )</th>
<th>( \epsilon_8 = 0.0875 )</th>
<th>( \epsilon_9 = 0.1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>0.008</td>
<td>6. ( \cdot 10^{26} )</td>
<td>3. ( \cdot 10^{26} )</td>
<td>2. ( \cdot 10^{26} )</td>
<td>9. ( \cdot 10^{25} )</td>
<td>5. ( \cdot 10^{25} )</td>
</tr>
<tr>
<td>0.02</td>
<td>0.016</td>
<td>2. ( \cdot 10^{22} )</td>
<td>1. ( \cdot 10^{22} )</td>
<td>5. ( \cdot 10^{21} )</td>
<td>3. ( \cdot 10^{21} )</td>
<td>1. ( \cdot 10^{21} )</td>
</tr>
<tr>
<td>0.03</td>
<td>0.024</td>
<td>4. ( \cdot 10^{19} )</td>
<td>2. ( \cdot 10^{19} )</td>
<td>1. ( \cdot 10^{19} )</td>
<td>6. ( \cdot 10^{18} )</td>
<td>3. ( \cdot 10^{18} )</td>
</tr>
<tr>
<td>0.04</td>
<td>0.032</td>
<td>6. ( \cdot 10^{17} )</td>
<td>3. ( \cdot 10^{17} )</td>
<td>2. ( \cdot 10^{17} )</td>
<td>8. ( \cdot 10^{16} )</td>
<td>4. ( \cdot 10^{16} )</td>
</tr>
<tr>
<td>0.05</td>
<td>0.04</td>
<td>2. ( \cdot 10^{16} )</td>
<td>9. ( \cdot 10^{15} )</td>
<td>5. ( \cdot 10^{15} )</td>
<td>3. ( \cdot 10^{15} )</td>
<td>1. ( \cdot 10^{15} )</td>
</tr>
<tr>
<td>0.06</td>
<td>0.048</td>
<td>1. ( \cdot 10^{15} )</td>
<td>6. ( \cdot 10^{14} )</td>
<td>3. ( \cdot 10^{14} )</td>
<td>2. ( \cdot 10^{14} )</td>
<td>8. ( \cdot 10^{13} )</td>
</tr>
<tr>
<td>0.07</td>
<td>0.056</td>
<td>1. ( \cdot 10^{14} )</td>
<td>5. ( \cdot 10^{13} )</td>
<td>3. ( \cdot 10^{13} )</td>
<td>1. ( \cdot 10^{13} )</td>
<td>7. ( \cdot 10^{12} )</td>
</tr>
<tr>
<td>0.08</td>
<td>0.064</td>
<td>1. ( \cdot 10^{13} )</td>
<td>6. ( \cdot 10^{12} )</td>
<td>3. ( \cdot 10^{12} )</td>
<td>2. ( \cdot 10^{12} )</td>
<td>8. ( \cdot 10^{11} )</td>
</tr>
<tr>
<td>0.09</td>
<td>0.072</td>
<td>2. ( \cdot 10^{12} )</td>
<td>1. ( \cdot 10^{12} )</td>
<td>5. ( \cdot 10^{11} )</td>
<td>3. ( \cdot 10^{11} )</td>
<td>1. ( \cdot 10^{11} )</td>
</tr>
<tr>
<td>0.1</td>
<td>0.08</td>
<td>3. ( \cdot 10^{11} )</td>
<td>2. ( \cdot 10^{11} )</td>
<td>9. ( \cdot 10^{10} )</td>
<td>5. ( \cdot 10^{10} )</td>
<td>2. ( \cdot 10^{10} )</td>
</tr>
</tbody>
</table>

### Tab. 8.24: Dependence of \( T_{\text{max}}/2\pi \) on \( \rho_{\text{col}} \) for various values of \( \epsilon \).

<table>
<thead>
<tr>
<th>( \rho_{\text{start}} )</th>
<th>( \rho_{\text{tol}} )</th>
<th>Interval of Eccentricity</th>
<th>( T_{\text{max}}/2\pi )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0006</td>
<td>0.02</td>
<td>([0, 0.108])</td>
<td>1. ( \cdot 10^{21} )</td>
</tr>
<tr>
<td>0.0009</td>
<td>0.03</td>
<td>([0, 0.112])</td>
<td>3. ( \cdot 10^{18} )</td>
</tr>
<tr>
<td>0.0012</td>
<td>0.04</td>
<td>([0, 0.116])</td>
<td>4. ( \cdot 10^{16} )</td>
</tr>
<tr>
<td>0.0015</td>
<td>0.05</td>
<td>([0, 0.12])</td>
<td>1. ( \cdot 10^{15} )</td>
</tr>
<tr>
<td>0.0018</td>
<td>0.06</td>
<td>([0, 0.124])</td>
<td>8. ( \cdot 10^{13} )</td>
</tr>
<tr>
<td>0.0021</td>
<td>0.07</td>
<td>([0, 0.128])</td>
<td>7. ( \cdot 10^{12} )</td>
</tr>
<tr>
<td>0.0024</td>
<td>0.08</td>
<td>([0, 0.132])</td>
<td>8. ( \cdot 10^{11} )</td>
</tr>
<tr>
<td>0.0027</td>
<td>0.09</td>
<td>([0, 0.136])</td>
<td>1. ( \cdot 10^{11} )</td>
</tr>
<tr>
<td>0.003</td>
<td>0.1</td>
<td>([0, 0.14])</td>
<td>2. ( \cdot 10^{10} )</td>
</tr>
</tbody>
</table>
It seems that this behavior is due to the fact that a branch of 6\pi-periodic solutions bifurcate from Branch I for $e \approx 0.26$ (cf. Figure 8.5).

**Nekoroshev-type Results**

We present results of Nekoroshev-type for 2\pi-periodic solutions corresponding to Branch I. In Table 8.25 we compare the results obtained by explicit computations and by a priori estimates for $e_0 = 0.1$. The optimal truncation order for the a priori estimates is given in the tables while the truncation order of the explicit computations is kept at $r = 15$. Note that for very small $\rho_{tol}$ the optimal truncation order $r_{opt}$ may become larger than the denominator $q_N$ of the best convergent of $\omega_1/\omega_0$. Thus the a priori method breaks down. Similar results may be obtained for $e_1, \ldots, e_8$. Note that the stability neighborhoods $U_{\rho_{tol}}(e_i)$ are too small to cover the whole interval $[0, 0.1]$. For decreasing $\rho_{tol}$ they collapse to single points.
8.6 The Influence of the Coordinate Transformation

In this section we discuss the influence of the coordinate transformation on the results on practical stability.

The Linear Transformation

For $\epsilon_0 = 0$ the linear transformation has the form

$$S = \begin{pmatrix} s_{11} & s_{12} \\ -\sqrt{3} s_{11} & \sqrt{3} s_{12} \end{pmatrix}.$$  

The results in the Section 8.3 are obtained for $s_{11} = 1$ and $s_{12} = 0$. At this place we study the influence of the particular choice of the transformation matrix $S$ on $T_{\text{max}}/2\pi$.

First we give an estimate for $T_{\text{max}}/2\pi$ using the norms computed in Section 8.3:

$$T_{\text{max}} = \frac{1}{2} R^2 (\tilde{p}_{\text{tol}}^2 - \tilde{p}_{\text{start}}^2) - 2 \sum_{k=3}^{2r} ||I^k||_R R^k \tilde{p}_{\text{tol}}^k$$

$$= \frac{1}{2} R^2 \left( ||S||_{R^k}^2 \tilde{p}_{\text{tol}}^2 - ||S^{-1}||^2_{R^k} \tilde{p}_{\text{start}}^2 \right) - 2 \sum_{k=3}^{2r} ||R^k||_R R^k ||S||_{R^k}^{-k} \tilde{p}_{\text{tol}}^k$$

$$\geq \frac{1}{2} R^2 \left( \frac{1}{(s_{11}^2 + s_{12}^2)^2} \left( \frac{1}{3} \tilde{p}_{\text{tol}}^2 - \frac{4}{3} \tilde{p}_{\text{start}}^2 \right) - 2 \sum_{k=3}^{\infty} \frac{1}{(s_{11}^2 + s_{12}^2)^2} \left( 2s_{11} + 2s_{12} \right)^{2k} \frac{1}{(2k)!} \frac{1}{(2k-1)!} \frac{1}{(s_{11}^2 + s_{12}^2)^2} R^{2k} \rho_{\text{tol}}^{2k} \right.$$
8. Stability of the Dumbbell Satellite

\[
\frac{1}{2} \left( \frac{1}{4} \rho_{\text{tol}}^2 - \frac{4}{3} \rho_{\text{start}}^2 \right)- \frac{1}{2} \sum_{k=3}^{\infty} \frac{1}{(2k)!} \frac{((|s_{11}| + |s_{12}|)^{2k}}{(s_{11}^2 + s_{12}^2)^k} R^{2k} \rho_{\text{tol}}^{2k} =: \tilde{T}_{\text{max}}.
\]

For \( \rho_{\text{tol}} \) small we further have

\[
\tilde{T}_{\text{max}} \approx \frac{1}{2} \left( \frac{1}{4} \rho_{\text{tol}}^2 - \frac{4}{3} \rho_{\text{start}}^2 \right) \frac{1}{(2^{\frac{1}{2}})^{2k} - 1}! \frac{((|s_{11}| + |s_{12}|)^{2k}}{(s_{11}^2 + s_{12}^2)^k} R^{2k} \rho_{\text{tol}}^{2k}.
\]

The term

\[
\frac{\frac{(|s_{11}| + |s_{12}|)^2}{s_{11}^2 + s_{12}^2}}
\]

is (i) minimal for either \( s_{11} = 0 \) or \( s_{12} = 0 \) and (ii) maximal for \( s_{11} = s_{12} \neq 0 \). Thus we find

\[
\frac{\tilde{T}_{\text{max}}^{(ii)}}{\tilde{T}_{\text{max}}^{(i)}} \approx 2^{-\frac{r+1}{2}}.
\]

For example let \( r = 20 \). Then in the worst case (ii) we roughly loose a factor \( 2 \cdot 10^3 \) compared with the best case (i). In reality the loss is a bit smaller. The difference results from the estimate of \( f^k \) that we used in our consideration.

In the general case \( e_0 \neq 0 \) the matrix \( G^{10} \) (cf. Theorem 4.2.1) has the form

\[
G^{10} = \begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix}.
\]

Then the linear transformation has the form

\[
S := \begin{pmatrix} s_{11} \\ -\frac{g_{11}}{g_{12}} s_{12} \\ \sqrt{-\frac{g_{11}}{g_{12}} s_{11}} \end{pmatrix}.
\]

We may therefore expect that the results obtained above still hold true.

The Floquet Transformation

In the non-autonomous case the Floquet-transformation removes the time-dependency of the linear part. Again the transformation matrix \( Q(t) \) is not uniquely determined. Every matrix of the form

\[
Q(t) := Q(t) \left( \begin{array}{cc} \cos(2jt) & \sin(2jt) \\ -\sin(2jt) & \cos(2jt) \end{array} \right), \quad j \in \mathbb{Z}.
\]

removes the time-dependency of the linear part as well (cf. Chapter 4). The simplicity of the case \( e_0 = 0 \) allows us to study the influence of the particular choice of \( Q \) on the results on practical stability.
8.6. The Influence of the Coordinate Transformation

To this end consider the system of differential equations

\[
\begin{align*}
\dot{x} &= \sqrt{3}y \\
\dot{y} &= -\sqrt{3}x - \left( \frac{\sqrt{3}}{2} \sin(2x) - \sqrt{3}x \right)
\end{align*}
\] (8.2)

Since this system is autonomous the Floquet-transformation may be chosen trivially \( Q(t) \equiv I \). Thus we have

\[
\dot{Q}(t) = \begin{pmatrix} \cos(2jt) & \sin(2jt) \\ -\sin(2jt) & \cos(2jt) \end{pmatrix}.
\]

For the left-hand side of (8.2) we have

\[
\frac{d}{dt} \begin{pmatrix} \cos(2jt) & \sin(2jt) \\ -\sin(2jt) & \cos(2jt) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \end{align*}

\[
= \begin{pmatrix} -2j \sin(2jt) & 2j \cos(2jt) \\ -2j \cos(2jt) & -2j \sin(2jt) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} \cos(2jt) & \sin(2jt) \\ -\sin(2jt) & \cos(2jt) \end{pmatrix} \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}
\]

and for the right-hand side

\[
\begin{pmatrix} 0 & \sqrt{3} \\ -\sqrt{3} & 0 \end{pmatrix} \begin{pmatrix} \cos(2jt) & \sin(2jt) \\ -\sin(2jt) & \cos(2jt) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} +
\]

\[
= \begin{pmatrix} -\sqrt{3} \sin(2jt) & \sqrt{3} \cos(2jt) \\ -\sqrt{3} \cos(2jt) & -\sqrt{3} \sin(2jt) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} +
\]

\[
+ \begin{pmatrix} 0 \\ -\sqrt{3} \sin(2jt) \\ \sqrt{3} \cos(2jt) \\ \sqrt{3} \cos(2jt) \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -\sqrt{3} \sin(2jt) \\ \sqrt{3} \cos(2jt) \end{pmatrix}.
\]

Thus we find

\[
\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} \cos(2jt) & -\sin(2jt) \\ \sin(2jt) & \cos(2jt) \end{pmatrix} \begin{pmatrix} \left( -\sqrt{3} - 2j \right) \sin(2jt) & \left( \sqrt{3} - 2j \right) \cos(2jt) \\ \left( -\sqrt{3} - 2j \right) \cos(2jt) & \left( \sqrt{3} - 2j \right) \sin(2jt) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} +
\]

\[
+ \begin{pmatrix} \left( -\sqrt{3} - 2j \right) \sin(2jt) & \left( \sqrt{3} - 2j \right) \cos(2jt) \\ \left( -\sqrt{3} - 2j \right) \cos(2jt) & \left( \sqrt{3} - 2j \right) \sin(2jt) \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\]

Therefore the transformed system reads

\[
\begin{align*}
\dot{x} &= (\sqrt{3} - 2j)y + \frac{\sqrt{3}}{2} \sin(2jt) \cdot \left( \sin(2\cos(2jt)x + \sin(2jt)y) - 2\cos(2jt)x + \sin(2jt)y \right) \\
\dot{y} &= - (\sqrt{3} - 2j)x - \frac{\sqrt{3}}{2} \cos(2jt) \cdot \left( \sin(2\cos(2jt)x + \sin(2jt)y) - 2\cos(2jt)x + \sin(2jt)y \right).
\end{align*}
\] (8.13)
Lemma 8.6.1

Let

\[ \dot{I}(x, y) := \frac{1}{2}(x^2 + y^2) - \frac{1}{4}(-1 + 2(\cos(2jt)x + \sin(2jt)y)^2 + \cos(2(\cos(2jt)x + \sin(2jt)y))). \]

Then \( \dot{I}(x, y) \) is a first integral of (8.13).

Proof: For the derivative of \( \dot{I}(x, y) \) we obtain

\[
\begin{align*}
\frac{d}{dt} \dot{I}(t, x, y) &= x\dot{x} + y\dot{y} + \frac{1}{2} \left( \sin(2(\cos(2jt)x + \sin(2jt)y)) - 2(\cos(2jt)x + \sin(2jt)y) \right) \cdot \\
&\quad \cdot \left( -2j \sin(2jt)x + 2j \cos(2jt)y + \cos(2jt)\dot{x} + \sin(2jt)\dot{y} \right) = \\
&= x(\sqrt{3} - 2j)y + \\
&\quad + \frac{\sqrt{3}}{2} x \sin(2jt) \left( -2j \sin(2jt)x + 2j \cos(2jt)y + \cos(2jt)\dot{x} + \sin(2jt)\dot{y} \right) - \\
&\quad - y(\sqrt{3} - 2j)x - \\
&\quad - \frac{\sqrt{3}}{2} y \cos(2jt) \left( -2j \sin(2jt)x + 2j \cos(2jt)y + \cos(2jt)\dot{x} + \sin(2jt)\dot{y} \right) + \\
&\quad + \frac{1}{2} \left( \sin(2(\cos(2jt)x + \sin(2jt)y)) - 2(\cos(2jt)x + \sin(2jt)y) \right) \cdot \\
&\quad \cdot \left( -2j \sin(2jt)x + 2j \cos(2jt)y + \\
&\quad \quad \quad + \cos(2jt)(\sqrt{3} - 2)y + \frac{\sqrt{3}}{2} \sin(2jt) \cos 2jt \cdot \\
&\quad \quad \quad \cdot \left( -2j \sin(2jt)x + 2j \cos(2jt)y + \cos(2jt)\dot{x} + \sin(2jt)\dot{y} \right) - \\
&\quad \quad \quad - \sin(2jt)(\sqrt{3} - 2)x - \frac{\sqrt{3}}{2} \cos(2jt) \sin 2jt \cdot \\
&\quad \quad \quad \cdot \left( -2j \sin(2jt)x + 2j \cos(2jt)y + \cos(2jt)\dot{x} + \sin(2jt)\dot{y} \right) \right) = \\
&= 0.
\end{align*}
\]
Thus \( \hat{I}(t, x, y) \) is a first integral of (8.13).

The truncation of \( \hat{I}(x, y) \) reads

\[
\hat{I}^{(r)}(t, x, y) = \sum_{k=2}^{r} \hat{I}^k(t, x, y) = \frac{1}{2} (x^2 + y^2) - \frac{1}{4} \sum_{k=2}^{\frac{r}{2}} (-1)^k \frac{(2 \cos(2jt)x + 2 \sin(2jt)y)^{2k}}{(2k)!} \]

and for the derivative we obtain

\[
\frac{d}{dt} \hat{I}^{(r)}(t, x, y) = \sum_{k=r+1}^{\infty} \hat{I}^k \neq 0,
\]

where the polynomials \( \hat{I}^k \) are given by (cf. Lemma 6.4.1)

\[
\hat{I}^k(t, x, y) = x\hat{h}_1^{k-1} + y\hat{h}_2^{k-1} + \sum_{l=3}^{r} \left( \frac{\partial \hat{I}^l}{\partial x}\hat{h}_1^{k-l+1} + \frac{\partial \hat{I}^l}{\partial y}\hat{h}_2^{k-l+1} \right),
\]

where

\[
\hat{h}_1(t, x, y)^{2k+1} = \frac{\sqrt{3}}{2} \sin(2jt)(-1)^k \frac{(2 \cos(2jt)x + 2 \sin(2jt)y)^{2k+1}}{(2k+1)!},
\]

\( \hat{h}_1(t, x, y)^{2k} = 0, \)

\[
\hat{h}_2(t, x, y)^{2k+1} = -\frac{\sqrt{3}}{2} \cos(2jt)(-1)^k \frac{(2 \cos(2jt)x + 2 \sin(2jt)y)^{2k+1}}{(2k+1)!},
\]

\( \hat{h}_2(t, x, y)^{2k} = 0 \)

are the terms of the power series expansion of the right-hand side of (8.13).

For the approximate first integral \( \hat{I}^{(r)}(x, y) \) we find

\[
\hat{I}^{(r)}(t, x, y) = \sum_{k=2}^{r} \hat{I}^k(x, y) = \frac{1}{2} (x^2 + y^2) - \frac{1}{4} \sum_{k=2}^{\frac{r}{2}} (-1)^k \frac{(2 \cos(2jt)x + 2 \sin(2jt)y)^{2k}}{(2k)!},
\]

Since \( \hat{h}^{2k}_i = 0 \) and \( \frac{\partial \hat{h}_i}{\partial x} = 0 \) if \( l \) is odd and \( \hat{h}_i^{(2k+1)-l+1} = 0 \) if \( l \) is even, we have \( \hat{I}^{2k+1} = 0. \)

For \( \hat{I}^{2k} \) we obtain for \( \hat{I}(x, y) \)

\[
\hat{I}^{2k} = x\hat{h}_1^{2k-1} + y\hat{h}_2^{2k-1} + \sum_{l=2}^{\frac{r}{2}} \left( \frac{\partial \hat{I}^{2l}}{\partial x}\hat{h}_1^{2k-2l+1} + \frac{\partial \hat{I}^{2l}}{\partial y}\hat{h}_2^{2k-2l+1} \right) = \]
\[ \dot{\mathbf{i}}(t, x, y) = \sum_{k=0}^{\infty} \dot{\mathbf{i}}^k(x, y) = \frac{\sqrt{3}}{2} (\sin(2jt)x - \cos(2jt)y) \sum_{k=\lfloor \frac{l}{2} \rfloor + 1}^{\infty} (-1)^k \frac{(\cos(2jt)x + \sin(2jt)y)^{2k-1}}{(2k-1)!}. \]

This leads to the following norms (cf. (8.3)):

\[ \| \dot{Q}(0) \|_R = \sqrt{2}, \quad \| \dot{Q}^{-1}(0) \|_R = \sqrt{2}, \]

\[ \| \dot{h}_{1}^{2k} \|_R = 0, \quad \| \dot{h}_{1}^{2k+1} \|_R = 2^{2k+1} \frac{\sqrt{3}}{2} \frac{2^{2k+1}}{(2k+1)!} = 2^{2k+1} \| h_{2}^{2k+1} \|_R, \]

\[ \| \dot{h}_{2}^{2k} \|_R = 0, \quad \| \dot{h}_{2}^{2k+1} \|_R = 2^{2k+1} \frac{\sqrt{3}}{2} \frac{2^{2k+1}}{(2k+1)!} = 2^{2k+1} \| h_{2}^{2k+1} \|_R, \]

\[ \| \dot{f}_{2k+1} \|_R = 0, \quad \| \dot{f}_{2k} \|_R = 2^{2k} \frac{1}{4} \frac{2^{2k}}{(2k)!} = 2^{2k} \| f_{2k} \|_R, \]

\[ \| \dot{f}_{2k+1} \|_R = 0, \quad \| \dot{f}_{2k} \|_R \leq 2^{2k} \frac{\sqrt{3}}{2} \frac{2^{2k-1}}{(2k-1)!} = 2^{2k} \| f_{2k} \|_R. \]
Let \( \tilde{\rho}_{\text{start}} = \frac{2}{3} \tilde{\rho}_{\text{tol}} \). Then \( T_{\text{max}} \) and \( \hat{T}_{\text{max}} \) read (cf. Theorem 7.1.2)

\[
T_{\text{max}} = \frac{1}{2} R^2 (\tilde{\rho}_{\text{tol}}^2 - \tilde{\rho}_{\text{start}}^2) - 2 \sum_{k=3}^{r} \| I_k^i \|_R \tilde{\rho}_{\text{tol}}^k = \frac{21}{50} \rho_{\text{tol}} - 2 \sum_{k=3}^{r} \| I_k^i \|_R \tilde{\rho}_{\text{tol}}^k.
\]

\[
\hat{T}_{\text{max}} = \frac{1}{2} R^2 (\tilde{\rho}_{\text{tol}}^2 - \tilde{\rho}_{\text{start}}^2) - 2 \sum_{k=3}^{r} \| \hat{I}_k^i \|_R \hat{\tilde{\rho}}_{\text{tol}}^k \geq \frac{21}{50} \rho_{\text{tol}} - 2 \sum_{k=3}^{r} \| I_k^i \|_R 2^{-\frac{k}{2}} \rho_{\text{tol}}^k.
\]

\[
= \frac{9}{100} \rho_{\text{tol}}^2 - 2 \sum_{k=3}^{r} 2^{-\frac{k}{2}} \rho_{\text{tol}}^k.
\]

For small \( \rho_{\text{tol}} \) the second term in the numerator is small compared with the first:

\[
2 \sum_{k=3}^{r} \| I_k^i \|_R \tilde{\rho}_{\text{tol}}^k \ll \frac{21}{50} \rho_{\text{tol}} \quad \text{and} \quad 2 \sum_{k=3}^{r} 2^{-\frac{k}{2}} \rho_{\text{tol}}^k \ll \frac{9}{100} \rho_{\text{tol}}^2.
\]

In the denominator the first non-trivial term of the series is dominating:

\[
\sum_{k=r+1}^{\infty} \| \hat{I}_k^i \|_R \hat{\tilde{\rho}}_{\text{tol}}^k \approx \| \hat{I}_r^{i+1} \|_R \hat{\tilde{\rho}}_{\text{tol}}^{r+1} \quad \text{and} \quad \sum_{k=r+1}^{\infty} 2^{-\frac{k}{2}} \rho_{\text{tol}}^k \approx 2^{\frac{r}{2}} \| \hat{I}_r^{i+1} \|_R \hat{\tilde{\rho}}_{\text{tol}}^{r+1}.
\]

Thus we have

\[
\frac{\hat{T}_{\text{max}}}{T_{\text{max}}} \approx \frac{2^{\frac{r}{2}} \| \hat{I}_r^{i+1} \|_R \hat{\tilde{\rho}}_{\text{tol}}^{r+1}}{\| \hat{I}_r^{i+1} \|_R \rho_{\text{tol}}^{r+1}} \approx \frac{3}{14} 2^{-\frac{r}{2} + 1}.
\]

For example let \( r = 20 \). Then with the transformation \( \hat{I}(t) \) instead of \( Q(t) \equiv I \) we roughly loose a factor \( 10^4 \). In reality the loss is smaller. The difference results from the use of the estimate of \( \hat{I}^k \) in our consideration.

In the general case \( e_0 \neq 0 \) we expect a similar influence of the particular choice of the Floquet transformation on \( T_{\text{max}} \). But in contrast to the case \( e_0 = 0 \) no choice of the Floquet transformation is a priori distinguished.

In Table 8.26 we compare \( T_{\text{max}}/2\pi \) obtained for two different choices of the Floquet transformation. For the eccentricity we choose \( e_0 = 0.05 \) and for starting neighborhood \( \rho_{\text{start}} := 0.4 \rho_{\text{tol}} \). We see that the loss is roughly a factor \( 3 \cdot 10^2 \).
8. Stability of the Dumbbell Satellite

**Tab. 8.26:** $T_{\text{max}}/2\pi$ for different choices of the Floquet transformation.

<table>
<thead>
<tr>
<th>$\rho_{\text{tol}}$</th>
<th>$T_{\text{max}}/2\pi$</th>
<th>$T_{\text{max}}/2\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-9}$</td>
<td>$1 \cdot 10^{164}$</td>
<td>$3 \cdot 10^{166}$</td>
</tr>
<tr>
<td>$10^{-8}$</td>
<td>$1 \cdot 10^{145}$</td>
<td>$3 \cdot 10^{147}$</td>
</tr>
<tr>
<td>$10^{-7}$</td>
<td>$1 \cdot 10^{126}$</td>
<td>$3 \cdot 10^{128}$</td>
</tr>
<tr>
<td>$10^{-6}$</td>
<td>$1 \cdot 10^{107}$</td>
<td>$3 \cdot 10^{109}$</td>
</tr>
<tr>
<td>$10^{-5}$</td>
<td>$1 \cdot 10^{88}$</td>
<td>$3 \cdot 10^{90}$</td>
</tr>
<tr>
<td>$10^{-4}$</td>
<td>$1 \cdot 10^{69}$</td>
<td>$3 \cdot 10^{71}$</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>$1 \cdot 10^{50}$</td>
<td>$3 \cdot 10^{52}$</td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>$1 \cdot 10^{31}$</td>
<td>$3 \cdot 10^{33}$</td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>$4 \cdot 10^{11}$</td>
<td>$2 \cdot 10^{14}$</td>
</tr>
</tbody>
</table>

**8.7 Computational Notes**

We close this chapter with some remarks on the use of the computer.

- All computations are performed in *Mathematica*® 3.0 or 4.0 running on Macintosh® G3 computers.
- Almost all computations are performed with machine precision, i.e. with 16 decimal digits. Critical parts are performed with 26 decimal digits.
- For solving initial value problems the built-in function *NDSolve* is used. The solution are considered to be exact despite the small discretisation and round-off errors.
- The $2\pi$-periodic functions are represented by a finite sequence of their Fourier-coefficients:

\[
f(t) = \sum_{k_0=-\infty}^{\infty} f^{k_0} e^{ik_0t} \rightarrow \{f_{-n}, f_{-(n-1)}, \ldots, f_0, \ldots, f_{n-1}, f_n\}.
\]

It turned out the for Branch I $n = 16$ is suitable, while for Branch II we chose $n = 32$. To justify these choices we performed some computations for various $n$. Tables 8.27 and 8.28 gives a survey for the times $T_{\text{max}}/2\pi$ obtained for $n = 4$, $n = 16$ and $n = 128$. The truncation order is $r = 15$.

- The monomials $f(t)x^py^q\Delta^s$ are represented by the sequence of Fourier-coefficients of $f$ and the exponents $p$, $q$ and $s$:

\[
f(t)x^py^q\Delta^s \rightarrow \text{m}[\text{flist}, p, q, s].
\]

*Mathematica*®'s built-in functions are extended to deal with these expressions.
### Tab. 8.27: \( T_{\max}/2\pi \) for different choices of \( n \) for the \( 2\pi \)-periodic solution for \( \epsilon_0 = 0.05 \) corresponding to Branch I.

<table>
<thead>
<tr>
<th>( n = 4 )</th>
<th>( n = 16 )</th>
<th>( n = 128 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_{\text{tol}} )</td>
<td>( T_{\max}/2\pi )</td>
<td>( T_{\max}/2\pi )</td>
</tr>
<tr>
<td>( 10^{-9} )</td>
<td>1.411908590478717 ( \times 10^{-26} )</td>
<td>1.500502290774992 ( \times 10^{-24} )</td>
</tr>
<tr>
<td>( 10^{-8} )</td>
<td>1.41190854366394 ( \times 10^{-12} )</td>
<td>1.500502250981572 ( \times 10^{-10} )</td>
</tr>
<tr>
<td>( 10^{-7} )</td>
<td>1.411908072243208 ( \times 10^{-9} )</td>
<td>1.500501853047391 ( \times 10^{-6} )</td>
</tr>
<tr>
<td>( 10^{-6} )</td>
<td>1.41190361013328 ( \times 10^{-4} )</td>
<td>1.500497873703354 ( \times 10^{-2} )</td>
</tr>
<tr>
<td>( 10^{-5} )</td>
<td>1.411856248915271 ( \times 10^{-1} )</td>
<td>1.500458080040467 ( \times 10^{1} )</td>
</tr>
<tr>
<td>( 10^{-4} )</td>
<td>1.411385148003706 ( \times 10^{4} )</td>
<td>1.50060118531987 ( \times 10^{4} )</td>
</tr>
<tr>
<td>( 10^{-3} )</td>
<td>1.406676142083291 ( \times 10^{10} )</td>
<td>1.496078304983568 ( \times 10^{10} )</td>
</tr>
<tr>
<td>( 10^{-2} )</td>
<td>1.359782740696077 ( \times 10^{14} )</td>
<td>1.456037840305800 ( \times 10^{14} )</td>
</tr>
<tr>
<td>( 10^{-1} )</td>
<td>9.072603966200360 ( \times 10^{13} )</td>
<td>1.033515568675945 ( \times 10^{12} )</td>
</tr>
</tbody>
</table>

### Tab. 8.28: \( T_{\max}/2\pi \) for different choices of \( n \) for the \( 2\pi \)-periodic solution for \( \epsilon_0 = 0.01 \) corresponding to Branch III.

<table>
<thead>
<tr>
<th>( n = 4 )</th>
<th>( n = 8 )</th>
<th>( n = 16 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_{\text{tol}} )</td>
<td>( T_{\max}/2\pi )</td>
<td>( T_{\max}/2\pi )</td>
</tr>
<tr>
<td>( 10^{-9} )</td>
<td>3.93662 ( \times 10^{10} )</td>
<td>6.24106 ( \times 10^{99} )</td>
</tr>
<tr>
<td>( 10^{-8} )</td>
<td>3.93661 ( \times 10^{10} )</td>
<td>6.24102 ( \times 10^{85} )</td>
</tr>
<tr>
<td>( 10^{-7} )</td>
<td>3.93653 ( \times 10^{76} )</td>
<td>6.24057 ( \times 10^{71} )</td>
</tr>
<tr>
<td>( 10^{-6} )</td>
<td>3.93572 ( \times 10^{62} )</td>
<td>6.23608 ( \times 10^{57} )</td>
</tr>
<tr>
<td>( 10^{-5} )</td>
<td>3.92763 ( \times 10^{48} )</td>
<td>6.19113 ( \times 10^{43} )</td>
</tr>
<tr>
<td>( 10^{-4} )</td>
<td>3.84651 ( \times 10^{34} )</td>
<td>5.74068 ( \times 10^{29} )</td>
</tr>
<tr>
<td>( 10^{-3} )</td>
<td>3.01356 ( \times 10^{20} )</td>
<td>1.13133 ( \times 10^{15} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( n = 32 )</th>
<th>( n = 64 )</th>
<th>( n = 128 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho_{\text{tol}} )</td>
<td>( T_{\max}/2\pi )</td>
<td>( T_{\max}/2\pi )</td>
</tr>
<tr>
<td>( 10^{-9} )</td>
<td>2.04775 ( \times 10^{102} )</td>
<td>2.25566 ( \times 10^{102} )</td>
</tr>
<tr>
<td>( 10^{-8} )</td>
<td>2.04773 ( \times 10^{88} )</td>
<td>2.25564 ( \times 10^{88} )</td>
</tr>
<tr>
<td>( 10^{-7} )</td>
<td>2.04758 ( \times 10^{74} )</td>
<td>2.25547 ( \times 10^{74} )</td>
</tr>
<tr>
<td>( 10^{-6} )</td>
<td>2.04604 ( \times 10^{60} )</td>
<td>2.25377 ( \times 10^{60} )</td>
</tr>
<tr>
<td>( 10^{-5} )</td>
<td>2.03060 ( \times 10^{46} )</td>
<td>2.23676 ( \times 10^{46} )</td>
</tr>
<tr>
<td>( 10^{-4} )</td>
<td>1.87587 ( \times 10^{32} )</td>
<td>2.06625 ( \times 10^{32} )</td>
</tr>
<tr>
<td>( 10^{-3} )</td>
<td>2.94529 ( \times 10^{17} )</td>
<td>3.24311 ( \times 10^{17} )</td>
</tr>
</tbody>
</table>
8. Stability of the Dumbbell Satellite
9

Chaotic Behavior for Small Eccentricities

We consider the motion of a satellite around its center of mass. It is described by

\[ \dot{z} = F(t, z, e), \quad (2.9) \]

where the function \( F \) is defined by

\[ F(t, z, e) := F(t, x, y, e) = \left( \frac{3}{2} \frac{\sin(2x)}{1-e \cos t} + y \frac{e \sin t}{1-e \cos t} + \frac{2\sqrt{1-e^2} \sin t}{(1-e \cos t)^2} \right). \]

The unperturbed equation admits equilibrium solutions \((x, y) = (k \cdot \frac{\pi}{2}, 0), k \in \mathbb{Z}\). For odd values of \( k \) the equilibrium solutions are stable while for \( k \) even the equilibrium solutions are unstable.

In the last chapter we saw that for small eccentricities solutions with initial values \((x_0, y_0)\) close to \((k \pi, 0), k \in \mathbb{Z}\), remain close to \((k \pi, 0)\) for large times.

Now we study the behavior of solutions with initial values \((x_0, y_0)\) close to \((\frac{\pi}{2} + k \pi, 0), k \in \mathbb{Z}\). We show that for small eccentricities \( e \) the behavior of (2.9) is chaotic in the sense of Smale (cf. [16]) using the so-called method of Melnikov.

The chapter is organized as follows:

**Section 9.1:** In the first section we sketch the method of Melnikov and state the most important definitions and theorems. Since the method is well-known we omit the proofs. A more detailed treatement as well as further applications may be found in [17], [15] or [9].

**Section 9.2:** In the second section we prove that (2.9) fulfills the assumptions required by the method of Melnikov. This guarantees chaotic behavior for small eccentricities.
9.1 The Method of Melnikov

In this section we sketch the theory of Melnikov. We start with the basic assumptions.

Basic Assumptions
We consider a plane periodic systems of differential equations of the form

\[ \dot{z} = F(t, z, e) := F^0(z) + eF^1(t, z, e), \quad z \in \mathbb{R}^2, \quad (9.1c) \]

where \( F^0 \) and \( F^1 \) are sufficiently regular, \( F^1 \) is \( 2\pi \)-periodic with respect to \( t \), and \( e \) is a small parameter.

For the unperturbed system (9.10) we make the following additional hypotheses:

(i) \( F^0 \) is Hamiltonian:
There is a scalar function \( H \) such that

\[ F^0(z) = \begin{pmatrix} \frac{\partial H(z)}{\partial z_2} \\ \frac{\partial H(z)}{\partial z_1} \end{pmatrix}. \]

(ii) \( z_0^* \) is a hyperbolic equilibrium of (9.10):
\( F^0(z_0^*) = 0 \) and the eigenvalues \( \lambda_{1,2} \) of \( DF^0(z_0^*) \) satisfy \( \lambda_1 < 0 < \lambda_2 = -\lambda_1 \).

(iii) \( z_0^* \) admits a separatrix solution:
There is a solution \( z_0^s \) of (9.10) such that

\[ \lim_{t \to \pm \infty} z_0^s(t) = z_0^* \]

as shown in Figure 9.1.

![Fig. 9.1: Phase portrait of (9.10).](image)

Note that Hypothesis (i) is not necessary. It simplifies the proofs of the theorems stated below, however, and it is frequently satisfied.

Before we present the basic idea of the method of Melnikov we recall some properties of the Poincaré map.
The Poincaré Map

Let $P$ denote the so-called period or Poincaré map of system (9.1,e):

$$P : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \ z \mapsto P(z) := \phi(2\pi, z, e),$$

where $\phi$ denotes the flow associated with (9.1,e). Because the vector field $F$ is $2\pi$-periodic in $t$ we have

$$\phi(k \cdot 2\pi, z, e) = P^k(z).$$

This relation implies that the global behavior of (9.1,e) follows from the corresponding properties of the discrete dynamical system generated by the Poincaré map $P$.

**Basic Properties of the Poincaré Map**

(i) For $e = 0$ we have a hyperbolic fixed point $z_0^*$:

$$P(z_0^*) = z_0^*$$

and the eigenvalues $\Lambda_{\pm}$ of the linearized map $DP(z_0^*)$ satisfy $0 < \Lambda_0^+ < 1 < \Lambda_0^-$. 

(ii) For $e$ small we have a hyperbolic fixed point $z_e^*$ close to $z_0^*$ (cf. Figure 9.2).

![Fig. 9.2: A hyperbolic fixed point for small $e$.](image)

(iii) For $e = 0$ the homoclinic solution curve is a subset of the stable and unstable manifolds $W_0^\pm$ of $z_0^*$ defined by

$$W_0^\pm := \left\{ z \in \mathbb{R}^2 \mid \lim_{k \to \pm\infty} P^k(z) = z_0^* \right\}.$$ 

Moreover they are generated by the separatrix solution $z_0^*$ of (9.1,e).

(iv) For $e > 0$ the stable and unstable manifolds $W_e^\pm$ of $z_e^*$ coincide no longer in general (cf. Figure 9.3).

![Fig. 9.3: A transversal homoclinic orbit.](image)
The existence of a single transversal intersection $z_0$ implies that there are infinitely many intersections $z_n := P^n(z_0)$. The orbit $(z_n)_{n \in \mathbb{Z}}$ is called a transversal homoclinic orbit of $z^*$ with respect to $P$.

A Notion of Chaos

We start this section with the definition of chaos in the sense of Smale. To this end we first introduce the so-called Bernoulli shift system.

Definition 9.1.1
Let $\Sigma$ be the set of all bi-infinite sequences $s$ of zeros and ones

$$\Sigma := \{ s = (\ldots, s_{-1}; s_0, s_1, \ldots) \mid s_i \in \{0,1\} \}$$

endowed with the metric

$$d(s,t) := \max_{n \in \mathbb{Z}} \left\{ 2^{-|n|} \right\}_{s_n \neq t_n}$$

and let the map $\sigma$ be defined by

$$\sigma : \Sigma \rightarrow \Sigma, \ s = (\ldots, s_{-1}; s_0, s_1, \ldots) \mapsto \sigma(s) := (\ldots, s_0; s_1, s_2, \ldots).$$

The discrete dynamical system $(\Sigma, \sigma)$ is called Bernoulli shift system and $\sigma$ is called the shift map.

The behavior of the shift system is unpredictable. It is therefore used as a prototype of chaotic behavior. This idea leads to the following definition.

Definition 9.1.2
A discrete dynamical system $(X,P)$ is called chaotic in the sense of Smale, if it admits the Bernoulli shift system as a subsystem, i.e. there exists a homeomorphism $\tau$ of $\Sigma$ onto $\tau(\Sigma)$ and a positive integer $N$ such that the following diagram commutes:

$$\begin{array}{ccc}
\Sigma & \xrightarrow{\sigma} & \Sigma \\
\tau \downarrow & & \downarrow \tau \\
\tau(\Sigma) \subset X & \xrightarrow{P^N} & \tau(\Sigma) \subset X
\end{array}$$

For a survey on several definitions of chaos cf. [16].

Now the idea is to construct for every $s = (\ldots, s_{-1}; s_0, s_1, \ldots) \in \Sigma$ an orbit of the Poincaré map with the following properties:
• if \( s_i = 0 \), then the orbit stays near the hyperbolic fixed point for a certain number \( m \) of iteration of the Poincaré map.

• if \( s_i = 1 \), then the orbit follows the transversal homoclinic orbit and returns to a neighborhood of the fixed point, the amount of time taken for this "round trip" being \( m \) iterations of the Poincaré map.

Before we can realize this idea, we need some more preparations.

### Hyperbolic Sets

The idea of a hyperbolic fixed point of a planar map \( P \) may be generalized to a compact subset of \( \mathbb{R}^2 \).

**Definition 9.1.3**

A compact set \( \Lambda \subset \mathbb{R}^2 \) which is invariant with respect to some map \( P \) is called hyperbolic for \( P \) if there are two continuous vector fields \( h^+ \) and \( h^- \) on \( \Lambda \) with the following properties:

(i) \( h^+(z) \) and \( h^-(z) \) are linearly independent for all \( z \in \Lambda \).

(ii) \( h^+ \) and \( h^- \) are invariant with respect to \( DP \), i.e. there are maps \( \lambda^\pm : \Lambda \to \mathbb{R} \) such that

\[
    h^\pm(\lambda^\pm(z)) = \frac{1}{\lambda^\pm(z)}DP(z)h^\pm(z)
\]

for all \( z \in \Lambda \).

(iii) The map \( P \) is contracting in the direction of \( h^+ \) and expanding in the direction of \( h^- \), i.e. there are constants \( c_1 \) and \( c_2 \) such that

\[
    \frac{1}{c_1} \leq |\lambda^+(z)| \leq c_2 < 1, \quad 1 < \frac{1}{c_2} \leq |\lambda^-(z)| \leq c_1.
\]

Obviously the set consisting of a single hyperbolic fixed point is a hyperbolic set. The next theorem shows that under certain assumptions this set may be enlarged by the transversal homoclinic orbit.

**Theorem 9.1.1**

Let the following hypotheses hold true:

**H**: \( P \) admits a hyperbolic fixed point \( z^* \). Moreover, there are two curves \( \Gamma^\pm := \{ \gamma^\pm(s) | s \in [a^\pm, b^\pm] \} \) in \( \mathbb{R}^2 \) with the following properties:

**H1**: The maps \( \gamma^\pm \) are injective, sufficiently regular and their derivatives do not vanish.

**H2**: There are \( \sigma^\pm \in [a^\pm, b^\pm] \) such that \( \gamma^\pm(\sigma^\pm) = z^* \) holds.
**H3:** $\Gamma^-$ is invariant with respect to $P^{-1}$, $\Gamma^+$ is invariant with respect to $P$. 
If $z \in \Gamma^-$ then $\lim_{n \to -\infty} P^n(z) = z^*$. 
If $z \in \Gamma^+$ then $\lim_{n \to \infty} P^n(z) = z^*$.

**H4:** There are $s_\pm \neq \sigma_\pm$ such that $\gamma_-(s_-) = \gamma_+(s_+)$ and $\gamma'_-(s_-)$ and $\gamma'_+(s_+)$ are linearly independent.

Consider the transversal homoclinic orbit $\{z_n\}_{n \in \mathbb{Z}}$ generated by $z_0 := \gamma_-(s_-) = \gamma_+(s_+)$. Then

$$\Lambda := \{z^*\} \cup \{z_n | n \in \mathbb{Z}\}$$

is a hyperbolic set for $P$.

In the next section we show why hyperbolic sets are so important.

**The Shadowing Lemma**

Let $\ell^\infty(\mathbb{Z}, \mathbb{R}^2)$ denote the space of bounded bi-infinite sequences $p := (p_n)_{n \in \mathbb{Z}}$ of points $p_n$ in $\mathbb{R}^2$, endowed with the supremum norm

$$\|p\|_\infty := \sup_{n \in \mathbb{Z}} |p_n|_2,$$

where $|.|_2$ is the Euclidian norm in $\mathbb{R}^2$.

The following definition introduces the notion of so-called pseudo orbits.

**Definition 9.1.4**

A bi-infinite sequence $q := (q_n)_{n \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z}, \mathbb{R}^2)$ is called a $\delta$-pseudo orbit of a map $P$ if

$$|q_{n+1} - P(q_n)| \leq \delta$$

holds for all $n \in \mathbb{Z}$.

![Fig. 9.4: A section of a $\delta$-pseudo orbit of $P$.](image)

If there is a real orbit near a pseudo orbit we speak of a shadowing orbit.
Definition 9.1.5
Let \( q := (q_n)_{n \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \) be a \( \delta \)-pseudo orbit for \( P \). A bi-infinite sequence \( p := (p_n)_{n \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \) is called \( \rho \)-shadowing orbit of \( q \) with respect to \( P \) if \( p \) is an orbit of \( P \) and
\[
\| p - q \|_\infty \leq \rho.
\]
The following theorem is referred to as the shadowing lemma.

Theorem 9.1.2 (Shadowing Lemma)
Let \( P \) admit a hyperbolic set \( \Lambda \). Then there is a \( \rho_0 > 0 \) such that for all \( \rho \) with \( 0 < \rho < \rho_0 \) there exists a \( \delta := \delta(\rho) \) such that for every \( \delta \)-pseudo orbit \( q := (q_n)_{n \in \mathbb{Z}} \subset \Lambda \) there is a unique \( \rho \)-shadowing orbit \( p := (p_n)_{n \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \).

The Theorem of Smale
Assume that the Poincaré map \( P \) of (2.9\(_e\)) admits a fixed point \( z^* \) and a transversal homoclinic orbit \( (z_n)_{n \in \mathbb{Z}} \) associated with \( z^* \). Then the theorem of Smale establishes a connection between the Bernoulli shift system and \((\mathbb{R}^2, P)\).

Theorem 9.1.3
Consider the system of differential equations (2.9\(_e\)) for some fixed eccentricity \( e \) and let \( P \) denote the Poincaré map.
If the stable and unstable manifolds \( W^\pm_e \) of the fixed point \( z^* \) of \( P \) admits a transversal intersection \( z_0 \), then
\[
\Lambda := \{z^*\} \cup \{z_n := P^n z_0 | n \in \mathbb{Z}\}
\]
is a hyperbolic set.
Let
\[
u := (z^*, z_{-N}, z_{-N+1}, \ldots, z_{-1}, z_0, z_1, \ldots, z_N, z^*)
\]
and
\[
v := (z^*, z^*, \ldots, z^*, z^*, \ldots, z^*, z^*)
\]
where the positive integer \( N \) is chosen such that \(|z^* - z_{-N}| \leq \delta\) and \(|z^* - z_{N+1}| \leq \delta\) for a positive constant \( \delta \leq \frac{1}{3}|z_0 - z^*| \).
Let

* \( \tau_{\text{pseudo}}: \)

\[ \Sigma \rightarrow S := \tau_{\text{pseudo}}(\sigma) \subset \ell^\infty(\mathbb{Z}, \mathbb{R}^2), \quad s \mapsto \tau_{\text{pseudo}}(s) := (\ldots, b_{-1}, b_0, b_1, \ldots), \]

where

\[ b_n := \begin{cases} u & \text{if } s_n = 0 \\ v & \text{if } s_n = 1, \end{cases} \]

* \( \tau_{\text{shadow}}: \)

\[ S \rightarrow S^* := \tau_{\text{shadow}}(S) \subset \ell^\infty(\mathbb{Z}, \mathbb{R}^2), \quad q \mapsto \tau_{\text{shadow}}(q) := p, \]

where \( p \) is the (unique) shadowing orbit associated with \( q \), according to the Shadowing Lemma,

* \( \pi: \)

\[ S^* \rightarrow \mathbb{R}^2, \quad p \mapsto \pi(p) := p_{N+1}. \]

Then the map

\[ \tau: \Sigma \rightarrow \mathbb{R}^2, \quad \tau := \pi \circ \tau_{\text{shadow}} \circ \tau_{\text{pseudo}} \]

has the following properties:

(i) \( \tau \) is a homeomorphism between \( \Sigma \) and \( \tau(\Sigma) \subset \mathbb{R}^2 \) such that the following diagram commutes:

\[
\begin{array}{ccc}
\Sigma & \xrightarrow{\sigma} & \Sigma \\
\downarrow{\tau} & & \downarrow{\tau} \\
\tau(\Sigma) & \xrightarrow{P^{2N+3}} & \tau(\Sigma) \subset \mathbb{R}^2
\end{array}
\]

(ii) There is \( \rho > 0 \) such that the following holds:

- If \( s_0 = 0 \) then \( |\tau(s) - z^*| \leq \rho \), whereas if \( s_0 = 1 \) then \( |\tau(s) - z^*| \geq 2\rho \).

- If \( s_i = 0 \) then \( |z_n - z^*| \leq \rho \) for \( n \in \{i \cdot (2N + 3), \ldots, (i + 1) \cdot (2N + 3)\} \).

(iii) \((\mathbb{R}^2, P^{N+1})\) admits the Bernoulli shift system \((\Sigma, \sigma)\) as a subsystem and is therefore chaotic in the sense of Smale.

It remains to provide conditions under which system (9.1e) admits a hyperbolic set.

The Melnikov Function

The Melnikov function \( d \) is a measure for the distance between the stable and the unstable manifold \( W^\pm_e \) of the hyperbolic fixed point \( z^*_e \) of the Poincaré map \( P \).
9.2. Chaotic Behavior for Small Eccentricities

**Theorem 9.1.4**

Consider system (2.9e) and let the following assumptions hold:

(i) The unperturbed system (2.90) is Hamiltonian.

(ii) $z^*_0$ is a hyperbolic equilibrium of (2.90).

(iii) $z^*_0$ admits a separatrix solution.

Assume that the so-called Melnikov function

$$d : \mathbb{R} \to \mathbb{R}, \sigma \mapsto d(\sigma) := \int_{-\infty}^{\infty} F^0(x_s(t)) \wedge F^1(t - \sigma, x_s(t), 0) dt$$

admits a simple zero. Then the following holds:

(i) For sufficiently small $e > 0$ the hypotheses $\textbf{H}$ hold and Theorem 9.1.1 implies that the transversal homoclinic orbit $A := \{z^*_0\} \cup \{z_n | n \in \mathbb{Z}\}$ is hyperbolic.

(ii) Moreover the theorem of Smale implies that the behavior of the system is chaotic.

9.2 Chaotic Behavior for Small Eccentricities

Now we are in a position to establish the chaotic behavior of the dumbbell satellite for small eccentricities $e$.

To this end equation (2.9) is rewritten as

$$\begin{align*}
\dot{x} &= y \\
\dot{y} &= -\frac{3}{2} \sin(2x) + e \left( \frac{3}{2} \frac{\cos t}{1 - e \cos t} \sin(2x) + y \frac{\sin t}{1 - e \cos t} + \frac{2\sqrt{1 - e^2} \sin t}{(1 - e \cos t)^2} \right)
\end{align*}$$

Thus the functions $F^0$ and $F^1$ are given by

$$F^0(z) := \begin{pmatrix} y \\ -\frac{3}{2} \sin(2x) \end{pmatrix}$$

and

$$F^1(t, z, e) := \begin{pmatrix} 0 \\ -\frac{3}{2} \frac{\cos t}{1 - e \cos t} \sin(2x) + y \frac{\sin t}{1 - e \cos t} + \frac{2\sqrt{1 - e^2} \sin t}{(1 - e \cos t)^2} \end{pmatrix}.$$

The flow of (9.3) is $\pi$-periodic with respect to $x$. Therefore the phase space is rather a cylinder than a plane (cf. Figure 9.5).
The Basic Hypotheses

Now we show that (9.3) fulfills the basic hypotheses \((i) - (iii)\) of Theorem 9.1.4.

- \(F^0\) is Hamiltonian:
  Let \(H(x, y) := \frac{1}{2}y^2 - \frac{3}{4}\cos(2x)\). Then we have
  \[
  \begin{align*}
  \dot{x} &= \frac{\partial H}{\partial y}, \\
  \dot{y} &= -\frac{\partial H}{\partial x}.
  \end{align*}
  \]

- \((\frac{\pi}{2}, 0)\) is a hyperbolic equilibrium solution of (2.9b):
  Obviously \((\frac{\pi}{2}, 0)\) is an equilibrium solution of (2.9). It remains to show that \((\frac{\pi}{2}, 0)\) is hyperbolic. The Jacobian of \(F^0\) reads
  \[
  DF^0(x, y) = \begin{pmatrix} 0 & 1 \\ -3\cos(2x) & 0 \end{pmatrix},
  \]
  which implies that
  \[
  DF^0\left(\frac{\pi}{2}, 0\right) = \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix}.
  \]
  For the eigenvalues we find \(\lambda_1 := -\sqrt{3} < 0 < \sqrt{3} =: \lambda_2\). Thus \((\frac{\pi}{2}, 0)\) is hyperbolic, indeed.

- The unperturbed system admits a separatrix solution \((x_0^a, y_0^a)\):
  Let \(y'\) denote the derivative of \(y\) with respect to \(x\). (9.3c) combined with \(\dot{y} = y' \cdot \dot{x}\) gives
  \[
  -\frac{3}{2}\sin(2x) = yy'.
  \]
  We integrate this equation and obtain
  \[
  y^2 = \frac{3}{2}\cos(2x) + c.
  \]
9.2. Chaotic Behavior for Small Eccentricities

From \( y\left(\frac{x}{2}\right) = 0 \) we conclude that \( c = \frac{3}{2} \). Therefore we have

\[
y^2 = \frac{3}{2} (\cos(2x) + 1) = \frac{3}{2} \left( (2\cos^2 x + 1) - 1 \right) = 3\cos^2 x.
\]

Together with \( y = \dot{x} \) we obtain

\[
\dot{x} = \sqrt{3}\cos x.
\]

We solve this equation first for \( t \):

\[
t = \frac{1}{\sqrt{3}} \int_0^x \frac{1}{\cos \xi} d\xi = \frac{1}{\sqrt{3}} \log \left| \tan \left( \frac{x}{2} + \frac{\pi}{4} \right) \right|.
\]

It follows that

\[
e^{\sqrt{3}t} = \tan \left( \frac{x}{2} + \frac{\pi}{4} \right) = \frac{1 + \tan \left( \frac{x}{2} \right)}{1 - \tan \left( \frac{x}{2} \right)}
\]

and furthermore

\[
\tan \left( \frac{x}{2} \right) = \frac{e^{\sqrt{3}t} - 1}{e^{\sqrt{3}t} + 1}.
\]

From

\[
\tan(x) = \frac{2\tan \left( \frac{x}{2} \right)}{1 - \tan^2 \left( \frac{x}{2} \right)}
\]

we conclude that

\[
\tan(x) = \frac{2e^{\sqrt{3}t} - 1}{e^{\sqrt{3}t} + 1} = \frac{2 \left( e^{\sqrt{3}t} - 1 \right) \left( e^{\sqrt{3}t} + 1 \right)}{\left( e^{\sqrt{3}t} + 1 \right)^2 - \left( e^{\sqrt{3}t} - 1 \right)^2} = \frac{2 \left( e^{2\sqrt{3}t} - 1 \right)}{4e^{\sqrt{3}t}} = \frac{1}{2} \left( e^{\sqrt{3}t} - e^{-\sqrt{3}t} \right) = \sinh(\sqrt{3}t).
\]

Thus we have

\[
x_0^* (t) = \arctan(\sinh(\sqrt{3}t))
\]

and

\[
y_0^* (t) = \dot{x}_0^* (t) = \frac{1}{1 + \sinh^2 (\sqrt{3}t)} \cdot \cosh (\sqrt{3}t) \cdot \sqrt{3} = \frac{\sqrt{3}}{\cosh (\sqrt{3}t)}.
\]

Obviously \((x_0^*, y_0^*)\) has the desired properties.
Computing the Melnikov Function

Now we compute the Melnikov function. We start with the wedge product

\[ F^0(t, z_\sigma, z_\sigma^*; 0) = -\frac{3}{2} \cos(t - \sigma) \sin(2x_0^*(t)) y_0^*(t) + \]

\[ + \sin(t - \sigma) (y_0^*)^2 + 2 \sin(t - \sigma) y_0^*(t). \]

With the identities

\[ \sin^2 x = \frac{\tan^2 x}{1 + \tan^2 x}, \quad \cos^2 x = \frac{1}{1 + \tan^2 x} \]

we obtain

\[ \sin(2x_0^*(t)) = 2 \sin(x_0^*(t)) \cos(x_0^*(t)) = 2 \frac{\sinh(\sqrt{3}t)}{1 + \sinh^2(\sqrt{3}t)} = 2 \frac{\sinh(\sqrt{3}t)}{cosh^2(\sqrt{3}t)}. \]

Inserting this into the wedge product we find

\[ F^0(t, z_\sigma, z_\sigma^*; 0) = \]

\[ = -\frac{3}{2} \cos(t - \sigma) \cdot 2 \frac{\sinh(\sqrt{3}t)}{cosh^2(\sqrt{3}t)} \cdot \frac{\sqrt{3}}{cosh(\sqrt{3}t)} + \]

\[ + \sin(t - \sigma) \cdot \frac{3}{cosh^2(\sqrt{3}t)} + 2 \sin(t - \sigma) \cdot \frac{\sqrt{3}}{cosh(\sqrt{3}t)} = \]

\[ = -3 \sqrt{3} \cos \sigma \cdot \frac{\cos t \sinh(\sqrt{3}t)}{cosh^3(\sqrt{3}t)} - 3 \sqrt{3} \sin \sigma \cdot \frac{\sin t \sinh(\sqrt{3}t)}{cosh^3(\sqrt{3}t)} + \]

\[ + 3 \cos \sigma \cdot \frac{\sin t}{cosh^2(\sqrt{3}t)} - 3 \sin \sigma \cdot \frac{\cos t}{cosh^2(\sqrt{3}t)} + \]

\[ + 2 \sqrt{3} \cos \sigma \cdot \frac{\sin t}{cosh(\sqrt{3}t)} - 2 \sqrt{3} \sin \sigma \cdot \frac{\cos t}{cosh(\sqrt{3}t)}. \]

Thus we have to compute the integrals

\[ I_1 := \int_{-\infty}^{\infty} \frac{\cos t \sinh(\sqrt{3}t)}{cosh^3(\sqrt{3}t)} dt, \]

\[ I_2 := \int_{-\infty}^{\infty} \frac{\sin t \sinh(\sqrt{3}t)}{cosh^3(\sqrt{3}t)} dt, \]

\[ I_3 := \int_{-\infty}^{\infty} \frac{\sin t}{cosh^2(\sqrt{3}t)} dt, \]

\[ I_4 := \int_{-\infty}^{\infty} \frac{\cos t}{cosh^2(\sqrt{3}t)} dt, \]
9.2. Chaotic Behavior for Small Eccentricities

\[ I_5 := \int_{-\infty}^{\infty} \frac{\sin t}{\cosh(\sqrt{3}t)} \, dt, \]

\[ I_6 := \int_{-\infty}^{\infty} \frac{\cos t}{\cosh(\sqrt{3}t)} \, dt. \]

\[ I_1 : \text{Since the integrand is odd in } t \text{ the integral vanishes:} \]

\[ I_1 = 0. \]

\[ I_2 : \text{Integration by parts leads to} \]

\[ I_2 = \int_{-\infty}^{\infty} \frac{\sin t \sinh(\sqrt{3}t)}{\cosh^3(\sqrt{3}t)} \, dt = -\left. \frac{\sin t}{2\sqrt{3} \cosh^2(\sqrt{3}t)} \right|_{-\infty}^{\infty} + \frac{1}{2\sqrt{3}} \int_{-\infty}^{\infty} \frac{\cos t}{\cosh^2(\sqrt{3}t)} \, dt = \]

\[ = \frac{1}{2\sqrt{3}} I_4. \]

With the integral \( I_4 \) computed below we obtain

\[ I_2 = \frac{\pi}{6\sqrt{3} \sinh\left(\frac{\pi}{2\sqrt{3}}\right)}. \]

\[ I_3 : \text{Since the integrand is odd in } t \text{ the integral vanishes:} \]

\[ I_3 = 0. \]

\[ I_4 : \text{We compute the integral by the calculus of residues. Let} \]

\[ g(z) := \frac{\cos z}{\cosh^2(\sqrt{3}z)}. \]

The poles of \( g \) are

\[ t = \frac{i\pi}{2k + 1}. \]

We integrate over the path \( \Gamma := \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4 \) indicated in the Figure 9.6. We obtain

\[ \int_{\Gamma} g(z) \, dz = 2\pi i \text{ res}_{z = i\frac{\pi}{2\sqrt{3}}} g(z). \]

It is well-known that the residue is equal to the coefficient \( g_{-1} \) of the Laurent series of \( g \) at \( z = i\frac{\pi}{2\sqrt{3}} \). Thus we have to compute \( g_{-1} \).

The Taylor expansions of \( \cos z \) and \( \cosh(\sqrt{3}z) \) at \( z = i\frac{\pi}{2\sqrt{3}} \) read

\[ \cos z = \sum_{k=0}^{\infty} \cos^{(k)}\left(\frac{\pi}{2\sqrt{3}}\right) \left(\frac{z - i\frac{\pi}{2\sqrt{3}}}{2\sqrt{3}}\right)^k = \]

\[ = \cos\left(i\frac{\pi}{2\sqrt{3}}\right) - \sin\left(i\frac{\pi}{2\sqrt{3}}\right) \left(\frac{z - i\frac{\pi}{2\sqrt{3}}}{2\sqrt{3}}\right) + O\left(\left(\frac{z - i\frac{\pi}{2\sqrt{3}}}{2\sqrt{3}}\right)^2\right) \]
Fig. 9.6: Integration on the rectangle $\Gamma$. 

and

$$
cosh(\sqrt{3}z) = \sum_{k=0}^{\infty} \frac{\cosh^{(k)}(\sqrt{3} \cdot i \frac{\pi}{2\sqrt{3}})}{k!} \left( z - i \frac{\pi}{2\sqrt{3}} \right)^k = 
$$

$$
= i\sqrt{3} \left( z - i \frac{\pi}{2\sqrt{3}} \right) + O \left( \left( z - i \frac{\pi}{2\sqrt{3}} \right)^2 \right).
$$

This leads to

$$
g_{-1} = \frac{-\sin \left( i \frac{\pi}{2\sqrt{3}} \right)}{-3} = \frac{i}{3} \sinh \left( \frac{\pi}{2\sqrt{3}} \right).
$$

Thus we have

$$
\int_{\Gamma_{-1}} g(z) dz = -\frac{2\pi}{3} \sinh \left( \frac{\pi}{2\sqrt{3}} \right).
$$

Let us now consider the integrals over the individual paths $\Gamma_k$. Obviously the integrals over $\Gamma_2$ and $\Gamma_4$ vanish for $R \to \infty$. For the integral over $\Gamma_3$ we obtain

$$
\int_{\Gamma_3} \frac{\cos z}{\cosh^2(\sqrt{3}z)} dz = \int_{-R}^{R} \frac{\cos \left( -t + i \frac{\pi}{\sqrt{3}} \right)}{\cosh^2 \left( \sqrt{3} \left( -t + i \frac{\pi}{\sqrt{3}} \right) \right)} (-dt) =
$$

$$
= \int_{-R}^{R} \frac{\cos(-t) \cos \left( i \frac{\pi}{\sqrt{3}} \right) - \sin(-t) \sin \left( i \frac{\pi}{\sqrt{3}} \right)}{\left( \cosh(-\sqrt{3}t) \cosh(i\pi) + \sin(-\sqrt{3}t) \sin(i\pi) \right)^2} (-dt) =
$$

$$
= -\cos \left( i \frac{\pi}{\sqrt{3}} \right) \int_{-R}^{R} \frac{\cos t}{\cosh^2(\sqrt{3}t)} dt - i \sin \left( i \frac{\pi}{\sqrt{3}} \right) \int_{-R}^{R} \frac{\sin t}{\cosh^2(\sqrt{3}t)} dt =
$$

$$
= -\cosh \left( \frac{\pi}{\sqrt{3}} \right) \int_{-R}^{R} \frac{\cos t}{\cosh^2(\sqrt{3}t)} dt.
$$
Combining these results we obtain

\[-\frac{2\pi}{3} \sinh \left( \frac{\pi}{2\sqrt{3}} \right) = \int_\Gamma g(z)dz = \left( 1 - \cosh \left( \frac{\pi}{\sqrt{3}} \right) \right) \int_{-R}^R \frac{\cos t}{\cosh(\sqrt{3}t)} dt.\]

Thus we have for \( R \to \infty \)

\[I_4 = \int_{-\infty}^\infty \frac{\cos t}{\cosh^2(\sqrt{3}t)} dt = -\frac{2\pi}{3} \frac{\sinh \left( \frac{\pi}{2\sqrt{3}} \right)}{1 - \cosh \left( \frac{\pi}{\sqrt{3}} \right)} = \]

\[= -\frac{2\pi}{3} \frac{\sinh \left( \frac{\pi}{2\sqrt{3}} \right)}{-2\sinh^2 \left( \frac{\pi}{2\sqrt{3}} \right)} = \frac{\pi}{3} \frac{1}{\sinh \left( \frac{\pi}{2\sqrt{3}} \right)}.\]

\(I_5:\) Since the integrand is odd in \( t \) the integral vanishes:

\[I_5 = 0.\]

\(I_6:\) Again we use the calculus of residues. Since the integrand is quite similar to that of \(I_4\) we proceed exactly in the same way and note only the differences.

The coefficient \( g_{-1} \) of the Laurent series now reads

\[g_{-1} = \frac{\cos \left( i \frac{\pi}{2\sqrt{3}} \right)}{i\sqrt{3}} = -\frac{i}{\sqrt{3}} \cosh \left( \frac{\pi}{2\sqrt{3}} \right).\]

Thus we have

\[\int_\Gamma g(z)dz = \frac{2\pi}{\sqrt{3}} \cosh \left( \frac{\pi}{2\sqrt{3}} \right).\]

For the integral over \( \Gamma_3 \) we obtain

\[\int_{\Gamma_3} \frac{\cos z}{\cosh(\sqrt{3}z)} dz = \int_{-R}^R \frac{\cos \left( -t + i \frac{\pi}{\sqrt{3}} \right)}{\cosh \left( \sqrt{3} \left( -t + i \frac{\pi}{\sqrt{3}} \right) \right)} (-dt) = \]

\[= \int_{-R}^R \frac{\cos(-t) \cos \left( i \frac{\pi}{\sqrt{3}} \right) - \sin(-t) \sin \left( i \frac{\pi}{\sqrt{3}} \right)}{\cosh(-\sqrt{3}t) \cosh(i\pi) + \sinh(-\sqrt{3}t) \sinh(i\pi)} (-dt) = \]

\[= \cos \left( i \frac{\pi}{\sqrt{3}} \right) \int_{-R}^R \frac{\cos t}{\cosh(\sqrt{3}t)} dt + \sin \left( i \frac{\pi}{\sqrt{3}} \right) \int_{-R}^R \frac{\sin t}{\cosh(\sqrt{3}t)} dt = \]

\[= \cosh \left( \frac{\pi}{\sqrt{3}t} \right) \int_{-R}^R \frac{\cos t}{\cosh(\sqrt{3}t)} dt.\]
Combining these results we obtain
\[
\frac{2\pi}{\sqrt{3}} \cosh \left( \frac{\pi}{2\sqrt{3}} \right) = \int_{\Gamma} g(z)dz = \left( 1 + \cosh \left( \frac{\pi}{\sqrt{3}} \right) \right) \int_{-\infty}^{\infty} \frac{\cos t}{\cosh(\sqrt{3}t)} dt.
\]

Thus we have for \( R \to \infty \)
\[
I_6 = \int_{-\infty}^{\infty} \frac{\cos t}{\cosh(\sqrt{3}t)} dt = \frac{2\pi}{\sqrt{3}} \cosh \left( \frac{\pi}{2\sqrt{3}} \right)
\]
\[
= \frac{2\pi}{\sqrt{3}} \frac{\cosh \left( \frac{\pi}{2\sqrt{3}} \right)}{1 + 2 \cosh^2 \left( \frac{\pi}{2\sqrt{3}} \right) - 1} = \frac{\pi}{\cosh \left( \frac{\pi}{2\sqrt{3}} \right)}. \]

We finally obtain for the Melnikov function
\[
d(\sigma) = -3\sqrt{3} \cos \sigma \cdot I_1 - 3\sqrt{3} \sin \sigma \cdot I_2 + 3 \cos \sigma \cdot I_3 - 3 \sin \sigma \cdot I_4 + 2\sqrt{3} \cos \sigma \cdot I_5 - 2\sqrt{3} \sin \sigma \cdot I_6 =
\]
\[
= -3\sqrt{3} \frac{\pi}{6\sqrt{3} \sinh \left( \frac{\pi}{2\sqrt{3}} \right)} \sin \sigma - 3 \frac{\pi}{3 \sinh \left( \frac{\pi}{2\sqrt{3}} \right)} \sin \sigma - 2\sqrt{3} \frac{\pi}{\sqrt{3} \cosh \left( \frac{\pi}{2\sqrt{3}} \right)} \sin \sigma =
\]
\[
= -\frac{\pi}{2} \left( \frac{3}{\sinh \left( \frac{\pi}{2\sqrt{3}} \right)} + \frac{4}{\cosh \left( \frac{\pi}{2\sqrt{3}} \right)} \right) \sin \sigma.
\]

The Main Result

Now we are in a position to state the main result of the chapter.

**Theorem 9.2.1**

Consider equation (9.3e). For all \( e > 0, e \) sufficiently small the following holds:
The Melnikov function \( d \) defined in (9.2) admits a simple zero. This implies that the restriction of a certain power of the Poincaré map to a suitable set in \( \mathbb{R}^2 \) defines a chaotic dynamical system in the sense of Smale.

**Proof:** The Melnikov function
\[
d(\sigma) = -\frac{\pi}{2} \left( \frac{3}{\sinh \left( \frac{\pi}{2\sqrt{3}} \right)} + \frac{4}{\cosh \left( \frac{\pi}{2\sqrt{3}} \right)} \right) \sin \sigma
\]

obviously admits simple zeroes. Thus the claim follows from Theorem 9.1.4. \( \square \)
Verification of Chaos for Large Eccentricities

In this chapter we present a method for verifying chaotic behavior of the dumbbell satellite problem without the restriction to small eccentricities $e$. As an example we prove chaotic behavior for the eccentricity $e = 0.3$.

We remember: The motion of the satellite around its center of mass is described by

$$\dot{z} = F(t, z, e),$$

where the function $F$ is defined by

$$F(t, z, e) := F(t, x, y, e) = \left(-\frac{3}{2} \frac{\sin(2\pi)}{1-e\cos t} + y \frac{\sin t}{1-e\cos t} + 2e\sqrt{1-e^2}\sin t\right).$$

As in the previous chapter we prove the existence of a homeomorphism $\tau$ between the Bernoulli shift system and a subsystem of $(\mathbb{R}^2, P)$, where $P$ denotes the Poincaré map of (2.9).

Since the eccentricities $e$ are no longer supposed to be small, perturbation methods like the Melnikov method are no longer applicable. Therefore we follow a rather different approach developed by Stoffer, Palmer and Kirchgraber based on numerical shadowing (cf. [22], [24] and [23]).

The main idea is as follows. Assume that we can find two approximate periodic orbits $\bar{u}$ and $\bar{v}$ which are very close at some point. This allows us to construct uncountably many pseudo orbits $q$ of $P$ by arbitrarily combining blocks $u$ and $v$. If applicable the Shadowing Lemma implies the existence of a shadowing orbit to each pseudo orbit. As in the previous chapter this leads to the homeomorphism $\tau$.

The chapter is organized as follows:

**Section 10.1:** We develop the theoretical background. In particular we prove a suitable form of the Shadowing Lemma.
Section 10.2: Step by step we describe an algorithm for the verification of chaotic behavior.

Section 10.3: As an example we apply the algorithm to the dumbbell satellite problem with eccentricity $e = 0.3$.

Section 10.4: We close the chapter with some remarks on the computations carried out on the computer.

### 10.1 Numerical Shadowing

In this section we prove a suitable form of the Shadowing Lemma and we provide computable estimates.

Trying to apply the approach of Stoffer, Palmer and Kirchgraber we were faced with some difficulties. On the one hand we do not easily succeed in finding suitable periodic orbits $\mathbf{u}$ and $\mathbf{v}$, since the dumbbell satellite problem is strongly hyperbolic. On the other hand we do not succeed in fulfilling a number of computable conditions required by the chosen approach.

A solution of both problems bases on the following idea: We subdivide the time interval $[0, T]$ into a number of subintervals of equal length and replace the Poincaré map $P$ by a sequence of maps $P_k$. Figure 10.1 gives a first impression of this idea. The details are presented in the subsequent section.

![Subdivision of the interval $[0, T]$.](image)

**Fig. 10.1:** Subdivision of the interval $[0, T]$.

### Preliminary Definitions

Let $\Omega := (\Omega_n)_{n \in \mathbb{Z}}$ be a sequence of open subsets of $\mathbb{R}^2$ and let $\mathbf{P} := (P_n)_{n \in \mathbb{Z}}$ be a sequence of functions

$$P_n : \Omega_n \to \mathbb{R}^2.$$
10.1. Numerical Shadowing

By the following definitions we generalize the notions introduced in Section 9.1.

**Definition 10.1.1**
A sequence \( p := (p_n)_{n \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \) is called an orbit of \( P := (P_n)_{n \in \mathbb{N}} \), if
\[
p_{n+1} = P_n(p_n) \quad \text{for all} \quad n \in \mathbb{Z}.
\]

A sequence \( q := (q_n)_{n \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \) is called a \( \delta \)-pseudo orbit of \( P := (P_n)_{n \in \mathbb{N}} \), if
\[
|q_{n+1} - P_n(q_n)| \leq \delta \quad \text{for all} \quad n \in \mathbb{Z}.
\]

A sequence \( p := (p_n)_{n \in \mathbb{Z}} \) is called a \( \rho \)-shadowing orbit of a pseudo orbit \( q := (q_n)_{n \in \mathbb{Z}} \), if \( p \) is an orbit and if
\[
\|q - p\|_\infty \leq \rho.
\]

As pointed out at the beginning we are interested in the case where the sequence of functions is generated from a block of finite length. We call such a sequence of functions periodic.

**Definition 10.1.2**
A sequence \( P := (P_n)_{n \in \mathbb{Z}} \) is called periodic with period \( K \) if
\[
P_{n+K} = P_n, \quad \text{for all} \quad n \in \mathbb{Z}.
\]

If the sequence of functions is periodic, it also makes sense to define periodic orbits.

**Definition 10.1.3**
Let \( P := (P_n)_{n \in \mathbb{N}} \) be periodic with period \( K \). An orbit \( p := (p_n)_{n \in \mathbb{Z}} \in \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \) of \( P \) is called periodic with period \( N \) if there exists a number \( N' \in \mathbb{N} \) such that
\[
p_{n+N} = p_n, \quad \text{for} \quad N := KN' \quad \text{and} \quad \text{for all} \quad n \in \mathbb{Z}.
\]

An analogous definition holds for periodic pseudo orbits.

**The Theorem of Newton-Kantorovich**
Consider the non-linear equation
\[
F(z) = 0.
\]

From calculus the so-called method of Newton is well known: Starting with some initial value \( z^{(0)} \) one computes successive approximations of a solution by the formula
\[
z^{(k+1)} := z^{(k)} - \frac{F(z^{(k)})}{F'(z^{(k)})}.
\]
The convergence of the sequence \((z^{(k)})_{k\in\mathbb{N}}\) is guaranteed under certain assumptions on the function \(F\) and on the starting value \(z^{(0)}\).

In this section we state a generalization of Newton’s method. It is due to L.V. Kantorovich and therefore referred to as Theorem of Newton-Kantorovich.

**Theorem 10.1.1**

Let \((E, \|\cdot\|)\) denote a Banach space and \(\Omega\) is an open subset of \(E\). Consider a function \(F \in C^1,\text{Lip}(\Omega, E)\), where \(\Omega\) is convex subset of \(\Omega\).

Suppose that there exists a \(q \in E\) such that

1. \(\|(DF(q))^{-1}F(q)\| \leq \alpha\),
2. \(\|(DF(q))^{-1}\| \leq \beta\),
3. \(\|DF(p) - F(p')\| \leq \gamma\|p - p'\|\) for \(p, p' \in \Omega\),

where \(\alpha, \beta\) and \(\gamma\) are positive constants.

Then the following holds:

**Part A:** If

\[
\alpha \beta \gamma \leq \frac{1}{2}
\]

and \(B_\rho(q) \subset \Omega\) for

\[
\rho \geq \rho_0 := \frac{1 - \sqrt{1 - 2\alpha \beta \gamma}}{\beta \gamma},
\]

then the equation

\[
F(p) = 0.
\]

has a solution \(p\) with \(\|p - q\| \leq \rho\) and the sequence

\[
p^{(0)} := q, \quad p^{(k+1)} := p^{(k)} - (DF(p^{(k)}))^{-1}F(p^{(k)})
\]

converges to \(p\).

**Part B:** If

\[
\alpha \beta \gamma < \frac{1}{2}
\]

and

\[
\rho_0 \leq \rho < \rho_1 := \frac{1 + \sqrt{1 - 2\alpha \beta \gamma}}{\beta \gamma},
\]

then \(p\) is the only solution with \(\|p - q\| \leq \rho\).

A complete treatment together with the proofs may be found in [19].
The Shadowing Lemma

In this section we state and prove a suitable form of the famous Shadowing Lemma.

**Theorem 10.1.2 (Shadowing Lemma)**

Let $\Omega := (\Omega_n)_{n \in \mathbb{Z}}$ be a sequence of open convex sets of $\mathbb{R}^2$ and let $q = (q_n)_{n \in \mathbb{Z}}$ be a sequence of points with $q_n \in \Omega_n$ for all $n \in \mathbb{Z}$.

Let $P = (P_n)_{n \in \mathbb{Z}}$ be a sequence of functions with the following properties:

(i) $P_n : \Omega_n \rightarrow \mathbb{R}^2$ is one-to-one for all $n \in \mathbb{Z}$.

(ii) $P_n$ is continuously differentiable for all $n \in \mathbb{Z}$.

(iii) the derivative $D_P n$ is Lipschitz-continuous on $\Omega_n$ with Lipschitz-constant $M > 0$, independent of $n$:

$$|D_P n(u) - D_P n(v)| \leq M|u - v| \quad \text{for all } u, v \in \Omega_n, \quad n \in \mathbb{Z}.$$ 

Define sequences $d = (d_n)_{n \in \mathbb{Z}}$ and $\theta = (\theta_n)_{n \in \mathbb{Z}}$ in $\ell^\infty(\mathbb{Z}, \mathbb{R}^2)$ and a sequence of bounded $2 \times 2$-matrices $A = (A_n)_{n \in \mathbb{Z}}$ such that for all $n \in \mathbb{Z}$ the following holds:

(iv) $q_{n+1} - P_{n}(q_n) = d_n + \theta_n$.

(v) $|\theta_n| \leq \delta_0$ some positive constant $\delta_0$.

(vi) $|A_n - D_P n(q_n)| \leq \delta_1$ for some positive constant $\delta_1$.

Assume that the operator $C : \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \rightarrow \ell^\infty(\mathbb{Z}, \mathbb{R}^2)$ has the following properties:

(vii) $C$ is invertible and $\|C^{-1}\| \leq l$ for some positive constant $l$.

(viii) $\|C^{-1}d\| \leq l_d$ for some positive constant $l_d$.

(ix) the constant $l$ satisfies

$$l \leq \frac{1}{\delta_1 + l_d M + \sqrt{2M \delta_0 + 2\delta_1 l_d M + (l_d M)^2}}.$$

Then the numbers

$$\rho_0 := \frac{2(l_d + \|C^{-1}\| \delta_0)}{1 - l\delta_1 + \sqrt{(1 - l\delta_1)^2 - 2lM(l_d + l\delta_0)}},$$

$$\rho_1 := \frac{1 - l\delta_1 + \sqrt{(1 - l\delta_1)^2 - 2lM(l_d + l\delta_0)}}{lM}$$

satisfy $0 < \rho_0 < \rho_1$. Finally for some $\rho^* \in [\rho_0, \rho_1]$ let $\overline{B_{\rho^*}}(q_n) \subset \Omega_n$ for all $n \in \mathbb{Z}$.

Then there is a $\rho^*$-shadowing orbit $p = (p_n)_{n \in \mathbb{Z}}$ of $q$. Moreover, there is no orbit $\tilde{p}$ other than $p$ with $\|\tilde{p} - q\|_\infty < \rho^*$. 


Proof: The goal is to apply the Theorem of Newton-Kantorovich. To this end we introduce the operator

$$\mathcal{F} : B_{p^*}(q) \to \ell^\infty(\mathbb{Z}, \mathbb{R}^2), \quad (\mathcal{F}(q))_n = q_{n+1} - P_n(q_n).$$

This allows us to restate the claim of the Shadowing Lemma in the following form: There is a unique $p \in B_{p^*}(q)$ such that $\mathcal{F}(p) = 0$ and $\|p - q\| \leq \rho_0$. It remains to show that the assumptions of the Theorem of Newton-Kantorovich are fulfilled.

- $\|D\mathcal{F}(q)^{-1}\mathcal{F}(q)\| \leq \alpha$:

  Introducing the linear operator $\mathcal{G}$ by

  $$\mathcal{G} : \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \to \ell^\infty(\mathbb{Z}, \mathbb{R}^2), \quad (\mathcal{G}(\mathcal{G})_n = (A_n - DP_n(q_n))\mathcal{G}_n.$$

  we may write

  $$\left(D\mathcal{F}(q)\mathcal{G}_n\right)_n = z_{n+1} - DP_n(q_n)z_n = (L\mathcal{G})_n + (\mathcal{G}\mathcal{G}_n).$$

  Since $(vii)$ holds we may further write $D\mathcal{F}(q) = L(I + L^{-1}\mathcal{G})(q)$. From $(vi)$ and $(ix)$ it follows that $\|L^{-1}\mathcal{G}\| \leq \|L^{-1}\|\delta_1 < 1$ and therefore that $D\mathcal{F}(q)$ is invertible, too:

  $$D\mathcal{F}(q)^{-1} = (I + L^{-1}\mathcal{G})^{-1}L^{-1}. \quad (10.1)$$

  Using $(iv)$ and $(v)$ we conclude that

  $$\|D\mathcal{F}(q)^{-1}\mathcal{F}(q)\| \leq \frac{1}{1 - \|L^{-1}\|\delta_1} (\|L^{-1}d\| + \|L^{-1}\|\delta_0) \leq \frac{l_d + l\delta_0}{1 - l\delta_1}.$$ 

  We therefore set

  $$\alpha := \frac{l_d + l\delta_0}{1 - l\delta_1}.$$

- $\|D\mathcal{F}(q)^{-1}\| \leq \beta$:

  From 10.1 it follows that

  $$\|D\mathcal{F}(q)^{-1}\| \leq \frac{1}{1 - \|L^{-1}\|\delta_1} \|L^{-1}\| \leq \frac{l}{1 - l\delta_1}.$$ 

  We therefore set

  $$\beta := \frac{l}{1 - l\delta_1}.$$

- $\|D\mathcal{F}(p') - D\mathcal{F}(p'')\| \leq M\|p' - p''\|$ for all $p', p'' \in B_{p^*}(q)$:

  From $(iii)$ it follows that

  $$\|((D\mathcal{F}(p') - D\mathcal{F}(p''))(\mathcal{G}))_n\| = \|(DP_n(p'_n) - DP_n(p''_n))\mathcal{G}_n\| \leq M|p'_n - p''_n| \|\mathcal{G}_n\| \leq M\|p' - p''\|\|\mathcal{G}_n\|.$$ 

  We therefore set

  $$\gamma := M.$$
Substituting the definitions of $\alpha$, $\beta$ and $\gamma$ into $\alpha \beta \gamma < \frac{1}{2}$ and using $l\delta_1 < 1$ we find

$$\alpha \beta \gamma < \frac{1}{2}$$

$$\frac{l_d + l\delta_0}{1 - l\delta_1} \cdot \frac{l}{1 - l\delta_1} \cdot M < \frac{1}{2}$$

$$0 < (1 - l\delta_1)^2 - 2lM(l_d + l\delta_0)$$

$$0 < l^2(\delta_1^2 - 2\delta_0M) - 2l(\delta_1 + l_dM) + 1 =: h(l).$$

Number and position of the zeroes of the function $h$ depend on its leading coefficient $\delta_1^2 - 2\delta_0M$. For $\delta_1^2 - 2\delta_0M = 0$ we have a linear function with the single zero:

$$l_+ := \frac{1}{2(\delta_1 + l_dM)}.$$

For $\delta_1^2 - 2\delta_0M \neq 0$ we have a quadratic function with two zeroes:

$$l_\pm = \frac{2(\delta_1 + l_dM) \pm \sqrt{4(\delta_1 + l_dM)^2 - 4(\delta_1^2 - 2\delta_0M)}}{2(\delta_1^2 - 2\delta_0M)} = \frac{1}{\delta_1 + l_dM \mp \sqrt{(\delta_1 + l_dM)^2 - (\delta_1^2 - 2\delta_0M)}}.$$

Figure 10.2 gives a survey.

$$\delta_1^2 - 2\delta_0M > 0 \quad \delta_1^2 - 2\delta_0M = 0 \quad \delta_1^2 - 2\delta_0M < 0$$

**Fig. 10.2:** Number and position of the zeroes of $h$.

Obviously Assumption (ix) is equivalent to $0 < l < l_\pm$. Therefore it implies that $h(l) > 0$ and furthermore $\alpha \beta \gamma < \frac{1}{2}$. Thus the assumptions of the Theorem of Newton-Kantorovich hold indeed.

Substituting $\alpha$, $\beta$ and $\gamma$ into

$$\frac{1 - \sqrt{1 - 2\alpha \beta \gamma}}{\beta \gamma} \quad \text{and} \quad \frac{1 + \sqrt{1 - 2\alpha \beta \gamma}}{\beta \gamma}$$
respectively, we obtain the the constants $p_0$ and $p_1$ defined in the assumptions of this theorem.

We conclude by the theorem of Newton-Kantorovich that $\mathcal{F}$ has a unique zero $p$ in $B_{p_0}(q)$ and that there is no other zero in $B_{p^*}(q) \setminus \{p\}$ for $p_0 \leq p^* < p_1$. □

**Estimates for $||\mathcal{L}^{-1}||$ and $||\mathcal{L}^{-1}d||$**

In this section we give computable estimates for $||\mathcal{L}^{-1}||$ and $||\mathcal{L}^{-1}d||$ under the following assumptions:

- The dimension of the underlying Euclidian space is 2.
- The sequence of functions $P = (P_n)_{n \in \mathbb{Z}}$ is $K$-periodic.
- The sequence $q = (q_n)_{n \in \mathbb{Z}}$ consists of blocks $u$ and $v$ of periodic pseudo orbits:
  
  $u = (u_1, u_2, \ldots, u_{N_u})$, \hspace{1cm} where \hspace{1cm} $N_u = K N_u', N_u' \in \mathbb{N}$
  
  $v = (v_1, v_2, \ldots, v_{N_v})$, \hspace{1cm} where \hspace{1cm} $N_v = K N_v', N_v' \in \mathbb{N}$

- $A_n$ is the computed value of $DP_n(q_n)$.

Under these assumptions the sequence $A = (A_n)_{n \in \mathbb{Z}}$ consists of blocks of length $N_u$ and $N_v$, respectively. Thus $A_n := A(u_i)$ for some $i$ or $A_n := A(v_j)$ for some $j$.

The matrices $A(u_i)$, $1 \leq i \leq N_u$ and $A(v_j)$, $1 \leq j \leq N_v$ may be **triangularized** in the following way: There are almost orthogonal matrices $Q(u_i)$ and $Q(v_j)$ and upper triangular matrices $R(u_i)$ and $R(v_j)$ such that

$A(u_i)Q(u_i) \approx Q(u_{i+1})R(u_i)$, \hspace{1cm} index mod $N_u$;

$A(v_j)Q(v_j) \approx Q(v_{j+1})R(v_j)$, \hspace{1cm} index mod $N_v$.

$R(u_i)$, $R(v_j)$ and $Q(u_i)$, $Q(v_j)$ may be computed iteratively. For details see Section 10.2. These matrices allow us to give computable estimates of $||\mathcal{L}^{-1}||$ and $||\mathcal{L}^{-1}d||$ mentioned above.

**Lemma 10.1.1**

Let $A = (A_n)_{n \in \mathbb{Z}}$, $Q = (Q_n)_{n \in \mathbb{Z}}$, $R = (R_n)_{n \in \mathbb{Z}}$ be bounded sequences of $2 \times 2$-matrices with the following properties:

(i) There exists a constant $\delta_2$ satisfying $0 \leq \delta_2 \leq 1$ such that

$|A_nQ_n - Q_{n+1}R_n| \leq \delta_2$ \hspace{1cm} for all \hspace{1cm} $n \in \mathbb{Z}$.

(ii) There exists a constant $\delta_3$ satisfying $0 \leq \delta_3 < 1$ such that

$|Q_n^TQ_n - I| \leq \delta_3$ \hspace{1cm} for all \hspace{1cm} $n \in \mathbb{Z}$. 
Define the operators \( \mathcal{L}, \mathcal{R}, \mathcal{Q} \) and \( \mathcal{Q} \) by

\[
\mathcal{L} : \ell^\infty(Z, \mathbb{R}^2) \to \ell^\infty(Z, \mathbb{R}^2), \quad (\mathcal{L} \zeta)_n := \zeta_{n+1} - A_n \zeta_n, \quad n \in \mathbb{Z},
\]

\[
\mathcal{R} : \ell^\infty(Z, \mathbb{R}^2) \to \ell^\infty(Z, \mathbb{R}^2), \quad (\mathcal{R} \zeta)_n := \zeta_{n+1} - R_n \zeta_n, \quad n \in \mathbb{Z},
\]

\[
\mathcal{Q} : \ell^\infty(Z, \mathbb{R}^2) \to \ell^\infty(Z, \mathbb{R}^2), \quad (\mathcal{Q} \zeta)_n := Q_n \zeta_n, \quad n \in \mathbb{Z},
\]

\[
\mathcal{Q} : \ell^\infty(Z, \mathbb{R}^2) \to \ell^\infty(Z, \mathbb{R}^2), \quad (\mathcal{Q} \zeta)_n := Q_{n+1} \zeta_n, \quad n \in \mathbb{Z}.
\]

If in addition

(iii) \( \mathcal{R} \) is invertible and

(iv) \( \|\mathcal{R}^{-1}\| \) satisfies

\[
\|\mathcal{R}^{-1}\| \delta_2 < \sqrt{1 - \delta_3},
\]

then the operator \( \mathcal{L} \) is invertible and the following estimates hold

\[
\|\mathcal{L}^{-1}\| \leq \frac{\sqrt{1 + \delta_3}}{\sqrt{1 - \delta_3} - \|\mathcal{R}^{-1}\| \delta_2} \|\mathcal{R}^{-1}\|,
\]

\[
\|\mathcal{L}^{-1} d\| \leq \frac{\sqrt{1 - \delta_3^2}}{\sqrt{1 - \delta_3} - \|\mathcal{R}^{-1}\| \delta_2} \|\mathcal{R}^{-1} \mathcal{Q}^{-1} d\|.
\]

**Proof:** We first show that \( \mathcal{Q} \) and \( \mathcal{Q} \) are invertible and give estimates for their norms. To this end we consider the identity

\[
Q_n^T Q_n = I + (Q_n^T Q_n - I).
\]  \hspace{1cm} (10.2)

On one hand the identity (10.2) implies, using assumption (ii), that

\[
|Q_n| \leq \sqrt{1 + \delta_3}
\]

and therefore

\[
\|Q\| = \|Q\| \leq \sqrt{1 + \delta_3}.
\]

On the other hand identity (10.2) allows us to expand \((Q_n^T Q_n)^{-1}\) into a von Neumann series:

\[
|(Q_n^T Q_n)^{-1}| = \left| (I + (Q_n^T Q_n - I))^{-1} \right| = \sum_{k=0}^{\infty} (-Q_n^T Q_n + I)^k \leq \sum_{k=0}^{\infty} |Q_n^T Q_n - I|^k \leq \sum_{k=0}^{\infty} \delta_3^k = \frac{1}{1 - \delta_3}.
\]
This implies that
\[ |Q_n^{-1}| \leq \frac{1}{\sqrt{1-\delta_3}} \]
holds. Thus \( Q^{-1} \) and \( \tilde{Q}^{-1} \) exist and the following estimates hold:
\[ \|Q^{-1}\| = \|\tilde{Q}^{-1}\| \leq \frac{1}{\sqrt{1-\delta_3}}. \quad (10.3) \]

Now let \( \zeta := (\zeta_n)_{n \in \mathbb{Z}} \) be any sequence in \( \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \). We show that \( \mathcal{L}z = \zeta \) has a unique solution \( z \) and we estimate \( \|z\| \) in terms of \( \|\zeta\| \).

To this end we apply the transformation \( z_n := Q_n \hat{z}_n \) to \( \mathcal{L}z = \zeta \). We obtain
\[ Q_{n+1} \hat{z}_{n+1} - A_n Q_n \hat{z}_n = \zeta_n, \quad n \in \mathbb{Z} \]
or
\[ Q_{n+1} (\hat{z}_{n+1} - R_n \hat{z}_n) - (A_n Q_n - Q_{n+1} R_n) \hat{z}_n = \zeta_n, \quad n \in \mathbb{Z}, \quad (10.4) \]
respectively. Defining the operator \( \mathcal{S} \) by
\[ \mathcal{S} : \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \to \ell^\infty(\mathbb{Z}, \mathbb{R}^2), \quad (\mathcal{S}z)_n := (A_n Q_n - Q_{n+1} R_n)z_n, \]
equation (10.4) may be written as
\[ (\mathcal{Q} \mathcal{R} - \mathcal{S}) \hat{z} = \zeta, \]
or by assumption (iii) as
\[ \tilde{Q} \mathcal{R} \left( I - \left( \tilde{Q} \mathcal{R} \right)^{-1} \mathcal{S} \right) \hat{z} = \zeta. \quad (10.5) \]

From assumptions (i) and (iv) and estimate (10.3), it follows that
\[ \| (\tilde{Q} \mathcal{R})^{-1} \mathcal{S} \| \leq \| \mathcal{R}^{-1} \| \cdot \| \tilde{Q}^{-1} \| \cdot \| \mathcal{S} \| \leq \frac{\sqrt{1-\delta_3}}{\delta_2} \cdot \frac{1}{\sqrt{1-\delta_3}} \cdot \delta_2 = 1. \]

Thus equation (10.5) may be solved for \( \hat{z} \) and we obtain for \( \|\hat{z}\|\):
\[ \|\hat{z}\| \leq \frac{1}{1 - \frac{\|\mathcal{R}^{-1}\| \delta_2}{\sqrt{1-\delta_3}}} \|\mathcal{R}^{-1} \tilde{Q}^{-1} \zeta\|. \]

Due to \( z = \mathcal{Q} \hat{z} \) and with the estimate on \( \|\mathcal{Q}\| \), see the formula following (10.2) and (10.3), we find
\[ \|\mathcal{L}^{-1} \zeta\| = \|z\| \leq \sqrt{1+\delta_3} \frac{1}{1 - \frac{\|\mathcal{R}^{-1}\| \delta_2}{\sqrt{1-\delta_3}}} \|\mathcal{R}^{-1} \tilde{Q}^{-1} \zeta\| = \]
\[ = \frac{\sqrt{1-\delta_3}}{\sqrt{1-\delta_3} - \delta_2 \|\mathcal{R}^{-1}\|} \|\mathcal{R}^{-1} \tilde{Q}^{-1} \zeta\| = \]
\[ \leq \frac{\sqrt{1+\delta_3}}{\sqrt{1-\delta_3} - \delta_2 \|\mathcal{R}^{-1}\|} \|\mathcal{R}^{-1}\| \cdot \|\zeta\|. \]
The last two lines correspond to the estimates in the claim. □

The next lemma provides an estimate for \( \| R^{-1} \| \).

**Lemma 10.1.2**

Consider the operators \( \tilde{Q} \) and \( R \) defined in Lemma 10.1.1 and let the sequence \( R = (R_n)_{n \in \mathbb{Z}} \) consist of blocks \((R(u_1), R(u_2), \ldots, R(u_{N_u})), (R(v_1), R(v_2), \ldots, R(v_{N_v}))\) of triangular matrices:

\[
R(u_i) = \begin{pmatrix}
a(u_i) & b(u_i) \\
0 & c(u_i)
\end{pmatrix}, \quad R(v_j) = \begin{pmatrix}
a(v_j) & b(v_j) \\
0 & c(v_j)
\end{pmatrix}.
\]

Assume that there are positive numbers \( \eta_i^u, 1 \leq i \leq N_u \) and \( \eta_j^v, 1 \leq j \leq N_v \) satisfying

\[
\eta_i^u = \eta_i^v \geq \max \{ |c(u_{N_u})| \eta_i^u, |c(v_{N_v})| \eta_j^v \} + 1,
\]

\[
\eta_i^u + 1 \geq |c(u_i)| \eta_i^u + 1, \quad 1 \leq i \leq N_u - 1,
\]

\[
\eta_j^v + 1 \geq |c(v_j)| \eta_j^v + 1, \quad 1 \leq j \leq N_v - 1.
\]

and assume that there are positive numbers \( \xi_i^u, 1 \leq i \leq N_u \) and \( \xi_j^v, 1 \leq j \leq N_v \) satisfying

\[
\xi_i^u \geq \frac{1}{|a(u_{N_u})|} \left( \max \{ \xi_i^u, \xi_j^v \} + |b(u_{N_u})| \eta_i^u + 1 \right),
\]

\[
\xi_i^u \geq \frac{1}{|a(v_{N_v})|} \left( \max \{ \xi_i^u, \xi_j^v \} + |b(v_{N_v})| \eta_j^v + 1 \right),
\]

\[
\xi_i^u \geq \frac{1}{|a(u_i)|} \left( \xi_i^u + 1 + |b(u_i)| \eta_i^u + 1 \right), \quad 1 \leq i \leq N_u - 1,
\]

\[
\xi_j^v \geq \frac{1}{|a(v_j)|} \left( \xi_j^v + 1 + |b(v_j)| \eta_j^v + 1 \right), \quad 1 \leq j \leq N_v - 1.
\]

Then the following estimate holds:

\[
\| R^{-1} \| \leq \max_{1 \leq i \leq N_u, 1 \leq j \leq N_v} \left\{ \sqrt{ (\xi_i^u)^2 + (\eta_i^u)^2 }, \sqrt{ (\xi_j^v)^2 + (\eta_j^v)^2 } \right\}.
\]

**Proof:** This proof is arranged in two parts. In the first part we show that the equation \( Rz = z' \) has a unique solution for any \( z' = (z'_n)_{n \in \mathbb{Z}} \) with \( \| z' \| = 1 \). In the second part we give an estimate for this solution and therefore for \( \| R^{-1} \| \).

Let \( z_n =: \begin{pmatrix} x_n \\ y_n \end{pmatrix} \) and \( z'_n =: \begin{pmatrix} x'_n \\ y'_n \end{pmatrix} \). Then \( Rz = z' \) may be written as

\[
z_{n+1} - R_n z_n = z'_n, \quad n \in \mathbb{Z}
\]

or

\[
x_{n+1} - a_n x_n - b_n y_n = x'_n, \quad y_{n+1} - c_n y_n = y'_n, \quad n \in \mathbb{Z}.
\]
We first concentrate on the \( y \)-equations which are independent of \( x \). By repeated substitution we formally obtain an explicit expression for \( y_n \):

\[
y_n = \sum_{k=1}^{\infty} y'_{n-k} \prod_{i=1}^{k-1} c_{n-i}.
\]

(10.7)

It remains to show that the right-hand side converges. We prove this in two steps.

- From the assumptions on the numbers \( \eta_i^u \) and \( \eta_j^v \) we conclude that

\[
\eta_i^u \geq \prod_{i=1}^{N_u} c(u_i) \cdot \eta_i^u + 1, \quad \eta_j^v \geq \prod_{j=1}^{N_v} c(v_j) \cdot \eta_j^v + 1,
\]

or

\[
\eta_i^u \left(1 - \left| \prod_{i=1}^{N_u} c(u_i) \right| \right) \geq 1, \quad \eta_j^v \left(1 - \left| \prod_{j=1}^{N_v} c(v_j) \right| \right) \geq 1.
\]

Since \( \eta_i^u, \eta_j^v > 1 \) it follows that

\[
\left| \prod_{i=1}^{N_u} c(u_i) \right| < 1, \quad \left| \prod_{j=1}^{N_v} c(v_j) \right| < 1. \quad (10.8)
\]

- Let

\[
y_n^{(m)} := \sum_{k=1}^{m_2} y'_{n-k} \prod_{i=1}^{k-1} c_{n-i}.
\]

We must show that for every \( \varepsilon > 0 \) there exists a number \( m_0 \in \mathbb{N} \) such that for every \( m_1, m_2 > m_0 \) we have

\[
|y_n^{(m_2)} - y_n^{(m_1)}| < \varepsilon.
\]

Let \( m_2 > m_1 \). Since \( |y'_{n}| \leq 1 \) for all \( n \in \mathbb{Z} \) we have

\[
|y_n^{(m_2)} - y_n^{(m_1)}| = \left| \sum_{k=m_1+1}^{m_2} y'_{n-k} \prod_{l=1}^{k-1} c_{n-l} \right| \leq \sum_{k=m_1+1}^{m_2} \left| \prod_{l=1}^{k-1} c_{n-l} \right| \leq \sum_{k=m_1+1}^{m_2} \left( \prod_{l=1}^{k-1} c_{n-l} \right) \cdot \left| \sum_{k=m_1+1}^{m_2} \prod_{l=1}^{k-1} c_{n-l} \right|.
\]

The products \( \prod_{l=m_0+1}^{k} c_{n-l} \) consist of blocks \( \prod_{i=1}^{N_u} c(u_i), \prod_{j=1}^{N_v} c(v_j) \) and at most of two sections of these blocks. While the absolute values of the sections may be large, the absolute values of the blocks are bounded by some positive number \( q < 1 \). We conclude that there exists a positive number \( C \) such that

\[
\sum_{k=m_1}^{m_2-1} \left| \prod_{l=m_0+1}^{k} c_{n-l} \right| < C.
\]
The leading term \( \prod_{i=1}^{m_0} c_i \) consists also of the blocks with \( \prod_{i=1}^{N_u} c(u_i) \) and \( \prod_{j=1}^{N_v} c(v_j) \) and at most two sections of these blocks. We conclude that for \( m_0 \) large enough, \( \prod_{i=1}^{m_0} c_i \) is smaller than \( \varepsilon/C \). Thus we have

\[
|y_n^{(m_2)} - y_n^{(m_1)}| < \frac{\varepsilon}{C} \cdot C = \varepsilon.
\]

That proves that the right-hand side of (10.7) is convergent. Moreover the sequence \( (y_n)_{n \in \mathbb{Z}} \) is uniformly bounded.

Now we concentrate on the recursive equation in the \( x \)-variable. To this end let \( x'_n := x'_n + b_n y_n \). By repeated substitution we formally obtain an explicit expression for \( x_n \):

\[
x_n = -\sum_{k=0}^{\infty} \frac{1}{\prod_{l=0}^{k} a_{n+l}} \left( \prod_{l=0}^{k} a_{n+l} \right)^{-1}.
\]

It remains to show that the right-hand side converges. The following two steps are similar to those above.

1. From the assumptions on the numbers \( \xi_i^u \) and \( \xi_j^v \) we conclude that

\[
\xi_i^u \geq \frac{1}{\prod_{i=1}^{N_u} a(u_i)} \cdot \eta_i^u + \frac{1}{|a(u_{N_u})|}, \quad \xi_j^v \geq \frac{1}{\prod_{j=1}^{N_v} c(v_j)} \cdot \eta_j^v + \frac{1}{|a(v_{N_v})|},
\]

or

\[
\xi_i^u \left( 1 - \frac{1}{\prod_{i=1}^{N_u} a(u_i)} \right) \geq \frac{1}{|a(u_{N_u})|}, \quad \xi_j^v \left( 1 - \frac{1}{\prod_{j=1}^{N_v} c(v_j)} \right) \geq \frac{1}{|a(v_{N_v})|}.
\]

Since \( \xi_i^u, \xi_j^v > 1 \) it follows that

\[
\prod_{i=1}^{N_u} a(u_i) > 1, \quad \prod_{j=1}^{N_v} c(v_j) > 1.
\]

2. Let

\[
x_n^{(m)} := -\sum_{k=0}^{m} \frac{1}{\prod_{l=0}^{k} a_{n+l}} \left( \prod_{l=0}^{k} a_{n+l} \right)^{-1}.
\]

Again we must show that for every \( \varepsilon > 0 \) there exists a number \( m_0 \in \mathbb{N} \) such that for every \( m_1, m_2 > m_0 \) we have

\[
|x_n^{(m_2)} - x_n^{(m_1)}| < \varepsilon.
\]

Since \( (x'_n)_{n \in \mathbb{Z}}, (b_n)_{n \in \mathbb{Z}} \) and \( (y_n)_{n \in \mathbb{Z}} \) are uniformly bounded sequences \( (\bar{x'}_n)_{n \in \mathbb{Z}} \) is also uniformly bounded. Thus there exist a number \( B \) such that \( |\bar{x}'_n| < B \) for all \( n \in \mathbb{Z} \).
344 10. Verification of Chaos for Large Eccentricities

Therefore we have

\[ |x^{(m_2)}_n - x^{(m_1)}_n| = \left| \sum_{k=m_1+1}^{m_2} \tilde{x}_{n+k} \cdot \left( \prod_{l=0}^{k} a_{n+l} \right)^{-1} \right| \leq B \cdot \sum_{k=m_1+1}^{m_2} \left| \prod_{l=0}^{k} a_{n+l} \right|^{-1} \leq B \cdot \prod_{l=m_1+1}^{m_2} a_{n+l}^{-1} \cdot \sum_{k=m_1+1}^{m_2} \left| \prod_{l=m_1+1}^{k} a_{n+l} \right|^{-1} . \]

The products \( \left( \prod_{l=m_1+1}^{k} a_{n+l} \right)^{-1} \) consist of blocks \( \left( \prod_{l=1}^{N_u} a(u_l) \right)^{-1} \), \( \left( \prod_{j=1}^{N_v} a(v_j) \right)^{-1} \) and at most two sections of these blocks. Since the blocks are bounded by some positive number smaller than 1, there exists a positive number \( C \) such that

\[ \sum_{k=m_1+1}^{m_2} \left| \prod_{l=m_1+1}^{k} a_{n+l} \right|^{-1} < C, \quad k \geq m_1. \]

The leading term \( |\prod_{l=0}^{m_1} a_{n+l}|^{-1} \) also consists of the blocks with \( |\prod_{l=1}^{N_u} c(u_l)| < 1 \), \( |\prod_{j=1}^{N_v} c(v_j)| < 1 \) and at most two sections of these blocks. We conclude that for \( m_0 \) large enough the product \( |\prod_{l=0}^{m_0} a_{n+l}|^{-1} \) is smaller than \( \varepsilon/BC \). Thus we have

\[ |x^{(m_2)}_n - x^{(m_1)}_n| < \frac{\varepsilon}{BC} = \varepsilon. \]

This proves that the right-hand side of (10.9) is convergent. Moreover the sequence \((y_n)_{n \in \mathbb{Z}}\) is uniformly bounded.

All this shows that the equation \( \mathcal{R}z = z' \) has a unique uniformly bounded solution. Thus \( \mathcal{R} \) is invertible. This completes the proof of the first part. It remains to give estimates on \( x_n \) and \( y_n \).

We define the sequences \((\xi_n)_{n \in \mathbb{Z}}\) and \((\eta_n)_{n \in \mathbb{Z}}\) by

\[ \xi_n := \begin{cases} \xi^u_i & \text{if } x_n = u_i \\ \xi^v_j & \text{if } x_n = v_j \end{cases}, \quad \eta_n := \begin{cases} \eta^u_i & \text{if } x_n = u_i \\ \eta^v_j & \text{if } x_n = v_j \end{cases}. \]

On the one hand it follows from (10.6) that

\[ |x_n| \leq \frac{1}{|a_n|} (|x_{n+1}| + |b_n| |y_n| + 1) , \]
\[ |y_n| \leq |c_{n-1}| |y_{n-1}| + 1. \]

On the other hand it follows from the assumptions of the lemma that

\[ \xi_n \geq \frac{1}{|a_n|} (\xi_{n+1} + |b_n| \eta_n + 1) , \quad \eta_n \geq |c_{n-1}| \eta_{n-1} + 1. \]
Taking the difference of the corresponding equations we obtain

\[ |x_n| - \xi_n \leq \frac{1}{|a_n|} (|x_{n+1}| - \xi_{n+1} + |b_n| (|y_n| - \eta_n)) \]

\[ |y_n| - \eta_n \leq |c_{n-1}| (|y_{n-1}| - \eta_{n-1}) \]

and further

\[ |x_{n+1}| - \xi_{n+1} \geq |a_n| (|x_n| - \xi_n) - |b_n| (|y_n| - \eta_n) \] (10.11)

\[ |y_{n-1}| - \eta_{n-1} \geq \frac{1}{|c_{n-1}|} (|y_n| - \eta_n) . \] (10.12)

Repeated use of (10.12) yields

\[ |y_{n-k}| - \eta_{n-k} \geq \frac{1}{\prod_{j=n-k}^{n-1} |c_j|} (|y_n| - \eta_n) . \] (10.13)

If \(|y_n| > \eta_n\) was true for some \(n\), then inequality (10.8) would imply that the right-hand side of (10.13) increases unboundedly for \(k \to \infty\) in contradiction to the invertibility of \(\mathcal{R}\). Thus we may conclude that the sequence \((|y_n|)_{n \in \mathbb{Z}}\) is majorized by \((\eta_n)_{n \in \mathbb{Z}}\).

Using this result we may simplify (10.11) to

\[ |x_{n+1}| - \xi_{n+1} \geq |a_n| (|x_n| - \xi_n) . \]

Now the same argument as above applies. Thus we conclude that the sequence \((|x_n|)_{n \in \mathbb{Z}}\) is majorized by \((\xi_n)_{n \in \mathbb{Z}}\). The estimate of \(\mathcal{R}^{-1}\) follows immediately.

The next lemma provides an estimate for \(\|\mathcal{R}^{-1} \hat{Q}^{-1} d\|\).

**Lemma 10.1.3**

*Consider the operators \(\hat{Q}\) and \(\mathcal{R}\) defined in Lemma 10.1.1 and let the sequence \(R = (R_n)_{n \in \mathbb{Z}}\) consist of blocks \((R(u_1), R(u), \ldots, R(u_{N_u}))\) and \((R(v_1), R(v), \ldots, R(v_{N_v}))\) of triangular matrices:*

\[ R(u) = \begin{pmatrix} a(u) & b(u) \\ 0 & c(u) \end{pmatrix}, \quad R(v) = \begin{pmatrix} a(v) & b(v) \\ 0 & c(v) \end{pmatrix} . \]

*Let \(d = (d_n)_{n \in \mathbb{Z}}\) be the sequence defined by*

\[ d_n = \begin{cases} \pm(u_1 - v_1), & \text{for } n = l_k, \text{ where } l_k \text{ are the indices at which the blocks change from } u \text{ to } v \text{ or vice versa,} \\ 0 & \text{else} \end{cases} \]

*and let \((\mathcal{Q}^u) = Q(u_1)^{-1}(u_1 - v_1)\) and \((\mathcal{Q}^v) = Q(v_1)^{-1}(u_1 - v_1)\).*

*Assume that there are numbers \(\eta_i^u, 1 \leq i \leq N_u\) and \(\eta_j^v, 1 \leq j \leq N_v\) satisfying*
10. Verification of Chaos for Large Eccentricities

\begin{align*}
\eta_i^u &\geq \max\{c(u_{N_u})|\eta^u_{N_u}|, |c(v_{N_u})|\eta^v_{N_u} + |y^u|\}, \\
\eta_i^v &\geq \max\{c(u_{N_u})|\eta^u_{N_u} + |y^u|, |c(v_{N_u})|\eta^v_{N_u}\}, \\
\eta_{i+1}^u &\geq |c(u_i)|\eta^u_i, \quad 1 \leq i \leq N_u - 1, \\
\eta_{j+1}^v &\geq |c(v_j)|\eta^v_j, \quad 1 \leq j \leq N_v - 1.
\end{align*}

and assume that there are numbers \(\xi_i^u, 1 \leq i \leq N_u\) and \(\xi_j^v, 1 \leq j \leq N_v\) satisfying

\begin{align*}
\eta^u_i &> \max\{|\xi_i^u|, |\xi_i^v| + |x^u|\} + |b(u_{N_u})|\eta^u_{N_u}, \\
\eta^v_j &> \max\{|\xi_j^v|, |\xi_j^u| + |x^v|\} + |b(v_{N_u})|\eta^v_{N_u}, \\
\xi_i^u &> \frac{1}{|a(u_{N_u})|} (|\xi_i^v|, |\xi_i^u| + |x^u|) + |b(u_{N_u})|\eta^u_{N_u}), \\
\xi_j^v &> \frac{1}{|a(v_{N_u})|} (|\xi_j^u|, |\xi_j^v| + |x^v|) + |b(v_{N_u})|\eta^v_{N_u})
\end{align*}

Then the following estimate holds:

\[
\|\mathcal{R}^{-1}\mathcal{Q}^{-1}\mathbf{d}\| \leq \max_{1 \leq i \leq N_u, 1 \leq j \leq N_v} \left\{ \sqrt{(\xi_i^u)^2 + (\eta_i^u)^2}, \sqrt{(\xi_j^v)^2 + (\eta_j^v)^2} \right\}.
\]

**Proof:** From assumption (ii) in Lemma 10.1.1 it follows that \(\mathcal{Q}\) is invertible. The invertibility of \(\mathcal{R}\) was proven in Lemma 10.1.2. Thus \(\mathcal{R}^{-1}\mathcal{Q}^{-1}\) exists and it only remains to show that the estimate holds.

Let \(z_n = (x_n, y_n)\). Then \(\mathcal{Q}\mathcal{R}z = \mathbf{d}\) may be written as

\[
z_{n+1} - R_n x_n = Q_{n+1}^{-1}\mathbf{d}_n, \quad n \in \mathbb{Z}
\]

or

\[
x_{n+1} - a_n x_n - b_n y_n = x_n', \\
y_{n+1} - c_n y_n = y_n', \quad n \in \mathbb{Z}.
\]

(10.14)

where

\[
\begin{pmatrix}
\begin{cases}
(x_n^u \\
y_n^u)
\end{cases}
& \text{if } n = l_k \text{ for some } k \text{ and if the blocks change from } u \text{ to } v \\
\begin{cases}
(x_n^v \\
y_n^v)
\end{cases}
& \text{if } n = l_k \text{ for some } k \text{ and if the blocks change from } v \text{ to } u \\
\begin{cases}
0 \\
0
\end{cases}
& \text{else}
\end{pmatrix}
\]

We define the sequences \((\xi_n)_{n \in \mathbb{Z}}\) and \((\eta_n)_{n \in \mathbb{Z}}\) by

\[
\xi_n := \begin{cases}
\xi_i^u & \text{if } x_n = u_i \\
\xi_j^v & \text{if } x_n = v_j
\end{cases}, \quad \eta_n := \begin{cases}
\eta_i^u & \text{if } x_n = u_i \\
\eta_j^v & \text{if } x_n = v_j
\end{cases}.
\]
10.2. An Algorithm for the Verification of Chaos

On the one hand it follows from (10.14) that

\[ |x_n| \leq \frac{1}{|a_n|} (|x_{n+1}| + |b_n||y_n| + |x'_n|) \]

\[ |y_n| \leq |c_{n-1}||y_{n-1}| + |y'_n| . \]

On the other hand it follows from the assumptions of the lemma that

\[ \tilde{\xi}_n \leq \frac{1}{|a_n|} (\tilde{\xi}_{n+1} + |b_n|\tilde{\eta}_n + |x'_n|) \]

\[ \tilde{\eta}_n \leq |c_{n-1}|\tilde{\eta}_{n-1} + |y'_n| . \]

Taking the difference of the corresponding equations we obtain

\[ |x_n| - \tilde{\xi}_n \leq \frac{1}{|a_n|} (|x_{n+1}| - \tilde{\xi}_{n+1} + |b_n|(|y_n| - \tilde{\eta}_n)) \]

\[ |y_n| - \tilde{\eta}_n \leq |c_{n-1}|(|y_{n-1}| - \tilde{\eta}_{n-1}) \]

and further

\[ |x_{n+1}| - \tilde{\xi}_{n+1} \geq |a_n|(|x_n| - \tilde{\xi}_n) - |b_n|(|y_n| - \tilde{\eta}_n) \] \hspace{1cm} (10.15)

\[ |y_{n-1}| - \tilde{\eta}_{n-1} \geq \frac{1}{|c_{n-1}|}(|y_n| - \tilde{\eta}_n) . \] \hspace{1cm} (10.16)

Repeated use of (10.16) yields

\[ |y_{n-k}| - \tilde{\eta}_{n-k} \geq \frac{1}{\prod_{j=n-k}^{n-1}|c_j|}(|y_n| - \tilde{\eta}_n) . \] \hspace{1cm} (10.17)

If \(|y_n| > \tilde{\eta}_n\) was true for some \(n\) then inequality (10.8) would imply that the right-hand side of (10.17) increases unboundedly for \(k \to \infty\) in contradiction to the invertibility of \( \mathcal{R} \). Thus we may conclude that the sequence \((|y_n|)_{n \in \mathbb{Z}}\) is majorized by \((\tilde{\eta}_n)_{n \in \mathbb{Z}}\).

Using this result we may simplify (10.15) to

\[ |x_{n+1}| - \tilde{\xi}_{n+1} \geq |a_n|(|x_n| - \tilde{\xi}_n) . \]

Now a similar argument may be applied. Therefore we conclude that the sequence \((|x_n|)_{n \in \mathbb{Z}}\) is majorized by \((\tilde{\xi}_n)_{n \in \mathbb{Z}}\).

The estimate of \(\| \mathcal{Q}^{-1} \mathcal{R}^{-1} \mathcal{d} \|\) now follows immediately. \(\square\)

10.2 An Algorithm for the Verification of Chaos

In this section we present an algorithm for the verification of chaos in the dumbell satellite problem for large eccentricities. In the following section we apply this algorithm for the case \(e = 0.3\).
Periodic Orbits

We describe an efficient method to compute a pair of periodic orbits that are very close at some point.

**Step 1:** We randomly choose a point \( r_1 \) and compute a large number \( n \) of iterations of the Poincaré operator of (2.9). We denote this orbit by \( r := (r_k)_{1 \leq k \leq n} \).

Remark: Due to the expected chaotic behavior of the system the resulting orbit depends strongly on the computer system used for the computations.

**Step 2:** We determine the minimal distance between any two points of \( r \), i.e.

\[
d_{\text{min}} := \min_{1 \leq i < j \leq n} |r_i - r_j|.
\]

**Step 3:** We identify blocks of \( r \) that are periodic up to an error of \( k \cdot d_{\text{min}} \), where \( k \) is a small integer. Of special interest are pairs of blocks that come close to each other at some point. Let \( u' \) and \( v' \) denote such a pair of blocks with block length \( N'_{u} \) and \( N'_{v} \), respectively.

**Step 4:** We subdivide the interval \([0, 2\pi]\) into \( K \) subintervals of equal length as described in Section 10.1 and replace the Poincaré map \( P \) by the \( K \)-periodic sequence \( P = (P_k)_{1 \leq k \leq K} \) of maps \( P_k(z) := \phi(t_k + h_0; t_k, z) \), where \( \phi(t_k + h_0; t_k, z) \) denotes the solution of (2.9) with initial values \((t_k, z)\) and \( h_0 = 2\pi/K \).

Let \( u^0 \) and \( v^0 \) denote the orbits of \( P \) corresponding to \( u' \) and \( v' \). These orbits are of length \( N_u := K \cdot N'_{u} \) and \( N_v := K \cdot N'_{v} \), respectively. Using a high-dimensional version of the well-known algorithm of Newton we solve the equations

\[
P_1(u_1) = u_2, \quad P_1(v_1) = v_2,
\]
\[
P_2(u_2) = u_3, \quad P_1(v_1) = v_3,
\]
\[
\vdots
\]
\[
P_{N_u}(u_{N_u}) = u_1, \quad P_{N_v}(v_{N_v}) = v_{N_v},
\]

numerically starting with \( u = u^0 \) and \( v = v^0 \).

Finally the resulting blocks \( u \) and \( v \) are rearranged such that the smallest “jump” from one block to the other occurs at \( u_1 \) and \( v_1 \). The periodic sequences \( \tilde{u} \) and \( \tilde{v} \) obtained this way are orbits up to some small discretisation and roundoff error.

**Step 5:** Since the different lengths of the blocks \( u \) and \( v \) would lead to technical problems, we work in the sequel with blocks \( \hat{u} := uv \) and \( \hat{v} := vu \) rather than with \( u \) and \( v \). Let \( N \) denote the dimension of these blocks:

\[
N := N_u + N_v = \dim \hat{u} = \dim \hat{v}.
\]
10.2. An Algorithm for the Verification of Chaos 349

For further purposes let $M$ denote the index where the largest "jump" between $\tilde{u}$ and $\tilde{v}$ occurs:

$$
\rho_2 := |\tilde{u}_M - \tilde{v}_M| := \max_{1 \leq i \leq N_u + N_v} |\tilde{u}_i - \tilde{v}_i|.
$$

(10.19)

Random pseudo orbits

As in Chapter 9 let $\Sigma$ denote the set of all bi-infinite sequences $s$ of zeros and ones endowed with the metric

$$
d(s, t) := \max_{n \in \mathbb{Z}} \{2^{-|n|}\}.\]

For every sequence $s \in \Sigma$ we may construct a sequence in $\ell^\infty(\mathbb{Z}, \mathbb{R}^2)$ by concatenating blocks $\tilde{u}$ and $\tilde{v}$:

$$
\tau_{\text{pseudo}} : \Sigma \rightarrow S := \tau_{\text{pseudo}}(\Sigma) \subset \ell^\infty(\mathbb{Z}, \mathbb{R}^2),
$$

where

$$
tau_{\text{pseudo}} : \Sigma \rightarrow S := \tau_{\text{pseudo}}(\Sigma) \subset \ell^\infty(\mathbb{Z}, \mathbb{R}^2),
$$

where

$$
\begin{cases} 
\tilde{u} & \text{if } s_n = 0 \\
\tilde{v} & \text{if } s_n = 1.
\end{cases}
$$

The sequences $\tau_{\text{pseudo}}(s)$ are $\delta$-pseudo orbits of $P = (P_k)_{k \in \mathbb{Z}}$ for $\delta := |u_1 - v_1| + \delta_0$, where $\delta_0$ is an upper bound for the discretisation and round-off errors. This follows from

$$
|u_i - P_{i-1}(u_{i-1})| \leq \delta_0 \leq \delta \mod N_u,
$$

$$
|v_j - P_{j-1}(v_{j-1})| \leq \delta_0 \leq \delta \mod N_v,
$$

$$
|u_1 - P_{N_u}(u_{N_u})| \leq |u_1 - v_1| + |u_1 - P_{N_u}(u_{N_u})| \leq |u_1 - v_1| + \delta_0 = \delta,
$$

$$
|v_1 - P_{N_v}(v_{N_v})| \leq |v_1 - u_1| + |v_1 - P_{N_v}(v_{N_v})| \leq |v_1 - u_1| + \delta_0 = \delta.
$$

The Existence of Shadowing Orbits

We verify that for every pseudo orbit $q := \tau_{\text{pseudo}}(s)$, $s \in \Sigma$ there exists a unique shadowing orbit $p := \tau_{\text{shadow}}(q)$ close to $q$. To this end we compute the constants and verify the estimates introduced in Lemma 10.1.1–10.1.3.

The Computation of $\delta_0$, $\delta_1$

The solution of the differential equation (2.9) admits the following Taylor expansion:

$$
\phi(t_n + h; t_n, z) = z + \Phi_1 h + \ldots \Phi_{p-1} h^{p-1} + R^{(p)},
$$

(10.20)

where

$$
R^{(p)} := \int_0^1 (1 - s)^{p-1} \phi^{(p)}(t_n + sh; t_n, z) ds h^p.
$$
The coefficients

\[ \Phi_j = \Phi_j(t_n, z) := \frac{1}{j!} \frac{d^j}{dt^j} \phi(t_n; t_n, z) \]

may be computed recursively (cf. [28]).

An interval inclusion for the remainder \( R^{(p)} \) may be obtained as follows. Assume that we have an interval \( D_n \) for which \( \phi(t_n + h; t_n, z) \) Figure \( n \) \( \in D_n \) for all \( h \in [0, h_0] \) (cf. Figure 10.3).

![Fig. 10.3: An a priori estimate \( D_n \) for the solution \( \phi(t_n + h; t_n, z) \) for \( h \in [0, h_0] \).](image)

For the derivative in the remainder \( R^{(p)} \) the following formula holds:

\[ \phi^{(p)}(t_n + sh; t_n, z) = \phi^{(p)}(t_n + sh; t_n + sh, \phi(t_n + sh; t_n, z)) = \]

\[ = p! \Phi_p(t_n + sh, \phi(t_n + sh; t_n, z)). \]

This implies

\[ R^{(p)} \in \Phi_p([t_n, t_{n+1}], D_n)h^p. \]

Using a software package that allows interval arithmetics such as Mathematica\textsuperscript{®}, it is straightforward to compute the coefficients

\[ \tilde{\Phi}_j := \Phi_j([t_n, t_{n+1}], D_n). \]

From (10.20) we conclude that

\[ \phi(t_n + h; t_n, z) \in z_n + \Phi_1 h + \ldots + \Phi_{p-1}h^{p-1} + \tilde{\Phi}_p h^p. \]

Thus the interval

\[ z_n + \Phi_1 h_0 + \ldots + \Phi_{p-1}h_0^{p-1} + \tilde{\Phi}_p h_0^p \]

certainly contains the image \( P_n(z_n) = \phi(t_n + h_0; t_n, z_n) \).

To obtain the desired a priori estimate \( D_n \) let

\[ B_\varepsilon(x) := [x_1 - \varepsilon, x_1 + \varepsilon] \times [x_2 - \varepsilon, x_2 + \varepsilon] \]

and put

\[ D_n := z_n + \Phi_1[0, h_0] + \ldots + \Phi_{p-1}[0, h_0^{p-1}] + B_\varepsilon(0). \]
10.2. An Algorithm for the Verification of Chaos

It turns out that for \( \varepsilon = 0.0001 \) the inclusion

\[
\phi(t_n + h; t_n, z_n) \in E_n := z_n + \Phi_1[0, h_0] + \ldots + \Phi_{p-1}[0, h_0^{p-1}] + \Phi_p[0, h_0^p] \subset D_n
\]

holds for all \( h \in [0, h_0] \).

Corresponding validated intervals for the Jacobian are obtained by integrating the variational equation

\[
\dot{w} = D_z F(t, \phi(t; t_n, z_n))w.
\]

The Determination of \( \delta_2 \) and \( \delta_3 \)

The matrices \( A(u_i), 1 \leq i \leq N_u \) and \( A(v_j), 1 \leq j \leq N_v \) introduced in Section 10.1 may be triangularized in the following way:

\[
\begin{align*}
Q^{(i)}(u_1) = I & \quad \rightarrow \quad Q^{(i)}(u_{i+1})R^{(i)}(u_i) = A(u_i)Q^{(i)}(u_i) \quad \rightarrow \quad Q^{(i)}(u_{N_u})R^{(i)}(u_{N_u}) = A(u_{N_u})Q^{(i)}(u_{N_u}) \\
Q^{(j)}(v_1) = I & \quad \rightarrow \quad Q^{(j)}(v_{j+1})R^{(j)}(v_j) = A(v_j)Q^{(j)}(v_j) \quad \rightarrow \quad Q^{(j)}(v_{N_v})R^{(j)}(v_{N_v}) = A(v_{N_v})Q^{(j)}(v_{N_v})
\end{align*}
\]

Within machine precision the limit is reached after a few iterations.

From the computation scheme it follows that the expressions \( |A_nQ_n - Q_{n+1}R_n| \) and \( |Q_nQ_n - I| \) are of the order of the machine precision as long as we stay within one single block \( u \) or \( v \), respectively. At the jumps from \( u \) to \( v \), or vice versa, however, these values are expected to be much larger. We put

\[
\delta_2 := \max \left\{ \max_{1 \leq i \leq N_u} \{|A(u_i)Q(u_i) - Q(v_{i+1})R(u_i)|\}, \max_{1 \leq j \leq N_v} \{|A(v_j)Q(v_j) - Q(u_{j+1})R(v_j)|\}, |A(u_{N_u})Q(u_{N_u}) - Q(v_1)R(u_{N_u})|, |A(v_{N_v})Q(v_{N_v}) - Q(u_1)R(v_{N_v})| \right\}.
\]

For \( \delta_3 \) we put

\[
\delta_3 := \max_{1 \leq i \leq N_u, 1 \leq j \leq N_v} \{|Q(u_i)^TQ(u_i) - I|, |Q(v_j)^TQ(v_j) - I| \}.
\]

The Determination of \( ||R^{-1}|| \) and \( ||R^{-1}\tilde{Q}^{-1}d|| \)

The sequences \( \eta^u := (\eta_i)_{1 \leq i \leq N_u}, \eta^v := (\eta_j)_{1 \leq j \leq N_v} \) defined in Lemma 10.1.2 may be computed by a similar scheme as above.
Again it turns out that the sequences $(\eta^{(k)})_{k \in \mathbb{N}}$ and $(\eta^{(k)})_{k \in \mathbb{N}}$ converge quickly.

The sequences $\xi^u := (\xi^u_i)_{1 \leq i \leq N_u}$ and $\xi^v := (\xi^v_j)_{1 \leq j \leq N_v}$ may be computed analogously and Lemma 10.1.2 then yields an estimate for $||\mathcal{R}^{-1}||$.

The sequences $\tilde{\eta}^u := (\tilde{\eta}^u_i)_{1 \leq i \leq N_u}, \tilde{\eta}^v := (\tilde{\eta}^v_j)_{1 \leq j \leq N_v}$ and $\tilde{\xi}^u := (\tilde{\xi}^u_i)_{1 \leq i \leq N_u}, \tilde{\xi}^v := (\tilde{\xi}^v_j)_{1 \leq j \leq N_v}$ are determined in the same way. Lemma 10.1.3 yields an estimate for $||\mathcal{R}^{-1}\mathcal{Q}^{-1}d||$.

The Determination of $||\mathcal{L}^{-1}||$ and $||\mathcal{L}^{-1}d||$

If the estimates for $||\mathcal{R}^{-1}||$ and $||\mathcal{R}^{-1}\mathcal{Q}^{-1}d||$ are known, condition (iv) in Lemma 10.1.1 may be checked. If it holds true, one may compute the estimates for $||\mathcal{L}^{-1}||$ and $||\mathcal{L}^{-1}d||$.

The Computation of $M$

For $p = 1$ Eq. (10.20) may be written in the following form:

$$\phi(t_n + t; t_n, z) = z + \int_0^t F(t_n + s; \phi(t_n + s; t_n, z), e) ds. \quad (10.21)$$

The derivatives of (10.21) with respect to $z$ read

$$D_z \phi(t_n + t; t_n, z) = I + \int_0^t F_z(t_n + s; \phi(t_n + s; t_n, z), e) D_z \phi(t_n + s; t_n, z) ds \quad (10.22)$$

and

$$D_z^2 \phi(t_n + t; t_n, z) = \int_0^t F_{zz}(t_n + s; \phi(t_n + s; t_n, z, e) D_z \phi(t_n + s; t_n, z))^2 +$$

$$+ \int_0^t F_z(t_n + s; \phi(t_n + s; t_n, z), e) D_z^2 \phi(t_n + s; t_n, z) ds. \quad (10.23)$$

Let $m_1$ and $m_2$ be constants such that

$$\sup_{t \in [0,h_0], z \in D_n} \left| F_z(t_n + t, z, e) \right| \leq m_1,$$

$$\sup_{t \in [0,h_0], z \in D_n} \left| F_{zz}(t_n + t, z, e) \right| \leq m_2.$$
10.2. An Algorithm for the Verification of Chaos

We introduce the functions

\[ M_1(t, z) := |D_z \phi(t_n + t; t_n, z)|, \]
\[ M_2(t, z) := |D_z^2 \phi(t_n + t; t_n, z)|. \]

Taking the norms in (10.22) we obtain

\[ M_1(t, z) \leq 1 + m_1 \int_0^t M_1(s, z) ds \]
and Growall's Lemma implies

\[ M_1(t, z) \leq e^{m_1 t}. \quad (10.24) \]

Taking norms in (10.23) we find

\[ M_2(t, z) \leq m_2 \int_0^t M_1(s, z)^2 ds + m_1 \int_0^t M_2(s, z) ds \]
and with (10.24)

\[ M_2(t, z) \leq m_2 \frac{e^{2m_1 t} - 1}{2m_1} + m_1 \int_0^t M_2(s, z) ds. \]

Applying again a version of Gronwall's Lemma we obtain

\[ M_2(t, z) \leq m_2 \frac{e^{m_1 t} - 1}{m_1} e^{m_1 t}. \]

Since the right hand side is independent of \( z \) we obtain an estimate for \( M \) by choosing \( t = h_0 \):

\[ M \leq \frac{m_2}{m_1} (e^{m_1 h_0} - 1) e^{m_1 h_0}. \]

The Determination of \( \rho_0 \) and \( \rho_1 \)

Using the constants computed above one may finally compute the constants

\[ \rho_0 := \frac{2(l_d + \| L^{-1} \| \delta_0)}{1 - l \delta_1 + \sqrt{(1 - l \delta_1)^2 - 2l M(l_d + l \delta_0)}} \]
and

\[ \rho_1 := \frac{1 - l \delta_1 + \sqrt{(1 - l \delta_1)^2 - 2l M(l_d + l \delta_0)}}{l M} \]
(cf. Theorem 10.1), and the following holds:
For every pseudo orbit \( q \) consisting of blocks \( \hat{u} \) and \( \hat{v} \) there exists a shadowing orbit \( p \) that is \( \rho_0 \)-close to \( q \) and there is no other orbit staying \( \rho_1 \)-close to \( q \).

This allows us to construct a map between \( S := \tau_{\text{pseudo}}(\Sigma) \) and \( S^* := \tau_{\text{shadow}}(S) \subset \ell^\infty(\mathbb{Z}, \mathbb{R}^2) \):

\[
\tau_{\text{shadow}} : S \longrightarrow S^* := \tau_{\text{shadow}}(S) \subset \ell^\infty(\mathbb{Z}, \mathbb{R}^2),
\]

where

\[
\tau_{\text{shadow}}(q) := p, \quad \text{where } p \text{ is the (unique:) shadowing orbit associated with } q,
\]

according to the Shadowing Lemma.

### The Projector on \( \mathbb{R}^2 \)

Let the projector

\[
\pi : S^* \longrightarrow \mathbb{R}^2
\]

be defined by

\[
\pi(q) := \pi((\ldots, p_{-1}, p_0, p_1, \ldots)) := p_M,
\]

where \( M \), introduced in (10.19), is the index for which the largest "jump" between \( \hat{u} \) and \( \hat{v} \) occurs.

### Verification of Chaos in the Sense of Smale

In Definition 9.1.2 we defined a dynamical system to be chaotic in the sense of Smale, if it admits the Bernoulli shift system as a subsystem. The corresponding map \( \tau \) is obtained by composing the maps \( \tau_{\text{pseudo}}, \tau_{\text{shadow}} \) and \( \pi \):

\[
\tau : \Sigma \longrightarrow \mathbb{R}^2, \quad \tau := \pi \circ \tau_{\text{shadow}} \circ \tau_{\text{pseudo}}.
\]

The following diagram describes the construction of \( \tau \) in some detail:
10.2. An Algorithm for the Verification of Chaos

It remains to show that $\tau$ is a homeomorphism, indeed. Before we do this, we prove an important property of the shift space $\Sigma$.

**Lemma 10.2.1**

The shift space $\Sigma$ endowed with the metric

$$d(s, t) := \max \left\{ 2^{-|n|} \right\}$$

is a compact space.

**Proof:** We start the proof with an auxiliary result. Let $s, s' \in \Sigma$ and $n \in \mathbb{N}$, then

$$s_i = s'_i \text{ for } |i| \leq n \iff d(s, s') \leq 2^{-(n+1)}. \quad (10.25)$$

This result follows immediately from the definition of the norm and the fact that the sequences $s$ and $s'$ consist only of zeroes and ones.

As it is well-known, to prove a normed space to be compact, it suffices to show that every sequence of elements admits a limit point.

Let $S$ be such a sequence of elements $s' \in \Sigma$. Obviously there exists a subsequence $S_1 := (s_{1n})_{n \in \mathbb{N}}$ of $S_0 := S$ such that $d(s, s') \leq 2^{-1}$ for all elements $s$ and $s'$ of $S_1$.

Similarly there exists a subsequence $S_2 := (s_{2n})_{n \in \mathbb{N}}$ of $S_1$ such that $\|s - s'\| \leq 2^{-2}$ for all elements $s$ and $s'$ of $S_2$.

Repeating this process we obtain a sequence $(S_n)_{n \in \mathbb{N}}$. Choosing the $n$th element from $S_n$ we obtain the "diagonal" sequence $(s_0, s_1, s_2, \ldots)$. By construction we have

$$d(s^{(n)}, s^{(m)}) \leq 2^{-n} \text{ for } m \geq n.$$

(10.25) implies that $s_i^{(n)} = s_i^{(m)}$ for $|i| < n \leq m$. This allows us to define an element $s \in \Sigma$ by

$$s_i := s_i^{(n)} \text{ for } |i| < n.$$

Again by (10.25) we have

$$d(s, s^{(m)}) \leq 2^{-n} \text{ for } m \geq n.$$

Thus the "diagonal" sequence converges to $s$. This implies that $S$ has a limit point. \qed

We are now in the position to prove that $\tau$ is a homeomorphism.

**Theorem 10.2.1**

Consider the system of differential equations (2.9) for some fixed eccentricity $\epsilon$ and let $\phi(t; t_0, z_0)$ denote its solution with initial conditions $(t_0, z_0)$. Let $K$ be a positive integer and let $t_0, \ldots, t_K$ with $t_k := k \cdot 2\pi/K$ be a subdivision of the interval $[0, 2\pi]$. Assume that there are blocks $\mathbf{u}$ and $\mathbf{v}$ of length $N$ with the following properties:
(i) \((\ldots, \tilde{u}; \tilde{u}, \ldots)\) and \((\ldots, \tilde{v}; \tilde{v}, \ldots)\) are periodic pseudo orbits of \(P = (P_k)_{k \in \mathbb{Z}}\), \(P_k(z) := \phi(t_k + h_0; t_k, z)\) up to discretisations and round-off errors.

(ii) Sequences \(q\) consisting of the blocks \(\tilde{u}\) and \(\tilde{v}\) are \(\delta\)-pseudo orbits of \(P\).

(iii) For every pseudo orbit \(q\) consisting of blocks \(\tilde{u}\) and \(\tilde{v}\) there exists a \(\rho_0\)-shadowing orbit \(p\) and there is no other orbit staying \(\rho_1\)-close to \(q\).

(iv) The largest “jump” \(\rho_2\) between \(\tilde{u}\) and \(\tilde{v}\) occurs for the index \(M\) and fulfills the condition \(\rho_2 > 2\rho_0\).

Let

- \(T\) pseudo

\[ \Sigma \longrightarrow S := \tau_{\text{pseudo}}(\Sigma) \subset \ell^\infty(\mathbb{Z}, \mathbb{R}^2), \quad s \longmapsto \tau_{\text{pseudo}}(s) := (\ldots, b_{-1}; b_0, b_1, \ldots) \]

where

\[ b_n := \begin{cases} \tilde{u} & \text{if } s_n = 0 \\ \tilde{v} & \text{if } s_n = 1. \end{cases} \]

- \(T\) shadow

\[ S \longrightarrow S^* := \tau_{\text{shadow}}(S) \subset \ell^\infty(\mathbb{Z}, \mathbb{R}^2), \quad q \longmapsto \tau_{\text{shadow}}(q) := p, \]

where \(p\) is the (unique) shadowing orbit associated with \(q\), according to the Shadowing Lemma.

- \(\pi\)

\[ S^* \longrightarrow \mathbb{R}^2, \quad p \longmapsto \pi(p) := p_M, \]

where \(M\), introduced in (10.19), is the index for which the largest “jump” between \(\tilde{u}\) and \(\tilde{v}\) occurs.

Then the map

\[ \tau : \Sigma \longrightarrow \mathbb{R}^2, \quad \tau := \pi \circ \tau_{\text{shadow}} \circ \tau_{\text{pseudo}} \]

has the following properties:

(i) The following diagram commutes:

\[
\begin{array}{ccc}
\Sigma & \xrightarrow{\sigma} & \Sigma \\
\tau \downarrow & & \tau \downarrow \\
\tau(\Sigma) \subset \mathbb{R}^2 & \xrightarrow{p^N} & \tau(\Sigma) \subset \mathbb{R}^2
\end{array}
\]

(ii) \(\tau\) is a homeomorphism between \(\Sigma\) and \(\tau(\Sigma) \subset \mathbb{R}^2\).

(iii) \((\mathbb{R}^2, p^N)\) admits the Bernoulli shift system \((\Sigma, \sigma)\) as a subsystem and is therefore chaotic in the sense of Smale.
Proof:

(i) Let \( s \in \Sigma \) and \( s' := \sigma(s) \) and let the corresponding pseudo orbits be denoted by \( q \) and \( q' \) and the shadowing orbits by \( p \) and \( p' \). On the one hand we find

\[
P^N \circ \tau(s) = P^N \circ \pi \circ \tau_{\text{shadow}} \circ \tau_{\text{pseudo}} ((\ldots, s_{-1}; s_0, s_1, \ldots)) =
\]

\[
= : P^N \circ \pi \circ \tau_{\text{shadow}} ((\ldots, q_{-1}; q_0, q_1, \ldots)) =
\]

\[
= : P^N \circ \pi ((\ldots, p_{-1}; p_0, p_1, \ldots)) =
\]

\[
= : P^N(p_0^{(M)})
\]

and on the other hand

\[
\tau \circ \sigma(s) = \pi \circ \tau_{\text{shadow}} \circ \tau_{\text{pseudo}} \circ \sigma ((\ldots, s_{-1}; s_0, s_1, \ldots)) =
\]

\[
= \pi \circ \tau_{\text{shadow}} \circ \tau_{\text{pseudo}} ((\ldots, s_0, s_1, s_2, \ldots)) =
\]

\[
= \pi \circ \tau_{\text{shadow}} ((\ldots, q_0; q_1, q_2, \ldots)) =
\]

\[
= \pi ((\ldots, p_0; p_1, p_2, \ldots)) =
\]

\[
= p_1^{(M)}.
\]

Since \( P^N(p_0^{(M)}) = p_1^{(M)} \) holds the diagram commutes indeed.

(ii) We must show that \( \tau \) is bijective and bi-continuous. We do this in several steps.

- \( \tau \) is injective: Let \( s, s' \in \Sigma \) with \( s \neq s' \). It follows that there exists an integer \( n \) such that \( s_n \neq s'_n \). Due to the definition of \( \tau_{\text{pseudo}} \) the corresponding pseudo orbits \( q \) and \( q' \) differ in the \( n \)th block. (10.19) implies that there exists an integer \( j \) such that \( |q_j - q'_j| = \rho_2 \) cf.(10.19)). For the corresponding \( \rho_0 \)-shadowing orbits \( p \) and \( p' \) we find

\[
\rho_2 = |q_j - q'_j| \leq |q_j - p_j| + |p_j - p'_j| + |p'_j - q'_j| \leq
\]

\[
\leq \rho_0 + |p_j - p'_j| + \rho_0
\]

and therefore

\[
|p_j - p'_j| \geq \rho_2 - 2\rho_0 > 0.
\]

We conclude that \( p_j \neq p'_j \). Since \( P \) is invertible we further have \( p_M \neq p'_M \) and therefore \( \tau(s) \neq \tau(s') \). Thus \( \tau \) is one-to-one.

- \( \tau \) is surjective:

This is obvious.

- \( \tau \) is continuous:
Let $S := (s^n)_{n \in \mathbb{N}}$ be a sequence of elements in $\Sigma$ which converges to $s \in \Sigma$. Let $(q^n)_{n \in \mathbb{N}}$, $q$ denote the corresponding pseudo orbits and $(p^n)_{n \in \mathbb{N}}$, $p$ the corresponding shadowing orbits. We must show that $p^n_M$ converges to $p_M$ for $n \to \infty$.

Since the pseudo orbits $q^n$ consist of bounded blocks $\bar{u}$ and $\bar{v}$ they are bounded themselves. Then the shadowing property $|p^n_M - q^n_M| \leq \rho_0$ implies that the shadowing orbits $p^n_M$ are bounded too. Thus $(p^n_M)_{n \in \mathbb{N}}$ is a bounded sequence in $\mathbb{R}^2$ and therefore has a limit point.

We show that $q_M$ is the only limit point of $(q^n_M)_{n \in \mathbb{N}}$. To this end let $q_M$ be any limit point. Then there exists a subsequence $(q^n_M)_{n \in \mathbb{N}}$ with $q^n_M \to q_M$ for $n \to \infty$. Now consider $q^n_M + m = \hat{P}^m(q^n_M)$. Obviously, we have

$$|\hat{P}^m(q^n_M) - q^n_M| = |\hat{P}^m(q^n_M) - q^n_M - q^n_M + q^n_M|,$$

The first term on the right-hand side converges to 0, the second is bounded by $\rho_0$ because of the shadowing property and the third term is equal to 0 for $n$ sufficiently large. The last statement follows from (10.25) and the definition of $\tau_{\text{pseudo}}$. We conclude that $|\hat{P}^m(q^n_M) - q^n_M| \leq \rho_0$. Thus $(\hat{P}^m(q^n_M))_{n \in \mathbb{N}}$ is a $\rho_0$-shadowing orbit of $p$. Since $q$ admits the unique $\rho_0$-shadowing orbit $q$ it follows that $p_M = p_M$. This proves that $\tau$ is continuous.

• Since $\tau$ is a bijective and continuous map from the compact space $\Sigma$ to the Hausdorff space $\tau(\Sigma)$, $\tau$ is even a homeomorphism (cf. [10]).

(iii) The existence of chaotic behavior follows immediately from (i) and (ii). 

10.3 Verification of Chaos for $e = 0.3$

In this section we apply the algorithm to the dumbbell satellite problem for eccentricity $e = 0.3$.

Periodic Orbits

For our example we found blocks $u'$ and $v'$ of length $N'_u = 30$ and $N'_v = 21$ that suit our purposes.

In Figure 10.4 we show the corresponding orbits of the Poincaré map $P$.

For the further computations we subdivide the interval $[0, 2\pi]$ into $K = 60$ subintervals and compute the blocks $u$ and $v$ of length $N_u = 1800$ and $N_v = 1260$.

The smallest "jump" from one block to the other is found to be

$$\delta := \min_{1 \leq i \leq N_u + N_v} |\hat{u}_i - \hat{v}_i| = 7.8167 \ldots \cdot 10^{-9}.$$

For the largest "jump" between $\hat{u}$ and $\hat{v}$ we find

$$\rho_2 := |\hat{u}_{1115} - \hat{v}_{1115}| := \max_{1 \leq i \leq N_u + N_v} |\hat{u}_i - \hat{v}_i| = 5.4327 \ldots.$$
Fig. 10.4: Two periodic orbits of the Poincaré map corresponding to the blocks \( u' \) and \( v' \).

### The Existence of Shadowing Orbits

For the computation of the constants \( \delta_0 \) and \( \delta_1 \) we choose the order of the Taylor expansion (10.20) to be 17, i.e. we put \( p = 18 \). This leads to

\[
\delta_0 = 5.44161 \ldots \times 10^{-10}, \\
\delta_1 = 6.4271 \ldots \times 10^{-9}.
\]

For the constants \( \delta_2 \) and \( \delta_3 \) we obtain

\[
\delta_2 := 3.08697 \ldots \times 10^{-8}, \\
\delta_3 := 7.61383 \ldots \times 10^{-16}.
\]

For \( \| R^{-1} \| \) and \( \| R^{-1} Q^{-1} d \| \) we find

\[
\| R^{-1} \| = 7344.78 \ldots , \\
\| R^{-1} Q^{-1} d \| = 3.76817 \ldots \times 10^{-8}.
\]

Since

\[
\| R^{-1} \| \delta_2 = 0.000226731 \ldots < \sqrt{1 - \delta_3} = 0.99999 \ldots
\]

the assumptions of Lemma 10.1.1 are fulfilled and we find

\[
\| L^{-1} \| \leq 7346.44 \ldots , \\
\| L^{-1} d \| \leq 3.76903 \ldots \times 10^{-8}.
\]
To estimate the norm of $F_z(t, z, e)$ we first compute

$$F_z(t, z, e) = \begin{pmatrix} 0 & 1 \\ 3 \cos(2x) & e \sin(t) \\ 1 - e \cos(t) & 1 - e \cos(t) \end{pmatrix}$$

and then

$$F_z(t, z, e)^T F_z(t, z, e) = \begin{pmatrix} \frac{9 \cos^2(2x)}{(1 - e \cos(t))^2} & -\frac{3e \cos(2x) \sin(t)}{(1 - e \cos(t))^2} \\ \frac{3e \cos(2x) \sin(t)}{(1 - e \cos(t))^2} & \frac{e^2 \sin^2(t)}{(1 - e \cos(t))^2} \end{pmatrix}.$$ 

The characteristic equation of $F_z(t, z, e)^T F_z(t, z, e)$ reads

$$p(\lambda) := \det \left( F_z(t, z, e)^T F_z(t, z, e) - \lambda I_2 \right) =$$

$$= \lambda^2 - \lambda \left( 1 + \frac{9 \cos^2(2x)}{(1 - e \cos(t))^2} + \frac{e^2 \sin^2(t)}{(1 - e \cos(t))^2} \right) + \frac{9 \cos^2(2x)}{(1 - e \cos(t))^2} =$$

$$= \lambda^2 - \lambda \left( 1 - e \cos(t) \right)^2 + \frac{9 \cos^2(2x) + e^2 \sin^2(t)}{(1 - e \cos(t))^2} + \frac{9 \cos^2(2x)}{(1 - e \cos(t))^2}.$$ 

This leads to the following eigenvalues:

$$\lambda_{1,2} = \frac{(1 - e \cos t)^2 + 9 \cos^2(2x) + e^2 \sin^2 t}{2(1 - e \cos t)^2} \pm$$

$$\pm \sqrt{\left( (1 - e \cos t)^2 + 9 \cos^2(2x) + e^2 \sin^2 t - 36 \cos^2(2x) (1 - e \cos t)^2 \right)} \frac{2(1 - e \cos t)^2}{2(1 - e \cos t)^2}.$$ 

A tedious computation shows that $\lambda_{1,2}$ may be rewritten as

$$\frac{1}{2} + \frac{9 \cos^2(2x) + e^2 \sin^2 t}{2(1 - e \cos t)^2} \pm \sqrt{\left( 9 \cos^2(2x) - 1 - e^2 + 2e \cos t \right)^2 + 36e^2 \cos^2(2x) \sin^2 t} \frac{2(1 - e \cos t)^2}{2(1 - e \cos t)^2}.$$ 

Now we replace $\cos(2x)$, $\sin t$ and $\cos t$ by 1 and put $e = 0.3$. This leads the following estimate:

$$\lambda_{1,2} < \lambda_* := \frac{1}{2} + \frac{9 + 0.3^2}{2(1 - 0.3)^2} + \sqrt{\frac{(9 - 1 - 0.3^2 + 2 \cdot 0.3)^2 + 36 \cdot 0.3^2}{2(1 - 0.3)^2}} =$$

$$= 18.65130 \ldots.$$ 

Thus we obtain the following estimate for $|F_z(t, z, e)|$

$$\sup_{t \in [0, h_0], z \in \mathbb{R}^n} |F_z(t_n + t, z, e)| \leq \sup_{t \in [0, h_0], z \in \mathbb{R}^2} |F_z(t_n + t, z, e)| \leq \sqrt{\lambda_*} = 4.3083 \ldots =: m_1.$$ 

The norm of $F_z(t, z, e)$ is estimated more easily since $F_z(t, z, e)$ has only one non-vanishing term

$$\frac{6 \sin(2x)}{1 - e \cos t}.$$
10.3. Verification of Chaos for $e = 0.3$

Thus we find

$$\sup_{t \in [0, b_0], z \in D_n} |F_z(t_n + t, z, e)| \leq \sup_{t \in \mathbb{R}, z \in \mathbb{R}^2} |F_z(t_n + t, z, e)| \leq \frac{6}{1 - 0.3} = 8.5714 \ldots =: m_2.$$ 

This leads to the following estimate for the constant $M$:

$$\frac{m_2}{m_1} (e^{m_1 b_0} - 1) e^{m_1 b_0} = 1.7839 \ldots =: M.$$ 

Using all the constants computed above, we finally obtain the following values $\rho_0$ and $\rho_1$ (cf. Theorem 10.1.2):

$$\rho_0 = 7.777063 \ldots \cdot 10^{-8}, \quad \rho_1 = 1.5278 \ldots \cdot 10^{-3}.$$ 

We may summarize the computations in the following theorem.

**Theorem 10.3.1**

Consider the dumbbell satellite problem (2.9) for $e = 0.3$. For every pseudo orbit $q$ consisting of blocks $\bar{u}$ and $\bar{v}$ there exists a shadowing orbit $0.000000078$-close to $q$ and there is no other orbit staying $0.0015$-close to $q$.

The Verification of Chaotic Behavior

Since

$$\rho_2 - 2\rho_0 = 5.4327 \ldots - 2 \cdot 7.8167 \ldots \cdot 10^{-9} > 0,$$

the conditions of Theorem 10.2.1 are fulfilled. Thus the following theorem holds.

**Theorem 10.3.2**

Consider the system of differential equations

$$\dot{z} = F(t, z, e),$$

with

$$F(t, z, e) := F(t, x, y, e) = \left( \begin{array}{c} y \\ \frac{\sin(2x)}{2 (1 - e \cos t)} + y \frac{\sin t}{1 - e \cos t} + \frac{2e \sqrt{1 - e^2} \sin t}{(1 - e \cos t)^2} \end{array} \right).$$

for eccentricity $e = 0.3$ and let $P$ denote the Poincaré map of (2.9).

Then there exists a positive integer $N$ such that $(\mathbb{R}^2, P^N)$ admits the Bernoulli shift system $(\Sigma, \sigma)$ as a subsystem and is chaotic in the sense of Smale.
An Illustration of the Chaotic Behavior

For the periodic solutions associated with the blocks \( \hat{u} \) and \( \hat{v} \) we plot in Figure 10.5 the first component, i.e. the angle \( x \), against the eccentric anomaly \( t \) of the Keplerian motion.

Assume now that the satellite points to a certain fixed star at the beginning of block \( \hat{u} \) or \( \hat{v} \), respectively. At the end of the blocks, after \( N := N_u + N_v = 51 \) (cf. (10.18)) revolutions of the satellite around the earth, the dumbbell will point at the same direction, due to the periodicity of the orbits. But after 21 revolutions it will point at almost opposite directions depending on the choice of the orbits:

\[
\frac{x(21 \cdot 2\pi) - x(0)}{2\pi} = 21.05 \quad \text{for the orbit } \hat{u}
\]

and

\[
\frac{x(21 \cdot 2\pi) - x(0)}{2\pi} = 20.5 \quad \text{for the orbit } \hat{u}
\]

This situation is illustrated in Figure 10.6.
The succession of these almost opposite directions is completely unpredictable, since there exists a shadowing orbit for every pseudo orbit consisting of blocks $\hat{u}$ and $\hat{v}$.

10.4 Computational Notes

We close this chapter with some remarks on the use of the computer:

- All computations are performed in *Mathematica*® 3.0 or 4.0 running on Macintosh® G3 computers with 300 GHz clock rate.
- Almost all computations are performed with machine precision, i.e. with 16 decimal digits. Critical parts are performed with 26 decimal digits.
- The interval arithmetics was performed using the built-in function *Interval*. As a test we recomputed the constants with MATHLAB® on a Spark®-workstation and obtained slightly different results. These differences are explained by the different implementation of the interval arithmetics as well as the slightly different implementation of the Taylor-algorithm for solving the differential equations.
- Given the blocks $u$ and $v$ the total computation time is about 15 minutes (300MHz).
10. Verification of Chaos for Large Eccentricities
Bibliography


Curriculum Vitae

Geboren 10. August 1964

Eltern

Ausbildung
1970 – 1971 Kindergarten Frauenfeld
1971 – 1976 Primarschule Frauenfeld
1977 – 1983 Thurgauische Kantonsschule Frauenfeld, Typus C
1984 – 1989 ETH Zürich, Abteilung IX Mathematik
1992 Fähigkeitsausweis für das höhere Lehramt

Beruf
1989 – 1997 Assistent für Mathematik und Didaktik, ETH Zürich
1990 – 1997 Lehrbeauftrager für Mathematik und Informatik,
Kantonsschule im Lee, Winterthur
ab 1997 Hauptlehrer für Mathematik und Informatik,
Kantonsschule Schaffhausen.