Technische Universität München
Fakultät für Mathematik

Control of planar pendulum systems

Martin Lehl

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Martin Brokate
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Jürgen Scheurle
	2.	Univ.-Prof. Dr. Claus Hillermeier, Universität der Bundeswehr München
	3.	Prof. Dr. Gregorio Díaz (schriftliche Beurteilung), Universidad Complutense de Madrid, Spanien

Die Dissertation wurde am 31.10.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 25.02.2013 angenommen.

Contents

List of Figures

Acknowledgements

I gratefully thank Prof. Scheurle for his guidance and understanding during my studies.
Also, I want to thank COTESYS[1] for supporting parts of this work.
Finally, I thank my parents Hans-Werner and Waltraud Lehl.

[1]COTESYS=**Co**gnition for **Te**chnical **Sys**tems, Cluster of Excellence at TU München

Chapter 1

Introduction

Control theory deals with dynamical systems which can be influenced by *controls*. Such a system is called *control system*. Typical questions are: Given an initial state x_0 for a control system, is there a suitable control such that a predefined final state x_f can be reached? Can this state be reached in finite or infinite time? Which states can be reached within a certain time. These types of questions belong to a part of control theory, which is called *controllability* theory. We will introduce to this topic in chapter 3.

Another important part of control theory is the *stabilization* problem. Suppose the uncontrolled system has an unstable equilibrium point. Stabilizing the control system around this equilibrium point means that we are looking for a suitable control such that this state becomes a stable equilibrium point of the controlled system. Balancing a ball on one's head is an example for a stabilization problem. We present some ideas of this topic in chapter 4. The control law we present in this work as our main original contribution belongs to this part of control theory.

Another area of research is *optimal control theory*. Suppose for example that for a given initial and final state the question of controllability has a positive answer - meaning there is a suitable control such that the solution of the controlled system starting in x_0 reaches x_f in finite time. Among all controls which are admissible and suitable to perform this task, optimal control theory is searching for those controls, which minimize (or maximize) a certain cost functional such as "energy consumption" or "time needed" until the final state is reached. Our control law is based on a minimum energy control law for linear time-varying systems [Cheng, 1979] which was extended to nonlinear systems in [Sastry et al., 1994]. We will give a brief overview about historical facts in optimal control theory for linear systems in chapter 2.

Throughout this work t will represent the time variable. We only regard finite-dimensional state spaces and restrict ourselves to $\mathbb{R}^n$ for some integer $n > 0$. The state vector will be denoted by x. The variable u is reserved for the vector-valued control, which – unless stated otherwise – will be taken from a subspace $U \subseteq \mathbb{R}^r$, $0 \leq r \leq n$, which will usually be bounded and having 0 as an inner point. Instead of simply "control" we will equally use the terms "control function", "control law" and "control input" or "input function".

Control systems we regard can be described as vector-differential equation of the form

$$\dot{x}(t) = f(x(t), u(t), t), \quad t \geq t_0. \tag{1.1}$$

The right hand side of (1.1) has to fit certain regularity assumptions. We will assume regularity for f such that there is no blow-up of the solution in finite time and that for every initial value x_0 at t_0 and every admissible control input $u(t) \in U$ there is a unique solution $x(t, u(t); x_0)$ for system (1.1) and $t \geq t_0$. For simplicity we will often denote the solution as $x(t)$ hiding the fact that it also depends on the initial value and the control input.

In many cases the control input $u(t)$ depends only on the state $x(t)$ of the system and thus we have

$$u = u(x(t)). \tag{1.2}$$

Control systems using such control inputs are called *closed-loop* or *feedback-control* systems. The function f on the right hand side of (1.1) can be linear with respect to the state and the control. In this case we obtain the control system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad t \geq t_0, \tag{1.3}$$

where $A(t) \in \mathbb{R}^{(n,n)}$ and $B(t) \in \mathbb{R}^{(n,r)}$ are matrix-valued functions which we assume to have at least locally integrable elements. If $A(t) \equiv A$ and $B(t) \equiv B$ are constant matrices, the control system is called *linear autonomous* control system or *linear constant* control system.
Control systems which are not linear are called *nonlinear* control systems. Linear control systems arise for example as linearization of nonlinear control systems. When linearizing around a point $(x_0, u_0) \in \mathbb{R}^n \times \mathbb{R}^r$ we obtain a linear constant control system. Linearizing along a curve of the state space (not necessarily a trajectory of the system) we obtain in general a linear time-varying control system. Linear control systems are often helpful in describing the local behaviour of a nonlinear system (see section 3.1).

About this work

In this work we present a novel nonlinear feedback control law for planar pendulum systems with any number of links which locally stabilizes the inverted pendulum position (i.e. all pendulum links point upward). It is also possible to locally stabilize those systems along trajectories of the uncontrolled pendulum, when the linearization along these trajectories is controllable. Numerical simulations indicate that the presented control law can actually be used to swing the pendulum up to the inverted position from rather far away. For example in cases of up to three links, we successfully managed to swing it up not only from the stable equilibrium point but also using initial conditions where the pendulum links had very high velocities.

This work is organized in the following way:

In **chapter 2** we study the time-optimal control problem for the linear simple pendulum equation as introductory example and historical review:

- First we outline a result by D. Bushaw [Bushaw, 1958] which dates back to the beginning of control theory. It is one of the first papers providing a precise mathematical description of a class of control problems and its solution. The main idea is to study the geometry of solution curves in the phase space belonging to a certain control input and finding an equivalent geometrical formulation for the problem. This work belongs to the part of optimal control theory as the goal is to find time-optimal solutions.
  As the input space for the control inputs can be very rich, the set of solution curves may also be very large. Bushaw avoids this problem by assuming the input space to be discrete and finite. Although this sounds like a serious restriction, in most cases it is not even a greater restriction than assuming the control inputs to be bounded. The reason for this is the so-called *bang-bang principle* (see appendix A) which roughly speaking says that everything that can be done with bounded controls can also be achieved by controls

assuming limiting values. Bushaw could not refer to this result as it was unknown at that time.

- In the second result the input space is bounded and no longer discrete and finite as in the result of D. Bushaw. We present the theory developed by Hermes and Lasalle [Hermes and LaSalle, 1969] rather detailed as it gives a good survey of linear system theory which will be needed later in this work. Applying their theory to the linear simple pendulum equation we recover the solution of D. Bushaw, although the input spaces will be chosen differently!

- A third result concerning the linear case is presented in appendix D and is based on Pontrjagin's maximum principle. The maximum principle is an important result in optimal control theory as it can be formulated in a very general setting for nonlinear systems and for different cost functionals (a precise formulation is echoed in appendix D). Applied to the linear model of a simple pendulum we obtain the same control as we already obtained before.

Bushaw's method is very restrictive in the sense that it is only applicable for a very small class of systems with low dimension and allowing only controls assuming discrete values. Its strength lies in the fact that it provides a constructive method for finding an optimal control.
The method of Hermes and Lasalle is more general as it solves the time-optimal control problem for linear time-varying systems of arbitrary, but finite dimension and in addition it admits bounded controls. The disadvantage of this method is that the solution contains the state transition matrix of the uncontrolled linear system and – unless the system is autonomous – can in general not be given explicitly. For practical purposes a numerical scheme is necessary to overcome this difficulty. We used an algorithm proposed in [Eaton, 1962] to show that this method actually works and demonstrated that in terms of the simple pendulum where the state transition matrix can be given explicitly. Thus we are able to compare our numerical results with the analytical solution.
The method of Pontrjagin is the most general of the three, as it is designed for optimal control problems, not necessarily time-optimal control problems. In particular it is applicable for nonlinear systems. The disadvantage is that it only provides a necessary condition for the optimal solution and it does not guarantee existence of a solution. Due to its simplicity, for our introductory example existence of a solution is not a hard problem but still something that has to be proved.

Chapter 3 is dedicated to the topic of controllability. For linear systems, time-varying or not, the question of controllability has been completely answered (cf. e.g. [Klamka, 1991] which is a monograph dealing exclusively with controllability of linear systems). The main theorem for linear controllability is theorem 3.9 in the present work, where the so called *controllability Gramian* is introduced. The controllability Gramian is a matrix which is invertible if and only if the linear system is controllable. For several reasons this theorem plays an extraordinary role not only in the theory of linear control systems. To mention some of them:
– The criterion provides a *necessary and sufficient* condition for global controllability.
– Allowing *unbounded* controls every state can be transferred to any arbitrary state in *any given finite time*.
– The proof of the theorem is *constructive* and uses the Gramian to solve the state transition problem.
– The Gramian appears in the solution of many *nonlinear* control problems as it does in our

control law.

Remark: In contrary to the Kalman criterion, which works for finite dimensions of the state space only, the method using the controllability Gramian can be generalized to control problems with *infinite* dimensional state spaces, but infinite dimensional state spaces will not be regarded in this work.

Theorem 3.9 gives an integral criterion for systems of the form (1.3), which on the one hand does not require much regularity of the system matrices $A(t)$ and $B(t)$ but on the other hand brings along the integration problem which in general has to be solved numerically. In case the matrices $A(t)$ and $B(t)$ are sufficiently often differentiable, there is a criterion for controllability which only makes use of derivatives (theorem 3.10). It should be mentioned that the regularity assumptions depend on the dimension of the matrix $A(t)$ and therefore these regularity assumptions are strong for large dimensions of the underlying linear system (if the dimension of $A(t)$ is n the matrices $A(t)$ and $B(t)$ have to be $n-1$ times differentiable).
This theorem is important for another reason. It is a generalization of the Kalman criterion to time-varying systems.
Although autonomous linear systems are a special case of linear time-varying systems and therefore the whole "time-varying theory" can be applied, several methods deserve to be mentioned on their own.
Autonomous linear systems are much more simple to handle than time-varying linear systems (constant rank, computable state transition matrix, ...). The first results in the field of controllability were obtained for this class of systems. The Kalman rank condition which was published in [Kalman, 1960] is an important result to be mentioned (here theorem 3.11). It reduces the question of controllability to some elementary matrix computations involving the system matrices A and B. Unfortunately this method is not constructive, but – if applicable – easier to handle than the integral criterion of theorem 3.9.
Following the presentation in [Klamka, 1991] we add another important method for checking controllability of linear autonomous systems. It is based on the transformation to Jordan-form. Controllability does not depend on this full rank linear transformation. This method has the advantage that it reveals the coupling of the state variables. It provides a lower bound for the size of the input vector and therefore gives conditions for the minimum number of actuators necessary to guarantee controllability.

For general nonlinear systems it is still hard to prove controllability. The huge variety of possible nonlinearities makes it difficult to find a general method in order to answer the question of controllability.
There have been several attempts to tackle the problem of controllability for nonlinear systems. In this work we will only restate some results using Lie-algebra techniques as these ones seem to be a good choice in investigating controllability for an important class of nonlinear control systems which are the so-called control affine systems (see 3.2.3). Without going into detail, in sloppy notation they are of the form:

$$\dot{x}(t) = f(x(t)) + \sum_{i \in I} g_i(x(t))u_i(t)$$

where I is a finite index set whose cardinality typically does not exceed the dimension of the state space. This class of nonlinear control systems covers many control systems motivated by real-life applications. The system dynamics enters as "drift term" f, whereas the actuated parts correspond to the terms $g_i(x(t)) \cdot u_i(t)$ where $u_i(t)$ denote the control inputs. The systems we are interested in – planar pendulum systems – can be written in this form.

The main observation for these systems is that one can not only steer in the direction of the vector fields g_i but also in directions spanned by the Lie-brackets of these vector fields and – under additional assumptions – the vector field f. A major problem is that in general one can only go forward in the direction of f and has to make more and strong assumptions on f to actually go "backward". It is unsatisfactory that even for the case where only one control input enters in the control affine problem the question of controllability for the nonlinear system can only be answered by very restrictive assumptions on the drift term f. We give a brief summary of the existing theory.

Knowing much about linear systems and few about nonlinear ones it seems natural to look at the linearization of the nonlinear system and trying to make conclusions for the original nonlinear systems. This standard procedure will – if at all – at first deliver local results, but – depending on the "strength" of the nonlinearity, the domain of validity can still be very large.

An important result is due to Lee and Markus [Lee and Markus, 1967]. It states that in case the linearization around an equilibrium point is controllable, the nonlinear system is controllable in a neighborhood around the equilibrium point which is determined by the inverse function theorem.

If the linearization is not controllable sometimes local controllability can still be proved by the Lie algebraic methods mentioned above. This important result of Lee and Markus can be generalized to any trajectory (cf. e.g. [Coron, 2007]) and we will apply this theorem to planar pendulum systems.

From the same book we state theorem 3.39 which can be viewed as a generalization of the Kalman criterion for nonlinear systems in the sense that it simplifies to the generalized Kalman criterion already mentioned if applied to a linear time-varying system (note that here additional regularity assumptions on the system matrices are required) and to the Kalman condition if applied to a linear autonomous system (which automatically fits the mentioned regularity assumptions). This relationship is important as theorem 3.39 is about controllability in directions of Lie-Brackets and therefore is a beautiful link between the theorems stated in our work.

Chapter 4 deals with the stabilization problem. Given a reference trajectory (which might be a single point of the state space) we are interested in stabilizing the system along this trajectory. We start by giving all the necessary definitions along with the most important stability theorems based on Lyapunov's second method. Here we mainly follow the presentation in [Sastry, 1999].

For autonomous linear system it is well known that stability properties can be characterized by the location of the eigenvalues belonging to the system matrix. Unfortunately there is no generalization to time-varying linear systems. For those and nonlinear systems Lyapunov theory will be used here.

The difficulty in applying Lyapunov theory lies in the fact that a suitable Lyapunov function (see 4.8) has to be found. For linear time-varying systems there is a standard procedure leading to a Riccati differential equation, which - for autonomous linear systems reduces to the *Lyapunov equation* – a system of linear algebraic equations – where the problem is reduced to finding a suitable "right hand side" for this equation.

For nonlinear systems finding such a Lyapunov function is much more difficult. At least there is a theorem (see e.g. [Poznjak, 2008, Zubov 1964]) providing a necessary condition by stating that a Lyapunov function for nonlinear systems exists, if the system can be stabilized.

In this chapter we present 4 results from literature which provide stabilizing control laws for linear autonomous, linear time-varying and nonlinear systems. These theorems are strongly connected as the law for nonlinear systems can be seen as direct generalization of the stabilization law for time-varying systems which in turn can be seen as generalization of the linear autonomous systems. These theorems are the basis for the novel control law we present at the end of this chapter. Next we give a brief overview of the presented stabilization methods:

– For linear autonomous systems Kleinmann [Kleinmann, 1970] found a way to (asymptotically) stabilize the zero-solution of the uncontrolled systems by what he called "an easy way". This method has several advantages:

- The method uses the Gramian computed over a finite time interval. The controllability question can be answered by checking the rank of this matrix.
- No transformation of variables is needed.
- No eigenvalues have to be determined.

The disadvantage of this method is that one cannot prescribe the rate of convergence a priori and therefore convergence might be extremely slow.

– To overcome this disadvantage we present another method for linear autonomous systems which goes back to R. W. Bass. Bass never published his result but it is contained in some lecture notes for a course he taught at the NASA Langley Research Center in August 1961. A summary can be found in [Russell, 1979] for example. It should be mentioned that the result of Bass is earlier than that of Kleinmann.

– R. W. Bass used a modified controllability Gramian allowing to adjust the rate of convergence in the sense that a minimal rate of convergence is guaranteed. This "convergence factor" appears in the generalization of Kleinmann's method to linear time-varying systems and later for nonlinear systems. Therefore this idea is an important step and thus we decided to mention it. Given a minimal rate of convergence by the method of Bass, the actual performance of the resulting control law turns out to be much better when applied to the planar pendulum systems with up to three links.

– 1979 V. Cheng generalized Kleinmann's method to time-varying linear systems (see [Cheng, 1979]). Using the idea of Bass a minimum rate of convergence can be established. Cheng proves that his method uniformly exponentially stabilizes a linear time-varying system at a prescribed rate.
One of the disadvantages is that at every instant of time the controllability Gramian has not only to be computed but also to be inverted. As in general it is not a sparse matrix this inversion is costly for large dimensions.
One crucial thing to mention is that the resulting control law works globally to stabilize the system to the zero-solution. This is a very strong result.

– Fifteen years later a group around S. Sastry used Cheng's method to stabilize nonlinear systems. They regarded the problem of stabilizing the nonlinear system along a trajectory. The linearization along this reference trajectory has to be controllable.
In principle Sastry et al. showed that the control law suggested by V. Cheng for linear time-varying systems can also be used for nonlinear systems where as time-varying system the linearization along the reference trajectory is used. As a result they obtain a local stabilizing control law for the nonlinear system. The proof relies on Taylor series expansion. Convergence of the

trajectory of the controlled system to the reference trajectory is proved almost literally as in Cheng's paper but one has to take into account the higher order terms from the Taylor series expansion.

– In the last part of chapter 4 we present a modification of the control law of Sastry et al.; it is designed to stabilize a nonlinear system along a feasible reference trajectory, i.e. a trajectory with controllable linearization.
The main idea is to avoid the already mentioned matrix inversion of the controllability Gramian at every instant of time and therefore during every numerical time step. Since the controllability Gramian depends continuously on the system matrix resulting from the linearization along the reference trajectory it is possible to keep it unchanged over some time interval. Numerical experiments with planar pendulum systems with up to three links indicate that these time intervals can actually be chosen pretty large, even if the underlying system is very sensitive as for example in case of the triple pendulum. In our numerical examples we chose intervals of length up to 2 time units which means, if the numerical time step is assumed to be 0.01 that we only have to invert the Gramian matrix in 0.5% of the numerical time steps. In particular if the reference trajectory reduces to a single point of the state space this method performs pretty well due to the fact that the linearization around this point results in a linear autonomous system. We conjecture, but unfortunately did not manage to proof that for reasonable systems the basin of attraction for stabilizing an equilibrium position is the whole state space. In fact Cheng's global stabilization result for the zero-solution of linear time-varying systems suggests that this global result remains true for the modified control law although the convergence rate will decrease. In our numerical simulations for planar pendulum systems we could not find any initial conditions which failed to converge to the inverted pendulum position.
In general, stabilizing the inverted pendulum position is performed in two parts. First, a control input is designed which swings the pendulum up. Second, linearization theory is used to balance the inverted pendulum. This procedure results in two different control laws and switching between them is needed at an appropriate moment. Although the modified control we are going to propose here seems to be suitable to swing up and stabilize the inverted pendulum position without switching, we suggest to use the dynamics of the uncontrolled pendulum system as much as possible in order to bring the pendulum closer to the desired equilibrium state as this is energetically more efficient and reduces the amount of input energy significantly as we show by some examples concerning the double and triple pendulum.
As a conclusion we derive a control law which considerably reduces the computational effort compared to the control law proposed in [Sastry et al., 1994] . At least for planar pendulum systems with up to three links this control law appears to be capable of swinging up and balancing the pendulum at its inverted position from any starting configuration.

The last chapter is devoted to the application of those methods to planar pendulum systems. First we derive the equations of motion for planar pendulum systems up to the triple pendulum. After a short discussion we apply different stabilization methods for the derived systems and compare the results. Finally we apply our modified control law for different scenarios including the use of the natural dynamics as described above. We compare the results by evaluating a suitable cost functional which penalizes both the total amount of input energy as well as large input values.
The obtained models for the planar pendulum systems have been implemented in MATLAB. Simulations include the dynamics of the uncontrolled pendulum and the controlled pendulum. Results were saved as video files. Each video frame shows three figures:

– the first figure shows the solution of the controlled system;
– the second figure shows the solution of the reference trajectory, which is a solution of the uncontrolled system;
– the third figure shows both solutions.
To present the results in this work we provide selected simulation results as a series of frames obtained from these video files. After each series of frames we add the plots showing the corresponding deviation in the single components of the state space (uncontrolled vs. controlled solution). Finally there are plots of each component of the control input.
For the planar pendulum system with n links an explicit general form of the equations of motion is derived in **appendix C**. In this appendix we generalize a result from [Lam and Davison, 2006] where controllability of the linearized n-pendulum around its inverted position is shown. We developed an explicit formula for the linearization around an equilibrium state of the n-pendulum and proof controllability of the linearization around an arbitrary equilibrium state. We show that theoretically it is sufficient to actuate only the first (or likewise the last) pendulum link to establish controllability.

Chapter 2

Time-optimal control for the linear simple pendulum equation - a historical review

One of the first analytical results in optimal control theory was given by D. Bushaw in his PhD-thesis written in 1952 and partially published in [Bushaw, 1958] where he investigated the following problem:
If g is a given function

$$g : \begin{cases} \mathbb{R}^2 \supset D \to \mathbb{R}, \quad 0 \in D \\ (x, y) \mapsto g(x, y) \end{cases} , \tag{2.1}$$

find a function $u(x, y)$ defined on D with the following properties

- $u(x, y)$ assumes only the values -1 and 1.
- For any point (x_0, y_0) a solution $x(t)$ of the differential system

  $$\ddot{x}(t) + g(x(t), \dot{x}(t)) = u(x(t), \dot{x}(t)), \quad x(0) = x_0,\ \dot{x}(0) = y_0 \tag{2.2}$$

  exists, and there is a (least) positive value of t, say t^*, such that for this solution $x(t^*) = \dot{x}(t^*) = 0$.
- For all points in D, t^* is minimal with respect to the class of functions u satisfying the first two properties.

With respect to some additional assumptions the first property is no real restriction. It would suffice to assume $|u| \leq 1$. The reason is the so called "bang-bang-principle", which roughly speaking states that under suitable assumptions one can replace every admissible bounded control by a control function, which only assumes the limiting values (see appendix A). The assumption of bounded control functions is often justified with the idea that in practice control inputs have to be realized by actuators which can only provide limited forces to the underlying mechanical system.
A scalar-valued bang-bang control only assumes two values and is therefore also called *on-off control*. In many practical problems bang-bang controls have to be avoided. For example no one wants to drive a car where the only options are maximum acceleration or maximum deceleration.

We will give a glimpse of Bushaw's idea in terms of the linearized simple pendulum (in this case we have $g(x, \dot{x}) = x$)

$$\ddot{x}(t) + x(t) = u(x, \dot{x}). \tag{2.3}$$

The control u only assumes the values $+1$ and -1. For $u = -1$ equation (2.3) becomes

$$\ddot{x}(t) + x(t) = -1, \tag{2.4}$$

which we will refer to as N-System. Its solution for the initial value $\begin{pmatrix} x(0) \\ \dot{x}(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ is given by

$$x(t) = y_0 \cdot \sin(t) + (x_0 - 1)\cos(t) - 1. \tag{2.5}$$

For $u = +1$ equation (2.3) simplifies to

$$\ddot{x}(t) + x(t) = 1, \tag{2.6}$$

which we will refer to as P-System. Its solution for $\begin{pmatrix} x(0) \\ \dot{x}(0) \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ can be expressed as

$$x(t) = y_0 \cdot \sin(t) + (x_0 - 1)\cos(t) + 1. \tag{2.7}$$

The solution trajectories in the phase space $(x, \dot{x}) =: (x, y)$ are concentric circles or parts of them ("arcs") with $(-1, 0)$ as center for the N-system and $(1, 0)$ for the P-system. The radius is determined by the initial value $(x(0), \dot{x}(0))$ and – as long as $x(0) \neq 0$ – depends on u.

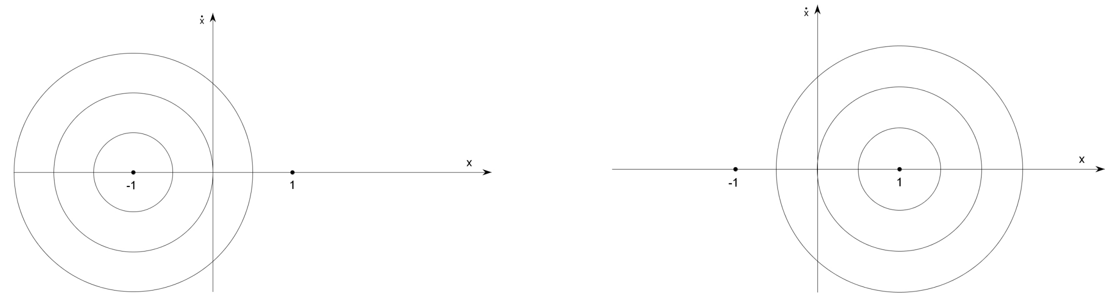


Figure 2.1: solution trajectories for N-system (left) and P-system (right)

For a given initial value we follow the circle containing this initial point in clockwise direction until u changes its sign. As every point in the phase space lies on a solution trajectory belonging to the N-system and a solution trajectory belonging to the P-system, changing the sign of u can be geometrically interpreted as changing from a P-arc to a N-arc or from a N-arc to a P-arc. On the other hand, such a change can only occur, when u changes sign. Such, finding a solution to problem (2.3) is equivalent to finding a connected path consisting of P-arcs and N-arcs which leads from the initial point to the origin.
Bushaw's solution can be summarized as follows, where we will only show the third item:

1. Solution trajectories consist of continuously assembled alternating P- and N-arcs (paths).
2. Except for possibly the first and last arc, all arcs are semicircles.
3. Above the x-axis there can only be transitions from N-arcs to P-arcs and below the x-axis from P-arcs to N-arcs.

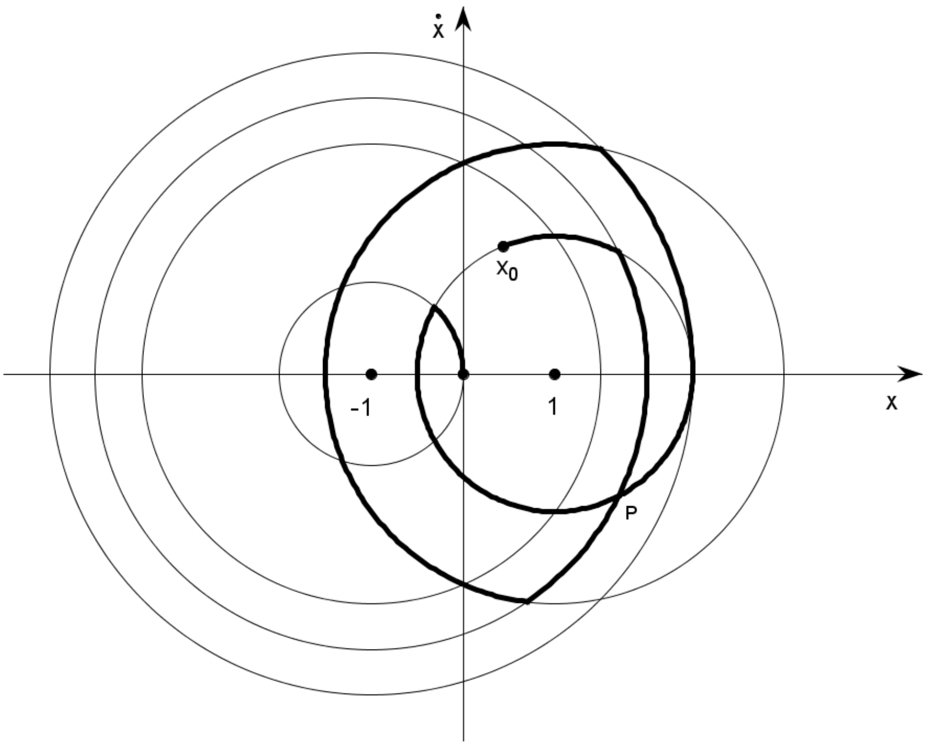


Figure 2.2: a possible solution path (not time-optimal!)

4. Existence follows from construction.

Figure (2.2) gives an example of a solution which is not time-optimal.
The N/P-system can be written as first order system

$$\frac{d}{dt}x = y \qquad (2.8)$$
$$\frac{d}{dt}y = -x \pm 1$$

Let t_{ABC} denote the time needed to move along the path ABC in figure (2.4) and t_{ADC} the time length for path ADC above the x-axis. Then we have due to (2.8) and separation of variables

$$t_{ABC} = \int_{ABC} y^{-1}\, dx < \int_{ADC} y^{-1}\, dx = t_{ADC} \qquad (2.9)$$

since for every fixed x (except for the point A and C) the corresponding y-value on the arc ABC is greater than the one belonging to the arc ADC. An analogous argument shows that below the x-axis there can only be transitions from P-arcs to N-arcs.
As indicated in figure (2.3) fewer time is needed to move from point P to point Q along a $P-$arc than along a $N-$arc, since the time needed is proportional to the angle corresponding to the arc.
With the help of these properties one can show existence of time-optimal controls for arbitrary initial data and the property that for time-optimal solution trajectories all arcs are semicircles except possibly the first and last one. We will omit the details of the proof, which can be found in [Bushaw, 1958].

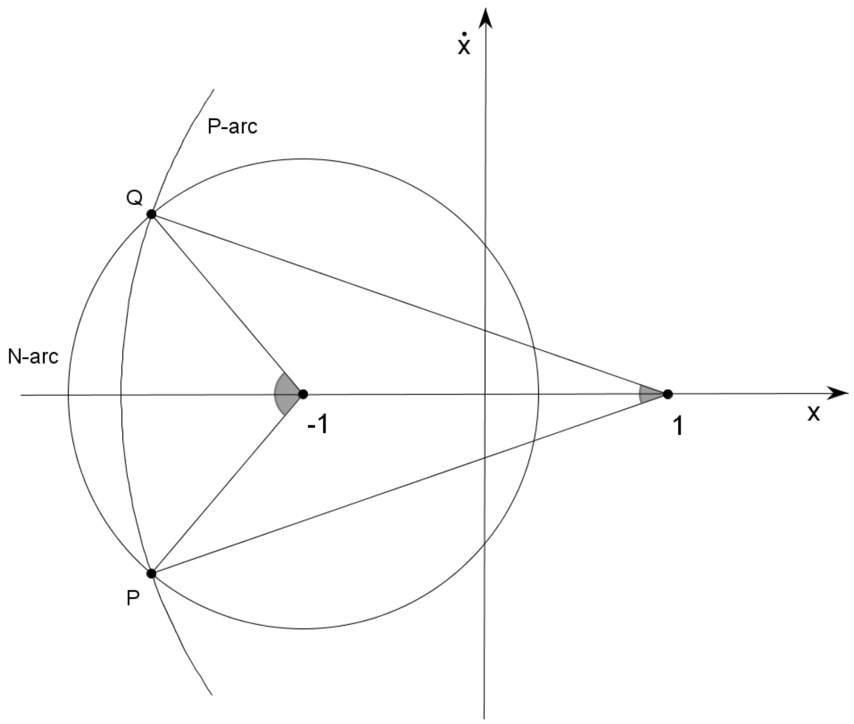


Figure 2.3: choosing the right arc from P to Q

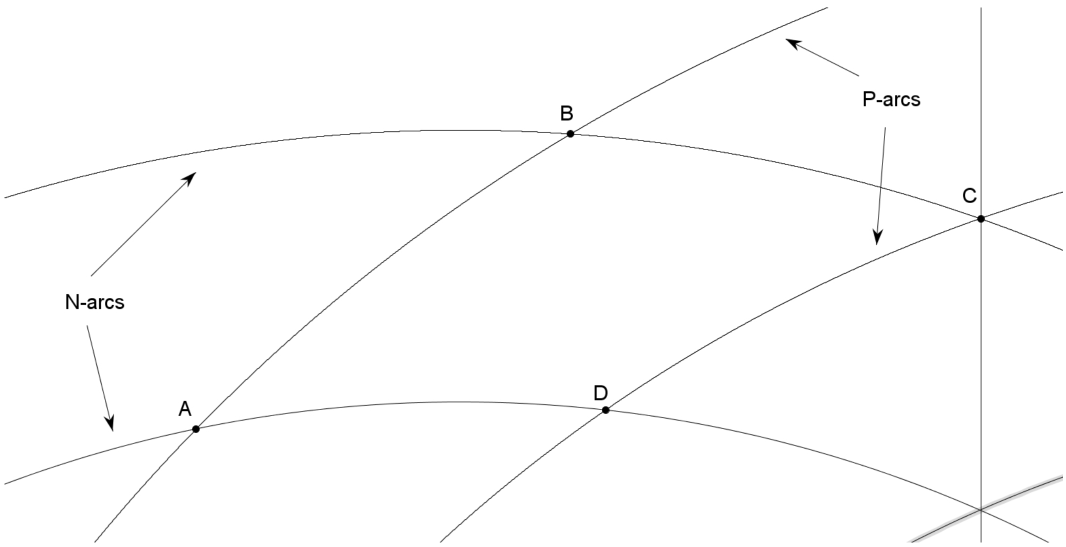


Figure 2.4: in the upper half plane there are only PN transitions possible

2.1 Linear time-optimal control theory

In this section we will introduce the linear time-optimal control problem and its solution.

2.1.1 Problem formulation

The systems we regard are of the form

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \qquad x(0) = x_0 \tag{2.10}$$

where for $r \leq n$, $r, n \in \mathbb{N}$, $x(t) \in \mathbb{R}^n$ the matrices $A(t) \in \mathbb{R}^{(n,n)}$, $B(t) \in \mathbb{R}^{(n,r)}$ have at least locally integrable elements. We are looking for a control function $u : \mathbb{R}_0^+ \to \Omega \subseteq \mathbb{R}^r$, $\Omega = [-1, 1]^r$, which brings the linear time-varying system (2.10) to the zero state $0 \in \mathbb{R}^n$ in minimum time t^*. The linearized pendulum equation (cf. section (5.1))

$$\ddot{z}(t) + z(t) = u(t) \tag{2.11}$$

can be written as first-order system with $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $x : \mathbb{R} \to \mathbb{R}^2$.
It will be shown with the methods of LaSalle [LaSalle, 1960, Hermes and LaSalle, 1969], that for every possible initial condition there is an admissible control u^* and a finite time t^* solving the above problem. The resulting control is *bang-bang* ($|u| = 1$) and unique.

2.1.2 Transforming into an equivalent problem

Let $\Phi(t, 0)$ denote the solution to $\dot{X}(t) = A(t)X(t)$ with initial condition $X(0) = I$ where $Y(t) := \Phi^{-1}(t, 0)B(t)$ and $A(t), B(t)$ are the system matrices of (2.10)[1]. The solution of (2.10) is given by

$$x(t, u) = \Phi(t, 0)x_0 + \Phi(t, 0) \int_0^t Y(\tau)u(\tau)d\tau \tag{2.12}$$

which can be directly verified (or cf. e.g. [Balakrishnan, 1976, Theorem 4.8.3]).
The solution formula (2.12) for equation (2.10) shows, that the control u only influences the term $\int_0^t Y(\tau)u(\tau)d\tau$. Since the final state shall be the stable equilibrium represented by the origin of the phase space. Both sides of (2.12) vanish at $t = t^*$. Therefore problem (2.10) is equivalent to finding an admissible u^* and a minimum time t^* for a given initial state x_0 such that:

$$-x_0 = \int_0^{t^*} Y(\tau)u^*(\tau)d\tau. \tag{2.13}$$

We define

$$\mathcal{R}(t) := \{\int_0^t Y(\tau)u(\tau)d\tau, u \in \Omega\} \tag{2.14}$$

as "reachability set". The solution to the original problem is then equivalent to finding an admissible control u so that $-x_0 \in \mathcal{R}(t^*)$ and $-x_0 \notin \mathcal{R}(t)$ for all t with $0 \leq t < t^*$.

[1] Since in this chapter we always assume that the initial time $t_0 = 0$ we write $Y(t)$ instead of $Y(t, 0)$ for simplicity

2.1.3 Time-optimal control

So far we do not know anything about the existence of a suitable time-optimal control function u. We state that the i-th component u_i^* of any time-optimal control u^* (if it exists) must be of the form

$$u_i^*(\tau) = \operatorname{sgn}(\eta^T Y(\tau))_i \quad \text{for } (\eta^T Y(\tau))_i \neq 0, \tag{2.15}$$

for some $\eta \in \mathbb{R}^n \setminus \{0\}$.

Let $u(\tau) \in \Omega$ be an admissible control, such that the system reaches the origin of the phase space in finite time. Because (2.13) is an equivalent formulation of problem (2.10) this means

$$-x_0 \in \mathcal{R}(t). \tag{2.16}$$

Even for discontinuous control functions u the functional $\int_0^t Y(\tau)u(\tau)d\tau$ and the set $\mathcal{R}(t)$ are continuous functions of t. For arbitrary positive times t_1 and t_2, with $t_1 < t_2$ we have $\mathcal{R}(t_1) \subset \mathcal{R}(t_2)$ meaning whenever we can reach the origin within time t_1 we can reach the origin in time t_2. To see this we could take the control function u_2 which is the same as u_1 in the interval $[0, t_1]$ and 0 elsewhere.
We are looking for a control function $\tilde{u}(\tau)$ such that $-x_0$ is contained in the set $\mathcal{R}(t^*)$ where t^* is as small as possible. Because $\mathcal{R}(t)$ is convex (cf. appendix A) and depends continuously on time, $-x_0$ has to be on $\partial\mathcal{R}(t^*)$.

There is a support hyperplane $\mathcal{H}$ containing $-x_0$, such that all points belonging to $\mathcal{R}(t^*)$ lie on the same side of $\mathcal{H}$ (e.g. [Eggleston, 1958]). Let η be the normal to $\mathcal{H}$ in $-x_0$ and pointing outwards of $\mathcal{R}(t^*)$[2]. For every control function $\tilde{u}$ bringing the system from the initial state x_0 to the origin in minimum time t^* we have:

$$\begin{aligned}
&\eta^T \int_0^{t^*} Y(\tau)u(\tau)d\tau \leq \eta^T \int_0^{t^*} Y(\tau)\tilde{u}(\tau)d\tau \; \forall u \in \Omega \iff \\
&\int_0^{t^*} \eta^T Y(\tau)u(\tau)d\tau \leq \int_0^{t^*} \eta^T Y(\tau)\tilde{u}(\tau)d\tau \; \forall u \in \Omega \implies \\
&\int_0^{t^*} \eta^T Y(\tau)\tilde{u}(\tau)d\tau = \int_0^{t^*} \left|\eta^T Y(\tau)\right| d\tau
\end{aligned}$$

Therefore time-optimal controls have to be of the form (2.15).

Remark: If $\eta^T Y(\tau)$ vanishes, (2.15) is not defined and does not provide any information about the control function. This does not necessarily mean that the control u has no effect as we can see from (2.12).
For the linear pendulum equation (2.11) a possibly time-optimal control has to be bang-bang because we can compute the explicit representation of $Y(\tau)$:

$$Y(\tau) = \Phi(\tau,0)^{-1}(\tau)B = \begin{pmatrix} \cos(\tau) & \sin(\tau) \\ -\sin(\tau) & \cos(\tau) \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin(\tau) \\ \cos(\tau) \end{pmatrix} \tag{2.17}$$

and since $\sin(\tau)$ and $\cos(\tau)$ are linearly independent functions $\eta^T Y(\tau)$ can only vanish on a set of measure zero.

[2]Note that η does not need to be a normal to $\partial\mathcal{R}(t^*)$ in $-x_0$ as the separating hyperplane $\mathcal{H}$ need not be unique

2.1.4 Uniqueness

The normal η is determined by the point $-x_0$ in the phase space and the position of the hyperplane $\mathcal{H}$ containing this point. The control function

$$\tilde{u}(\tau) = \operatorname{sgn}(\eta^T Y(\tau)) \tag{2.18}$$

need not be unique since for example it could be possible to define more than just one hyperplane containing $-x_0$. In this case there would be at least two different suitable normal vectors. It is clear that uniqueness of the (time-optimal) control function is strongly linked with the geometry of the reachability set. We already mentioned that (2.18) is only a necessary condition. Under certain circumstances this necessary condition turns out to be sufficient.

Definition 2.1. *(i) A point x is said to be an* extreme *point of the convex set M if $x \in M$ and there are no two points $x_1, x_2 \in M$, $x_1 \neq x_2$ such that x can be expressed as linear combination of x_1 and x_2.*

(ii) A point x is called exposed *point of the convex set M if $x \in M$ and there is a support hyperplane to M that meets M in the single point x.*

For problem (2.10) the necessary condition (2.15) is also sufficient. We formulate the following theorem which holds for all linear time-varying systems of the form

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \tag{2.19}$$

where $x(t)$ is a n-vector, $A(t)$ a square matrix of size $n \times n$ with integrable elements $a_{ij}(t)$. The control vector $u(t)$ is a r-vector, where r is at most n and $B(t)$ is a $n \times r$-matrix with integrable elements. Before stating the theorem we will introduce some useful notions following [Hermes and LaSalle, 1969]:

Definition 2.2.

(i) Two control functions $u = (u_1, \dots, u_r)^T$ and $v = (v_1, \dots, v_r)^T$ are said to be essentially equal *on $[0, t]$ if for every $j = 1, \dots, r$ we have $u_j = v_j$ almost everywhere on $[0, t]$ where the j-th column of B is different from the zero-vector: $b_{*j}(t) \neq 0$.*

(ii) The control to reach a point q in time t is said to be essentially unique *if all controls to reach this point are essentially equal.*

The first part means that the controls u and v are equal almost everywhere whenever they are effective.
We will also use the self explaining terms *essentially bang-bang* or *essentially determined*. The term *essentially* can be thought of as an abbreviation for *whenever the control is effective.*

Theorem 2.3. *For system (2.19) we have*

1. *q is an extremal point of $\mathcal{R}(t) \Longleftrightarrow$ there is a unique trajectory from the origin to q [Hermes and LaSalle, 1969, Theorem 14.2].*

2. *The control function u^* to reach q^* in minimum time t^* is determined* essentially unique *by $\operatorname{sgn}(\eta^T Y(t)) = \operatorname{sgn}(\eta^T \Phi^{-1}(t, 0)B), \eta \neq 0$ if and only if q^* is an exposed point of $\mathcal{R}(t^*)$ [Hermes and LaSalle, 1969, Theorem 15.1].*

Proof. 1. Suppose there are two control functions u and v, bringing the system (2.19) to state q in the same time t on different solution trajectories. Then there is a time $s \in (0,t)$ such that

$$p_u := \int_0^s Y(\tau)u(\tau)d\tau \neq \int_0^s Y(\tau)v(\tau)d\tau =: p_v. \tag{2.20}$$

Define q_1 and q_2 as follows:

$$\begin{aligned} q_1 &= p_u + (q - p_v) \\ q_2 &= p_v + (q - p_u). \end{aligned}$$

As constructed q_1 will be reached when using the control function u for the time interval $[0,s]$ and the control function v for $(s,t]$. q_2 will be reached when using first the control function v and then u. Since both points can be reached in total time t they belong to the reachability set $\mathcal{R}(t)$. Since $q = \frac{1}{2}q_1 + \frac{1}{2}q_2$ it is not an extreme point of $\mathcal{R}(t)$. This means whenever q is an extreme point of the reachability set $\mathcal{R}(t)$ the trajectory to this point is unique. We will now show the other direction. Suppose q is contained in $\mathcal{R}(t)$ but it is not an extreme point. Then q can be represented as convex combination of two points, say q_1 and q_2 of $\mathcal{R}(t)$, $q_1 \neq q_2$:

$$q = \frac{1}{2}q_1 + \frac{1}{2}q_2, \quad q_1, q_2 \in \mathcal{R}(t).$$

Every point in $\mathcal{R}(t)$ can be reached by using a bang-bang control. Therefore there are bang-bang controls u and v, such that q_1 can be reached with control u in time t and q_2 with control v also in time t. We can reach q in time t by using the control $w := \frac{1}{2}u + \frac{1}{2}v$. The control function w is not bang-bang, since u and v are different bang-bang controls. This means, that at least one component, say w_j is not bang-bang. Due to the bang-bang theorem there is a bang-bang control $\hat{w}_j$ such the following equation holds:

$$y_j(t,w) := \int_0^t Y_{*j}(\tau)w_j(\tau)d\tau = \int_0^t Y_{*j}(\tau)\hat{w}_j(\tau)d\tau =: y_j(t,\hat{w}) \tag{2.21}$$

For all other components we take $\hat{w}_i := w_i$. Then we have

$$y(t,w) := \int_0^t Y(\tau)w(\tau)d\tau = \int_0^t Y(\tau)\hat{w}(\tau)d\tau =: y(t,\hat{w}) \tag{2.22}$$

We can now find a time $s < t$ such that $y(s,w) \neq y(s,\hat{w})$:

$$\begin{aligned} y(t,w) - y(t,\hat{w}) &= \int_0^t Y(\tau)w(\tau)d\tau - \int_0^s Y(\tau)\hat{w}(\tau)d\tau \\ &= \int_0^t Y(\tau)(w(\tau) - \hat{w}(\tau))d\tau \\ &= \int_0^t \sum_{i=1}^n Y_{*i}(\tau)(w_i(\tau) - \hat{w}_i(\tau))d\tau \\ &= \int_0^t Y_{*j}(\tau)(w_j(\tau) - \hat{w}_j(\tau))d\tau \end{aligned}$$

Because w_j and $\hat{w}_j$ are not identical and the left hand side is absolutely continuous there must be a time $s < t$ such that $y(s,w) - y(s,\hat{w}) \neq 0$. But this means, the trajectory is not unique, completing the proof of the first part.

2. Every exposed point is also an extreme point. So the trajectory to an exposed point is unique. We will show, that for exposed points the control function is uniquely determined by the necessary condition (2.18).
   Let $u^*(t) = \mathrm{sgn}(\eta^T Y(t))$ be essentially determined on the interval $[0, t^*]$. The point q^* which will be reached in time t^* by using this control is a boundary point of $\mathcal{R}(t^*)$ having η as normal to a support plane $\mathcal{H}^*$. Points on $\mathcal{H}^*$ can only be reached by controls of the form $\mathrm{sgn}(\eta^T Y(t))$. Since any other control of this form is essentially equal to u^* the point q^* is the only point of the reachability set $\mathcal{R}(t^*)$ which lies on $\mathcal{H}^*$ which means that q^* is an exposed point.
   Now let q^* be an exposed point of the reachability set $\mathcal{R}(t^*)$. For any support plane $\mathcal{H}$ for $\mathcal{R}(t^*)$ at q^* we have $\mathcal{H} \cap \mathcal{R}(t^*) = \{q^*\}$. In particular there is a normal η such that for $u^*(t) = \mathrm{sgn}(\eta^T Y(t))$ and q^* we have $q^* = y(t^*, u^*) := \int_0^{t^*} \eta^T Y(t) dt$.
   q^* is also an extreme point which means that the trajectory to reach this point is unique. From the proof of the first part of this theorem we know that controls have to be bang-bang. Let u and v be bang-bang-controls, such that q^* is reached in time t^*. Then we can reach q^* also by using the the control $w := \frac{1}{2}u + \frac{1}{2}v$. This control function has to be bang-bang as well. But this is only possible, when u and v are essentially equal which completes the proof.

□

Before giving an example we present a theorem about existence of a time-optimal solution for linear systems:

2.1.5 Existence

Theorem 2.4. *[Hermes and LaSalle, 1969, Th. 13.1]*
Regard system (2.10). *If for a given state x_1 and an initial value x_0 there is a control u and a time τ such that the solution of the controlled system starting in x_0 reaches x_1 in time τ, then there is a time-optimal control.*

Proof. Define $w(t) := \Phi^{-1}(t, 0)x_1 - x_0$, then we have $x_1 = x(\tau, u)$ is equivalent to $w(\tau) \in \mathcal{R}(\tau)$ which follows directly from (2.12). For $t^* = \inf_t\{w(t) \in \mathcal{R}(t)\}$ we clearly have $0 \leq t^* \leq t$. Let $y(t, \hat{u})$ denote the elements of $\mathcal{R}(t)$ corresponding to the control input $\hat{u}$.
There is a sequence t_n converging to t^* and a sequence of control inputs u_n such that for $w(t_n) := y(t_n, u_n)$ we have

$$\|w(t^*) - y(t^*, u_n)\| \leq \|w(t^*) - w(t_n)\| + \|y(t_n, u_n) - y(t^*, u_n)\| \tag{2.23}$$

$$\leq \|w(t^*) - w(t_n)\| + \left\| \int_{t^*}^{t_n} Y(\tau) u_n(\tau) d\tau \right\| \tag{2.24}$$

where for continuity reasons the right hand side converges to 0 as n tends to infinity. Therefore $y(t^*, u_n)$ converges to $w(t^*) = y(t^*, u^*)$ which lies in $\mathcal{R}(t^*)$ as the reachability set is closed (a proof for this fact can be found in [Halmos, 1948],[Lyapunov, 1940]) as well as the input space. Such a time-optimal control u^* exists. □

2.1.6 Summary example: A time-optimal control for a linear pendulum model

For the linear pendulum optimal controls are of the form (2.18) and using (2.17) it reads as

$$u^*(t) = \operatorname{sgn}(\eta^T Y(t)) = \operatorname{sgn}(\eta^T \Phi^{-1}(t,0)B) = \operatorname{sgn}(-\eta_1 \sin(t) + \eta_2 \cos(t)). \quad (2.25)$$

Because $\sin(t)$ and $\cos(t)$ are linearly independent functions and η is nontrivial by being a normal, a possibly time-optimal control is uniquely determined by (2.25). So every boundary point of the reachability set is an exposed point due to theorem (2.3) which means, the reachability set is strictly convex. We also know that is strictly increasing with t.

Remark: For the linear pendulum equation (2.10) the set $\{t|\eta^T Y(t) = 0, t \in [0, t^*]\}$ has measure zero. Linear time-varying systems having this property are called *normal* [Hermes and LaSalle, 1969, corollary 15.1]. For normal system time-optimal controls are essentially unique determined by a bang-bang control. The reachability sets are strictly convex.

We will give a geometric approximative solution to the problem of finding a time-optimal control for the linearized pendulum equation (2.10) for a concrete initial value:

$$\begin{aligned} \dot{x}(t) &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t) \\ x(0) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} =: x_0 \end{aligned}$$

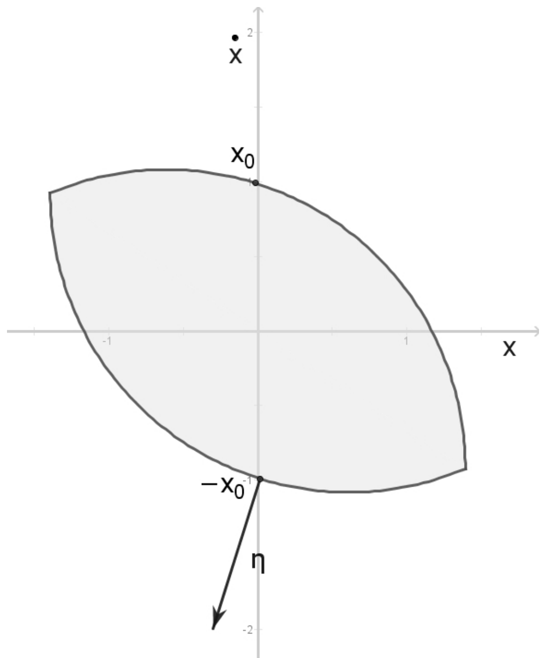


Figure 2.5: finding the normal η for the linearized pendulum

A time-optimal control function for this problem has to be bang-bang as showed above. Since $-x_0$ is contained in $\mathcal{R}(\pi)$ we know, that the optimal control function will change sign at most one time. With the help of Cinderella[3] one can visualize all reachability sets for $t \leq \pi$ ($\rightsquigarrow$ interactive Java-applet: http://home.in.tum.de/~lehl/pendel.html). So we choose a time t such that $-x_0$ is a boundary point of the reachability set $\mathcal{R}(t)$. The normal to the support hyperplane containing $-x_0$ is given by η. If we choose a vector ξ such that $\eta = \xi/\|\xi\|$ then $u^*(t) = \operatorname{sgn}(\eta^T Y(t)) = \operatorname{sgn}(\xi^T Y(t)/\|\xi\|) = \operatorname{sgn}(\xi^T Y(t))$. We may choose an approximation for a suitable ξ as $(-6, -19)^T$.

The (approximative) time-optimal control is then given by:

$$u^*(t) = \operatorname{sgn}(\xi^T Y(t)) \approx \operatorname{sgn}(\begin{pmatrix} -6 \\ -19 \end{pmatrix}^T \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}) = \operatorname{sgn}(6\sin(t) - 19\cos(t)). \quad (2.26)$$

We start with $u^* = -1$ and follow a circle around $(-1, 0)$. At time $t \approx 1.27$ we change to $+1$ and follow a circle around $(1, 0)$ until the origin is reached after 0.72 further time units. The

[3] http://www.cinderella.de by Prof. Dr. Jürgen Richter-Gebert and Prof. Dr. Ulrich Kortenkamp

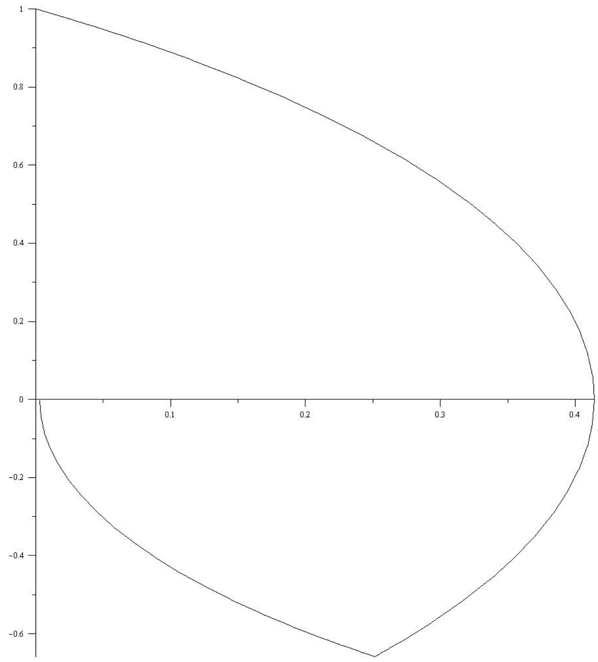


Figure 2.6: solution trajectory in the $(x, \dot{x})$ space

solution is shown in figure (2.6) where we can see that we only slightly miss the origin.

Remark: Using a numerical algorithm (see appendix B and [Eaton, 1962]) we obtain 2.0 as t^* and $\eta^* = \begin{pmatrix} -0.2924\ldots \\ -0.9562\ldots \end{pmatrix}$ where the final distance to the origin is less than 4/1000. Nevertheless the rude geometrical guess led to an acceptable result ($t^* = 1.99$).

Remark: There is a powerful tool to obtain this result called *Pontrjagin's maximum principle* (see appendix (D)), which gives a necessary condition for a trajectory to be (time-)optimal. The method presented above is also sufficient.

Chapter 3

Controllability

The method presented in the previous chapter by Hermes and LaSalle gives *necessary* conditions for time-optimal controls. Only for special cases these necessary conditions are also sufficient, e.g. for linear time-varying systems which are *normal*. From LaSalle's theory we already know that for linear systems (time-varying or not) the existence of an admissible control implies the existence of a time-optimal control.
In this chapter we will give some results about controllability of linear and nonlinear systems.

Consider the dynamical system

$$\dot{x} = f(x, u); \tag{3.1}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ the control and f is a locally Lipschitzian function which maps into $\mathbb{R}^n$. Given an initial value $x(t_0) = x_0$ and an admissible control $u(t)$, $t \geq t_0$ we denote the corresponding solution of (3.1) – if it exists – as $x(t, u; x_0)$ for $t \geq t_0$. In many times we will omit the initial value x_0 and write $x(t, u)$ for simplicity.

Definition 3.1. *Reachability set*
Given a state $q \in \mathbb{R}^n$ we define the reachability set *$\mathcal{R}(q, T)$ to be the set of all states $p \in \mathbb{R}^n$ for which there exists an admissible control u such that $x(T, u, q) = p$*

Remark: For $q = 0$ we also write $\mathcal{R}(T)$ instead of $\mathcal{R}(0, T)$.

Definition 3.2. *Reachable set*
The reachable set *of state q at time T is defined as*

$$\mathcal{R}_T(q) = \bigcup_{t_0 \leq t \leq T} \mathcal{R}(q, t). \tag{3.2}$$

Definition 3.3. *Controllability*
Given system (3.1) *and points x_0, x_1 of the state space. The system is said to be* controllable from x_0 to x_1 *if there is an admissible control u such that the corresponding solution starts in x_0 and ends up in x_1 in finite time. System* (3.1) *is said to be* asymptotically controllable from x_0 to x_1 *if there is an admissible control u such that x_1 is reached in infinite time. Or to be more precise, for every $\varepsilon > 0$ there is a finite time T such that the trajectory which starts in x_0 and corresponds to the control u hits the ball around x_1 with radius ε and remains there for all time $t > T$.*

Definition 3.4. *Local controllability*
Given system (3.1) *and a point x_0 of the state space system* (3.1) *is said to be* locally controllable

in x_0 if there is an environment of x_0 in the state space, such that x_0 is controllable to every point in this environment in finite time.

Definition 3.5. *Small-time local controllability (STLC)[Sussmann, 1983b]*
A control system (3.1) *is said to be* small-time locally controllable *(STLC) from a point p if, for every time $T > 0$, p is in the interior of the set of points q that can be reached from p in time not greater than T.*

Figure (3.1) shows the reachability sets for a two-dimensional-system at times $t = h, 2h, 3h$ and $4h$, which is locally controllable, but which is not small-time locally controllable at the center of the circle.

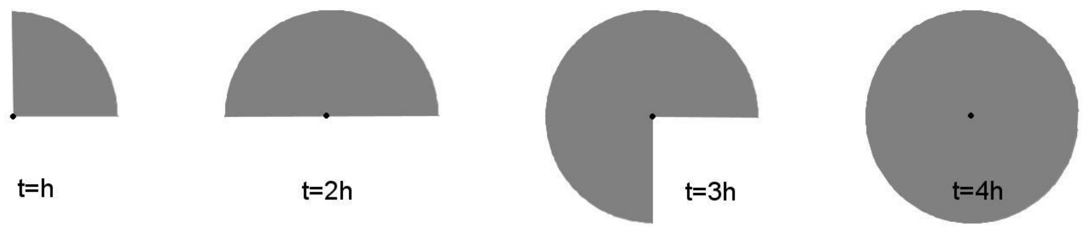


Figure 3.1: Example for a locally but not small-time locally controllable system

The reachability sets shown in (3.1) have a remarkable property. Although the first three reachability sets do not contain an environment of the center of the circle they all contain a nonempty open set in the state space. This motivates the following definition:

Definition 3.6. *Accessibility*
System (3.1) *is said to be accessible from state $q \in \mathbb{R}^n$ if for every $T > t_0$ the reachable set $\mathcal{R}_T(p)$ contains a nonempty open set.*

Definition 3.7. *Global controllability*
System (3.1) *is said to be* globally controllable *(or* completely controllable*) if for any two points x_0, x_1 there exists an admissible control u that steers x_0 to x_1 along a trajectory of the system in finite time.*
System (3.1) *is said to be* globally asymptotically controllable *if for any two points x_0, x_1 and any $\varepsilon > 0$ there exists an admissible control u that steers x_0 to an ε-environment of x_1 along a trajectory of the system in a finite time T and the solution trajectory remains there for all times $t > T$.*

Definition 3.8. *Null-controllability*
A state x_0 is said to be null-controllable *if there exists an admissible control u that steers x_0 to the origin in finite time.*
System (3.1) *is said to be* globally null-controllable *if every state x_0 is null-controllable.*
State x_0 is asymptotically null-controllable *if there is an admissible control steering x_0 to the origin in infinite time.*

3.1 Controllability of linear systems

As linear autonomous systems are a special case of linear time-varying systems all controllability criteria which apply for the latter also apply for the former. We will start with linear time-varying systems and will then give the simplified versions for the time-invariant cases as well as some criteria, which do apply for this case only. The main references for this chapter are [Chen, 1970] and [Klamka, 1991].
Before we start with the controllability criteria, we will formulate two lemmata:

Lemma 3.1. *[Chen, 1970, Th. 5-2],[Klamka, 1991, Lemma 1.3.2]*
Assume that the functions $f_1, \dots, f_n : \mathbb{R} \to \mathbb{R}^m$ have continuous derivatives up to order $n-1$ on the interval $[t_0, t_2]$. Let F be the $n \times m$ matrix with f_i as its i-th row and let $F^{(k)}$ be the k-th derivative of F. If there is some $t_1 \in [t_0, t_2]$ such that the $n \times n \cdot m$ matrix

$$[F(t_1)|F^{(1)}(t_1)|\dots|F^{(n-1)}(t_1)] \tag{3.3}$$

has rank n, then the functions f_i are linearly independent on $[t_0, t_2]$ over the field of real numbers.

Proof. by contradiction:
Suppose there is some t_1 in $[t_0, t_2]$ such that

$$\text{rank}[F(t_1)|F^{(1)}(t_1)|\dots|F^{(n-1)}(t_1)] = n$$

but the functions f_i are linearly dependent on $[t_0, t_2]$. Then there is a nonzero vector $\alpha \in \mathbb{R}^n$ such that $\alpha^T F(t) = 0^T$ for all $t \in [t_0, t_2]$. Building the k-th time derivative up to order $n-1$ yields

$$\alpha^T F^{(k)}(t) = 0^T \text{ for all } t \in [t_0, t_2] \text{ and } k = 1, \dots, n-1$$

Therefore we have

$$\alpha^T [F(t_1)|F^{(1)}(t_1)|\dots|F^{(n-1)}(t_1)] = 0^T \tag{3.4}$$

which means that the n rows of $[F(t_1)|F^{(1)}(t_1)|\dots|F^{(n-1)}(t_1)]$ are linearly dependent contradicting the assumption that $[F(t_1)|F^{(1)}(t_1)|\dots|F^{(n-1)}(t_1)]$ has rank n. □

Remark: Lemma (3.1) is a sufficient but not necessary condition for a set of functions to be linearly independent. For example the functions $f_1(t) = t^5$ and $f_2(t) = |t^5|$ are linearly independent on $[-1, 1]$ but fail to match the condition of lemma (3.1).

Lemma 3.2. *[Klamka, 1991, Lemma 1.3.1]*
The functions $f_1, \dots, f_n$ are linearly independent on $[t_0, t_1]$ if and only if the $n \times n$ matrix defined by

$$G(t_0, t_1) := \int_{t_0}^{t_1} F(t) F^T(t) dt \tag{3.5}$$

has full rank.

Proof. Necessity: (by contradiction)
Assume that the functions $f_1, \dots, f_n$ are linearly independent on $[t_0, t_1]$ but the matrix $G(t_0, t_1)$

is singular. Then there is a nonzero vector $\alpha \in \mathbb{R}^n$ such that $\alpha^T G(t_0, t_1) = 0^T$ a.e. in $[t_0, t_1]$. We then have

$$
\begin{aligned}
0 &= \alpha^T G(t_0, t_1)\alpha \\
&= \alpha^T \int_{t_0}^{t_1} F(t)F^T(t)dt\alpha \\
&= \int_{t_0}^{t_1} \alpha^T F(t)F^T(t)\alpha dt \\
&= \int_{t_0}^{t_1} \left(\alpha^T F(t)\right)\left(\alpha^T F(t)\right)^T dt
\end{aligned} \tag{3.6}
$$

Since the integrand is nonnegative, $\alpha^T F(t) = 0^T$ almost everywhere. But this is a contradiction to the assumption that the rows $f_i(t)$, $i = 1, \ldots, n$ of the matrix $F(t)$ are linearly independent.
Sufficiency: (by contradiction)
Assume that $G(t_0, t_1)$ is nonsingular and the functions $f_1, \ldots, f_n$ are linear dependent. Then there is a nonzero vector $\alpha \in \mathbb{R}^n$ such that $\alpha^T F(t) = 0^T$ a.e. in $[t_0, t_1]$. We then have

$$
0^T = \int_{t_0}^{t_1} \alpha^T F(t)F^T(t)dt = \alpha^T G(t_0, t_1) \tag{3.7}
$$

which is a contradiction to the nonsingularity of $G(t_0, t_1)$. □

3.1.1 Linear time-varying systems

Now we will give some controllability criteria for time-varying linear systems.
We consider the system

$$
\frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t) \tag{3.8}
$$

where $x(t) \in \mathbb{R}^n$ denotes the state vector, $u \in L^1_{\text{loc}}([t_0, \infty], \mathbb{R}^m)$[1] will be admissible controls and $A(t)$ is an $n \times n$ matrix with locally Lebesgue integrable elements $a_{ij} \in L^1_{\text{loc}}([t_0, \infty], \mathbb{R})$ for $i = 1, \ldots, n;\ j = 1, \ldots, n$. The matrix $B(t)$ has size $n \times m$ where $m \leq n$ and $b_{ij} \in L^1_{\text{loc}}([t_0, \infty], \mathbb{R})$ for $i = 1, \ldots, n;\ j = 1, \ldots, m$.
For a given control function $u(t)$ and initial value $x(t_0) = x_0$ there is a unique solution of equation (3.8) denoted by $x(t, x(t_0), u)$ which is absolutely continuous[2] (see for example [Desoer and Zadeh, 1963]).
The solution can be expressed as

$$
x(t, x(t_0), u) = \Phi(t, t_0)x(t_0) + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)d\tau, \quad t \geq t_0 \tag{3.9}
$$

where $\Phi(t, t_0)$ is the the solution of $\dot{X}(t) = A(t)X(t)$, $X(t_0) = I$. It is called *fundamental matrix* or *transition matrix* and defined for all t, t_0 in $(-\infty, \infty)$. The fundamental matrix has

[1] in [Klamka, 1991] u is assumed to be $L^2_{\text{loc}}([t_0, \infty], \mathbb{R}^m)$ allowing a proof that uses Hilbert space techniques, nevertheless there is no need to make this strong regularity assumption on the control input

[2] a function $f : \mathbb{R} \supset [a, b] \to \mathbb{R}$ is said to be *absolutely continuous* if it has a derivative f' a.e. which is Lebesgue integrable and for all $x \in [a, b]$ we have $f(x) = f(a) + \int_a^x f'(t)dt$.

the following properties

$$\operatorname{rank}\Phi(t,t_0) = n \qquad t_0, t \in \mathbb{R} \text{ arbitrary} \tag{3.10}$$
$$\Phi(t,t) = I \qquad t \in (-\infty,\infty) \tag{3.11}$$
$$\Phi^{-1}(t,t_0) = \Phi(t_0,t) \qquad t, t_0 \in (-\infty,\infty) \tag{3.12}$$
$$\Phi(t_2,t_1)\Phi(t_1,t_0) = \Phi(t_2,t_1) \qquad t_2, t_1, t_0 \in (-\infty,\infty). \tag{3.13}$$

The next theorem will give a necessary *and sufficient* controllability criterion for linear time-varying system. This result was first published in [Kalman et al., 1963]. There is a constructive proof of this theorem in [Klamka, 1991, Th. 1.3.1].

Theorem 3.9. *The dynamical system* (3.8) *with the above made regularity assumptions is globally controllable if and only if there is a time* $t_1 > t_0$ *such that the* $n \times n$ *matrix defined by*

$$W(t_0,t_1) := \int_{t_0}^{t_1} \Phi(t_1,t)B(t)B^T(t)\Phi^T(t_1,t)dt \tag{3.14}$$

is nonsingular.

Proof. Sufficiency: Let $x(t_0) \in \mathbb{R}^n$ be an arbitrary initial value of system (3.8). Suppose there exists some finite time $t_1 > t_0$ such that $W(t_0,t_1)$ is invertible. Given the above regularity assumption, the control law defined by

$$u(t) := B^T(t)\Phi^T(t_1,t)W^{-1}(t_0,t_1)(x_1 - \Phi(t_1,t_0)x(t_0)) \tag{3.15}$$

in the interval $[t_0,t_1]$ belongs to the class $L^1_{\text{loc}}([t_0,t_1],\mathbb{R})$ and steers x_0 to x_1 in time $t_1 - t_0$ along the trajectory $x(t,x(t_0),u)$, $t \in [t_0,t_1]$. This can be directly verified using solution formula (3.9):

$$\begin{aligned} x(t_1,x(t_0),u) &= \Phi(t_1,t_0)x(t_0)+ \\ &\quad + \int_{t_0}^{t_1} \Phi(t_1,t)B(t)B^T(t)\Phi^T(t_1,t)dt \cdot W^{-1}(t_0,t_1)(x_1 - \Phi(t_1,t_0)x(t_0)) \\ &= \Phi(t_1,t_0)x_0 + W(t_0,t_1)W^{-1}(t_0,t_1)(x_1 - \Phi(t_1,t_0)x_0) \\ &= x_1 \end{aligned} \tag{3.16}$$

Necessity: (by contradiction)
Suppose system (3.8) with initial value $x(t_0) = x_0$ is controllable to an arbitrary state x_1 in some finite time $t_1 > t_0$ (which does not depend on x_1). Assume that $W(t_0,t_1)$ is not invertible. By lemma (3.1) the rows of the matrix $\Phi(t_1,t)B(t)$ are linearly dependent on $[t_0,t_1]$ meaning there exists a nonzero vector $\alpha \in \mathbb{R}^n$ such that

$$\alpha^T\Phi(t_1,t)B(t) = 0 \quad \forall t \in [t_0,t_1]. \tag{3.17}$$

From the solution formula (3.9) and the latter equation we have for $x_0 = 0$ using control (3.15)

$$x_1 = \int_{t_0}^{t_1} \Phi(t_1,t)B(t)u(t)dt. \tag{3.18}$$

By means of (3.17) we obtain $\alpha^T x_1 = 0$ which leads to a contradiction to the assumption $\alpha \neq 0$ by choosing $x_1 = \alpha$ conluding the second part of the proof. □

Remark (cf. e.g. [Coron, 2007][p. 6f]): For its importance in linear control theory and its special structure the matrix $W(t_0, t_1)$ defined in (3.14) is also called *controllability Gramian* of the control system (3.8). This matrix plays an important role to steer solutions of (3.8) to a certain state as well as in stabilization theory.
Since for every $x \in \mathbb{R}^n$ we have

$$x^T W(t_0, t_1)x = \int_{t_0}^{t_1} \|B^T(t)\Phi^T(t_1, t)x\|^2 \, dt \tag{3.19}$$

the controllability Gramian is a nonnegative symmetric matrix and it is invertible if and only if there is a constant $c > 0$ such that

$$x^T W(t_0, t_1)x \geq c\|x\|^2, \; \forall x \in \mathbb{R}^n, \tag{3.20}$$

which is a strong hint that the left hand side might be helpful in finding a suitable Lyapunov function for the controlled system. We will later see that for some control inputs it actually is a Lyapunov function.

If the system matrices $A(t)$, $B(t)$ in (3.8) show more regularity, a criterion similar to that of Kalman for linear autonomous systems can be established:
Let $A(t)$ and $B(t)$ be $(n-1)$ times continuously differentiable. Then we define n matrices $M_0(t), \ldots, M_{n-1}(t)$ of size $n \times m$ as follows:

$$\begin{aligned} M_0(t) &= B(t) \\ M_{k+1}(t) &= -A(t)M_k(t) + \frac{d}{dt}M_k(t), \quad k = 0, \ldots, n-2. \end{aligned} \tag{3.21}$$

Theorem 3.10. *[Chen, 1970, Th. 5-5],[Klamka, 1991, Th. 1.3.2.]*
Assume the matrices $A(t), B(t)$ in (3.8) are $n-1$ times continuously differentiable. The dynamical system (3.8) is globally controllable if there exists some time $t_1 > t_0$ such that

$$\operatorname{rank}[M_0(t_1)|M_1(t_1)|\ldots|M_{n-1}(t_1)] = n \tag{3.22}$$

where $M_i(t)$, $i = 0, \ldots, n-1$ are the above defined matrices.

Proof. We have

$$\begin{aligned} \Phi(t_0, t)B(t) &= \Phi(t_0, t)M_0(t) \\ \frac{d}{dt}\Phi(t_0, t)B(t) &= \frac{d}{dt}\Phi^{-1}(t, t_0)B(t) \\ &= -\Phi^{-1}(t, t_0)\left(\frac{d}{dt}\Phi(t, t_0)\right)\Phi^{-1}(t, t_0)B(t) + \Phi^{-1}(t, t_0)\left(\frac{d}{dt}B(t)\right) \\ &= -\Phi^{-1}(t, t_0)A(t)\Phi(t, t_0)\Phi^{-1}(t, t_0)M_0(t) + \Phi^{-1}(t, t_0)\left(\frac{d}{dt}M_0(t)\right) \\ &= -\Phi^{-1}(t, t_0)\left(A(t)M_0(t) + \frac{d}{dt}M_0(t)\right) \\ &= \Phi(t_0, t)M_1(t). \end{aligned} \tag{3.23}$$

For higher derivatives we get

$$\frac{d^k}{dt^k}\Phi(t_0, t)B(t) = \Phi(t_0, t)M_k(t), \qquad k = 2, \ldots, n-1 \tag{3.24}$$

Therefore

$$\begin{aligned}
&\operatorname{rank}[M_0(t_1)|M_1(t_1)|\dots|M_{n-1}(t_1)] \\
&= \operatorname{rank}\Phi(t_0,t_1)[M_0(t_1)|M_1(t_1)|\dots|M_{n-1}(t_1)] \\
&= \operatorname{rank}\left[\Phi(t_0,t_1)B(t_1)|\ \frac{d}{dt}\Phi(t_0,t)B(t)\Big|_{t=t_1}|\dots|\ \frac{d^{n-1}}{dt^{n-1}}\Phi(t_0,t)B(t)\Big|_{t=t_1}\right]
\end{aligned} \tag{3.25}$$

Due to lemma (3.1) the rows of $\Phi(t_0,t)B(t)$ are linearly independent for $t \in [t_0,T], T \geq t_1$. From lemma (3.2) and theorem (3.9) the theorem then follows. □

3.1.2 Linear autonomous systems

We are given the linear autonomous (control) system

$$\frac{d}{dt}x(t) = Ax(t) + Bu(t) \tag{3.26}$$

where A and B are constant matrices of dimension $n \times n$ and $n \times m$, $x \in \mathbb{R}^n$ denotes the state vector and $u \in U \subset \mathbb{R}^m$, $m \leq n$ denotes the control. The set U is a bounded subset in $\mathbb{R}^m$ and control the control components $u_1,\dots,u_m$ are assumed to be Lebesgue integrable.
The next theorem is a direct consequence of theorem (3.10).

Theorem 3.11. *Kalman's controllability criterion [Kalman, 1960, Corollary 5.5]*
System (3.26) *is globally controllable if and only if*

$$\operatorname{rank}[B|AB|A^2B|\dots|A^{n-1}B] = n \tag{3.27}$$

Proof. Since the constant matrices A and B are infinitely many times continuously differentiable, theorem (3.10) can be applied and the matrices $M_i(t)$ reduce to A^iB for $i = 0,\dots,n-1$ and our theorem follows. □

The next theorem sometimes reduces the effort for the Kalman controllability criterion (3.11). If B is a $n \times m$ matrix of rank $r < m$ one has only to check the rank of a $n \times nr$ matrix instead of a $n \times nm$ matrix.

Theorem 3.12. *When the rank of the matrix B of the linear autonomous system* (3.26) *is $r < n$ the system is globally controllable if and only if the rank of the matrix*

$$[B|AB|A^2B|\dots|A^{n-r}B] \tag{3.28}$$

is n.

Proof. If we can show that $\operatorname{rank}[B|AB|A^2B|\dots|A^{n-r}B] = \operatorname{rank}[B|AB|\dots|A^{n-1}B]$ the theorem follows from the Kalman controllability criterion (3.11). Therefore let W_i denote the $n \times ni$ dimensional matrix $[B|AB|\dots|A^iB]$ for $i \in \mathbb{N}$. Now let us suppose for a moment that

$$\operatorname{rank} W_i = \operatorname{rank} W_{i+1}$$

for some $i \in \mathbb{N}$. Since the columns of W_i are in W_{i+1} the assumption $\operatorname{rank} W_i = \operatorname{rank} W_{i+1}$ implies that every column of the matrix $A^{i+1}B$ is linearly dependent on the columns of the matrices $B, AB, \dots, A^iB$. By induction it follows that for every $k > i+1$ the columns of the matrix A^kB are linearly dependent on the columns of $B, AB, \dots, A^iB$. Therefore we get $\operatorname{rank} W_k = \operatorname{rank} W_i$.

In other words, if there is a number $i \in \mathbb{N}$ such that $\operatorname{rank} W_i = \operatorname{rank} W_{i+1}$ then the rank of all matrices having an higher index as i does not increase any more.
Now we start with the matrix $W_0 = [B]$ which has rank r. Since the maximum rank of W_{n-1} is n it suffices to append at most $n-r$ submatrices – which are the matrices $AB, A^2B, \dots A^{n-r}B$. Therefore $\operatorname{rank} W_{n-r} = \operatorname{rank} W_{n-1}$ and theorem (3.11) can be applied to conclude the proof. □

Corollary 3.1. *If the matrix B in* (3.26) *has rank r the linear autonomous system* (3.26) *is globally controllable if and only if the $n \times n$ dimensional matrix*

$$W_{n-r} W_{n-r}^T \tag{3.29}$$

is nonsingular.

Proof. Since for any real matrix A we have $\operatorname{rank} AA^T = \operatorname{rank} A$ (this can be easily seen by proving equality of their nullspaces: $AA^T x = 0 \iff x^T AA^T x = |A^T x|^2 = 0$ and using the fact that $\operatorname{rank} A = \operatorname{rank} A^T$) the corollary is a direct consequence of theorem (3.12). □

Theorem 3.13. *System* (3.26) *is globally controllable if and only if the rows of the matrix*

$$\exp^{A \cdot t} B \tag{3.30}$$

are linearly independent on $[0, \infty)$.

Proof. For a constant matrix A the fundamental matrix is given by $\exp^{A(t-t_0)}$ (e.g. [Desoer and Zadeh, 1963, Chapter 6.2]). Using theorem (3.9) system (3.26) is controllable if and only if $\exp^{A(t-t_0)} B$ has linearly independent rows. Since $\exp^{A(t-t_0)} B$ is an analytic function for $t \in (-\infty, \infty)$ linear independence of its rows in a certain time interval is equivalent to linear independence in each time interval. □

Corollary 3.2. *The linear autonomous system* (3.26) *is globally controllable if and only if it is locally controllable.*

Proof. That the notion of local and global controllability is the same for linear autonomous systems follows directly from the proof of the latter theorem (3.13). □

Theorem 3.14. *The linear autonomous system* (3.26) *is globally controllable if and only if the rows of the matrix*

$$(sI - A)^{-1} B \tag{3.31}$$

are linearly independent on $[0, \infty)$ *over the field of complex numbers.*

Proof. Since $(sI - A)^{-1} B = \mathcal{L}(\exp^{At} B)(s)$, where $\mathcal{L}$ is the Laplace transform,[3] and s is a complex number, the theorem follows from the property of the Laplace transform as it is a one-to-one linear operator (e.g. [Engelberg, 2005, Chapter 1]). □

Theorem 3.15. *Linear feedback control equivalence*
If the control function u is a linear feedback control, i.e. can be written in the form

$$u(t) = Fx(t) + v(t) \tag{3.32}$$

[3]For a real valued function $f(t)$ the Laplace transform is defined as $\mathcal{L}(f(t)) = \int_0^\infty \exp^{st} f(t) dt$, where s is a complex number (e.g. [Engelberg, 2005, Chapter 1]).

where F is called feedback matrix, *the linear autonomous dynamical system* (3.26) *is globally controllable if and only if the linear state-feedback dynamical system*

$$\dot{x}(t) = (A + BF)x(t) + Bv(t) \tag{3.33}$$

is globally controllable with respect to the new control function v.

Proof. For a proof see for example [Chen, 1970, Chapter 7.3]. □

Controllability of Jordan-form dynamical systems

Changing the basis of the state space shouldn't affect the controllability of a linear system, which will be shown in the next theorem.

Theorem 3.16. *Invariance under linear equivalence transformations [Klamka, 1991, Lemma 1.5.1]*
The controllability of the dynamical system (3.26) *is invariant under any linear equivalence transformation $x = Tz$, where $x, z \in \mathbb{R}^n$ and T is a $n \times n$ regular matrix.*

Proof. Since T is a regular matrix, the inverse T^{-1} exists and using the the transformation $x = Tz$ we obtain from (3.26) the transformed system

$$\dot{z}(t) = T^{-1}ATz(t) + T^{-1}Bu(t). \tag{3.34}$$

Defining $J := T^{-1}AT$ and $G := T^{-1}B$ we obtain

$$\dot{z}(t) = Jz(t) + Gu(t). \tag{3.35}$$

System (3.35) is globally controllable if and only if

$$\operatorname{rank}[G|JG|\dots|J^{n-1}G] = n. \tag{3.36}$$

Because

$$\operatorname{rank}[G|JG|\dots|J^{n-1}G] = \operatorname{rank} T[B|AB|\dots|A^{n-1}B] \tag{3.37}$$

this is case if and only if the the original system (3.26) is globally controllable. □

Remark: Theorem (3.16) also holds for time-varying linear dynamical system if there is a transformation $T(t)$ which is nonsingular for all $t \in (-\infty, \infty)$.

Transforming a linear autonomous system to its Jordan canonical form will allow us to answer the question of controllability almost by inspection.
Let us assume, the dynamical system (3.26) has k distinct eigenvalues $\lambda_1, \dots, \lambda_k$ with multiplicities $n_1, \dots, n_k$, where $\sum_{i=1}^{k} n_i = n$ gives the dimension of the state space. Then there is a nonsingular transformation matrix T such that we can transform system (3.26) to its Jordan canonical form[4]

$$\dot{z}(t) = Jz(t) + Gu(t). \tag{3.38}$$

[4] which is unique except for the sequence of the Jordan blocks

where the matrices J and G can be arranged in the following way:

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}, \qquad G = \begin{pmatrix} G_1 \\ G_2 \\ \vdots \\ G_k \end{pmatrix}; \tag{3.39}$$

$$J_i = \begin{pmatrix} J_{i1} & & & \\ & J_{i2} & & \\ & & \ddots & \\ & & & J_{ir(i)} \end{pmatrix}, \qquad G_i = \begin{pmatrix} G_{i1} \\ G_{i2} \\ \vdots \\ G_{ir(i)} \end{pmatrix} \qquad i = 1, \dots, k; \tag{3.40}$$

$$J_{ij} = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & 1 & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}, \qquad G_{ij} = \begin{pmatrix} G_{ij1} \\ G_{ij2} \\ \vdots \\ G_{ijn_{ij}} \end{pmatrix} \qquad \begin{matrix} i = 1, \dots, k \\ j = 1, \dots, r(i) \end{matrix}; \tag{3.41}$$

(3.42)

	missing entries are equal to zero,
J_i	is the $n_i \times n_i$ Jordan block belonging to the eigenvalue λ_i
G_i	is the $n_i \times m$ submatrix of G corresponding to the Jordan block J_i,
$r(i)$	is the number of Jordan blocks in the submatrix J_i,
J_{ij}	are the $n_{ij} \times n_{ij}$ Jordan blocks belonging to the eigenvalue λ_i, for $i = 1, \dots, k$; $j = 1, \dots, r(i)$,
G_{ij}	are the $n_{ij} \times m$ submatrices of G corresponding to J_{ij},
$G_{ijn_{ij}}$	are the rows of G_{ij} corresponding to the rows of J_{ij},

where we have

$$n = \sum_{i=1}^{k} n_i = \sum_{i=1}^{k} \sum_{j=1}^{r(i)} n_{ij} \tag{3.43}$$

For $s \neq \lambda_i$ the inverse of $(sI - J_{ij})$ is given by

$$(sI - J_{ij})^{-1} = \begin{pmatrix} (s-\lambda_i)^{-1} & (s-\lambda_i)^{-2} & \dots & (s-\lambda_i)^{-n_{ij}} \\ 0 & (s-\lambda_i)^{-1} & \dots & (s-\lambda_i)^{-n_{ij}+1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (s-\lambda_i)^{-1} \end{pmatrix} \tag{3.44}$$

which can be directly verified.

Lemma 3.3. *[Klamka, 1991, Lemma 1.5.2],[Chen, 1970, Ch. 5.5]*
The rows of the matrix $(sI - J)^{-1}G$ are linearly independent over the field of complex numbers if and only if, for every $i = 1, \dots, k$, the rows of the matrices $(sI - J_i)^{-1}G_i$ are linearly independent over the field of complex numbers.

Proof. Since we have

$$(sI-J)^{-1}G = \begin{pmatrix} (sI-J_1)^{-1}G_1 \\ (sI-J_2)^{-1}G_2 \\ \vdots \\ (sI-J_k)^{-1}G_k \end{pmatrix} \tag{3.45}$$

$$(sI-J_i)^{-1}G_i = \begin{pmatrix} (sI-J_{i1})^{-1}G_{i1} \\ (sI-J_{i2})^{-1}G_{i2} \\ \vdots \\ (sI-J_{ir(i)})^{-1}G_{ir(i)} \end{pmatrix}; \quad i=1,\dots,k, \tag{3.46}$$

$$(sI-J_{ij})^{-1}G_{ij} = \begin{pmatrix} (s-\lambda_i)^{-1} & (s-\lambda_i)^{-2} & \dots & (s-\lambda_i)^{-n_{ij}} \\ 0 & (s-\lambda_i)^{-1} & \dots & (s-\lambda_i)^{-n_{ij}+1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (s-\lambda_i)^{-1} \end{pmatrix} \begin{pmatrix} G_{ij1} \\ G_{ij2} \\ \vdots \\ G_{ijn_{ij}} \end{pmatrix}; \quad \begin{matrix} i=1,\dots,k; \\ j=1,\dots,r(i) \end{matrix}; \tag{3.47}$$

the rows of $(sI-J_i)^{-1}G_i$ are linear combinations of $(s-\lambda_i)^{-1}$. Therefore the rows of $(sI-J_i)^{-1}G_i$ are linearly independent if and only if the rows of $(sI-J)^{-1}G$ are linearly independent. □

Theorem 3.17. *[Klamka, 1991, Th. 1.5.1],[Chen, 1970, Ch. 5.5]*
The dynamical system (3.26) *is globally controllable if and only if for each* $i=1,\dots,k$ *the rows* $G_{i1n_{i1}}, G_{i2n_{i2}}, \dots, G_{ir(i)n_{ir(i)}}$ *of the matrix* G *are linearly independent over the field of complex numbers.*

Proof. Theorem 3.14 states that (3.26) is globally controllable if and only if the rows of $(sI-J)^{-1}G$ are linearly independent over the field of complex numbers.
Necessity: From lemma (3.3) we know that the matrix $(sI-J_i)^{-1}G_i$ contains the $r(i)$ rows

$$[(s-\lambda_i)^{-1}G_{i1n_{i1}}],\ [(s-\lambda_i)^{-1}G_{i2n_{i2}}],\ \dots,\ [(s-\lambda_i)^{-1}G_{ir(i)n_{ir(i)}}]. \tag{3.48}$$

Now, if the rows $G_{i1n_{i1}}, G_{i2n_{i2}}, \dots, G_{ir(i)n_{ir(i)}}$ are not linearly independent then the rows of $(sI-J_i)^{-1}G_i$ are not linearly independent and therefore the rows of $(sI-J)^{-1}G$ cannot be linearly independent.
Sufficiency: Assume $G_{i1n_{i1}}, G_{i2n_{i2}}, \dots, G_{ir(i)n_{ir(i)}}$ are linearly independent. From (3.47) we see that the rows of $(sI-J_{ij})^{-1}G_{ij}$ depend on $G_{ijn_{ij}}$. The l-th row of $(sI-J_{ij})^{-1}G_{ij}$ contains a term of the form $(s-\lambda_i)^{-n_{ij}+l-1}G_{ijn_{ij}}$ showing the linear independence of the rows of $(sI-J_{ij})^{-1}G_{ij}$. Since $G_{i1n_{i1}}, G_{i2n_{i2}}, \dots, G_{ir(i)n_{ir(i)}}$ are linearly independent, all rows of $(sI-J_i)^{-1}G_i$ are linearly independent for all $i=1,\dots,k$ and therefore also $(sI-J)^{-1}G$ has linearly independent rows and the theorem follows. □

Corollary 3.3. *If the dynamical system* (3.26) *is globally controllable then for each* $i=1,\dots,k$ *we have*

$$r(i) \le m, \tag{3.49}$$

where m *is the number of columns of the matrix* B *in* (3.26).

Proof. This is a direct consequence of theorem 3.17 since otherwise it wouldn't be possible that $G_{i1n_{i1}}, G_{i2n_{i2}}, \dots, G_{ir(i)n_{ir(i)}}$ are linearly independent. □

From the proof of the last corollary we obtain a necessary condition for the size of the matrix B: it must have at least as many columns as the maximum number of Jordan blocks belonging to an eigenvalue of the matrix A. This gives a lower bound for underactuated systems, since m is the number of control components in the control vector u. For example a system like

$$\dot{x}(t) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} \tag{3.50}$$

is globally controllable by using a control function with a single control component. For example $B = (1, 0, 1, 0, 1)^T$ gives a completely controllable system.
But if there are at least two Jordan blocks belonging to the same eigenvalue, this is not longer the case

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}. \tag{3.51}$$

Here the matrix B must have at least two columns, for example

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.52}$$

leads to a completely controllable system.
Jordan blocks belonging to different eigenvalues can be handled by a single column of B. The lower bound of columns the matrix B must have is determined by the highest number of Jordan blocks all belonging to the same eigenvalue. This number is a necessary condition for global controllability and due to the invariance under a change of variables it is not hard to see that it is also sufficient.
From the foregoing we have the corollary which is a special case of the last corollary (3.3):

Corollary 3.4. *If* (3.26) *is globally controllable and B has size $n \times 1$ all eigenvalues of the matrix A have to belong to at most one Jordan block.*

3.2 Controllability of nonlinear systems

In this chapter we will restrict ourselves to conclusions we can make from the linearization of the system and those nonlinear systems, which have a very special form - so called control affine systems. For practical purposes this selection seems to be a good choice, as in many cases the control enters linearly. Even for this case we will omit in many cases lengthy and technical proofs but will refer to the publications where they can be found. This chapter is only intended to recall some results which are helpful in tackling the problem of controllability for nonlinear systems.

Many of the cited theorems are taken from the book [Coron, 2007] which can be seen as the " "state-of-the-art" in this area.
We consider the nonlinear control system

$$\dot{x}(t) = f(x(t), u(t)), \tag{3.53}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \Omega \subseteq \mathbb{R}^m$ the control where Ω is closed and bounded. For technical reasons we define $\hat{\Omega}$ as a bounded open set containing Ω. This is necessary in order to define $\frac{\partial f}{\partial u}$ where u is a boundary point of Ω. As an admissible control function we will assume u to be a (locally) Lebesgue integrable function, which maps from $\mathbb{R}$ into $\mathbb{R}^m$ and $m \leq n$.
In the first part, when talking about the linearization of (3.53), we will assume $f \in C^1(\mathbb{R}^n \times \hat{\Omega}, \mathbb{R}^n)$. Later in this chapter we will make stronger assumptions for the right hand side of (3.53). We will have to assume that f is complete, i.e. for every bounded admissible control and every initial value the solution exists for all times t. Furthermore it will be at least C^∞, in many cases we will even assume it to be analytic. See (3.104) for an example where this distinguishment is important.

For the notation of the solution operator of (3.53) – if it exists – we will make the following convention, which for example is used in [Brockett, 1976]: We denote the solution operator of (3.53) as $(\exp(t - t_0) f(x, u))$, or in the short form $(\exp(t - t_0) f)$.
We use the bracketing $(\exp tf)$ to distinguish the solution operator of (3.53) from the exponential $\exp(tf)$, whenever this expression is defined. In addition we will use $(\cdot)-$exponential when talking about $(\exp tf)$. If f is linear constant we have $(\exp tf) = \exp(tf)$.

3.2.1 Linearization of the nonlinear system

A first attempt to study controllability of nonlinear systems is to consider its linearization. Even if the linearization at some state $\tilde{x}$ and some control $\tilde{u}$ is *not* controllable, it may happen that the nonlinear system is locally controllable – in this case the linearization provides us with no information about the nonlinear system. A positive result was given in [Lee and Markus, 1967]: If the linearization at a stationary point of the nonlinear system is controllable, the system is locally controllable at this point.

Definition 3.18. *Linearization along a trajectory*
The linearized control system *along the trajectory $\bar{x}(t)$ corresponding to the control $\bar{u}(t)$ is given by the linear time-varying control system*

$$\dot{x}(t) = \frac{\partial f}{\partial x}(\bar{x}(t), \bar{u}(t))x(t) + \frac{\partial f}{\partial u}(\bar{x}(t), \bar{u}(t))u(t). \tag{3.54}$$

The linearization at a point in $D := (\mathbb{R}^n \times \hat{\Omega})$ is a special case, as it is the linearization along a constant trajectory $\bar{x}(t) \equiv \tilde{x}$.

Definition 3.19. *Linearization at a point*
The linearized control system *at a point $\tilde{x}$ corresponding to the control $\tilde{u}$ is given by the linear autonomous control system*

$$\dot{x}(t) = \frac{\partial f}{\partial x}(\tilde{x}, \tilde{u})x(t) + \frac{\partial f}{\partial u}(\tilde{x}, \tilde{u})u(t). \tag{3.55}$$

The following theorem is due to Lee and Markus. In [Lee and Markus, 1967] they regard the case where the zero-state is an equilibrium point of the uncontrolled system.

Theorem 3.20. *Let* $(\hat{x}, \hat{u}) \in D$ *and an equilibrium point of* (3.53)*, where* $\hat{u}$ *is a constant control. If the the linearized control system at* $(\hat{x}, \hat{u})$ *is completely controllable then the nonlinear control system* (3.53) *is locally controllable at* $(\hat{x}, \hat{u})$*.*

Proof. Since D is open and $\hat{x}$ is an equilibrium point with respect to the control $\hat{u}$, for every control $u(t)$ with $\|u(t) - \hat{u}\| < \varepsilon$, where $0 < \varepsilon \ll 1$, the solution $x(t)$ is defined at least for a small time interval. Let us assume, it is defined for $0 \leq t \leq 1$, which can be achieved by choosing ε suitably small. Defining $A := \frac{\partial f}{\partial x}(\hat{x}, \hat{u})$ and $B := \frac{\partial f}{\partial u}(\hat{x}, \hat{u})$ we know from the controllability of the linearized system, that there are admissible controls $u_1(t), \ldots, u_n(t)$ steering the system

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{3.56}$$

to n linearly independent directions.
We now introduce a new parameter $\xi = (\xi_1, \ldots, \xi_n)^T \in \mathbb{R}^n$ and define a control function

$$u(t, \xi) = \hat{u} + \sum_{i=1}^{n} \xi_i u_i(t) \tag{3.57}$$

using the reference control function $\hat{u}$ and the control functions $u_1, \ldots, u_n$ from the linearized system. Because $u_i(t)$ are assumed to be Lebesgue integrable and $[0, 1]$ is a compact subset of $\mathbb{R}$ there is a $u_{max} < \infty$ such that

$$\max_{\substack{1 \leq i \leq n \\ 0 \leq t \leq 1}} \|u_i(t)\| < u_{max}.$$

If we choose ξ such that

$$\max_{1 \leq i \leq n} |\xi_i| \leq \frac{\varepsilon}{n \cdot u_{max}},$$

we have

$$\|u(t, \xi) - \hat{u}\| = \|\sum_{i=1}^{n} \xi_i u_i(t)\| \leq \sum_{i=1}^{n} \|\xi_i u_i(t)\| \leq \frac{\varepsilon}{n \cdot u_{max}} \sum_{i=1}^{n} \|u_i(t)\|$$

and for $0 \leq t \leq 1$ we have

$$\|u(t, \xi) - \hat{u}\| \leq \frac{\varepsilon}{n \cdot u_{max}} n \cdot u_{max} = \varepsilon.$$

If we choose ε suitably small, $u(t, \xi)$ is an admissible control. Let $x(t, \xi)$ denote the corresponding solution to the control function $u(t, \xi)$. We will now show, that $x(1, \xi)$ covers an open set in $\mathbb{R}^n$ when ξ varies near zero, which will proof the theorem. Therefore consider the differentiable map

$$\xi \mapsto x(1, \xi). \tag{3.58}$$

Using the implicit function theorem all we have to show is that

$$Z(t) := \frac{\partial x}{\partial \xi}(t, \xi)\bigg|_{\xi=0} = \frac{\partial x}{\partial \xi}(t, 0) \tag{3.59}$$

is nonsingular.
Using

$$\frac{\partial x}{\partial t}(t, \xi) = f(x(t, \xi), u(t, \xi)), \quad x(0, \xi) = \hat{x} \tag{3.60}$$

and
$$\frac{\partial}{\partial \xi}\frac{\partial x}{\partial t}(t,x) = \frac{\partial f}{\partial x}(x(t,\xi),u(t,\xi))\frac{\partial x}{\partial \xi}(t,\xi) + \frac{\partial f}{\partial u}(x(t,\xi),u(t,\xi))\frac{\partial u}{\partial \xi}(t,\xi) \tag{3.61}$$
together with $x(t,0) = \hat{x}$, $u(t,0) = \hat{u}$ and (3.59) we have
$$\dot{Z}(t) = AZ(t) + B[u_1,\dots,u_n] \tag{3.62}$$
If $z_i(t)$ denotes the i-th column of the matrix $Z(t)$ we have
$$\dot{z}_i(t) = Az_i(t) + Bu_i(t), \quad z_i(0) = \hat{x} \tag{3.63}$$
Since u_i steers the the linearized system to n linearly independent directions for $0 \le t \le 1$ we have $\operatorname{span}\{z_1(1),\dots,z_n(1)\} = \mathbb{R}^n$. Therefore $Z(1) = \frac{\partial x}{\partial \xi}(1,0)$ is nonsingular and the theorem follows. □

Remark: Note that in the proof of theorem (3.20) the control functions $u_1,\dots,u_n$ for the linearized system do not have to lie in the set of admissible control functions for the nonlinear problem. They only have to be (locally) Lebesgue integrable. If we choose ξ suitably small, the composite control function $u(t,\xi)$ which will be close to the reference control function $\hat{u}$ and because D is open we obtain an admissible control u for the nonlinear problem. An important result from stabilization theory says that if the linearized system around an equilibrium point of a nonlinear system is controllable, then there is a *smooth* feedback law stabilizing this equilibrium point for the nonlinear system (cf. [Coron, 2007][Theorem 10.13, p.218]).

The following 3-dimensional system shows that local controllability of the nonlinear system can be possible although the linearization is not controllable.
$$\dot{x}(t) = \begin{pmatrix} \sin(x_3(t)) \\ \cos(x_3(t)) \\ 0 \end{pmatrix} u_1(t) + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} u_2(t) \tag{3.64}$$
This example is taken from [Nijmeijer and van der Schaft, 1990, p. 52f] and is a simplified model of driving a car (in a plane domain). The position of the car at time t is given by $(x_1(t), x_2(t)) \in \mathbb{R}^2$ and the steering angle at time t is $x_3(t)$. So the control component u_1 stands for driving the car forwards (or backwards) and the control component u_2 models the steering of the car. So it makes sense to assume that the set of admissible controls is bounded and contains the origin (as an inner point). For $(x_1, x_2, x_3, u_1, u_2) = (0,0,0,0,0)$ the linearized system is uncontrollable:
$$\dot{x}(t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}, \tag{3.65}$$
since the first component $x_1(t)$ is clearly uncontrollable. That the nonlinear system is locally controllable will be shown later. In this case the linearization misses to give information for the nonlinear model. We need to introduce some mathematical notions.

3.2.2 Vector fields, integral manifolds, Lie algebras and distributions

Controllability theory for nonlinear systems is still in progress and only for some classes results have been obtained. These results have only local character. Due to many different kinds of nonlinearities which can occur many different approaches have been made. We will restrict

ourselves here to the special case of control affine systems and the approach related to the theory of vector fields and Lie algebras.
We will begin with some mathematical preliminaries (see e.g. [Sastry, 1999]).

Definition 3.21. *vector fields in* $\mathbb{R}^n$
A vector field *is a map* $f : \mathbb{R}^n \to \mathbb{R}^n$ *which assigns every vector* x *of an open subset* $U \subset \mathbb{R}^n$ *a vector* $f(x) \in \mathbb{R}^n$*. For* $k \geq 1$ *it is called* C^k*-vector field if* f *is k-times continuously differentiable with respect to each of its components* $x_1, \ldots, x_n$*.*

Whenever we use the term *smooth vector field* we assume the vector field is continuously differentiable as many times as necessary in the context where it is used.

The Lie derivative is the derivative of a scalar function along a vector field:

Definition 3.22. *Lie derivative*
Let f *be a vector field in* $\mathbb{R}^n$ *and* $V : \mathbb{R}^n \to \mathbb{R}$ *be a scalar-valued function. Then the Lie derivative of* V *in the direction of* f *is defined as*

$$\mathcal{L}_f V := \sum_{i=1}^{n} f_i \frac{\partial V}{\partial x_i} = \langle \nabla V(x), f(x) \rangle . \tag{3.66}$$

The derivative of a vector field with respect to another vector field is given by the Lie bracket of these vector fields:

Definition 3.23. *Lie bracket*
Given to vector fields f, g *in* $\mathbb{R}^n$ *the* Lie bracket $[f, g]$ *defines a new vector field in* $\mathbb{R}^n$*. Its components are given by*

$$[f, g]_j(x) := \sum_{k=1}^{n} f_k(x) \frac{\partial g_j}{\partial x_k}(x) - g_k(x) \frac{\partial f_j}{\partial x_k}(x), \quad \forall j = 1, \ldots, n, \ \forall x \in \mathbb{R}^n \tag{3.67}$$

Definition 3.24. *Commuting vector fields*
Two vector fields f, g *in* $\mathbb{R}^n$ *are called* commuting vector fields *if their Lie bracket vanishes:*

$$[f, g] = 0.$$

Remark: Since $[g, f] = -[f, g]$ for commuting vector fields we also have $[g, f] = 0$.

Definition 3.25. $\mathbb{R}$*-algebra*
A $\mathbb{R}$-algebra *is a vector space* A *over the field* $\mathbb{R}$ *together with a bilinear operation*

$$\begin{cases} A \times A \to A \\ (x, y) \mapsto x \cdot y = xy \end{cases} \tag{3.68}$$

called multiplication such that

- $(x + y) \cdot z = x \cdot z + y \cdot z$
- $x(y + z) = x \cdot y + x \cdot z$
- $\mu \cdot (x \cdot y) = (\mu \cdot x) \cdot y = x \cdot (\mu \cdot y)$

for all $x, y, z \in A$ *and for all* $\mu \in \mathbb{R}$.

Definition 3.26. *Lie Algebra*
A Lie algebra $(\mathfrak{G}, F, \{\cdot,\cdot\})$ *is a vector space* $\mathfrak{G}$ *over some field* F *together with a binary operation*

$$\{\cdot,\cdot\} : \mathfrak{G} \times \mathfrak{G} \to \mathfrak{G} \tag{3.69}$$

such that $\{\cdot,\cdot\}$ *has the following properties*

- *bilinearity:*

$$\{ax + by, z\} = a\{x, z\} + b\{y, z\}$$
$$\{z, ax + by\} = a\{z, x\} + b\{z, y\}$$

  for all $a, b \in F$ *and all* $x, y, z \in \mathfrak{G}$.

- *antisymmetry:*

$$\{x, y\} = -\{y, x\} \quad \forall x, y \in \mathfrak{G}$$

- *Jacobi-Identity*

$$\{x, \{y, z\}\} + \{y, \{z, x\}\} + \{z, \{x, y\}\} = 0, \quad \forall x, y, z \in \mathfrak{G}$$

The space of C^∞-vector fields in $\mathbb{R}^n$ over the field of real numbers together with the Lie bracket $[\cdot,\cdot]$ is a Lie algebra. We define the *adjoint action* which is helpful in representing higher order Lie brackets:

Definition 3.27. *Adjoint Action*
Let $(\mathfrak{G}, F, \{\cdot,\cdot\})$ *be a Lie algebra. Given* $f \in \mathfrak{G}$ *one defines, by induction on* $k \in \mathbb{N}$, *the* adjoint action *of* f *on* $\mathfrak{G}$ *by*

$$\mathrm{ad}_f^0\, g = g, \tag{3.70}$$
$$\mathrm{ad}_f^{k+1}\, g = \{f, \mathrm{ad}_f^k\, g\} \tag{3.71}$$

for all $k \in \mathbb{N}$ *and all* $g \in \mathfrak{G}$.

For example if we take $f, g \in (C^\infty(\mathbb{R}^n, \mathbb{R}^n), \mathbb{R}, [\cdot,\cdot])$ we have

$$\mathrm{ad}_f^1\, g = [f, g] \tag{3.72}$$
$$\mathrm{ad}_f^2\, g = [f, [f, g]] \tag{3.73}$$
$$\mathrm{ad}_f^3\, g = [f, [f, [f, g]]]. \tag{3.74}$$

Definition 3.28. *Distribution*
Given a set of smooth vector fields $g_1, g_2, \ldots, g_m$ *we define the* distribution $\triangle$ *as*

$$\triangle = \mathrm{span}\{g_1, \ldots, g_m\}. \tag{3.75}$$

Throughout this chapter "span" is meant over the ring of smooth functions, i.e. elements of $\triangle(x)$ which denotes $\triangle$ at x are of the form

$$\alpha_1(x)g_1(x) + \alpha_2(x)g_2(x) + \ldots + \alpha_m(x)g_m(x) \tag{3.76}$$

where the $\alpha_i(x)$ are smooth scalar-valued functions of x for all $i = 1, \ldots, m$.

Definition 3.29. *Involutive Distribution*
A distribution $\triangle$ is called involutive distribution*, if for any two vector fields $g_1, g_2 \in \triangle$ their Lie bracket $[g_1, g_2]$ is also in $\triangle$.*

Definition 3.30. *Involutive Closure*
Given a distribution $\triangle$, the involutive closure $\bar{\triangle}$ *denotes the smallest involutive distribution containing $\triangle$.*

For the following two definitions let us assume we have a distribution $\triangle$ on $\mathbb{R}^n$ and a submanifold N of $\mathbb{R}^n$

Definition 3.31. *integral manifold*
The submanifold N of $\mathbb{R}^n$ is called integral manifold *of $\triangle$ if for every $x \in N$ the tangent space T_xN coincides with $\triangle(x)$.*

Definition 3.32. *maximal integral manifold*
Let N be an integral manifold of $\triangle$. It is called maximal integral manifold *if it is connected and every other connected integral manifold of $\triangle$ which contains N coincides with N.*

The following example is taken from [Sastry, 1999, p. 512f] and motivates the use of Lie brackets when talking about controllability of nonlinear systems.

Consider the system

$$\dot{x}(t) = g_1(x)u_1 + g_2(x)u_2 \tag{3.77}$$

where $u = (u_1, u_2)^T$, $u_1, u_2 : \mathbb{R} \to \mathbb{R}$ is the control vector and g_1, g_2 are smooth vector fields in $\mathbb{R}^n$. Given some initial value $x_0 \in \mathbb{R}^n$ we can go in any direction which is in

$$\text{span}\{g_1(x_0), g_2(x_0)\}. \tag{3.78}$$

As we will see we could also steer the system along the direction defined by the vector field $[g_1, g_2]$ (which in general is not in $\text{span}\{g_1(x_0), g_2(x_0)\}$). To see this we use the piecewise constant control

$$u(t) = \begin{cases} (1,0)^T & t \in [0, h[\\ (0,1)^T & t \in [h, 2h[\\ (-1,0)^T & t \in [2h, 3h[\\ (0,-1)^T & t \in [3h, 4h[\end{cases} \tag{3.79}$$

for small h and evaluate the Taylor series expansion for x about x_0 up to order 2:

$$\begin{aligned} x(h) &= x(0) + h\dot{x}(0) + \frac{1}{2}h^2\ddot{x}(0) + \dots \\ &= x_0 + hg_1(x_0) + \frac{1}{2}h^2 \left.\frac{\partial g_1(x)}{\partial x}\right|_{x=x_0} g_1(x_0) + \dots \\ &= x_0 + hg_1(x_0) + \frac{1}{2}h^2 \frac{\partial g_1(x_0)}{\partial x} g_1(x_0) + \dots \end{aligned}$$

where $\frac{\partial g_1(x)}{\partial x}$ denotes the Jacobian and we use the abbreviation $\frac{\partial g_1(x_0)}{\partial x}$ for $\left.\frac{\partial g_1(x)}{\partial x}\right|_{x=x_0}$.

$$
\begin{aligned}
x(2h) &= x(h) + h\dot{x}(h) + \frac{1}{2}h^2\ddot{x}(h) + \dots \\
&= x(h) + hg_2(x(h)) + \frac{1}{2}h^2\frac{\partial g_2(x(h))}{\partial x}g_2(x(h)) + \dots \\
&= x_0 + hg_1(x_0) + \frac{1}{2}h^2\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + \\
&\quad + hg_2(x_0 + hg_1(x_0) + \frac{1}{2}h^2\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + \dots) + \\
&\quad + \frac{1}{2}h^2\frac{\partial g_2(x_0 + hg_1(x_0) + \frac{1}{2}h^2\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + \dots)}{\partial x} \cdot \\
&\quad \cdot g_2(x_0 + hg_1(x_0) + \frac{1}{2}h^2\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + \dots) + \dots \\
&= x_0 + hg_1(x_0) + \frac{1}{2}h^2\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + h(g_2(x_0) + h\frac{\partial g_2(x_0)}{\partial x}g_1(x_0) + \dots) + \\
&\quad + \frac{1}{2}h^2\frac{\partial g_2(x_0)}{\partial x}g_2(x_0) + \dots \\
&= x_0 + h\left(g_1(x_0) + g_2(x_0)\right) + \frac{1}{2}h^2\left(\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + 2\cdot\frac{\partial g_2(x_0)}{\partial x}g_1(x_0) + \frac{\partial g_2(x_0)}{\partial x}g_2(x_0)\right) + \dots
\end{aligned}
$$

$$
\begin{aligned}
x(3h) &= x(2h) - hg_1(x(2h)) + \frac{1}{2}h^2\frac{\partial g_1(x(2h))}{\partial x}g_1(x(2h)) + \dots \\
&= x_0 + h\left(g_1(x_0) + g_2(x_0)\right) + \frac{1}{2}h^2\left(\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + 2\cdot\frac{\partial g_2(x_0)}{\partial x}g_1(x_0) + \frac{\partial g_2(x_0)}{\partial x}g_2(x_0)\right) + \\
&\quad - hg_1(x_0 + h(g_1(x_0) + g_2(x_0)) + \dots) + \frac{1}{2}h^2\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + \dots \\
&= x_0 + h\left(g_1(x_0) + g_2(x_0)\right) + \frac{1}{2}h^2\left(\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + 2\cdot\frac{\partial g_2(x_0)}{\partial x}g_1(x_0) + \frac{\partial g_2(x_0)}{\partial x}g_2(x_0)\right) + \\
&\quad - hg_1(x_0) - h^2\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) - h^2\frac{\partial g_1(x_0)}{\partial x}g_2(x_0) + \frac{1}{2}\frac{\partial g_1(x_0)}{\partial x}g_1(x_0) + \dots \\
&= x_0 + hg_2(x_0) + h^2\left(\frac{\partial g_2(x_0)}{\partial x}g_1(x_0) - \frac{\partial g_1(x_0)}{\partial x}g_2(x_0) + \frac{1}{2}\frac{\partial g_2(x_0)}{\partial x}g_2(x_0)\right) + \dots
\end{aligned}
$$

$$
\begin{aligned}
x(4h) &= x(3h) - hg_2(x(3h)) + \frac{1}{2}h^2\frac{\partial g_2(x(3h))}{\partial x}g_2(x(3h)) + \dots \\
&= x_0 + hg_2(x_0) + h^2\left(\frac{\partial g_2(x_0)}{\partial x}g_1(x_0) - \frac{\partial g_1(x_0)}{\partial x}g_2(x_0) + \frac{1}{2}\frac{\partial g_2(x_0)}{\partial x}g_2(x_0)\right) - \\
&\quad - h(g_2(x_0 + hg_2(x_0) + \dots)) + \frac{1}{2}h^2\frac{\partial g_2(x_0)}{\partial x}g_2(x_0) + \dots \qquad (3.80) \\
&= x_0 + hg_2(x_0) + h^2\left(\frac{\partial g_2(x_0)}{\partial x}g_1(x_0) - \frac{\partial g_1(x_0)}{\partial x}g_2(x_0) + \frac{1}{2}\frac{\partial g_2(x_0)}{\partial x}g_2(x_0)\right) - \\
&\quad - hg_2(x_0) - h^2\frac{\partial g_2(x_0)}{\partial x}g_2(x_0) + \frac{1}{2}h^2\frac{\partial g_2(x_0)}{\partial x}g_2(x_0) + \dots \\
&= x_0 + h^2\left(\frac{\partial g_2(x_0)}{\partial x}g_1(x_0) - \frac{\partial g_1(x_0)}{\partial x}g_2(x_0)\right) + \dots \qquad (3.81) \\
&= x_0 + h^2\left([g_1(x_0), g_2(x_0)]\right) + \dots \qquad (3.82)
\end{aligned}
$$

We return to example (3.64) – the model of driving a car – to show that this procedure actually defines a new direction, in which - at least approximately - the system can be steered. The vector fields $\begin{pmatrix}\sin x_3\\ \cos x_3\\ 0\end{pmatrix}$ and $\begin{pmatrix}0\\0\\1\end{pmatrix}$ do not commute since

$$\left[\begin{pmatrix}\sin x_3\\ \cos x_3\\ 0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right]=\begin{pmatrix}0&0&0\\0&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}\sin x_3\\ \cos x_3\\ 0\end{pmatrix}-\begin{pmatrix}0&0&\cos x_3\\0&0&-\sin x_3\\0&0&0\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}-\cos x_3\\ \sin x_3\\ 0\end{pmatrix}\neq 0. \tag{3.83}$$

We apply the control function (3.79) to obtain according to the above calculations the approximations

$$\begin{aligned}
x(h)&=\begin{pmatrix}x_1(0)\\x_2(0)\\x_3(0)\end{pmatrix}+h\begin{pmatrix}\sin x_3(0)\\ \cos x_3(0)\\ 0\end{pmatrix}\dots\\
x(2h)&=\begin{pmatrix}x_1(0)\\x_2(0)\\x_3(0)\end{pmatrix}+h\begin{pmatrix}\sin x_3(0)\\ \cos x_3(0)\\ 1\end{pmatrix}\dots\\
x(3h)&=\begin{pmatrix}x_1(0)\\x_2(0)\\x_3(0)\end{pmatrix}+h\begin{pmatrix}\sin x_3(0)-\sin(x_3(0)+h)\\ \cos x_3(0)-\cos(x_3(0)+h)\\ 1\end{pmatrix}\dots\\
x(4h)&=\begin{pmatrix}x_1(0)\\x_2(0)\\x_3(0)\end{pmatrix}+h\begin{pmatrix}\sin x_3(0)-\sin(x_3(0)+h)\\ \cos x_3(0)-\cos(x_3(0)+h)\\ 0\end{pmatrix}\dots
\end{aligned}$$

If we use a first order approximation

$$\begin{aligned}
\sin(x_3(0)+h)&\approx\sin x_3(0)+h\cos x_3(0)\\
\cos(x_3(0)+h)&\approx\cos x_3(0)+h\sin x_3(0)
\end{aligned}$$

we obtain

$$x(4h)=\begin{pmatrix}x_1(0)\\x_2(0)\\x_3(0)\end{pmatrix}+h^2\begin{pmatrix}-\cos x_3(0)\\ \sin x_3(0)\\ 0\end{pmatrix}+\ldots=\begin{pmatrix}x_1(0)\\x_2(0)\\x_3(0)\end{pmatrix}+h^2\left[\begin{pmatrix}\sin x_3(0)\\ \cos x_3(0)\\ 0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right]+\ldots= \tag{3.84}$$

$$=\begin{pmatrix}x_1(0)\\x_2(0)\\x_3(0)\end{pmatrix}+h^2\begin{pmatrix}-\cos x_3(0)\\ \sin x_3(0)\\ 0\end{pmatrix}+\ldots \tag{3.85}$$

where $\begin{pmatrix}-\cos x_3(0)\\ \sin x_3(0)\\ 0\end{pmatrix}$ which is not an element of span$\{\begin{pmatrix}\sin x_3(0)\\ \cos x_3(0)\\ 0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\}$.

The example above shows that Lie brackets can answer the controllability question in case the linearization is not controllable.

In the next section we will introduce drift-free control affine systems - to which the example above belongs - where Lie brackets are a powerful tool as they can establish a sufficient criterion for controllability.

3.2.3 Drift-free and control affine systems

We will consider control systems of the form

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} g_j(x(t))u_j(t), \quad m \leq n \tag{3.86}$$

which are called *affine* control systems. The functions $f, g_1, \ldots, g_m$ are assumed to be vector fields of $\mathbb{R}^n$ and – unless stated otherwise – are supposed to be analytic. The vector field f is called *drift vector field* or simply *drift term* whereas the g_i are referred to as *input vector fields*. As always $x \in \mathbb{R}^n$ denotes the state vector and u_i are the control functions which can be summarized to a control vector $u = (u_1, \ldots, u_m)$.

Definition 3.33. *Accessibility algebra*
Given a control system (3.86) *the* accessibility algebra $\mathcal{A}$ *is defined as the smallest Lie algebra of vector fields in* $\mathbb{R}^n$ *which contains the vector fields* $f, g_1, \ldots, g_m$.

Definition 3.34. *Accessibility distribution*
The accessibility distribution $\mathcal{C}$ *of system* (3.86) *is defined as the distribution which is generated by the elements of the accessibility algebra* $\mathcal{A}$ *of* (3.86).

Definition 3.35. *Let* $\mathrm{diff}(\mathbb{R}^n)$ *denote the group of diffeomorphisms of* $\mathbb{R}^n$. *For a given set* X *of complete vector fields in* $\mathbb{R}$ *we denote by* $\{\exp X\}$ *the smallest subgroup of* $\mathrm{diff}(\mathbb{R}^n)$ *which contains* $(\exp tf)$ *for* $f \in X$ *and* $t \in \mathbb{R}$.

The accessibility distribution is equal to $\{\exp \bar{\triangle}\}$ where $\triangle$ is generated by the vector fields appearing on the right hand side of (3.86).
Remark: Without proof we mention the fact that every element of $\mathcal{C}$ can be expressed as linear combination of terms which have the following form

$$[h_1[h_2, [\ldots [h_k, h_{k+1}] \ldots]]], \text{ with } k \in \mathbb{N},\ h_i \in \{f_1, g_1, \ldots g_m\}. \tag{3.87}$$

For a proof see for example [Nijmeijer and van der Schaft, 1990, Proposition 3.8].

Theorem 3.36. *(Version of Frobenius theorem)[Brockett, 1976, Theorem 1]*
Let $\triangle_I$ *be an involutive distribution generated by some vector fields* $h_1, \ldots, h_n$ *in* $\mathbb{R}^n$.

1. *If* $h_1, \ldots, h_n$ *are analytic on* $\mathbb{R}^n$ *– then given any point* $x_0 \in \mathbb{R}^n$ *there is a maximal submanifold* N *of* $\mathbb{R}^n$ *containing* x_0 *such that* $\triangle_I$ *spans the tangent space of* N *at each point of* N.

2. *If* $h_1, \ldots, h_n \in C^\infty$ *on* $\mathbb{R}^n$ *with constant dimension of* $\triangle_I(x)\ \forall x \in \mathbb{R}^n$ *– then given any point* $x_0 \in \mathbb{R}^n$ *there is a maximal submanifold* N *containing* x_0 *such that for every* $x \in \mathbb{R}^n$ *we have* $\triangle_I(x)$ *spans the tangent space of* N *at* x.

For a proof see for example [Sastry, 1999, Ch. 8.3].

Drift-free control systems

Throughout this subsection the drift term $f(x)$ in (3.86) vanishes and all our control systems are of the form

$$\dot{x}(t) = \sum_{i=1}^{m} g_j(x(t))u_j(t). \tag{3.88}$$

Example (3.64) of driving a car in a plane domain belongs to this class of control systems where the vector fields are $g_1(x(t)) = (\sin x_3(t), \cos x_3(t), 0)^T$ and $g_2(x(t)) = (0, 0, 1)^T$. As we have already seen, when starting in point p of the state space we can not only steer in all directions of $\text{span}\{g_1(p), g_2(p)\}$ but we can also approximately steer in the direction defined by the Lie bracket $[g_1(p), g_2(p)]$ of the two vector fields. This suggests, that the set of available directions contains the accessibility algebra $\mathcal{A}$ generated by the vector fields $g_1, \ldots, g_m$. Using more sophisticated control functions than (3.79) allows us to steer also in the directions $[g_1(x_0), [g_1(x_0), g_2(x_0)]]$ or $[[g_1(x_0), g_2(x_0)], [g_1(x_0), [g_1(x_0), g_2(x_0)]]]$ [Sastry, 1999, p. 513]. Actually these directions are already contained in the higher order terms of the Taylor expansion (3.82) [Nijmeijer and van der Schaft, 1990, p.78]. So also "brackets of brackets" and their linear combinations define available directions, suggesting, that the accessibility distribution $\mathcal{C}$ is contained in the set of available directions in the case of a driftless system.
For systems with drift term higher order brackets do not necessarily define new directions. The "intuitive" reason is that the additional drift term allows only to follow the direction $+f_0(x)$ but not the direction $-f_0(x)$. For example the two-dimensional system

$$\dot{x}(t) = \underbrace{\begin{pmatrix} x_2^2(t) \\ 0 \end{pmatrix}}_{f_0(x)} + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{f_1(x)} u(t) \tag{3.89}$$

is nondecreasing in its first component, meaning that $-(1, 0)^T$ for example is not an available direction to steer the system to. But the iterated Lie bracket

$$[[f_1, f_0], f_1](x) = -\begin{pmatrix} 2 \\ 0 \end{pmatrix} \tag{3.90}$$

suggests exactly this direction.

The accessibility distribution $\mathcal{C} = \{\exp \bar{\triangle}\}$ seems to be much larger than $\{\exp \triangle\}$. It is an astonishing result of Chow that for sufficiently smooth vector fields these sets are actually equal which we will restate here without proof:

Theorem 3.37. *(Version of Chow's theorem)[Brockett, 1976, Theorem 2]*
Let $\triangle$ be a distribution and $\bar{\triangle}$ its involutive closure.

1. *If the elements of $\bar{\triangle}$ are analytic on $\mathbb{R}^n$ – then given any point $x_0 \in \mathbb{R}^n$ there is a maximal submanifold N of $\mathbb{R}^n$ containing x_0 such that $N = \{\exp \triangle\}x_0 = \{\exp \bar{\triangle}\}x_0$.*

2. *If the elements of $\bar{\triangle}$ are C^∞ on $\mathbb{R}^n$ with $\dim(\bar{\triangle})$ constant on $\mathbb{R}^n$ – then given any point $x_0 \in \mathbb{R}^n$ there is a maximal submanifold N containing x_0 such that $N = \{\exp \triangle\}x_0 = \{\exp \bar{\triangle}\}x_0$.*

From the theorem of Chow we can now deduce a controllability rank criterion for local controllability:

Theorem 3.38. *Controllability rank condition for drift-free control systems*
If for some state $x_0 \in \mathbb{R}^n$ the accessibility distribution $\mathcal{C}$ of (3.88) at x_0 has dimension n , then (3.88) is locally controllable in x_0.

Proof. Chow's theorem. □

Control systems with drift term

Control affine systems have the special form (3.86)

$$\dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} g_i(x(t))u_i(t).$$

They can be regarded as a special case of drift-free control systems in the following sense: Maybe in a drift-free control system there is a failure of a control component, say $u_{m+1}(t)$, which for example could freeze and send a constant output signal u_{m+1}. Then we can interpret the term $g_{m+1}(x(t))u_{m+1}(t) = g_{m+1}(x(t))u_{m+1}$ as drift term $f(x(t))$

$$\dot{x}(t) = \sum_{i=1}^{m+1} g_i(x(t))u_i(t) = \sum_{i=1}^{m} g_i(x(t))u_i(t) + \underbrace{(g_{m+1}(x(t))u_{m+1})}_{f(x(t))} \tag{3.91}$$

which results in a system of type (3.86).

The following theorem gives a sufficient criterion for small-time local controllability for nonlinear control affine systems. It will turn out, that this theorem will be a generalization of the Kalman criterion (3.11) of the linear control problem (3.26). We consider the control affine problem (3.86) for $t \geq 0$:

Theorem 3.39. *(cf. [Coron, 2007][pp. 131-133])*
Consider

$$\dot{x}(t) = F(x,u) = f(x(t)) + \sum_{i=1}^{m} g_i(x(t))u_i(t) \tag{3.92}$$

where u_i are scalar input functions, f, g_i are analytic functions. Let $x = a$ be an equilibrium point of the uncontrolled problem, i.e. $F(a,0) = f(a) = 0$. Then, if

$$\operatorname{span}\{\operatorname{ad}_f^k g_j(a);\ k \in \mathbb{N},\ j \in \{1,\dots,m\}\} = \mathbb{R}^n \tag{3.93}$$

holds the control affine system is small-time locally controllable at $x = a$.

We will omit the proof for this theorem and refer to Halina Frankowska who proofed this result in [Frankowska, 2005] by using the Brouwer fixed point theorem in a more general setting regarding differential inclusions.
Instead we will show that the directions defined by $\operatorname{ad}_f^k g_j(a);\ k \in \mathbb{N},\ j \in \{1,\dots,m\}$ are in fact available directions for the control affine problem (3.86):
Fix $j \in \{1,\dots,m\}$ and define an auxiliary function Φ such that

$$\Phi \in C^k([0,1]) \tag{3.94}$$

$$\Phi^{(l)}(0) = \Phi^{(l)}(1) = 0, \quad \forall l \in \{0,\dots,k-1\}. \tag{3.95}$$

For $\eta \in (0,1]$ and $\varepsilon \in [0,1]$ define $u(t) : [0,\eta] \to \mathbb{R}^m$ by choosing its components

$$u_i := \begin{cases} [0,\eta] \to \mathbb{R}^n \\ t \mapsto \delta_{ij} \cdot \varepsilon \cdot \Phi^{(k)}(\frac{t}{\eta}), \end{cases} \tag{3.96}$$

where $\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$ is the Kronecker delta. Defining $A := \frac{\partial F}{\partial x}(a)$ and $B := \frac{\partial F}{\partial u}(a, 0) = (g_1(a), \ldots, g_m(a))$ we consider the two initial value problems

$$\dot{x}(t) = F(x, u) = f(x(t)) + \sum_{i=1}^{m} g_i(x(t))u_i(t), \quad x(0) = a \tag{3.97}$$

$$\dot{y}(t) = Ay(t) + Bu(t), \quad y(0) = 0. \tag{3.98}$$

Using Gronwall's lemma (4.1) there is some $C > 0$ and $\varepsilon_0 > 0$ such that for all $\varepsilon \in [0, \varepsilon_0]$ and $\eta \in (0, 1]$ we obtain

$$|x(t) - a - y(t)| = |\int_0^t f(x(\tau), u(\tau)) - Ay(\tau) - Bu(\tau)d\tau| \leq C\varepsilon^2. \tag{3.99}$$

For the solution of the linear system (3.98) we obtain by integration by parts

$$\begin{aligned} y(\eta) &= \int_0^\eta \mathrm{e}^{(\eta-t)A} Bu(t)\, dt \\ &= \int_0^\eta \mathrm{e}^{(\eta-t)A} \varepsilon \cdot \Phi^{(k)}(\frac{t}{\eta}) \cdot g_j(a)\, dt \\ &= -\varepsilon\eta \int_0^\eta \mathrm{e}^{(\eta-t)} A \cdot \Phi^{(k-1)}(\frac{t}{\eta}) dt \\ &\vdots \\ &= (-1)^k \eta^k \varepsilon \int_0^\eta \mathrm{e}^{(\eta-t)A} A^k \Phi(\frac{t}{\eta}) g_j(a)\, dt \\ &= (-1)^k \eta^{k+1} \varepsilon \int_0^1 \Phi(\tau) \mathrm{e}^{\eta(1-\tau)A} A^k g_j(a)\, d\tau \end{aligned} \tag{3.100}$$

Hence, for some finite constant D sufficiently large but independent of $\varepsilon \geq 0$ and $\eta \in (0, 1]$ we have

$$\left| y(\eta) - (-1)^k A^k g_j(a) \eta^{k+1} \varepsilon \int_0^1 \Phi(\tau)\, d\tau \right| \leq D\varepsilon\eta^{k+2}. \tag{3.101}$$

From the definition of the adjoint action (3.27) one easily obtains the formula

$$\operatorname{ad}_f^k g_j(a) = (-1)^k A^k g_j(a) \tag{3.102}$$

which together with the estimates (3.99) and (3.101) gives

$$x(\eta) \approx a + \varepsilon \cdot \eta^{k+1} (-1)^k \operatorname{ad}_f^k g_j(a) \int_0^1 \Phi(\tau)\, d\tau, \tag{3.103}$$

which shows, that one can actually move in the direction of $(-1)^k \operatorname{ad}_f^k g_j(a)$. Also the direction $-(-1)^k \operatorname{ad}_f^k g_j(a)$ is available, simply by changing the sign of the auxiliary function Φ.
In the linear case condition (3.93) above reduces to the Kalman controllability criterium (3.11). A necessary condition for analytic right hand sides of (3.86) (or even general nonlinear systems) is due to [Nagano, 1966] can be formulated as follows: If the (control affine/general nonlinear) system is small-time controllable at $x = x_0$ then the corresponding accessibility distribution has rank n at $x = x_0$ (which often is called the *Lie algebra condition*. Note that for this necessary

condition analyticity is crucial. A one-dimensional counterexample if this condition does not hold is given by

$$\dot{x}(t) = f(x(t), u(t)) : \begin{cases} \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\ (x(t), u(t)) \mapsto \begin{cases} u(t)\mathrm{e}^{-\frac{1}{u^2(t)}} & \text{for } u(t) \neq 0 \\ 0 & \text{for } u(t) \end{cases} \end{cases} \tag{3.104}$$

The right hand side is infinitely many times differentiable. But although it is controllable at $x = 0$ the accessibility distribution at this point has rank 0, such that analyticity is crucial for Nagano's theorem.
Remark: Nagano's theorem is important as a necessary condition for nonlinear analytic systems. But there are important cases where with less assumptions the Lie algebra condition is (although not necessary) a sufficient condition establishing (small-time) controllability. For linear autonomous systems for example once more we rediscover the Kalman condition (3.11) for controllability. For driftless control systems with vector fields in C^∞, we already saw from Chow's theorem (3.37), that the Lie algebra condition turns out to be sufficient.
As seen in theorem (3.39) there are Lie-brackets which are useful in determining controllability properties, but example (3.89) shows, that there are also Lie-brackets, which are not helpful in determining controllability properties or even worse they can be an obstruction to controllability. As we have seen, Lie-brackets naturally appear as limiting directions and therefore the effects of "bad brackets" have to be neutralized in some way. A natural question is, which of the Lie-brackets are "good ones" and which of them not. A more ambitious question then is to ask how the effects of the "bad brackets" can possibly be healed. This interplay of "good" and "bad" brackets has not been fully understood. It is still an open problem and even the scalar input case ($m = 1$) seems to be quite difficult.
Let us consider this special case where $m = 1$ in (3.86):

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t), \quad t \geq 0. \tag{3.105}$$

We assume that f and g are smooth vector fields and that $|u| \leq 1$.
Let $x = x_0$ be an equilibrium point of the uncontrolled system, i.e. $f(x_0) = 0$. Henry Hermes conjectured in [Hermes, 1976] that Lie-brackets with an even number of $g's$ are "bad ones". Let $\mathcal{S}^k(f, g)$ denote the linear span of all brackets built of the vector fields f and g such that the number of $g's$ is at most k. Then the Hermes conjecture which was proofed by Sussmann 1983 can be formulated as theorem:

Theorem 3.40. *Hermes-Sussmann*
Regard system (3.105)*, where $x \in \mathbb{R}^n$ and f, g are smooth vector fields mapping into $\mathbb{R}^n$. Let $x_0 \in \mathbb{R}^n$ be such that $f(x_0) = 0$ and $\mathcal{S}^k(f, g)(x_0) = \mathbb{R}^n$ for some $k \geq 1$ and condition*

$$\mathcal{S}^{2p}(f, g)(x_0) = \mathcal{S}^{2p+1}(f, g)(x_0) \tag{3.106}$$

holds for every natural p. Then (3.105) *is small-time locally controllable at x_0.*

Unfortunately this condition is not necessary, as could be shown by Jacubczyck in [Sussmann, 1987]. For the proof which is quite lengthy we refer to the original work [Sussmann, 1987] where also the Jacubczyck example was presented. An outline of the proof, which still is quite lengthy, can be found in [Sussmann, 1983a]. The main idea was approximating the original nonlinear system by systems containing Lie-brackets to obtain more information than one would get from the linearization of the system. As seen in the motivating example control variations play an important role. When studying the solutions of the approximating system, the Baker-Campbell-Hausdorff formula plays an important role and Gianna Stefani conjectured in [Stefani, 1985] that symmetries in this formula have to be studied to identify the real bad brackets. Real bad brackets, because we already know that condition (3.106) is not necessary and Stefani could show in the same paper, that not all brackets with an even number of $g's$ are obstructions to controllability. In this paper a 3-dimensional example of type (3.105) was regarded, where the local controllability could be established with the help of Lie-brackets containing the term g four times.

Chapter 4

Stability and stabilizing control laws

In this chapter we introduce an algorithm which stabilizes a nonlinear system along a given reference trajectory. This is a local method and uses the (controllable) linearization along this trajectory. To motivate the control law we will first give some results from the linear theory. We start with stability notions, where we mainly follow the presentation in [Sastry, 1999, p. 86ff]:

4.1 Stability notions

Given a general nonlinear differential equation

$$\dot{x}(t) = f(x(t), t), \quad x(t_0) = x_0,\ t \geq t_0 \tag{4.1}$$

where $x \in \mathbb{R}^n$ and unless stated otherwise, f is $C^\infty(\mathbb{R}^n, \mathbb{R}^n)$. We assume there is an equilibrium point x_e. Without loss of generality we may assume that this equilibrium point is the origin $x = 0$ of the state space (which can always be achieved by a suitable change of variables).

Lemma 4.1. *(cf. e.g. [Sastry, 1999][p. 86]) Bellman-Gronwall Lemma*
Let $z(\cdot), a(\cdot), u(\cdot) : \mathbb{R}_0^+ \to \mathbb{R}^+$ be given positive functions and $T > 0$. Then, if for all $t \leq T$ we have

$$z(t) \leq u(t) + \int_0^t a(\tau) z(\tau) d\tau, \tag{4.2}$$

if follows that for $t \in [0, T]$ the following inequality holds:

$$z(t) \leq u(t) + \int_0^t a(\tau) u(\tau) \mathrm{e}^{\int_\tau^t a(\sigma) d\sigma} d\tau. \tag{4.3}$$

Proof. Define

$$r(t) := \int_0^t a(\tau) z(\tau) d\tau,$$

differentiating and using (4.2) yields

$$\dot{r}(t) = a(t) z(t) \leq a(t) u(t) + a(t) r(t)$$

which means that for some positive function $s(t)$ we have

$$\dot{r}(t) = a(t) u(t) + a(t) r(t) - s(t).$$

This is an inhomogeneous linear differential equation and solving it with initial condition $r(0) = 0$ yields

$$r(t) = \mathrm{e}^{\int_0^t a(\sigma)d\sigma} \cdot \int_0^t \mathrm{e}^{-\int_0^\tau a(\sigma)d\sigma}(a(\tau)u(\tau) - s(\tau))d\tau = \int_0^t \mathrm{e}^{\int_\tau^t a(\sigma)d\sigma} a(\tau)u(\tau)d\tau - \int_0^t \mathrm{e}^{\int_\tau^t a(\sigma)d\sigma} s(\tau)d\tau$$

Since $s \geq 0$ and the exponential is positive, inequality (4.3) follows concluding the proof. □

Proposition 4.1. *(cf. [Sastry, 1999]) Rate of convergence (growth/decay rate)*
Regard system (4.1) *and assume that f is Lipschitz continuous in x with Lipschitz constant L and piecewise constant with respect to t. Further assume that $x = 0$ is an equilibrium state of the uncontrolled system, then - as long as $x(t)$ remains in a ball around the equilibrium point $x = 0$ - the solution $x(t)$ satisfies*

$$\|x(t)\| \leq \|x_0\| \mathrm{e}^{L(t-t_0)}. \tag{4.4}$$

Proof. Since $\tilde{x}(t) \equiv 0$ is a trivial solution for the initial value problem with $\tilde{x}(t_0) = 0 =: \tilde{x}_0$ we have

$$\begin{aligned} \|x(t) - \tilde{x}(t)\| &\leq \|x_0 - \tilde{x}_0\| + \int_{t_0}^t \|f(x(\tau), \tau) - f(\tilde{x}(\tau), \tau)\| d\tau \\ &\leq \|x_0 - \tilde{x}_0\| + L \int_{t_0}^t \|x(\tau) - \tilde{x}(\tau)\| d\tau \\ \|x(t)\| &\leq \|x_0\| + L \int_{t_0}^t \|x(\tau)\| d\tau \end{aligned}$$

Applying the Bellman-Gronwall Lemma (4.1) with $a(t) \equiv \|x_0\|, z(t) = \|x(t)\|, u(t) \equiv L$ leads to the desired inequality whenever $\|x_0\| \neq 0$. For $\|x_0\| = 0$ inequality (4.4) trivially holds since then $x(t) \equiv 0$. □

Remark: The Lipschitz constant L serves as growth rate if positive and decay rate if negative.

Definition 4.1. *[Sastry, 1999][Def. 5.4] Stability in the sense of Lyapunov*
The equilibrium point $x = 0$ is called stable equilibrium point of (4.1) *(in the sense of Lyapunov) if for all $t_0 \in \mathbb{R}$ and $\varepsilon > 0$, there exists $\delta(\varepsilon, t_0)$ such that*

$$\|x_0\| < \delta(\varepsilon, t_0) \Longrightarrow \|x(t)\| < \varepsilon \quad \forall t \geq t_0, \tag{4.5}$$

where $x(t)$ is the solution of (4.1) *with initial value $x(t_0) = x_0$.*

Definition 4.2. *[Sastry, 1999][Def. 5.5] Uniform stability*
The equilibrium point $x = 0$ is called a uniformly stable equilibrium point of (4.1) *if in the preceding definition (4.1) δ can be chosen independent of t_0.*

Definition 4.3. *[Sastry, 1999][Def. 5.6] Asymptotic stability*
The equilibrium point $x = 0$ is an asymptotically stable equilibrium point of (4.1) *if*

- $x = 0$ *is a stable equilibrium point of* (4.1),
- $x = 0$ *is attractive, i.e. for all $t_0 \in \mathbb{R}$ there exists a $\delta(t_0)$ such that*

$$\|x_0\| < \delta(t_0) \Longrightarrow \lim_{t \to \infty} \|x(t)\| = 0,$$

  where $x(t)$ is the solution of (4.1) *with initial value $x(t_0) = x_0$.*

Definition 4.4. *[Sastry, 1999][Def. 5.10] Exponential stability*
The equilibrium point $x = 0$ *is an exponentially stable equilibrium point of* (4.1) *if there exist* $m, \alpha > 0$ *such that for the solution of* (4.1) *with initial value* $x(t_0) = x_0$ *we have*

$$\|x(t)\| \leq m\|x_0\|e^{-\alpha(t-t_0)} \tag{4.6}$$

for all x_0 *in an environment of* 0*. The constant* α *is called the* rate of convergence *or* decay rate*.*

4.1.1 Stability of linear time-varying systems

We consider the linear time-varying system

$$\dot{x}(t) = A(t)x(t), \quad x(t_0) = x_0,\ t \geq t_0 \tag{4.7}$$

where $A(t) \in \mathbb{R}^{n\times n}$ is a piecewise continuous bounded function. As usual we denote the state transition matrix as $\Phi(t, t_0)$.

Theorem 4.5. *(cf. [Sastry, 1999][Th. 5.32]) Stability of linear systems*
The right-hand-side of the following table gives the stability conclusions of the equilibrium point $x = 0$ *of the linear time-varying system* (4.7)*:*

	Conditions on $\mathbf{\Phi(t, t_0)}$	***Stability conclusions***
1.	$\sup_{t\geq t_0} \|\Phi(t, t_0)\| < M(t_0) < \infty$	*stable*
2.	$\sup_{t_0} \sup_{t\geq t_0} \|\Phi(t, t_0)\| < \infty$	*uniformly stable*
3.	$\lim_{t\to\infty} \|\Phi(t, t_0)\| = 0$	*asymptotically stable*

where $\|\Phi(t, t_0)\| = \max\{\|\Phi(t, t_0)x\| : x \in \mathbb{R}^n, \|x\| = 1\}$*.*

Proof. 1. Assume that

$$\|\Phi(t, t_0)\| < M(t_0) < \infty \quad \forall t \geq t_0. \tag{4.8}$$

We therefore have

$$\|x(t)\| = \|\Phi(t, t_0)x_0\| \leq \|\Phi(t, t_0)\|\|x_0\| \leq M(t_0)\|x_0\| \quad \forall t \geq t_0. \tag{4.9}$$

Thus, given an arbitrary $\varepsilon > 0$ we have for $\delta := \varepsilon/M(t_0)$

$$\|x_0\| < \delta = \varepsilon/M(t_0) \Longrightarrow \|x(t)\| \leq M(t_0)\|x_0\| = M(t_0)\frac{\varepsilon}{M(t_0)} = \varepsilon. \tag{4.10}$$

showing stability of the equilibrium point $x = 0$.
Suppose now that (4.8) does not hold. I.e. there is at least one element in $\Phi(t, t_0)$ which in absolute value takes on arbitrarily large values. Without loss of generality let us assume, this element is $\Phi(t, t_0)_{ik}$, $1 \leq i, k \leq n$. Choose the vector e_k - which has zeros everywhere except the $k-$th entry which is one - as initial vector x_0. Then the $i-$th component of the state vector x at time $t \geq t_0$ is given by

$$x_i(t) = \Phi(t, t_0)_{ik} \tag{4.11}$$

and - as $\Phi(t, t_0)$ was assumed to be unbounded, so is x_i, showing that the equilibrium state $x = 0$ is unstable. Thus stability implies uniform boundedness of $\|\Phi(t, t_0)\|$.

2. (cf. e.g. [DaCunha, 2005]) We assume that $x = 0$ is a uniformly stable equilibrium point of (4.7). Then there is a $M > 0$ such that for any t_0 and $x(t_0)$ we have

$$\|x(t)\| \leq M\|x(t_0)\|, \quad t \geq t_0. \tag{4.12}$$

Given any t_0 and $t^* \geq t_0$, we can choose a state x^* such that

$$\|x^*\| = 1, \quad \|\Phi(t^*, t_0)x^*\| = \|\Phi(t^*, t_0)\|\|x^*\| = \|\Phi(t^*, t_0)\| \tag{4.13}$$

Such a state always exists since $\|\Phi(t^*, t_0)\| = \max\{\Phi(t^*, t_0)x, \ x \in \mathbb{R}^n, \ \|x\| = 1\}$.
Now we apply (4.12) to the solution of (4.7) at time t^* with initial state $x_0 = x^*$, which gives

$$\begin{aligned} \|x(t^*)\| = \|\Phi(t^*, t_0)x^*\| &= \|\Phi(t^*, t_0)\|\|x^*\| \leq M\|x^*\| \\ \Longrightarrow \quad \|\Phi(t^*, t_0)\| &\leq M \end{aligned} \tag{4.14}$$

showing one direction.
Now suppose that there is $M > 0$ such that $\sup_{t_0} \sup_{t \geq t_0} \|\Phi(t, t_0)\| < M$. For any t_0 and $x(t_0) = x_0$ we have

$$\|x(t)\| = \|\Phi(t, t_0)x_0\| \leq \|\Phi(t, t_0)\|\|x_0\| \leq M\|x_0\|, \quad t \geq t_0 \tag{4.15}$$

which shows uniform stability.

3. Assume that $\lim_{t\to\infty} \|\Phi(t, t_0)\| = 0$ holds. Due to continuity we have $\|\Phi(t, t_0)\| < M(t_0)$ showing stability using 1. Moreover we have

$$\|x(t)\| = \|\Phi(t, t_0)x_0\| \leq \|\Phi(t, t_0)\|\|x_0\| \overset{t\to\infty}{\longrightarrow} 0 \Longrightarrow \lim_{t\to\infty} \|x(t)\| = 0, \quad t \geq t_0. \tag{4.16}$$

showing asymptotic stability.
Now assume $x = 0$ is an asymptotic stable equilibrium point of (4.7). Then there exists a x_0 with $\|x_0\| = 1$ such that $\|\Phi(t, t_0)x_0\| = \|\Phi(t, t_0)\|$. Choose a basis $\{z^{(i)}\}_{1\leq i\leq n}$ of $\mathbb{R}^n$ such that $\|z^{(i)}\| < \delta(t_0)$ $\forall i \in \{1, \ldots, n\}$. Then there are $\xi_1, \ldots \xi_n \in \mathbb{R}^n$ such that $\sum_{i=1}^n \xi_i z^{(i)} = x_0$. It follows that

$$\begin{aligned} \|\Phi(t, t_0)\| &= \|\Phi(t, t_0)x_0\| \\ &= \|\Phi(t, t_0)\left(\sum_{i=1}^n \xi_i z^{(i)}\right)\| \\ &= \|\sum_{i=1}^n \xi_i \Phi(t, t_0)z^{(i)}\| \\ &\leq n \cdot \max\{|\xi_1|, \ldots, |\xi_n|\} \cdot \max_{1\leq i\leq n}\{\|\Phi(t, t_0)z^{(i)}\|\} \end{aligned}$$

Since $\lim_{t\to\infty} \max_{1\leq i\leq n}\{\|\Phi(t, t_0)z^{(i)}\|\} = 0$ due to asymptotic stability, we have $\|\Phi(t, t_0)\| = 0$ concluding the proof.

□

The next theorem shows that for a linear time-varying system uniform asymptotic stability and exponential stability are the same:

Theorem 4.6. *Exponential and uniform asymptotic stability*
The point $x = 0$ is a uniform exponentially stable equilibrium point of (4.7) *if and only if $x = 0$ is an exponentially stable equilibrium point of* (4.7).

Proof. The equilibrium point 0 of (4.7) is uniformly asymptotically stable if it is uniformly stable and $\|\Phi(t, t_0)\| \Longrightarrow 0$ as $t \to \infty$, uniformly in t_0. The direction

$$\text{exponential stability} \implies \text{uniform asymptotic stability}$$

follows from the definition of the stability notions. To show the other direction we start with the assumption of uniform asymptotic stability. I.e. $\forall t_1$ there exist m_0, T such that

$$\Phi(t, t_0) \leq m_0, \quad t \geq t_1. \tag{4.17}$$

Uniform convergence of $\Phi(t, t_0)$ to 0 implies that

$$\|\Phi(t, t_1)\| \leq \frac{1}{2}, \quad \forall t \geq t_1 + T. \tag{4.18}$$

Given any t, t_0 choose k such that

$$t_0 + kT \leq t \leq t_0 + (k+1)T. \tag{4.19}$$

We have

$$\begin{aligned}
\|\Phi(t, t_0)\| &= \|\Phi(t, t_0 + kT) \prod_{j=1}^{k} \Phi(t_0 + (k+1-j)T, t_0 + j)\| \\
&\leq \|\Phi(t, t_0 + kT)\| \prod_{j=1}^{k} \|\Phi(t_0 + (k+1-j)T, t_0 + j)\| \\
&\leq m_0 2^{-k} \leq 2m_0 2^{\frac{-(t-t_0)}{T}} \leq 2m_0 \mathrm{e}^{-\log \frac{2}{T}(t-t_0)}
\end{aligned}$$

showing that $x = 0$ is an exponentially stable equilibrium point of (4.7). □

To apply the previous two theorems (4.5) and (4.6) an estimation for the norm of the state transition matrix is needed. If upper and lower bounds for the (time-varying) eigenvalues of the symmetric part of $A(t)$ can be found, these can be used to estimate the norm of the state transition matrix. We sketch this method without details following [Conti, 1976].
Let S denote the symmetric part of $A(t)$ defined by

$$S(t) := \frac{1}{2}(A(t) + A^T(t)). \tag{4.20}$$

For any solution $x(t)$ of (4.7) we have

$$\frac{1}{2}\frac{d}{dt} x^T(t)x(t) = x^T(t)S(t)x(t). \tag{4.21}$$

For $s, t \geq t_0$ we obtain

$$x^T(t)x(t) - x^T(s)x(s) = 2\int_s^t x^T(\tau)S(\tau)x(\tau)d\tau \tag{4.22}$$

such that

$$\Phi^T(t,s)\Phi(t,s) - I = 2\int_s^t \Phi^T(\tau,s)S(\tau)\Phi(\tau,s)d\tau. \tag{4.23}$$

Since $S(t)$ is a symmetric matrix with entries in $\mathbb{R}$ for every $t \geq t_0$ we obtain a minimum eigenvalue $\lambda(t) \in \mathbb{R}$ and a maximal eigenvalue $\mu(t) \in \mathbb{R}$ for every $t \in \mathbb{R}, t_0 \leq t$ such that for every vector v we have

$$\lambda(t)v^T v \leq v^T S(t) v \leq \mu(t) v^T v, \quad t \geq t_0. \tag{4.24}$$

From (4.21) and (4.24) we obtain the Wintner-Wasžewski inequality

$$\mathrm{e}^{\left(2\int_s^t \lambda(\tau)d\tau\right)} x^T(s)x(s) \leq x^T(t)x(t) \leq \mathrm{e}^{\left(2\int_s^t \mu(\tau)d\tau\right)} x^T(s)x(s), \quad t_0 \leq s \leq t \tag{4.25}$$

where $x(t)$ is a solution of (4.7). For an arbitrary vector $v \in \mathbb{R}^n$ we therefore have

$$\mathrm{e}^{\left(2\int_s^t \lambda(\tau)d\tau\right)} v^T v \leq v^T\Phi^T(t,s)\Phi(t,s)v \leq \mathrm{e}^{\left(2\int_s^t \mu(\tau)d\tau\right)} v^T v, \quad t_0 \leq s \leq t \tag{4.26}$$

which is equivalent to

$$\mathrm{e}^{-\left(2\int_s^t \mu(\tau)d\tau\right)} v^T v \leq v^T\Phi^T(s,t)\Phi(s,t)v \leq \mathrm{e}^{-\left(2\int_s^t \lambda(\tau)d\tau\right)} v^T v, \quad t_0 \leq s \leq t \tag{4.27}$$

Since $\|S(t)\| \leq \|A(t)\|$ which follows from the triangle inequality and the definition of S we obtain for $v \in \mathbb{R}^n$

$$\begin{aligned}
\mathrm{e}^{\left(-2\int_s^t \|A(\tau)\|d\tau\right)} v^T v &\leq \mathrm{e}^{\left(-2\int_s^t \|S(\tau)\|d\tau\right)} v^T v \\
&\leq v^T\Phi^T(t,s)\Phi(t,s)v \cdot v^T\Phi^T(s,t)\Phi(s,t)v \\
&\leq \mathrm{e}^{\left(2\int_s^t \|S(\tau)\|d\tau\right)} v^T v \\
&\leq \mathrm{e}^{\left(2\int_s^t \|A(\tau)\|d\tau\right)} v^T v, \quad t_0 \leq s \leq t.
\end{aligned} \tag{4.28}$$

For the state transition matrix we have now the inequality

$$\mathrm{e}^{\left(\int_s^t \lambda(\tau)d\tau\right)} \leq \|\Phi(t,s)\| \leq \mathrm{e}^{\left(\int_s^t \mu(\tau)d\tau\right)}, \quad t_0 \leq s \leq t. \tag{4.29}$$

And if there are $\lambda, \mu < \infty$ such that

$$\lambda \leq \lambda(t) \leq \mu(t) \leq \mu, \quad t_0 \leq t \tag{4.30}$$

this reduces to

$$\mathrm{e}^{\lambda(t-s)} \leq \|\Phi(t,s)\| \leq \mathrm{e}^{\mu(t-s)}, \quad t_0 \leq s \leq t. \tag{4.31}$$

For continuity reasons such upper and lower bounds exist for every finite interval $[s,t]$ with $t_0 \leq s \leq t$.

The following theorem contains a method to show stability of an equilibrium point without making use of the fundamental solution. It is often called Lyapunov's second method. We need the following definition:

Definition 4.7. *A symmetric matrix $C(t) \in \mathbb{R}^{n\times n}$ is* positive definite*, denoted as $C(t) > 0$, if for each $t \in \mathbb{R}$, $x^T C(t) x > 0 \, \forall x \in \mathbb{R}^n \setminus \{0\}$.*
A symmetric matrix $C(t) = C^T(t) \in \mathbb{R}^{n\times n}$ is uniformly positive definite*, if $C(t) - \alpha I$ is positive definite for some constant $\alpha > 0$ (denoted as $C(t) - aI > 0$ or simply $C(t) > aI$).*

Remark: Constant positive definite matrices are always uniformly positive definite which we will show in lemma (4.4).

Theorem 4.8.
Assume that $A(t)$, $t \geq t_0$ is bounded. If for some $\alpha > 0$ there is a $Q(t) \geq \alpha I \quad \forall t \geq t_0$ such that

$$\mathbb{R}^{n\times n} \ni P(t) := \int_t^\infty \Phi^T(\tau,t)Q(\tau)\Phi(\tau,t)d\tau, \quad t \geq t_0 \tag{4.32}$$

is bounded, then the origin is an uniformly asymptotically stable equilibrium point of (4.7).

Proof. We can show that $P(t)$ is uniformly positive definite for $t \geq t_0$, i.e. there exists $\beta > 0$ such that

$$\beta x^T x \leq x^T P(t)x, \quad \forall x \in \mathbb{R}^n, t \geq t_0. \tag{4.33}$$

To proof this fact we need the following inequality, where k is the bound for $\|A(t)\|$:

$$\mathrm{e}^{-k(\tau-t)}\|x\| \leq \|\Phi(\tau,t)x\|. \tag{4.34}$$

Inequality (4.33) can be obtained as follows:

$$x^T P(t)x = \int_t^\infty x^T\Phi^T(\tau,t)Q(\tau)\Phi(\tau,t)x\, d\tau \tag{4.35}$$

$$\geq \alpha \int_t^\infty \|\Phi(\tau,t)\|^2 d\tau \tag{4.36}$$

$$\geq \alpha \int_t^\infty x^T x \mathrm{e}^{-2k(\tau-t)} d\tau \tag{4.37}$$

$$= \underbrace{\frac{\alpha}{2k}}_{=:\beta} x^T x. \tag{4.38}$$

Since $P(t)$ is bounded by assumption, we have for some $\gamma > 0$

$$\beta\|x\|^2 \leq x^T P(t)x \leq \gamma\|x\|^2. \tag{4.39}$$

Defining $v(x(t),t) := x^T(t)P(t)x(t)$ we have $v(0,t) = 0$ and with (4.39) that $v(x(t),t)$ is a decreasing positive definite function. Moreover we have

$$\begin{aligned}\dot{v}(x(t),t) &= x^T(t)(\dot{P}(t) + A^T(t)P(t) + P(t)A(t))x(t)\\ &= -x^T(t)Q(t)x(t)\\ &\leq -\alpha\|x\|^2\end{aligned}$$

along solutions $x(t)$ of system (4.7). A function $v(x(t),t)$ with these properties is called Lyapunov function for system (4.7). Together we have

$$\begin{aligned}\frac{\dot{v}(x(t),t)}{v(x(t),t)} &\leq -2k\\ v(x(t),t) &\leq v(x(t_0),t_0)\mathrm{e}^{-2k(t-t_0)} \quad t_0 \leq t\\ \beta\|x\|^2 &\leq v(x(t_0),t_0)\mathrm{e}^{-2k(t-t_0)}\\ \|x\| &\leq v(x(t_0),t_0)^{\frac{1}{2}}\beta^{-\frac{1}{2}}\mathrm{e}^{-k(t-t_0)}\end{aligned}$$

showing exponential stability of the origin and with theorem (4.6) we have uniform asymptotic stability of the origin concluding the proof. □

This theorem gives a sufficient condition for uniform asymptotic stability by means of a certain class of Lyapunov functions. The Lyapunov functions used in this theorem are of very special nature and it should be mentioned that we could formulate a more general theorem by enlarging the set of Lyapunov function candidates. If one could find any (differentiable) function such that this function is a Lyapunov function with respect to (4.7) one can show stability of the origin which follows directly from the proof. We will cite such a result for nonlinear systems (Zubov 1964).

The advantage of using Lyapunov functions is that one does not need to know the solution of the underlying system. The disadvantage of the method of using Lyapunov functions is, that finding a suitable Lyapunov function is often a very difficult task. For linear time-varying systems the theorem above gives a "constructive condition" in the sense that it describes a possibility to construct such a Lyapunov function. It should be mentioned that the state transition matrix – which is needed in this construction – can in general not be obtained in an explicit form. One has to use approximations, for example with Chebyshev polynomials [Sinha and Chou, 1976].

4.1.2 Stability of linear autonomous systems

In this section we regard linear autonomous systems of the form

$$\dot{x}(t) = Ax(t), \quad x(t_0) = x_0, \tag{4.40}$$

where $A \in \mathbb{R}^{n\times n}$ is a constant matrix. Linear autonomous systems of the form (4.40) are special cases of linear time-varying systems (4.7). Therefore all stability results of the latter section hold for systems of the form (4.40) as well. In addition we will give some stability results using the eigenvalues of the matrix A. From the stability results of the time-varying case, we will restate the Lyapunov result (4.8) for the autonomous case. The results presented in this section are mainly taken from [Grüne and Junge, 2009] and [Leigh, 1980].

Lemma 4.2. *Equivalence of stability and uniform stability*
The equilibrium point $x = 0$ is a stable equilibrium point of (4.40) if and only if $x = 0$ is an uniformly stable equilibrium point of (4.40).

Proof. The equilibrium point $x = 0$ is an stable equilibrium point of (4.40) if and only if

$$\sup_{t\geq t_0} \|\Phi(t, t_0)\| < \infty. \tag{4.41}$$

Since $\Phi(t, t_0) = \mathrm{e}^{A(t-t_0)}$ this condition holds if and only if

$$\sup_{t_0} \sup_{t\geq t_0} \|\Phi(t, t_0)\| < \infty, \tag{4.42}$$

which is the criterion for uniform stability. □

Lemma 4.3. *Invariance with respect to coordinate transformations*
Let $T \in \mathbb{R}^{n\times n}$ be an invertible constant matrix. Use the coordinate transformation $y = T^{-1}x$ and define $\tilde{A} := T^{-1}AT$. The transformed system then reads:

$$\dot{y}(t) = \tilde{A}y(t), \quad y(t_0) = T^{-1}x_0. \tag{4.43}$$

The equilibrium point $x = 0$ of system (4.40) has the same stability properties as the equilibrium point $y = 0$ of the transformed system.

Proof. We start with the stability properties of the equilibrium point $y=0$ of system (4.43) and show that then the equilibrium point $x=0$ of system (4.40) has the same stability properties. The solution of system (4.40) is given by $x(t)=\Phi_A(t,t_0)x_0=\mathrm{e}^{A(t-t_0)}x_0$, the solution of system (4.43) is given by $y(t)=\Phi_{\tilde{A}}(t,t_0)y_0=\mathrm{e}^{T^{-1}AT(t-t_0)}y_0=T^{-1}\mathrm{e}^{A(t-t_0)}Ty_0$. From $Tx(t)=y(t)$ we have

$$\Phi_A(t,t_0)x_0=T\Phi_{\tilde{A}}(t,t_0)(T^{-1}x_0). \tag{4.44}$$

Let $y=0$ be an stable equilibrium point of (4.43). Then for $\tilde{\varepsilon}=\varepsilon/\|T\|>0$ there is a $\tilde{\delta}>0$ such that $\|y_0\|\leq\tilde{\delta}$ implies

$$\|\Phi_{\tilde{A}}(t,t_0)y_0\|\leq\tilde{\varepsilon}\quad\forall t\geq t_0. \tag{4.45}$$

Now $\|x_0\|\leq\delta$ implies

$$\|\Phi_A(t,t_0)x_0\|\overset{(4.44)}{=}\|T\Phi_{\tilde{A}}(t,t_0)(T^{-1}x_0)\|\leq\|T\|\|\Phi_{\tilde{A}}(t,t_0)y_0\|\overset{(4.45)}{\leq}\|T\|\frac{\varepsilon}{\|T\|}=\varepsilon \tag{4.46}$$

showing the stability of the equilibrium point $x=0$ of system (4.40).
Interchanging T with T^{-1} in the above argumentation shows that stability of the equilibrium point $x=0$ of system (4.40) implies stability of the equilibrium point $y=0$ of system (4.43). The same technique shows that $x=0$ is an asymptotically stable/unstable equilibrium point of system (4.40) if and only if $y=0$ is an asymptotically stable/unstable equilibrium point of (4.43). The "duality" for exponential stability follows from the equivalence theorem (4.6) for uniform asymptotic stability and exponential stability. □

Theorem 4.9. *(cf. e.g. [Grüne and Junge, 2009][cf. Th. 8.6]*[1] *Eigenvalue criteria*
Let $\lambda_1,\dots,\lambda_n\in\mathbb{C}$ denote the eigenvalues of the matrix A of system (4.40)*, where $a_l\in\mathbb{R}$ denotes the real part of the eigenvalue λ_l and $b_l\in\mathbb{R}$ denotes the imaginary part of the eigenvalue λ_l. We then have: The equilibrium point $x=0$ of system* (4.40) *is*

1. *stable if and only if all eigenvalues have non-positive real part and those eigenvalues, which have real part zero are semi-simple, i.e. given a Jordan normal representation of A, the Jordan blocks belonging to these eigenvalues have dimension* 1*.*

2. *unstable if and only if there is an eigenvalue λ_r, $1\leq r\leq n$ such that $\Re(\lambda_r)>0$ or $\Re(\lambda_r)=0$ and given a Jordan representation of A, the corresponding Jordan block J_r has dimension at least* 2×2*.*

3. *asymptotically stable if and only if all eigenvalues have negative real part.*

Proof. Due to lemma (4.3) if suffices to show the theorem for a system, where the matrix A has Jordan normal form. We regard the system

$$\dot{x}(t)=Jx(t),\quad x(t_0)=x_0, \tag{4.47}$$

where J is the a Jordan normal representation of the matrix A. (For convenience, we denoted the state variable of system (4.47) with $x(t)$ although it is not the same as the state variable $x(t)$ of system (4.40), as we have to change coordinates to obtain system (4.47)). To simplify the proof we make use of the 1-norm, denoted by $\|\cdot\|$ and defined by $\|x\|_1=\sum_{i=1}^n|x_i|$ for $x\in\mathbb{R}^n$.

[1]in [Grüne and Junge, 2009] the proof is incomplete: not all of the stated implications are shown

Without loss of generality we can assume $t_0 = 0$. The fundamental solution $\Phi(t, t_0 = 0)$ is then given by e^{Jt} and we have

$$\|e^{Jt}\|_1 = \sum_{k=1}^{m} \|e^{J_k t} x^{(k)}\|_1, \tag{4.48}$$

where $x^{(k)} \in \mathbb{R}^n$ are those vectors, which are built by those components of the vector x, which belong to the (generalized) eigenspace of the corresponding Jordan block J_k - the remaining entries are set equal to zero. Therefore it suffices to proof the stated stability conclusions for every Jordan block.
Eigenvalue criteria of 1. $\Longrightarrow$ stability / Eigenvalue criteria of 3. $\Longrightarrow$ asymptotic stability
Each Jordan block J_k is a square matrix of dimension $d \leq n$ and can be decomposed as

$$J_k = \lambda_k \cdot I_k + N_k, \tag{4.49}$$

where I_k is the unit matrix in $\mathbb{R}^{d\times d}$ and $N_k \in \mathbb{R}^{d\times d}$ is the nilpotent matrix

$$N_k = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}. \tag{4.50}$$

Further we have $N_k^d = 0$ and $\lambda_k I_k N_k = N_k \lambda I_k$, which are well-known facts from linear algebra (or can be verified by direct computation) as well as the representation

$$e^{J_k t} = e^{\lambda_k t} e^{N_k t} = e^{\lambda_k t}\left(I_k + tN_k + \ldots + \frac{t^{d-1}}{(d-1)!}N_k^{d-1}\right). \tag{4.51}$$

Using $|e^{\lambda_k t}| = |e^{(a_k+ib_k)t}| = e^{a_k t}$ and matrix norm induced by the 1-norm we obtain

$$\|e^{J_k t}\|_1 \leq |e^{\lambda_k t}|\|e^{N_k t}\|_1 \leq e^{a_k t}\left(1 + t\|N_k\|_1 + \ldots + \frac{t^{d-1}}{(d-1)!}\|N_k\|_1^{d-1}\right). \tag{4.52}$$

Case-by-case analysis – $a_k = \Re(\lambda_k) = 0$:
If $\Re(\lambda_k) = 0$, then by assumption $d = 1$ and we have

$$\|e^{J_k t}\|_1 = |e^{a_k t}| = e^0 = 1 \tag{4.53}$$

which establishes stability due to

$$\|\Phi(t,0)x_0\|_1 = \|e^{J_k t}x_0\|_1 \leq \|e^{J_k t}\|_1\|x_0\|_1 = \|x_0\|_1. \tag{4.54}$$

$a_k = \Re(\lambda_k) < 0$:
For any $\gamma > 0$ and any $p \in \mathbb{N}$ we have $\lim_{t\to\infty} e^{-\gamma t}t^p = 0$ such that for some $c > 0$ we have

$$e^{-\gamma t}\left(1 + t\|N_k\|_1 + \ldots + \frac{t^{d-1}}{(d-1)!}\|N_k\|_1^{d-1}\right) \leq c, \quad \forall t \geq 0. \tag{4.55}$$

Then for every $\sigma \in]0, -a[$ we can choose $\gamma = -a - \sigma > 0$ to obtain

$$\|e^{J_k t}\|_1 \leq e^{-\sigma t}e^{-\gamma t}\left(1 + t\|N_k\|_1 + \ldots + \frac{t^{d-1}}{(d-1)!}\|N_k\|_1^{d-1}\right) \leq ce^{-\sigma t}, \tag{4.56}$$

resulting in

$$\|\Phi(t,0)x_0\|_1 = \|\mathrm{e}^{J_k t}x_0\|_1 \leq \|\mathrm{e}^{J_k t}\|_1\|x_0\|_1 = c\mathrm{e}^{-\sigma t}\|x_0\|_1 \tag{4.57}$$

showing exponential stability which implies (uniform) asymptotic stability and therefore stability.

We now show, that the eigenvalue criteria of 2. lead to the instability of the equilibrium point $x = 0$ of system (4.40).

For some λ_k we have $a_k = \Re(\lambda_k) > 0$ or $a_k = \Re(\lambda_k) = 0 \wedge d \geq 2$.

Case-by-case analysis – $a_k = \Re(\lambda_k) = 0$:

Then for $e_1 \in \mathbb{R}^d$ and every $\varepsilon > 0$ we have

$$\|\mathrm{e}^{J_k t}(\varepsilon e_1)\|_1 = |\mathrm{e}^{\lambda_k t}|\varepsilon = \mathrm{e}^{a_k t}\varepsilon \tag{4.58}$$

which tends to infinity as $t \to \infty$. Therefore we can find arbitrary small initial values such that the solution is unbounded in time which proofs instability in the case where $a_k \geq 0$.

$a_k = \Re(\lambda_k) = 0 \wedge d \geq 2$:

For $e_2 \in \mathbb{R}^d$ we have due to (4.49)

$$\mathrm{e}^{\lambda_k t}e_2 = \mathrm{e}^{\lambda_k t}(te_1 + e_2) \tag{4.59}$$

and therefore for every $\varepsilon > 0$

$$\|\mathrm{e}^{J_k t}(\varepsilon e_2)\|_1 = |\mathrm{e}^{\lambda_k t}|\varepsilon(1+t) = \varepsilon(1+t) \tag{4.60}$$

which tends to infinity as $t \to \infty$ which shows instability of the equilibrium point $x = 0$ of (4.40).

Next we show that if the eigenvalue criteria of 3. do not hold, then $x = 0$ is not an asymptotically stable equilibrium point of (4.40).

If the eigenvalue criteria of 3. do not hold, there is at least one eigenvalue, say λ_r such that $\Re(\lambda_r) \geq 0$. If $\Re(\lambda_r) > 0$ or $\Re(\lambda_r) = 0$ and the corresponding Jordan block has dimension at least 2 then we have instability due to 2. If $\Re(\lambda_r) = 0$ and the dimension of the corresponding Jordan block is 1 we have for arbitrary $\varepsilon > 0$

$$\|\Phi_k(t,t_0)\varepsilon\|_1 = |\mathrm{e}^{\lambda_r t}|\varepsilon = \varepsilon \tag{4.61}$$

which does not converge to zero when $t \to \infty$ such that $x = 0$ fails to be an asymptotic stable equilibrium point of system (4.40) since at least the component corresponding to the Jordan block J_r does not meet the necessary condition of converging to zero as $t \to \infty$ for suitably small initial values.

The remaining directions now follow from the already proofed implications. □

There is no such generalization of this theorem for time-varying systems. The following example is due to Markus (cf. e.g. [Leigh, 1980][p. 69]). It is a two-dimensional example of a time-varying linear system having a complex pair of eigenvalues with negative real part. Nevertheless the system admits unbounded solutions:

$$\dot{x}(t) = \underbrace{\begin{pmatrix} a\cos^2(t) - 1 & 1 - a\sin(t)\cos(t) \\ -1 - a\sin(t)\cos(t) & a\sin^2(t) - 1 \end{pmatrix}}_{A(t)} x(t), \quad t \geq 0 \tag{4.62}$$

where $x(t) \in \mathbb{R}^2$. The fundamental matrix is given by

$$\Phi(t,0) = \begin{pmatrix} e^{(a-1)t}\cos(t) & e^{-t}\sin(t) \\ -e^{(a-1)t}\sin(t) & e^{-t}\cos(t) \end{pmatrix} \tag{4.63}$$

and for $1 < a < 2$ the solution

$$x(t) = \Phi(t,0)x_0, \quad x_0 \neq 0 \tag{4.64}$$

is clearly unstable. However the eigenvalues of $A(t)$ are independent of t and are given by

$$\lambda_{1/2} = \frac{a-2}{2} \pm i\frac{\sqrt{4-a^2}}{2} \tag{4.65}$$

where $\Re(\lambda_1) = \Re(\lambda_2) = \frac{a-2}{2} < 0$ due to $1 < a < 2$ completing the "counter-example".

Lemma 4.4. *(cf. e.g. [Grüne and Junge, 2009][Ch. 9]*
Let $P \in \mathbb{R}^{n\times n}$ be a positive definite matrix. Then there are constants $\alpha, \beta \in \mathbb{R}$, $0 < \alpha \leq \beta$ such that

$$\alpha\|x\|^2 \leq x^T P x \leq \beta\|x\|^2, \quad \forall x \in \mathbb{R}^n. \tag{4.66}$$

Proof. For $y = \frac{x}{\|x\|} \in \mathbb{R}^n$ we have $x^T P x = \|x\|^2 y^T P y$. Since $F(y) := y^T P y$ is a continuous function of $\{y \mid \exists\, x : y = \frac{x}{\|x\|}\}$ is compact, F attains its maximum $F_{\max} =: \beta$ and minimum $F_{\min} := \alpha$. □

The following theorem gives a stability criterion for linear autonomous systems via Lyapunov functions. This result can found in [Grüne and Junge, 2009], [Sastry, 1999] or in any other textbook about linear control theory:

Theorem 4.10. *Consider the linear autonomous system*

$$\dot{x}(t) = Ax(t), \quad x(t_0) = x_0. \tag{4.67}$$

Suppose there exists a matrix $P \in \mathbb{R}^{n\times n}$ and a constant $\alpha > 0$ such that

$$x^T(t)(A^T P + PA)x(t) \leq -\alpha\|x(t)\|^2 \tag{4.68}$$

for all $x \in \mathbb{R}^n$. Then $x = 0$ is an asymptotically stable equilibrium point of system (4.40) *if and only if P is positive definite.*

Proof. We define

$$V(x(t)) = x^T(t)Px(t) \tag{4.69}$$

and obtain

$$\dot{V}(x(t)) = x^T(t)(A^T P + PA)x(t) \leq -\alpha\|x(t)\|^2. \tag{4.70}$$

Assume now, that P is positive definite. Due to lemma (4.4) there is a $\beta > 0$ such that

$$\frac{\beta}{\alpha}\|x\|^2 \leq x^T P x, \quad \forall x \in \mathbb{R}^n. \tag{4.71}$$

Since now $V(x) > 0$, $\dot{V}(x) \leq 0$ for all $x \neq 0$ and $V(0) = 0$ we have that $V(x(t))$ is a Lyapunov-function and exponential stability follows as in the proof of (4.8):

$$\begin{aligned}\frac{\dot{V}(x(t))}{V(x(t))} &\leq -\beta \\ v(x(t)) &\leq v(x_0)\mathrm{e}^{-\beta(t-t_0)}, \quad t_0 \leq t \\ \frac{\beta}{\alpha}\|x(t)\|^2 &\leq v(x_0)\mathrm{e}^{-\beta(t-t_0)}\|x(t)\| \leq v(x_0)^{1/2}\alpha^{1/2}\beta^{-1/2}\mathrm{e}^{-\frac{1}{2}\beta(t-t_0)}\end{aligned}$$

where exponential stability implies asymptotic stability.
To show necessity of the positive definiteness of P let us assume, $P \not> 0$ but (4.68) still holds. Then there is a $x_0 \in \mathbb{R}^n$, with $\|x_0\| \neq 0$ and $V(x_0) \leq 0$. Since $\Phi(t, t_0)x_0 = \mathrm{e}^{A(t-t_0)}x_0 \neq 0$ for all $t \geq t_0$ we have due to (4.68) that $V(\Phi(t, t_0)x_0)$ is strictly monotonically decreasing. Therefore for some $\delta > 0$ we have

$$\|\Phi(t,t_0)x_0\|^2\|P\| \geq \|V(\Phi(t,t_0)x_0)\| \geq \delta \Longrightarrow \|\Phi(t,t_0)x_0\| \geq \frac{\delta}{\|P\|} > 0 \tag{4.72}$$

for all $t > T$ where T denotes the time where we have $V(\Phi(T, t_0)x_0) = -\delta$. (Due to (4.68) the case $\|P\| = 0$ cannot occur by assumption). The last inequality shows, that if P is not positive definite, $x = 0$ can not be an asymptotic stable equilibrium point of (4.40). □

4.1.3 Stability for nonlinear systems

Lyapunov showed in his doctor thesis 1892 that the existence of a suitable Lyapunov function for the equilibrium point $x = 0$ is sufficient to show stability. Zubov could show 1962 that if the zero-state is locally stable then - at least locally - there is a Lyapunov function for the underlying system (cf. e.g. [Poznjak, 2008]). The following theorem therefore is a necessary and sufficient condition for stability in the sense of Lyapunov.

Theorem 4.11. *(cf. e.g. [Poznjak, 2008][Th. 20.1.]) Local stability by Lyapunov's criterion (Zubov 1964)*
The equilibrium $x = 0$ of (4.1) is locally stable if and only if there exists a function $V(x, t)$, called Lyapunov function, satisfying the following conditions:

1. *$V(x, t)$ is defined for $\|x\| \leq h$ and $t \geq t_0$, where h is some small positive number.*
2. *$V(0, t) = 0$ for all $t \geq t_0$ and is continuous in x for all $t \geq t_0$ in the point $x = 0$.*
3. *$V(x, t)$ is positive definite.*
4. *$V(x(t), t)$, where $x(t)$ is a solution of (4.1), does not increase in t for $t \geq t_0$ where x_0 satisfies $\|x_0\| \leq h$.*

Proof. Sufficiency: Suppose there exists a function $V(x, t)$ such that conditions $1.-4.$ of theorem (4.11) hold. Due to positive definiteness (condition 3.) there is a function $W(x)$ such that

$$V(x,t) \geq W(x), \quad \forall t \geq t_0 \tag{4.73}$$

$$W(0) = 0, W(x) > 0 \quad \forall\, x \,:\, \|x\| \neq 0. \tag{4.74}$$

For $0 < \varepsilon < h$ consider the compact set of all states x satisfying $\|x\| = \varepsilon$. Then

$$\inf_{\{x:\|x\|=\varepsilon\}} W(x) =: \lambda > 0. \tag{4.75}$$

There is a number $\delta = \delta(t_0, \varepsilon)$ such that $\|x\| < \delta$ implies $V(x,t) < \lambda$ (conditions 1. and 2.). Due to property 4. we have for $\|x_0\| < \delta$ we have

$$V(x(t),t) \leq V(x_0,t_0) < \lambda, \quad \text{for } t \geq t_0, \tag{4.76}$$

and therefore $\|x(t)\| < \varepsilon$ for all solutions where the corresponding initial state x_0 satisfies $\|x_0\| < \delta$ which shows stability of the equilibrium point $x = 0$.
Necessity: Suppose the state $x = 0$ is a stable equilibrium point of system (4.1). Let $x(t)$ be the solution of (4.1) where the initial state x_0 satisfies $\|x_0\| \leq h$. Define

$$V(x,t) := \sup_{s \geq t} \|x(s)\|, \tag{4.77}$$

where $x(s)$ is the solution corresponding to the initial state $x(t)$
Since $x = 0$ is a stable equilibrium point of (4.1) we have $V(0,t) = 0$ for all $t \geq t_0$. Stability of the zero-state together with continuity of the solution $x(\cdot)$ and continuous dependence on its initial data guarantees that conditions 1. and 2. are satisfied by $V(x,t)$.
For x_0 with $\|x_0\| \neq 0$ we have

$$V(x_0,t_0) = \sup_{t \geq t_0} \|x(t)\| \geq \|x_0\| =: W(x_0) > 0 \tag{4.78}$$

and therefore $V(x,t)$ is positive definite satisfying condition 3.
Condition 4. easily follows from the definition of V. for $s \geq t (\geq t_0)$ we have

$$V(x(t),t) = \sup_{\tilde{t} \geq t} \|x(\tilde{t})\| \geq \sup_{\tilde{t} \geq s} \|x(\tilde{t})\| = V(x(s),s) \tag{4.79}$$

showing that $V(x(t),t)$ is nonincreasing in $t \geq t_0$ along solutions of (4.1). □

4.2 Stabilizing control laws

David Kleinmann [Kleinmann, 1970] used the Gramian introduced in (3.14) to stabilize linear constant systems. We assume the same assumptions as for (3.26), where we introduced linear constant systems.

4.2.1 Linear constant systems I (Kleinmann)

Theorem 4.12. *If a linear constant system of the form* (3.26)

$$\dot{x}(t) = Ax(t) + Bu(t)$$

is controllable, the control law

$$u(t) = -B^T \left(\int_0^T \mathrm{e}^{-A\tau} B B^T \mathrm{e}^{-A^T \tau} d\tau \right)^{-1} x(t), \quad T > 0 \tag{4.80}$$

stabilizes the system around the origin of the state space.

Proof. The zero state is an equilibrium point of the uncontrolled system.
We define

$$S(0,T) := \int_0^T \mathrm{e}^{-A\tau} B B^T \mathrm{e}^{-A^T \tau} d\tau \tag{4.81}$$

which is very similar to the matrix $W(0,T)$ defined in formula (3.14) of theorem (3.9). The relationship between both is given by

$$S(0,T) = \Phi(T,0)W(0,T)\Phi^T(T,0) \tag{4.82}$$

Since we assumed controllability of the nonautonomous linear system we have invertibility of $W(0,T)$ for every $T > 0$ due to theorems (3.9) and (3.13). With relationship (4.82) we have that $S(0,T)$ is also invertible. We will proof now that

$$v(t,x(t)) := x^T(t)S(0,T)^{-1}x(t) \tag{4.83}$$

is a suitable Lyapunov function showing that the origin of the state space is a stable equilibrium point of the controlled system in the sense of Lyapunov. Since $S(0,T)$ is a constant matrix, we will use the abbreviation S instead of $S(0,T)$.
With

$$\dot{x}(t) = Ax(t) + (-BB^TS^{-1})x(t) = (A - BB^TS^{-1})x(t) \tag{4.84}$$

we have

$$\dot{v}(t,x(t)) = x^T(t)(A^T - S^{-1}BB^T)S^{-1}x(t) + x^T(t)S^{-1}(A - BB^TS^{-1})x(t) \tag{4.85}$$

$$= x^T(t)\left[(A^T - S^{-1}BB^T)S^{-1} + S^{-1}(A - BB^TS^{-1})\right]x(t). \tag{4.86}$$

So we have to show that $\left[(A^T - S^{-1}BB^T)S^{-1} + S^{-1}(A - BB^TS^{-1})\right]$ is negative definite which – since S is a regular symmetric matrix – is equivalent to showing that

$$S\left[(A^T - S^{-1}BB^T)S^{-1} + S^{-1}(A - BB^TS^{-1})\right]S < 0 \tag{4.87}$$

(confer for example [Wigner, 1963]). The left hand side of (4.87) then reduces to

$$\begin{aligned} &S(A^T - S^{-1}BB^T)S^{-1}S + SS^{-1}(A - BB^TS^{-1})S \\ &= SA^T - BB^T + AS - BB^T \\ &= AS + SA^T - 2BB^T = -\mathrm{e}^{-AT}BB^T\mathrm{e}^{-A^TT} - BB^T \end{aligned} \tag{4.88}$$

which is negative definite.
For the last step (4.88) we used

$$AS + SA^T = -\mathrm{e}^{-AT}BB^T\mathrm{e}^{-A^TT} + BB^T \tag{4.89}$$

which holds because both sides of this equation are a representation of

$$-\int_0^T \frac{d}{d\tau}\left(\mathrm{e}^{-A\tau}BB^T\mathrm{e}^{-A^T\tau}\right)d\tau. \tag{4.90}$$

If we use the abbreviation $\bar{A}$ for the controlled system we have due to (4.84) $\bar{A} = (A - BB^TS^{-1})$ and because

$$-\mathrm{e}^{-AT}BB^T\mathrm{e}^{-A^TT} - BB^T = \bar{A}S + S\bar{A}^T \tag{4.91}$$

we have shown that the origin of the state space is a stable equilibrium point of the controlled system in the sense of Lyapunov. □

Remark: Kleinmann's method guarantees that the controlled system is exponential stable. It does not tell something about the rate of convergence. Before giving the generalization of Kleinmann's theorem to linear time-varying systems by Victor Cheng we shortly sketch a method to exponentially stabilize a controllable constant linear system where we can give a lower bound for the rate of convergence.

4.2.2 Linear constant systems II (Bass)

This method was presented in [Russell, 1979] where it is stated that it was first introduced by R. W. Bass in some lecture notes of the NASA Langley Research Center in August 1961.
In contrast to the method of Kleinmann we need not compute matrix exponentials or integrals. The problem of stabilizing a linear constant system is reduced to some linear equations coming from a Lyapunov equation.

Theorem 4.13. *The linear constant control system*

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{4.92}$$

is controllable if and only if the linear constant control system

$$\dot{x}(t) = (A - \lambda I)x(t) + Bu(t) \tag{4.93}$$

is controllable, where $\lambda \in \mathbb{R}$ is arbitrary.

Proof. Follows from the Kalman controllability criterion (3.11) and

$$[B|(A-\lambda I)B|(A-\lambda I)^2 B|\dots|(A-\lambda)^{n-1}B] =$$

$$= [B|AB|\dots|A^{n-1}B] \cdot \begin{bmatrix} I & -\lambda I & \lambda^2 I & \dots & (-1)^{n-1}\lambda^{n-1} I \\ 0 & I & -2\lambda I & \dots & (-1)^{n-2}(n-1)\lambda^{n-2} I \\ 0 & 0 & I & \dots & (-1)^{n-3}\binom{n-1}{n-3}\lambda^{n-3} I \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & I \end{bmatrix}$$

Therefore $[B|(A-\lambda I)B|(A-\lambda I)^2B|\dots|(A-\lambda)^{n-1}B]$ has full rank if and only if $[B|AB|\dots|A^{n-1}B]$ has full rank. ☐

R. W. Bass proposed the following method. Choose $\lambda > 0$ large enough such that $-A - \lambda I$ is stable in the sense of (4.9). This can be accomplished by choosing for example

$$\lambda > \max_i \sum_{j=1}^{n} |a_{ij}| \quad \text{or} \quad \lambda > \max_j \sum_{i=1}^{n} |a_{ij}|. \tag{4.94}$$

Regard

$$(A + \lambda I)P + P(A + \lambda I)^T = BB^T \tag{4.95}$$

which is a Lyapunov equation and has a unique positive definite symmetric solution P. Equation (4.95) is equivalent to

$$(A + \lambda I - BB^T P^{-1})P + P(A + \lambda I - BB^T P^{-1})^T + BB^T = 0. \tag{4.96}$$

Due to theorem (4.10) the matrix $(A + \lambda I - BB^T P^{-1})$ is stable in the sense of (4.9) and since $\lambda > 0$ we conclude that $(A - BB^T P^{-1})$ is stable where its eigenvalues lie in the left half plane and have distance at least λ from the imaginary axis such that we have exponential stability at a decay rate at least $-\lambda$.
We will apply the method of Bass for the linearized pendulum in section (5.1.2).

Remark: If we combine Kleinmann's method and the idea of Bass to stabilize the α-shifted system we obtain an additional factor in the formula for $S(0,T)$:

$$\begin{aligned} S_\alpha(0,T) &= \int_0^T \mathrm{e}^{-(A+\alpha I)\tau} BB^T \mathrm{e}^{-(A+\alpha I)^T\tau} d\tau \\ &= \int_0^T \mathrm{e}^{-\alpha I\tau}\mathrm{e}^{-A\tau} BB^T \mathrm{e}^{-A^T\tau}\mathrm{e}^{-\alpha I\tau} d\tau \\ &= \int_0^T \mathrm{e}^{-2\alpha\tau}\mathrm{e}^{-A\tau} BB^T \mathrm{e}^{-A^T\tau} d\tau. \end{aligned} \tag{4.97}$$

This allows us to preadjust the rate of decay of the controlled system. We will see that the generalization of Kleinmann's method for linear time-varying systems by Victor Cheng uses such a "factor for convergence".

4.2.3 Stabilizing under a time-varying nonlinearity - a sufficient criterion

Given a nonlinear system, consisting of a constant linear system with a time-varying nonlinearity $x^0(t)$, which is assumed to be at least continuously differentiable, we will give a condition under which the system can be transformed into a linear constant system and therefore can be stabilized with the methods presented above:

Theorem 4.14. *Regard the nonlinear system*

$$\dot{x}(t) = Ax(t) + Bu(t) + x^0(t). \tag{4.98}$$

Suppose there exists a $G \in GL(n)$ such that

$$\frac{d}{dt}x^0(t) = G^{-1}(AG - I)x^0(t), \tag{4.99}$$

then the system (4.98) can be stabilized along the trajectory $-Gx^0(t)$.

Proof. After the variable transformation

$$y(t) := x(t) + Gx^0(t) \tag{4.100}$$

system (4.98) becomes

$$\begin{aligned} \dot{y}(t) &= Ax(t) + Bu(t) + x^0(t) + G\dot{x}^0(t) \\ &= Ay(t) + Bu(t) - AGx^0(t) + x^0(t) + G\frac{d}{dt}x^0(t). \end{aligned} \tag{4.101}$$

Now if condition (4.99) holds we simply obtain a linear constant system

$$\dot{y}(t) = Ay(t) + Bu(t) \tag{4.102}$$

which can be stabilized with a linear constant feedback. □

Remarks:

- If A is invertible we can choose $G = A^{-1}$ to stabilize around a given point.
- Condition (4.99) is very restrictive in the sense that the "nonlinearity" of (4.98) is the solution of a linear system of the form $\dot{x}(t) = G^{-1}(AG - I)x(t)$ for some $G \in GL(n)$.

The following theorem by Victor Cheng is a generalization of Kleinmann's method for linear time-varying systems.

4.2.4 Linear time-varying systems (Cheng)

Consider the linear time-varying system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \quad t \geq t_0 \tag{4.103}$$

where we assume that $A(t) : \mathbb{R} \to \mathbb{R}^{n\times n}$ and $B(t) : \mathbb{R} \to \mathbb{R}^{n\times n}$ are piecewise continuous.
We define

$$H_\alpha(t_0, t) := \int_{t_0}^{t} \mathrm{e}^{4\alpha(t_0-\tau)} \Phi(t_0, \tau)B(\tau)B^T(\tau)\Phi^T(t_0, \tau)d\tau, \quad \alpha \geq 0 \tag{4.104}$$

which for $\alpha = 0$ is the time-varying generalization of (4.81).

Theorem 4.15. *[Cheng, 1979, Th. C1] If there is a $\delta > 0$ and $h_M \geq h_m > 0$ such that*

$$0 < h_m I \leq H_0(t, t+\delta) \leq h_M I \quad \forall t \geq t_0 \tag{4.105}$$

then for any $\alpha > 0$ and $\gamma : \mathbb{R}^+ \to [\frac{1}{2}, \infty[$ which is piecewise continuous, the linear time-varying control law

$$u(t) := -\gamma(t)B^T(t)H_\alpha(t, t+\delta)^{-1}x(t) \quad \forall t \in \mathbb{R}^+ \tag{4.106}$$

makes the zero-solution of the controlled system

$$\dot{x}(t) = \left(A(t) - \gamma(t)B(t)B^T(t)H_\alpha(t, t+\delta)^{-1}\right)x(t) \tag{4.107}$$

uniformly exponentially stable at a rate greater than α.

Proof. Instead of showing stability of the origin for the controlled system (4.107) we show stability for the α-shifted controlled system:

$$\dot{x}(t) = \left(A(t) + \alpha I - \gamma(t)B(t)B^T(t)H_\alpha(t, t+\delta)^{-1}\right)x(t) \tag{4.108}$$

as stability of the former is a direct consequence of stability of the latter (which for linear constant systems was introduced as the method of Bass). We will show that

$$v(t, x(t)) := x(t)^T H_\alpha(t, t+\delta)^{-1}x(t) \tag{4.109}$$

is a suitable Lyapunov function. Because $H_\alpha(t, t+\delta)$ is symmetric and positive definite the inverse $H_\alpha(t, t+\delta)^{-1}$ is also positive definite, therefore we have $v(t, x(t)) \geq 0$. For the zero-solution we have $v(t, 0) = 0$. It remains to show that the time-derivative of $v(t, x(t))$ is negative whenever $x(t)$ is not the zero-solution. We will use the following identities

$$\frac{d}{dt}H_\alpha(t, t+\delta)^{-1} = -H_\alpha(t, t+\delta)^{-1}\frac{d}{dt}H_\alpha(t, t+\delta)H_\alpha(t, t+\delta)^{-1} \tag{4.110}$$

$$(H_\alpha(t, t+\delta)^{-1})^T = H_\alpha(t, t+\delta)^{-1} \tag{4.111}$$

$$\frac{d}{dt}\Phi(t, \tau) = A(t)\Phi(t, \tau) \tag{4.112}$$

$$\begin{aligned}\frac{d}{dt}H_\alpha(t, t+\delta) &= 4\alpha H_\alpha(t, t+\delta) + A(t)H_\alpha(t, t+\delta) + H_\alpha(t, t+\delta)A^T(t) + \\ &\quad + \mathrm{e}^{-4\alpha\delta}\Phi(t, t+\delta)B(t+\delta)B^T(t+\delta)\Phi^T(t, t+\delta) - B(t)B^T(t)\end{aligned} \tag{4.113}$$

to compute the time-derivative of $v(t, x(t))$:

$$\begin{aligned}
\frac{d}{dt}v(t,x(t)) &= x^T(t)\left[A^T(t)H_\alpha^{-1}(t,t+\delta) - H^T(t,t+\delta)^{-1}B(t)B^T(t)\gamma(t)H_\alpha^{-1}(t,t+\delta)+\right.\\
&\quad +\alpha H_\alpha^{-1}(t,t+\delta) - H_\alpha^{-1}(t,t+\delta)\frac{d}{dt}H(t,t+\delta)H_\alpha^{-1}(t,t+\delta) + H_\alpha^{-1}(t,t+\delta)A(t)-\\
&\quad \left.-H_\alpha^{-1}(t,t+\delta)\gamma(t)B(t)B^T(t)H_\alpha^{-1}(t,t+\delta) + \alpha H_\alpha^{-1}(t,t+\delta)\right]x(t)\\
&= -x^T(t)H_\alpha^{-1}(t,t+\delta)\left[-H_\alpha(t,t+\delta)A^T(t) + \gamma(T)B(t)B^T(t) - \alpha H_\alpha(t,t+\delta)+\right.\\
&\quad +4\alpha H_\alpha(t,t+\delta) + A(t)H_\alpha(t,t+\delta) + H_\alpha(t,t+\delta)A^T(t) + \mathrm{e}^{-4\alpha\delta}\Phi(t,t+\delta)B(t+\delta)\cdot\\
&\quad \cdot B^T(t+\delta)\Phi(t,t+\delta) - B(t)B^T(t) - A(t)H_\alpha(t,t+\delta) + \gamma(t)B(t)B^T(t) - \left.\alpha H_\alpha(t,t+\delta)\right]\cdot\\
&\quad \cdot H_\alpha^{-1}(t,t+\delta)x(t)\\
&= -x^T(t)H_\alpha^{-1}(t,t+\delta)\left(\underbrace{(2\gamma-1)B(t)B^T(t)}_{I}+\right.\\
&\quad \left.+\underbrace{\mathrm{e}^{-4\alpha\delta}\Phi(t,t+\delta)B(t+\delta)B^T(t+\delta)\Phi^T(t,t+\delta)}_{II}+2\alpha H_\alpha(t,t+\delta)\right)H_\alpha^{-1}(t,t+\delta)x(t)
\end{aligned}$$

and since $II \geq 0$ and $I \geq 0$ because $\frac{1}{2} \leq \gamma(t)$ we can omit them to obtain

$$\frac{d}{dt}v(t,x(t)) \leq -2\alpha x^T(t)H_\alpha^{-1}(t,t+\delta)x(t) \leq -\frac{2\alpha}{h_M}\|x(t)\|^2 \leq 0 \tag{4.114}$$

which gives us

$$\begin{aligned}
\frac{\dot v(t,x(t))}{v(t,x(t))} &\leq \frac{-2\alpha h_M^{-1}}{\mathrm{e}^{4\alpha\delta}h_m^{-1}} = -2\alpha\frac{h_m}{h_M}\mathrm{e}^{-4\alpha\delta}\\
\ln|v(t,x(t))| - \ln|v(t_0,x_0)| &\leq -2\alpha\frac{h_m}{h_M}\mathrm{e}^{-4\alpha\delta}(t-t_0)\\
|v(t,x(t))| &\leq |v(t_0,x_0)|\mathrm{e}^{\left(-2\alpha\frac{h_m}{h_M}\mathrm{e}^{-4\alpha\delta}(t-t_0)\right)}\\
h_M^{-1}\|x(t)\|^2 &\leq \mathrm{e}^{4\alpha\delta}h_m^{-1}\|x(t_0)\|^2\mathrm{e}^{\left(-2\alpha\frac{h_m}{h_M}\mathrm{e}^{-4\alpha\delta}(t-t_0)\right)}\\
\|x(t)\| &\leq \|x_0\|\sqrt{\frac{h_M}{h_m}}\mathrm{e}^{2\alpha\delta}\mathrm{e}^{\left(-\alpha\frac{h_m}{h_M}\mathrm{e}^{-4\alpha\delta}(t-t_0)\right)}
\end{aligned}$$

showing that the zero-solution of the $\alpha-$shifted system is uniformly exponentially stable at a rate at least $\alpha\frac{h_m}{h_M}\mathrm{e}^{-4\alpha\delta}$. Therefore the original system is uniformly exponentially system at a rate at least $\alpha\frac{h_m}{h_M}\mathrm{e}^{-4\alpha\delta}$ which concludes the proof. □

4.2.5 Nonlinear systems I (Sastry et al.)

Based on the work of Kleinmann and Cheng a similar control law was proposed for nonlinear systems, stabilizing it to a given reference trajectory, chosen as a bounded solution of the uncontrolled system. The following theorem was proposed in a paper of G. Walsh, D. Tilbury, S. Sastry, R. Murray and J. P. Laumond in [Sastry et al., 1994]

Theorem 4.16. *[Sastry et al., 1994][Prop. 1]*
Let

$$\dot x(t) = f(x(t),u(t)) \quad t \geq t_0 \tag{4.115}$$

be a nonlinear system which belongs to the class of C^2-functions with regard to x and u. Given a bounded reference curve $x^0(t)$ of (4.115) *as solution of the uncontrolled system, i.e. a solution of $\dot{x}(t) = f(x(t), 0)$, define*

$$A(t) := \frac{\partial f}{\partial x}(x^0(t), 0) \tag{4.116}$$

$$B(t) := \frac{\partial f}{\partial u}(x^0(t), 0). \tag{4.117}$$

$$\tag{4.118}$$

Let $\Phi(t, t_0)$ denote the state transition matrix of $A(t)$ corresponding to the initial time t_0. For $\alpha \geq 0$ we slightly modify the definition of $H_\alpha(t_0, t)$:

$$H_\alpha(t_0, t) := \int_{t_0}^{t} \mathrm{e}^{6\alpha(t_0-\tau)} \Phi(t_0, \tau) B(\tau) B^T(\tau) \Phi^T(t_0, \tau) d\tau. \tag{4.119}$$

If there exists a $\delta > 0$ such that $H_\alpha(t, t+\delta)$ is bounded away from singularity and numbers h_m, h_M such that

$$0 < h_M < H_\alpha^{-1}(t, t+\delta) < h_m \quad \forall t \tag{4.120}$$

then, for any function $\gamma : t \to [\frac{1}{2}, \infty)$, continuous and bounded, the control function

$$u(t) := -\gamma(t) B^T(t) H_\alpha^{-1}(t, t+\delta)(x(t) - x^0(t)) \quad t \geq t_0 \tag{4.121}$$

locally, uniformly, exponentially stabilizes the system (4.115) *to the reference trajectory $x^0(t)$ at a rate at least $\alpha h_m h_M^{-1} \mathrm{e}^{-6\alpha\delta}$.*

Proof. We introduce a new variable

$$\tilde{x}(t) = x(t) - x^0(t) \tag{4.122}$$

$$\tag{4.123}$$

The Taylor series expansion of system (4.115) along the reference trajectory then gives

$$\dot{\tilde{x}}(t) = A(t)\tilde{x}(t) + B(t)u(t) + o(\tilde{x}(t), u(t), t), \tag{4.124}$$

where the higher order terms are $o(\tilde{x}(t), u(t), t)$ since we assumed $f \in C^2$. The suggested control u is a feedback control law, so the higher order terms actually depend only on $\tilde{x}(t)$ and t.
Since $x^0(t)$ was assumed to be bounded we know that $B(t)$ is bounded. We also assumed that H is bounded away from singularity meaning that H^{-1} is bounded from above. Together with the assumption that $\gamma(t)$ is bounded this results in

$$\|u(t)\| = \|\gamma(t) B^T(t) H_\alpha^{-1}(t, t+\delta)\tilde{x}(t)\| \leq C\|\tilde{x}(t)\|, \quad C < \infty \tag{4.125}$$

From this follows

$$\lim_{\|\tilde{x}(t)\| \to 0} \sup_{t \geq t_0} \frac{\|o(\tilde{x}(t), t)\|}{\|\tilde{x}(t)\|} = 0. \tag{4.126}$$

The proof is almost the same as in theorem 4.15:
We show that

$$v(\tilde{x}(t), t) = \tilde{x}^T(t) H_\alpha^{-1}(t, t+\delta)\tilde{x}(t) \tag{4.127}$$

The computation is pretty much the same as in the proof of theorem 4.15 with the only difference that we get the additional terms

$$\tilde{x}^T(t)H_\alpha^{-1}(t,t+\delta)o(\tilde{x}(t),t) + o(\tilde{x}(t),t)H_\alpha^{-1}(t,t+\delta)\tilde{x}(t). \tag{4.128}$$

The time-derivative of $v(\tilde{x}(t),t)$ is given by

$$\begin{aligned}\frac{d}{dt}v(\tilde{x}(t),t) = -\tilde{x}^T(t)H_\alpha^{-1}(t,t+\delta)\Big((2\gamma-1)B(t)B^T(t) + \mathrm{e}^{-6\alpha\delta}\cdot \\ \cdot\Phi(t,t+\delta)B(t+\delta)B^T(t+\delta)\Phi^T(t,t+\delta) + 4\alpha H_\alpha(t,t+\delta)\Big)H_\alpha^{-1}(t,t+\delta)\tilde{x}(t)+ \\ \tilde{x}^T(t)H_\alpha^{-1}(t,t+\delta)o(\tilde{x}(t),t) + o(\tilde{x}(t),t)H_\alpha^{-1}(t,t+\delta)\tilde{x}(t)\end{aligned}$$

Similar to (4.114) the first two lines are bounded by $-\frac{4\alpha}{h_M}\|\tilde{x}(t)\|^2$ whereas the last line is bounded from above by $\frac{\alpha}{h_M}\|\tilde{x}(t)\|^2$ for sufficiently small $\tilde{x}(t)$ which is guaranteed by (4.126). Together we have

$$\begin{aligned}\dot{v}(\tilde{x}(t),t) &\le -4\alpha h_M^{-1}\|\tilde{x}(t)\| \\ \frac{\dot{v}(\tilde{x}(t),t)}{v(\tilde{x}(t),t)} &\le \frac{-4\alpha h_M^{-1}\|\tilde{x}(t)\|}{\mathrm{e}^{6\alpha\delta}h_m^{-1}\|\tilde{x}(t)\|} = -4\alpha h_m h_M^{-1}\mathrm{e}^{-6\alpha\delta} \\ v(\tilde{x}(t),t) &\le v(\tilde{x}(t_0),t_0)\mathrm{e}^{-4\alpha h_m h_M^{-1}\mathrm{e}^{-6\alpha\delta}(t-t_0)} \quad \text{for } 0\le t_0\le t \\ h_M^{-1}\|\tilde{x}(t)\|^2 &\le \mathrm{e}^{6\alpha\delta}h_m^{-1}\|\tilde{x}(t_0)\|^2\mathrm{e}^{-2\alpha h_m h_M^{-1}\mathrm{e}^{-6\alpha\delta}(t-t_0)} \\ \|\tilde{x}(t)\| &\le \|\tilde{x}(t_0)\|h_M^{1/2}h_m^{-1/2}\mathrm{e}^{3\alpha\delta}\mathrm{e}^{-2\alpha h_m h_M^{-1}\mathrm{e}^{-6\alpha\delta}(t-t_0)}\end{aligned}$$

concluding the proof. □

4.2.6 Nonlinear systems II - a modified control law

The control law designed by Sastry et al. has the disadvantage that in every single numerical time step the integral for the matrix $H_\alpha(t,t+\delta)$ has to be solved, which in general will be done numerically. Moreover this matrix has to be inverted, which again will in general be done numerically.
One can think of a car driver, who – at every moment – looks ahead a short distance to the course of the road to adjust velocity and steering angle. Of course this method seems to be most adequate but experience shows that under certain circumstances it suffices to look on the road only from time to time. A short distraction where for some time we do not look on the road will in many cases not lead to a catastrophe. This is the motivation for the following modified control law. Instead of computing the matrix $H_\alpha^{-1}(t,t+\delta)$ in every single time step, we will compute it only once, keep it constant for a certain amount of time and then repeat this step for the next time interval, not necessarily of same length.
One might assume that this kind of "control" strongly depends on the "smoothness of the road". For the idealized example of a straight road we would even expect to do not worse than with looking on the road at every time. We will show by the example of the double and triple pendulum system that even for those very sensitive systems our approach leads to very good results for even relatively large time intervals.

The scalar control function u is designed as feedback control very similar to (4.121). Let $t_0 < t_1 < t_2 < \ldots < t_i < t_j < \ldots$ be a strictly increasing sequence. Define the positive numbers $\delta_1 := t_2 - t_1$, $\delta_2 = t_3 - t_2$ and in general $\delta_k = t_{k+1} - t_k$. For some constant $\alpha \geq 0$ we define the following matrix similar to (4.104):

$$\tilde{H}_\alpha(t_k, \delta_k) := \int_{t_k}^{t_k+\delta_k=t_{k+1}} \mathrm{e}^{4\alpha(t_k-\tau)} \Phi(t_k, \tau) B B^T \Phi^T(t_k, \tau) d\tau. \tag{4.129}$$

Assume that the matrix $\tilde{H}_\alpha(t_k, \delta_k)$ is uniformly bounded away from singularity in $\{t_i\}_{i\in\mathbb{N}}$, i.e. there is a positive constant h independent of k such that for all $k \geq 0$ we have $0 < hI < \tilde{H}_\alpha(t_k, \delta_k)$. We need this condition for the invertibility of $\tilde{H}$ and the existence of an upper bound $\tilde{h}$ for its inverse.

For a finite interval $[a, b]$, where $a < b < \infty$, we define the control function u as:

$$u(t) := \sum_{k=\max_i\{t_i \leq a\}}^{\min_i\{b \leq t_i\}} (-1) \cdot B^T \tilde{H}_\alpha^{-1}(t_k, \delta_k) \chi_{[t_k, t_{k+1})}(t)(x(t) - x^0(t)), \tag{4.130}$$

where $\chi_{[t_k,t_{k+1})}(t) = \begin{cases} 1 & \text{if } t \in [t_k, t_{k+1}) \\ 0 & \text{otherwise} \end{cases}$ is the characteristic function and $x^0(t)$ is the reference trajectory.

Chapter 5

Application to plane pendulum systems up to three links

5.1 Simple Pendulum

5.1.1 System dynamics

In this section we will deduce the equation of motion for the mathematical pendulum using the Euler-Lagrange formalism and show that its linearization around one of its equilibrium points is given by equation (2.11).

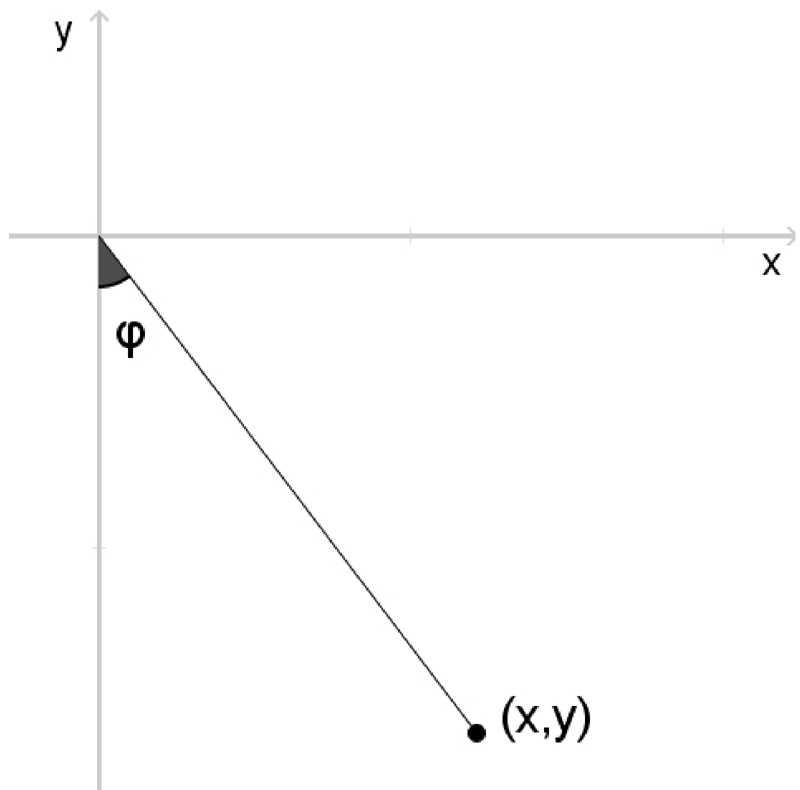


Figure 5.1: simple pendulum

Consider a simple mathematical pendulum, where we make simplifications such as massless rigid link, no friction, no gravity, etc.
The angular point of the simple pendulum is centered at the origin of the coordinate center and the angle is measured from the downright position. For simplicity we set the length of the link equal to 1 as well as the (point) mass centered at the bob, which we will also assume to be 1. At time t we have

$$x(t) = \sin \varphi(t) \tag{5.1}$$
$$y(t) = \cos \varphi(t) \tag{5.2}$$
$$\tag{5.3}$$

and for the velocity $v(t)$

$$v^2(t) = \dot{x}^2(t) + \dot{y}^2(t) = \dot{\varphi}^2(t) \sin^2 \varphi(t) + \dot{\varphi}(t) \cos^2 \varphi(t) = \dot{\varphi}^2(t). \tag{5.4}$$

The kinetic energy of the pendulum system is given by

$$E_{\text{kin}}(\varphi, \dot{\varphi}) = \frac{1}{2}\dot{\varphi}^2(t). \tag{5.5}$$

If the x-axis marks the zero level for the potential energy we can measure it by

$$E_{\text{pot}}(\varphi, \dot{\varphi}) = -\cos \varphi(t). \tag{5.6}$$

The Lagrangian $L(\varphi, \dot{\varphi})$ is then given by

$$L(\varphi, \dot{\varphi}) = E_{\text{kin}}(\varphi, \dot{\varphi}) - E_{\text{pot}}(\varphi, \dot{\varphi}) = \frac{1}{2}\dot{\varphi}^2(t) + \cos \varphi. \tag{5.7}$$

The Lagrangian for the n-pendulum can be found in appendix C.
The equation of motion can now be derived by the Euler-Lagrange equation:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}}(\dot{\varphi}, \varphi) - \frac{\partial L}{\partial \varphi} = 0, \tag{5.8}$$

which gives

$$\frac{d}{dt}\dot{\varphi}(t) + \sin \varphi(t) = 0 \Longleftrightarrow \ddot{\varphi}(t) + \sin \varphi(t) = 0. \tag{5.9}$$

Writing the differential equation (5.9) as first order system we obtain the (nonlinear) pendulum equation

$$\begin{aligned} \dot{\varphi}(t) &= \psi(t) \\ \dot{\psi}(t) &= -\sin(\varphi(t)). \end{aligned} \tag{5.10}$$

5.1.2 Discussion of the linearized simple pendulum

The linearization along a trajectory $(\tilde{\varphi}(t), \tilde{\psi}(t))^T$ can be computed by formula (3.54) of definition (3.18):

$$\frac{d}{dt}\begin{pmatrix} \varphi(t) \\ \psi(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\cos(\tilde{\varphi}(t)) & 0 \end{pmatrix} \begin{pmatrix} \varphi(t) \\ \psi(t) \end{pmatrix}. \tag{5.11}$$

The linearized model (2.11) we used so far corresponds to the linearization at the equilibrium point $(0, 0)$. Equation (2.11) can be obtained be setting $(\tilde{\varphi}(t), \tilde{\psi}(t)) = (0, 0)$ in (5.11).

Controllability of the linearization around an arbitrary point

The linearization around a given point $(\varphi^*, \psi^*)^T$ is given by

$$\frac{d}{dt}\begin{pmatrix}\varphi(t)\\ \psi(t)\end{pmatrix} = \underbrace{\begin{pmatrix}0 & 1\\ -\cos(\varphi^*) & 0\end{pmatrix}}_{=:A_{(\varphi^*,\psi^*)}}\begin{pmatrix}\varphi(t)\\ \psi(t)\end{pmatrix}. \tag{5.12}$$

We regard the control system

$$\frac{d}{dt}\begin{pmatrix}\varphi(t)\\ \psi(t)\end{pmatrix} = \begin{pmatrix}0 & 1\\ -\cos(\varphi^*) & 0\end{pmatrix}\begin{pmatrix}\varphi(t)\\ \psi(t)\end{pmatrix} + \underbrace{\begin{pmatrix}0\\ 1\end{pmatrix}}_{=:B}u(t), \tag{5.13}$$

where u is a scalar-valued control input.
(5.13) is a linear autonomous control system. Due to corollary (3.2) local controllability is equivalent to global controllability. The global controllability can be checked via the Kalman rank condition (3.11):

$$\operatorname{rank}[B|A_{(\varphi^*,\psi^*)}B] = \operatorname{rank}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = 2, \tag{5.14}$$

Therefore the linearization of the nonlinear pendulum equation around an arbitrary point is globally controllable.

Example: state-transition in finite time

We regard the linearization around the unstable equilibrium point $(\pi, 0)^T$ of the nonlinear pendulum equation. We will use theorem (3.9) to show that any state $x_0 = (\pi - \varepsilon, \delta)^T$ for some $\varepsilon, \delta \in \mathbb{R}$ can be transferred to the unstable equilibrium $(\pi, 0)^T$ in any positive finite time $T > 0$. Without loss of generality we assume $t_0 = 0$.
Theorem (3.9) provides the control law performing this task:

$$u^*(t) = -B^T e^{A^T_{(\pi,0)}(T-t)}\left(\int_0^T e^{A_{(\pi,0)}(T-\tau)}BB^T e^{A^T_{(\pi,0)}(T-t)}d\tau\right)^{-1}\left(e^{A_{(\pi,0)}T}\begin{pmatrix}\pi-\varepsilon\\ \delta\end{pmatrix} - \begin{pmatrix}\pi\\ 0\end{pmatrix}\right) \tag{5.15}$$

for $0 \le t \le T$ and which can be easily verified by inserting into the solution formula (3.9) for linear control systems:

$$\begin{aligned}x(T, x_0, u^*) &= e^{A_{(\pi,0)}T}\begin{pmatrix}\pi-\varepsilon\\ \delta\end{pmatrix} - \int_0^T e^{A_{(\pi,0)}(T-\tau)}BB^T e^{A^T_{(\pi,0)}(T-\tau)}d\tau\cdot \\ &\cdot\left(\int_0^T e^{A_{(\pi,0)}(T-\tau)}BB^T e^{A^T_{(\pi,0)}(T-t)}d\tau\right)^{-1}\left(e^{A_{(\pi,0)}T}\begin{pmatrix}\pi-\varepsilon\\ \delta\end{pmatrix} - \begin{pmatrix}\pi\\ 0\end{pmatrix}\right)\\ &= \begin{pmatrix}\pi\\ 0\end{pmatrix}.\end{aligned} \tag{5.16}$$

Stability of the equilibria

The nonlinear pendulum equation (5.10) has the equilibria $(0, 0)^T$ and $(\pi, 0)^T$. We are interested in the stability of the linearized pendulum equation around these equilibria.

For the equilibrium $(0,0)^T$ we obtain

$$\frac{d}{dt}\begin{pmatrix}\varphi(t)\\ \psi(t)\end{pmatrix} = \underbrace{\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}}_{=:A_{(0,0)}}\begin{pmatrix}\varphi(t)\\ \psi(t)\end{pmatrix}. \tag{5.17}$$

Stability of (5.17) can be determined with the eigenvalue criteria (4.9). The system matrix $A_{(0,0)}$ has eigenvalues $-i$ and i. Since they are semi-simple we can conclude stability but no asymptotic stability.

For the equilibrium $(\pi,0)^T$ we obtain

$$A_{(\pi,0)} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \tag{5.18}$$

which has eigenvalues -1 and 1 and due to theorem (4.9) the corresponding linear system is unstable.

If the state transition matrix of the linear system is known, we can use theorem (4.5) to determine the stability properties:
For $A_{(0,0)}$ the state transition matrix for $t \geq t_0$ is given by

$$\Phi_{(0,0)}(t,t_0) = \begin{pmatrix}\cos(t-t_0) & \sin(t-t_0)\\ -\sin(t-t_0) & \cos(t-t_0)\end{pmatrix}. \tag{5.19}$$

In order to apply theorem (4.5) we have to compute the norm of (5.19).

$$\begin{aligned}\|\Phi_{(0,0)}(t,t_0)\|_2 &= \max_{x\in\mathbb{R}^2,\|x\|_2=1}\|\Phi_{(0,0)}(t,t_0)x\|_2 = \left\|\begin{pmatrix}\cos(t-t_0)x_1+\sin(t-t_0)x_2\\ -\sin(t-t_0)x_1+\cos(t-t_0)x_2\end{pmatrix}\right\|_2\\ &= \max_{x\in\mathbb{R}^2,\|x\|_2=1}\sqrt{2\cos^2(t-t_0)x_1^2+2\sin^2(t-t_0)x_2^2}\\ &\leq \sqrt{2(\cos^2(t-t_0)+\sin^2(t-t_0))\max_{x\in\mathbb{R}^2,\|x\|_2=1}\{x_1^2,x_2^2\}}\\ \Longrightarrow \|\Phi_{(0,0)}(t,t_0)\|_2 &\leq \sqrt{2}\leq\infty,\end{aligned} \tag{5.20}$$

therefore the origin of the system

$$\dot{x}(t) = A_{(0,0)}x(t) \tag{5.21}$$

is uniformly stable. To show that the origin is *not* asymptotically stable with respect to system (5.21) if suffices to take $x_1 = x_2 = \sin\left(\frac{\pi}{4}\right)$ since then we have

$$\|\Phi_{(0,0)}(t,t_0)(x_1,x_2)^T\|_2 \geq \sqrt{2}\min\{|x_1|,|x_2|\} = \sqrt{2}\sin\left(\frac{\pi}{4}\right) > 0. \tag{5.22}$$

Because $\|(x_1,x_2)^T\|_2 = 1$ this implies that $\|\Phi_{(0,0)}(t,t_0)\|_2 \not\to 0$ as $t\to\infty$. For linear autonomous systems exponential stability is equivalent to uniform asymptotic stability. For lack of asymptotic stability, we therefore do not have exponential stability either.

For system $\dot{x}(t) = A_{(\pi,0)}x(t)$ the state transition matrix for $t \geq t_0$ is given by

$$\Phi_{(\pi,0)}(t,t_0) = \begin{pmatrix}1/2\,\mathrm{e}^{-(t-t_0)}+1/2\,\mathrm{e}^{t-t_0} & 1/2\,\mathrm{e}^{t-t_0}-1/2\,\mathrm{e}^{-(t-t_0)}\\ 1/2\,\mathrm{e}^{t-t_0}-1/2\,\mathrm{e}^{-(t-t_0)} & 1/2\,\mathrm{e}^{-(t-t_0)}+1/2\,\mathrm{e}^{t-t_0}\end{pmatrix} = \begin{pmatrix}\cosh(t-t_0) & \sinh(t-t_0)\\ \sinh(t-t_0) & \cosh(t-t_0)\end{pmatrix} \tag{5.23}$$

and since $\sinh(t-t_0)$ and $\cosh(t-t_0)$ tend to infinity as $t \to \infty$ we see that because of $\Phi(t,t_0)_{(\pi,0)}(0,1)^T = (\sinh(t-t_0), \cosh(t-t_0))^T$ the norm of (5.21) is unbounded which implies instability.

Stabilization - Balancing the upright position

The upright position $(\pi,0)^T$ is a fixed point of the nonlinear system (5.41). The linearization in $(\pi,0)^T$ is given by

$$\frac{d}{dt}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\cos(\pi) & 0 \end{pmatrix}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}. \tag{5.24}$$

We will assume in this example that we have a scalar control u which enters linearly. The controlled system is assumed to be of the form

$$\frac{d}{dt}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{A}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} + \underbrace{\begin{pmatrix} 0 \\ 1 \end{pmatrix}}_{B} u(t) \tag{5.25}$$

In the following we will use the pole-shifting theorem (3.15), Kleinmann's method (4.12) and the method of Bass (4.2.2) to compute the stabilization control law.

First we use the pole-shifting theorem (3.15) to stabilize the linear system by a constant feedback.
We want to stabilize the origin in the sense that the equilibrium point $(0,0)^T$ becomes asymptotically stable which, if the controlled system remains autonomous and linear, together with (5.20) implies exponential stability.
Motivated by theorem (3.15) we make the following ansatz for the control function u:

$$u(t) := F \cdot \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \tag{5.26}$$

where $F \in \mathbb{R}^{(1,2)}$. We obtain the controlled system

$$\frac{d}{dt}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + BF\right)\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \tag{5.27}$$

where we are now looking for suitable choices F_{11} and F_{12} such that the origin is an asymptotically stable equilibrium point of (5.27). Using theorem (4.9) we can achieve asymptotic stability by choosing F such that $\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + BF\right)$ has only negative eigenvalues.
If we choose $F_{11} = -2$ and $F_{12} = -2$ we obtain -1 as eigenvalue with multiplicity 2. The controlled system is given by

$$\frac{d}{dt}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \tag{5.28}$$

and for the initial values $x(0) = x_0$, $y(0) = y_0$ the solution is given by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} (x_0 + y_0 \cdot t)\mathrm{e}^{-t} \\ -(x_0 - y_0 + y_0 \cdot t)\mathrm{e}^{-t} \end{pmatrix} \tag{5.29}$$

which for any choice of $(x_0, y_0)^T \in \mathbb{R}^2$ tends to $(0,0)^T$ as time tends to infinity.

For applying Kleinmann's method (4.12) we need to compute the inverse of the integral

$$S(0,T) := \int_0^T \mathrm{e}^{-A\tau} BB^T \mathrm{e}^{-A^T\tau} d\tau \tag{5.30}$$

where without loss of generality we set $t_0 = 0$ and T is some positive time such that (5.30) is invertible.
For $A_{(\pi,0)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ the state transition matrix is given by formula (5.23) and we obtain

$$S_{(\pi,0)}(0,T) = \left(\begin{array}{cc} -1/8\,\left(1 + 4\,T\mathrm{e}^{2\,T} - \mathrm{e}^{4\,T}\right)\mathrm{e}^{-2\,T} & 1/8\,\left(2\,\mathrm{e}^{2\,T} - 1 - \mathrm{e}^{4\,T}\right)\mathrm{e}^{-2\,T} \\ 1/8\,\left(2\,\mathrm{e}^{2\,T} - 1 - \mathrm{e}^{4\,T}\right)\mathrm{e}^{-2\,T} & 1/8\,\left(\mathrm{e}^{4\,T} + 4\,T\mathrm{e}^{2\,T} - 1\right)\mathrm{e}^{-2\,T} \end{array} \right) \tag{5.31}$$

which has determinant

$$\begin{aligned} \det S_{(\pi,0)}(0,T) &= \frac{1}{8}\left(\cosh(2T) - 1\right) - \frac{1}{4}T^2 \\ &= \frac{1}{8}\left(\left(\sum_{k=0}^{\infty} \frac{2^{2k}}{(2k)!}T^{2k}\right) - 1 - 2T^2\right) \\ &= \frac{1}{8}\left(\sum_{k=2}^{\infty} \frac{2^{2k}}{(2k)!}T^{2k}\right) \end{aligned} \tag{5.32}$$

which for every $T > 0$ is greater than 0 and therefore $S_{(\pi,0)}(0,T)$ is invertible for $T > 0$ (which actually follows directly from the assumption of controllability). The inverse can be computed as

$$S^{-1}_{(\pi,0)}(0,T) = \left(\begin{array}{cc} -2\,\frac{\mathrm{e}^{4\,T} + 4\,T\mathrm{e}^{2\,T} - 1}{2\,\mathrm{e}^{2\,T} + 4\,T^2\mathrm{e}^{2\,T} - 1 - \mathrm{e}^{4\,T}} & 2\,\frac{2\,\mathrm{e}^{2\,T} - 1 - \mathrm{e}^{4\,T}}{2\,\mathrm{e}^{2\,T} + 4\,T^2\mathrm{e}^{2\,T} - 1 - \mathrm{e}^{4\,T}} \\ 2\,\frac{2\,\mathrm{e}^{2\,T} - 1 - \mathrm{e}^{4\,T}}{2\,\mathrm{e}^{2\,T} + 4\,T^2\mathrm{e}^{2\,T} - 1 - \mathrm{e}^{4\,T}} & 2\,\frac{1 + 4\,T\mathrm{e}^{2\,T} - \mathrm{e}^{4\,T}}{2\,\mathrm{e}^{2\,T} + 4\,T^2\mathrm{e}^{2\,T} - 1 - \mathrm{e}^{4\,T}} \end{array} \right), \quad T > 0. \tag{5.33}$$

The control law due to Kleinmann is

$$u(t) = (A - BB^T S^{-1}_{(\pi,0)}(0,T))x(t) \tag{5.34}$$

For $T = 2$ for example the matrix $A - BB^T S^{-1}_{(\pi,0)}(0,2)$ has a pair of complex eigenvalues with real part less than approximately -1.27.

Finally we will show that with the method of Bass (4.2.2) we can obtain a stabilizing feedback control without having too much computational effort.
In order to stabilize the zero-state of system (5.24) we first have to choose an adequate $\lambda > 0$ as indicated in (4.94). We obtain

$$\max_i \left(\sum_{j=1}^{2} |a_{ij}|\right) = \max_j \left(\sum_{i=1}^{2} |a_{ij}|\right) = 1 \tag{5.35}$$

and conclude that $\lambda \geq 2$ is a sufficient choice expecting that the zero state is an exponential stable equilibrium point of the controlled system with decay rate at least $-\lambda$. For $\lambda = 3$ we

obtain
We solve the Lyapunov equation (4.95) for system (5.24):

$$\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} P + P \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{5.36}$$

where P is a symmetric matrix which can be computed as

$$P = \begin{pmatrix} 1/96 & -1/32 \\ -1/32 & 17/96 \end{pmatrix} \tag{5.37}$$

with inverse

$$P^{-1} = \begin{pmatrix} 204 & 36 \\ 36 & 12 \end{pmatrix}. \tag{5.38}$$

The control law then is given by

$$u(t) = -B^T P^{-1} x(t) = \begin{pmatrix} 0 & 0 \\ -36 & -12 \end{pmatrix} x(t) \tag{5.39}$$

and the controlled system is given by

$$\dot{x}(t) = (A - BB^T P^{-1})x(t) = \underbrace{\begin{pmatrix} 0 & 1 \\ -35 & -12 \end{pmatrix}}_{\tilde{A}} x(t) \tag{5.40}$$

where $\tilde{A}$ has eigenvalues -5 and -7. Therefore the zero-state is an exponential stable equilibrium point of the controlled system with decay rate -5 which is better than the expected value -3 which is guaranteed by the design of the method. Remark: The zero-state of the λ-shifted system is exponential stable with decay rate -2.

5.1.3 Nonlinear simple pendulum

Consider the nonlinear pendulum system

$$\ddot{x}(t) + \sin x(t) = 0, \qquad t \geq 0 \tag{5.41}$$

which was obtained via the Lagrange formalism and where we renamed the variables for convenience. Any external forces enter this equation on the right hand side of (5.41). So the controlled nonlinear pendulum equation reads:

$$\ddot{x}(t) + \sin x(t) = u(t), \qquad t \geq 0, \tag{5.42}$$

Writing system (5.41) and (5.42) as first order system we obtain

$$\begin{aligned} \dot{x}(t) &= y(t) \\ \dot{y}(t) &= -\sin(x(t)) \end{aligned} \tag{5.43}$$

for the uncontrolled system and

$$\begin{aligned} \dot{x}(t) &= y(t) \\ \dot{y}(t) &= -\sin(x(t)) + u(t) \end{aligned} \tag{5.44}$$

for the controlled system.
To get a first glimpse of the dynamics of the system (5.43) we plot the vector field in the (x, y)-space along the coordinate axes, which is shown in figure (5.2):
The complete phase plot is shown in figure (5.4). The origin $(0, 0)$ of the coordinate center is a fixed point of the system. It is stable in the sense of Lyapunov but not asymptotically stable. So one might be interested in a control $u(t)$ which for example makes the origin asymptotically stable. The control u enters only in the second equation of (5.44). We can only change the velocity, i.e. the second component of $(x, y)^T$. Figure (5.2) shows the natural choice in stabilizing the origin. This solution shown in figure (5.3) has the physical interpretation of adding friction in form of a damping term. From figure (5.3) we see that a feedback function $u(t) = -ky(t)$ is a suitable choice for any $k > 0$. One can also expect that choosing u as feedback function depending on x will not make the origin asymptotically stable, but – when u is depending on x and y – might influence the decay rate.

Equation (5.41) can be solved exactly in terms of elliptic integrals. We will distinguish four kinds of motion:

1. Pendulum swings without reaching the upright position, changing the direction of its movement periodically
2. limit case where the pendulum tends up to the upright position without reaching it in finite time and without changing its direction (separatrix solution),
3. pendulum rotates in one direction
4. pendulum is at rest.

The first three cases can be distinguished by using the size of the initial energy of system (5.41). We can do this, because the total energy of the system does not depend on the time (i.e. the total energy is a *conservation law* with respect to (5.41)):

$$\frac{d}{dt}E = \frac{d}{dt}(E_{\text{kin}} + E_{\text{pot}}) = \dot{x}(t)\ddot{x}(t) + \dot{x}(t)\sin x(t) = \dot{x}(\ddot{x}(t) + \sin x(t)) \overset{(5.41)}{=} 0. \tag{5.45}$$

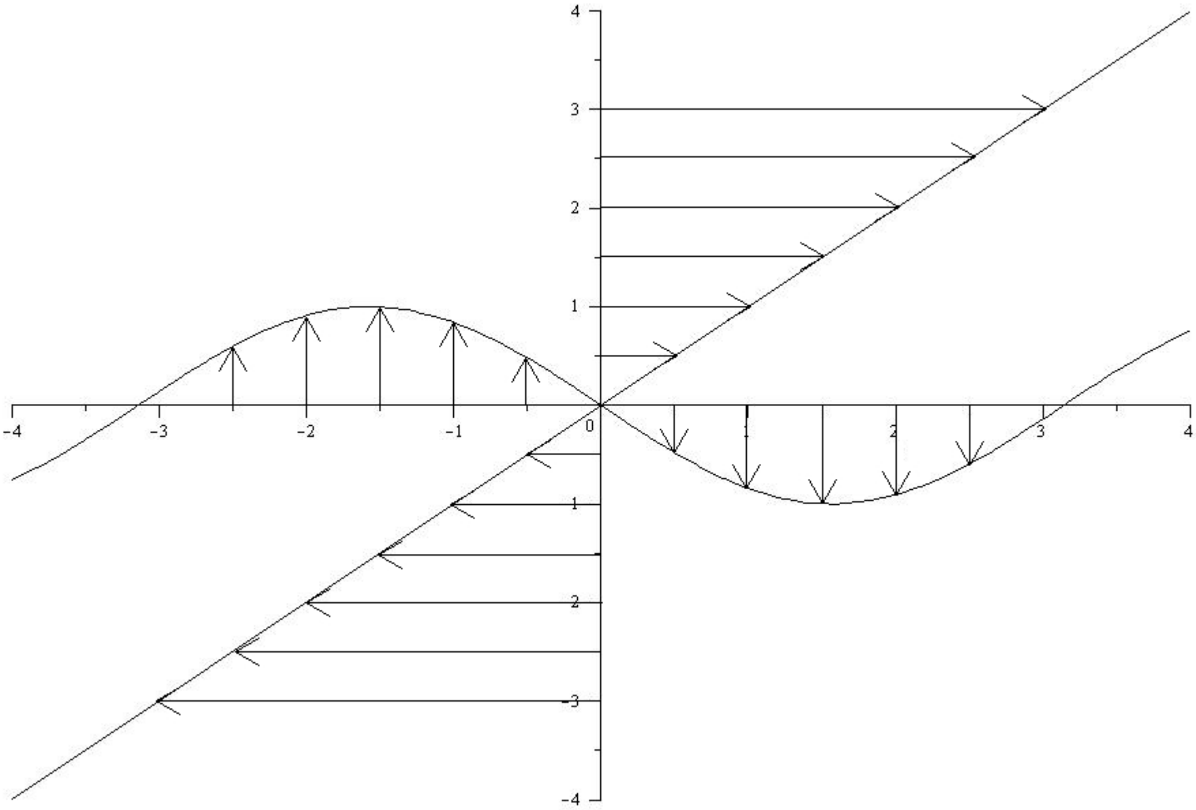


Figure 5.2: vector field of the nonlinear pendulum along the axes

Regarding the initial value problem

$$\ddot{x}(t) + \sin x(t) = 0; \quad x(0) = x_0,\ \dot{x}(0) = \dot{x}_0 \tag{5.46}$$

and denoting the initial energy E_0 we have due to (5.45) $E = E_0$ and we can make the following classification:

1. $\underline{E_0 < 1 \text{ and } E_0 \neq 0}$: pendulum swings,
2. $\underline{E = 1}$ and $x(0) \neq \pi$: separatrix solution,
3. $\underline{E > 1}$: rotation

The solutions where the pendulum is at rest are obtained by direct integration. One obtains the stable equilibrium point $x = 0$ and the unstable equilibrium point $x = \pi$.
Figure (5.4) shows different level sets for the energy of system (5.41). These level curves were obtained numerically by using a Runge-Kutta-Fehlberg method of order five. In figure (5.4) the separatrix solution consists of the two curves connecting $(-\pi, 0) \to (\pi, 0)$. Since the variable x is 2π periodic, the points $(\pi, 0)$ and $(-\pi, 0)$ denote the same state of system (5.41). The upright position is an equilibrium point, so one has to start at another point of the trajectory. The pendulum will move in one of the two direction towards the upright position, where it slows down as it comes closer to it. In finite time it will get arbitrarily close to the upright position, but it will never reach it.
The closed orbits correspond to the swinging pendulum whereas the remaining solution curves

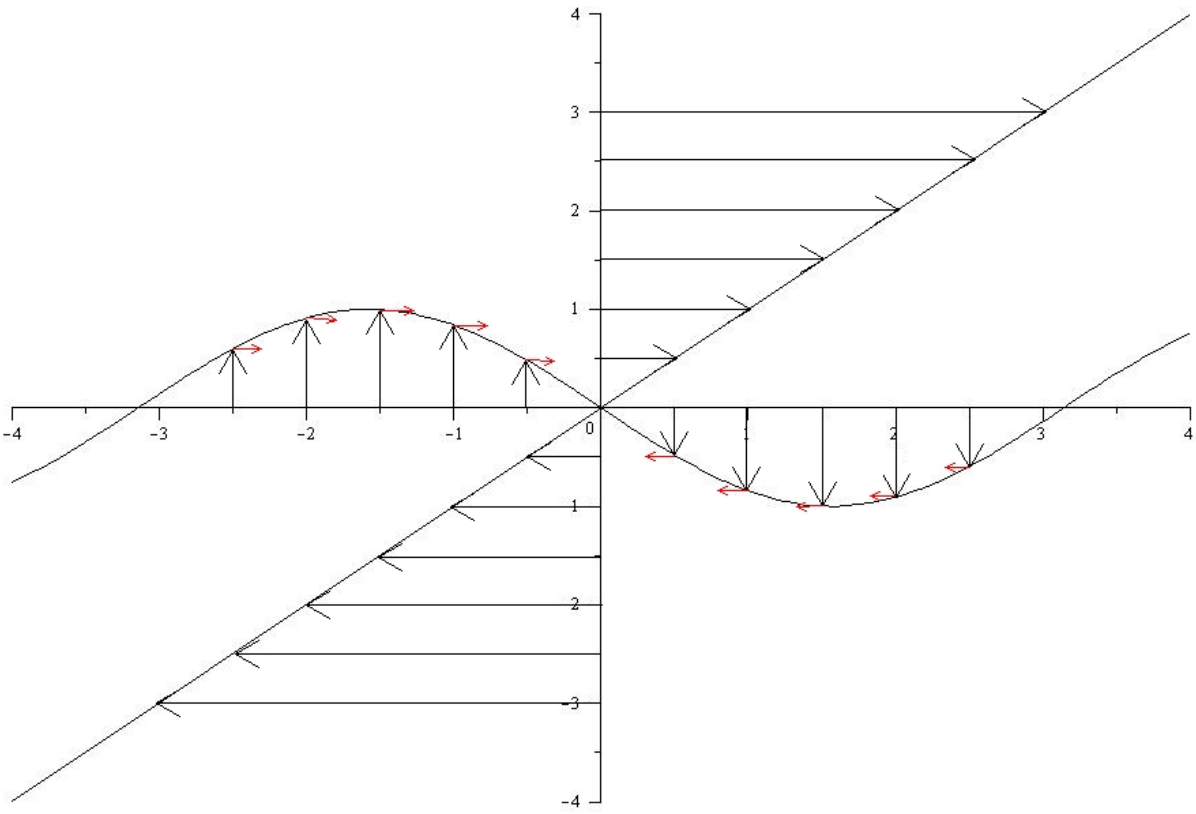


Figure 5.3: stabilizing the origin of the nonlinear pendulum equation

belong to the case, where the pendulum rotates about the angular point.

Solution as elliptic integral

We regard case 1 where $E_0 < 1$ and $E_0 \neq 0$ and give a mathematical description for solutions belonging to the swinging pendulum which does not pass the upright equilibrium point: Multiplying equation (5.41) with $2 \cdot \dot{x}(t)$ and integrating with respect to t gives

$$\dot{x}^2(t) - 2 \cdot \cos x(t) = C, \tag{5.47}$$

where C is some integration constant. We can determine C in the following way. The pendulum changes its direction when $\dot{x}(t) = 0$. Let the position be given by $\pm x_{\max}$. Since $\cos x_{\max} = \cos -x_{\max}$ we have

$$C = -2 \cos x_{\max}. \tag{5.48}$$

We obtain from (5.47)

$$\dot{x}^2(t) - 2\cos(x(t)) = 2\cos x_{\max} \tag{5.49}$$

$$\Longrightarrow \dot{x}(t) = \pm\sqrt{2\cos x(t) - 2\cos x_{\max}}, \tag{5.50}$$

where different signs correspond to different directions of the pendulum movement. Separation of variables – where we avoid an integration constant by assuming $x(0) = 0$ – leads to:

$$t = \int_0^x \frac{1}{\pm\sqrt{2\cos\tilde{x} - 2\cos x_{\max}}} d\tilde{x} \tag{5.51}$$

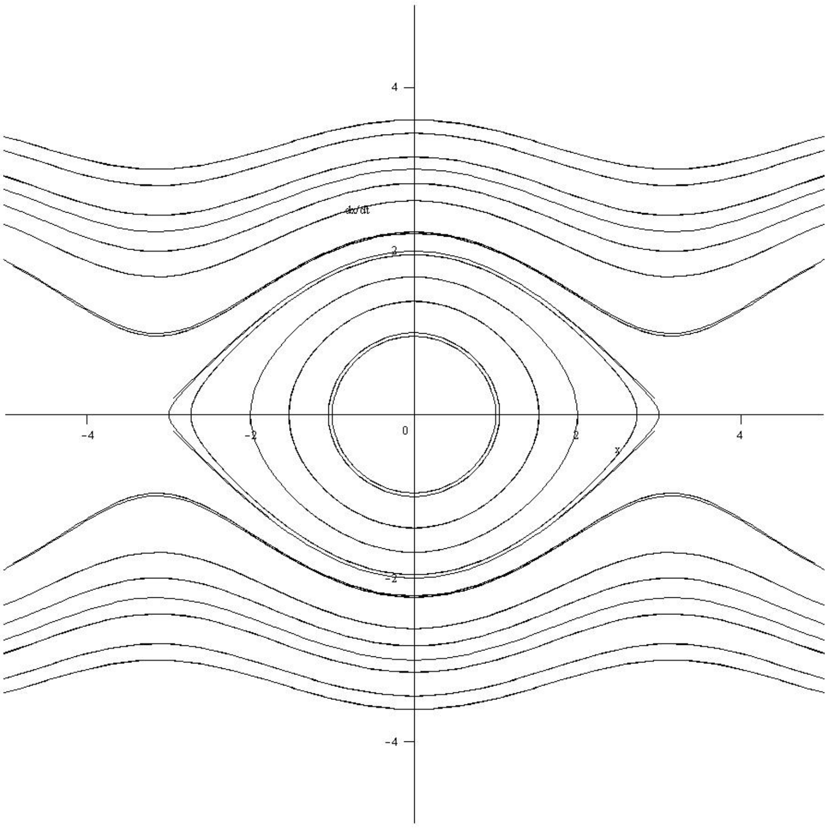


Figure 5.4: level sets of the pendulum

The (non-negative) radicand can be expressed as

$$2\cos\tilde{x} - 2\cos x_{\max} = 2\left(\cos^2\left(\frac{\tilde{x}}{2}\right) - \sin^2\left(\frac{\tilde{x}}{2}\right)\right) - 2\left(\cos^2\left(\frac{x_{\max}}{2}\right) - \sin^2\left(\frac{x_{\max}}{2}\right)\right). \tag{5.52}$$

Using the change of variables $k := \sin\left(\frac{x_{\max}}{2}\right)$ and $k\sin\psi := \sin\left(\frac{\tilde{x}}{2}\right)$ one obtains

$$\begin{aligned} 2\cos\tilde{x} - 2\cos x_{\max} =& 2(1 - 2k^2\sin^2\psi) - 2(1 - k^2) \\ =& 4k^2(1 - \sin^2\psi) \\ =& (2k\cos\psi)^2 \qquad (5.53) \\ \frac{d\tilde{x}}{d\psi} =& \frac{d}{d\psi} 2\arcsin(k\sin\psi) \\ d\tilde{x} =& 2\frac{k\cos\psi}{\sqrt{1 - k^2\sin^2\psi}} d\psi \qquad (5.54) \end{aligned}$$

By means of (5.53), (5.54) we can transform the integral in (5.51) into

$$\int_0^{\arcsin\frac{\sin\left(\frac{x}{2}\right)}{k}} \frac{\pm\operatorname{signum} k}{\sqrt{1 - k^2\sin^2\psi}} d\psi \tag{5.55}$$

which is an elliptic integral of the variables k and ψ.

Separatrix solution

For the energy level $E_0 = 1$ we obtain the separatrix solution. From equation (5.51) we can deduct

$$t = \int_0^x \frac{1}{\pm\sqrt{2\cos\tilde{x}+2}} d\tilde{x} = \int_0^x \frac{1}{\pm\sqrt{4\cos^2\left(\frac{\tilde{x}}{2}\right)}} d\tilde{x}. \tag{5.56}$$

For simplicity we drop the "$-$" sign and regard only the equation

$$t = \int_0^x \frac{1}{\sqrt{4\cos^2\left(\frac{\tilde{x}}{2}\right)}} d\tilde{x} \tag{5.57}$$

Since $\tilde{x} \in (-\pi, \pi)$ we have $\cos\left(\frac{\tilde{x}}{2}\right) > 0$ and therefore

$$t = \int_0^x \frac{1}{\left|2\cos\left(\frac{\tilde{x}}{2}\right)\right|} d\tilde{x} = \int_0^x \frac{1}{2\cos\left(\frac{\tilde{x}}{2}\right)} d\tilde{x} \tag{5.58}$$

The integral can be explicitly solved with Bronstein formula 325 (cf. [Bronstein and Semendjajew, 1991]) and we obtain

$$t = \ln\left(\sec\left(\frac{x}{2}\right) + \tan\left(\frac{x}{2}\right)\right) \tag{5.59}$$

The argument of the logarithm on the right hand side can be simplified to $\tan\left(\frac{x}{4} + \frac{\pi}{4}\right)$:

$$\begin{aligned}
\cos\left(\frac{x}{4}\right)\sin\left(\frac{x}{4}\right)\cos\left(\frac{x}{4}\right) &= \sin\left(\frac{x}{4}\right)\cos^2\left(\frac{x}{4}\right) \iff \\
\cos\left(\frac{x}{4}\right)\left(\cos^2\left(\frac{x}{4}\right) - \cos^2\left(\frac{x}{4}\right) + \sin\left(\frac{x}{4}\right)\cos\left(\frac{x}{4}\right)\right) &= \sin\left(\frac{x}{4}\right)\cos^2\left(\frac{x}{4}\right) \iff \\
\cos\left(\frac{x}{4}\right)\left(1 + \sin\left(\frac{x}{4}\right)\cos\left(\frac{x}{4}\right) - \sin^2\left(\frac{x}{4}\right) - \cos^2\left(\frac{x}{4}\right)\right) &= \sin\left(\frac{x}{4}\right)\left(1 - \sin^2\left(\frac{x}{4}\right)\right).
\end{aligned}$$

Expanding both sides and rearranging them gives

$$\begin{aligned}
\cos\left(\frac{x}{4}\right) + 2\sin\left(\frac{x}{4}\right)\cos^2\left(\frac{x}{4}\right) - \sin\left(\frac{x}{4}\right) - 2\sin^2\left(\frac{x}{4}\right)\cos\left(\frac{x}{4}\right) &= \\
\cos^2\left(\frac{x}{4}\right)\sin\left(\frac{x}{4}\right) - \sin^3\left(\frac{x}{4}\right) + \cos^3\left(\frac{x}{4}\right) - \sin^2\left(\frac{x}{4}\right)\cos\left(\frac{x}{4}\right) &\iff \\
\left(1 + 2\sin\left(\frac{x}{4}\right)\cos\left(\frac{x}{4}\right)\right)\left(\cos\left(\frac{x}{4}\right) - \sin\left(\frac{x}{4}\right)\right) &= \\
\left(\cos^2\left(\frac{x}{4}\right) - \sin^2\left(\frac{x}{4}\right)\right)\left(\sin\left(\frac{x}{4}\right) + \cos\left(\frac{x}{4}\right)\right) &\iff \\
\left(1 + \sin\left(\frac{x}{2}\right)\right)\cos\left(\frac{x+\pi}{4}\right) = \cos\left(\frac{x}{2}\right)\sin\left(\frac{x+\pi}{4}\right) &\iff \\
\frac{1 + \sin\left(\frac{x}{2}\right)}{\cos\left(\frac{x}{2}\right)} = \frac{\sin\left(\frac{x+\pi}{4}\right)}{\cos\left(\frac{x+\pi}{4}\right)} &\iff \\
\sec\left(\frac{x}{2}\right) + \tan\left(\frac{x}{2}\right) = \tan\left(\frac{x+\pi}{4}\right)
\end{aligned}$$

So equation (5.59) simplifies to

$$t = \ln\left(\sec\left(\frac{x}{2}\right) + \tan\left(\frac{x}{2}\right)\right) = \ln\left(\tan\left(\frac{x+\pi}{4}\right)\right) \tag{5.60}$$

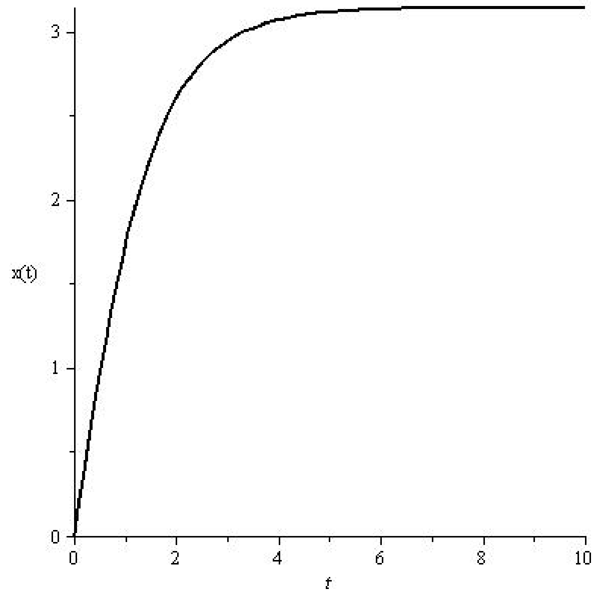


Figure 5.5: separatrix solution

and we obtain the explicit formula for the trajectory $x(t)$:

$$x(t) = -\pi + 4\arctan(\mathrm{e}^t). \tag{5.61}$$

Figure (5.5) shows the solution for $t \in [0, 10]$.
Using the *Gudermannian function* gd which is defined as

$$\mathrm{gd}(t) := \int_0^t \frac{1}{\cosh(\tau)} d\tau \tag{5.62}$$

we obtain

$$x(t) = 2\,\mathrm{gd}(t) \tag{5.63}$$

since

$$\begin{aligned}\mathrm{gd}(t) = &\int_0^t \frac{1}{\cosh(\tau)} d\tau = \arcsin(\tanh(t)) \\ =& \arctan(\sinh(t)) = 2\arctan(\tanh(t/2)) \\ =& 2\arctan(\mathrm{e}^t) - \frac{1}{2}\pi. \end{aligned} \tag{5.64}$$

Historical remark: The Gudermannian function is named after the german mathematician Christoph Gudermann, who was a student of Gauß and later one of Weierstrass's teachers. It links trigonometric functions and hyperbolic functions without using complex numbers.
The concept of *uniform convergence* appears for the first time in literature in one of Gudermann's papers about elliptic functions in the year 1838 (cf. e.g. [Schlote, 2002]).

Linearization along the separatrix

We can give an explicit form for the state transition matrix of the linearized system along the separatrix. Let $x^0(t)$ denote the separatrix solution $-\pi + 4\arctan(\exp(t))$. Then the linearized system along $x^0(t)$ is given by

$$\ddot{x}(t) + \cos(x^0(t))x(t) = 0 \tag{5.65}$$

Since for every smooth differential equation

$$\dot{x}(t) = f(x(t)), \quad x \in \mathbb{R}^n \tag{5.66}$$

and any solution $z(t)$ of (5.66) we have

$$\ddot{z}(t) = \frac{\partial f}{\partial x}(z(t))\dot{z}(t) \tag{5.67}$$

we know that $\dot{z}(t)$ is a solution of the linearized system of (5.66) along the trajectory $z(t)$. Therefore $\dot{x}^0(t)$ is a solution of (5.65).

$$\dot{x}^0(t) = \frac{4\mathrm{e}^t}{1+\mathrm{e}^{2t}} = \frac{4}{\mathrm{e}^t + \mathrm{e}^{-t}} = 2\frac{1}{\cosh(t)} = 2\,\mathrm{sech}(t). \tag{5.68}$$

With the reduction method of d'Alembert we can construct a second solution for (5.65). First we note that $\mathrm{sech}(t) > 0\,\forall t \geq 0$. We make the following ansatz

$$y(t) := 2\,\mathrm{sech}(t)z(t) \tag{5.69}$$

and obtain

$$\ddot{y}(t) = \left(2\frac{d^2}{dt^2}\,\mathrm{sech}(t) + 2\cos(x^0(t))\right)z(t) + 4\frac{d}{dt}\,\mathrm{sech}(t)\frac{d}{dt}z(t) + 2\,\mathrm{sech}(t)\frac{d^2}{dt^2}z(t) = 0 \tag{5.70}$$

and since $2\,\mathrm{sech}(t)$ is a solution of (5.65) differential equation (5.70) simplifies to

$$2\,\mathrm{sech}(t)\frac{d^2}{dt^2}z(t) + 4\frac{d}{dt}\,\mathrm{sech}(t)\frac{d}{dt}z(t) = 0 \tag{5.71}$$

The substitution $v(t) = \frac{d}{dt}z(t)$ leads to the scalar differential equation

$$2\,\mathrm{sech}(t)\frac{d}{dt}v(t) + 4\frac{d}{dt}\,\mathrm{sech}(t)v(t) = 0 \tag{5.72}$$

which is equivalent to

$$\frac{\dot{v}(t)}{v(t)} = -2\frac{\frac{d}{dt}\,\mathrm{sech}(t)}{\mathrm{sech}(t)} \tag{5.73}$$

which has as solution

$$v(t) = \cosh^2(t) \tag{5.74}$$

and therefore

$$z(t) = \int \cos^2(t)dt = \frac{1}{2}(t + \sinh(t)\cosh(t)) + C, \tag{5.75}$$

where $C \in \mathbb{R}$ is an integration constant.
A second solution of (5.65) is given by

$$y(t) = 2\,\mathrm{sech}(t)\frac{1}{2}(t + \sinh(t)\cosh(t)) + D. \tag{5.76}$$

We can choose D such that the initial condition $y(0) = (0,1)^T$ is fulfilled and obtain for system

$$\dot{x}(t) = \begin{pmatrix} 0 & 1 \\ -\cos(x^0(t)) & 0 \end{pmatrix} x(t), \quad x \in \mathbb{R}^2,\ t \geq 0 \tag{5.77}$$

the state transition matrix

$$\Phi(t,0) = \begin{pmatrix} \mathrm{sech}(t) & \frac{1}{2}\,\mathrm{sech}(t)(t + \sinh(t)\cosh(t)) - \frac{1}{2} \\ -\tanh(t)\,\mathrm{sech}(t) & \frac{1}{2}\,\mathrm{sech}(t)(\cosh^2(t) - t\tanh(t) + 1) \end{pmatrix}. \tag{5.78}$$

Since the linearization of the nonlinear pendulum equation about an arbitrary point is controllable, we expect the linearization along the separatrix to be controllable. The method of choice for smooth linear time-varying systems is given in theorem (3.9). In our example there must be a $t > 0$ such that

$$[M_0(t)|M_1(t)] \tag{5.79}$$

has rank two, where

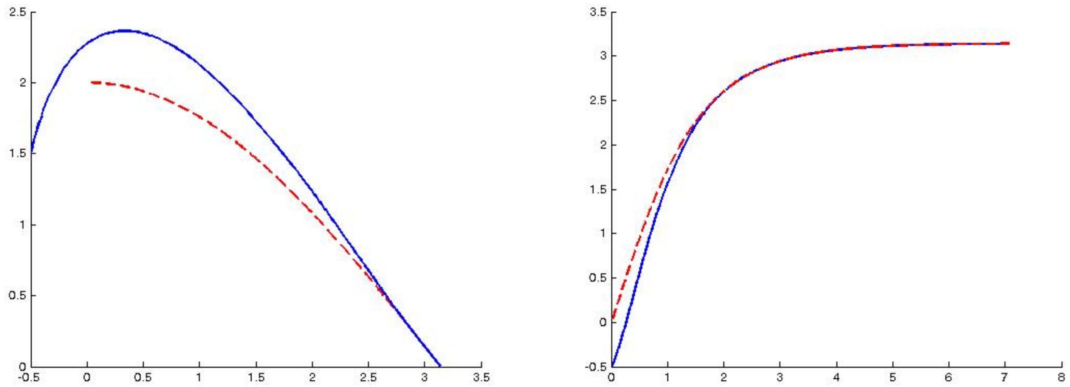
$$\begin{aligned} M_0(t) :=& B \\ M_1(t) :=& -\begin{pmatrix} 0 & 1 \\ -\cos(x^0(t))0 & \end{pmatrix} M_0(t) + \frac{d}{dt} M_0(t) \end{aligned}$$

are recursively defined proposed in theorem (3.9). The matrix (5.79) simplifies to $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ which has full rank.

Numerical results for the stabilization along the separatrix

We used the modified control law presented in section (4.2.6) to stabilize the nonlinear pendulum equation along the separatrix solution.

The exact solution starts from $(0,2)^T$, the trajectory to be controlled starts from $(-0.5,1.5)^T$. The stabilization method used is the modified control law with $\alpha = 0$:



(a) in phase space (b) first component (angle) vs. time

Figure 5.6: the controlled solution and the reference trajectory (broken line, separatrix)

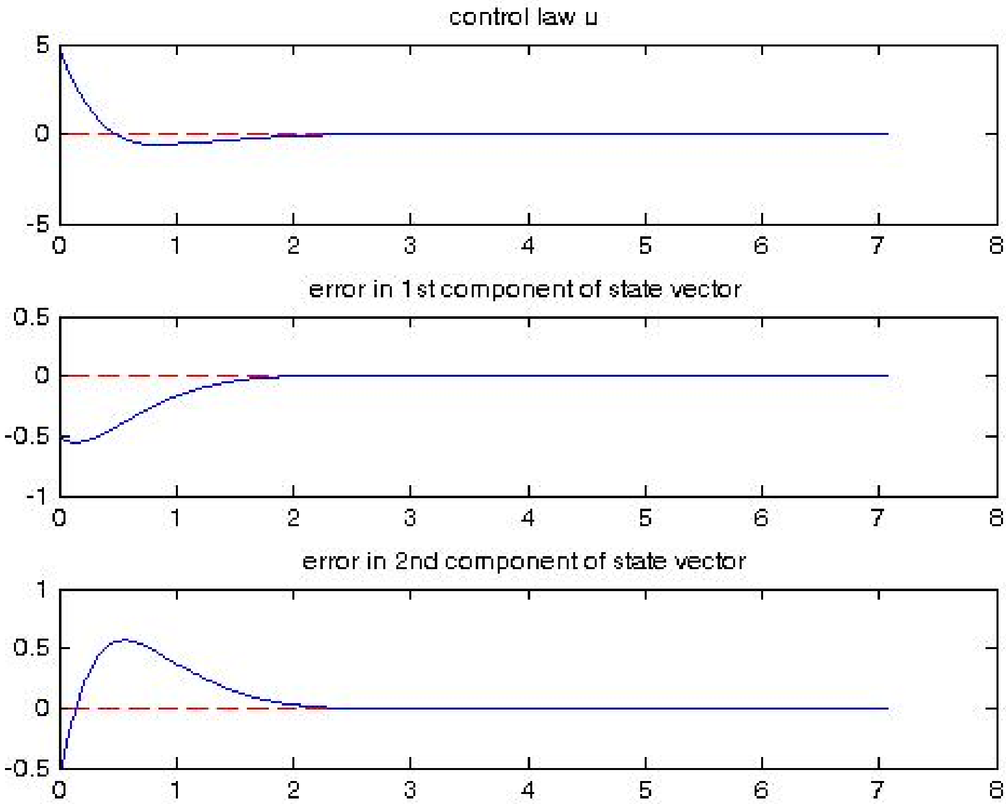
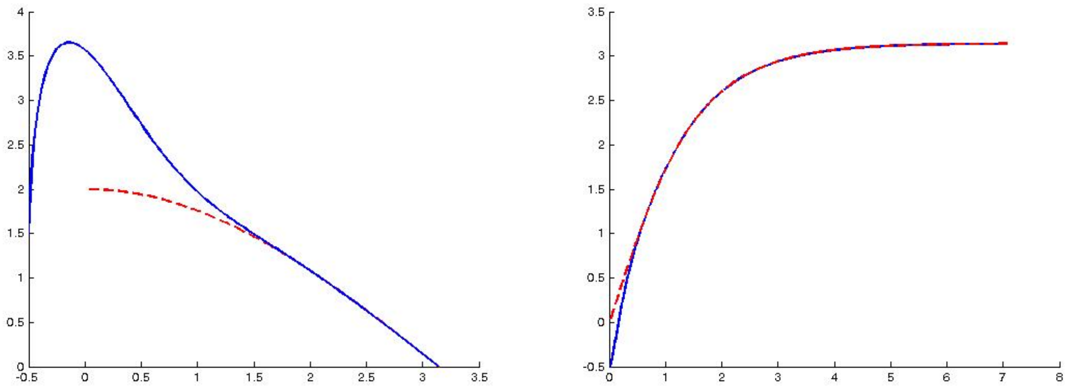


Figure 5.7: Control law and deviation in the single components

The exact solution starts from $(0, 2)^T$, the trajectory to be controlled starts from $(-0.5, 1.5)^T$. The stabilization method used is the modified control law with $\alpha = 2$:



(a) in phase space (b) first component (angle) vs. time

Figure 5.8: the controlled solution and the reference trajectory (broken line, separatrix)

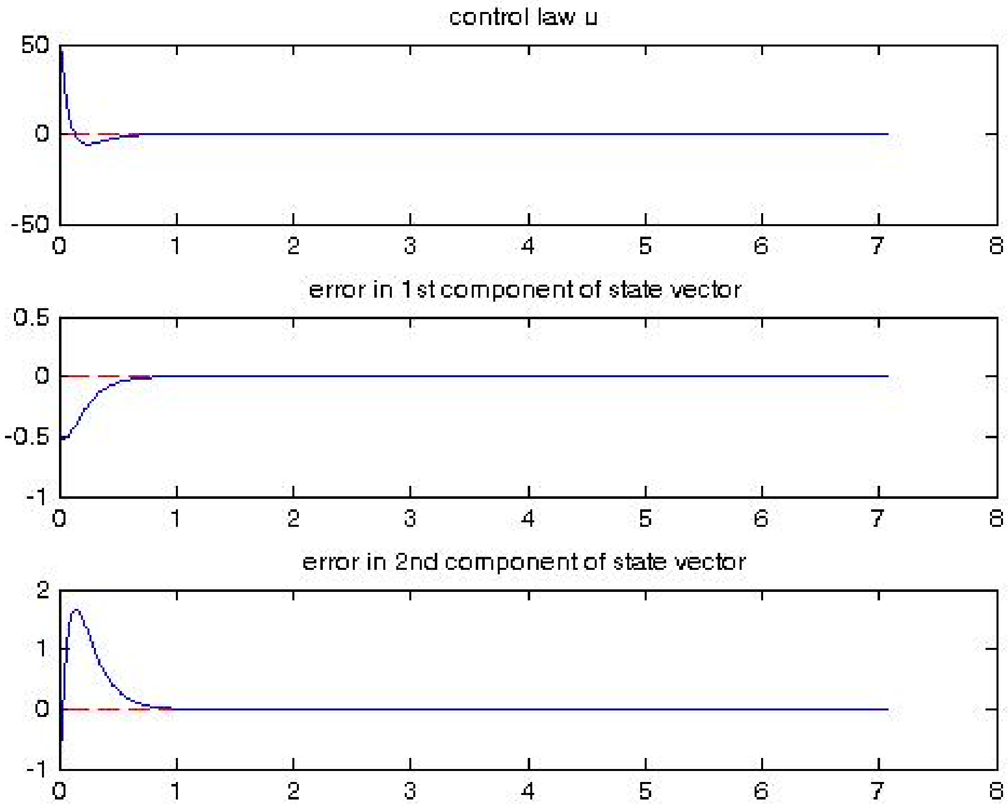
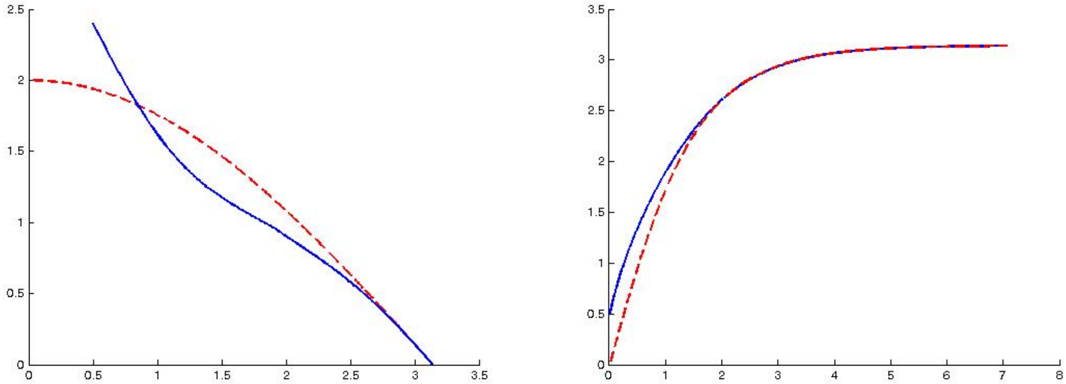


Figure 5.9: Control law and deviation in the single components

The exact solution starts from $(0, 2)^T$, the trajectory to be controlled starts from $(0.5, 2.4)^T$. The stabilization method used is the modified control law with $\alpha = 0$:



(a) in phase space (b) first component (angle) vs. time

Figure 5.10: the controlled solution and the reference trajectory (broken line, separatrix)

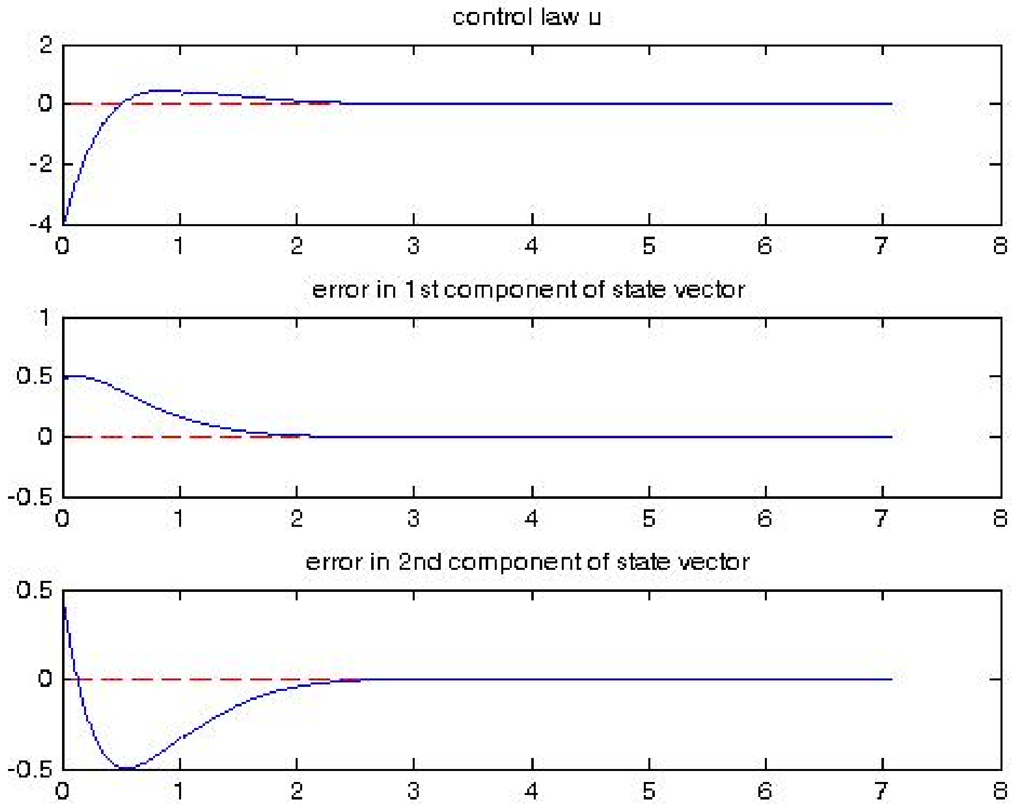
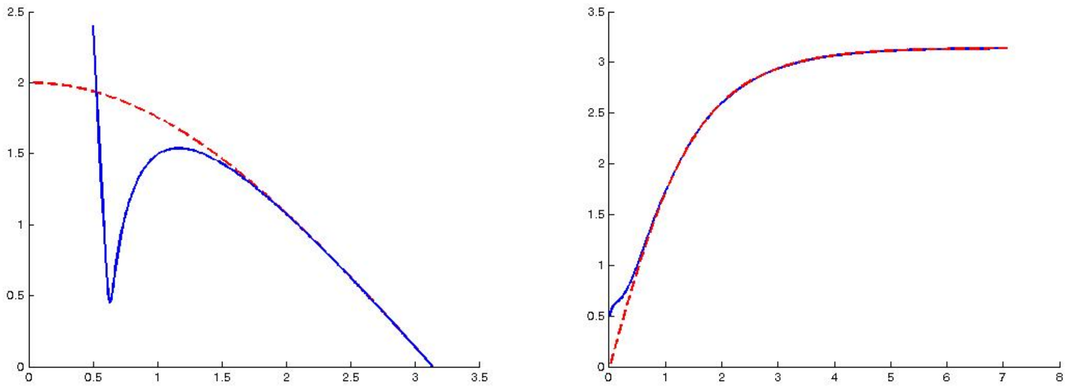


Figure 5.11: Control law and deviation in the single components

The exact solution starts from $(0, 2)^T$, the trajectory to be controlled starts from $(0.5, 2.4)^T$. The stabilization method used is the modified control law with $\alpha = 2$:



(a) in phase space (b) first component (angle) vs. time

Figure 5.12: the controlled solution and the reference trajectory (broken line, separatrix)

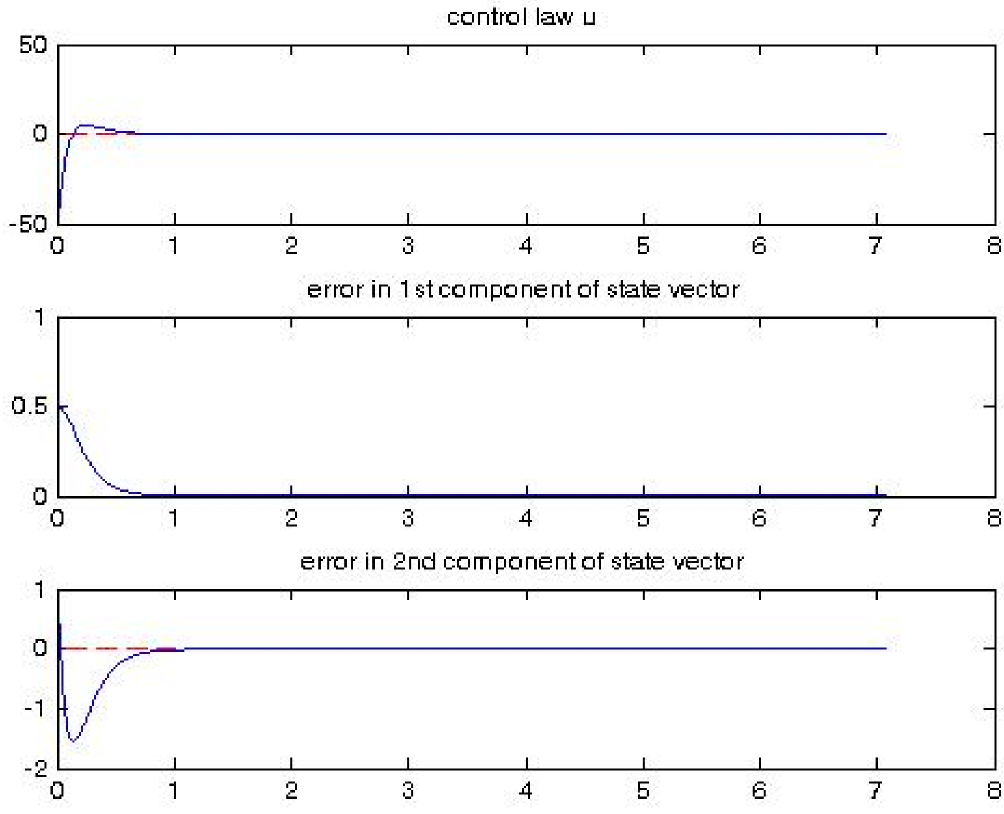
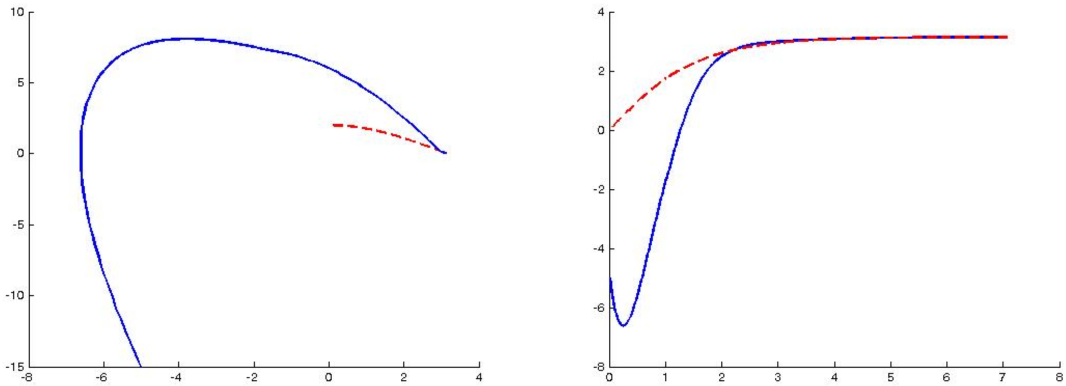


Figure 5.13: Control law and deviation in the single components

The exact solution starts from $(0, 2)^T$, the trajectory to be controlled starts from $(-5, -15)^T$. The stabilization method used is the modified control law with $\alpha = 0$:



(a) in phase space

(b) first component (angle) vs. time

Figure 5.14: the controlled solution and the reference trajectory (broken line, separatrix)

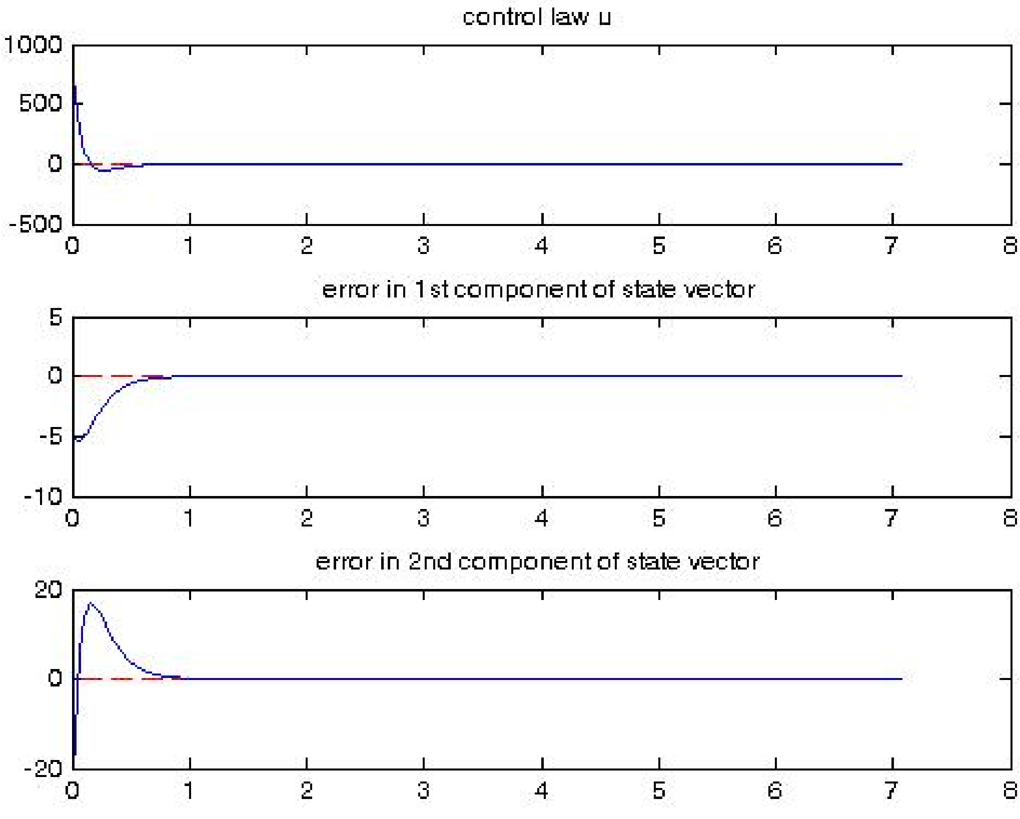
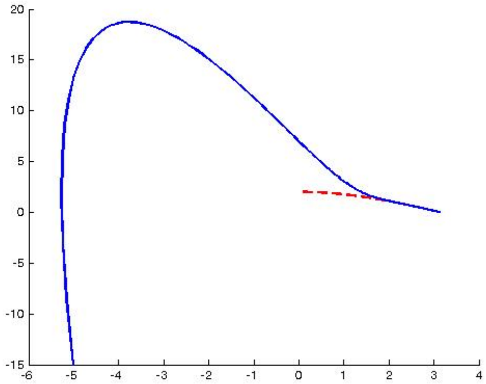
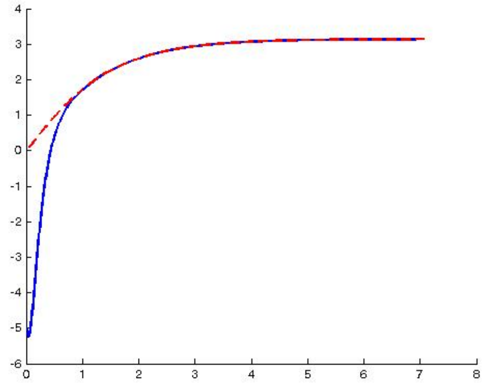


Figure 5.15: Control law and deviation in the single components

The exact solution starts from $(0, 2)^T$, the trajectory to be controlled starts from $(-5, -15)^T$. The stabilization method used is the modified control law with $\alpha = 2$:



(a) in phase space



(b) first component (angle) vs. time

Figure 5.16: the controlled solution and the reference trajectory (broken line, separatrix)

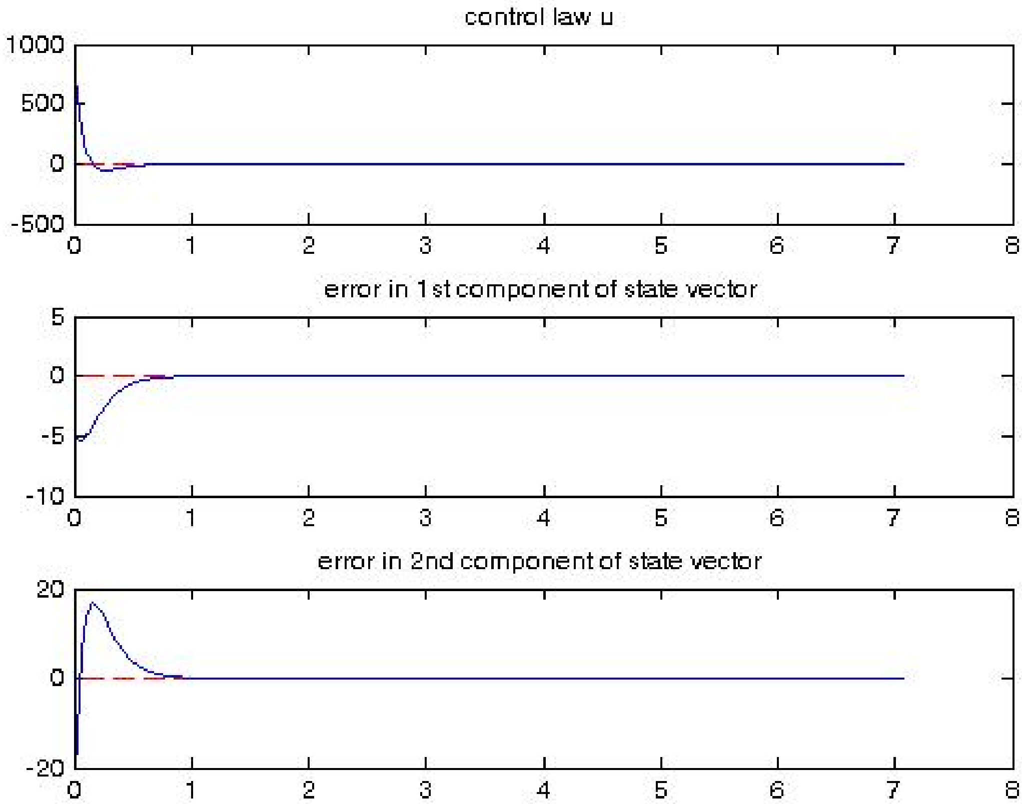


Figure 5.17: Control law and deviation in the single components

5.2 Double Pendulum

5.2.1 System dynamics

We consider a mathematical double pendulum making the usual assumption (no friction, massless pendulum links, motion occurs in a plane, no gravity, ...). The two (massless) links have the

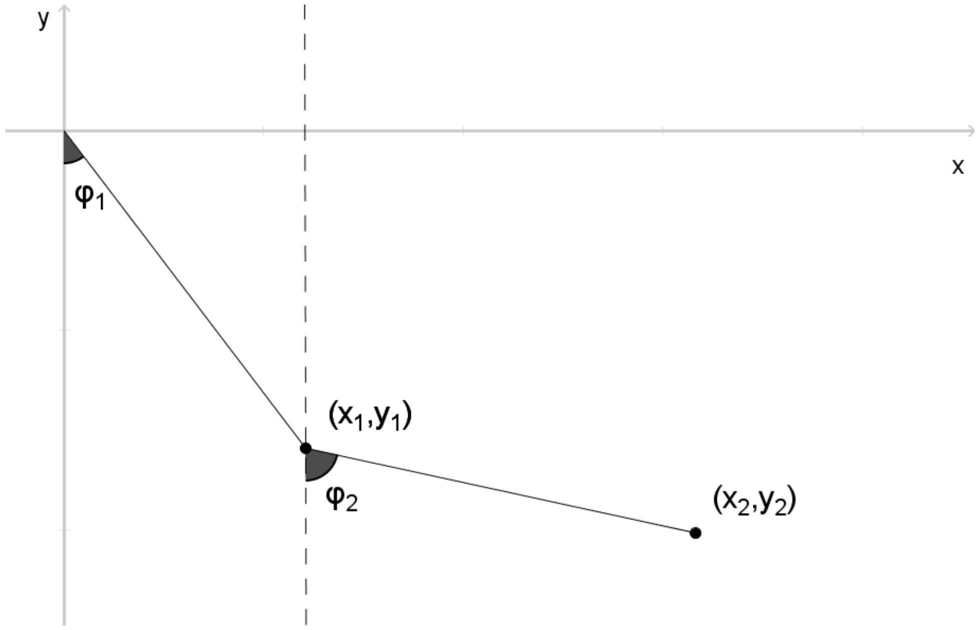


Figure 5.18: mathematical double pendulum

same length which are set equal to 1. The double pendulum is fixed at the origin of the coordinate center. The first pendulum ends at the point (x_1, y_1), the second one rotates about (x_1, y_1) and ends at the point (x_2, y_2). The angles are measured with respect to the negative y-axis and φ_1 denotes the angle corresponding to the first pendulum, φ_2 denotes the angle corresponding to the second pendulum. We derive the equations of motion via the Euler-Langrange formulation.

From figure (5.18) we have the simple relations

$$x_1 = \sin \varphi_1 \tag{5.80}$$
$$y_1 = \cos \varphi_1 \tag{5.81}$$
$$x_2 = \sin \varphi_1 + \sin \varphi_2 \tag{5.82}$$
$$y_2 = \cos \varphi_1 + \cos \varphi_2 \tag{5.83}$$

such that for the velocity v_1 of the first pendulum and the velocity v_2 of the second one the following relations hold

$$v_1^2 = \dot{x}_1^2 + \dot{y}_1^2 = \dot{\varphi}_1^2 \sin^2 \varphi_1 + \dot{\varphi}_1^2 \cos^2 \varphi_1 = \dot{\varphi}_1^2 \tag{5.84}$$
$$\begin{aligned} v_2^2 &= \dot{x}_2^2 + \dot{y}_2^2 = \left(\dot{\varphi}_1^2 \cos^2 \varphi_1 + \dot{\varphi}_2^2 \cos^2 \varphi_2\right)^2 + (\dot{\varphi}_1(-\sin \varphi_1) + \dot{\varphi}_2(-\sin \varphi_2))^2 \\ &= \dot{\varphi}_1^2 + \dot{\varphi}_2^2 + 2\dot{\varphi}_1\dot{\varphi}_2 \cos(\varphi_1 - \varphi_2). \end{aligned} \tag{5.85}$$

The kinetic energy E_{kin} and the potential energy E_{pot} of the double pendulum system are given by

$$\begin{aligned} E_{\mathrm{kin}} &= \frac{1}{2}v_1^2 + \frac{1}{2}v_2^2 = \frac{1}{2}\dot{\varphi}_1^2 + \frac{1}{2}\left(\dot{\varphi}_1^2 + \dot{\varphi}_2^2 + 2\dot{\varphi}_1\dot{\varphi}_2\cos(\varphi_1 - \varphi_2)\right) \\ &= \dot{\varphi}_1^2 + \frac{1}{2}\dot{\varphi}_2^2 + \dot{\varphi}_1\dot{\varphi}_2\cos(\varphi_1 - \varphi_2) \qquad (5.86) \\ E_{\mathrm{pot}} &= -\cos\varphi_1 - (\cos\varphi_1 + \cos\varphi_2) \\ &= -2\cos\varphi_1 - \cos\varphi_2. \qquad (5.87) \end{aligned}$$

The Langrangian L of the double pendulum system is defined as

$$L(\varphi_1, \varphi_2, \dot{\varphi}_1, \dot{\varphi}_2) := E_{\mathrm{kin}} - E_{\mathrm{pot}} = \dot{\varphi}_1^2 + \frac{1}{2}\dot{\varphi}_2^2 + \dot{\varphi}_1\dot{\varphi}_2\cos(\varphi_1 - \varphi_2) + 2\cos\varphi_1 + \cos\varphi_2. \qquad (5.88)$$

Using $\varphi = (\varphi_1, \varphi_2)^T$ and $\dot{\varphi} = (\dot{\varphi}_1, \dot{\varphi}_2)^T$ we can derive the equations of motion by evaluating the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\varphi}}(\varphi, \dot{\varphi})\right) - \frac{\partial L}{\partial \varphi}(\varphi, \dot{\varphi}) = 0. \qquad (5.89)$$

We have

$$\begin{aligned} \frac{\partial L}{\partial \dot{\varphi}_1}(\varphi, \dot{\varphi}) &= 2\dot{\varphi}_1 + \dot{\varphi}_2\cos(\varphi_1 - \varphi_2) \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}_1}(\varphi, \dot{\varphi}) &= 2\ddot{\varphi}_1 + \ddot{\varphi}_2\cos(\varphi_1 - \varphi_2) - \dot{\varphi}_2\sin(\varphi_1 - \varphi_2)(\dot{\varphi}_1 - \dot{\varphi}_2) \\ \frac{\partial L}{\partial \varphi_1}(\varphi, \dot{\varphi}) &= -\dot{\varphi}_1\dot{\varphi}_2\sin(\varphi_1 - \varphi_2) + 2\sin\varphi_1 \\ \frac{\partial L}{\partial \dot{\varphi}_2}(\varphi, \dot{\varphi}) &= \dot{\varphi}_2 + \dot{\varphi}_1\cos(\varphi_1 - \varphi_2) \\ \frac{d}{dt}\frac{\partial L}{\partial \dot{\varphi}_2}(\varphi, \dot{\varphi}) &= \ddot{\varphi}_2 + \ddot{\varphi}_1\cos(\varphi_1 - \varphi_2) + \dot{\varphi}_1\sin(\varphi_1 - \varphi_2)(\dot{\varphi}_2 - \dot{\varphi}_1) \\ \frac{\partial L}{\partial \varphi_2}(\varphi, \dot{\varphi}) &= -\sin\varphi_2 + \dot{\varphi}_1\dot{\varphi}_2\sin(\varphi_1 - \varphi_2) \end{aligned}$$

and with the Euler-Lagrange equation (5.89) we obtain the equations of motion of the mathematical double pendulum:

$$2\ddot{\varphi}_1 + \ddot{\varphi}_2\cos(\varphi_1 - \varphi_2) + \dot{\varphi}_2^2\sin(\varphi_1 - \varphi_2) + 2\sin\varphi_1 = 0 \qquad (5.90)$$

$$\ddot{\varphi}_2 + \ddot{\varphi}_1\cos(\varphi_1 - \varphi_2) - \dot{\varphi}_1^2\sin(\varphi_1 - \varphi_2) + \sin\varphi_2 = 0. \qquad (5.91)$$

which can be written as

$$D(\varphi)\ddot{\varphi} + C(\varphi, \dot{\varphi})\dot{\varphi} + g(\varphi) = 0 \qquad (5.92)$$

where

$$\begin{aligned} D(\varphi) &= \begin{pmatrix} 2 & \cos(\varphi_1 - \varphi_2) \\ \cos(\varphi_1 - \varphi_2) & 1 \end{pmatrix} \qquad (5.93) \\ C(\varphi, \dot{\varphi}) &= \begin{pmatrix} 0 & \dot{\varphi}_2\sin(\varphi_1 - \varphi_2) \\ -\dot{\varphi}_1\sin(\varphi_1 - \varphi_2) & 0 \end{pmatrix} \qquad (5.94) \\ g(\varphi) &= \begin{pmatrix} 2\sin\varphi_1 \\ \sin\varphi_2 \end{pmatrix}. \qquad (5.95) \end{aligned}$$

Remark: Note that $D(\varphi)$ is symmetric and $D(\varphi)_{11} > 0$. Since $\det(D(\varphi)) = 2 - \cos^2(\varphi_1 - \varphi_2)$ is positive as well, $D(\varphi)$ is positive definite for all φ. This form is useful in showing that local controllability around the equilibrium points with a scalar control input. We give a more general proof for the pendulum with n links in theorem C.1 of appendix C.3.

Introducing new variables $\omega_1 = \dot{\varphi}_1$ and $\omega_2 = \dot{\varphi}_2$ in order to rewrite the equations of motion as system of first order.
From (5.91) we have

$$\dot{\omega}_2 = -\sin\varphi_2 + \omega_1^2 \sin(\varphi_1 - \varphi_2) - \dot{\omega}_1 \cos(\varphi_1 - \varphi_2) \tag{5.96}$$

and inserting into (5.90) yields

$$\begin{aligned}\dot{\omega}_1(2 - \cos^2(\varphi_1 - \varphi_2)) =& \sin\varphi_2 \cos(\varphi_1 - \varphi_2) - \omega_1^2 \cos(\varphi_1 - \varphi_2)\sin(\varphi_1 - \varphi_2) - \\ & - \omega_2^2 \sin(\varphi_1 - \varphi_2) - 2\sin\varphi_1.\end{aligned} \tag{5.97}$$

Using the trigonometric relations

$$2\cos(\varphi_1 - \varphi_2)\sin\varphi_2 = \sin\varphi_1 - \sin(\varphi_1 - 2\varphi_2) \tag{5.98}$$

$$2\cos^2(\varphi_1 - \varphi_2) = 1 + \cos(2\varphi_1 - 2\varphi_2) \tag{5.99}$$

and solving for $\dot{\omega}_1$ we obtain

$$\dot{\omega}_1 = \frac{-3\sin\varphi_1 - \sin(\varphi_1 - 2\varphi_2) - 2\sin(\varphi_1 - \varphi_2)(\omega_2^2 + \omega_1^2 \cos(\varphi_1 - \varphi_2))}{3 - \cos(2\varphi_1 - 2\varphi_2)} \tag{5.100}$$

From (5.90) we have

$$\dot{\omega}_1 = -\frac{1}{2}\dot{\omega}_2 \cos(\varphi_1 - \varphi_2) - \frac{1}{2}\omega_2^2 \sin(\varphi_1 - \varphi_2) - \sin\varphi_1 \tag{5.101}$$

and inserting into (5.91) yields

$$\begin{aligned}\dot{\omega}_2(\cos^2(\varphi_1 - \varphi_2) - 2) =& -\omega_2^2 \cos(\varphi_1 - \varphi_2)\sin(\varphi_1 - \varphi_2) - 2\cos(\varphi_1 - \varphi_2)\sin\varphi_1 - \\ & - 2\omega_1^2 \sin(\varphi_1 - \varphi_2) + 2\sin\varphi_2.\end{aligned} \tag{5.102}$$

Together with the trigonometric relations (5.98) and (5.99) we obtain

$$\dot{\omega}_2 = \frac{2\sin(\varphi_1 - \varphi_2)(2\omega_1^2 + 2\cos\varphi_1 + \omega_2^2 \cos(\varphi_1 - \varphi_2)}{3 - \cos(2\varphi_1 - 2\varphi_2)} \tag{5.103}$$

and the equations of motions are given by the system of first order differential equations

$$\begin{cases} \dot{\varphi}_1 &= \omega_1 \\ \dot{\varphi}_2 &= \omega_2 \\ \dot{\omega}_1 &= \frac{-3\sin\varphi_1 - \sin(\varphi_1 - 2\varphi_2) - 2\sin(\varphi_1 - \varphi_2)(\omega_2^2 + \omega_1^2\cos(\varphi_1 - \varphi_2))}{3 - \cos(2\varphi_1 - 2\varphi_2)} \\ \dot{\omega}_2 &= \frac{2\sin(\varphi_1 - \varphi_2)(2\omega_1^2 + 2\cos\varphi_1 + \omega_2^2\cos(\varphi_1 - \varphi_2))}{3 - \cos(2\varphi_1 - 2\varphi_2)}. \end{cases} \tag{5.104}$$

5.2.2 Discussion of the linearized double pendulum

Linearization along a trajectory $(\tilde\varphi_1, \tilde\varphi_2, \tilde\omega_1, \tilde\omega_2)^T$ yields the linear differential equation

$$\frac{d}{dt}\begin{pmatrix}\varphi_1\\ \varphi_2\\ \omega_1\\ \omega_2\end{pmatrix} = \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ a_{31} & a_{32} & a_{33} & a_{34}\\ a_{41} & a_{42} & a_{43} & a_{44}\end{pmatrix}\begin{pmatrix}\varphi_1\\ \varphi_2\\ \omega_1\\ \omega_2\end{pmatrix} \tag{5.105}$$

where the entries a_{ij}, $3 \le i \le 4$, $1 \le j \le 4$ are given by

$$\begin{aligned} a_{31} =& -2\,\frac{\left(-3\,\sin(\tilde\varphi_1) - \sin(\tilde\varphi_1 - 2\,\tilde\varphi_2) - 2\,\sin(\tilde\varphi_1 - \tilde\varphi_2)\left(\tilde\omega_2^2 + \tilde\omega_1^2\cos(\tilde\varphi_1 - \tilde\varphi_2)\right)\right)}{(3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2))^2\,(\sin(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2))^{-1}} + \\ &+ \frac{-2\,\cos(\tilde\varphi_1 - \tilde\varphi_2)\left(\tilde\omega_2^2 + \tilde\omega_1^2\cos(\tilde\varphi_1 - \tilde\varphi_2)\right) + 2\,(\sin(\tilde\varphi_1 - \tilde\varphi_2))^2\,\omega_1{}^2}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} - \\ &- \frac{3\,\cos(\tilde\varphi_1) + \cos(\tilde\varphi_1 - 2\,\tilde\varphi_2)}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} \end{aligned} \tag{5.106}$$

$$\begin{aligned} a_{32} =& \frac{2\,\cos(\tilde\varphi_1 - 2\,\tilde\varphi_2) + 2\,\cos(\tilde\varphi_1 - \tilde\varphi_2)\left(\tilde\omega_2^2 + \tilde\omega_1^2\cos(\tilde\varphi_1 - \tilde\varphi_2)\right) - 2\,(\sin(\tilde\varphi_1 - \tilde\varphi_2))^2\,\tilde\omega_1^2}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} + \\ &+ 2\,\frac{\left(-3\,\sin(\tilde\varphi_1) - \sin(\tilde\varphi_1 - 2\,\tilde\varphi_2) - 2\,\sin(\tilde\varphi_1 - \tilde\varphi_2)\left(\tilde\omega_2^2 + \tilde\omega_1^2\cos(\tilde\varphi_1 - \tilde\varphi_2)\right)\right)}{(3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2))^2\,(\sin(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2))^{-1}} \end{aligned} \tag{5.107}$$

$$a_{33} = -4\,\frac{\sin(\tilde\varphi_1 - \tilde\varphi_2)\,\tilde\omega_1\cos(\tilde\varphi_1 - \tilde\varphi_2)}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} \tag{5.108}$$

$$a_{34} = -4\,\frac{\sin(\tilde\varphi_1 - \tilde\varphi_2)\,\tilde\omega_2}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} \tag{5.109}$$

$$\begin{aligned} a_{41} =& 2\,\frac{\cos(\tilde\varphi_1 - \tilde\varphi_2)\left(2\,\tilde\omega_1^2 + 2\,\cos(\tilde\varphi_1) + \tilde\omega_2^2\cos(\tilde\varphi_1 - \tilde\varphi_2)\right)}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} + \\ &+ 2\,\frac{\sin(\tilde\varphi_1 - \tilde\varphi_2)\left(-2\,\sin(\tilde\varphi_1) - \tilde\omega_2^2\sin(\tilde\varphi_1 - \tilde\varphi_2)\right)}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} - \\ &- 4\,\frac{\sin(\tilde\varphi_1 - \tilde\varphi_2)\left(2\,\tilde\omega_1^2 + 2\,\cos(\tilde\varphi_1) + \tilde\omega_2^2\cos(\tilde\varphi_1 - \tilde\varphi_2)\right)\sin(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)}{(3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2))^2} \end{aligned} \tag{5.110}$$

$$\begin{aligned} a_{42} =& -2\,\frac{\cos(\tilde\varphi_1 - \tilde\varphi_2)\left(2\,\tilde\omega_1^2 + 2\,\cos(\tilde\varphi_1) + \tilde\omega_2^2\cos(\tilde\varphi_1 - \tilde\varphi_2)\right)}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} + \\ &+ 2\,\frac{(\sin(\tilde\varphi_1 - \tilde\varphi_2))^2\,\tilde\omega_2^2}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} + \\ &+ 4\,\frac{\sin(\tilde\varphi_1 - \tilde\varphi_2)\left(2\,\tilde\omega_1^2 + 2\,\cos(\tilde\varphi_1) + \tilde\omega_2^2\cos(\tilde\varphi_1 - \tilde\varphi_2)\right)\sin(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)}{(3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2))^2} \end{aligned} \tag{5.111}$$

$$a_{43} = 8\,\frac{\sin(\tilde\varphi_1 - \tilde\varphi_2)\,\tilde\omega_1}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)} \tag{5.112}$$

$$a_{44} = 4\,\frac{\sin(\tilde\varphi_1 - \tilde\varphi_2)\,\tilde\omega_2\cos(\tilde\varphi_1 - \tilde\varphi_2)}{3 - \cos(2\,\tilde\varphi_1 - 2\,\tilde\varphi_2)}. \tag{5.113}$$

The matrix resulting from the linearization along a trajectory usually depends on the time. In case the reference trajectory reduces to a single point, the matrix of the linearization is independent of the time.

Controllability of the linearization around an arbitrary point

The linearization around a given point $(\varphi_1^*, \varphi_2^*, \omega_1^*, \omega_2^*)^T$ can be seen as a special case of linearization along a trajectory with $(\tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\omega}_1, \tilde{\omega}_2)^T = (\varphi_1^*, \varphi_2^*, \omega_1^*, \omega_2^*)^T$ and we obtain the linear control system which can be written as

$$\frac{d}{dt}\begin{pmatrix}\varphi_1(t)\\ \varphi_2(t)\\ \omega_1(t)\\ \omega_2(t)\end{pmatrix} = \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ a_{31}^* & a_{32}^* & a_{33}^* & a_{34}^*\\ a_{41}^* & a_{42}^* & a_{43}^* & a_{44}^*\end{pmatrix}\begin{pmatrix}\varphi_1(t)\\ \varphi_2(t)\\ \omega_1(t)\\ \omega_2(t)\end{pmatrix} + \begin{pmatrix}0 & 0\\ 0 & 0\\ 1 & 0\\ 0 & 1\end{pmatrix} u(t). \tag{5.114}$$

It is easy to see that due to the Kalman rank condition (3.11) we have global controllability of the linearization independent of the choice of $(\varphi_1^*, \varphi_2^*, \omega_1^*, \omega_2^*)^T$ since we have

$$\operatorname{rank}[B|A_{(\varphi_1^*, \varphi_2^*, \omega_1^*, \omega_2^*)}B|\ldots] = \operatorname{rank}\begin{bmatrix}0 & 0 & 1 & 0 & \ldots\\ 0 & 0 & 0 & 1 & \ldots\\ 1 & 0 & 0 & 0 & \ldots\\ 0 & 1 & 0 & 0 & \ldots\end{bmatrix} = 4. \tag{5.115}$$

Due to theorem (3.20) the nonlinear system (5.104) is locally controllable at its equilibrium points.

5.2.3 Stability of the equilibria

For $(\varphi_1^e, \varphi_2^e, \omega_1^e, \omega_2^e)$ to be an equilibrium point the right-hand side of (5.104) must vanish, which is the case if and only if

$$\omega_1^e = \omega_2^e = 0 \tag{5.116}$$

$$-3\sin(\varphi_1^e) - \sin(\varphi_1^e - 2\varphi_2^e) = 0 \tag{5.117}$$

$$\sin(\varphi_1^e - \varphi_2^e)\cos(\varphi_1^e) = 0 \tag{5.118}$$

From the second equation we can see that either

$$\varphi_1^e = k \cdot \pi + \varphi_2^e, \quad k \in \mathbb{Z} \tag{5.119}$$

must hold or

$$\varphi_1^e = \frac{\pi}{2} + k \cdot \pi, \quad k \in \mathbb{Z}. \tag{5.120}$$

Since for (5.120) the first term $-3\sin(\varphi_1^e)$ of (5.117) attains its maximum in absolute value which is always greater than the second term $-\sin(\varphi_1^e - 2\varphi_2^e)$. Therefore an equilibrium point $(\varphi_1^e, \varphi_2^e)$ must satisfy (5.119).
Inserting (5.119) into (5.117) leads to

$$\begin{aligned} 0 &= -3\sin(\varphi_2^e + k\pi) - \sin(\varphi_2^e + k\pi - 2\varphi_2^e)\\ \Longleftrightarrow 0 &= -3\sin(\varphi_2^e + k\pi) + \sin(\varphi_2^e - k\pi)\\ \Longleftrightarrow 0 &= -2\underbrace{\cos(k\pi)}_{=(-1)^k}\sin(\varphi_2^e) \end{aligned} \tag{5.121}$$

$$\tag{5.122}$$

and since φ_2^e is 2π-periodic it has to be either 0 or π such that we obtain the equilibrium points

$$(0,0,0,0), \quad (0,\pi,0,0), \quad (\pi,0,0,0), \quad (\pi,\pi,0,0) \tag{5.123}$$

which can be physically interpreted as follows:

- $(0,0,0,0)$ both pendulum links point downward,
- $(0,\pi,0,0)$ the first pendulum link points downward, whereas the second one points upward,
- $(\pi,0,0,0)$ the first pendulum link points upward, whereas the second one points downward,
- $(\pi,\pi,0,0)$ both pendulum links point upward.

For the equilibrium point $(0,0,0,0)^T$ we obtain the linear system

$$\frac{d}{dt}\begin{pmatrix}\varphi_1(t)\\ \varphi_2(t)\\ \omega_1(t)\\ \omega_2(t)\end{pmatrix} = \underbrace{\begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -2 & 1 & 0 & 0\\ 2 & -2 & 0 & 0\end{pmatrix}}_{=:A_{(0,0,0,0)}}\begin{pmatrix}\varphi_1(t)\\ \varphi_2(t)\\ \omega_1(t)\\ \omega_2(t)\end{pmatrix} \tag{5.124}$$

where the matrix $A_{(0,0,0,0)}$ has eigenvalues $\pm\sqrt{2-\sqrt{2}}\cdot i, \pm\sqrt{2+\sqrt{2}}\cdot i \in \mathbb{C}$ and due to theorem (4.9) the zero-solution of (5.124) is stable but not asymptotically stable.

For the equilibrium points $(\pi,0,0,0)^T$, $(0,\pi,0,0)^T$, $(\pi,\pi,0,0)^T$ the system matrices of the linear systems are given by

$$A_{(\pi,0,0,0)} := \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 2 & -1 & 0 & 0\\ 2 & -2 & 0 & 0\end{pmatrix} \tag{5.125}$$

$$A_{(0,\pi,0,0)} := \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ -2 & 1 & 0 & 0\\ -2 & 2 & 0 & 0\end{pmatrix} \tag{5.126}$$

$$A_{(\pi,\pi,0,0)} := \begin{pmatrix}0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\\ 2 & -1 & 0 & 0\\ -2 & 2 & 0 & 0\end{pmatrix} \tag{5.127}$$

where

- $A_{(\pi,0,0,0)}$ has eigenvalues $\pm 2^{\frac{1}{4}}, \pm 2^{\frac{1}{4}}\cdot i$ and therefore the zero-solution is unstable,
- $A_{(0,\pi,0,0)}$ has eigenvalues $\pm 2^{\frac{1}{4}}, \pm 2^{\frac{1}{4}}\cdot i$, the zero-solution is unstable,
- $A_{(\pi,\pi,0,0)}$ has eigenvalues $\pm\sqrt{2+\sqrt{2}}, \pm\sqrt{2-\sqrt{2}}$ and again the zero-solution is unstable.

If we have an explicit form of the state transition matrix, theorem (4.5) can be used to determine stability properties.

For $A_{(0,0,0,0)}$ the state transition matrix $\Phi_{A_{(0,0,0,0)}}(t,t_0)$ for $t \geq t_0$ has entries

$$\Phi_{11} = 1/2\cos\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) + 1/2\cos\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) \tag{5.128}$$

$$\Phi_{12} = 1/4\sqrt{2}\left(-\cos\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) + \cos\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\right) \tag{5.129}$$

$$\begin{aligned}\Phi_{13} = {}& (\sqrt{2-\sqrt{2}})^{-1}\left(1/2\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right)\sqrt{2} - 1/2\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right)\right) - \\ & - 1/4\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\sqrt{2} + \sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)(\sqrt{2-\sqrt{2}})^{-1}\end{aligned} \tag{5.130}$$

$$\Phi_{14} = -1/4\sqrt{2-\sqrt{2}}\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) + 1/4\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) \tag{5.131}$$

$$\Phi_{21} = 1/2\sqrt{2}\left(-\cos\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) + \cos\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\right) \tag{5.132}$$

$$\Phi_{22} = 1/2\cos\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) + 1/2\cos\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) \tag{5.133}$$

$$\Phi_{23} = -1/2\sqrt{2-\sqrt{2}}\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) + 1/2\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) \tag{5.134}$$

$$\begin{aligned}\Phi_{24} = {}& (\sqrt{2-\sqrt{2}})^{-1}\left(1/2\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right)\sqrt{2} - 1/2\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right)\right) - \\ & - 1/4\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\sqrt{2} + \sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)(\sqrt{2-\sqrt{2}})^{-1}\end{aligned} \tag{5.135}$$

$$\begin{aligned}\Phi_{31} = {}& (\sqrt{2-\sqrt{2}})^{-1}\left(-2\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) - 1/2\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right)\sqrt{2}\right) \\ & + 1/2\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) + 1/2\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\sqrt{2}\end{aligned} \tag{5.136}$$

$$\begin{aligned}\Phi_{32} = {}& (\sqrt{2-\sqrt{2}})^{-1}\left(-1/2\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\sqrt{2-\sqrt{2}} + \sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\right) + \\ & + 1/2\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right)(\sqrt{2-\sqrt{2}})^{-1} - 1/4\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\sqrt{2}\end{aligned} \tag{5.137}$$

$$\Phi_{33} = 1/2\cos\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) + 1/2\cos\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) \tag{5.138}$$

$$\Phi_{34} = 1/4\sqrt{2}\left(-\cos\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) + \cos\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\right) \tag{5.139}$$

$$\Phi_{41} = -\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) - 1/2\,\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\sqrt{2}+ \tag{5.140}$$
$$+\,(2\,\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) + \sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right))(\sqrt{2-\sqrt{2}})^{-1}$$

$$\Phi_{42} = \left(-2\,\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) - 1/2\,\sin\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right)\sqrt{2}\right)(\sqrt{2-\sqrt{2}})^{-1}+ \tag{5.141}$$
$$+\,1/2\,\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) + 1/2\,\sqrt{2+\sqrt{2}}\sin\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\sqrt{2}$$

$$\Phi_{43} = 1/2\,\sqrt{2}\left(-\cos\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) + \cos\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right)\right) \tag{5.142}$$

$$\Phi_{44} = 1/2\,\cos\left(\sqrt{2-\sqrt{2}}\,(t-t_0)\right) + 1/2\,\cos\left(\sqrt{2+\sqrt{2}}\,(t-t_0)\right) \tag{5.143}$$

Since both the sine and cosine function are bounded, every single component of the state transition matrix is bounded. There exists an uniform upper bound for all entries of the state transition matrix, say C:

$$|\Phi_{ij}| \leq C < \infty,\ i,j \in \{1,\ldots,4\}. \tag{5.144}$$

We then have

$$\|\Phi_{A_{(0,0,0,0)}}(t,t_0)\|_2 = \max_{x\in\mathbb{R}^4,\|x\|_2=1} \|\Phi_{A_{(0,0,0,0)}}(t,t_0)x\|_2 \tag{5.145}$$
$$\leq \max_{x\in\mathbb{R}^4,\|x\|_2=1} \sqrt{4(Cx_1+Cx_2+Cx_3+Cx_4)^2} \tag{5.146}$$
$$\leq 2C \max_{x\in\mathbb{R}^4,\|x\|_2=1} \|x\|_2 = 2C < \infty \tag{5.147}$$

showing that the origin is uniformly stable with respect to the system

$$\dot{x}(t) = A_{(0,0,0,0)}x(t). \tag{5.148}$$

The origin is not asymptotically stable, since Φ_{11} is nonvanishing for $t \to \infty$. We have

$$\|\Phi_{A(0,0,0,0)}(t,t_0)\|_2 = \max_{x\in\mathbb{R}^4,\|x\|_2=1} \|\Phi_{A(0,0,0,0)}(t,t_0)x\|_2 \tag{5.149}$$
$$\geq \|\Phi_{A(0,0,0,0)}(t,t_0)(1,0,0,0)^T\|_2 \geq |\Phi_{11}(t,t_0)| \tag{5.150}$$

which implies that $\|\Phi_{A(0,0,0,0)}(t,t_0)\|_2 \not\to 0$ as $t \to \infty$.

For $A_{(\pi,0,0,0)}$ the state transition matrix $\Phi_{A_{(\pi,0,0,0)}}$ has entries

$$\Phi_{11} = 1/2\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) + 1/4\,e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + 1/4\,e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - 1/2\sqrt{2}\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) + \\ + 1/4\sqrt{2}e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} + 1/4\sqrt{2}e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} \tag{5.151}$$

$$\Phi_{12} = 1/8\sqrt{2}\left(2\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0}\right) \tag{5.152}$$

$$\Phi_{13} = 1/8\sqrt[4]{2}\left(2\sqrt{2}\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - \sqrt{2}e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + \sqrt{2}e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - 2\,e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0}\right) + \\ + 1/8\sqrt[4]{2}\left(-4\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) + 2\,e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) \tag{5.153}$$

$$\Phi_{14} = -1/8\sqrt[4]{2}\left(-2\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) \tag{5.154}$$

$$\Phi_{21} = -1/4\sqrt{2}\left(2\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0}\right) \tag{5.155}$$

$$\Phi_{22} = 1/2\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) + 1/4\,e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + 1/4\,e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} + 1/2\sqrt{2}\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - \\ - 1/4\sqrt{2}e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - 1/4\sqrt{2}e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} \tag{5.156}$$

$$\Phi_{23} = 1/4\sqrt[4]{2}\left(-2\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) \tag{5.157}$$

$$\Phi_{23} = 1/8\sqrt[4]{2}\left(2\sqrt{2}\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - \sqrt{2}e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + \sqrt{2}e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} + 2\,e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0}\right) + \\ + 1/8\sqrt[4]{2}\left(4\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - 2\,e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) \tag{5.158}$$

$$\Phi_{31} = 1/4\sqrt[4]{2}\left(2\sqrt{2}\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} - 2\sin\left(\sqrt[4]{2}t + \sqrt[4]{2}t_0\right) + e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) + \\ + 1/4\sqrt[4]{2}\left(-\sqrt{2}e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + \sqrt{2}e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) \tag{5.159}$$

$$\Phi_{32} = -1/8\,2^{3/4}\left(2\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) \tag{5.160}$$

$$\Phi_{33} = 1/2\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) + 1/4\,e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + 1/4\,e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - 1/2\sqrt{2}\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) + \\ + 1/4\sqrt{2}e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} + 1/4\sqrt{2}e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} \tag{5.161}$$

$$\Phi_{34} = 1/8\sqrt{2}\left(2\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0}\right) \tag{5.162}$$

$$\Phi_{41} = 1/4\,2^{3/4}\left(2\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) \tag{5.163}$$

$$\Phi_{42} = -1/4\sqrt[4]{2}\left(2\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) + e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} - e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} + 2\sqrt{2}\sin\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right)\right) - \\ - 1/4\sqrt[4]{2}\left(-\sqrt{2}e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + \sqrt{2}e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0}\right) \tag{5.164}$$

$$\Phi_{43} = -1/4\sqrt{2}\left(2\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0}\right) \tag{5.165}$$

$$\Phi_{44} = 1/2\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) + 1/4\,e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} + 1/4\,e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} + 1/2\sqrt{2}\cos\left(\sqrt[4]{2}t - \sqrt[4]{2}t_0\right) - \\ - 1/4\sqrt{2}e^{\sqrt[4]{2}t-\sqrt[4]{2}t_0} - 1/4\sqrt{2}e^{-\sqrt[4]{2}t+\sqrt[4]{2}t_0} \tag{5.166}$$

Because of

$$\|\Phi_{A(\pi,0,0,0)}(t,t_0)\|_2 = \max_{x\in\mathbb{R}^4,\|x\|_2=1} \|\Phi_{A(\pi,0,0,0)}(t,t_0)x\|_2 \tag{5.167}$$

$$\geq \|\Phi_{A(\pi,0,0,0)}(t,t_0)(1,0,0,0)^T\|_2 \tag{5.168}$$

Now since Φ_{11} is unbounded in t we have $\sup_{t\geq t_0} \|\Phi_{A(\pi,0,0,0)}\|_2 \not< \infty$ and therefore the origin is an unstable equilibrium point with respect to the system

$$\dot{x}(t) = A_{(\pi,0,0,0)}x(t). \tag{5.169}$$

In an analogous way one can show that the origin is unstable with respect to the systems

$$\dot{x}(t) = A_{(0,\pi,0,0)}x(t) \tag{5.170}$$

and

$$\dot{x}(t) = A_{(\pi,\pi,0,0)}x(t). \tag{5.171}$$

5.2.4 Stabilization

We regard the control problem

$$\dot{x}(t) = A_{(\pi,0,0,0)}x(t) + Bu(t) \tag{5.172}$$

with $B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ and the usual assumptions on u.

Method of Bass

First we will apply the method of Bass (4.2.2) to find a suitable (feedback) control $\tilde{u}$ such that the zero-solution is stable with respect to the system

$$\dot{x}(t) = A_{(\pi,0,0,0)}x(t) + B\tilde{u}(t). \tag{5.173}$$

First we need to choose a $\lambda \in \mathbb{R}$ such that (4.94) holds:

$$\lambda > \max_i \sum_{j=1}^{4} |a_{ij}| = 4$$

proposing that $\lambda = 5$ is a suitable choice. Now solving equation

$$(A_{(\pi,0,0,0)} + \lambda I)P + P(A_{(\pi,0,0,0)} + \lambda I)^T = BB^T \tag{5.174}$$

with $\lambda = 5$ and the assumption $P = P^T$ we obtain (numerically)

$$P = \begin{pmatrix} 0.0022 & 0.0001 & -0.0109 & -0.0006 \\ 0.0001 & 0.0019 & 0.0000 & -0.0093 \\ -0.0109 & 0.0000 & 0.1044 & 0.0014 \\ -0.0006 & -0.0093 & 0.0014 & 0.0965 \end{pmatrix} \tag{5.175}$$

and according to the theory of Bass (4.2.2) $\tilde{u}(t) = -B^T P^{-1} x(t)$ is a stabilizing control law and the controlled linear system with system matrix $(A_{(\pi,0,0,0)} + \lambda I - BB^T P^{-1})$ has an exponentially stable zero-solution and with system matrix $A_{(\pi,0,0,0)} - BB^T P^{-1}$ the zero-solution is stable with decay rate at least $-\lambda$ which is -5 here.
A numerical verification shows that all eigenvalues of $(A_{(\pi,0,0,0)} + \lambda I - BB^T P^{-1})$ have real part smaller than -3 and all eigenvalues of $A_{(\pi,0,0,0)} - BB^T P^{-1}$ have real part smaller than -8. As for the the simple pendulum the (exponential) decay rate obtained by applying the method of Bass is better than expected from theory (which would be -5 here).

Pole-shifting

Another way of designing a stabilizing feedback control $\tilde{u}(t) = Fx(t)$ is the pole-shifting method. Instead of giving an upper bound for the eigenvalues' real parts as in the method of Bass we will predefine the eigenvalues in this method. We have seen now in two examples that a system stabilized with the method of Bass may decay much more faster than predicted by the theory. The method of pole-shifting avoids this problem:
We are looking for a 2×4 matrix F such that all eigenvalues of $A_{(\pi,0,0,0)} + BF$ have negative real part. For example we could try to choose F such that for the characteristic polynomial we obtain

$$\rho_{A_{(\pi,0,0,0)}+BK}(\lambda) \stackrel{!}{=} (\lambda + 1)^4, \tag{5.176}$$

with -1 as an eigenvalue of multiplicity 4. Together with $F = \begin{pmatrix} f_{11} & f_{12} & f_{13} & f_{14} \\ f_{21} & f_{22} & f_{23} & f_{24} \end{pmatrix}$ equation (5.176) reads as

$$\begin{aligned} \lambda^4 + (-f_{24} - f_{13})\,\lambda^3 + (-f_{22} - f_{23}f_{14} + f_{24}f_{13} - f_{11})\,\lambda^2 + (-2\,f_{14} - f_{14}f_{21} + f_{23} - f_{23}f_{12} - 2\,f_{13} + {} \\ f_{13}f_{22} + 2\,f_{24} + f_{24}f_{11})\,\lambda - 2 - 2\,f_{12} + f_{21} - f_{21}f_{12} - 2\,f_{11} + 2\,f_{22} + f_{22}f_{11} \\ \stackrel{!}{=} \lambda^4 + 4\,\lambda^3 + 6\,\lambda^2 + 4\,\lambda + 1 \end{aligned} \tag{5.177}$$

Equating coefficients gives 4 equations for 8 unknowns. So the problem of pole placement via state feedback is underdetermined. In [Kautsky et al., 1985] the extra degrees of freedom are used to minimize the sensitivities of the closed-loop poles to perturbations in $A_{(\pi,0,0,0)}$ and K. The only restriction is that the multiplicities of the eigenvalues of $A_{(\pi,0,0,0)} + BF$ are at most the rank of B.
Since B has rank 2 we want $A_{(\pi,0,0,0)} + BF$ to have eigenvalues -1 and -2 with multiplicity 2 each. The algorithm proposed in [Kautsky et al., 1985] then leads to

$$F = \begin{pmatrix} -4 & 1 & -3 & 0 \\ -2 & 0 & 0 & -3 \end{pmatrix} \tag{5.178}$$

and the feedback control is then given by

$$u(t) = Fx(t) = \begin{pmatrix} -4 & 1 & -3 & 0 \\ -2 & 0 & 0 & -3 \end{pmatrix} x(t) \tag{5.179}$$

which makes the zero-solution stable with respect to (5.173).

5.2.5 Swing-up and balancing with a single control law

We will distinguish between "position" and "configuration" when talking about the state of the system. Figuratively speaking, "position" is a snapshot of the double pendulum system, it is a static description of the positions of the pendulum links. With "configuration" we want to describe the exact state of the system including the velocities of the pendulum links. A position at time t can be described by the vector $(\varphi_1(t), \varphi_2(t))^T$ and the corresponding configuration is given by the state vector $(\varphi_1(t), \varphi_2(t), \omega_1(t), \omega_2(t))^T$. The control law we use swings up the double pendulum from different starting configurations *and* balances the position where both links of the double pendulum point upright. We will refer to this position as "upup-position" and in general "direction1direction2-position" will denote the position where the first link points in direction1 and the second link points in direction2.

The mathematical model of consideration is obtained by adding a control term to the nonlinear model derived in the previous section:

$$\frac{d}{dt}\begin{pmatrix}\varphi_1(t)\\ \varphi_2(t)\\ \omega_1(t)\\ \omega_2(t)\end{pmatrix} = \begin{pmatrix}\omega_1\\ \omega_2\\ \frac{-3\sin\varphi_1-\sin(\varphi_1-2\varphi_2)-2\sin(\varphi_1-\varphi_2)(\omega_2^2+\omega_1^2\cos(\varphi_1-\varphi_2))}{3-\cos(2\varphi_1-2\varphi_2)}\\ \frac{2\sin(\varphi_1-\varphi_2)(2\omega_1^2+2\cos\varphi_1+\omega_2^2\cos(\varphi_1-\varphi_2))}{3-\cos(2\varphi_1-2\varphi_2)}\end{pmatrix} + \underbrace{\begin{pmatrix}0 & 0\\ 0 & 0\\ 1 & 0\\ 0 & 1\end{pmatrix}}_{=:B} u(t) =: f(z,u). \tag{5.180}$$

Here u is a vector-valued function with two components and $z(t) = (\varphi_1(t), \varphi_2(t), \omega_1(t), \omega_2(t))^T$. From the physical point of view it makes sense to add the input to the variables ω_1 and ω_2, which represent the velocities of the links. It does not make much sense to add the control input to the variables φ_1 and φ_2 which represent the positions of the links. A simple reason is for example that φ_1 and φ_2 do not contain information about whether the corresponding pendulum link is swinging up or down whereas the variables ω_1 and ω_2 contain this information in form of a positive or negative sign. We will discuss the choice of the matrix B in (5.2.6). We will also discuss in (5.2.6) whether it is necessary for B to have two columns (fully actuated) or whether it would suffice to use a suitable B with only one column (underactuated case). We will use the modified control law presented in (4.2.6):
We are interested in swinging up the double pendulum and balancing it at its upright position $(\pi, \pi, 0, 0)$. We will use the modified control law (4.2.6) to perform this task. In a first attempt we just use the equilibrium point $(\pi, \pi, 0, 0)^T$ as target "trajectory". In this case our control law becomes very simple since linearizing the uncontrolled system (5.180) about the equilibrium point $(\pi, \pi, 0, 0)^T$ gives $A := \frac{\partial f}{\partial z}(z,u)\Big|_{z^T=(\pi,\pi,0,0)}$ as system matrix of the linearization.

We will discuss these results before in a second attempt we try to use the dynamics of the uncontrolled double pendulum and compare the results with the results of the first attempt.

5.2.6 Simulation results and discussion

Before discussing the results we will present the simulation results obtained by applying the control law presented above. We used different initial data for our computations and ran each simulation with two different values for the parameter α. In particular we chose $\alpha = 0$ and $\alpha = 2$. Each simulation runs for 10 time units and the matrix H_α^{-1} used in our control is updated every 2 time units ($\delta_k = \delta = 2 \,\forall\, k \in \mathbb{N}_0$). The simulation results will be presented in a series of figures each consisting of three subfigures. For the first part we will only show the results for the first 5 time units, since all the interesting dynamics take place within this time interval.

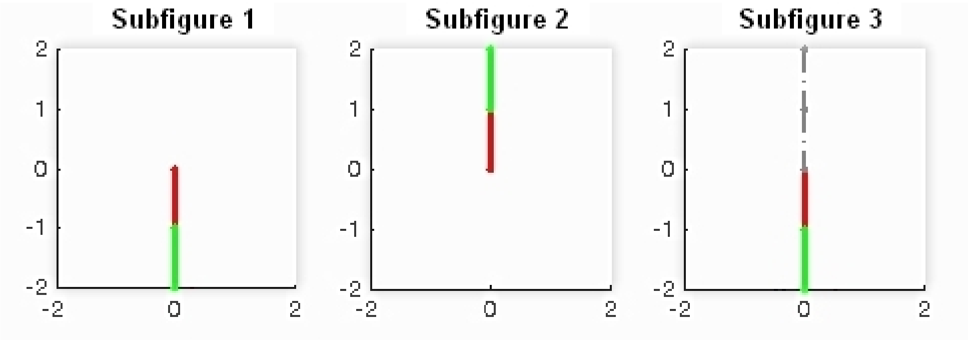


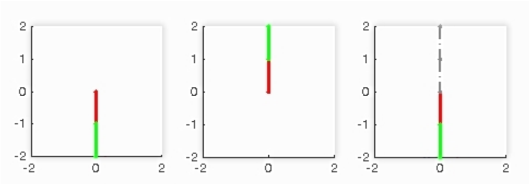
Figure 5.19: setup of the single figures

Figure (5.19) shows one of these figures consisting of three subfigures. It is organized in the following way:

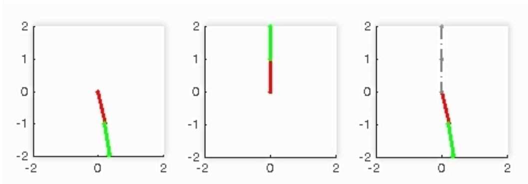
- Subfigure 1 shows the solution of the controlled problem,
- Subfigure 2 shows the solution of the ”target state” or ”trajectory” . In our case the ”target state” will be the upup-position of the double pendulum. As ”target trajectory” or reference trajectory we will choose the solution of uncontrolled nonlinear model of the double pendulum with initial data $(1, 2, 0, 0)^T$.
- Subfigure 3 shows both the solution of the controlled and uncontrolled problem. The target solution is shown as gray broken line ”$- \cdot -$”.

Each figure consisting of these three subfigures shows the state of the system at a certain time. We will call such a figure ”frame”. For better readability and comparability of the results belonging to the same initial data but to different values for α we rescaled the time. For the graphic presentation of the simulation results we will refer to frames rather then speaking of time (1 time unit $\hat{=}$ 200 frames).
Originally the simulation results were obtained by MATLAB and saved as movie. The numbering of the frames presented here is the same as that in the movies obtained by the matlab simulation. The first frame of each series will always show the the position at the beginning of the simulation (frame 0), the last frame of each series will show the position at the end of the simulation.

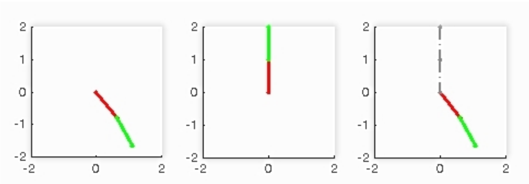
initial data for controlled system: $(0, 0, 0, 0)^T$, simulation time: 5 time units, $\alpha = 0$:



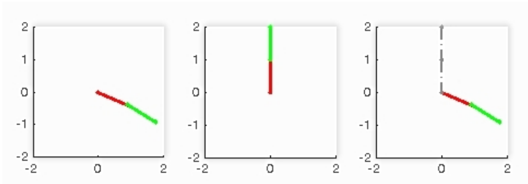
(a) frame 0



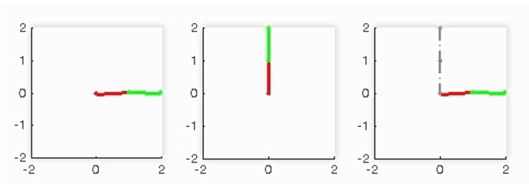
(b) frame 25



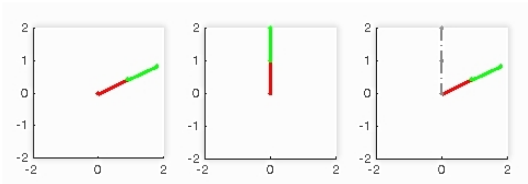
(c) frame 50



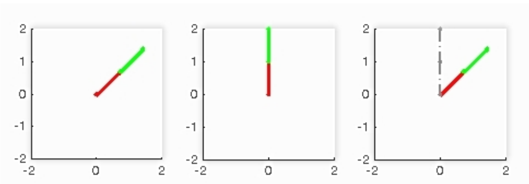
(d) frame 75



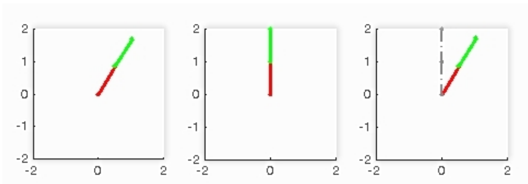
(e) frame 100



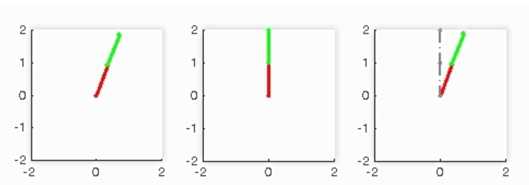
(f) frame 125



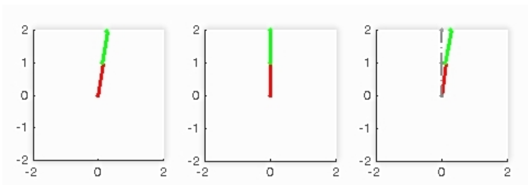
(g) frame 150



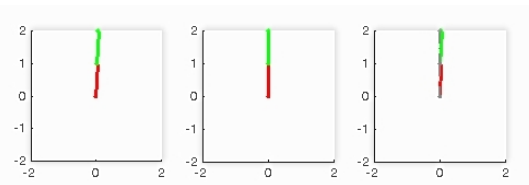
(h) frame 175



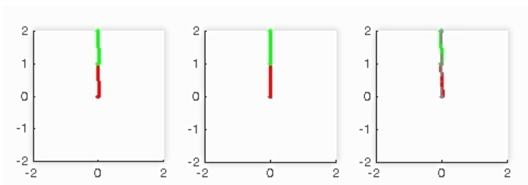
(i) frame 200



(j) frame 250



(k) frame 300



(l) frame 1000

Figure 5.20: Swing-up from downdown-position to upup-position and balancing ($\alpha = 0$)

Initial data $(0, 0, 0, 0)^T$ and $\alpha = 0$

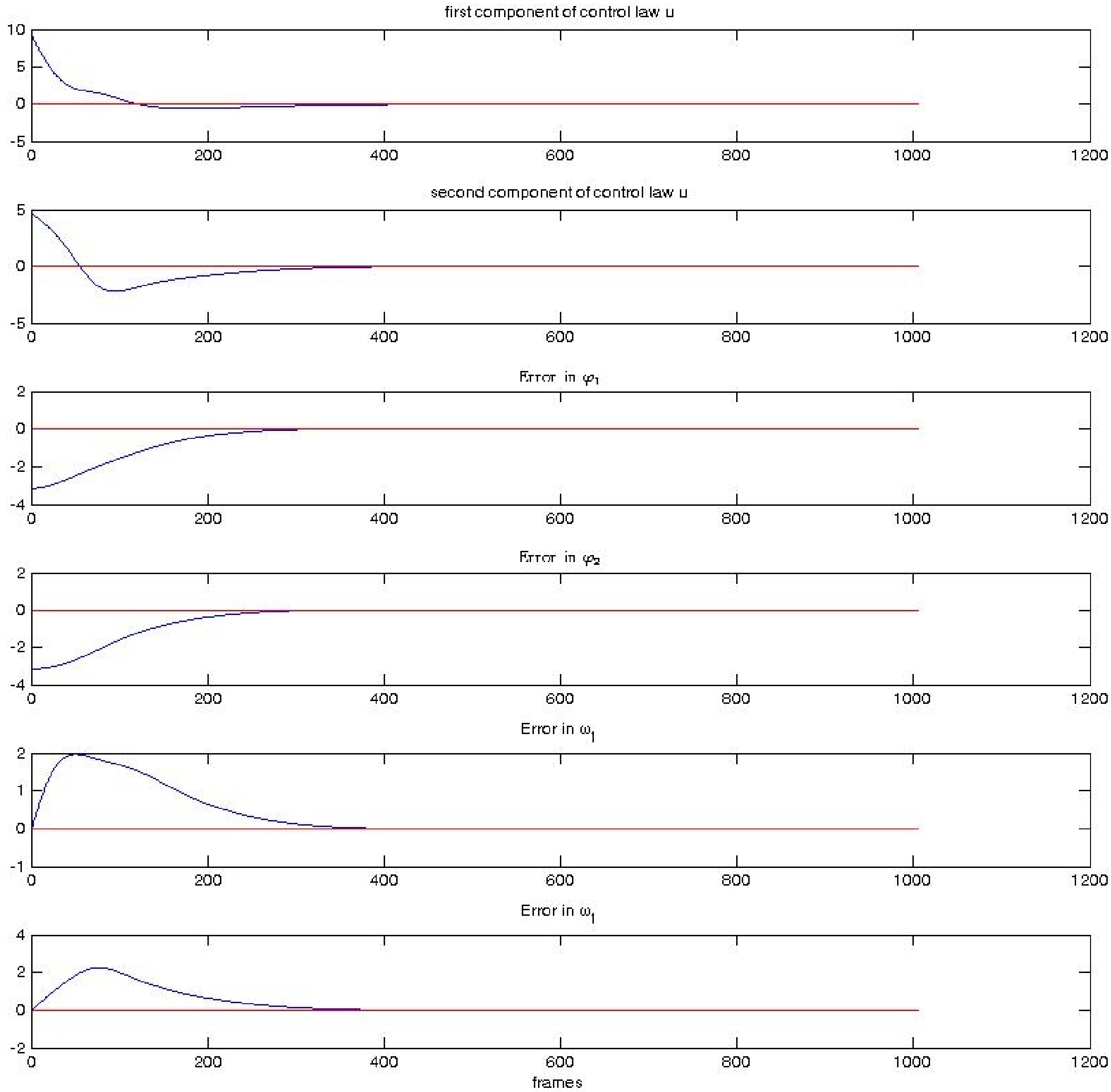
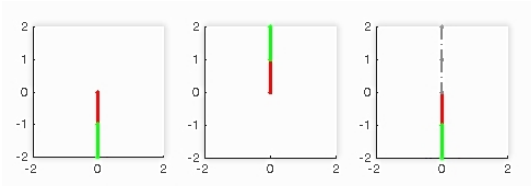
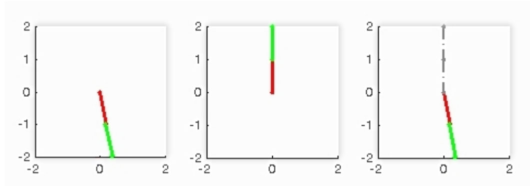


Figure 5.21: Control law and deviation in the single components

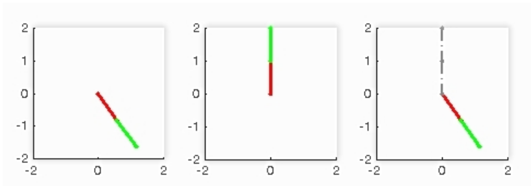
initial data for controlled system: $(0,0,0,0)^T$, simulation time: 5 time units, $\alpha = 2$:



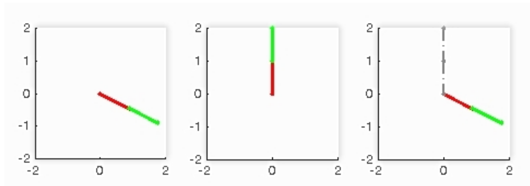
(a) frame 0



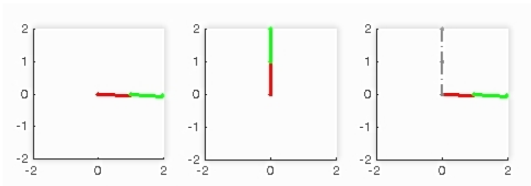
(b) frame 5



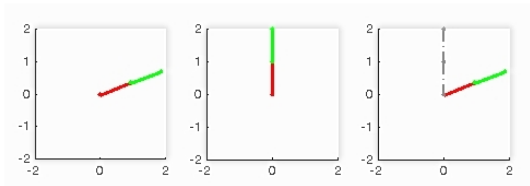
(c) frame 10



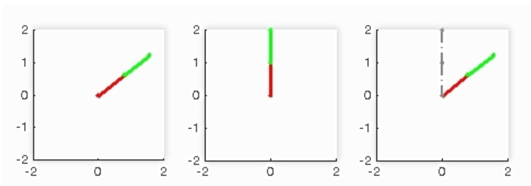
(d) frame 15



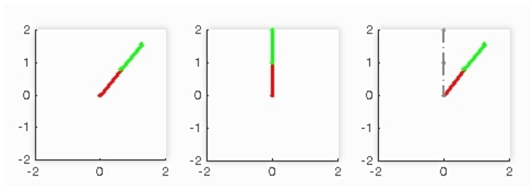
(e) frame 20



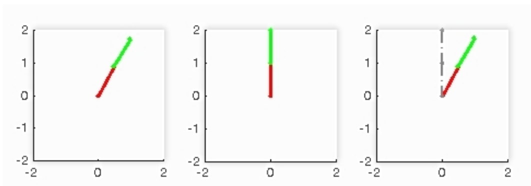
(f) frame 25



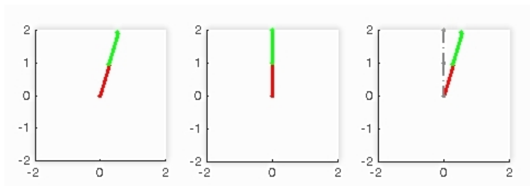
(g) frame 30



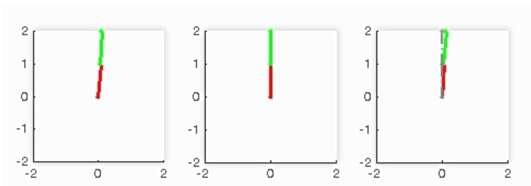
(h) frame 35



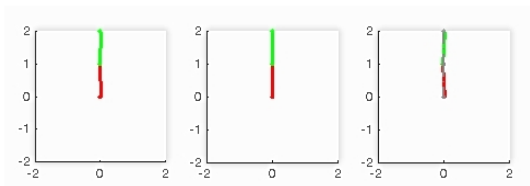
(i) frame 40



(j) frame 50



(k) frame 75



(l) frame 1000

Figure 5.22: Swing-up from downdown-position to upup-position and balancing ($\alpha = 2$)

Initial data $(0,0,0,0)^T$ and $\alpha = 2$

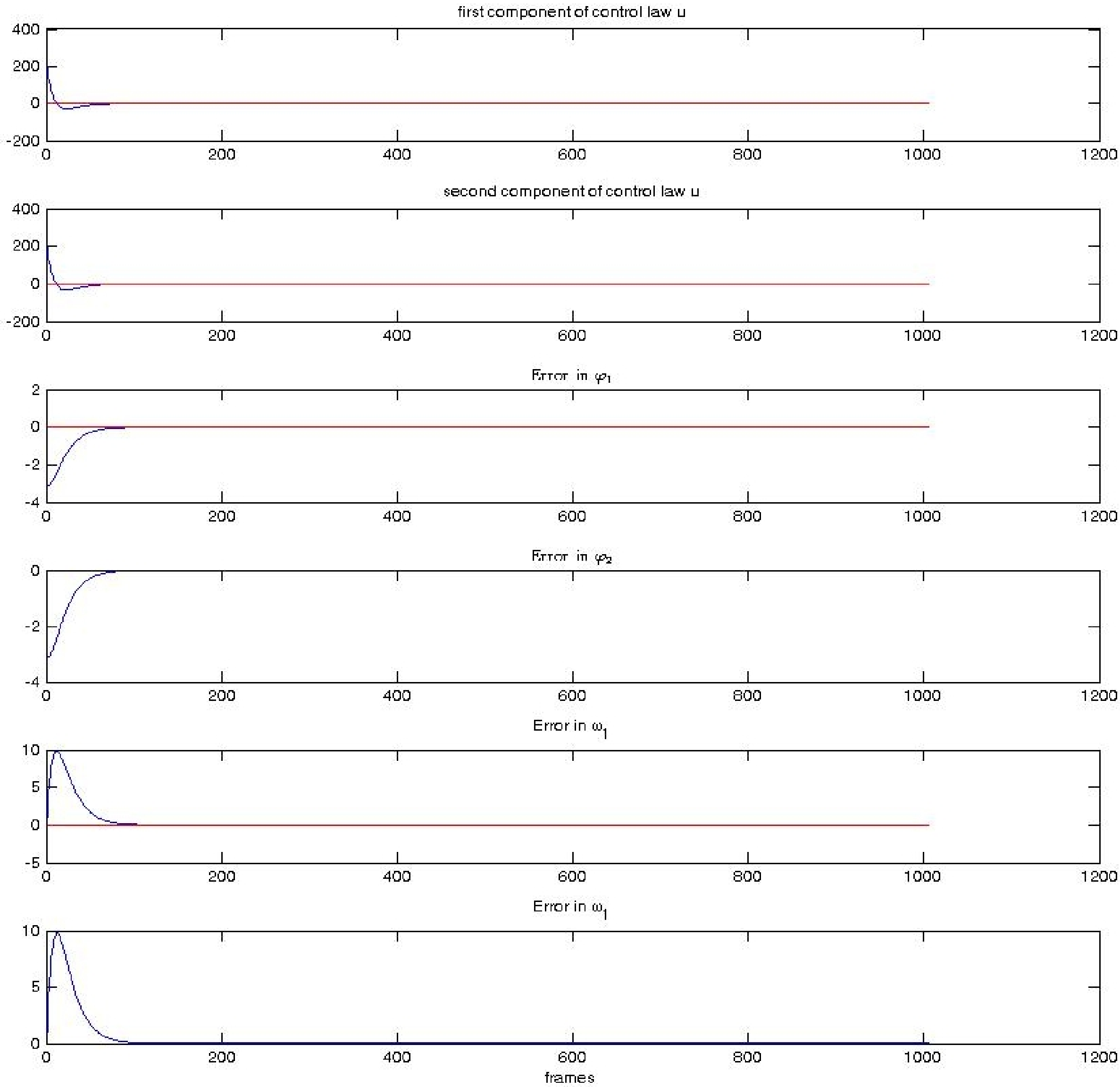
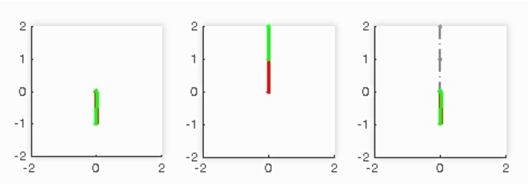
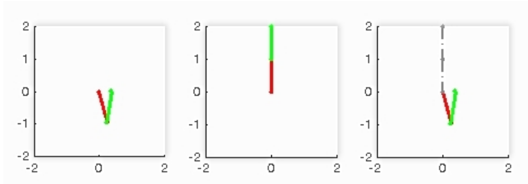


Figure 5.23: Control law and deviation in the single components

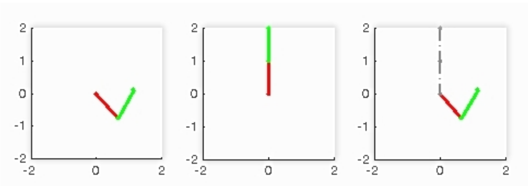
initial data for controlled system: $(0, \pi, 0, 0)^T$, simulation time: 5 time units, $\alpha = 0$:



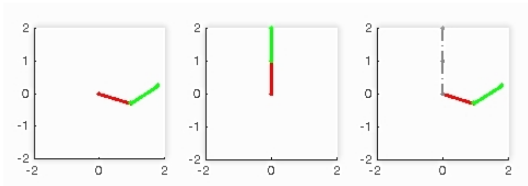
(a) frame 0



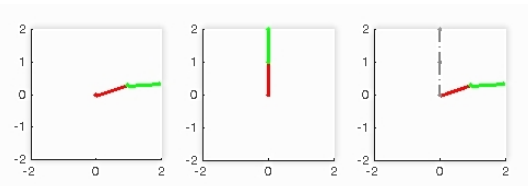
(b) frame 20



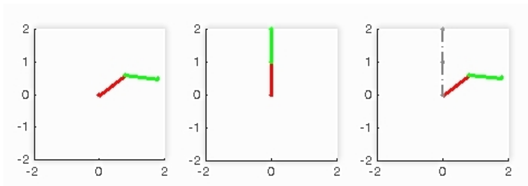
(c) frame 40



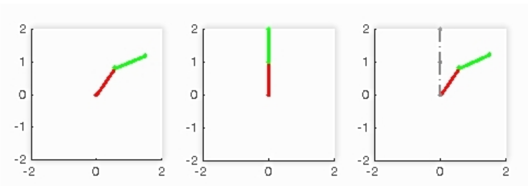
(d) frame 60



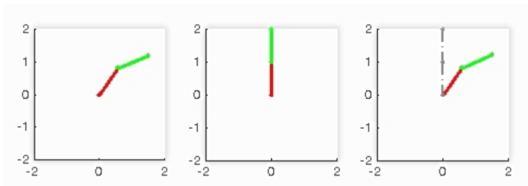
(e) frame 80



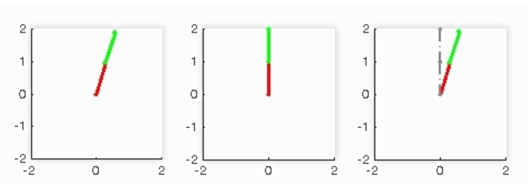
(f) frame 100



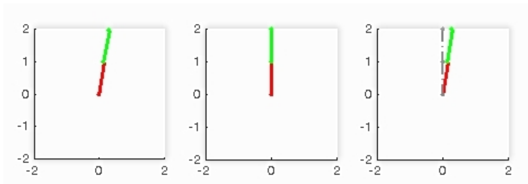
(g) frame 150



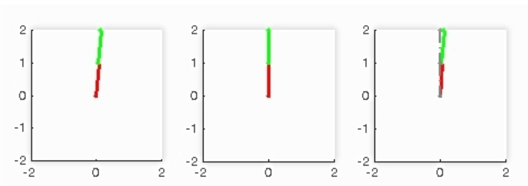
(h) frame 200



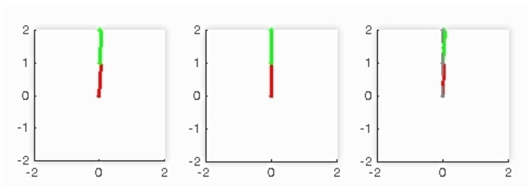
(i) frame 250



(j) frame 300



(k) frame 350



(l) frame 1000

Figure 5.24: Swing-up from downup-position to upup-position and balancing ($\alpha = 0$)

Initial data $(0, \pi, 0, 0)^T$ and $\alpha = 0$

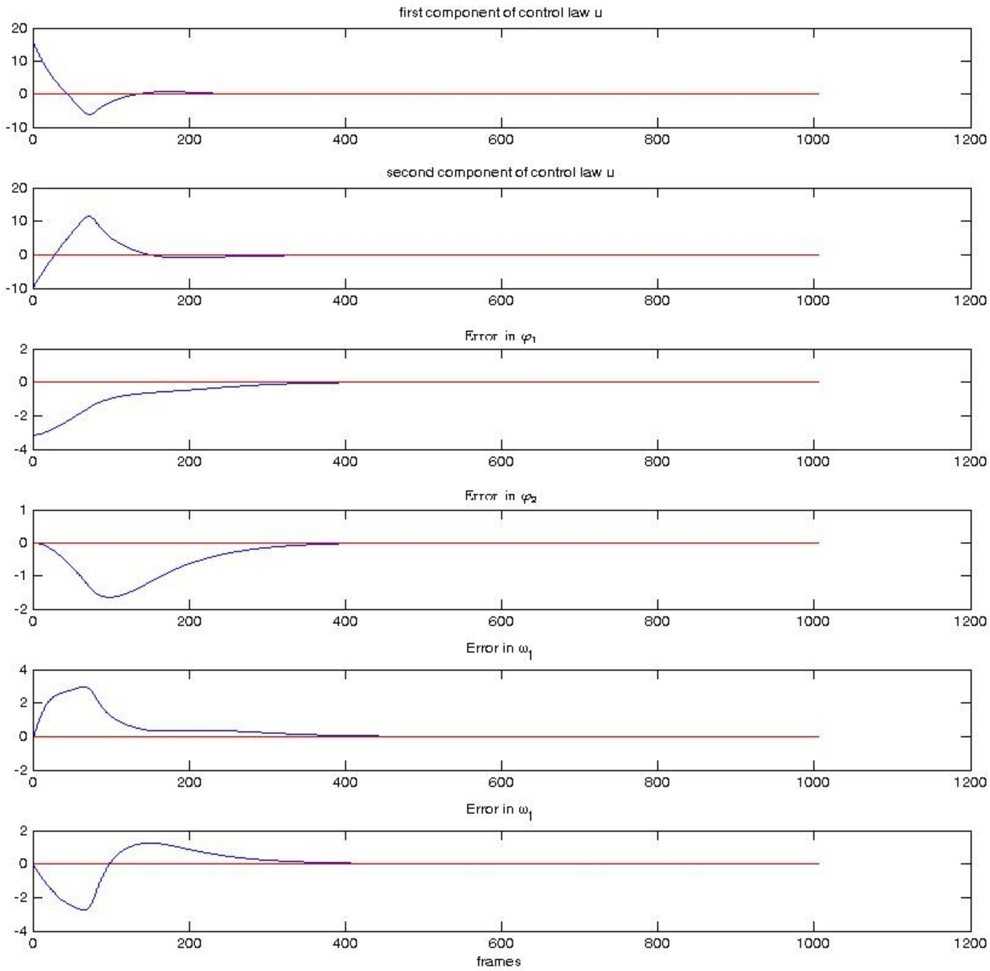
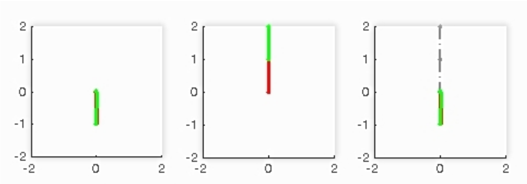
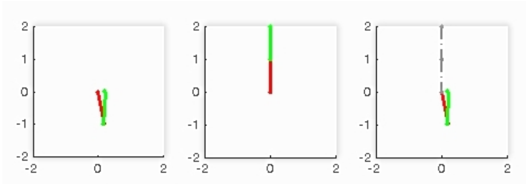


Figure 5.25: Control law and deviation in the single components

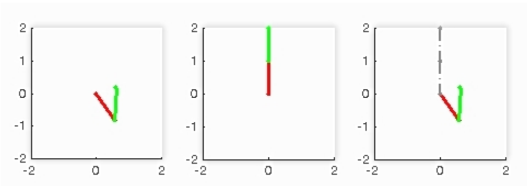
initial data for controlled system: $(0, \pi, 0, 0)^T$, simulation time: 5 time units, $\alpha = 2$:



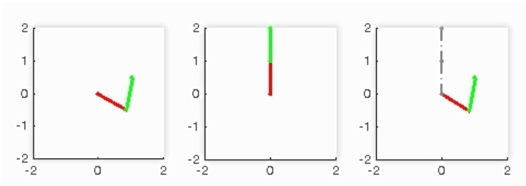
(a) frame 0



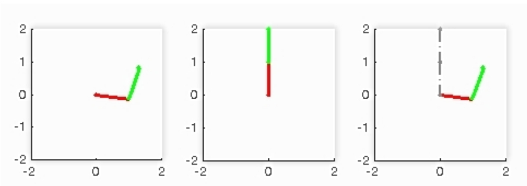
(b) frame 5



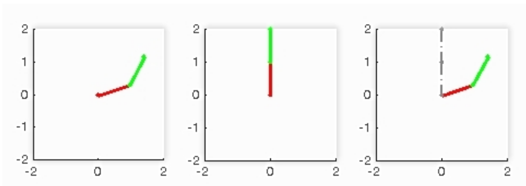
(c) frame 10



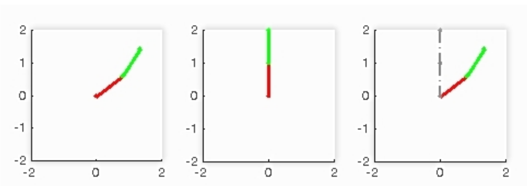
(d) frame 15



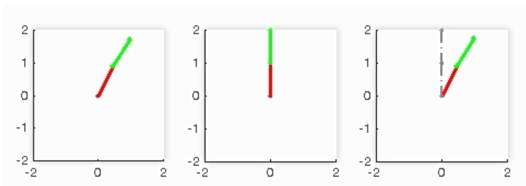
(e) frame 20



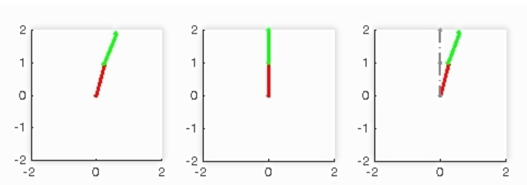
(f) frame 25



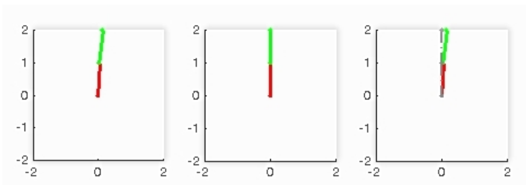
(g) frame 30



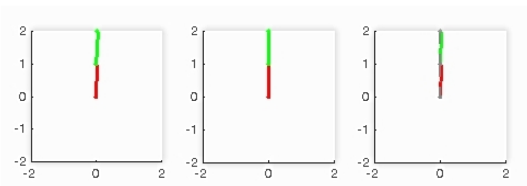
(h) frame 40



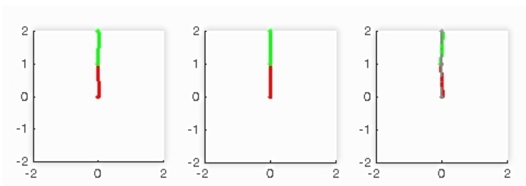
(i) frame 50



(j) frame 75



(k) frame 100



(l) frame 1000

Figure 5.26: Swing-up from downup-position to upup-position and balancing ($\alpha = 2$)

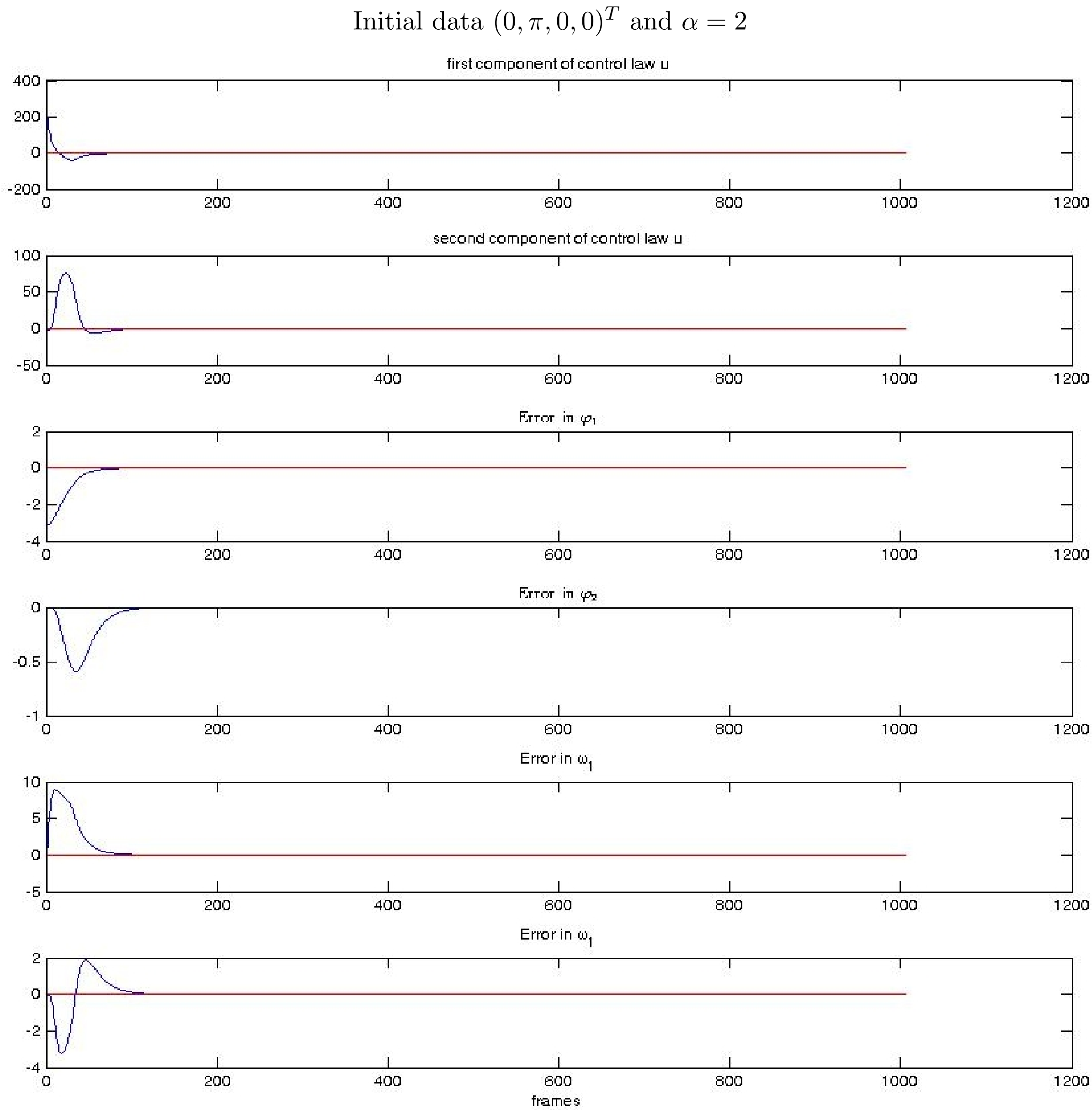


Figure 5.27: Control law and deviation in the single components

For our simulation we chose two different initial configurations. The stable equilibrium point, where both pendula point downward (downdown-position) and the unstable equilibrium position where the first pendulum points downward and the second one upward (downup-position). For these examples the initial velocities of the pendulum links were assumed to be zero.
For each starting configuration we ran the simulation with two different values for the parameter α. We chose 0 and 2 for the simulations.
We succeeded in swinging up the double pendulum and balance it in the upup-position with a *single* control law. As expected a higher value for the parameter α leads to a faster "convergence" to the upup-position. Starting from the downdown-position and choosing $\alpha = 2$ the double pendulum can be stabilized at the upup-position about 4.5 times faster than with $\alpha = 0$. The price for the faster convergence to the upper equilibrium point is a significant increase in the magnitude of the control input u. Comparing figures (5.21) and (5.23) we see that for $\alpha = 2$ the maximum value of the control function u is roughly about 20 times higher than the maximum value attained for $\alpha = 0$. In fact in this example the parameter α can only be used in a very narrow range to get better results with respect to the time needed to stabilize the upup-position, since exorbitantly high control inputs would be needed. The following table shows the largest value of $\|u\|$ for different values for α. With $\text{frame}_{\max}$ we give roughly the time [in frames] which is needed until the upper equilibrium state is reached in good approximation:

α	0	2	4	8	10
$\max_t \|u(t)\|$	9	222	950	4473	7604
$\text{frame}_{\max}$	300	75	40	20	15

The control law we propose is based on control laws designed for linear autonomous systems (see [Kleinmann, 1970] and (4.12) and linear time-varying systems (see [Cheng, 1979] and (4.15)). We recall that the nonlinear system (5.180) is locally controllable around the equilibrium point $(\pi, \pi, 0, 0)$, since its linearization around this equilibrium point is completely controllable (cf. theorem (3.20)), which was shown via the Kalman controllability criteria (3.11): The system matrix of the linearization at $(z^0)^T = (\pi, \pi, 0, 0)^T$ is given by

$$A := \frac{\partial f}{\partial z}(z, u)\bigg|_{z=(\pi,\pi,0,0)^T} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & -1 & 0 & 0 \\ -2 & 2 & 0 & 0 \end{pmatrix} \tag{5.181}$$

and the matrix

$$[B|AB|A^2B|A^3B] = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -2 & 2 \\ 1 & 0 & 0 & 0 & 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -2 & 2 & 0 & 0 \end{bmatrix} \tag{5.182}$$

has rank 4 and therefore the linear control system

$$\frac{d}{dt}z(t) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & -1 & 0 & 0 \\ -2 & 2 & 0 & 0 \end{pmatrix} z(t) + \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} u(t), \quad z \in \mathbb{R}^4,\ t \in \mathbb{R} \tag{5.183}$$

is completely controllable due to Kalman and the nonlinear system (5.180) is locally controllable in an environment around $(\pi, \pi, 0, 0)^T$.

Local stability of the controlled system

Once the double pendulum is close to the equilibrium point z^0 where both links point upright, the proposed control law (4.130) guarantees convergence to this equilibrium point:
Since z^0 is an equilibrium point of the uncontrolled nonlinear system $f(z,0)$ we obtain by Taylor series expansion (cf. proof of theorem (4.16))

$$\dot{z}(t) = f(z,u) - f(z^0,0) \tag{5.184}$$
$$= A(z(t)-z^0) + Bu(t) + o(z(t)-z^0, u(t), t) \tag{5.185}$$

Inserting our control law (4.130) and using the fact that in our special case $\tilde{H}_\alpha(t_k,\delta_k) = H_\alpha$ for all $k \in \mathbb{N}_0$, leads to

$$\dot{z}(t) = (A - BB^T\tilde{H}_\alpha^{-1})(z(t)-z^0) + o(z(t)-z^0(t), t) \tag{5.186}$$

where our regularity assumptions guarantee that

$$\lim_{\|z(t)-z^0\|\to 0} \sup_{t\geq t_0} \frac{\|o(z(t)-z^0(t),t)\|}{\|z(t)-z^0\|} = 0. \tag{5.187}$$

All eigenvalues of $A - BB^T\tilde{H}_\alpha^{-1}$ have real part < -1 such that with the help of theorem (4.9) this establishes (global) stability of the system

$$\dot{z}(t) = (A - BB^T\tilde{H}_\alpha^{-1})(z(t)-z^0) \tag{5.188}$$

and therefore local stability around z^0 (and $u^0 = 0$) of the system

$$\dot{z} = f(z,u). \tag{5.189}$$

A numerical Lyapunov function candidate for global convergence

We will propose a Lyapunov function candidate and give some numerical results. Unfortunately I am not able to proof mathematically that the proposed function is actually a Lyapunov function. Instead we will discuss some properties of the suggested function in relation with the divergence of f.

For the divergence of the right hand side f of the controlled double pendulum system (5.180) we can give the following upper bound for $x \neq (\pi,\pi,0,0)^T$:

$$\begin{aligned}
\operatorname{div} f &= -\operatorname{trace} BB^T H_\alpha^{-1} + 4(\omega_2-\omega_1)\frac{\sin(\varphi_1-\varphi_2)\cos(\varphi_1-\varphi_2)}{3-\cos(2\varphi_1-2\varphi_2)} \\
&= -\operatorname{trace} BB^T H_\alpha^{-1} + 4(\omega_2-\omega_1)\frac{\sin(2\varphi_1-2\varphi_2)}{6-2\cos(2\varphi_1-2\varphi_2)} \\
&\leq -\operatorname{trace} BB^T H_\alpha^{-1} + (\omega_2-\omega_1)
\end{aligned}$$

Since $BB^TH_\alpha^{-1}$ is symmetric and positive definite we have

$$-\operatorname{trace} BB^T H_\alpha^{-1} < 0 \tag{5.190}$$

and therefore $\operatorname{div} f < 0$ if

$$\omega_2 - \omega_1 < \operatorname{trace} BB^T H_\alpha^{-1}. \tag{5.191}$$

Since $H_\alpha^{-1} = H_0^{-1}\mathrm{e}^{4\alpha\xi}$ for some $\xi \in [0,2]$ this can always be guaranteed by choosing a suitable $\alpha > 0$.

Note that although for every initial condition we can choose an appropriate $\alpha > 0$ such that for our control law u and the initial value z_0 we have $\operatorname{div} f(z_0, u) < 0$ it may happen that for later times condition (5.191) is violated.

We propose the following function as Lyapunov function candidate for the equilibrium point $(\pi,\pi,0,0)^T$ with respect to system (5.180)

$$V(z(t)) := \frac{1}{2}\|z_1(t) - \pi\|^2 + \frac{1}{2}\|z_2(t) - \pi\|^2. \tag{5.192}$$

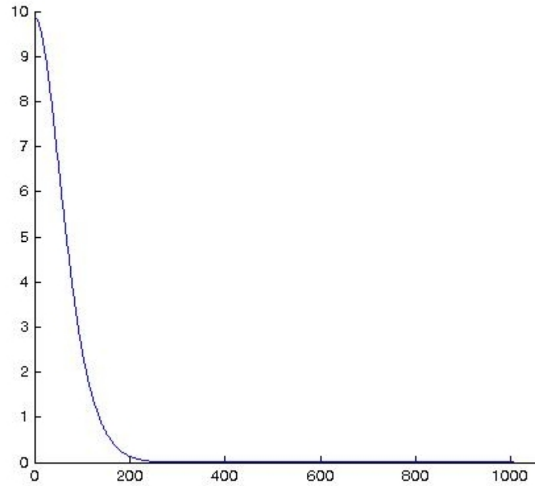
The condition $V(z(t)) > 0$ for $z(t) \in \mathbb{R}^4 \setminus \{(\pi,\pi,0,0)^T\}$ and $V((\pi,\pi,0,0)^T) = 0$ holds. For the derivative we have

$$\dot{V}(z(t)) = \left\langle \begin{pmatrix} z_1(t) - \pi \\ z_2(t) - \pi \end{pmatrix}, \begin{pmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{pmatrix} \right\rangle \tag{5.193}$$

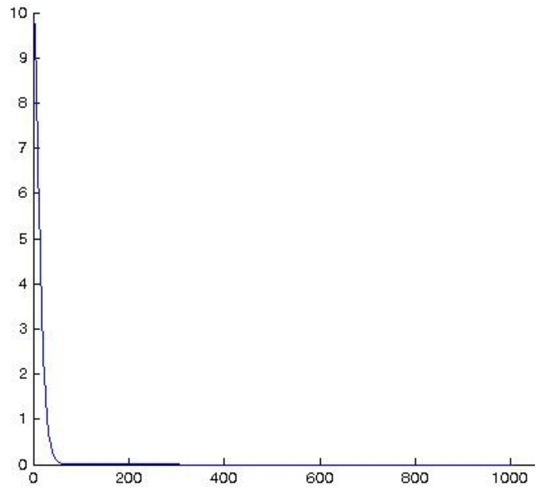
or using the variables $(\varphi_1(t), \varphi_2(t), \omega_1(t), \omega_2(t))^T = z(t)$

$$\dot{V}(\varphi_1(t), \varphi_2(t), \omega_1(t), \omega_2(t)) = \left\langle \begin{pmatrix} \varphi_1(t) - \pi \\ \varphi_2(t) - \pi \end{pmatrix}, \begin{pmatrix} \omega_1(t) \\ \omega_2(t) \end{pmatrix} \right\rangle \tag{5.194}$$

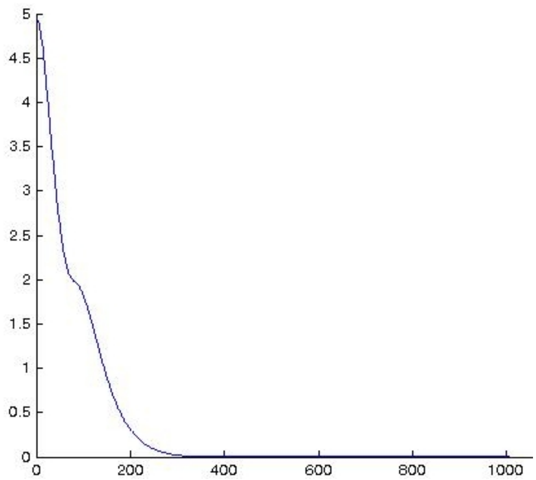
An analytical estimation of this derivative is complicated because for estimating the terms $\omega_1(t)$, $\omega_2(t)$ we would need an analytical representation of H_α^{-1}. For our simulation we obtain the following numerical results for $V(z(t))$:



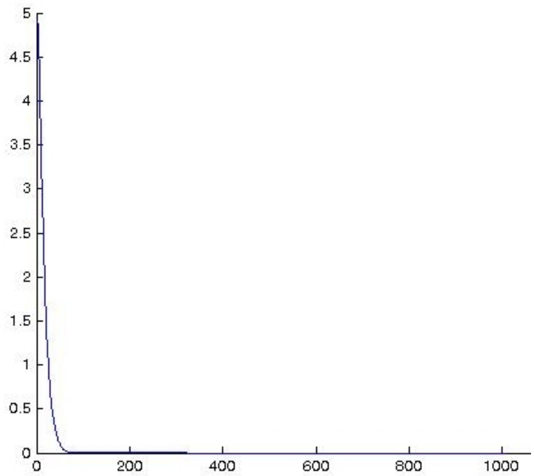
(a) initial condition $z_0 = (0, 0, 0, 0)^T, \ \alpha = 0$



(b) initial condition $z_0 = (0, 0, 0, 0)^T, \ \alpha = 2$



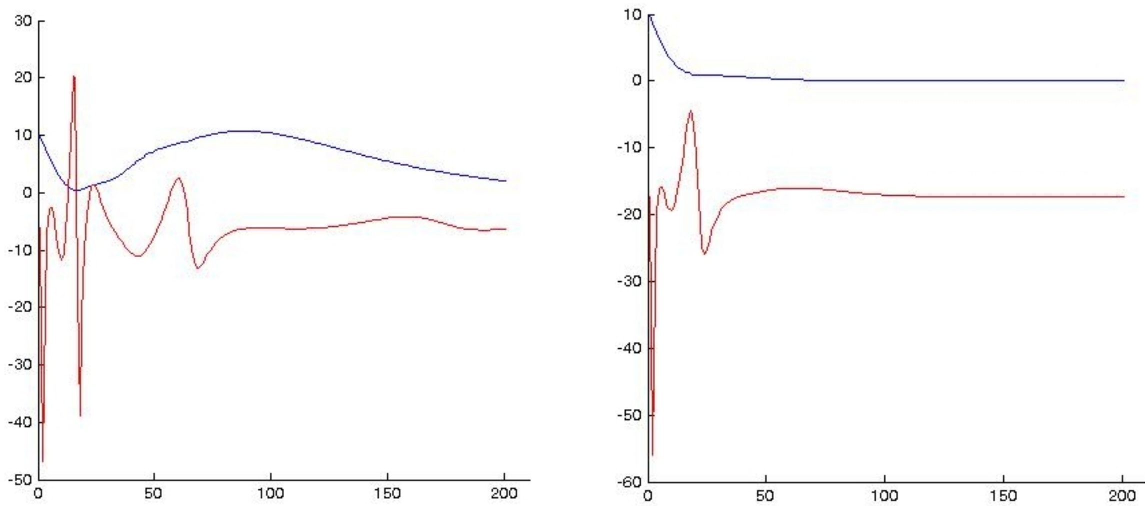
(c) initial condition $z_0 = (0, \pi, 0, 0)^T, \ \alpha = 0$



(d) initial condition $z_0 = (0, \pi, 0, 0)^T, \ \alpha = 2$

Figure 5.28: Lyapunov function candidate

We will give a further example where our Lyapunov function candidate fails to be a decaying function. We chose the initial value $z_0^T = (0, 0, -10, 50)$ with $\alpha = 0$ and $\alpha = 1$. One can observe from the numerical computations that for the case $\alpha = 0$ the condition (5.191) does not hold for all times and that our Lyapunov function candidate fails to be decaying. So our assumption is that $\dot{V}(z(t)) < 0$ if (5.191) holds for all times. For $\alpha = 1$ condition (5.191) is not violated and our Lyapunov function candidate is decaying everywhere:



(a) initial condition $z_0 = (0, 0, -10, 50)^T$, $\alpha = 0$ (b) initial condition $z_0 = (0, 0, .10, 50)^T$, $\alpha = 1$

Figure 5.29: Lyapunov function candidate and divergence of f

Bad condition as an obstruction for the underactuated case

Nonlinear underactuated control is an important field of control theory. There are several reasons why one might be interested in underactuated control. For example if one or even more of the actuators fail one is interested in the question whether the remaining control components are sufficient to control the system. Practical aspects may also play an important role. For example a double pendulum system, where forces to both pendulum links can be applied, is much more complicated to realize, then a double pendulum system, where only a force to the first link is applied. Such an underactuated control for the nonlinear double pendulum system was designed in [Fantoni and Lozano, 2002][pp. 53-72]. It is based on a Lyapunov approach and results in an oscillating control law (see figure 5.30), which brings the double pendulum close to the upup-position. For the balancing part a linear controller was used.

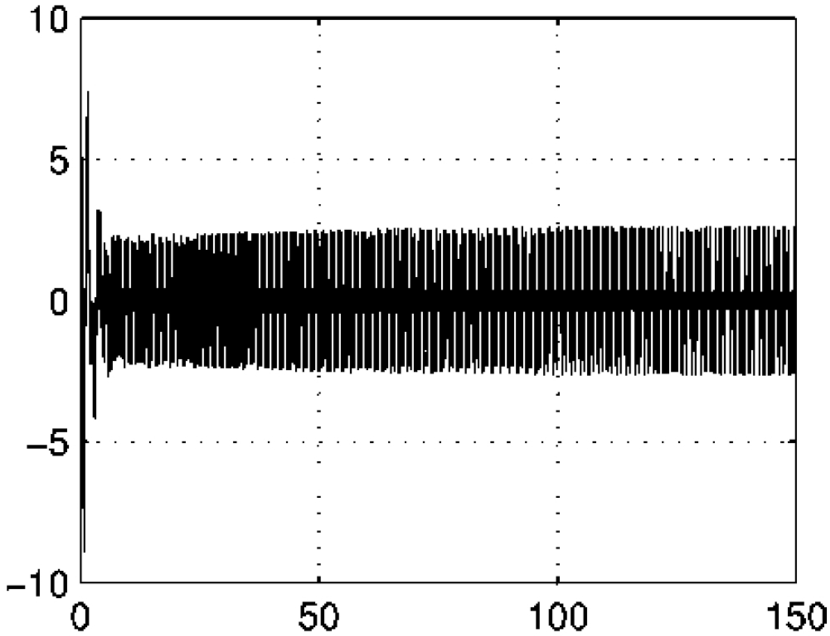


Figure 5.30: Nonlinear underactuated control by Fantoni / Lozano for swinging up the double pendulum. The figure shows the scalar valued control which represents the force applied to the first link. [Fantoni and Lozano, 2002][p.70].

The oscillations are due to the controller design. Fantoni and Lozano wanted to design a controller, which does not need high gains. The highest value for the control is attained for swinging up the first link to its upright position. Then the first link more or less stays in its upright position while the second pendulum link is brought to its upright position by adding only a small portion of energy so that with every "swing" the second link becomes closer to the upright position. When it is finally close enough to the upright position a linear feedback controller is applied for the balancing part.

This example shows that an underactuated control of the nonlinear double pendulum system is possible. On the other hand the control law is designed in two steps, the first brings the pendulum close to the equilibrium where both pendulum links point upright (swinging up part) - the second balances the double pendulum at its upup-position (balancing part). Up to my

knowledge there is no underactuated control for the double pendulum, which can achieve both parts with a single control law.
We wanted to design such a control, which swings the double pendulum up and balances it with only one control law, where no switching would be needed. Unfortunately in the underactuated case the controllability Gramian turns out to be ill conditioned - being a real obstacle for our approach to work:
The linearized system

$$\frac{d}{dt}z(t) = \underbrace{\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & -1 & 0 & 0 \\ -2 & 2 & 0 & 0 \end{pmatrix}}_{A} z(t) + \underbrace{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{B} u(t), \quad z \in \mathbb{R}^4,\ t \in \mathbb{R} \tag{5.195}$$

is completely controllable, since the Kalman matrix

$$[B|AB|A^2B|A^3B] = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & -2 \\ 1 & 0 & 2 & 0 \\ 0 & 0 & -2 & 0 \end{bmatrix} \tag{5.196}$$

has full rank, which means that the nonlinear underactuated double pendulum system

$$\frac{d}{dt}\begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \omega_1(t) \\ \omega_2(t) \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \frac{-3\sin\varphi_1 - \sin(\varphi_1 - 2\varphi_2) - 2\sin(\varphi_1-\varphi_2)(\omega_2^2 + \omega_1^2\cos(\varphi_1-\varphi_2))}{3-\cos(2\varphi_1 - 2\varphi_2)} \\ \frac{2\sin(\varphi_1-\varphi_2)(2\omega_1^2 + 2\cos\varphi_1 + \omega_2^2\cos(\varphi_1-\varphi_2))}{3-\cos(2\varphi_1-2\varphi_2)} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} u(t). \tag{5.197}$$

is locally controllable around its equilibrium point $(\pi, \pi, 0, 0)^T$.
For our algorithm to work it is crucial that the matrix $H_\alpha(t_i, \delta_i)$ is invertible for every $t_i, i = 1, \ldots$. Since we are interested in steering the system to an equilibrium point, this condition reduces to finding a positive δ such that $H_\alpha(0, \delta)$ is invertible. We will show that for $\alpha = 0$ such an $\delta > 0$ exists: Note that the matrix A has eigenvalues $\pm\sqrt{2-\sqrt{2}}$, $\pm\sqrt{2+\sqrt{2}}$. Therefore a regular transformation matrix T exists such that $A = TDT^{-1}$

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 2 & -1 & 0 & 0 \\ -2 & 2 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/2\,\frac{1+\sqrt{2}}{2+\sqrt{2}} & -1/2\,\frac{1+\sqrt{2}}{2+\sqrt{2}} & 1/2\,\frac{1+\sqrt{2}}{2+\sqrt{2}} & -1/2\,\frac{1+\sqrt{2}}{2+\sqrt{2}} \\ -1/4\,\frac{-2+\sqrt{2}}{\sqrt{2-\sqrt{2}}} & 1/4\,\frac{\sqrt{2}}{\sqrt{2-\sqrt{2}}} & -1/2\,\frac{1}{\sqrt{2-\sqrt{2}}(2+\sqrt{2})} & -1/2\,\frac{1+\sqrt{2}}{\sqrt{2-\sqrt{2}}(2+\sqrt{2})} \\ 1/2\,\frac{-1+\sqrt{2}}{\sqrt{2-\sqrt{2}}} & -1/2\,\frac{1}{\sqrt{2-\sqrt{2}}} & -1/2\,\frac{-1+\sqrt{2}}{\sqrt{2-\sqrt{2}}} & 1/2\,\frac{1}{\sqrt{2-\sqrt{2}}} \end{pmatrix}. \tag{5.198}$$

$$\cdot \begin{pmatrix} \sqrt{2-\sqrt{2}}t & 0 & 0 & 0 \\ 0 & \sqrt{2+\sqrt{2}}t & 0 & 0 \\ 0 & 0 & -\sqrt{2-\sqrt{2}}t & 0 \\ 0 & 0 & 0 & -\sqrt{2+\sqrt{2}}t \end{pmatrix}.$$

$$
\cdot \begin{pmatrix}
1 & 1/2\,\frac{2+\sqrt{2}}{1+\sqrt{2}} & -1/2\,\frac{(2+\sqrt{2})(-2+\sqrt{2})}{\sqrt{2-\sqrt{2}}} & -1/2\,\frac{(1+\sqrt{2})(-2+\sqrt{2})}{\sqrt{2-\sqrt{2}}} \\
1/2\,\frac{\sqrt{2}(2+\sqrt{2})}{1+\sqrt{2}} & -1/2\,\frac{2+\sqrt{2}}{1+\sqrt{2}} & -1/2\,\frac{(2+\sqrt{2})(-2+\sqrt{2})}{\sqrt{2-\sqrt{2}}(1+\sqrt{2})} & 1/2\,\frac{-2+\sqrt{2}}{\sqrt{2-\sqrt{2}}} \\
1 & 1/2\,\frac{2+\sqrt{2}}{1+\sqrt{2}} & 1/2\,\frac{(2+\sqrt{2})(-2+\sqrt{2})}{\sqrt{2-\sqrt{2}}} & 1/2\,\frac{(1+\sqrt{2})(-2+\sqrt{2})}{\sqrt{2-\sqrt{2}}} \\
1/2\,\frac{\sqrt{2}(2+\sqrt{2})}{1+\sqrt{2}} & -1/2\,\frac{2+\sqrt{2}}{1+\sqrt{2}} & 1/2\,\frac{(2+\sqrt{2})(-1+\sqrt{2})(-2+\sqrt{2})}{\sqrt{2-\sqrt{2}}} & -1/4\,\frac{\sqrt{2}(2+\sqrt{2})(-2+\sqrt{2})}{\sqrt{2-\sqrt{2}}(1+\sqrt{2})}
\end{pmatrix}
$$

where D is a diagonal matrix. For $\alpha = 0$ we have

$$
\begin{aligned}
H_0(0,t) &= \int_0^t \mathrm{e}^{-A\tau} B B^T \mathrm{e}^{-A^T\tau}\, d\tau && (5.199)\\
&= T \int_0^t \mathrm{e}^{-D\tau} T B B^T T^{-T} \mathrm{e}^{-D\tau}\, d\tau T^T && (5.200)
\end{aligned}
$$

such that $H_0(0,t)$ is invertible if and only if $\int_0^t \mathrm{e}^{-D\tau} T B B^T T^{-T} \mathrm{e}^{-D\tau}\, d\tau$ is invertible. This integral can be explicitly solved with the symbolic toolbox of matlab, the entries of this matrix $T^{-1} H_0(0,t) T^{-T}$ are given by

$$
\begin{aligned}
\left(T^{-1}H_0(0,t)T^{-T}\right)_{11} &= -1/4\sqrt{2-\sqrt{2}}\left(-3-2\sqrt{2}+3\,\mathrm{e}^{-2\sqrt{2-\sqrt{2}}t}+2\,\mathrm{e}^{-2\sqrt{2-\sqrt{2}}t}\sqrt{2}\right)\\
\left(T^{-1}H_0(0,t)T^{-T}\right)_{21} &= -1/4\,\frac{-3-2\sqrt{2}+3\,\mathrm{e}^{-t\left(\sqrt{2+\sqrt{2}}+\sqrt{2-\sqrt{2}}\right)}+2\,\mathrm{e}^{-t\left(\sqrt{2+\sqrt{2}}+\sqrt{2-\sqrt{2}}\right)}\sqrt{2}}{(1+\sqrt{2})\left(2-\sqrt{2}\right)^{-3/2}}\\
\left(T^{-1}H_0(0,t)T^{-T}\right)_{31} &= \left(-1/2+1/4\sqrt{2}\right)\left(2+\sqrt{2}\right)^2 t\\
\left(T^{-1}H_0(0,t)T^{-T}\right)_{41} &= -1/4\sqrt{2-\sqrt{2}}\left(-2-\sqrt{2}+2\,\mathrm{e}^{-t\left(\sqrt{2-\sqrt{2}}-\sqrt{2+\sqrt{2}}\right)}+\mathrm{e}^{-t\left(\sqrt{2-\sqrt{2}}-\sqrt{2+\sqrt{2}}\right)}\sqrt{2}\right)\\
\left(T^{-1}H_0(0,t)T^{-T}\right)_{12} &= -1/4\,\frac{-3-2\sqrt{2}+3\,\mathrm{e}^{-t\left(\sqrt{2+\sqrt{2}}+\sqrt{2-\sqrt{2}}\right)}+2\,\mathrm{e}^{-t\left(\sqrt{2+\sqrt{2}}+\sqrt{2-\sqrt{2}}\right)}\sqrt{2}}{\left(2-\sqrt{2}\right)^{-3/2}(1+\sqrt{2})}\\
\left(T^{-1}H_0(0,t)T^{-T}\right)_{22} &= -1/8\,\frac{\left(2+\sqrt{2}\right)^{5/2}\left(-1+\mathrm{e}^{-2\sqrt{2+\sqrt{2}}t}\right)}{\left(3+2\sqrt{2}\right)\left(1+\sqrt{2}\right)^2}\\
\left(T^{-1}H_0(0,t)T^{-T}\right)_{32} &= 1/4\sqrt{2-\sqrt{2}}\left(-2-\sqrt{2}+2\,\mathrm{e}^{t\left(\sqrt{2-\sqrt{2}}-\sqrt{2+\sqrt{2}}\right)}+\mathrm{e}^{t\left(\sqrt{2-\sqrt{2}}-\sqrt{2+\sqrt{2}}\right)}\sqrt{2}\right)\\
\left(T^{-1}H_0(0,t)T^{-T}\right)_{42} &= 1/4\,\frac{\left(-1+\sqrt{2}\right)\left(2+\sqrt{2}\right)^2\left(-2+\sqrt{2}\right)t}{1+\sqrt{2}}
\end{aligned}
$$

$$\left(T^{-1}H_0(0,t)T^{-T}\right)_{13} = \left(-1/2 + 1/4\sqrt{2}\right)\left(2+\sqrt{2}\right)^2 t$$
$$\left(T^{-1}H_0(0,t)T^{-T}\right)_{23} = 1/4\sqrt{2-\sqrt{2}}\left(-2-\sqrt{2}+2\,\mathrm{e}^{t\left(\sqrt{2-\sqrt{2}}-\sqrt{2+\sqrt{2}}\right)} + \mathrm{e}^{t\left(\sqrt{2-\sqrt{2}}-\sqrt{2+\sqrt{2}}\right)}\sqrt{2}\right)$$
$$\left(T^{-1}H_0(0,t)T^{-T}\right)_{33} = 1/4\sqrt{2-\sqrt{2}}\left(-3-2\sqrt{2}+3\,\mathrm{e}^{2\sqrt{2-\sqrt{2}}t} + 2\,\mathrm{e}^{2\sqrt{2-\sqrt{2}}t}\sqrt{2}\right)$$
$$\left(T^{-1}H_0(0,t)T^{-T}\right)_{43} = 1/4\sqrt{2}\sqrt{2-\sqrt{2}}\left(-1+\mathrm{e}^{t\left(\sqrt{2+\sqrt{2}}+\sqrt{2-\sqrt{2}}\right)}\right)$$
$$\left(T^{-1}H_0(0,t)T^{-T}\right)_{14} = -1/4\sqrt{2-\sqrt{2}}\left(-2-\sqrt{2}+2\,\mathrm{e}^{-t\left(\sqrt{2-\sqrt{2}}-\sqrt{2+\sqrt{2}}\right)} + \mathrm{e}^{-t\left(\sqrt{2-\sqrt{2}}-\sqrt{2+\sqrt{2}}\right)}\sqrt{2}\right)$$
$$\left(T^{-1}H_0(0,t)T^{-T}\right)_{24} = 1/4\,\frac{\left(-1+\sqrt{2}\right)\left(2+\sqrt{2}\right)^2\left(-2+\sqrt{2}\right)t}{1+\sqrt{2}}$$
$$\left(T^{-1}H_0(0,t)T^{-T}\right)_{34} = 1/4\sqrt{2}\sqrt{2-\sqrt{2}}\left(-1+\mathrm{e}^{t\left(\sqrt{2+\sqrt{2}}+\sqrt{2-\sqrt{2}}\right)}\right)$$
$$\left(T^{-1}H_0(0,t)T^{-T}\right)_{44} = -1/4\sqrt{2+\sqrt{2}}\left(3-2\sqrt{2}-3\,\mathrm{e}^{2\sqrt{2+\sqrt{2}}t} + 2\,\mathrm{e}^{2\sqrt{2+\sqrt{2}}t}\sqrt{2}\right)$$

which – due to maple – has rank 4 for every $t > 0$. Unfortunately, this matrix is ill conditioned. The following table gives some values for the condition number $\kappa(H_0(0,t)) = \|H_0^{-1}(0,t)\|_\infty \|H_0(0,t)\|_\infty$ for some values $t > 0$

t	0.5	1	1.5	2
$\kappa(H_0(0,t))$	$3.7 \cdot 10^6$	$2.5 \cdot 10^5$	$1.5 \cdot 10^5$	$2.3 \cdot 10^5$

Since $H_\alpha(0,t) = H_0(0,t) \cdot \mathrm{e}^{4\xi}$ for some $\xi \in [-t, 0]$ we have that all $H_\alpha(0,t)$ for $\alpha \geq 0$ are ill conditioned. Therefore our attempt fails to work in the underactuated case.

Starting nearby a solution which comes close to the upper equilibrium point by its natural dynamics

We regard a trajectory which starts in position $(0,0)^T$ but with velocities different from zero such that the uncontrolled motion of the double pendulum comes at least close to the upper equilibrium configuration. This "natural movement" of the double pendulum is used as target/reference trajectory for the control law. The idea behind this approach is that once we are close to the reference trajectory we can use the dynamics of the uncontrolled double pendulum to swing it up by just stabilizing the solution of the controlled pendulum equation along this solution.

Remark: Since it is not possible to find a solution for the nonlinear double pendulum such that the upper equilibrium point is reached in finite time (it does not exist!), we choose an initial condition, which brings the double pendulum close to the desired equilibrium point. When the solution is close enough, we will take as new reference trajectory the upper equilibrium state. In this way we guarantee that the controlled solution does not only swing-up but balances the pendulum in the upup-position.

For the solution of the uncontrolled pendulum equation we used the initial condition $z(0) = (0,0,-3.02,4.76)^T$. The tracks of the two pendulum bobs are shown in figure (5.31):

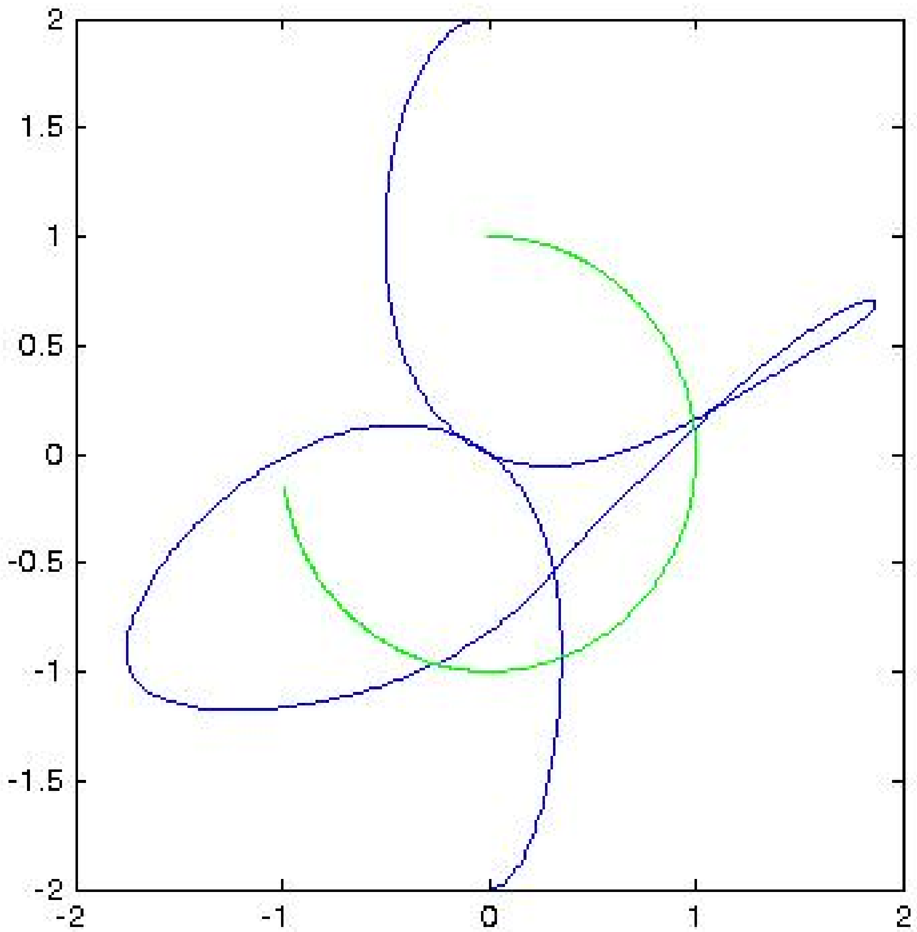
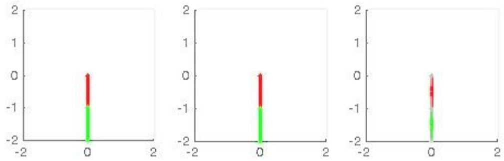
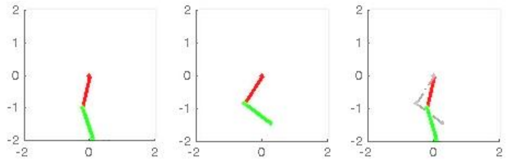


Figure 5.31: tracks of the pendulum bobs of the reference trajectory

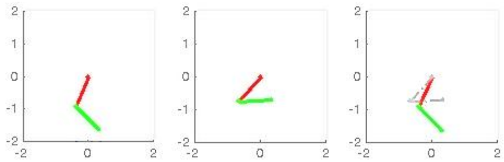
initial data for controlled system: $(0,0,0,0)^T$, simulation time: 8 time units (800 frames), $\alpha = 0$:



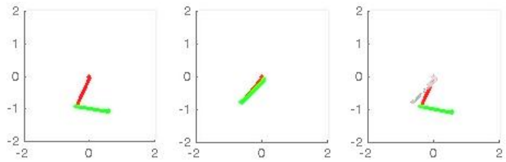
(a) frame 0



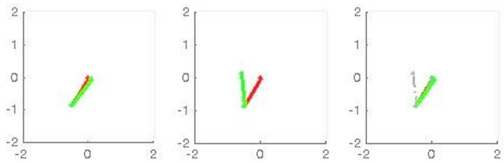
(b) frame 25



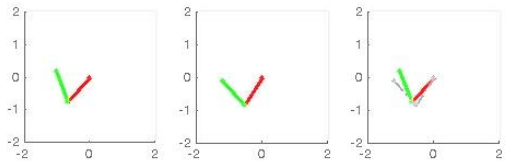
(c) frame 50



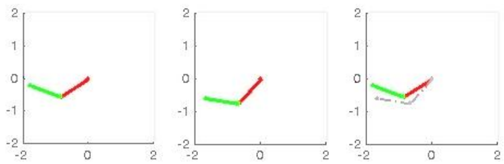
(d) frame 75



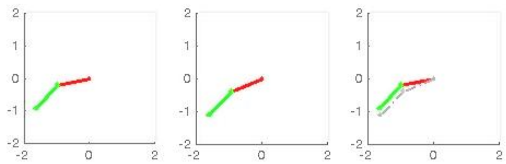
(e) frame 100



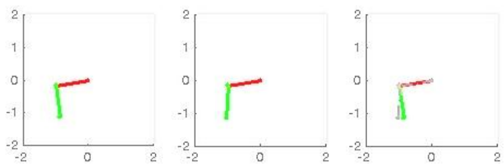
(f) frame 125



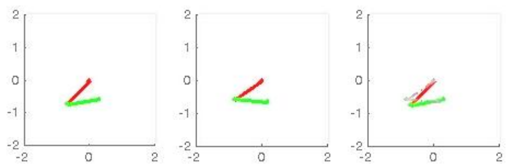
(g) frame 150



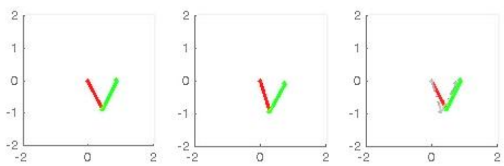
(h) frame 175



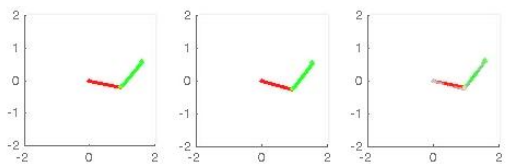
(i) frame 200



(j) frame 250



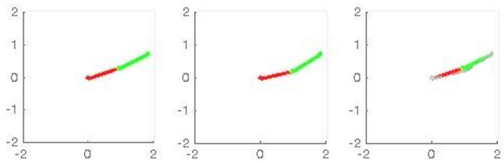
(k) frame 300



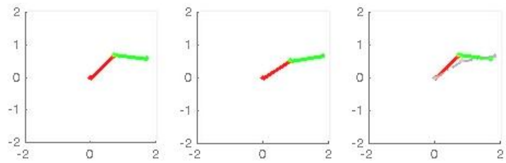
(l) frame 350

Figure 5.32: Part I: Swing-up along a trajectory and balancing ($\alpha = 0$)

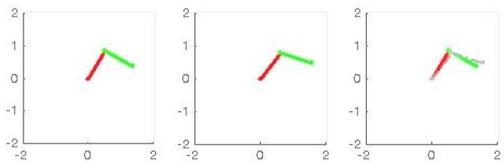
initial data for controlled system: $(0,0,0,0)^T$, simulation time: 8 time units (800 frames), $\alpha = 0$:



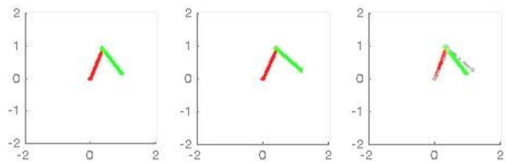
(a) frame 375



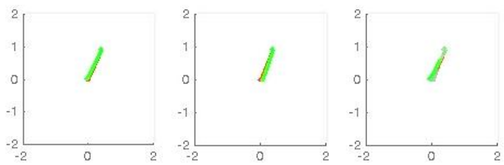
(b) frame 400



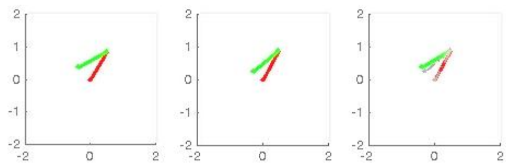
(c) frame 425



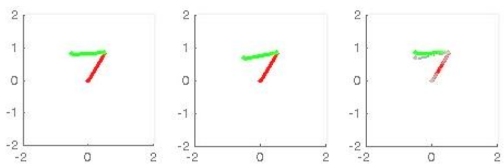
(d) frame 450



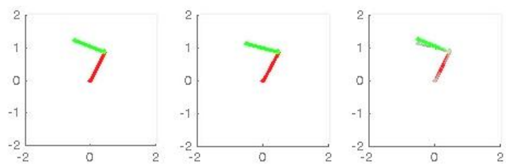
(e) frame 500



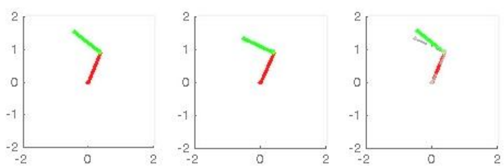
(f) frame 525



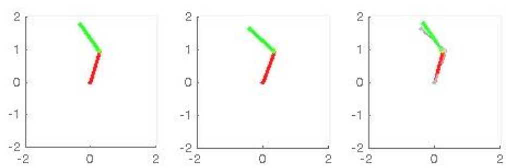
(g) frame 550



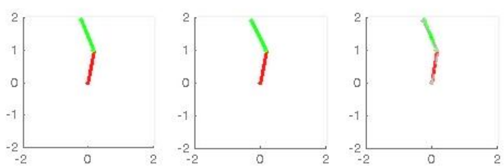
(h) frame 575



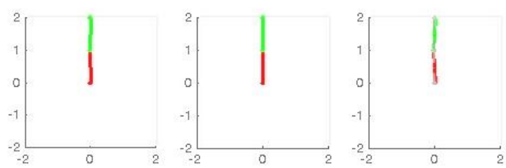
(i) frame 600



(j) frame 625



(k) frame 650



(l) frame 800

Figure 5.33: Part II: Swing-up along a trajectory and balancing ($\alpha = 0$)

Initial data $(0, 0, 0, 0)^T$ and $\alpha = 0$

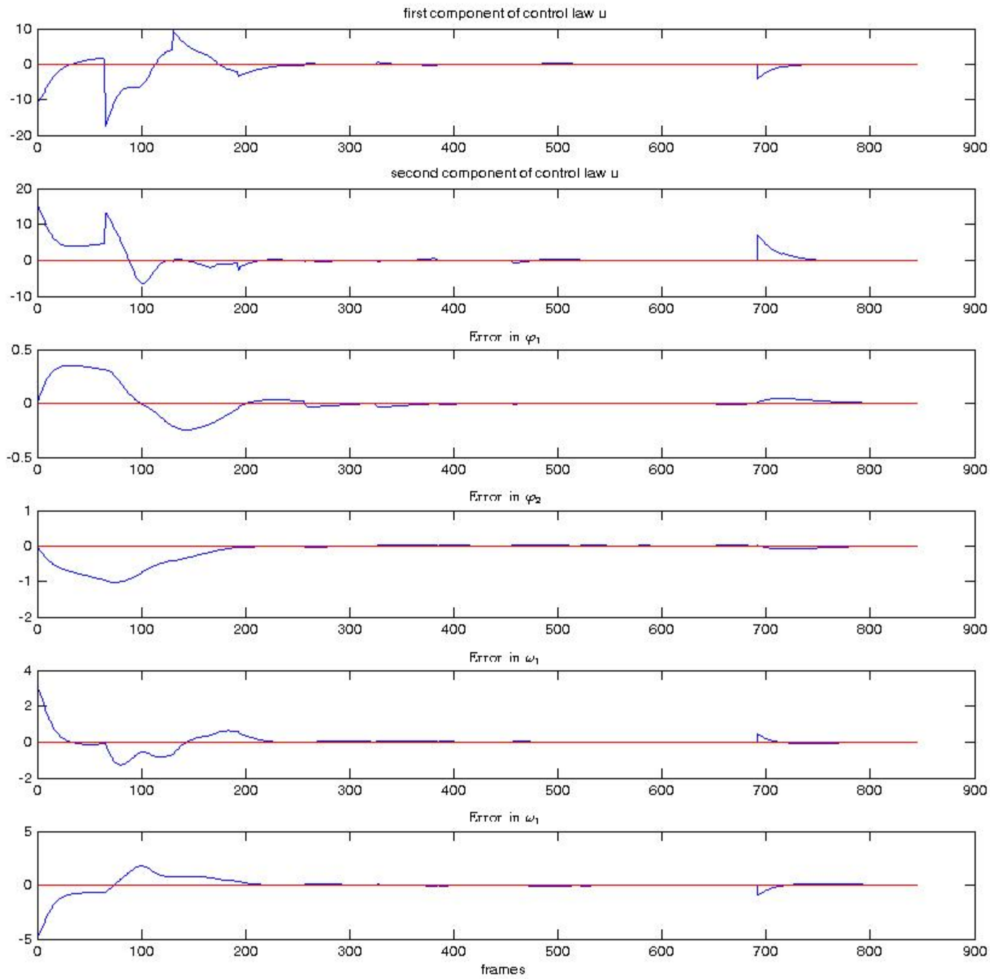
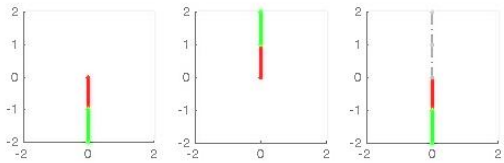
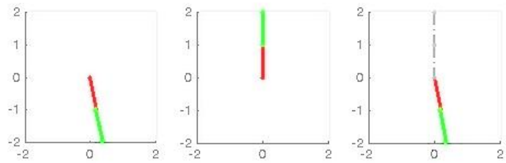


Figure 5.34: Control law and deviation in the single components

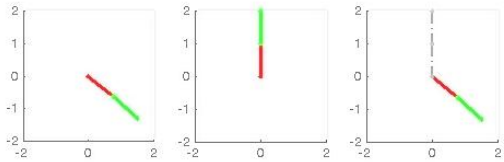
initial data for controlled system: $(0,0,0,0)^T$, simulation time: 8 time units (800 frames), $\alpha = 0$:



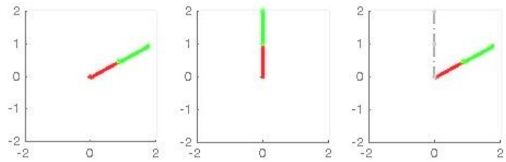
(a) frame 0



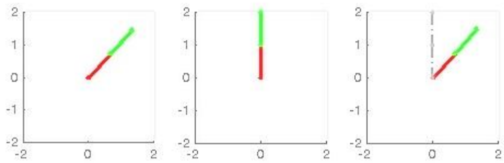
(b) frame 10



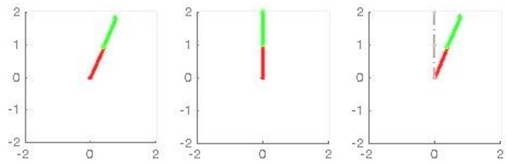
(c) frame 25



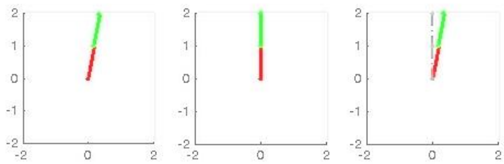
(d) frame 50



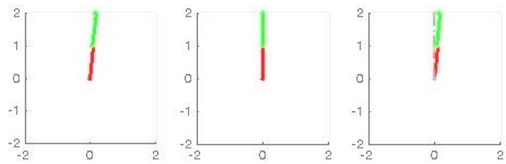
(e) frame 60



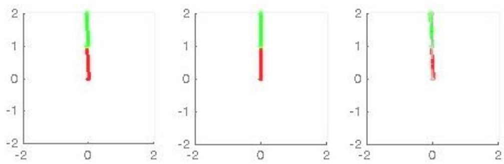
(f) frame 75



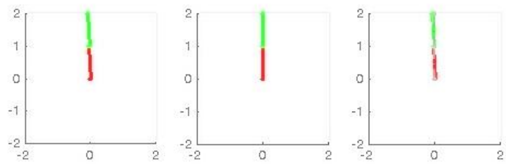
(g) frame 90



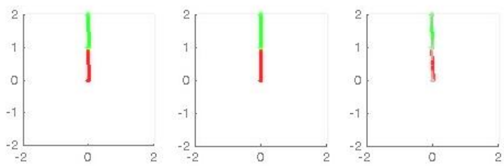
(h) frame 100



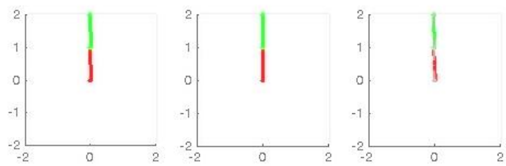
(i) frame 125



(j) frame 150



(k) frame 200



(l) frame 800

Figure 5.35: Swing-up and balancing ($\alpha = 0$)

Initial data $(0,0,0,0)^T$ and $\alpha = 0$

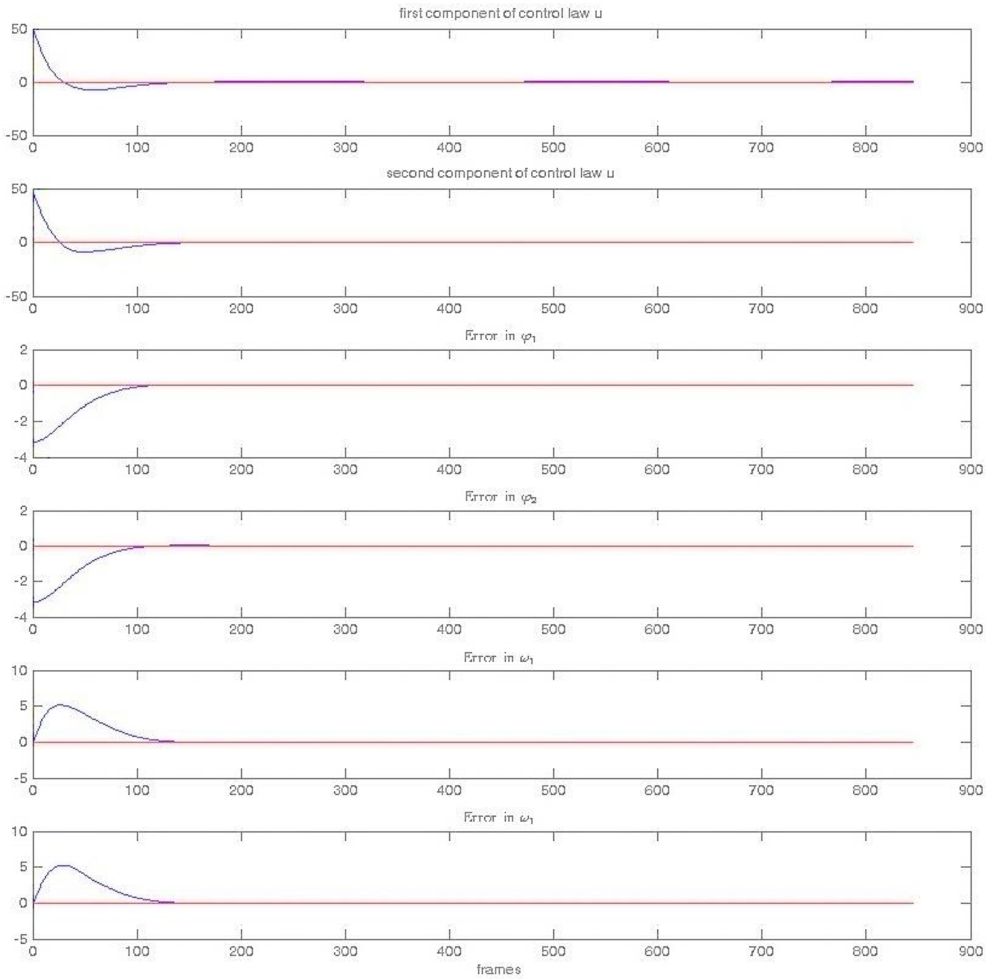


Figure 5.36: Control law and deviation in the single components

5.3 Triple Pendulum

5.3.1 System dynamics

We consider a mathematical triple pendulum with the same simplifications as for the double pendulum. We obtain the triple pendulum by "'adding"' a further pendulum link of length 1 to

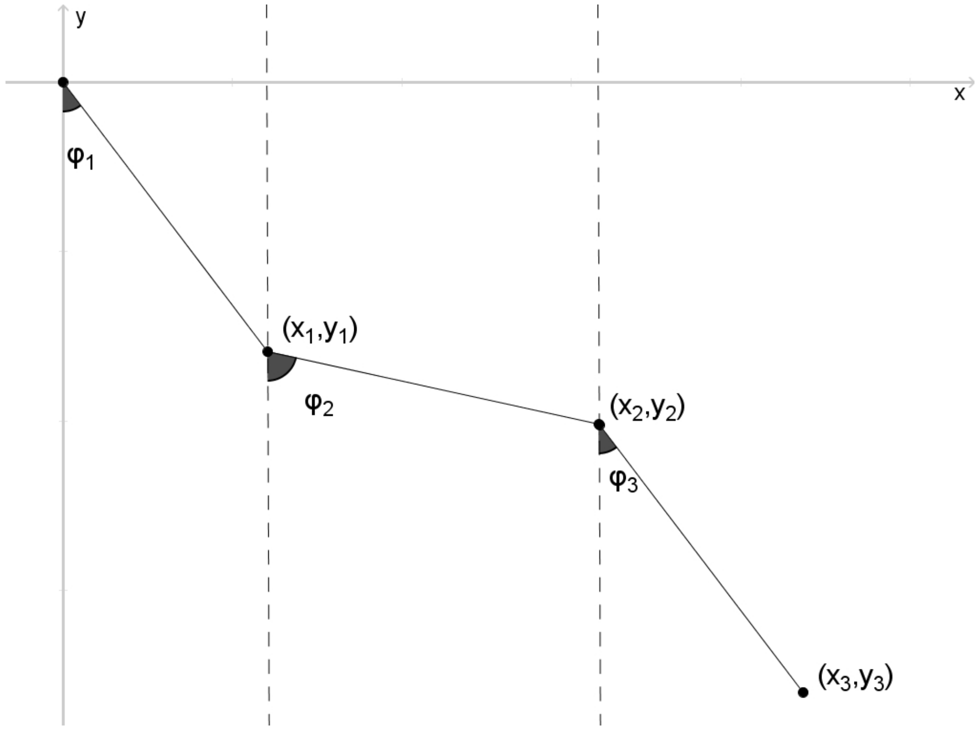


Figure 5.37: mathematical triple pendulum

the double pendulum described above. The equations of motion can be derived as follows:

From figure (5.37) we have

$$x_1 = \sin\varphi_1 \tag{5.201}$$
$$y_1 = \cos\varphi_1 \tag{5.202}$$
$$x_2 = \sin\varphi_1 + \sin\varphi_2 \tag{5.203}$$
$$y_2 = \cos\varphi_1 + \cos\varphi_2 \tag{5.204}$$
$$x_3 = \sin\varphi_1 + \sin\varphi_2 + \sin\varphi_3 \tag{5.205}$$
$$y_3 = \cos\varphi_1 + \cos\varphi_2 + \cos\varphi_3 \tag{5.206}$$

such that for the velocities we obtain

$$\begin{aligned} v_1^2 = \dot{x}_1^2 + \dot{y}_1^2 =& \dot{\varphi}_1^2 \\ v_2^2 = \dot{x}_2^2 + \dot{y}_2^2 =& \dot{\varphi}_1^2 + \dot{\varphi}_2^2 + 2\dot{\varphi}_1\dot{\varphi}_2\cos(\varphi_1 - \varphi_2) \\ v_2^3 = \dot{x}_3^2 + \dot{y}_3^2 =& \dot{\varphi}_1^2 + \dot{\varphi}_2^2 + \dot{\varphi}_3^2 + \\ &+ 2\dot{\varphi}_1\dot{\varphi}_2(\cos\varphi_1\cos\varphi_2 + \sin\varphi_1\sin\varphi_2) \\ &+ 2\dot{\varphi}_1\dot{\varphi}_3(\cos\varphi_1\cos\varphi_3 + \sin\varphi_1\sin\varphi_3) \\ &+ 2\dot{\varphi}_2\dot{\varphi}_3(\cos\varphi_2\cos\varphi_3 + \sin\varphi_2\sin\varphi_3) \end{aligned} \tag{5.207}$$

The kinetic energy E_{kin} and the potential energy E_{pot} of the triple pendulum system are given by

$$\begin{aligned} E_{\text{kin}} =& \frac{1}{2}v_1^2 + \frac{1}{2}v_2^2 + \frac{1}{2}v_3^2 \qquad (5.208)\\ =& \frac{3}{2}\dot\varphi_1^2 + \dot\varphi_2^2 + \frac{1}{2}\dot\varphi_3^2 + \\ &+ 2\dot\varphi_1\dot\varphi_2\cos(\varphi_1-\varphi_2) + \dot\varphi_1\dot\varphi_3\cos(\varphi_1-\varphi_3) + \dot\varphi_2\dot\varphi_3\cos(\varphi_2-\varphi_3)\\ E_{\text{pot}} =& -3\cos\varphi_1 - 2\cos\varphi_2 - \cos\varphi_3. \qquad (5.209) \end{aligned}$$

The Langrangian L of the triple pendulum system is then defined as

$$\begin{aligned} L(\varphi_1,\varphi_2,\dot\varphi_1,\dot\varphi_2) :=& E_{\text{kin}} - E_{\text{pot}}\\ =& \frac{3}{2}\dot\varphi_1^2 + \dot\varphi_2^2 + \frac{1}{2}\dot\varphi_3^2 + 2\dot\varphi_1\dot\varphi_2\cos(\varphi_1-\varphi_2) + \dot\varphi_1\dot\varphi_3\cos(\varphi_1-\varphi_3) + \qquad (5.210)\\ &+ \dot\varphi_2\dot\varphi_3\cos(\varphi_2-\varphi_3) + 3\cos\varphi_1 + 2\cos\varphi_2 + \cos\varphi_3 \end{aligned}$$

which could as well have been obtained by the general formula for the n-pendulum where $n = 3$ as derived in appendix C as equation (C.4).
Using $\varphi = (\varphi_1, \varphi_2, \varphi_3)^T$ and $\dot\varphi = (\dot\varphi_1, \dot\varphi_2, \dot\varphi_3)^T$ we can derive the equations of motion by evaluating the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot\varphi}(\varphi,\dot\varphi)\right) - \frac{\partial L}{\partial \varphi}(\varphi,\dot\varphi) = 0. \qquad (5.211)$$

We have

$$\begin{aligned} \frac{\partial L}{\partial \dot\varphi_1}(\varphi,\dot\varphi) =& 3\dot\varphi_1 + 2\dot\varphi_2\cos(\varphi_1-\varphi_2) + \dot\varphi_3\cos(\varphi_1-\varphi_3)\\ \frac{d}{dt}\frac{\partial L}{\partial \dot\varphi_1}(\varphi,\dot\varphi) =& 3\ddot\varphi_1 + 2\ddot\varphi_2\cos(\varphi_1-\varphi_2) + 2\dot\varphi_2\sin(\varphi_1-\varphi_2)(\dot\varphi_2-\dot\varphi_1) + \\ &+ \ddot\varphi_3\cos(\varphi_1-\varphi_3) + \dot\varphi_3\sin(\varphi_1-\varphi_3)(\dot\varphi_3-\dot\varphi_1)\\ \frac{\partial L}{\partial \varphi_1}(\varphi,\dot\varphi) =& -2\dot\varphi_1\dot\varphi_2\sin(\varphi_1-\varphi_2) - \dot\varphi_1\dot\varphi_3\sin(\varphi_1-\varphi_3) - 3\sin\varphi_1\\ \frac{\partial L}{\partial \dot\varphi_2}(\varphi,\dot\varphi) =& 2\dot\varphi_2 + 2\dot\varphi_1\cos(\varphi_1-\varphi_2) + \dot\varphi_3\cos(\varphi_2-\varphi_3)\\ \frac{d}{dt}\frac{\partial L}{\partial \dot\varphi_2}(\varphi,\dot\varphi) =& 2\ddot\varphi_2 + 2\ddot\varphi_1\cos(\varphi_1-\varphi_2) + 2\dot\varphi_1(\dot\varphi_2-\dot\varphi_1)\sin(\varphi_1-\varphi_2) + \\ &+ \ddot\varphi_3\cos(\varphi_2-\varphi_3) + \dot\varphi_3(\dot\varphi_3-\dot\varphi_2)\sin(\varphi_2-\varphi_3)\\ \frac{\partial L}{\partial \varphi_2}(\varphi,\dot\varphi) =& 2\dot\varphi_1\dot\varphi_2\sin(\varphi_1-\varphi_2) - \dot\varphi_2\dot\varphi_3\sin(\varphi_2-\varphi_3) - 2\sin\varphi_2\\ \frac{\partial L}{\partial \dot\varphi_3}(\varphi,\dot\varphi) =& \dot\varphi_3 + \dot\varphi_1\cos(\varphi_1-\varphi_3) + \dot\varphi_2\cos(\varphi_2-\varphi_3)\\ \frac{d}{dt}\frac{\partial L}{\partial \dot\varphi_3}(\varphi,\dot\varphi) =& \ddot\varphi_3 + \ddot\varphi_1\cos(\varphi_1-\varphi_3) + \dot\varphi_1(\dot\varphi_3-\dot\varphi_1)\sin(\varphi_1-\varphi_3) + \\ &+ \ddot\varphi_2\cos(\varphi_2-\varphi_3) + \dot\varphi_2(\dot\varphi_3-\dot\varphi_2)\sin(\varphi_2-\varphi_3)\\ \frac{\partial L}{\partial \varphi_3}(\varphi,\dot\varphi) =& \dot\varphi_1\dot\varphi_3\sin(\varphi_1-\varphi_3) + \dot\varphi_2\dot\varphi_3\sin(\varphi_2-\varphi_3) - \sin\varphi_3 \end{aligned}$$

and with the Euler-Lagrange equation (5.211) we obtain the equations of motion of the mathematical triple pendulum:

$$3\ddot{\varphi}_1+2\ddot{\varphi}_2\cos(\varphi_1-\varphi_2)+2\dot{\varphi}_2^2\sin(\varphi_1-\varphi_2)+\ddot{\varphi}_3\cos(\varphi_1-\varphi_3)+\dot{\varphi}_3^2\sin(\varphi_1-\varphi_3)+3\sin\varphi_1=0 \quad (5.212)$$

$$2\ddot{\varphi}_2+2\ddot{\varphi}_1\cos(\varphi_1-\varphi_2)-2\dot{\varphi}_1^2\sin(\varphi_1-\varphi_2)+\ddot{\varphi}_3\cos(\varphi_2-\varphi_3)+\dot{\varphi}_3^2\sin(\varphi_2-\varphi_3)+2\sin\varphi_2=0 \quad (5.213)$$

$$\ddot{\varphi}_3+\ddot{\varphi}_1\cos(\varphi_1-\varphi_3)-\dot{\varphi}_1^2\sin(\varphi_1-\varphi_3)+\ddot{\varphi}_2\cos(\varphi_2-\varphi_3)-\dot{\varphi}_2^2\sin(\varphi_2-\varphi_3)+\sin\varphi_3=0 \quad (5.214)$$

Introducing the variables $\omega_i(t)=\frac{d}{dt}\varphi_i(t)$ for $i\in\{1,2,3\}$ we obtain the first order differential equation system

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\varphi_1\\ \varphi_2\\ \varphi_3\\ \omega_1\\ \omega_2\\ \omega_3\end{pmatrix}=\begin{pmatrix}
\omega_1\\
\omega_2\\
\omega_3\\
-\big(2\cos^2(\varphi_1-\varphi_3)-6-4\cos(\varphi_1-\varphi_2)\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_3)+4\cos^2(\varphi_1-\varphi_2)\\
+3\cos^2(\varphi_2-\varphi_3)\big)^{-1}\big(-4\omega_2^2\sin(\varphi_1-\varphi_2)+2\omega_2^2\sin(\varphi_1-\varphi_2)\cos^2(\varphi_2-\varphi_3)\\
-2\omega_3^2\sin(\varphi_1-\varphi_3)+\omega_3^2\sin(\varphi_1-\varphi_3)\cos^2(\varphi_2-\varphi_3)-6\sin\varphi_1+3\sin\varphi_1\cos^2(\varphi_2-\varphi_3)\\
-4\cos(\varphi_1-\varphi_2)\omega_1^2\sin(\varphi_1-\varphi_2)+2\cos(\varphi_1-\varphi_2)\omega_3^2\sin(\varphi_2-\varphi_3)+4\cos(\varphi_1-\varphi_2)\sin\varphi_2\\
+2\cos(\varphi_1-\varphi_2)\cos(\varphi_2-\varphi_3)\omega_1^2\sin(\varphi_1-\varphi_3)+2\cos(\varphi_1-\varphi_2)\cos(\varphi_2-\varphi_3)\omega_3^2\sin(\varphi_2-\varphi_3)\\
-2\cos(\varphi_1-\varphi_2)\cos(\varphi_2-\varphi_3)\sin\varphi_3-2\cos(\varphi_1-\varphi_3)\omega_1^2\sin(\varphi_1-\varphi_3)+2\cos(\varphi_1-\varphi_3)\\
\cdot\cos(\varphi_2-\varphi_3)\omega_1^2\sin(\varphi_1-\varphi_2)-\cos(\varphi_1-\varphi_3)\cos(\varphi_2-\varphi_3)\omega_3^2\sin(\varphi_2-\varphi_3)\\
-2\cos(\varphi_1-\varphi_3)\cos(\varphi_2-\varphi_3)\sin\varphi_2-2\cos(\varphi_1-\varphi-3)\omega_2^2\sin(\varphi_2-\varphi_3)\\
+2\cos(\varphi_1-\varphi_3)\sin\varphi_3\big)\\
-\big(2\cos^2(\varphi_1-\varphi_3)-6-4\cos(\varphi_1-\varphi_2)\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_3)+4\cos^2(\varphi_1-\varphi_2)\\
+3\cos^2(\varphi_2-\varphi_3)\big)^{-1}\big(-3\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_3)\sin\varphi_1-3\cos(\varphi_2-\varphi_3)\omega_2^2\sin(\varphi_2-\varphi_3)\\
-3\cos(\varphi_2-\varphi_3)\omega_1^2\sin(\varphi_1-\varphi_3)-2\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_3)\omega_2^2\sin(\varphi_1-\varphi_2)\\
-\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_3)\omega_3^2\sin(\varphi_1-\varphi_3)+3\cos(\varphi_2-\varphi_3)\sin\varphi_3+2\sin\varphi_2\cdot\\
\cdot\cos^2(\varphi_1-\varphi_3)-3\omega_3^2\sin(\varphi_2-\varphi_3)+6\omega_1^2\sin(\varphi_1-\varphi_2)-6\sin\varphi_2-2\omega_1^2\sin(\varphi_1-\varphi_2)\\
\cdot\cos^2(\varphi_1-\varphi_3)+2\cos(\varphi_1-\varphi_2)\cos(\varphi_1-\varphi_3)\omega_2^2\sin(\varphi_2-\varphi-3)+2\cos(\varphi_1-\varphi_2)\\
\cdot\omega_3^2\sin(\varphi_1-\varphi_3)+4\cos(\varphi_1-\varphi_2)\omega_2^2\sin(\varphi_1-\varphi_2)-2\cos(\varphi_1-\varphi_2)\cos(\varphi_1-\varphi_3)\sin\varphi_3\\
+2\cos(\varphi_1-\varphi_2)\cos(\varphi_1-\varphi_3)\omega_1^2\sin(\varphi_1-\varphi_3)+6\cos(\varphi_1-\varphi_2)\sin\varphi_1\\
+\omega_3^2\sin(\varphi_2-\varphi_3)\cos^2(\varphi_1-\varphi_3)\big)\\
\big(2\cos^2(\varphi_1-\varphi_3)-6-4\cos(\varphi_1-\varphi_2)\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_3)+4\cos^2(\varphi_1-\varphi_2)\\
+3\cos^2(\varphi_2-\varphi_3)\big)^{-1}\big(6\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_2)\sin\varphi_1+6\cos(\varphi_2-\varphi_3)\omega_1^2\sin(\varphi_1-\varphi_2)\\
-3\cos(\varphi_2-\varphi_3)\omega_3^2\sin(\varphi_2-\varphi_3)+4\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_2)\omega_2^2\sin(\varphi_1-\varphi_2)\\
+2\cos(\varphi_2-\varphi_3)\cos(\varphi_1-\varphi_2)\omega_3^2\sin(\varphi_1-\varphi_3)-6\cos(\varphi_2-\varphi_3)\sin\varphi_2+2\cos(\varphi_1-\varphi_3)\\
\cdot\cos(\varphi_1-\varphi_2)\omega_3^2\sin(\varphi_2-\varphi_3)\\
+4\omega_1^2\sin(\varphi_1-\varphi_3)\cos^2(\varphi_1-\varphi_2)-4\cos(\varphi_1-\varphi_3)\omega_2^2\sin(\varphi_1-\varphi_2)+4\omega_2^2\sin(\varphi_2-\varphi_3)\\
\cdot\cos^2(\varphi_1-\varphi_2)+4\cos(\varphi_1-\varphi_3)\cos(\varphi_1-\varphi_2)\sin\varphi_2-4\sin\varphi_3\cos^2(\varphi_1-\varphi_2)\\
-2\cos(\varphi_1-\varphi_3)\omega_3^2\sin(\varphi_1-\varphi_3)-6\cos(\varphi_1-\varphi_3)\sin\varphi_1-4\cos(\varphi_1-\varphi_3)\\
\cdot\cos(\varphi_1-\varphi_2)\omega_1^2\sin(\varphi_1-\varphi_2)\\
-6\omega_2^2\sin(\varphi_2-\varphi_3)-6\omega_1^2\sin(\varphi_1-\varphi_3)+6\sin\varphi_3\big)
\end{pmatrix}$$

where we denote the right hand side as $f(\varphi_1,\varphi_2,\varphi_3,\omega_1,\omega_2,\omega_3)$.

The steady states of this equation can be obtained by setting ω_1,ω_2 and ω_3 as well as their time-derivatives equal to zero. An easier way to find the steady states is to use formula (C.13), which directly shows that the condition to be satisfied is

$$\sin \varphi_i = 0 \Longleftrightarrow \varphi_i \in \{0, \pi\} \text{ for } i \in \{1, 2, 3\} \tag{5.215}$$

since $\varphi_i \in [0, 2\pi)$, such that the equilibrium states are

$$\begin{gathered}(0,0,0),\ (\pi,0,0),\ (0,\pi,0),\ (0,0,\pi),\\ (\pi,\pi,0),\ (\pi,0,\pi),\ (0,\pi,\pi),\ (\pi,\pi,\pi).\end{gathered}$$

This means the mathematical triple pendulum can only be permanently at rest if and only if all the pendulum links point either up or down.

5.3.2 Discussion of the linearized triple pendulum

Linearizing along the trajectory $(\tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\varphi}_3, \tilde{\omega}_1, \tilde{\omega}_2, \tilde{\omega}_3)$ yields the linear differential equation

$$\frac{d}{dt}(\varphi_1, \varphi_2, \varphi_3, \omega_1, \omega_2, \omega_3)^T = A_{(\tilde{\varphi}_1,\tilde{\varphi}_2,\tilde{\varphi}_3,\tilde{\omega}_1,\tilde{\omega}_2,\tilde{\omega}_3)}(\varphi_1, \varphi_2, \varphi_3, \omega_1, \omega_2, \omega_3)^T \tag{5.216}$$

where

$$A_{(\tilde{\varphi}_1,\tilde{\varphi}_2,\tilde{\varphi}_3,\tilde{\omega}_1,\tilde{\omega}_2\tilde{\omega}_3)} = \frac{\partial f}{\partial(\varphi_1, \varphi_2, \varphi_3, \omega_1, \omega_2, \omega_3)}(\tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\varphi}_3, \tilde{\omega}_1, \tilde{\omega}_2\tilde{\omega}_3) \tag{5.217}$$

is in general a non-autonomous 6×6 matrix where the first 3 rows are given by

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ & \ddots & & & & \\ & & \ddots & & & \\ & & & \ddots & & \end{pmatrix}$$

Controllability of the linearization around a point

For the special case that the trajectory $(\tilde{\varphi}_1, \tilde{\varphi}_2, \tilde{\varphi}_3, \tilde{\omega}_1, \tilde{\omega}_2\tilde{\omega}_3)$ reduces to a single point $(\varphi_1^*, \varphi_2^*, \varphi_3^*, \omega_1^*, \omega_2^*, \omega_3^*)$ the constant linear control system

$$\frac{d}{dt}(\varphi_1, \varphi_2, \varphi_3, \omega_1, \omega_2, \omega_3)^T = A_{(\varphi_1^e,\varphi_2^e,\varphi_3^e,\omega_1^e,\omega_2^e,\omega_3^e)}(\varphi_1, \varphi_2, \varphi_3, \omega_1, \omega_2, \omega_3)^T + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} u(t) \tag{5.218}$$

is completely controllable as shown in appendix (C.4).

5.3.3 Stability of the equilibria

For $(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e)^T = (0,0,0,0,0,0)^T$ we obtain

$$A_(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -3 & 2 & 0 & 0 & 0 & 0 \\ 3 & -4 & 1 & 0 & 0 & 0 \\ 0 & 2 & -2 & 0 & 0 & 0 \end{pmatrix} \tag{5.219}$$

with 3 pairs of complex conjugate eigenvalues each with real part zero. Therefore the zero-solution of the linear system (5.216) is stable but not asymptotically stable.

For $(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e)^T = (\pi,0,0,0,0,0)^T$ we obtain

$$A_(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3 & -2 & 0 & 0 & 0 & 0 \\ 3 & -4 & 1 & 0 & 0 & 0 \\ 0 & 2 & -2 & 0 & 0 & 0 \end{pmatrix} \tag{5.220}$$

with two pairs of complex conjugate eigenvalues with real part zero and two real eigenvalues, one of them less than zero, such that the zero-solution of (5.216) is unstable in this case.

For $(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e)^T = (0,\pi,0,0,0,0)^T$ we obtain

$$A_(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -3 & 2 & 0 & 0 & 0 & 0 \\ -3 & 4 & -1 & 0 & 0 & 0 \\ 0 & 2 & -2 & 0 & 0 & 0 \end{pmatrix} \tag{5.221}$$

again with two pairs of complex conjugate eigenvalues with real part zero and two real eigenvalues, one of them less than zero, such that the zero-solution of (5.216) is unstable in this case as well.

For $(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e)^T = (0,0,\pi,0,0,0)^T$ we obtain

$$A_(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -3 & 2 & 0 & 0 & 0 & 0 \\ 3 & -4 & 1 & 0 & 0 & 0 \\ 0 & -2 & 2 & 0 & 0 & 0 \end{pmatrix} \tag{5.222}$$

with two pairs of complex conjugate eigenvalues with real part zero and two real eigenvalues, one of them less than zero, such that the zero-solution of (5.216) is unstable.

For $(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e)^T = (\pi, \pi, 0, 0, 0, 0)^T$ we obtain

$$A_(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3 & -2 & 0 & 0 & 0 & 0 \\ -3 & 4 & -1 & 0 & 0 & 0 \\ 0 & 2 & -2 & 0 & 0 & 0 \end{pmatrix} \tag{5.223}$$

with a pair of complex conjugate eigenvalues with real part zero and four real eigenvalues, two of them less than zero, such that the zero-solution of (5.216) is unstable.

For $(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e)^T = (\pi, 0, \pi, 0, 0, 0)^T$ we obtain

$$A_(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3 & -2 & 0 & 0 & 0 & 0 \\ 3 & -4 & 1 & 0 & 0 & 0 \\ 0 & -2 & 2 & 0 & 0 & 0 \end{pmatrix} \tag{5.224}$$

again with a pair of complex conjugate eigenvalues with real part zero and four real eigenvalues, two of them less than zero, such that the zero-solution of (5.216) is unstable.

For $(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e)^T = (0, \pi, \pi, 0, 0, 0)^T$ we obtain

$$A_(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -3 & 2 & 0 & 0 & 0 & 0 \\ -3 & 4 & -1 & 0 & 0 & 0 \\ 0 & -2 & 2 & 0 & 0 & 0 \end{pmatrix} \tag{5.225}$$

with a pairs of complex conjugate eigenvalues with real part zero and four real eigenvalues, two of them less than zero, such that the zero-solution of (5.216) is unstable.

For the last equilibrium state $(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e)^T = (\pi, \pi, \pi, 0, 0, 0)^T$ we obtain

$$A_(\varphi_1^e, \varphi_2^e, \varphi_3^e, \omega_1^e, \omega_2^e, \omega_3^e) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 3 & -2 & 0 & 0 & 0 & 0 \\ -3 & 4 & -1 & 0 & 0 & 0 \\ 0 & -2 & 2 & 0 & 0 & 0 \end{pmatrix} \tag{5.226}$$

with no complex eigenvalues and three of the six real eigenvalues less than zero, therefore system (5.216) is unstable.

The number of eigenvalues with real part smaller than zero corresponds to the number of pendulum links which are pointed upward.

5.3.4 Stabilization

The model of consideration is the linear control system

$$\dot{x}(t) = A_{(\pi,0,0,0,0,0)}x(t) + Bu(t) \tag{5.227}$$

with $B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and the usual assumptions on u.

Method of Bass

We apply the method of Bass (4.2.2) to find a feedback control $\tilde{u}$ such that the zero-solution is stable with respect to the system

$$\dot{x}(t) = A_{(\pi,0,0,0,0,0)}x(t) + B\tilde{u}(t). \tag{5.228}$$

First we need to choose a $\lambda \in \mathbb{R}$ such that (4.94) holds:

$$\lambda > \max_i \sum_{j=1}^{6} |a_{ij}| = 8$$

proposing that $\lambda = 9$ is a suitable choice. Now solving equation

$$(A_{(\pi,0,0,0,0,0)} + \lambda I)P + P(A_{(\pi,0,0,0,0,0)} + \lambda I)^T = BB^T \tag{5.229}$$

with $\lambda = 9$ and the assumption $P = P^T$ we obtain (numerically)

$$P = \begin{pmatrix} 0.0004 & 0.0000 & 0.0000 & -0.0032 & -0.0001 & 0.0000 \\ 0.0000 & 0.0003 & 0.0000 & 0.0000 & -0.0029 & -0.0001 \\ 0.0000 & 0.0000 & 0.0003 & 0.0000 & 0.0000 & -0.0030 \\ -0.0032 & 0.0000 & 0.0000 & 0.0566 & 0.0002 & 0.0000 \\ -0.0001 & -0.0029 & 0.0000 & 0.0002 & 0.0543 & 0.0005 \\ 0.0000 & -0.0001 & -0.0030 & 0.0000 & 0.0005 & 0.0549 \end{pmatrix} \tag{5.230}$$

and according to the theory of Bass (4.2.2) $\tilde{u}(t) = -B^T P^{-1} x(t)$ is a stabilizing control law and the controlled linear system with system matrix $(A_{(\pi,0,0,0,0,0)} + \lambda I - BB^T P^{-1})$ has an exponentially stable zero-solution and with system matrix $A_{(\pi,0,0,0,0,0)} - BB^T P^{-1}$ the zero-solution is stable with decay rate at least $-\lambda$ which is -9 here.
A numerical verification shows that all eigenvalues of $(A_{(\pi,0,0,0,0,0)} + \lambda I - BB^T P^{-1})$ have real part smaller than -7.5 and all eigenvalues of $A_{(\pi,0,0,0,0,0)} - BB^T P^{-1}$ have real part smaller than -16.5. As for the the simple and double pendulum the (exponential) decay rate obtained by applying the method of Bass is better than expected from theory (which would be -9 here).

Pole-shifting

The method of Bass guarantees a minimum decay rate. It does not provide an upper bound for this rate, which in practical purposes may be undesirable. Whenever technical systems interact with human beings it is essential to limit physical forces. For example one should set an upper

bound for elevators as well for the velocity as for the acceleration to avoid accidents.
The pole-shifting method allows us to predefine the location of the eigenvalues of the controlled system and therefore the decay rate. We look for a feedback control $\tilde{u}(t) = Fx(t)$ with a 3×6 matrix F such that all eigenvalues of $A_{(\pi,0,0,0,0,0)} + BF$ have negative real part, which can be chosen in advance. For example we could try to choose F such that for the characteristic polynomial we obtain

$$\rho_{A_{(\pi,0,0,0,0,0)}+BK}(\lambda) \stackrel{!}{=} (\lambda+1)^6, \tag{5.231}$$

with -1 as an eigenvalue of multiplicity 4. Together with $F = \begin{pmatrix} f_{11} & f_{12} & f_{13} & f_{14} & f_{15} & f_{16} \\ f_{21} & f_{22} & f_{23} & f_{24} & f_{25} & f_{26} \\ f_{31} & f_{32} & f_{33} & f_{34} & f_{35} & f_{36} \end{pmatrix}$ equation (5.231) becomes

$$\begin{aligned}
&\lambda^6 + (-f_{36} - f_{25} - f_{14})\,\lambda^5 + (-f_{35}f_{26} - f_{34}f_{16} + f_{25}f_{14} + f_{36}f_{14} - f_{24}f_{15} - f_{22} + 3 - f_{33} \\
&+f_{36}f_{25} - f_{11})\,\lambda^4 + (-f_{24}f_{12} - f_{34}f_{13} + f_{25}f_{11} + f_{33}f_{14} + f_{14}f_{22} + f_{36}f_{22} - f_{26}f_{32} - f_{15}f_{21} \\
&-f_{35}f_{23} + f_{36}f_{11} - f_{31}f_{16} + f_{33}f_{25} + f_{36}f_{24}f_{15} - f_{34}f_{15}f_{26} + f_{35}f_{26}f_{14} - f_{35}f_{24}f_{16} - f_{36}f_{25}f_{14} \\
&-6\,f_{14} - 3\,f_{15} + 2\,f_{24} + f_{25} - 2\,f_{26} - f_{35} - f_{36} + f_{34}f_{16}f_{25})\,\lambda^3 + (-6 - 3\,f_{35}f_{16} - 2\,f_{24}f_{16} \\
&+2\,f_{34}f_{26} - 3\,f_{36}f_{25} + 3\,f_{36}f_{15} + 2\,f_{14}f_{26} - 2\,f_{24}f_{15} + 2\,f_{25}f_{14} + 4\,f_{36}f_{14} - f_{34}f_{15} - 4\,f_{34}f_{16} \\
&-2\,f_{36}f_{24} + 3\,f_{35}f_{26} + f_{35}f_{14} + f_{33}f_{22} + f_{33}f_{11} + f_{31}f_{16}f_{25} - f_{31}f_{15}f_{26} - f_{34}f_{15}f_{23} - f_{35}f_{24}f_{13} \\
&-f_{35}f_{16}f_{21} + f_{35}f_{26}f_{11} + f_{34}f_{16}f_{22} + f_{36}f_{24}f_{12} + f_{26}f_{32}f_{14} - f_{36}f_{14}f_{22} - f_{36}f_{25}f_{11} - f_{34}f_{26}f_{12} \\
&+f_{35}f_{23}f_{14} + f_{34}f_{13}f_{25} + f_{36}f_{15}f_{21} - f_{33}f_{25}f_{14} - f_{32}f_{24}f_{16} + f_{33}f_{24}f_{15} - 6\,f_{11} - 3\,f_{12} + 2\,f_{21} \\
&+f_{22} - 2\,f_{23} - f_{32} - f_{33} + f_{22}f_{11} - f_{21}f_{12} - f_{31}f_{13} - f_{32}f_{23})\,\lambda^2 + (-2\,f_{24}f_{12} - 4\,f_{34}f_{13} \\
&-2\,f_{24}f_{13} + f_{32}f_{14} + 2\,f_{34}f_{23} - 2\,f_{36}f_{21} + 2\,f_{11}f_{26} - 2\,f_{16}f_{21} - 3\,f_{32}f_{16} + 3\,f_{33}f_{15} + 2\,f_{25}f_{11} \\
&-f_{31}f_{15} + 4\,f_{33}f_{14} + 2\,f_{14}f_{22} - f_{34}f_{12} - 3\,f_{36}f_{22} - 2\,f_{33}f_{24} + 3\,f_{36}f_{12} + 3\,f_{26}f_{32} - 2\,f_{15}f_{21} \\
&+3\,f_{35}f_{23} + 2\,f_{31}f_{26} + 2\,f_{14}f_{23} + 4\,f_{36}f_{11} + f_{35}f_{11} - 3\,f_{35}f_{13} - 4\,f_{31}f_{16} - 3\,f_{33}f_{25} + f_{31}f_{16}f_{22} \\
&-f_{31}f_{26}f_{12} - f_{31}f_{15}f_{23} + f_{36}f_{21}f_{12} + f_{31}f_{13}f_{25} + f_{26}f_{32}f_{11} - f_{32}f_{16}f_{21} - f_{32}f_{24}f_{13} + f_{32}f_{23}f_{14} \\
&-f_{34}f_{12}f_{23} - 6\,f_{14} - 6\,f_{15} - 6\,f_{16} + 4\,f_{24} + 6\,f_{25} + 6\,f_{26} + 2\,f_{34} + 3\,f_{35} + 6\,f_{36} - f_{35}f_{21}f_{13} \\
&+f_{34}f_{13}f_{22} + f_{33}f_{24}f_{12} + f_{33}f_{15}f_{21} + f_{35}f_{23}f_{11} - f_{33}f_{25}f_{11} - f_{33}f_{14}f_{22} - f_{36}f_{22}f_{11})\,\lambda \\
&- 6 + f_{33}f_{21}f_{12} - f_{31}f_{12} + 2\,f_{31}f_{23} - 3\,f_{32}f_{13} - f_{33}f_{22}f_{11} + 4\,f_{33}f_{11} + 3\,f_{33}f_{12} - 2\,f_{33}f_{21} \\
&- 3\,f_{33}f_{22} + f_{32}f_{23}f_{11} - f_{31}f_{12}f_{23} - f_{32}f_{21}f_{13} - 6\,f_{11} - 6\,f_{12} - 6\,f_{13} + 4\,f_{21} + 6\,f_{22} + 6\,f_{23} \\
&+ 2\,f_{31} + 3\,f_{32} + 6\,f_{33} + f_{31}f_{13}f_{22} + 2\,f_{22}f_{11} - 2\,f_{21}f_{13} + 2\,f_{23}f_{11} + f_{32}f_{11} - 2\,f_{21}f_{12} - 4\,f_{31}f_{13} \\
&+ 3\,f_{32}f_{23} \stackrel{!}{=} \lambda^6 + 6\,\lambda^5 + 15\,\lambda^4 + 20\,\lambda^3 + 15\,\lambda^2 + 6\,\lambda + 1
\end{aligned} \tag{5.232}$$

Equating coefficients gives 6 equations for 18 unknowns, which means that the problem of pole placement here is underdetermined. We use the matlab command "place" which uses an algorithm described in [Kautsky et al., 1985]. It uses the extra degrees of freedom to minimize the sensitivities of the closed-loop poles to perturbations in $A_{(\pi,0,0,0,0,0)}$ and F. In this method the only restriction is that the multiplicities of the eigenvalues of $A_{(\pi,0,0,0,0,0)} + BF$ are at most the rank of B.
Since B has rank 3 we want $A_{(\pi,0,0,0,0,0)} + BF$ to have eigenvalues -1 and -2 with multiplicity 3 each. The algorithm proposed in [Kautsky et al., 1985] then leads to

$$F = \begin{pmatrix} -5 & 2 & 0 & -3 & 0 & 0 \\ 3 & -6 & 1 & 0 & -3 & 0 \\ 0 & -2 & 0 & 0 & 0 & -3 \end{pmatrix} \tag{5.233}$$

and the feedback control is then given by

$$u(t) = Fx(t) = \begin{pmatrix} -5 & 2 & 0 & -3 & 0 & 0 \\ 3 & -6 & 1 & 0 & -3 & 0 \\ 0 & -2 & 0 & 0 & 0 & -3 \end{pmatrix} x(t) \tag{5.234}$$

which makes the zero-solution stable with respect to (5.228).

5.3.5 Stabilizing upright position via the linear model

Since we are interested in swinging up the pendulum to its unstable position where all three pendulum links point upward, we will apply the method of Bass and the pole shifting method to the linear control problem

$$\dot{x}(t) = A_{(\pi,\pi,\pi,0,0,0)}x(t) + Bu(t). \tag{5.235}$$

with the same B as in (5.228). Using the linear model for stabilizing around the equilibrium point $(\pi, \pi, \pi, 0, 0, 0)$ makes only sense if the state of the triple pendulum is already close to this equilibrium.

Method of Bass

Condition (4.94) suggests as for the example above that

$$\lambda > \max_i \sum_{j=1}^{6} |a_{ij}| = 8$$

$\lambda = 9$ is a suitable choice.
We have to solve the Lyapunov equation

$$(A_{(\pi,\pi,\pi,0,0,0)} + \lambda I)P + P(A_{(\pi,\pi,\pi,0,0,0)} + \lambda I)^T = BB^T \tag{5.236}$$

with $\lambda = 9$ and assuming P is symmetric. We obtain

$$P = \begin{pmatrix} 0.0004 & 0.0000 & 0.0000 & -0.0032 & 0.0001 & 0.0000 \\ 0.0000 & 0.0004 & 0.0000 & 0.0001 & -0.0033 & 0.0001 \\ 0.0000 & 0.0000 & 0.0004 & 0.0000 & 0.0001 & -0.0032 \\ -0.0032 & 0.0001 & 0.0000 & 0.0566 & -0.0009 & 0.0000 \\ 0.0001 & -0.0033 & 0.0001 & -0.0009 & 0.0570 & -0.0006 \\ 0.0000 & 0.0001 & -0.0032 & 0.0000 & -0.0006 & 0.0563 \end{pmatrix} \tag{5.237}$$

and according to the theory of Bass (4.2.2) $\tilde{u}(t) = -B^T P^{-1} x(t)$ is a stabilizing control law and the controlled linear system with system matrix $(A_{(\pi,\pi,\pi,0,0,0)} + \lambda I - BB^T P^{-1})$ has an exponentially stable zero-solution and with system matrix $A_{(\pi,0,0,0,0,0)} - BB^T P^{-1}$ the zero-solution is stable with decay rate at least $-\lambda$ which is -9 here.
A numerical verification shows that all eigenvalues of $(A_{(\pi,\pi,\pi,0,0,0)} + \lambda I - BB^T P^{-1})$ have real part smaller than -6 and all eigenvalues of $A_{(\pi,\pi,\pi,0,0,0)} - BB^T P^{-1}$ have real part smaller than -15.

Pole-shifting

As in the example above we want the controlled system to have the triple eigenvalues -1 and -2. By the pole-shifting method we obtain the feedback matrix

$$F = \begin{pmatrix} -5 & 2 & 0 & -3 & 0 & 0 \\ 3 & -6 & 1 & 0 & -3 & 0 \\ 0 & 2 & -4 & 0 & 0 & -3 \end{pmatrix} \tag{5.238}$$

and the stabilizing feedback control

$$\tilde{u}(t) = Fx(t). \tag{5.239}$$

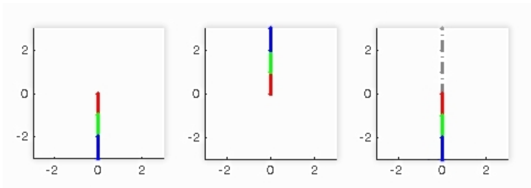
In order to compare the method of Bass with the pole-shifting method, we use the (numerical) eigenvalues of $A - BB^TP^{-1}$ for the design of the feedback matrix F_B:

$$F_B = \begin{pmatrix} -366.8477 & 2.0000 & 0 & -38.1595 & 0 & 0 \\ 3.0000 & -290.1057 & 1.0000 & 0 & -33.8405 & 0 \\ 0 & 2.0000 & -319.7099 & 0 & 0 & -36.0000 \end{pmatrix} \tag{5.240}$$

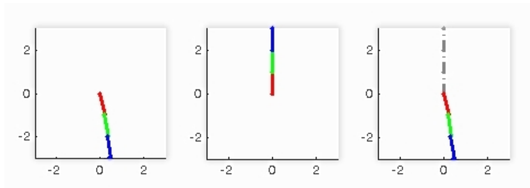
Now $A - BB^TP^{-1}$ has infinity norm ≈ 360 whereas $A + BF_B$ has infinity norm ≈ 402. Comparing the input gains we have to compare the infinity norms of $-B^TP^{-1}$ and F_B where we obtain $\|B^TP^{-1}\|_\infty \approx 362$ whereas $\|F_B\|_\infty \approx 407$.

5.3.6 Simulation

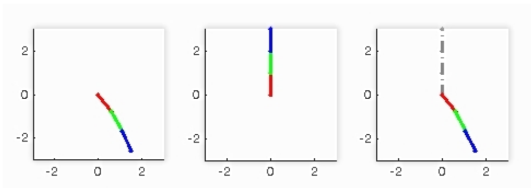
In the following examples we want to swing up the triple pendulum to its unstable equilibrium point $(\pi, \pi, \pi, 0, 0, 0)^T$ and balance it at this position, a problem that is often referred to as inverted triple pendulum. We use the modified control law presented in (4.130) and in a first attempt we choose the desired equilibrium point as reference trajectory $x^0(t)$. As for the double pendulum we choose as initial configuration the stable equilibrium point $(0, 0, 0, 0, 0, 0)^T$ running the matlab simulation for 10 time units with $\alpha = 0$ and $\alpha = 2$. The matrix $\tilde{H}_\alpha^{-1}$ is updated every two time units and therefore it has to be computed only 5 times for each simulation ($\delta_k = 2 \,\forall\, k \implies t_k = 2k,\ k = 0, 1, \ldots, 4$).
We repeat the simulation with the same parameters (simulation time 10 time units, $\delta_k = 2$; $t_k = 2k,\ k = 0, 1, \ldots, 4,\ \alpha \in \{0, 2\}$) for the "semistable" starting configuration $(0, \pi, 0, 0, 0, 0)^T$ to observe how the triple pendulum "unfolds" from its downupdown-position to swing up to the upupup-position.



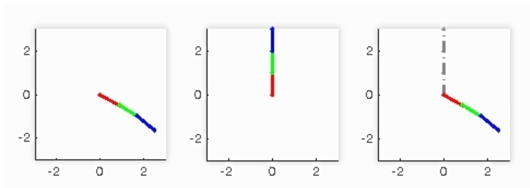
(a) frame 0



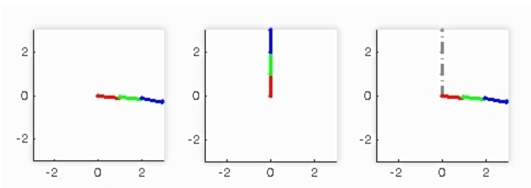
(b) frame 25



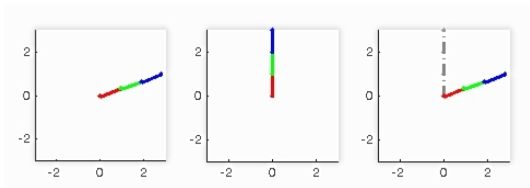
(c) frame 50



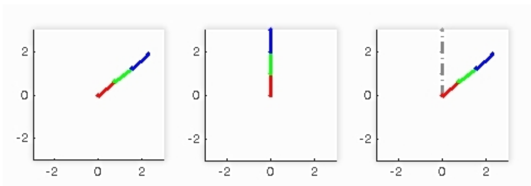
(d) frame 75



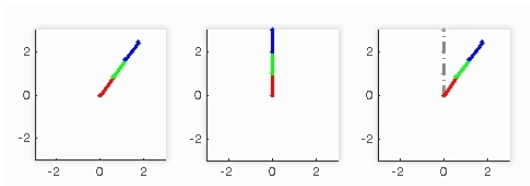
(e) frame 100



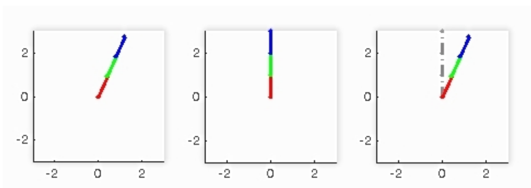
(f) frame 125



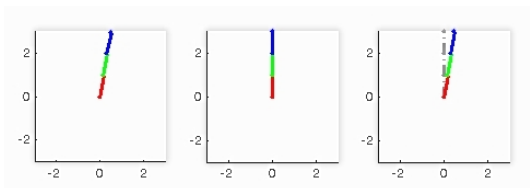
(g) frame 150



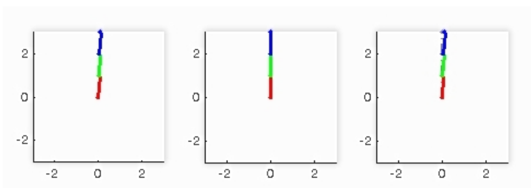
(h) frame 175



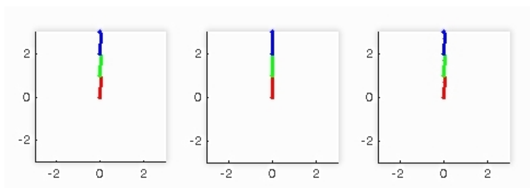
(i) frame 200



(j) frame 250



(k) frame 300



(l) frame 1000

Figure 5.38: Swing-up from downdowndown-position to upupup-position and balancing ($\alpha = 0$)

Initial data $(0, 0, 0, 0, 0, 0)^T$ and $\alpha = 0$

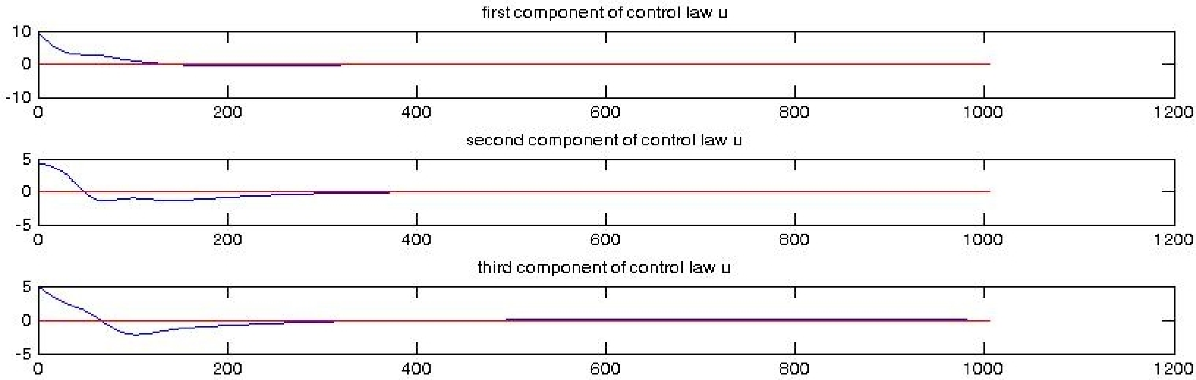


Figure 5.39: Control law u

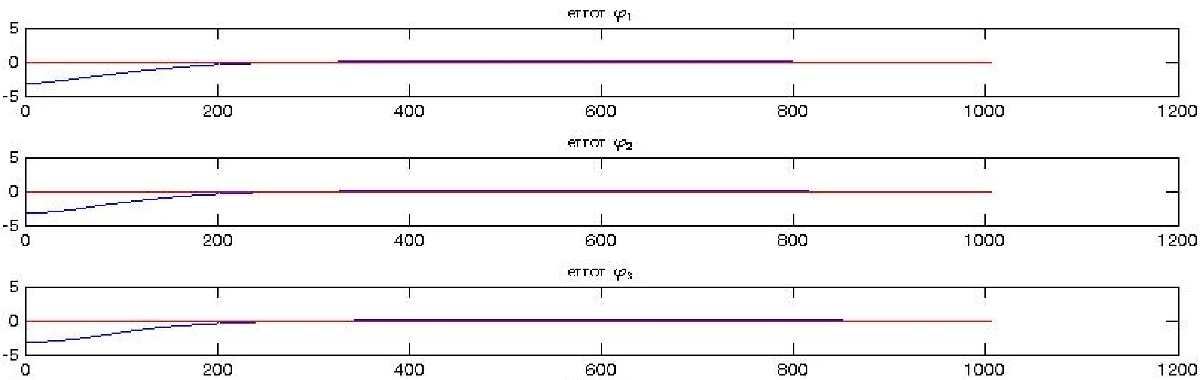


Figure 5.40: Error in state variables u

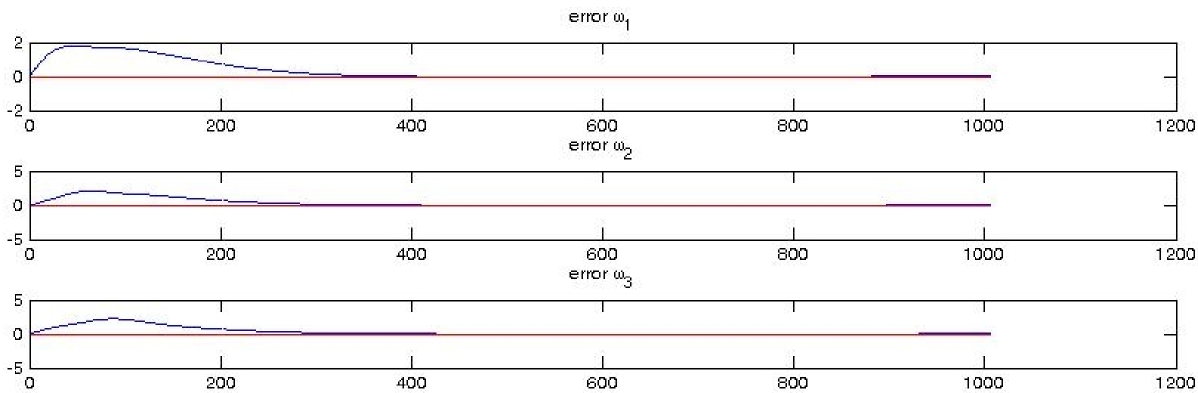
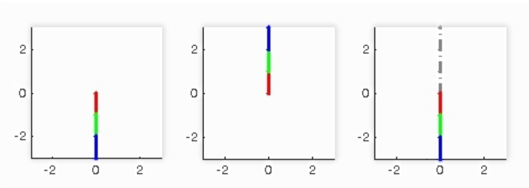
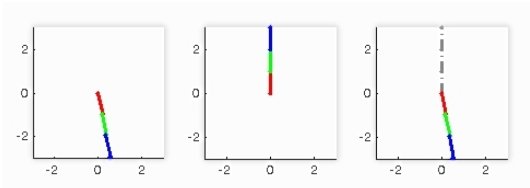
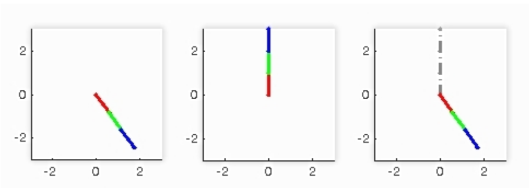
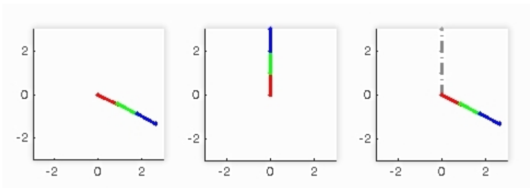


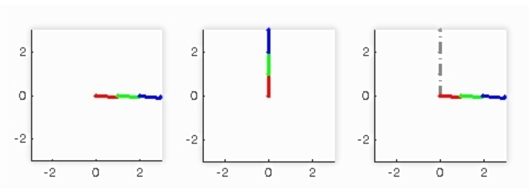
Figure 5.41: Error in velocity variables

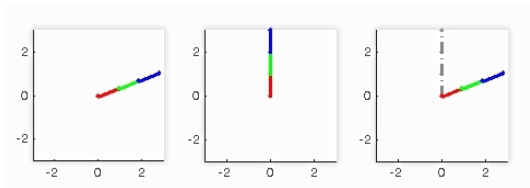

(a) frame 0

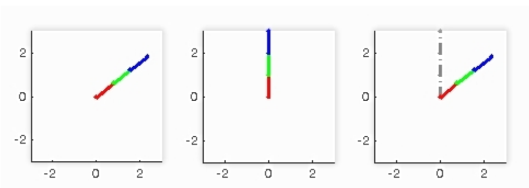

(b) frame 5

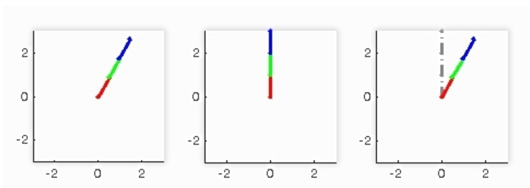

(c) frame 10

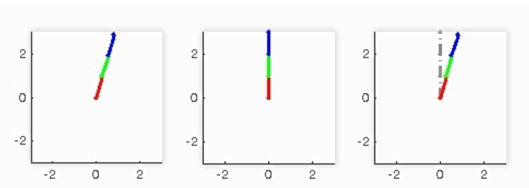

(d) frame 15

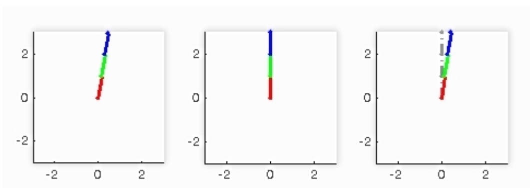

(e) frame 20

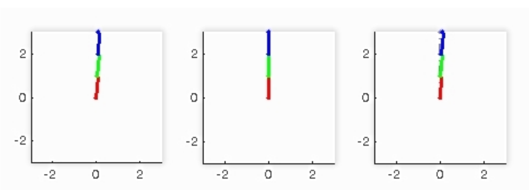

(f) frame 25


(g) frame 30


(h) frame 40


(i) frame 50


(j) frame 60


(k) frame 80

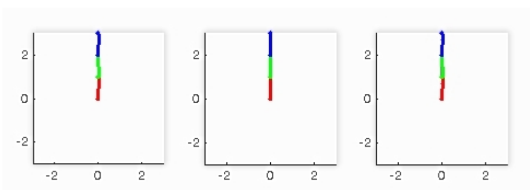

(l) frame 1000

Figure 5.42: Swing-up from downdowndown-position to upupup-position and balancing ($\alpha = 2$)

Initial data $(0,0,0,0,0,0)^T$ and $\alpha = 2$

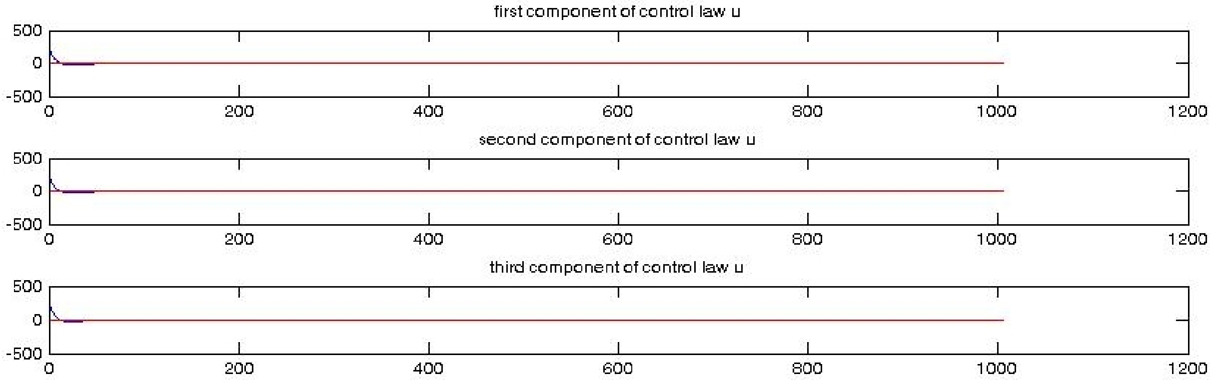


Figure 5.43: Control law u

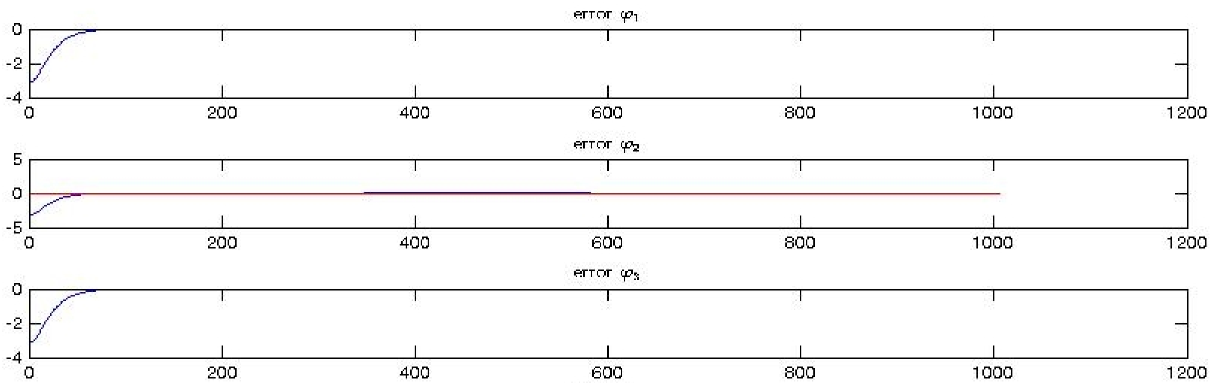


Figure 5.44: Error in state variables u

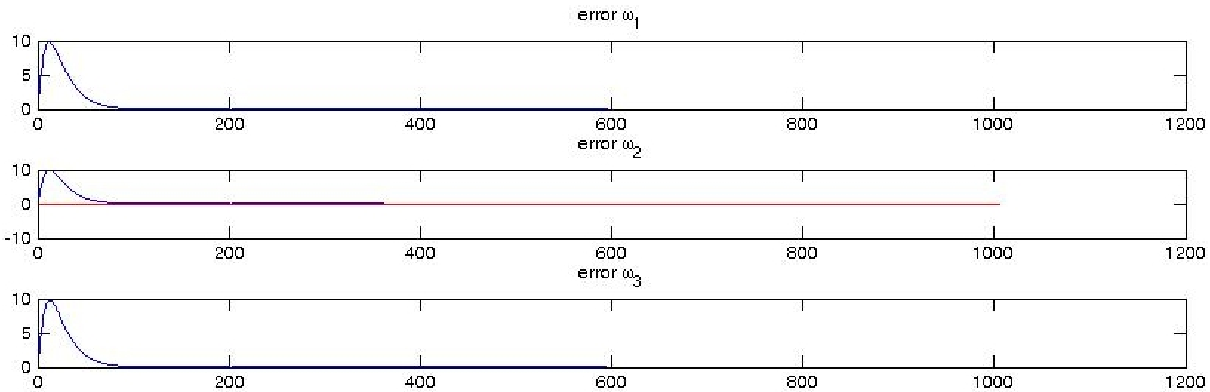
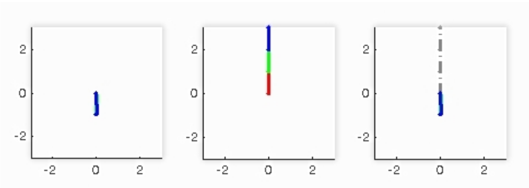
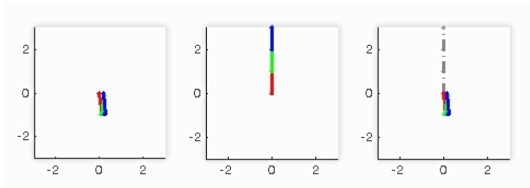


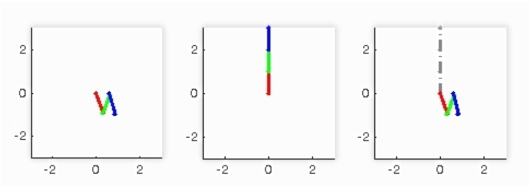
Figure 5.45: Error in velocity variables



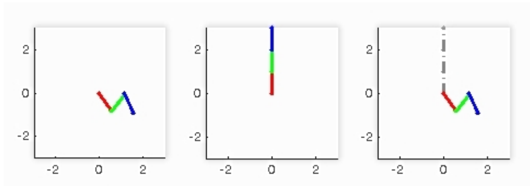
(a) frame 0



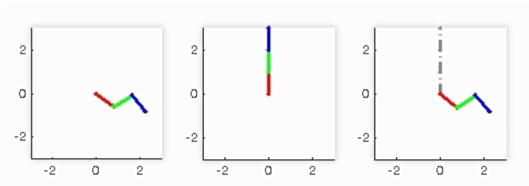
(b) frame 10



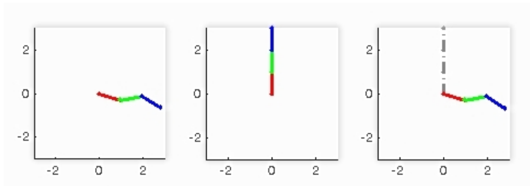
(c) frame 20



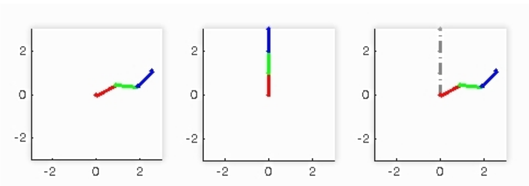
(d) frame 30



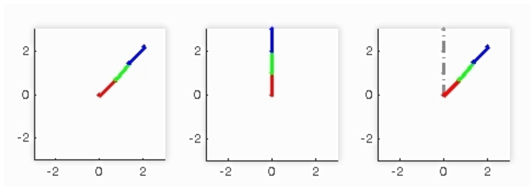
(e) frame 40



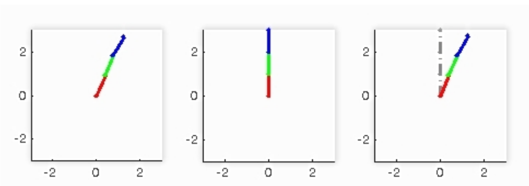
(f) frame 50



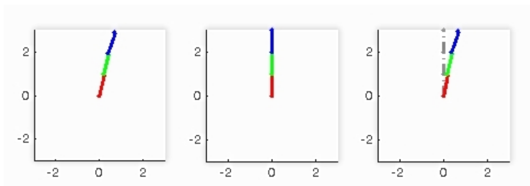
(g) frame 100



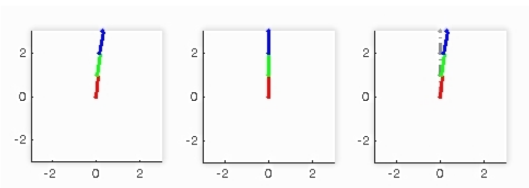
(h) frame 150



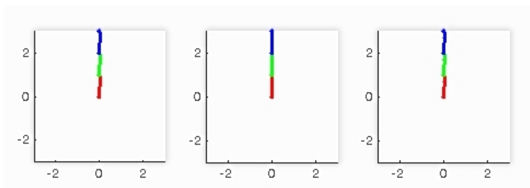
(i) frame 200



(j) frame 250



(k) frame 300



(l) frame 1000

Figure 5.46: Swing-up from downupdown-position to upupup-position and balancing ($\alpha = 0$)

Initial data $(0, \pi, 0, 0, 0, 0)^T$ and $\alpha = 0$

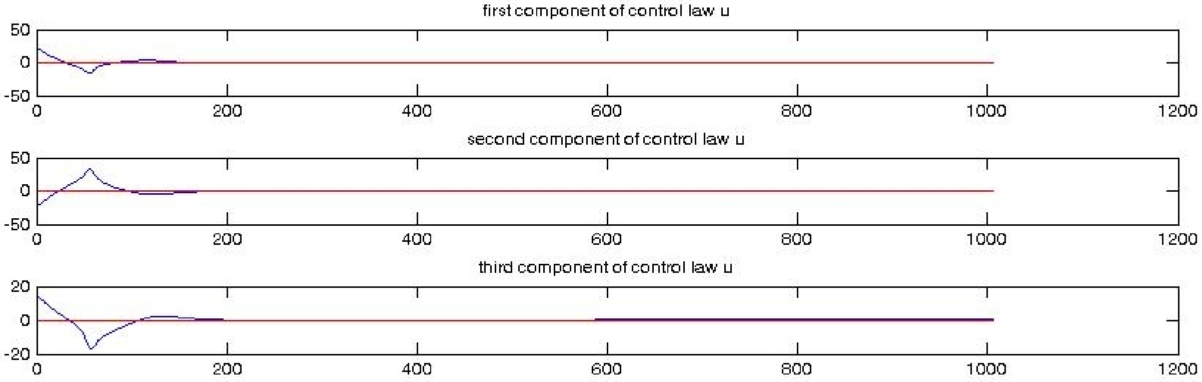


Figure 5.47: Control law u

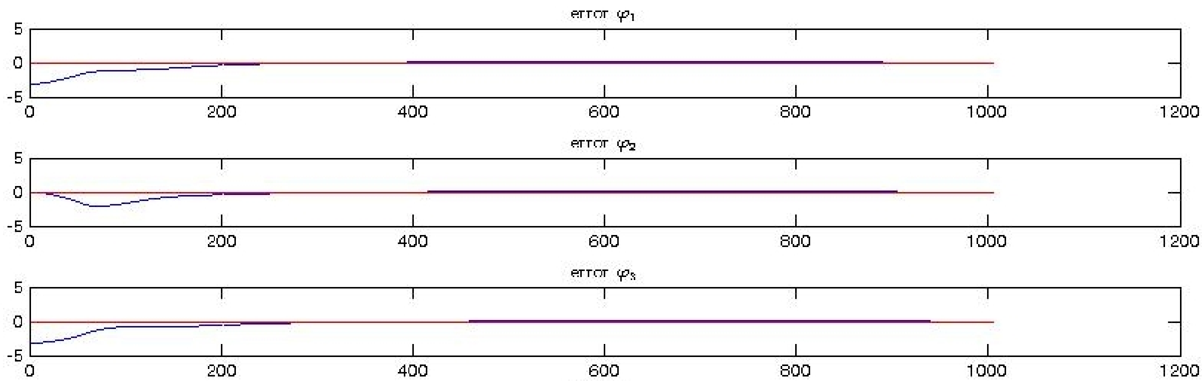


Figure 5.48: Error in state variables u

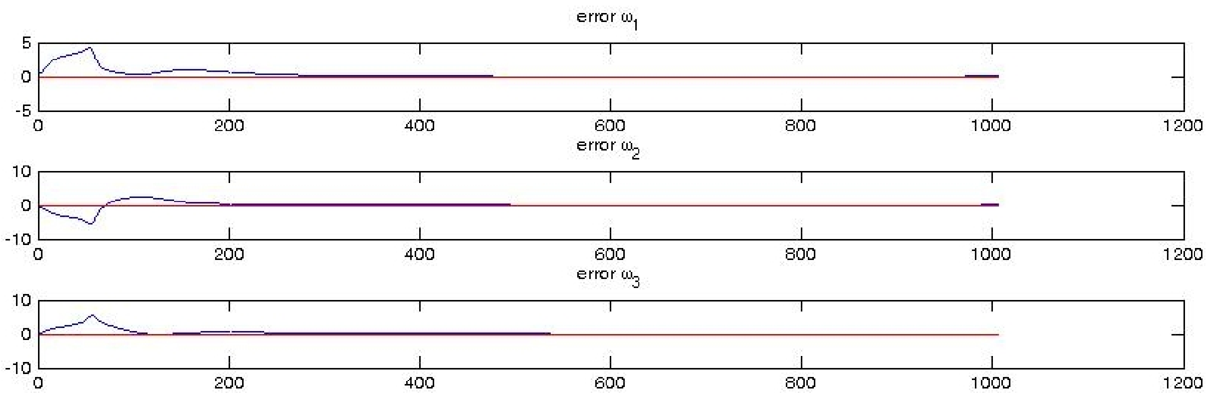
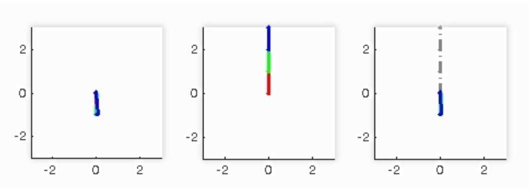
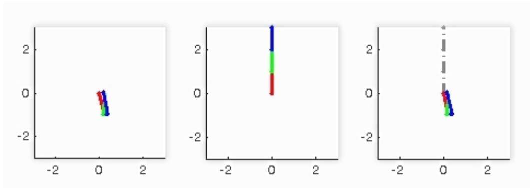


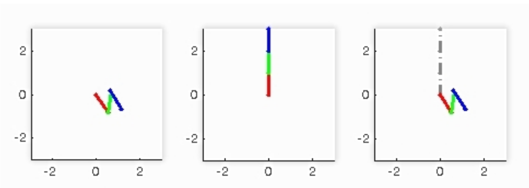
Figure 5.49: Error in velocity variables



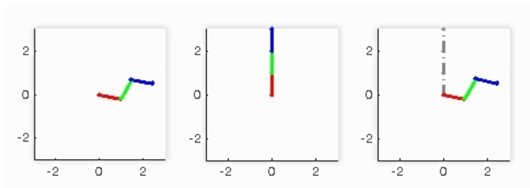
(a) frame 0



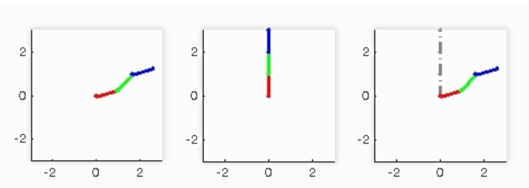
(b) frame 5



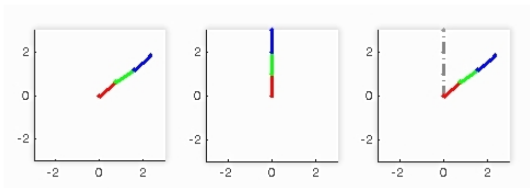
(c) frame 10



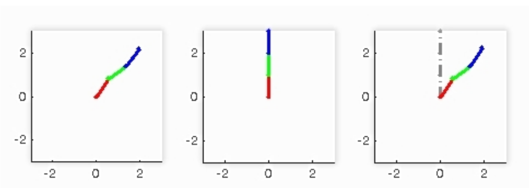
(d) frame 20



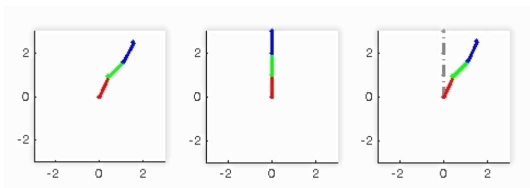
(e) frame 25



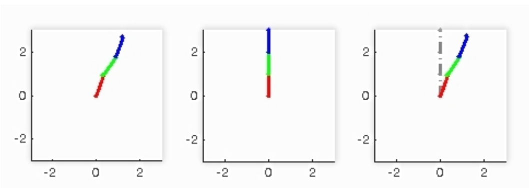
(f) frame 30



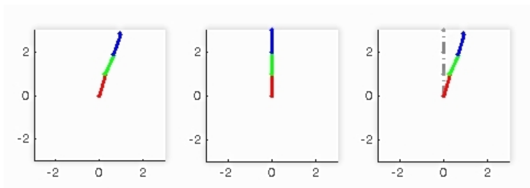
(g) frame 35



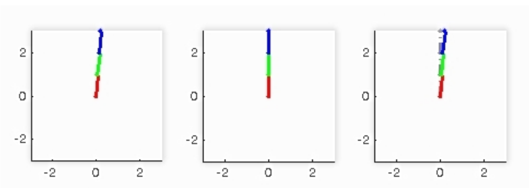
(h) frame 40



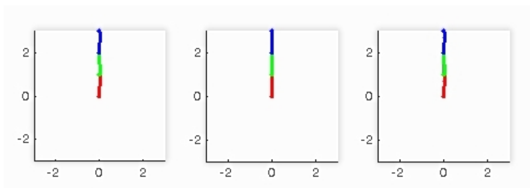
(i) frame 45



(j) frame 50



(k) frame 75



(l) frame 1000

Figure 5.50: Swing-up from downupdown-position to upupup-position and balancing ($\alpha = 2$)

Initial data $(0, \pi, 0, 0, 0, 0)^T$ and $\alpha = 2$

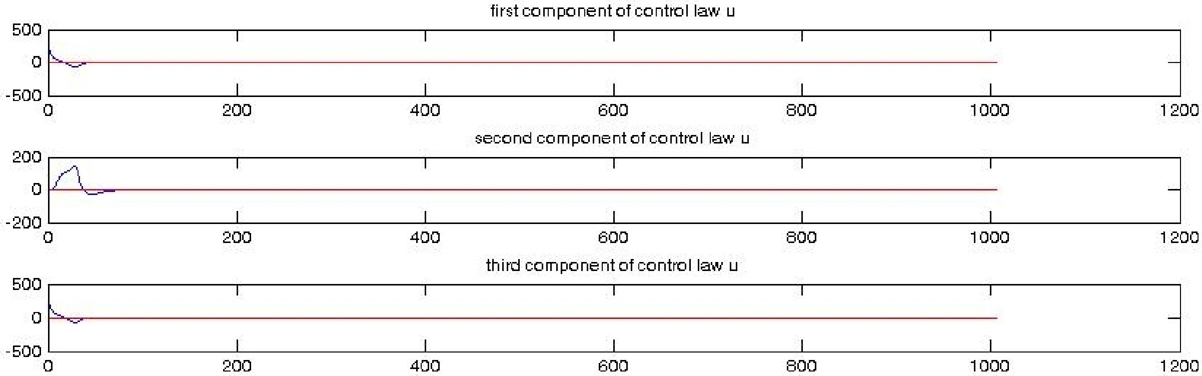


Figure 5.51: Control law u

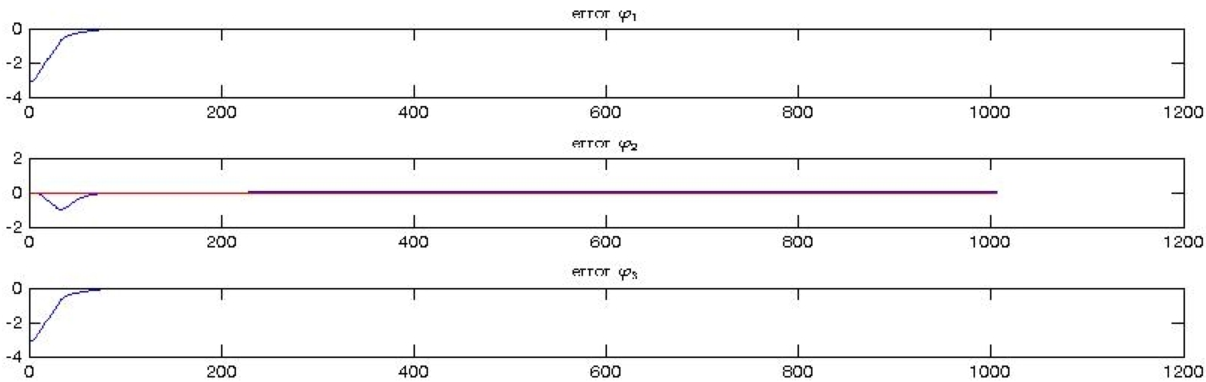


Figure 5.52: Error in state variables u

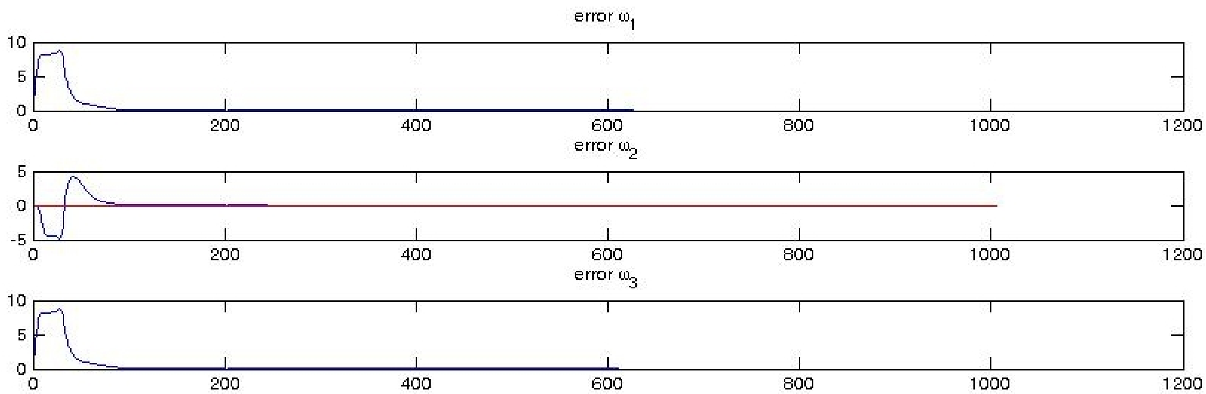


Figure 5.53: Error in velocity variables

The modified control law (4.130) successfully swings up the triple pendulum for these two initial configurations, showing that actually the basin of attraction for the stabilization of this trivial trajectory is pretty large. A higher value of α leads to a faster convergence to the desired state at the cost of higher control inputs.
In the next examples we try to use the natural dynamics of the triple pendulum for the swingup part, meaning we choose an initial configuration where all pendulum links point downward having initial velocities such that the motion of the uncontrolled triple pendulum comes close to the desired equilibrium state (which of course cannot be reached in finite time since it is an equilibrium state). Finding the proper initial velocities can be done for example by a numerical brute force method by randomly guessing initial velocities on the manifold implicitly defined by different potential energy levels for the starting configuration and the desired equilibrium state.

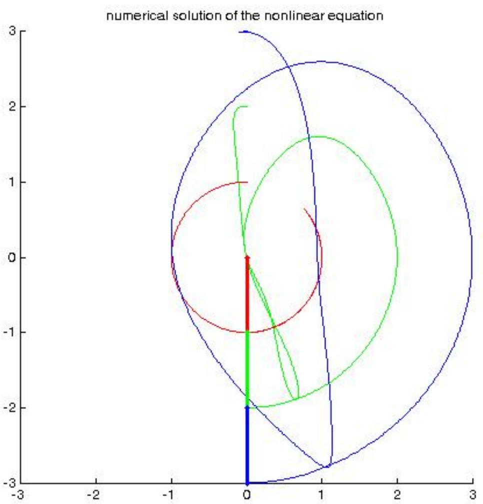


Figure 5.54: Motion of an upswinging triple pendulum without control

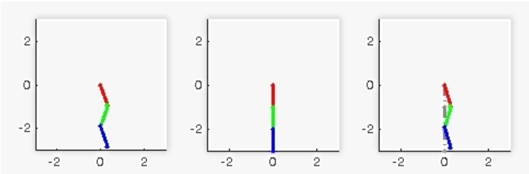
Figure (5.54) shows the motion $\tilde{x}$ of the uncontrolled pendulum with initial velocities $(1.1720, 0.9876, 2.1372)^T$. To reach the almost upright position two "swings" are performed. After the first swing (see frame 300) the triple pendulum is close to the upright position but needs another "dive" before actually becoming really close to it (frame 900).

We will take two different reference trajectories, the first one following the first swing only (until time approximately 3.26) and the second one following both swings until time reaches approximately 9.34.
We compare the energy needed by evaluating the integral

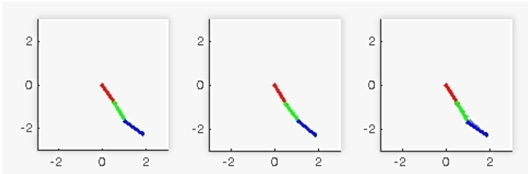
$$\int_0^{12} u(t)^T u(t) dt \tag{5.241}$$

and finally we compare how these solution perform "against" the trivial reference trajectory consisting of the equilibrium point.

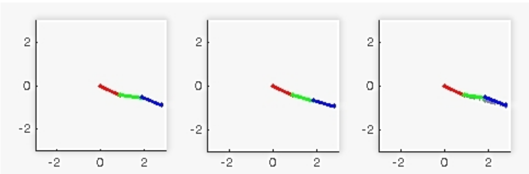
The simulation is ran for 12 time units with $\delta_k = 1 \, \forall k \Longrightarrow t_k = k$ for $k = 0, \ldots, 11$ such that the matrix $\tilde{H}_k^{-1}$ is updated 12 times for each simulation. We chose δ_k smaller than in the simulations before since the reference trajectory is more complicated in this case (although it would even work with $\delta_k = 2$).



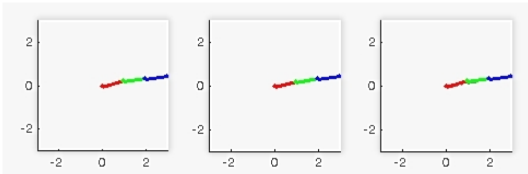
(a) frame 0



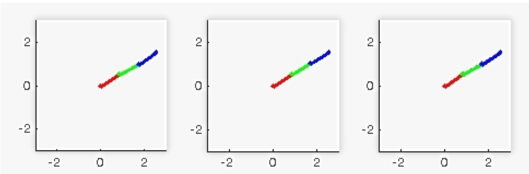
(b) frame 50



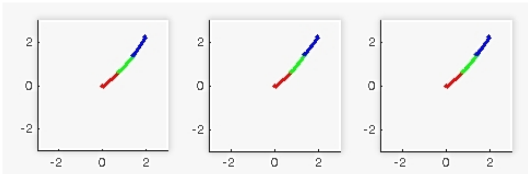
(c) frame 100



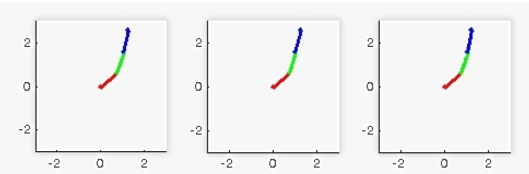
(d) frame 150



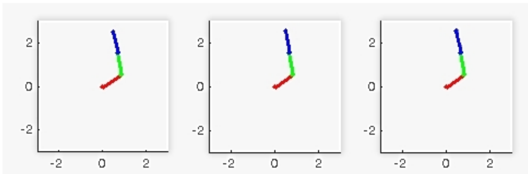
(e) frame 200



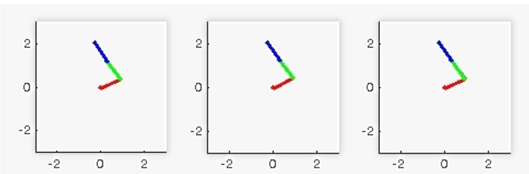
(f) frame 250



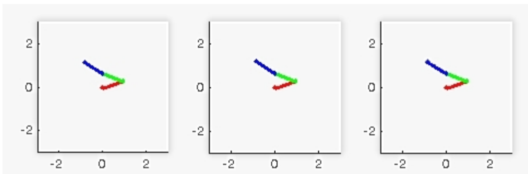
(g) frame 300



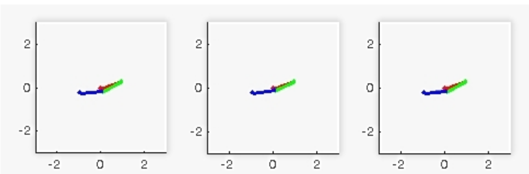
(h) frame 350



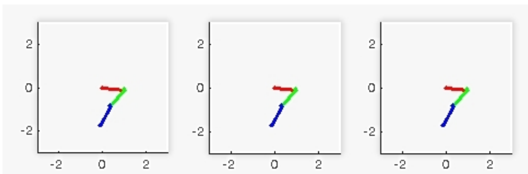
(i) frame 400



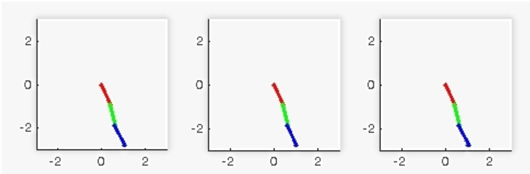
(j) frame 450

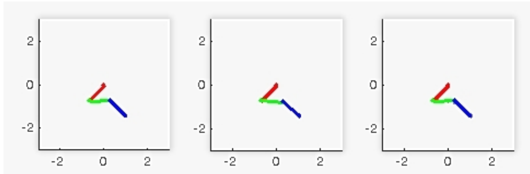


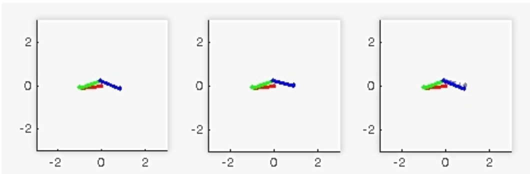
(k) frame 500

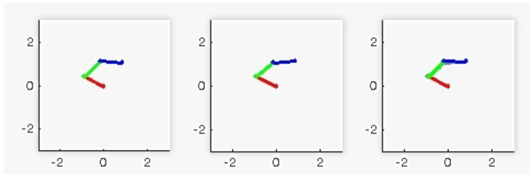


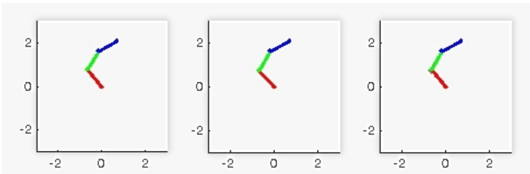
(l) frame 550

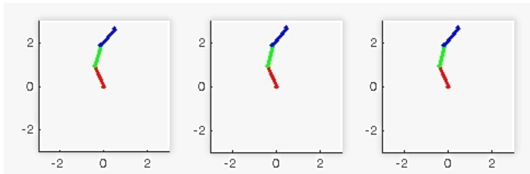

(a) frame 600

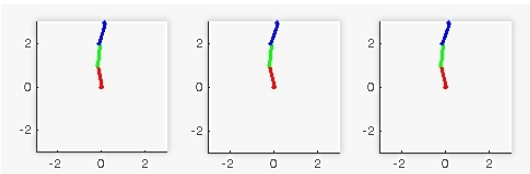

(b) frame 650

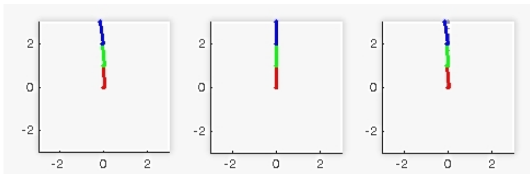

(c) frame 700

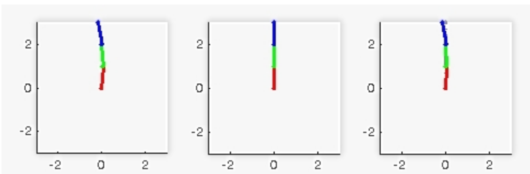

(d) frame 750

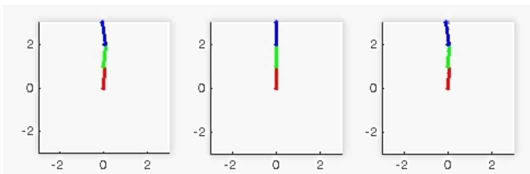

(e) frame 800

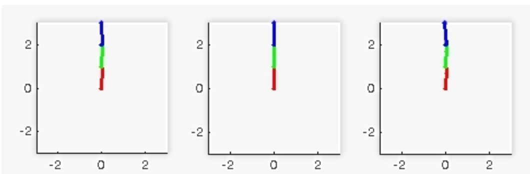

(f) frame 850


(g) frame 900


(h) frame 950


(i) frame 1000


(j) frame 1050


(k) frame 1100

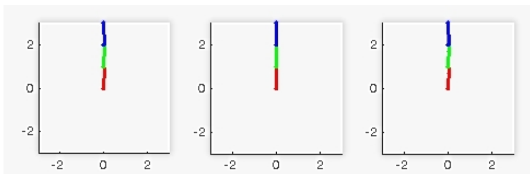

(l) frame 1200

Figure 5.55: Swing-up to upupup-position and balancing ($\alpha = 0$) along trajectory belonging to initial data $(0, 0, 0, 1.1720, 0.9876, 2.1372)^T$

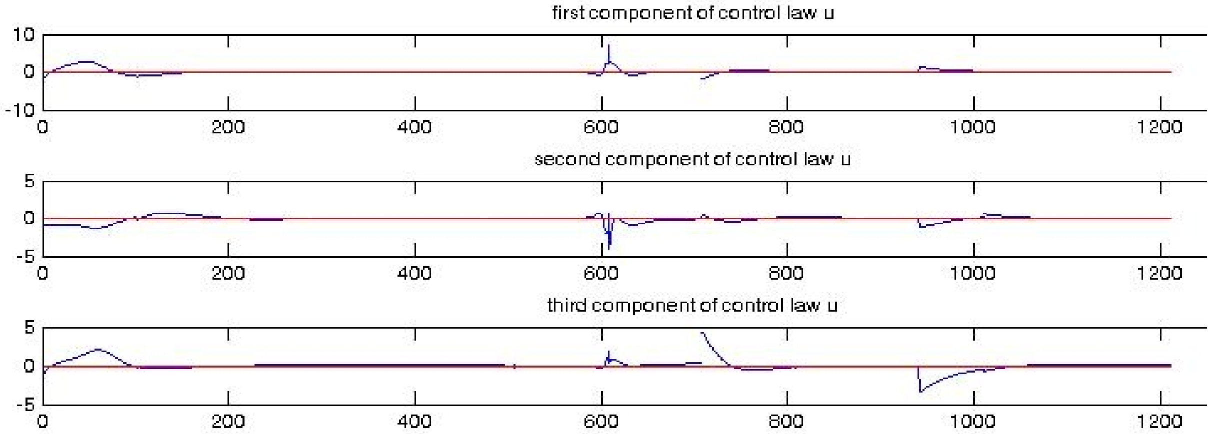


Figure 5.56: Control law u

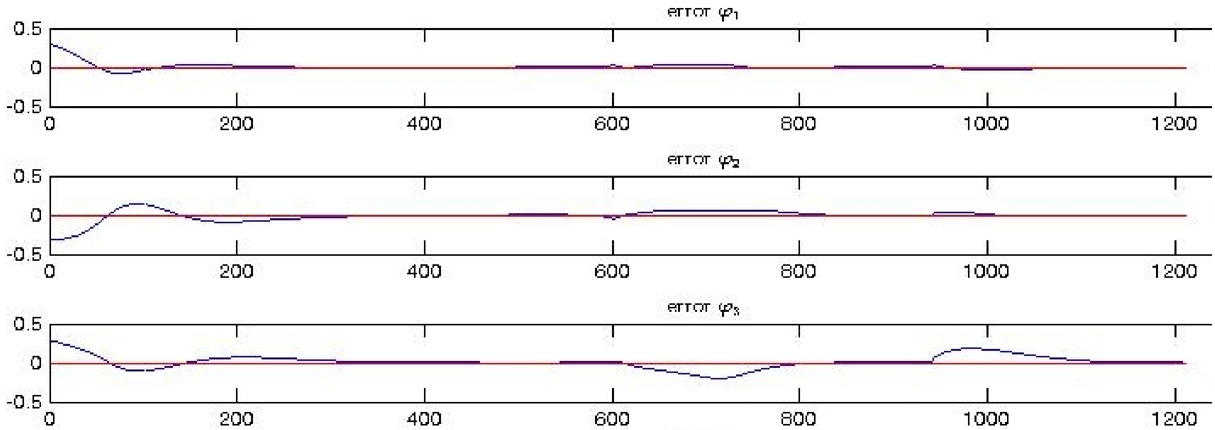


Figure 5.57: Error in state variables u

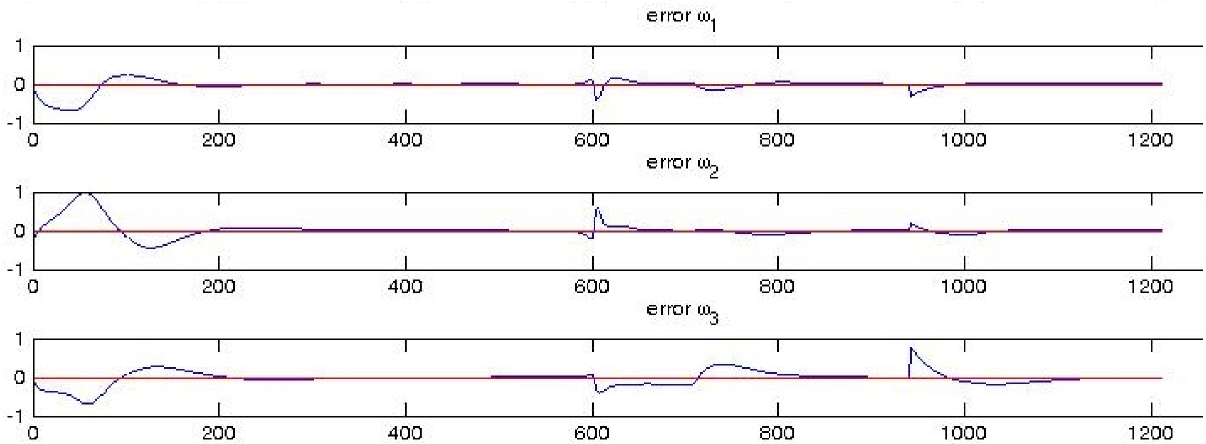
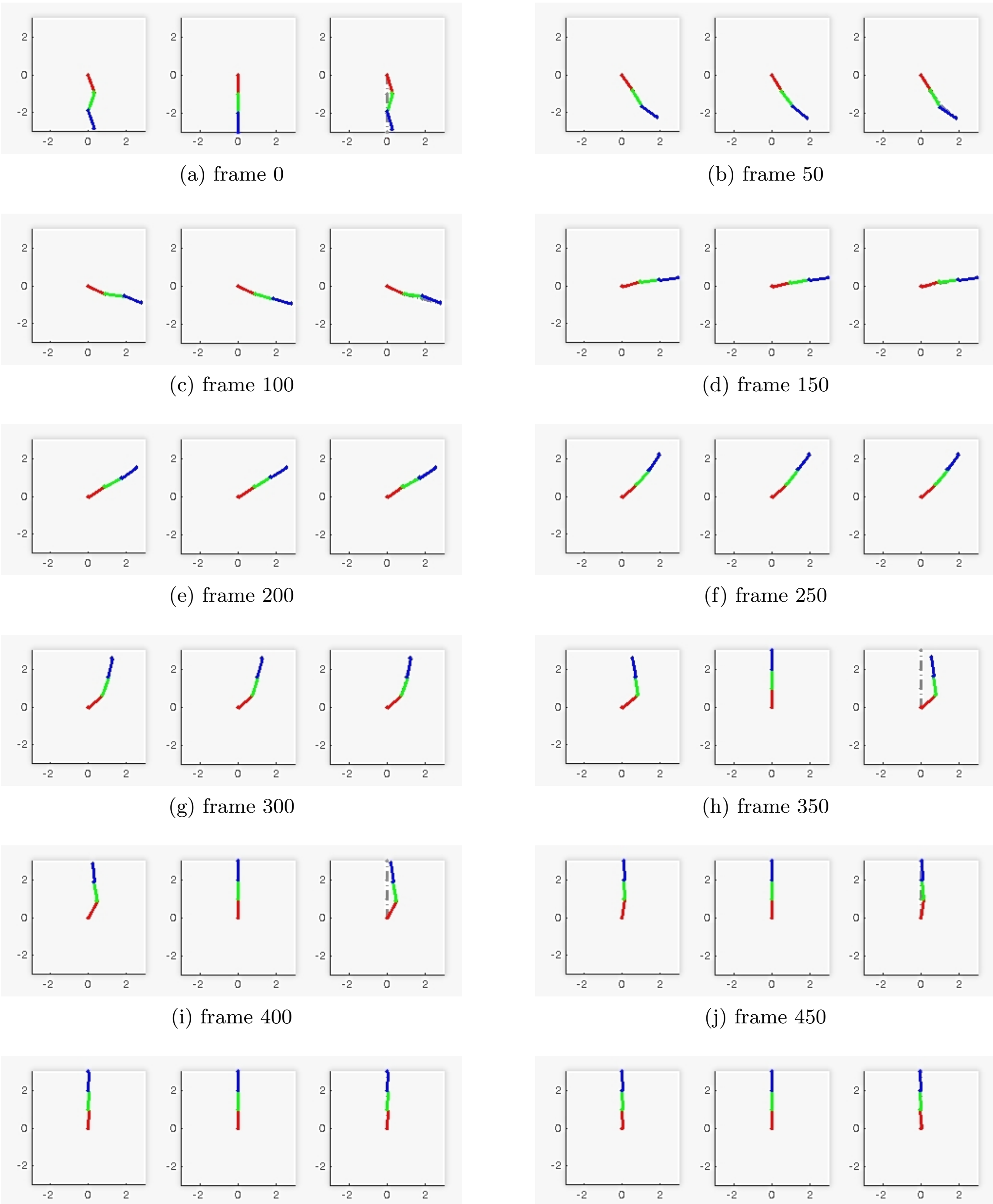


Figure 5.58: Error in velocity variables

Initial data $(0.3, -0.3, 0.3, 1, 0.8, 2)^T$ and $\alpha = 0$



(a) frame 0 (b) frame 50 (c) frame 100 (d) frame 150 (e) frame 200 (f) frame 250 (g) frame 300 (h) frame 350 (i) frame 400 (j) frame 450 (k) frame 500 (l) frame 1200

Figure 5.59: Swing-up to upupup-position and balancing by following a solution of the uncontrolled system until frame 324; initial data $(0, 0, 0, 1.1720, 0.9876, 2.1372)^T$

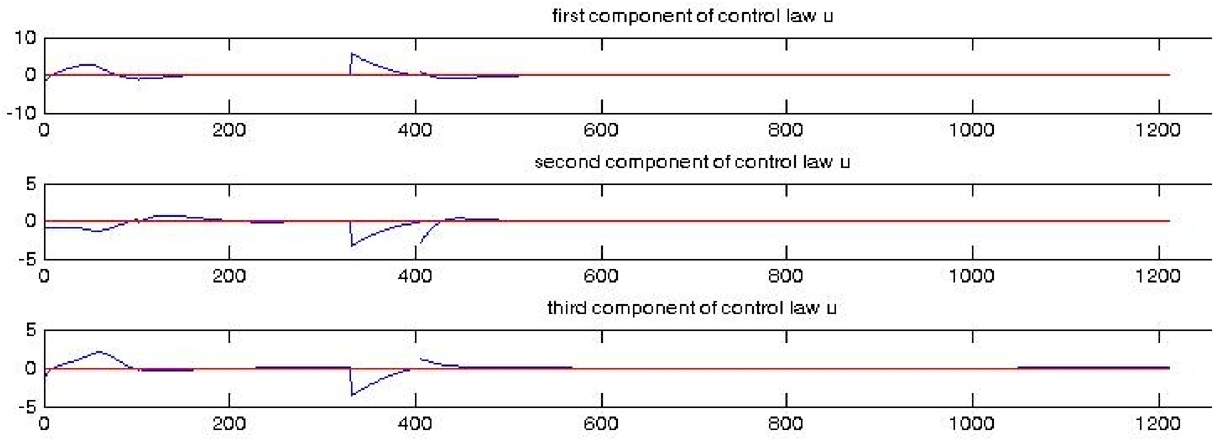


Figure 5.60: Control law u

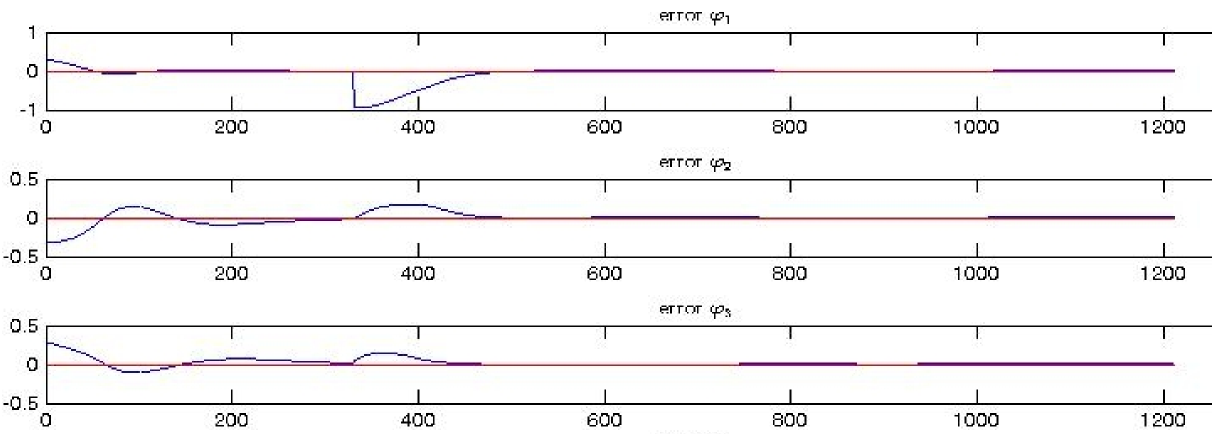


Figure 5.61: Error in state variables u

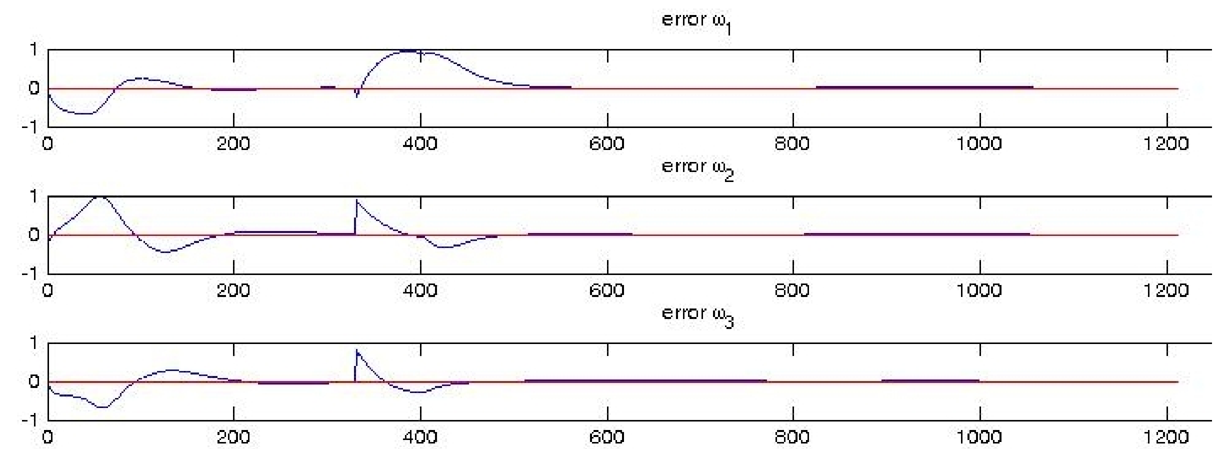


Figure 5.62: Error in velocity variables

We want to compare these two methods with the method by choosing as reference trajectory simply the equilibrium point $(\pi, \pi, \pi, 0, 0, 0)^T$. The motion of the controlled triple pendulum with the same initial condition as with simulations run before is shown in figure 5.63 where it can be seen that it moves up like a simple pendulum more or less. The components of the control input

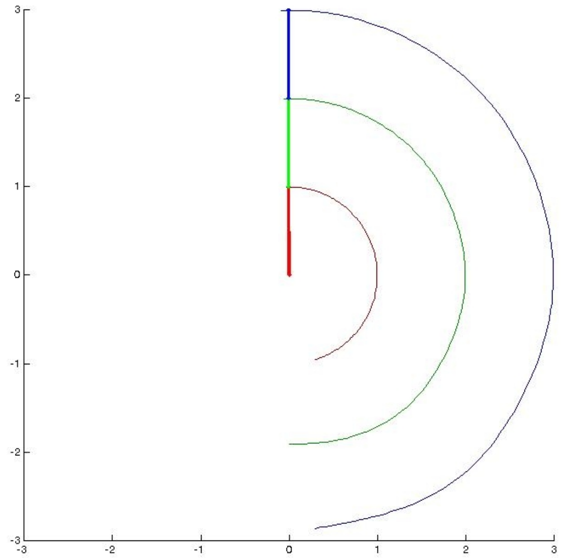


Figure 5.63: motion of controlled triple pendulum to the inverted position

u are shown in 5.64 showing that neglecting the natural dynamics of the triple pendulum leads to high input gains due to the feedback term $(x(t) - x^0(t))$ in the control law where $x^0(t)$ is the feedback term. This term is small when $x^0(t)$ is a solution of the uncontrolled system starting near the initial configuration of the controlled problem, why it is large, when taking as reference trajectory the inverted pendulum and the initial configuration of the controlled problem is not close to it. Figure (5.64) reflects the latter situation while (5.56), (5.60) show that only small input gains are necessary in the first case.
To compare the energy consumption of these methods we used the cost functional (5.241)

$$\int_{\text{simulation time}} u^T(t)u(t)dt \tag{5.242}$$

which not only measures the consumed energy but also penalizes for high values of the control input u. The results for the three different kinds of motion are given in table (5.1).
We recall that $\tilde{x}$ is the solution trajectory of the uncontrolled pendulum with initial data $(0, 0, 0, 1.1720, 0.9876, 2.1372)^T$.

reference trajectory	$x^* = (\pi, \pi, \pi, 0, 0, 0)^T$	$\tilde{x}$ until 3.26 then x^*	$\tilde{x}$ until 9.34 then x^*
energy consumption	170.36	16.67	11.91

Table 5.1: Energy consumption for $x_0 = (0.3, -0.3, 0.3, 1, 0.8, 2)^T$

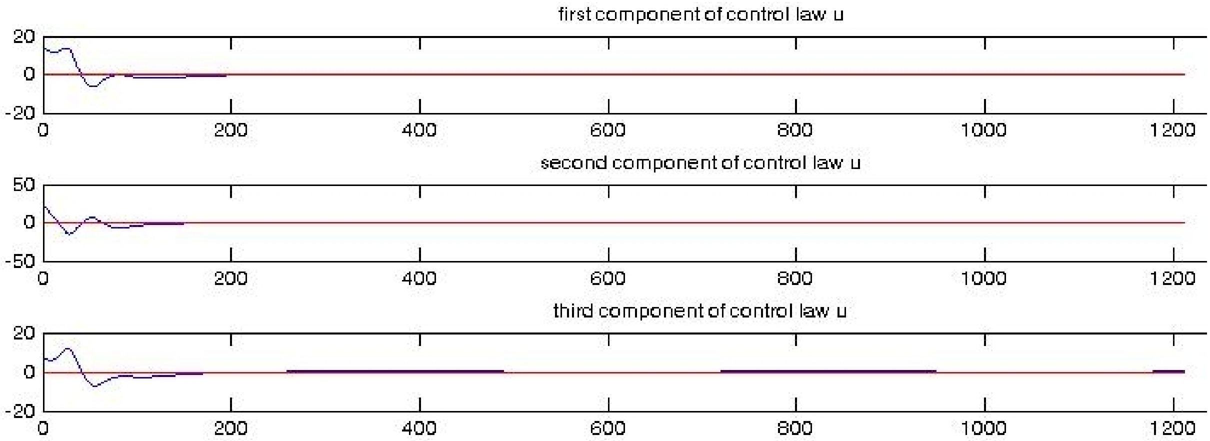


Figure 5.64: control input without using the "natural dynamics"

Using the natural dynamics of the triple pendulum can significantly reduce the amount of energy needed to swing it up. Even if we only partially use a solution of the uncontrolled system we can get pretty good results: using only the first of the two swings we can swing up the triple pendulum much more faster for only moderate higher costs. It is a "feature" of this special solution trajectory that it exhibits the possibility to "get off earlier". In this example we started with an initial value near the initial value for reference trajectory. If we start in the stable equilibrium point we obtain

reference trajectory	$x^* = (\pi, \pi, \pi, 0, 0, 0)^T$	$\tilde{x}$ until 3.26 then x^*	$\tilde{x}$ until 9.34 then x^*
energy consumption	155.74	37.94	33.15

Table 5.2: Energy consumption for $x_0 = (0, 0, 0, 0, 0, 0)^T$

Appendix A

The Bang-Bang Principle

A motivation for the bang-bang-principle is the following scenario. Whenever one tries to change a systems state with limited resources as fast as possible it seems to be useful to use as much of the available resources as possible at every time. For many applications this is certainly true, for example heating water until it cooks or bringing a car from A to B on a straight road. But there are also examples of optimal controls, which are not bang-bang. Nevertheless whenever there is a suitable admissible control, then there is also a suitable admissible bang-bang control. Although the bang-bang-principle can be applied more generally, we will only show it using the example of the linearized pendulum. For a general theory cf. e.g. [LaSalle, 1960]. We show that whenever there is a control $u(\tau) \in \Omega$ solving (2.10), then there is a bang-bang control $\hat{u}(\tau) \in \hat{\Omega}$ solving (2.10).
We define the sets

$$\Omega := \{u, |u| \leq 1\}, \qquad \hat{\Omega} := \{u, |u| = 1\}. \tag{A.1}$$

We restrict ourselves to measurable controls u in the interval $[0, t_{\max}]$. So $\hat{\Omega} \subset \Omega$ is the set of admissible bang-bang-controls. Using the equivalence between problem (2.10) and problem (2.13) it suffices to show that the following sets are equal

$$\mathcal{R}(t) = \{\int_0^t Y(\tau)u(\tau)d\tau,\ u \in \Omega\}, \quad \hat{\mathcal{R}}(t) := \{\int_0^t Y(\tau)\hat{u}(\tau)d\tau,\ \hat{u} \in \hat{\Omega}\}. \tag{A.2}$$

The new introduced set $\hat{\mathcal{R}}(t)$ denotes the reachability set under the restriction that only admissible bang-bang controls are used. Because of the trivial relation $\hat{\Omega} \subsetneq \Omega$ it is clear that $\hat{\mathcal{R}}(t)$ is included in the reachability set $\mathcal{R}(t)$. To show that we actually have set equivalence we have to show the other inclusion $\mathcal{R}(t) \subset \hat{\mathcal{R}}(t)$.

Let E be a measurable subset of the interval $[0, t]$, where $t_{\max} \geq t > 0$ is arbitrary but fixed. In addition, let $\mathfrak{E}$ be the Borel-σ-algebra in $[0, t]$. Furthermore $Y(\tau) = \Phi^{-1}(\tau, 0)B \in \mathbb{R}^2$ an integrable vector valued function defined on $[0, t]$. Then

$$\mu : \mathfrak{E} \to \mathbb{R}^2, \qquad \mu(E) \mapsto \int_E Y(\tau)d\tau$$

defines a vector measure. We denote the range of this vector measure as $R_\mu(t)$. Due to a theorem of Lyapunov [Lyapunov, 1940], [Halmos, 1948] the range of every countable additive and finite vector measure which maps into a finite dimensional euclidean space is closed and convex. Therefore $R_\mu(t)$ is closed and convex. We will see that closedness and convexity are also properties of the set $\hat{\mathcal{R}}(t)$:

For every element $\hat{u} \in \hat{\Omega}$ there is a set $E \in \mathfrak{E}$ so that $\hat{u}$ can be represented as $2\chi_E - 1$ where χ_E is the characteristic function of E. Such a set E always exists because $\hat{u}$ was supposed to be measurable. For example one could choose the set $\{\hat{u} = 1\}$ which belongs to $\mathfrak{E}$.
Multiplication with $Y(\tau)$ and integration from 0 to t leads to

$$\begin{aligned}\hat{\mathcal{R}}(t) \ni \int_0^t \hat{u}(\tau)Y(\tau)d\tau &= \int_0^t 2\chi_E(\tau)Y(\tau)d\tau - \int_0^t Y(\tau)d\tau \\ &= 2\int_E Y(\tau)d\tau - \int_0^t Y(\tau)d\tau \\ &= 2\mu(E) - \int_0^t Y(\tau)d\tau.\end{aligned}$$

With $\bar{Y}(t) := \int_0^t Y(\tau)d\tau$ we have for arbitrary, but fixed t:

$$\hat{\mathcal{R}}(t) = 2R_\mu(t) - \bar{Y}(t). \tag{A.3}$$

It follows from the equality of these sets that $\hat{\mathcal{R}}(t)$ is closed and convex.

We now show the desired inclusion $\mathcal{R}(t) \subset \hat{\mathcal{R}}(t)$. For every element $z(t) \in \mathcal{R}(t)$ we will construct a sequence $\{z_m(t)\}_{m\in\mathbb{N}} \subset \hat{\mathcal{R}}(t)$ converging to $z(t)$.

Let $z(t) \in \mathcal{R}(t)$ be arbitrary. We can represent $z(t)$ as

$$z(t) = \int_0^t Y(\tau)u(\tau)d\tau, \qquad u \in \Omega. \tag{A.4}$$

Now define $\beta(\tau) := \frac{1}{2}(u(\tau)+1)$ and $\bar{z}(t) := \frac{1}{2}(z(t)+\bar{Y}(t))$. The auxiliary function $\beta(\tau)$ assumes only values lying in the interval $[0,1]$. It is chosen so that the following equality holds

$$\bar{z}(t) = \int_0^t Y(\tau)\beta(\tau)d\tau. \tag{A.5}$$

We now construct a sequence converging to $\bar{z}(t)$ which will be used to construct the desired sequence converging to $z(t)$. With

$$E_j(t) := \left\{\tau \mid \frac{j-1}{m}t \le \beta(\tau) \le \frac{j}{m}t\right\} \quad \forall j \in \{1..m\} \tag{A.6}$$

we define

$$\bar{z}_m(t) := \sum_{j=0}^{m} \frac{j}{m}\int_{E_j} Y(\tau)d\tau, \quad m \in \mathbb{N}$$

converging to $\bar{z}(t)$:

$$\begin{aligned}\lim_{m\to\infty} |\bar{z}(t) - \bar{z}_m(t)| &\le \lim_{m\to\infty} \sum_{j=1}^{m} \int_{E_j} \left(\frac{j}{m} - \beta(\tau)\right) |Y(\tau)|d\tau \\ &\le \lim_{m\to\infty} \frac{1}{m} \sum_{j=0}^{m} \int_{E_j} |Y(\tau)|d\tau = \lim_{m\to\infty} \frac{1}{m} \int_0^t |Y(\tau)|d\tau = 0,\end{aligned}$$

since $Y(\tau)$ is integrable on $[0,t]$. When $\bar{z}_m(t)$ converges to $\bar{z}(t)$ then $z_m(t) := 2\bar{z}_m(t) - \bar{Y}(t)$ converges to $z(t)$. It remains to show that $z_m(t)$ actually lie in $\hat{\mathcal{R}}(t)$. As $z_m(t)$ were defined with the help of $\bar{z}_m(t)$, it suffices to show that $\bar{z}_m(t) \in R_\mu(t)$. To see this we define

$$F_j(t) := \bigcup_{i=j}^{m} E_i$$

and because of $\bigcup_{j=1}^m F_j = \bigcup_{j=1}^m jE_j$ we obtain

$$\bar{z}_m(t) = \sum_{j=1}^{m} \frac{j}{m} \int_{E_j} Y(\tau)d\tau = \frac{1}{m} \sum_{j=1}^{m} \int_{F_j} Y(\tau)d\tau \tag{A.7}$$

which is a convex combination of elements of $R_\mu(t)$ and therefore $\bar{z}(t) \in R_\mu(t)$. With (A.3) we see that the limit $z(t)$ of $\{z_m(t)\}$ has to lie in $\hat{\mathcal{R}}(t)$ since all terms of $\{z_m(t)\}$ lie in this (closed) set. Since $z(t)$ was an arbitrary element of $\mathcal{R}(t)$ this means that we have verified

$$\mathcal{R}(t) \subset \hat{\mathcal{R}}(t) \tag{A.8}$$

and together with the trivial inclusion $\hat{\mathcal{R}}(t) \subset \mathcal{R}(t)$ the two sets are equal, which concludes the proof.

Appendix B

An iterative procedure to determine the time-optimal control for linear systems

We consider the linear time-varying control *normal* system

$$\frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t) \tag{B.1}$$

with initial data $x(0) = x_0$, where x is a n-vector, A a $n \times n$-matrix, B a $n \times r$-matrix both with continuous entries and u is a r-dimensional vector-valued control function, where $r \leq n$. The set of admissible controls is a compact nonempty subset U of $\mathbb{R}^r$. Eaton proposes a numerical algorithm to determine this time-optimal control u^* (cf. [Eaton, 1962]). Let z be a n-dimensional continuous vector-valued function depending on t. Suppose there is a control function $\tilde{u}$ such that $z(\tilde{t}) = x(\tilde{t}, \tilde{u}(\tilde{t}))$ for a finite time $\tilde{t}$. Then, there exists a finite time $t^* \leq \tilde{t}$ and a control function u^* such that $z(t^*) = x(t^*, u^*(t^*))$. In this case we have

$$\Phi^{-1}(t^*, 0)z(t^*) - x_0 = \int_0^{t^*} Y(\tau)u^*(\tau)d\tau. \tag{B.2}$$

As shown in the introduction (cf. [LaSalle, 1960, Hermes and LaSalle, 1969]) u^* is essentially unique. For the control function u^* and a normal $\eta^* \in \mathbb{R}^n$ we must have:

$$\left\langle \int_0^{t^*} Y(\tau)u^*(\tau)d\tau, \eta^* \right\rangle = \max_{u \in U} \left\langle \int_0^{t^*} Y(\tau)u(\tau)d\tau, \eta^* \right\rangle = \max_{u \in U} \int_0^{t^*} (\eta^{*T} Y(\tau)u(\tau))d\tau. \tag{B.3}$$

which is maximal for

$$u^*(\tau) = \operatorname{sgn} \eta^{*T} Y(\tau). \tag{B.4}$$

The algorithm of Eaton approximates the normal η^* and therefore determines the optimal control u^* which gives the optimal time t^*. The main idea is to construct a sequence of support planes to the boundary of the reachability set at time t_m which contains the trajectory point of the moving target at time t_m. The normal vector to the boundary point of $\mathcal{R}(t_m)$ is given by η_m. This normal vector is used to construct the next support plane at time t_{m+1}. The sequence $\{t_m\}_{m \in \mathbb{N}}$ will be a to t^* convergent nondecreasing sequence such that the sequence $\{\eta_m\}_{m \in \mathbb{N}}$ will converge to η^*.

The left hand side of (B.2) can be understood as a continuous time dependent function $g(t)$. Since optimal controls are of the form

$$u(t) = \operatorname{sgn}\left(\eta^T Y(\tau)\right), \qquad \eta \in \mathbb{R}^n \text{ with } \|\eta\| = 1$$

one can define for these control functions the right hand side of (B.2) as function v depending continuously on t and on η:

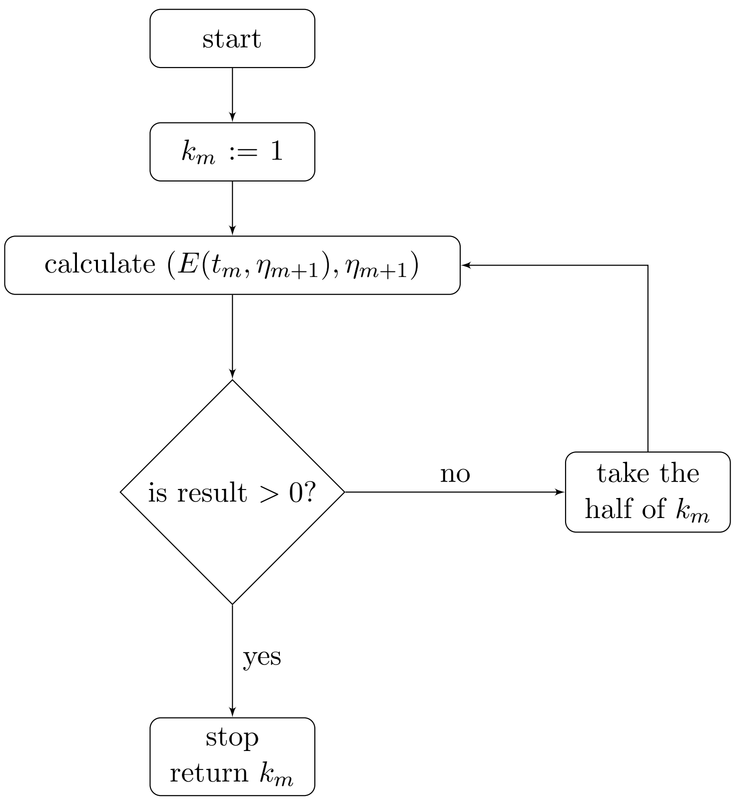
$$g(t) := \Phi^{-1}(t,0)z(t) - x_0, \qquad v(t,\eta) := \int_0^t Y(\tau)u(\tau)d\tau = \int_0^t Y(\tau)\operatorname{sgn}(\eta^T Y(\tau))d\tau. \tag{B.5}$$

The difference between left and right hand side defines a function $E(t,\eta) := g(t) - v(t,\eta)$ which vanishes in t^*. The reachability set $\mathcal{R}_v(t)$ of v is continuous, strictly monotonically increasing and convex. In addition this set is closed and its boundary moves continuously in time.
For the moment η is an arbitrary normed vector in $\mathbb{R}^n$. Since $u(\tau) = \operatorname{sgn}(\eta^T Y(\tau))$ maximizes $\eta^T v(t,\eta)$ we know that $v(t,\eta)$ is a boundary point of $\mathcal{R}_v(t)$ with η an outward normal to a support hyperplane through $v(t,\eta)$. With (B.1) being *normal* we also know, that the reachability set $\mathcal{R}_v(t)$ and the support hyperplane have only this point $v(t,\eta)$ in common. Our initial guess for η_1 has to be chosen such that the support plane to $\mathcal{R}(t_1)$ contains $g(t_1)$. We omit the details and just give the iteration rule.

The algorithm: The construction rule is given by

$$\eta_{m+1} = \frac{\eta_m + k_m E(t_m,\eta_m)}{\|\eta_m + k_m E(t_m,\eta_m)\|}, \; m \geq 1 \tag{B.6}$$

where k_m can be determined as follows:



Appendix C

The multi-pendulum with n links

C.1 A Lagrangian for the n-pendulum

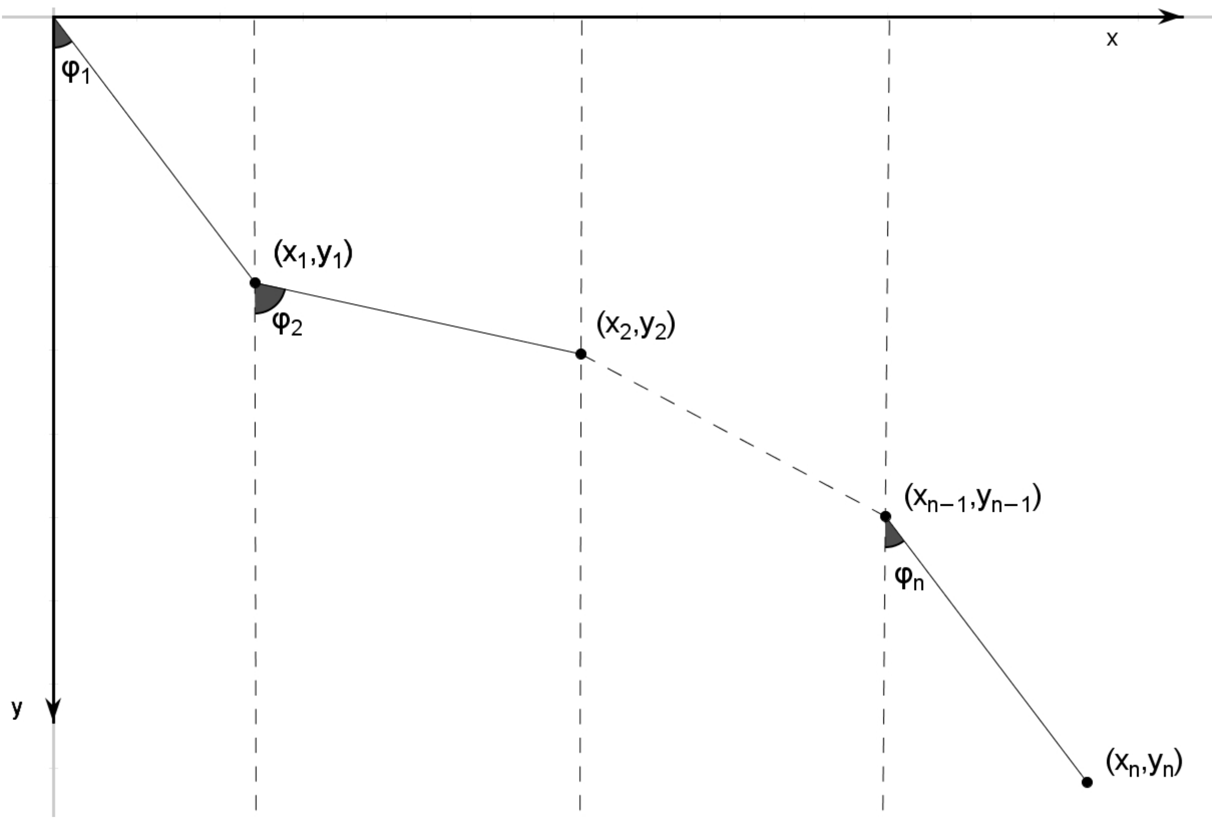


Figure C.1: n-pendulum

For the i-th pendulum the position is given by the coordinates

$$x_i = \sum_{j=1}^{i} \sin \varphi_j, \qquad y_i = \sum_{j=1}^{i} \cos \varphi_j. \tag{C.1}$$

The velocity of the i-th pendulum is given by

$$v_i^2 = \sum_{j=1}^{i} \dot{\varphi}_j^2 + \sum_{j=1}^{i-1} \sum_{k=j+1}^{i} 2\dot{\varphi}_j \dot{\varphi}_k \cos(\varphi_j - \varphi_k). \tag{C.2}$$

The kinetic energy E_{kin} and potential energy E_{pot} are given by:

$$E_{\text{kin}} = \frac{1}{2} \sum_{j=1}^{n} v_j^2, \qquad E_{\text{pot}} = -\sum_{j=1}^{n} (n+1-j) \cos \varphi_j. \tag{C.3}$$

The Lagrangian for the n-pendulum can be explicitly expressed as

$$\begin{aligned} L_n(\varphi_1, \dots, \varphi_n, \dot{\varphi}_1, \dots, \dot{\varphi}_n) =& \frac{1}{2} \sum_{i=1}^{n} (n+1-i) \left(\dot{\varphi}_i^2 + 2\cos \varphi_i \right) + \\ &+ \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (n+1-j) \dot{\varphi}_i \dot{\varphi}_j \cos(\varphi_i - \varphi_j) \end{aligned} \tag{C.4}$$

Proof. by induction over n:
For $n = 1$ (simple pendulum) we obtain

$$L_1(\varphi_1, \dot{\varphi}_1) = \frac{1}{2} \dot{\varphi}_1^2 + \cos \varphi_1. \tag{C.5}$$

in accordance to (5.7).
Let us assume that formula (C.4) holds for the n-pendulum and let us show that the Lagrangian of the $(n+1)$-pendulum can be expressed with formula (C.4).

$$L_{n+1} = L_n + \frac{1}{2} v_{n+1}^2 + \sum_{i=1}^{n+1} \cos \varphi_i \tag{C.6}$$

$$\begin{aligned} =& \frac{1}{2} \sum_{i=1}^{n} (n+1-i) \left(\dot{\varphi}_i^2 + 2\cos \varphi_i \right) \\ &+ \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (n+1-j) \dot{\varphi}_i \dot{\varphi}_j \cos(\varphi_i - \varphi_j) + \\ =& \frac{1}{2} \sum_{i=1}^{n+1} \dot{\varphi}_i^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n+1} \dot{\varphi}_i \dot{\varphi}_j \cos(\varphi_i - \varphi_j) + \sum_{i=1}^{n+1} \cos \varphi_i \end{aligned} \tag{C.7}$$

$$\begin{aligned} =& \frac{1}{2} \sum_{i=1}^{n+1} (n+2-i) \left(\dot{\varphi}_i^2 + 2\cos \varphi_i \right) + \\ &+ \sum_{i=1}^{n} \sum_{j=i+1}^{n+1} (n+2-j) \dot{\varphi}_i \dot{\varphi}_j \cos(\varphi_i - \varphi_j) \end{aligned} \tag{C.8}$$

□

C.2 Euler-Lagrange equations

The Euler-Lagrange equations for the n-pendulum are given by

$$\frac{d}{dt}\left(\frac{\partial L_n}{\partial \dot{\varphi}}(\varphi,\dot{\varphi})\right) - \frac{\partial L_n}{\partial \varphi}(\varphi,\dot{\varphi}) = 0. \tag{C.9}$$

where $L_n(\varphi,\dot{\varphi})$ is the just computed Lagrangian (C.4).

For $i \in \{1,\dots,n\}$ we obtain

$$\begin{aligned}\frac{\partial L_n}{\partial \dot{\varphi}_i} =&(n+1-i)\dot{\varphi}_i + \sum_{l=i+1}^{n}(n+1-l)\dot{\varphi}_l\cos(\varphi_i-\varphi_l)+ \\ &+\sum_{k=1}^{i-1}(n+1-i)\dot{\varphi}_k\cos(\varphi_k-\varphi_i)\end{aligned} \tag{C.10}$$

where the sum $\sum_{l=i+1}^{n}(n+1-l)\dot{\varphi}_l\cos(\varphi_i-\varphi_l) = 0$ for $i = n$.

$$\begin{aligned}\frac{d}{dt}\frac{\partial L_n}{\partial \dot{\varphi}_i} =&(n+1-i)\ddot{\varphi}_i + \sum_{l=i+1}^{n}(n+1-l)\left(\ddot{\varphi}_l\cos(\varphi_i-\varphi_l) - \dot{\varphi}_l\sin(\varphi_i-\varphi_l)(\dot{\varphi}_i-\dot{\varphi}_l)\right)+ \\ &+\sum_{k=1}^{i-1}(n+1-i)\left(\ddot{\varphi}_k\cos(\varphi_k-\varphi_i) - \dot{\varphi}_k\sin(\varphi_k-\varphi_i)(\dot{\varphi}_k-\dot{\varphi}_i)\right)\end{aligned} \tag{C.11}$$

$$\begin{aligned}\frac{\partial L_n}{\partial \varphi_i} =&-(n+1-i)\sin\varphi_i - \sum_{l=i+1}^{n}(n+1-l)\dot{\varphi}_i\dot{\varphi}_l\sin(\varphi_i-\varphi_l)+ \\ &+\sum_{k=1}^{i-1}(n+1-i)\dot{\varphi}_k\dot{\varphi}_i\sin(\varphi_k-\varphi_i).\end{aligned} \tag{C.12}$$

The i-th component of (C.9) then reads

$$\begin{aligned}&(n+1-i)(\ddot{\varphi}_i+\sin\varphi_i)+ \\ &+\sum_{l=i+1}^{n}(n+1-l)\left(\ddot{\varphi}_l\cos(\varphi_i-\varphi_l)+\dot{\varphi}_l^2\sin(\varphi_i-\varphi_l)\right)+ \\ &+\sum_{k=1}^{i-1}(n+1-i)\left(\ddot{\varphi}_k\cos(\varphi_k-\varphi_i)-\dot{\varphi}_k^2\sin(\varphi_k-\varphi_i)\right)=0.\end{aligned} \tag{C.13}$$

C.3 A first order system for the n-pendulum

The Euler-Lagrange equations consist of n differential equations of second order which can be written in the form

$$D(\varphi)\ddot{\varphi} + C(\varphi,\dot{\varphi})\dot{\varphi} + g(\varphi) = 0 \tag{C.14}$$

where with the abbreviations c_{km} and s_{km} for $\cos(\varphi_k - \varphi_m)$ and $\sin(\varphi_k - \varphi_m)$ respectively it follows from (C.13) that

$$D(\varphi) = \begin{pmatrix} n & (n-1)c_{12} & (n-2)c_{13} & \dots & \dots & c_{1n} \\ (n-1)c_{21} & (n-1) & (n-2)c_{23} & \dots & \dots & c_{2n} \\ (n-2)c_{31} & (n-2)c_{32} & (n-2) & (n-3)c_{34} & \dots & c_{3n} \\ \vdots & \vdots & \vdots & \ddots & \dots & \dots \\ 2c_{(n-1)1} & \vdots & \vdots & 2c_{(n-1)(n-2)} & 2 & c_{(n-1)n} \\ c_{n1} & c_{n2} & c_{n3} & \dots & \dots & c_{nn} \end{pmatrix} \tag{C.15}$$

$$C(\varphi,\dot\varphi) = \begin{pmatrix} 0 & (n-1)s_{12}\dot\varphi_2 & (n-2)s_{13}\dot\varphi_3 & \dots & \dots & s_{1n}\dot\varphi_n \\ -(n-1)s_{21}\dot\varphi_1 & 0 & (n-2)s_{23}\dot\varphi_3 & \dots & \dots & s_{2n}\dot\varphi_n \\ -(n-2)s_{31}\dot\varphi_1 & -(n-2)s_{32}\dot\varphi_2 & 0 & (n-3)s_{34}\dot\varphi_4 & \dots & s_{3n}\dot\varphi_n \\ \vdots & \vdots & \vdots & \ddots & \dots & \dots \\ -2s_{(n-1)1}\dot\varphi_1 & \vdots & \vdots & 2s_{(n-1)(n-2)}\dot\varphi_{n-2} & 0 & s_{(n-1)n}\dot\varphi_n \\ -s_{n1}\dot\varphi_1 & -s_{n2}\dot\varphi_2 & -s_{n3}\dot\varphi_3 & \dots & \dots & 0 \end{pmatrix} \tag{C.16}$$

$$g(\varphi) = \begin{pmatrix} n\sin\varphi_1 \\ (n-1)\sin\varphi_2 \\ \vdots \\ 2\sin\varphi_{n-1} \\ 1\sin\varphi_n \end{pmatrix} \tag{C.17}$$

Now, since $c_{km} = c_{mk}$, $D(\varphi)$ is a symmetric matrix. Furthermore it is positive definite (see e.g. [Murray et al., 1994][Lemma 4.2.]) and therefore invertible. We can write the equations of motion as a first order system:

$$\frac{d}{dt}\begin{pmatrix} \varphi \\ \dot\varphi \end{pmatrix} = \begin{pmatrix} \dot\varphi \\ -D^{-1}(\varphi)\left(C(\varphi,\dot\varphi)\dot\varphi + g(\varphi)\right) \end{pmatrix} = f(\varphi,\dot\varphi). \tag{C.18}$$

C.4 The linearized n-pendulum

The Jacobian of the right hand side of (C.18) is given by

$$J_f(\varphi,\dot\varphi) = \begin{pmatrix} 0 & I \\ -\frac{\partial}{\partial\varphi}[D^{-1}(\varphi)(C(\varphi,\dot\varphi)\dot\varphi + g(\varphi)] & \frac{\partial}{\partial\dot\varphi}[D^{-1}(\varphi)(C(\varphi,\dot\varphi)\dot\varphi + g(\varphi)] \end{pmatrix}. \tag{C.19}$$

which for an equilibrium point φ^* simplifies to

$$\begin{pmatrix} 0 & I \\ -D^{-1}(\varphi^*)\frac{\partial}{\partial\varphi}g(\varphi^*) & 0 \end{pmatrix} \tag{C.20}$$

since we have

$$\varphi_i^* \in \{0,\pi\}\ \forall\, i \in \{1,\dots,n\} \tag{C.21}$$

due to (C.13). Inserting (C.21) into (C.16) and (C.17) obtain

$$C(\varphi^*,\cdot) = 0 \tag{C.22}$$

as well as

$$g(\varphi^*) = 0. \tag{C.23}$$

The linearized system around a fixed point φ^* can be obtained via Taylor series expansion and a change of variables $\psi = \varphi - \varphi^*$:

$$\frac{d}{dt}\begin{pmatrix}\psi\\ \dot\psi\end{pmatrix} = \begin{pmatrix}0 & I\\ -D^{-1}(\varphi^*)\frac{\partial}{\partial\varphi}g(\varphi^*) & 0\end{pmatrix}\begin{pmatrix}\psi\\ \dot\psi\end{pmatrix}. \tag{C.24}$$

Theorem C.1. *The linear control system*

$$\frac{d}{dt}\begin{pmatrix}\psi\\ \dot\psi\end{pmatrix} = \begin{pmatrix}0 & I\\ -D^{-1}(\varphi^*)\frac{\partial}{\partial\varphi}g(\varphi^*) & 0\end{pmatrix}\begin{pmatrix}\psi\\ \dot\psi\end{pmatrix} + \begin{pmatrix}0\\0\\ \vdots\\ 0\\ 1\end{pmatrix}u \tag{C.25}$$

is globally controllable for any fixed point φ^* of (C.18) *with a scalar control input u.*

Before beginning with the proof let us introduce the operator "form":

Definition C.2. *operator* form

$$\text{form} := \begin{cases}\mathbb{R}^{(n,m)} \to \{0,1\}^{(n,m)}\\ S \mapsto \text{form}\, S = \begin{cases}0 & \textit{for } s_{ij} = 0\\ 1 & \textit{otherwise}\end{cases}, 1 \le i \le n,\ 1 \le j \le m.\end{cases} \tag{C.26}$$

We will need the following properties of the form-operator:
For $S \in \mathbb{R}^{(n,n)}$ and any diagonal matrix $\Lambda \in \mathbb{R}^{(n,n)}$ with nonzero diagonal elements we have

$$\text{form}(\Lambda S) = \text{form}(S) \tag{C.27}$$
$$\text{form}(S\Lambda) = \text{form}(S) \tag{C.28}$$
$$\tag{C.29}$$

Note that

$$\text{form}(S) = \text{form}\,\tilde S \not\Rightarrow \text{form}(S\tilde S) = \text{form}(S^2) \tag{C.30}$$
$$\text{form}(S\Lambda S) \ne \text{form}(S^2) \tag{C.31}$$

which can be seen by choosing for example $n = 2$ and $S = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ and $\begin{pmatrix}1 & -1\\ 0 & 1\end{pmatrix}$ and $\Lambda = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$.
But if S is symmetric we actually have

$$\text{form}(S^T\Lambda S) = \text{form}(S\Lambda S) = \text{form}(S^2) \tag{C.32}$$
$$\text{form}((\Lambda S)^n) = \text{form}(S^n) \tag{C.33}$$
$$\text{form}((S\Lambda)^n) = \text{form}(S^n) \tag{C.34}$$

Let $T \in \mathbb{R}^{(n,n)}$ be a tridiagonal matrix with nonzero elements on the super- and subdiagonal and e_i the i-th unit vector in $\mathbb{R}^n$ then

$$\text{form}(T^s \cdot e_n) = \text{form}(e_n + e_{n-1} + \ldots + e_{n-s}),\ \text{for}\ 1 \le s \le n-1 \tag{C.35}$$

Proof of theorem C.1.
In [Lam and Davison, 2006] it is proofed that the linearization of the inverted n-pendulum is controllable. We follow the argumentation of this proof and "generalize" the result for an arbitrary equilibrium state of the n-pendulum.

For linear autonomous systems like (C.25) we use the Kalman controllability test, which states that (C.25) is globally controllable if and only if for

$$A := \begin{pmatrix} 0 & I \\ -D^{-1}(\varphi^*)\frac{\partial}{\partial\varphi}g(\varphi^*) & 0 \end{pmatrix}; \quad B := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^{(1,2n)} \tag{C.36}$$

the matrix

$$K := [B|AB|A^2B|\dots|A^{n-1}B] \tag{C.37}$$

has rank $2n$.
Define

$$\tilde{A} := -D^{-1}(\varphi^*)\frac{\partial}{\partial\varphi}g(\varphi^*) \in \mathbb{R}^{(n,n)}; \quad \tilde{B} := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^{(1,n)} \tag{C.38}$$

then – due to the special structure of A and B – we have

$$[B|AB|A^2B|\dots|A^{n-1}B] = \left[\begin{array}{c|c|c|c|c|c|c|c|c} 0 & \tilde{B} & 0 & \tilde{A}\tilde{B} & 0 & \tilde{A}^2\tilde{B} & \dots & 0 & \tilde{A}^{n-1}\tilde{B} \\ \tilde{B} & 0 & \tilde{A}\tilde{B} & 0 & \tilde{A}^2\tilde{B} & 0 & \dots & \tilde{A}^{n-1}\tilde{B} & 0 \end{array}\right] \tag{C.39}$$

which directly implies that it will be sufficient to check if the rank of

$$\tilde{K} = \left[\ \tilde{B} \mid \tilde{A}\tilde{B} \mid \tilde{A}^2\tilde{B} \mid \dots \mid \tilde{A}^{n-1}\tilde{B}\ \right] \tag{C.40}$$

is n.

We will show that the matrix $\tilde{K}$ has the form

$$\begin{pmatrix} 0 & 0 & \dots & \dots & * \\ 0 & 0 & \dots & * & * \\ \vdots & \ddots & \dots & * & * \\ 0 & * & \dots & * & * \\ * & * & \dots & * & * \end{pmatrix} \tag{C.41}$$

where $*$ are nonzero entries. Therefore $\tilde{K}$ has rank n which will conclude the proof.

We define the following matrices:

$$D_+ = \begin{pmatrix} n & (n-1) & (n-2) & \dots & \dots & 1 \\ (n-1) & (n-1) & (n-2) & \dots & \dots & 1 \\ (n-2) & (n-2) & (n-2) & (n-3) & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \dots & \dots \\ 2 & \vdots & \vdots & \dots & 2 & 1 \\ 1 & 1 & 1 & \dots & \dots & 1 \end{pmatrix} \in \mathbb{R}^{(n,n)} \tag{C.42}$$

$$L_+ = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 1 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 1 & 1 & \dots & 1 & 0 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} \in \mathbb{R}^{(n,n)} \tag{C.43}$$

$$U_+ = L^T \tag{C.44}$$

$$D_{-i} = \text{diag}(1, 1, \dots, \underbrace{-1}_{i\text{–th position}}, \dots, 1) \in \mathbb{R}^{(n,n)} \tag{C.45}$$

$$g_+ = \text{diag}(n, n-1, \dots, 1) \in \mathbb{R}^{(n,n)} \tag{C.46}$$

For the n-pendulum the physical interpretation of an equilibrium point is that all pendulum links are pointing either downward (angle 0) or pointing upward (angle π) which in combination gives 2^n different equilibrium states.
For the "most unstable" equilibrium point $\varphi^+ = (\pi, \pi, \pi, 0, 0, 0)^T$, where all links point upward, we have $D(\varphi^+) = D_+$ and $\frac{\partial}{\partial \varphi} g(\varphi^+) = g_+$. The inverse D_+^{-1} is given by the tridiagonal matrix

$$D_+^{-1} = \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & & & -1 \\ & & & & -1 & 2 \end{pmatrix} \tag{C.47}$$

and since g_+ is a diagonal matrix with nonvanishing diagonal elements we have

$$\text{form}(D_+^{-1}) = \text{form}(-D_+^{-1} g_+). \tag{C.48}$$

and due to (C.35) for the Kalman matrix $\tilde{K}_+$ for $A_+ = -D_+^{-1} g_+$ and $\tilde{B} = e_n$ we obtain

$$\text{form}(\tilde{K}_+) = \begin{pmatrix} 0 & 0 & \dots & \dots & 1 \\ 0 & 0 & \dots & 1 & 1 \\ \vdots & \vdots & \ddots & 1 & 1 \\ 0 & 1 & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 1 \end{pmatrix} \in \{0,1\}^{(n,n)} \tag{C.49}$$

In [Seber, 2007] it is stated that the upper left entry of (C.47) is a 2 which is wrong due to the following observation which proofs the correctness of (C.47):

$$D_+ = U_+ \cdot L_+ \tag{C.50}$$
$$D_+^{-1} = (U_+ \cdot L_+)^{-1} = L_+^{-1} \cdot U_+^{-1} \tag{C.51}$$
$$= \begin{pmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & -1 & 1 & \\ & & & & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & -1 & 1 & \\ & & & & -1 & 1 \end{pmatrix}^T \tag{C.52}$$
$$= \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & & & -1 \\ & & & & -1 & 2 \end{pmatrix} \tag{C.53}$$

Starting from the inverted n-pendulum let us change the k-th pendulum position from the upright to the downright position. From the equations (C.15) and (C.17) we can see that for the new equilibrium state $\hat{\varphi}$ we obtain

$$\hat{D} = D(\hat{\varphi}) = D_{-k} D_+ D_{-k} \tag{C.54}$$
$$\hat{g} = g(\hat{\varphi}) = D_{-k} g_+ \tag{C.55}$$

and we obtain

$$\hat{D}^{-1} = (D_{-k} D_+ D_{-k})^{-1} \tag{C.56}$$
$$= D_{-k} D_+^{-1} D_{-k} \text{ since } D_{-k}^{-1} = D_{-k} \tag{C.57}$$
$$\text{form } \hat{D} = \text{form}(D_{-k} D_+ D_{-k}) = \text{form}(D_+) \tag{C.58}$$
$$\text{form}(\hat{g}) = \text{form}(D_{-k} g_+) = \text{form}(g_+) \tag{C.59}$$
$$\text{form}(-\hat{D}^{-1}\hat{g}) = \text{form}(-D_+^{-1} g_+) \tag{C.60}$$

such that $-\hat{D}^{-1}\hat{g}$ is still a tridiagonal matrix with nonzero entries on the diagonal and on the sub- and superdiagonal. For the Kalman matrix $\hat{K}$ belonging to the pair $(-\hat{D}^{-1}\hat{g}, \tilde{B})$ we therefore still obtain

$$\text{form}(\hat{K}) = \text{form}(K_+) \tag{C.61}$$

which shows controllability of the linearization around those equilibrium points of the n-pendulum where all pendulum links point upward except one pointing downward.
Now let $\bar{\varphi}$ be an arbitrary equilibrium point of the n-pendulum. Let $\mathcal{I}$ denote the set containing the number s if the s-th pendulum link points downward.
Define

$$P := \prod_{s \in \mathcal{I}} D_{-s} \tag{C.62}$$

which is well defined since the factors commute and therefore the ordering is not important.
Again we have $P^{-1} = P$ as $D_{-s}^{-1} = D_{-s}$ for $1 \leq s \leq n$.

Now for $\bar{\varphi}$ we obtain

$$\bar{D} = D(\bar{\varphi}) = PD_+P \tag{C.63}$$

$$\bar{g} = g(\bar{\varphi}) = Pg_+ \tag{C.64}$$

and we obtain

$$\bar{D}^{-1} = (PD_+P)^{-1} \tag{C.65}$$

$$= PD_+^{-1}P \text{ since } P^{-1} = P \tag{C.66}$$

$$\text{form}\,\bar{D} = \text{form}(PD_+P) = \text{form}(D_+) \tag{C.67}$$

$$\text{form}(\bar{g}) = \text{form}(Pg_+) = \text{form}(g_+) \tag{C.68}$$

$$\text{form}(-\bar{D}^{-1}\bar{g}) = \text{form}(-D_+^{-1}g_+) \tag{C.69}$$

such that $-\bar{D}^{-1}\bar{g}$ is tridiagonal with nonzero entries on the diagonal and on the sub- and superdiagonal. For $\bar{K}$ belonging to the pair $(-\bar{D}^{-1}\bar{g}, \tilde{B})$ we obtain

$$\text{form}(\bar{K}) = \text{form}(K_+) \tag{C.70}$$

showing that $\bar{K}$ has full rank concluding the proof. □

Remark: A consequence of the above theorem is that the n-pendulum is locally controllable near its equilibrium points by controlling the n-th pendulum link only. From the practical point of view it is nearly impossible to control the n-th pendulum link and one is rather interested in controlling the first pendulum link instead. The result of the above theorem still holds if we take $B = e_1^T$ and can be proved in the same way.

Theorem C.3. *Controllability of the fully actuated linearized n-pendulum*
The linear control system

$$\frac{d}{dt}(\varphi_1, \ldots, \varphi_n, \omega_1, \ldots, \omega_n)^T = A(\varphi_1, \ldots, \varphi_n, \omega_1, \ldots, \omega_n)^T + Bu, \tag{C.71}$$

is completely controllable, where A is the linearization of (C.18) *around an arbitrary point of the phase space,* $B = \begin{pmatrix} 0 \cdot I_n \\ I_n \end{pmatrix}$ *and u is an n-dimensional control input with the usual assumptions.*

Proof. From (C.18) and (C.13) we see that for all points of the state space the first n rows of the linearization matrix are given by $[0 \cdot I_n | I_n]$. But then $[B|AB]$ already has rank $2n$ and therefore due to Kalman (3.11) we have complete controllability. □

Appendix D

Pontrjagin's maximum principle

The time-optimal solution (2.15) for linear time-varying systems can also be obtained via the Pontrjagin maximum principle. Following the work of Pontrjagin, Gamkrelidze et al. [Pontrjagin et al., 1962] we introduce the Hamiltonian:

$$\mathcal{H}(\psi, x, u) = \langle \psi, A(t)x + B(t)u \rangle \tag{D.1}$$

and consider

$$\dot{x} = \frac{\partial \mathcal{H}}{\partial \psi} = A(t)x + B(t)u \tag{D.2}$$

$$\dot{\psi} = \frac{\partial \mathcal{H}}{\partial x} = -\psi A(t) \tag{D.3}$$

with all assumptions made in (B.1). If we take again $\Phi(t, 0)$ as solution of $\dot{X}(t) = A(t)X(t), X(0) = I$ then it is not difficult to see that $\Phi^{-1}(t, 0)$ is a solution of (D.3) and the general solution is $\eta^T \Phi^{-1}(t, 0)$.
We define

$$\mathcal{M}(\psi, x) := \max_{u \in U} \mathcal{H}(\psi, x, u) \tag{D.4}$$

The Pontrjagin maximum principle states, that if u^* is a time-optimal control we have for some nontrivial solution of (D.3)

$$\mathcal{H}(\psi(t), x(t, u^*), u^*(t)) = \mathcal{M}(\psi(t), x(t, u^*)) \quad a.e. \tag{D.5}$$

Using the definition of $\mathcal{M}$ and recalling $Y(t) = \Phi^{-1}(t, 0)B(t)$ we have

$$\begin{aligned} \mathcal{M}(\psi, x) &= \psi(t)A(t)x + \max_{u \in U} \psi(t)B(t)u \\ &= \psi(t)A(t)x + \psi(t)B(t)u^*(t) \text{ where} \\ u^*(t) &= \operatorname{sgn}(\psi(t)B(t)) = \operatorname{sgn}(\eta^T \Phi^{-1}(t, 0)B(t)) = \operatorname{sgn}(\eta^T Y(t)) \end{aligned}$$

maximizes $\mathcal{M}$ giving the same solution as in (B.1). The vector η is determined by the initial value.

In fact the Pontrjagin maximum principle is much stronger and even holds for some nonlinear problems. We will state Pontrjagin's result here for the *autonomous* problem. For a proof or the nonautonomous case refer to [Pontrjagin et al., 1962].

D.1 Problem statement

We are given a system of differential equations

$$\frac{dx_i}{dt} = f_i(x_1, \ldots, x_n, u_i, \ldots, u_r) = f_i(x, u), \quad i = 1, \ldots, n;\ r \leq n \tag{D.6}$$

representing for example the law of motion of a certain object with coordinates $x_1, \ldots x_n$ which are functions of the time t. They define the state vector $x = (x_1, \ldots, x_n)$ of system (D.6) as elements of an n-dimensional vector space X. The controls u are of the form $u = u(t) = (u_1(t), \ldots, u_r(t))$, where u_i are scalar-valued functions of the time t. Here we call a control admissible if it is piecewise continuous with range in a predefined set U. In vector notation we have

$$\frac{dx}{dt} = f(x, u) \tag{D.7}$$

where $f(x, u)$ is the vector with components $f_1(x, u), \ldots, f_n(x, u)$. Here and in (D.6) the functions f_i are defined for $x \in X$ (for example $X = \mathbb{R}^n$) and for $u \in U$ (e.g. $U = [-1, 1]^r \subset \mathbb{R}^r$). We will assume that the functions f_i and $\frac{\partial f_i}{\partial x_j}$, $i, j = 1, \ldots, n$ exist and are continuous on the direct product $X \times U$. Then - given a certain control $u = u(t)$ and an initial value $x(t_0) = x_0$ - the solution $x(t)$ is uniquely determined, continuous and piecewise differentiable. Sometimes to show the explicit dependence on the chosen control function u it will be denoted as $x(t, u)$.
We say the control $\tilde{u}$ defined on $[t_1, t_2]$ transfers the system from x_1 to x_2 if the solution $x(t, \tilde{u})$ is defined for all $t \in [t_1, t_2]$ and we have $x(t_1, \tilde{u}) = x_1$ and $x(t_2, \tilde{u}) = x_2$.
Given an additional function

$$f_0(x_1, \ldots, x_n, u_1, \ldots, u_r) = f_0(x, u) \tag{D.8}$$

such that f_0 and $\frac{\partial f_0}{\partial x_i}$, $i = 1, \ldots, n$ are well-defined and continuous on all of $X \times U$ then the optimal control problem reads as [Pontrjagin et al., 1962][p. 13]:

We are given two points x_1 and x_2 in the phase space X. If there are admissible control functions $u = u(t)$ which transfer system (D.6) *(or which is the same* (D.7)*) from state x_1 to state x_2 in finite time, find a control function for which the functional*

$$J := \int_{t_0}^{t_1} f_0(x(t), u(t))dt \tag{D.9}$$

takes on the least possible value. Here, $x(t)$ is the solution of (D.7) *with initial condition $x(t_0) = x_0$ and corresponding to control function $u = u(t)$. t_1 is the time at which the solution takes on the value x_2.*

Remarks:

1. For $f_0(x, u) \equiv 1$ the functional J is equal to $t_1 - t_0$ and minimizing J means minimizing the transition time from x_0 to x_1.

2. For fixed states x_1 and x_2 the upper and lower limits t_0 and t_1 are not fixed numbers but depend on the choice of the control function $u(t)$.

3. Since dealing with an autonomous system, we can relocate the initial time t_0 for the control function everywhere on the time-axis. This is because $u(t + h)$ defined on $[t_0 - h, t_1 - h]$ has the same effect as $u(t)$ defined on $[t_0, t_1]$ for all real amounts h.

4. As a consequence any part of an optimal trajectory is an optimal trajectory: Let $u(t)$ be an optimal control on the time interval $[t_0, t_1]$ bringing the system from state x_1 to state x_2. Then for any $\tau_1, \tau_2 \in [t_0, t_1]$ with $t_0 < \tau_1 < \tau_2 < t_1$ the control $u(t)$ on the interval $[\tau_1, \tau_2]$ is an optimal control bringing the system from state $x(\tau_1)$ to state $x(\tau_2)$.

Pontrjagin's maximum principle gives a necessary condition for control functions which are optimal in the above sense. It will be convenient to reformulate the problem as follows:
We adjoin a new coordinate x_0 varying according to the law

$$\frac{dx_0}{dt} = f_0(x_1, \ldots, x_n, u_1, \ldots, u_r) \tag{D.10}$$

where f_0 is the above introduced function. Adjoining this differential equation to the system of differential equations (D.6) gives

$$\frac{dx_i}{dt} = f_i(x_1, \ldots, x_n, u_i, \ldots, u_r) = f_i(x, u), \quad i = 0, \ldots, n;\ r \leq n \tag{D.11}$$

where the right hand side does not depend on x_0. The enhanced state vector is $\tilde{x} = (x_0, x_1, \ldots, x_n) = (x_0, x)$ which is an element of the enhanced $(n+1)$-dimensional vector space $\tilde{X}$. In vector notation (D.11) reads as

$$\frac{d\tilde{x}}{dt} = \tilde{f}(x, u) \tag{D.12}$$

where $\tilde{f} = (f_0, f) \in \tilde{X}$ is the vector with coordinates $f_0, f_1, \ldots, f_n$.
Define $\tilde{x}_0 = (0, x_0)$ and again let $u(t)$ be a control transferring system (D.7) from state x_1 at time t_0 to state x_2 at time t_1. Then the solution of (D.12) with initial condition $\tilde{x}(t_0) = \tilde{x}_0$ corresponding to this control function $u(t)$ is defined on the time interval $[t_0, t_1]$ and we have

$$\begin{aligned} x_0 &= \int_{t_0}^{t} f_0(x(\tau), u(\tau))d\tau \\ x &= x(t). \end{aligned}$$

In particular for $t = t_1$ this results in

$$\begin{aligned} x_0 &= \int_{t_0}^{t_1} f_0(x(\tau), u(\tau))d\tau = J \\ x &= x(t_1) = x_2 \end{aligned}$$

meaning the solution $\tilde{x}(t)$ with initial condition $\tilde{x}(t_0) = \tilde{x}_0$ passes through the point $(J, x_1) \in \tilde{X}$ at $t = t_1$. We define G to be the line in $\tilde{X}$ which is parallel to the x_0-axis (x_0 denotes the coordinate here and not the initial value of the original problem formulation) and goes through the point $(0, x_1) \in \tilde{X}$. We can then reformulate the optimal control problem as follows [Pontrjagin et al., 1962][p. 15]:

In the $(n+1)$-dimensional phase space $\tilde{X}$ the point $\tilde{x}_0 = (0, x_0)$ and the line G are given. Among all admissible controls $u = u(t)$ having the property that the corresponding solution $\tilde{x}(t)$ of (D.12) with initial condition $\tilde{x}(t_0) = \tilde{x}_0$ intersects G, find one whose point of intersection with G has the smallest coordinate x_0.

D.2 The maximum principle

Consider the following system of differential equations

$$\frac{dx_i}{dt} = f_i(x_1, \ldots, x_n, u_1, \ldots, u_r) = f_i(x, u), \qquad i = 0, \ldots, n; \quad r \leq n \tag{D.13}$$

$$\frac{d\psi_i}{dt} = -\sum_{\alpha=0}^{n} \frac{\partial f_\alpha(x, u)}{\partial x_i} \psi_\alpha, \qquad i = 0, \ldots, n \tag{D.14}$$

where $\psi_1, \ldots, \psi_n$ are auxiliary variables. Regarding a control function $u = u(t)$ on $[t_0, t_1]$ and the uniquely determined solution $\tilde{x}(t)$ of (D.13) with initial value $\tilde{x}(t_0) = \tilde{x}_0$ system (D.14) becomes

$$\frac{d\psi_i}{dt} = -\sum_{\alpha=0}^{n} \frac{\partial f_\alpha(x(t), u(t))}{\partial x_i} \psi_\alpha, \qquad i = 0, \ldots, n \tag{D.15}$$

which is a linear homogeneous differential equation admitting for any initial condition a unique solution $\psi = (\psi_1, \ldots, \psi_n)$. As $u(t)$ was supposed to be piecewise continuous, the functions ψ_i are piecewise continuously differentiable.
Introducing a new function

$$\mathcal{H}(\psi, x, u) = \langle \psi, \tilde{f}(x, u) \rangle = \sum_{\alpha=0}^{n} \psi_\alpha \cdot f_\alpha(x, u), \tag{D.16}$$

we can combine (D.13) and (D.14) as Hamiltonian system with Hamiltonian $\mathcal{H}$:

$$\frac{dx_i}{dt} = \frac{\partial \mathcal{H}}{\partial \psi_i}, \quad i = 0, 1, \ldots, n \tag{D.17}$$

$$\frac{d\psi_i}{dt} = -\frac{\partial \mathcal{H}}{\partial x_i}, \quad i = 0, 1, \ldots, n \tag{D.18}$$

Regarding $\mathcal{H}$ as a function of u we define

$$\mathcal{M}(\psi, x) = \sup_{u \in U} \mathcal{H}(\psi, x, u) \tag{D.19}$$

Theorem D.1. *Pontrjagin's maximum principle*
Let $u(t)$, $t_0 \leq t \leq t_1$ be an admissible control such that the corresponding solution $\tilde{x}(t)$ of (D.13) *starts in $\tilde{x}_0$ at time t_0 and passes at time t_1 through a point on the line $G \subset \tilde{X}$. In order that $u(t)$ is an optimal control and $\tilde{x}(t)$ is an optimal trajectory in the sense that the functional J in* (D.9) *is minimized - it is necessary that there is a nonzero continuous solution $\psi(t) = (\psi_1(t), \ldots, \psi_n(t))$ of* (D.14) *such that*

1. *For all $t \in [t_0, t_1]$ the function $\mathcal{H}(\psi(t), x(t), u)$ of the variable $u \in U$ attains its maximum:*

$$\mathcal{H}(\psi(t), x(t), u(t)) = \mathcal{M}(\psi(t), x(t))$$

2. *$\forall t \in [t_0, t_1]$ the following relations hold:*

$$\psi_0(t) \leq 0 \qquad \mathcal{M}(\psi(t), x(t)) = 0.$$

D.3 Example: Time-optimal solution for a linear autonomous system

We consider the two-dimensional linear autonomous system

$$\begin{aligned} \frac{dx_1}{dt} &= x_2 \\ \frac{dx_2}{dt} &= u \qquad |u| \leq 1. \end{aligned} \tag{D.20}$$

Here the state vector $x = (x_1, x_2)$ is an element of $X = \mathbb{R}^2$ and the control u is a function of the time t with range in the compact set $U = [-1, 1] \subset \mathbb{R}$. We want to bring the system from a given initial value x_0 to the origin $(0, 0)$ of the phase space in minimum time.
Due to (D.16) the Hamiltonian then is

$$\mathcal{H}(\psi, x, u) = \psi_1 \cdot x_2 + \psi_2 \cdot u \tag{D.21}$$

and the differential system for the auxiliary variables are

$$\begin{aligned} \frac{d\psi_1}{dt} &= 0 \\ \frac{d\psi_2}{dt} &= -\psi_1 \end{aligned} \tag{D.22}$$

which can be solved by direct integration to obtain

$$\begin{aligned} \psi_1 &= \eta_1 \\ \psi_2 &= \eta_1 - \eta_2 \cdot t, \qquad \eta_1, \eta_2 \in \mathbb{R} \text{ constants.} \end{aligned} \tag{D.23}$$

The Hamiltonian attains is maximum for

$$u(t) = \operatorname{sgn} \psi_2(t) = \operatorname{sgn}(\eta_1 - \eta_2 \cdot t). \tag{D.24}$$

So we know that a time-optimal control for (D.20) (if it exists) has to be bang-bang, changing sign at most one time.
Using the techniques of Hermes and LaSalle from chapter (1) we obtain the following results: Reformulating problem (D.20) in vector notation yields

$$\frac{dx}{dt} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t) \tag{D.25}$$

where the system matrices are $A := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $B := \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. The fundamental solution for $\dot{X}(t) = AX(t)$ with $X(0) = I$ is given by $\Phi(t, 0) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ with everywhere defined inverse $\Phi^{-1}(t, 0) = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix}$ such that $Y(t) := \Phi^{-1}(t, 0)B = \begin{pmatrix} -t \\ 1 \end{pmatrix}$. Hence system (D.25) is *normal* and time-optimal controls exist and are bang-bang. The reachability set is strictly convex and the initial value defines a normal vector η such that the optimal control has the form $u(t) = \operatorname{sgn} \eta^T Y(t) = \eta_1 - \eta_2 \cdot t$ in accordance to the result of Pontrjagin's maximum principle.

Index

Bibliography

A.V. Balakrishnan. *Applied Functional Analysis*, volume 3 of *Applications of Mathematics*. Springer, 1976.

R. W. Brockett. Nonlinear systems and differential geometry. *IEEE Transactions on Automatic Control*, 64(1):61–72, 1976.

I.N. Bronstein and K.A. Semendjajew. *Taschenbuch der Mathematik, 25*. 1991.

D. Bushaw. Optimal discontinuous forcing terms. *Contributions to the Theory of Nonlinear Oscillations*, IV:29–52, 1958.

C.-T. Chen. *Introduction to Linear System Theory*. Holt McDougal, 1970.

V. Cheng. A direct way to stabilize continuous-time and discrete-time linear time-varying systems. *IEEE Transactions on Automatic Control*, 24(4):641–643, 1979.

R. Conti. *Linear differential equations and control*. Instituto Nazionale di Alta Matematica, Bologna, Italy, 1976.

J.-M. Coron. *Control and Nonlinearity*. American Mathematical Society, Boston, MA, USA, 2007. ISBN 0821836684, 9780821836682.

J. DaCunha. Stability for time varying linear dynamic systems on time scales. *Journal of Computational and Applied Mathematics*, 176:381–410, 2005.

C. A. Desoer and L. A. Zadeh. *Linear System Theory: The State Space Approach*. McGraw Hill, New York, San Francisco, Toronto, London, 1th edition, 1963.

J.H. Eaton. An iterative solution to time-optimal control. *Journal of Mathematical Analysis and Applications*, 5:329–344, 1962.

H.G. Eggleston. *Convexity*, volume 47 of *Cambridge Tracts in Mathematics and Mathematical Physics*. Cambridge University Press, 1958.

Schlomo Engelberg. *A Mathematical Introduction to Control Theory*. Imperial College Press, London, 2005.

I. Fantoni and R. Lozano. *Non-linear Control for Underactuated Mechanical Systems*. Springer, Heidelberg, Berlin, Paris, New York, 2002.

H. Frankowska. Local controllability of control systems with feedback. *Journal of Optimization Theory and Applications*, 60:277–296, 2005.

L. Grüne and O. Junge. *Gewöhnliche Differentialgleichungen.* Vieweg+Teubner, Wiesbaden, 1st edition, 2009.

P.R. Halmos. The range of a vector measure. *Bulletin of the American Mathematical Society*, 54:416–421, 1948.

H. Hermes and J.P. LaSalle. *Functional Analysis and Time-Optimal Control.* Academic Press, New York, London, 1st edition, 1969.

Henry Hermes. Local controllability and sufficient conditions in singular problems. ii. *SIAM Journal on Control and Optimization*, 14(6):1049–1062, 1976.

R. E. Kalman. Contributions to the theory of optimal control. *Boletin de la Sociedad Matematica Mexicana*, 5:102–119, 1960.

R. E. Kalman, Y.-C. Ho, and K. S. Narendra. *Controllability of linear dynamical systems.* Contributions to differential equations I, 1th edition, 1963.

J. Kautsky, N. K. Nichols, and P. can Dooren. Robust pole assignment in linear state feedback. *International Journal of Control*, 41(5):1129–1155, 1985.

Jerzy Klamka. *Controllability of dynamical systems.* Kluwer Academic Publishers, Dordrecht, Boston, London, 1th edition, 1991.

D. L. Kleinmann. An easy way to stabilize a linear constant system. *IEEE Transactions on Automatic Control*, 15:692, 1970.

S. Lam and E.J. Davison. The real stabilizability radius of the multi-link inverted pendulum. In *American Control Conference, 2006*, page 6 pp., june 2006. doi: 10.1109/ACC.2006.1656483.

J.P. LaSalle. The time-optimal control problem. *Contributions to the Theory of Nonlinear Oscillations*, V:1–24, 1960.

E. B. Lee and L. Markus. *Foundations of optimal control theory.* The SIAM series in applied mathematics. Wiley, New York, 1th edition, 1967.

J. R. Leigh. *Functional Analysis and Linear Control Theory.* Academic Press, New York, San Francisco, Toronto, London, Sydney, 1980.

A. Lyapunov. Sur les fonctions-vecteurs completément additives. *Bulle. Acad. Sci. URSS. Sér. Math.*, 4:465–478, 1940.

R. M. Murray, Z. Li, and S. Sastry. A mathematical introduction to robotic manipulation, 1994.

T. Nagano. Linear differential systems with singularities and an application to transitive lie algebras. *J. Math. Soc. Japan*, 18:398–404, 1966.

H. Nijmeijer and A. van der Schaft. *Nonlinear Dynamical Control Systems.* Springer, New York, Heidelberg, Berlin, 1990.

L.S. Pontrjagin, V. G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko. *The Mathematical Theory of Optimal Processes.* Wiley, New York, 1962.

A. S. Poznjak. *Advanced Mathematical Tools for Automatic Control Engineers.* Elsevier, Oxford, 2008.

D. L. Russell. *Mathematics of finite-dimensional control systems.* Marcel Dekker Inc., 1979.

S. Sastry. *Nonlinear Systems - Analysis, Stability and Control.* Springer, New York, Heidelberg, Berlin, 1999.

S. Sastry, G. Walsh, D. Tilbury, R. Murray, and J. P. Laumond. Stabilization of trajectories for systems with nonholonomic constraints. *IEEE Transactions on Automatic Control*, 39(1): 216–222, 1994.

K.-H. Schlote. *Chronologie der Naturwissenschaften.* Harri Deutsch Verlag, Frankfurt am Main, 2002.

A. F. Seber. *A Matrix Handbook for Statisticians.* Wiley-Interscience, New York, NY, USA, 1st edition, 2007. ISBN 0471748692, 9780471748694.

S.C. Sinha and C.C. Chou. An approximate analysis of transient response of time-dependent linear systems by use of orthogonal polynomials. *Journal of Sound and Vibration*, 49(3):309 – 326, 1976. ISSN 0022-460X. doi: DOI:10.1016/0022-460X(76)90423-5.

G. Stefani. Local controllability of non-linear systems: An example. *Systems and Control Letters*, 6(2):123 – 125, 1985.

H.-J. Sussmann. Lie brackets, real analyticity and geometric control. In R. S. Millman R. W. Brockett and H. J. Sussmann, editors, *Differential Geometric Control Theory*, pages 1–116. Birkhauser, Boston, 1983a.

H. J. Sussmann. Small-time local controllability and continuity of the optimal time function for linear systems. *SIAM Journal of Control and Optimization*, 21(5):686–713, 1983b.

H. J. Sussmann. Lie brackets and local controllability: a sufficient condition for scalar-input systems. *Journal of Optimization Theory and Applications*, 53(2):281–296, 1987.

E. Wigner. On weakly positive matrices. *Canadian Journal of Mathematics*, 15:313–317, 1963.