Controllability of the Stresses in Multimode Viscoelastic Fluid of Upper Convected Maxwell Type

by
Evgeny Savel'ev

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mathematics

Michael Renardy, Chair
Robert C. Rogers,
David L. Russell,
Shu-Ming Sun.

June 19, 2009
Blacksburg, Virginia

Keywords: Controllability; Viscoelastic flow; Upper convected Maxwell fluid; Control of stresses

Controllability of the Stresses in Multimode Viscoelastic Fluid of Upper Convected Maxwell Type

Evgeny Savel'ev

(ABSTRACT)

Viscoelastic fluids, or Non-Newtonian fluids, are those that do not have a linear algebraic relation between the velocity field and the stresses arising in the media. Such fluids exhibit properties of both solids and liquids, and therefore cannot be modeled with methods of elasticity or Newtonian fluid mechanics. The popular models of viscoelasticity differ from each other only by the differential equation that describes the constitutive law for the fluid. Also, the media can have several relaxation modes, such as fluid mixes. This means that the stresses are determined as the sum of the stresses for each individual relaxation mode, which are described by corresponding differential equations evolving independently.

The question of controllability of the equations that describe the evolution of viscoelastic fluids is largely open. The presence of the non-algebraic constitutive relation makes the analysis unfeasible in general setup. The presence of several relaxation modes makes the problem even more complicated. Another issue is the necessity of controlling the stresses, since they are not determined by the momentary velocity field, thus they need to be included as the controlled states. In this work we are concentrating on the controllability of the stresses arising in the viscoelastic fluid that has its motion constrained to be of the shearing type. This restriction allows us to concentrate on the stresses only and assign the shearing rate to be the control. We consider only the Upper Convected Maxwell fluid which has several relaxation modes present. The results demonstrate that contrary to the one relaxation mode case the normal stresses cannot be driven arbitrary close to the exponentially decaying regime, unless the shearing stresses satisfy certain requirements, while the shear stresses remain exactly controllable.

To my beloved daughter

Acknowledgments

"Oh I'm gonna try with a little help from my friends".
-The Beatles

I, first of all, wish to express my gratitude to my advisor Prof. Michael Renardy, who patiently guided me throughout my studies and work at Virginia Tech. Without his knowledge and insight that he shared with me this work would have never been written. I am and forever will be thankful to him as my advisor and my teacher.

I also would like to thank my committee members for their time and effort. In particular, I thank Prof. David Russell for sharing with me his resources and insight into the control theory, I thank Prof. Robert Rogers for his patience and willingness to help, I especially appreciate his effort in perfecting the language of this work. I also thank Prof. Shu-Ming Sun for being an excellent teacher.

I would like to thank the faculty members of the Department of Mathematics at Virginia Tech, who provided me with knowledge that forms the core of this work. Especially, I would like to thank Prof. Martin Day, Prof. Peter Haskell and Prof. Martin Klaus who have always patiently answered my questions. I am also very thankful to my friends and fellow graduate students, who never hesitated to share their opinions and ideas with me. I appreciate especially the help I received from Austin Amaya for endlessly correcting my English and Denis Kovacs for his technical advice.

I am most grateful to my family for their selfless support and love. I am forever indebted to my wife and my daughter for keeping my spirits up and helping me to overcome the most difficult challenges.

Blacksburg, June 2009 Evgeny Savel'ev

Foreword

Controllability refers to the possibility of steering a system from given initial conditions to a desired final state with a given class of control inputs. In the context of continuum mechanics, such a control input can, for instance, be given by a body force or prescribed motion of the boundary.

The issue of controllability has been extensively studied in classical fields of continuum mechanics such as Newtonian fluids and linear elasticity. Complex nonlinear materials, however, pose new challenges. The behavior of such systems involves an interaction between deformation and microstructure. In general, this interaction puts constraints on the state of the system which cannot be altered by "macroscopic" control inputs. The study of the nature of such constraints is currently in its infancy and their full characterization poses a quite formidable problem.

This thesis addresses a special case, namely homogeneous shear flows of a multimode upper convected Maxwell fluid. The question we study is which stresses can be achieved when the shear rate is allowed to vary in an arbitrary fashion. The special structure of the upper convected Maxwell fluid allows a reformulation of this problem which is accessible to methods of the calculus of variations.

Blacksburg, June 2009 Michael Renardy

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Controllability of a system means it is possible to force the system to achieve a desired state, using a given collection of admissible controls, through which we affect the behavior of the system. The question of controllability is a part of branch of science called Control Theory. In addition to this question, control theory concerns questions such as optimal/automatic control, stability, robustness of control and many others. In early stages of the theory, when only linear constant coefficient ODE models were considered, controllability as a topic did not appear as exciting or rich. However, in case of time varying coefficients or PDE modeled processes, controllability is not that trivial a matter anymore. Consider linear, one spatial variable PDEs as an example. It is well known that the linear wave equation is controllable to any configuration through boundary control, if sufficient time is given [12, 20, 33]. The heat equation, by contrast, is not controllable exactly to arbitrary states through boundary control. This is due to its smoothing nature. Though it is exactly controllable to zero [33]. Even in the case of such elementary PDEs we can see that controllability is not a topic that has a uniform theory. Different classes of control problems require different approaches. Below we will attempt to provide a brief summary on the state of the question of controllability.

We first start with a simple control problem which is not of particularly practical interest but serves as a good test bed to demonstrate the main ideas and issues. We recommend the reader to consult the Appendix A pg. 48 for the definitions of various function spaces and notation we are going to use throughout the work. Consider a linear evolutionary problem:

$$\begin{aligned} \dot{x} &= A\,x + f(t), \\ x(0) &= x_0, \\ x(\cdot) &\in C([0,T];\mathbb{R}^n) \cap C^1([0,T];\mathbb{R}^n), \\ f(\cdot) &\in C([0,T];\mathbb{R}^n), \\ A &\in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^n). \end{aligned} \tag{1.1}$$

Here f is a control. In this setting it is trivial to control x from its initial state x_0 to any given final state. We simply connect the initial and finite state by any differentiable trajectory, then we assign $f = \dot{x} - A\,x$. Also, we can see that we can make x follow any prescribed trajectory. The conclusion is then that this system is controllable to any state over any nonzero time interval. We can also see that the control function f might turn out to have very large C norm if $x(t)$ is changing rapidly. This, in real life, would translate into large energy spending. However, in real life we do not have access to infinite resources. Therefore it is more common to have the C norm of f bounded by an *a priori* set constant. This severely restricts the dynamics of x. The system (1.1) has a solution:

$$x(t) = x_0 e^{-At} + \int_0^t e^{-A(t-s)} f(s) ds. \tag{1.2}$$

Therefore for any finite T the reachable final states of x form a bounded set. Moreover, since the operator $L(x) = \dot{x} - A\,x$ is unbounded, the question arises whether this reachable set of final states has nonempty interior. Luckily, in this simple setting it is not difficult to show that the reachable set not only does have nonempty interior, it is also closed and simply connected if the set of allowable control functions is a closed ball in $C([0,T];\mathbb{R}^n)$. However, the list of difficulties we can run into doesn't stop here. It can happen that some final states become completely unreachable. Consider for example the system below, for which the state $x(t) = 2$ can never be achieved:

$$\dot{x} = -x + u, \quad x(0) = 0, u \in (-1, 1).$$

In this example we can see that controllability is affected by the nature of the system (i.e. the differential equations that describe system's dynamics) and the restrictions imposed on the system.

Let us now consider a more practical problem, which arises quite often in automatic control applications. Suppose the dynamics of the system is governed by the following ODE:

$$\begin{aligned} \dot{x} &= A\,x + B\,u, \quad x(t) \in \mathbb{R}^n,\ u(t) \in \mathbb{R}^m, \\ x(0) &= x_0. \end{aligned} \tag{1.3}$$

Here A and B are matrices of appropriate dimensions. The problem (1.3) is called *controllable* (over the interval $[0,T]$) if for any pair of points x_0, x_T there exists a control $u \in L^2([0,T];\mathbb{R}^m)$ so that:

$$x(0) = x_0,\ x(T) = x_T.$$

The following well known result provides a simple technical condition for controllability:

Theorem 1.1 (Kalman Rank Condition.)**.** *The system* (1.3) *is controllable if and only if the following holds:*

$$\operatorname{rank}[B, AB, A^2B, \cdots, A^{n-1}B] = n. \tag{1.4}$$

The condition (1.4) does not refer to a specific interval $[0,T]$. The system is either controllable or not, regardless of alloted time. For the proofs we refer the reader to [4] and [2]. Directly following from this controllability criterion is the Popov-Belevitch-Hautus test, which says that the system (1.3) is controllable if and only if for all eigenvalues λ_i of A,

$$\text{rank}[\lambda_i I - A,\, B] = n.$$

One of the most famous results of control theory for linear constant coefficient ODE problems of the form (1.3) is a duality of the controllability and observability problems.

Definition 1.2. The system

$$\begin{array}{rcll} \dot{y} & = & C\,y, & y \in \mathbb{R}^n \\ \omega & = & H\,y, & \omega \in \mathbb{R}^m \end{array} \tag{1.5}$$

is called *observable* if for any nonzero solution y of the ODE above, the "observation" ω is nonzero.

There exists a result similar to (1.4):

Theorem 1.3. *The system* (1.5) *is observable if and only if*

$$\text{rank} \begin{bmatrix} H \\ HC \\ HC^2 \\ \vdots \\ HC^{n-1} \end{bmatrix} = n. \tag{1.6}$$

The main result is the following theorem [4], [2]:

Theorem 1.4. *The system* (1.3) *is controllable if and only if the system* (1.5) *with* $C = A^T$, $H = B^T$ *is observable.*

As was noted in [33], this result is merely an application of the well known functional analysis result relating the range of the operator and the kernel of its adjoint. However the result is important in control theory since it permits the generalization to infinite dimensional cases.

If we replace the matrices A and B with continuously varying linear operators $A(t)$ and $B(t)$ then things are not so elementary. For example, we loose the independence of controllability on the time interval $[0, T]$ (in case of constant coefficients controllability property depends only on A and B). The system may also become controllable to only a subset of possible final states. We refer interested readers to [34] where a condition similar to Kalman rank condition (1.4) for such linear ODE systems is provided.

Now we shall finish the discussion of ODEs with a quick look at the controllability of nonlinear ones. One can derive the controllability from the controllability of the linearized system, however this approach presents tremendous challenge in general cases, and is usually reserved for relatively simple special cases. Fortunately, there is a generalization of the linear systems result obtained by Kalman. In the linear case we use matrix rank as an indicator of linear independence of the possible directions that $x(t)$ can follow. We also need such a tool for the nonlinear vector functions. The tools necessary do exist, namely Lie algebras and Lie brackets. The result, however, is somewhat weaker than controllability. We will be able to obtain accessibility condition for systems of the form $\dot{x} = f(x, u)$.

Before we introduce the results, we note for the future that the control $u(t)$ that appears on the right hand side of $\dot{x} = f(x, u)$ is assumed the be the member of $L^\infty([0, T]; \mathcal{U})$, where $\mathcal{U}$ is a known vector space.

The ideas we are going to discuss were pioneered in [11], [8], [9]. For detailed overview of the current state of the field and proofs of theorems we refer to [34]. Here we only briefly outline the main ideas.

Definition 1.5. We shall call a (nonlinear) mapping f a *vector field* if it is (infinitely) differentiable and defined as follows:

$$f = (f_1, f_2, \ldots, f_n) : O \subset \mathbb{R}^n \to \mathbb{R}^n \tag{1.7}$$

The set of all vector fields $\mathbb{V}(O)$ defined on the set O is a vector space.
The *Lie bracket* of $f,\ g \in \mathbb{V}(O)$, denoted as $[f, g]$, is a vector field defined as:

$$[f, g] = (f \cdot \nabla)g - (g \cdot \nabla)f. \tag{1.8}$$

The *Lie algebra* $S \subset \mathbb{V}(O)$ is a vector space of vector fields which is closed under the Lie bracket operation, i.e. $\forall f, g \in S$ we have $[f, g] \in S$. Suppose A is a subset of $\mathbb{V}(O)$, then the *Lie algebra* A_{LA} *generated by* A is the intersection of all Lie algebras that contain A, or, in other words, a smallest Lie algebra containing A.

We now have the sufficient toolbox to establish accessibility condition. Consider a collection of vector fields obtained by putting all admissible controls in f:

$$A = \{f(\cdot, u) | u \in L^\infty([0, T]; \mathcal{U})\}. \tag{1.9}$$

We denote the set of all possible states $x(t)$ that are reachable in time less than or equal to T from the initial state x_0 as the following:

$$\mathcal{R}^{t \leqslant T}(x_0) = \{z_0 \in \mathbb{R}^n | \exists t \in [0, T],\ u \in L^\infty([0, T]; \mathcal{U}) : x(0) = x_0 \text{ and } x(t) = z_0\}. \tag{1.10}$$

In a similar way, we denote the set of all states that are controllable to z_0 as:

$$\mathcal{C}^{t \leqslant T}(z_0) = \{x_0 \in \mathbb{R}^n | \exists t \in [0, T],\ u \in L^\infty([0, T]; \mathcal{U}) : x(0) = x_0 \text{ and } x(t) = z_0\}. \tag{1.11}$$

Here for both definitions $x(t)$ is a solution for $\dot{x} = f(x, u)$.

Definition 1.6. The Lie algebra A_{LA} generated by (1.9) is called the *accessibility Lie algebra*, and we say that the *accessibility rank condition (ARC)* holds at x_0 if $A_{LA}(x_0) = \mathbb{R}^n$.

We shall now state the main result here.

Theorem 1.7. *Suppose that the ARC holds at x_0 for the system, then the sets $\mathcal{R}^{t \leqslant T}(x_0)$ and $\mathcal{C}^{t \leqslant T}(x_0)$ both have nonempty interiors.*

Consider the situation when ARC holds at both x_0, z_0. This implies that we can start at x_0 and reach any point in some small neighborhood of z_0, or we can start in some small neighborhood of x_0 and reach z_0. However, this does not necessarily imply that we can find a small neighborhood of x_0 so that if we start from inside of it we can reach any point in some small neighborhood of z_0. Thus accessibility does not imply controllability in general. However, for some restricted classes of evolutional equations the equivalence of two can be shown. For example accessibility and controllability are equivalent for reversible problems, or for the problems which have $f(\cdot, u)$ analytic for every admissible u.

The results for linear ODEs presented above permit generalization to the infinite dimensional cases. Naturally the framework changes: the matrices are replaced with linear operators and finite dimensional Euclidian spaces are replaced with Hilbert spaces. Also, the transition to infinite dimensions brings qualitative changes. First of all, a dense subspace of a Hilbert space does not need to be the whole space. This necessitates the notion of approximate controllability. Secondly, PDE problems assume boundary conditions. Consequently, the control problem may be formulated for distributed control as well as for boundary located control, thus producing two distinct control problems which may differ significantly.

In general the linear control problem can be schematically presented as the mapping diagram. In the Fig 1.1, see [33]. In figure 1.1 we have X, U, Z are Hilbert spaces; $C : \mathcal{D}(C) \to Z$ and

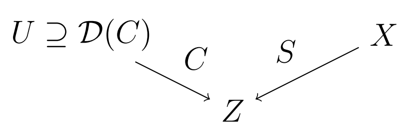


Figure 1.1: Linear control problem

$S : \mathcal{D}(S) \to Z$ are bounded linear operators with dense domains.

As was mentioned before, the infinite dimensional nature of the problem destroys the equivalence between dense subspaces and the whole space. This motivates the following controllability notation:

Definition 1.8. The control system on Fig. 1.1 is called

1. *(exactly) controllable* if $\mathcal{R}(S) \subseteq \mathcal{R}(C)$,

2. *approximately controllable* if $\mathcal{R}(S) \subseteq \overline{\mathcal{R}(C)}$.

As an example of how this framework fits PDE control problems we look at the following problem found in [33]:

$$
\begin{aligned}
\dot{z} &= A\,z + B\,u, \\
z(0) &= z_0, \quad z(t) \in H,\ u(t) \in V,
\end{aligned} \tag{1.12}
$$

where H and V are Hilbert spaces, the operator $A : H \to H$ is a generator of a C_0 semigroup $e^{A\,t}$ and $B : V \to H$ is a bounded linear operator. The variation of parameters formula yields a solution:

$$
z(t) = e^{A\,t} z_0 + \int_0^t e^{A(t-s)} B\,u\,ds. \tag{1.13}
$$

We can assign meaning to the above mentioned spaces X, U, Z and operators C, S in the following way:

$$
\begin{aligned}
X &= H \times H, \\
U &= L^2([0,T]; V), \\
Z &= H, \\
S &: X \to Z\ : S(z_1, z_0) = z_1 - e^{A\,T} z_0, \\
C &: U \to Z\ : C(u) = \textstyle\int_0^T e^{A(T-s)} B\,u\,ds.
\end{aligned} \tag{1.14}
$$

Now, we can see that if this system is exactly controllable in the sense of the definition we gave above, then there is a control u that drives the solution of (1.12) from the initial state z_0 to the final state $z(T) = z_1$ in time T.

As in the case of finite dimensional systems, there is a duality relation between controllability and observability problems. The density issues also cause slight alteration to the way we define observability.

Again, we present the problem schematically in Fig. 1.2

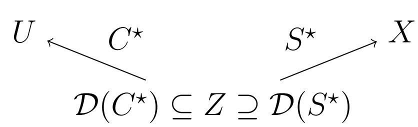


Figure 1.2: Linear observed problem

Definition 1.9. The system in the Fig. 1.2 is called

1. *distinguishable* if $\ker(C^\star) \subseteq \ker(S^\star)$,

2. *observable* if in addition to the above we have for $z \in \mathcal{D}(C^\star)$ $\|C^\star z\|_U \geqslant K \|S^\star z\|_X$ for some positive constant K.

In [5] the following result has been proved:

Theorem 1.10. *The control system in Fig. 1.1 is exactly (approximately) controllable if the adjoint system on Fig. 1.2 is observable (distinguishable).*

In many cases the observability of the adjoint problem is easier to show than the controllability of the actual problem. Thus observability analysis is often used as a convenient tool for analyzing controllability. However, the result presented in [33] provides an even easier criterion for controllability. The result is based on the property of the system called *stabilizability*, which arises in many automatic (feedback) control applications.

We first start with the definition of stabilizability.

Definition 1.11. The system (1.3) is called *stabilizable* (or *forward stabilizable*) if there exists a matrix K so that every eigenvalue of $A + BK$ has a negative real part.

This can be interpreted in the following way: there exists a feedback control $u = K\,x$ that makes all solutions $x(t|x_0)$ tend to zero as $t \to \infty$. It is true that controllable systems are stabilizable. However the converse doesn't hold unless additional assumptions are made. In particular, if so called backward stabilizability holds then the system is controllable.

Theorem 1.12. *If the system* (1.3) *can be stabilized both forward and backward, i.e. there exist two matrices K^-, K^+ so that the eigenvalues of $A + B\,K^-$ have their real part negative, while the eigenvalues of $A + B\,K^+$ have real part positive, then the system is controllable.*

This result has been proven in [32] and generalized to the infinite dimensional systems in [33], in which it was used to show boundary controllability of hyperbolic systems. In [3] this method is applied to obtain the controllability of a nonlinear PDE control problem for operators generating a nonlinear semigroup.

The strong interconnection between controllability, observbility and stabilizability forms the core of most of the controllability research. However, the methods that are used for establishing them are different for different classes of evolutionary problems. There is always an operator theoretical framework outlined above in figures 1.1, 1.2. However the level of abstraction sometimes makes the analysis unfeasible. Below we attempt to give a brief review of most commonly used methods for establishing controllability.

For hyperbolic problems with boundary located control the powerful method has been developed in [10, 20]. One of the authors has named it the *Hilbert uniqueness method* (see [20]). Other authors refer to it as the *multiplier method* (see [10,12]). In essence the main idea is to construct special Hilbert space so that the operator A in (1.12) is a Riesz-Fréchet mapping to the dual space. This setting is helpful since it makes the proof of uniqueness of the solution of the adjoint observability problem to be a technical matter.

Another method is based on “harmonic” analysis. It allows one to reduce the controllability problem to the problem of moments for certain exponential systems. The method can be successfully applied to hyperbolic and parabolic problems [1, 31, 33]. The most up to date and complete theory on this method can be found in [1, 13].

In 1990 the result announced by Lions [21] stimulated great interest for researchers in questions of controllability of the equations of fluid flow. The number of controllability results for Navier-Stokes and Euler equations is so vast that any attempt to produce a more or less detailed overview of the subject would easily take more space than we have. We refer interested readers to the brief historical introduction into the subject given in [7]. Most of the research is concerned with controllability of the velocity field, and the results are obtained via solving the auxiliary linear problem (for example using the equations for streamline functions), controllability of which implies the controllability of the original equations. The results include boundary (local exact or approximate) controllability of the solution to the steady state solution or to the solution of the problem with homogeneous boundary condition.

Going beyond Newtonian fluids in investigation of controllability is linked with numerous complications. One needs to deal with nonlinear equation of motion coupled with (in many cases nonlinear) constitutive equation. Coupled PDEs, even linear, presents new challenges. As an interesting reading on the topic, we refer the reader to the lecture notes [15], where the coupling of linear problems of acoustics and elastics are considered. Naturally, the coupling of nonlinear PDEs raises the bar even higher. So far no literature is available on the topic where the research is done in this generality.

In this work, we attempt to shed some light on the problem of controllability of the viscoelastic fluids, which are a subclass of non-Newtonian fluids. We are concerned with controllability of the stresses in upper convected Maxwell (UCM) fluid, which is one of the simplest models of nonlinear viscoelasticity.

The viscoelastic fluids take a special place in the theory of fluid dynamics. While exposing the properties of liquids, they also possess some qualities of elastic solids. The distinguishing feature that viscoelastic media possess is the fact that stresses arising in the motion depend not only on the current state of the body (i.e. the velocity field and displacement) but also on the history of the changes that the body was subjected to. In other words, viscoelastic media has “memory”.

The second half of the 20th century was marked by the rising interest in Non-Newtonian fluids. Mostly this interest was due to the development of the plastic industry, as the conventional Newtonian fluid models and elastic solids models could not account for the effects observed in the experiments. In the works of Oldroyd, Rivlin, Ericksen, Truesdell and Noll ([24, 30, 35]) a rigorous and universal theory was developed. Nevertheless, despite the great progress that has been achieved, the vast majority of questions, such as global existence and uniqueness of solutions in general settings, remains unanswered. For a brief overview of the field we refer the reader to [29].

One of these unanswered questions is the question of controllability of the governing equations. The challenges these problems present force the researcher to restrict oneself to the consideration of very narrow special cases. For example one might consider only very special boundary conditions.

There have been several attempts to attack viscoelastic cases. In [6, 25] the authors analyze the case where the constitutive equation is a linear ODE and the control is distributed over the space domain and included in the right hand side of the equation of motion. Only shear stresses were considered as the linear models of viscoelasticity predict zero normal stresses. In this setting exact controllability is obtained if control is distributed over the whole domain. If the control is distributed only over a strict subset of the domain the approximate controllability holds if inertia is included. In older papers [16–18, 22] the stresses are not considered as a controlled state. The systematic study of stress controllability of viscoelastic fluids in simple setups for various constitutive equations can be found in [25–28]. In [26] the obtained results show that normal stresses are controllable to restricted states only for various nonlinear one relaxation mode models when the velocity field is restricted to be of homogeneous shearing type. All popular models have lower bounds on the achievable normal stresses, some even having upper bounds. In the subsequent paper [27] the effects of inertia were included, spatially distributed control was assumed and only the upper convected Maxwell fluid was considered. The results show that the control needs to be distributed over the whole domain in order to control the normal stresses. Nevertheless, the stress tensor still needs to satisfy the positive definiteness constraint. In the latest work [28], arbitrary homogeneous velocity fields were allowed, while inertia related effects were discarded. Even with arbitrary velocity fields, for some models this still does not remove the constraints on the determinant of the stress tensor.

The recent advance in the area of numerical simulation of viscoelastic flows has allowed the investigations of control problems through numerical experiments [14, 23]. The approach these authors utilized is flexible enough to produce the control in situations where the numerical simulation does converge to a true solution. It certainly has great practical value. However, the approximated equations still elude attempts to analyze them. Also, we point out that the research does not touch the question of controllability at all. The authors themselves admit that the numerical procedure they developed yields sub-optimal control, and only concentrates on the control of the velocity field. Thus the question of achievable states in terms of both velocities and stresses is still open. Nevertheless, this result is interesting in the sense that it is a one of the first attempts to control a viscoelastic flow with nonlinear constitutive equation in a very general setup.

The difficulties arise from the above-mentioned feature that, unlike Newtonian fluids, the state of viscoelastic fluid is not defined completely by the velocity field at the present time alone. Therefore it is not sufficient to control only the velocity field, one needs to pay attention to the dynamics of the arising stresses at the constitutive equations level. The most popular models of viscoelasticity have their constitutive equations expressed in the form of nonlinear differential equations. This adds another dimension of complexity for controllability problems and as a result this area of fluid mechanics remains largely untouched.

To overcome these difficulties we are going to employ a number of simplifications. First, we consider only the case where the fluid is subjected to a homogenous shearing motion, with a variable shearing rate that acts as a control. Second, we assume that the fluid adapts to changes in shearing motion rate immediately, i.e. the fluid possesses no inertia. This simplification is definitely unphysical. However, for slowly changing flows the effects of inertia are small, thus we have a more or less justified reason to neglect them. The reasons we chose shearing motion are as follows:

1. Shearing motion is simple enough to permit feasible analysis. Also, because of its simplicity, we can concentrate on the controllability of the stresses without having to pay attention to the controllability of the velocity field. Therefore we are able to observe the effects of control on the stresses that are due only to the viscoelastic nature of the media, excluding the effects that the different arbitrary velocity patterns induce in general.

2. Even in the case of such simple setup, we are able to observe phenomena that are very different from the behavior of Newtonian fluids. In the case of laminar flow of Newtonian fluid one only needs to analyze the cases of shearing motion and elongation flow to be able to predict the effects inside any general flow pattern. Despite that this is only partially true for viscoelastic fluids, the shearing motion case does provide enough information to reveal the differences between Newtonian and Non-Newtonian fluids.

3. Viscometric flows, such as shearing flow or flows similar to shearing flows, are the basis of many viscometers that are used very often in industrial processes. Thus we deem it reasonable to begin the research in this direction.

This work is a continuation of research originated by Renardy in [26]. Renardy discusses the controllability of the various models of viscoelastic fluid under shearing motion. In this paper only the cases of one relaxation mode were considered. We attempt to extend the results for the case of several relaxation modes present.

The equations do not allow the use of the methods of controllability analysis described above, since they require the states of the system to form a vector space. The shear stresses (off-diagonal elements in the stress tensor) do satisfy this requirement. However the normal stresses do not. For this reason, the analysis of the achievable stresses concentrates mostly on normal stresses, while shear stresses (actually only the shear stresses of a particular relaxation mode) will be considered as a control. This approach is justified since the exact controllability relation between the shear stresses and shear rate holds in both ways.

Chapter 2

Simple homogeneous shear flow

In this chapter we introduce the basic equations for the fluid model and present some preliminary results.

2.1 Equations of motion

In this work we only consider the shearing flow case, in which the velocity has the form

$$\boldsymbol{u}(t,x,y,z) = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} u_1(t,y) \\ 0 \\ 0 \end{pmatrix}. \tag{2.1}$$

The dependence of $u_1(t,y)$ on y is linear. Thus $u_1(t,y) = \dot{\gamma}(t) \cdot y$. The quantity $\dot{\gamma}(t)$ is called shear rate and considered to be the only control available. We can easily check that the incompressibility condition holds:

$$\mathrm{div}\boldsymbol{u} = 0.$$

Consequently the velocity gradient has the form

$$\nabla \boldsymbol{u} = \begin{pmatrix} 0 & \dot{\gamma}(t) & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.2}$$

The particle positioned at $\boldsymbol{x}(t) = (x,y,z)$ at time t will occupy at time s the position $\boldsymbol{x}(s) = (x + y(\gamma(s) - \gamma(t)), y, z)$, where $\gamma(s) - \gamma(t) = \int_t^s \dot{\gamma}$. Therefore the relative deformation gradient has the form

$$\boldsymbol{F}(\boldsymbol{x},t,s) = \begin{pmatrix} 1 & (\gamma(s) - \gamma(t)) & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.3}$$

The relative Cauchy strain tensor has the form

$$\boldsymbol{C}(\boldsymbol{x},t,s) = \begin{pmatrix} 1 & (\gamma(s)-\gamma(t)) & 0 \\ (\gamma(s)-\gamma(t)) & 1+(\gamma(s)-\gamma(t))^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.4}$$

Since the Cauchy strain tensor and its relation to the stress tensor have to be isotropic we deduce that the stress tensor has the form

$$\boldsymbol{T} = \begin{pmatrix} T_{11} & T_{12} & 0 \\ T_{12} & T_{22} & 0 \\ 0 & 0 & T_{33} \end{pmatrix}. \tag{2.5}$$

In this work we are concerned with the case when the stress is a sum of individual relaxation mode contributions:

$$\boldsymbol{T} = \sum_{i=1}^{n} \boldsymbol{T}_i. \tag{2.6}$$

Each relaxation mode satisfies its own constitutive equation:

$$\dot{\boldsymbol{T}}_i - (\nabla \boldsymbol{u})\boldsymbol{T}_i - \boldsymbol{T}_i(\nabla \boldsymbol{u})^T + \lambda_i \boldsymbol{T}_i = \mu_i \left(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^T\right), i = 1, 2 \ldots, n. \tag{2.7}$$

Here all λ_i are assumed to be distinct and positive. To simplify notation we denote the individual components of the stress tensor as

$$\boldsymbol{T}_i = \begin{pmatrix} \sigma_i & \tau_i & 0 \\ \tau_i & \psi_i & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.8}$$

Here we used the symmetry of the stress tensor and the fact that $T_{33} = 0$ in this setup. Since $\nabla \boldsymbol{u}$ is known in advance, we can plug (2.8) into (2.7) and arrive at

$$\begin{aligned} \dot{\sigma}_i - 2\tau_i\dot{\gamma} + \lambda_i\sigma_i &= 0, \\ \dot{\tau}_i - \psi_i\dot{\gamma} + \lambda_i\tau_i &= \mu_i\dot{\gamma}, \\ \dot{\psi}_i + \lambda_i\psi_i &= 0. \end{aligned} \tag{2.9}$$

The last equation, for ψ, has solution $\psi_i(t) = \psi_i(0)e^{-\lambda_i t}$, which we cannot control at all. Assuming that since the beginning of the motion (when all the stresses were zero) the flow was of shearing type we can put $\psi_i(t) = 0$. Hence we are left with two sets of equations:

$$\begin{aligned} \dot{\sigma}_i - 2\tau_i\dot{\gamma} + \lambda_i\sigma_i &= 0, \\ \dot{\tau}_i + \lambda_i\tau_i &= \mu_i\dot{\gamma}. \end{aligned} \tag{2.10}$$

The collection of n pairs of equations (2.10) forms our control problem.

2.2 Controllability of the shear stresses

The problem we need to analyze involves the control of $2n$ equations (n is the number of the relaxation modes involved). Here we present an argument that would allow us to reduce the problem to controlling only the equations for σ_i while looking at equations for τ_i as a restriction on a control problem.

We can write the equations for τ_i as the following system:

$$\boldsymbol{\tau} = \Lambda\,\boldsymbol{\tau} + \mathrm{M}\,\dot{\gamma}. \tag{2.11}$$

Here Λ is a diagonal matrix with $-\lambda_i$ on the diagonal and M is the vector with μ_i as the components. For physical reasons all $\mu_i > 0$. Since we assumed that λ_i are all distinct and positive, the determinant of the square matrix

$$\det\left[\mathrm{M}, \Lambda\mathrm{M}, \Lambda^2\mathrm{M}, \ldots, \Lambda^{n-1}\mathrm{M}\right] \neq 0. \tag{2.12}$$

This means that the Kalman rank condition is satisfied, and therefore the system is controllable from any set of initial states $\tau_i(0)$ to any finite states $\tau_i(T)$.

Later in this work we assume that τ_i are elements of $L^2[0,T]$, which renders the initial and final values of τ_i meaningless. However this result is important since we can exclude the control $\dot{\gamma}$ from the problem and consider one of the shear stress function τ_i or any linear combination of them as a control. Moreover, further down we shall find $n-1$ independent linear combinations of τ_i that belong to $H^1[0,T]$, for which the initial and final conditions will make sense because of continuity of H^1 functions. Most importantly, the exact controllability result obtained here will be inherited.

The question may arise whether we lose all degrees of freedom in the process of controlling τ_i, and requiring a certain final values of $\tau_i(T)$ leaves ourselves no choice but to let normal stresses σ_i follow passively. To answer this question we look at the solution for τ_i obtained by variation of parameters:

$$\tau_i(T) - e^{-\lambda_i T}\tau_i(0) = \int_0^T \mu_i e^{-\lambda_i (T-t)}\dot{\gamma}\,dt. \tag{2.13}$$

The intersection of the kernels of the linear functionals $l_i(\dot{\gamma}) = \int_0^T \mu_i e^{-\lambda_i(T-t)}\dot{\gamma}\,dt$ is clearly infinite-dimensional. Thus there are infinitely many possibilities to control σ_i even after we restrict ourselves by choosing $\tau_i(0)$, $\tau_i(T)$ *a priori*.

Chapter 3

Controllability of normal stresses.

We showed in the previous chapter that we can reduce the problem from controlling all stresses to controlling only the normal stresses. In this chapter we develop the methods for investigating the controllability of the normal stresses. In [28] it has been shown that the stress tensor has to be positive definite. This means that the set of all achievable stresses does not form a vector space and the many tools available for controllability analysis cannot be applied here. However, since the set of achievable stresses now has a boundary, we can apply the methods of optimal control and calculus of variations to discover it. Following the derivation of the equations in the previous chapter we start with the two sets of equations:

$$\begin{aligned} \dot{\sigma}_i - 2\tau_i\dot{\gamma} + \lambda_i\sigma_i &= 0, \\ \dot{\tau}_i + \lambda_i\tau_i &= \mu_i\dot{\gamma}. \end{aligned}$$

For a given control $\dot{\gamma}$, we are fortunate to find the solution in closed form:

$$\tau_i(T) = \tau_i(0)e^{-\lambda_i T} + e^{-\lambda_i T}\int_0^T \mu_i e^{\lambda_i t}\dot{\gamma}dt \tag{3.1}$$

$$\mu_i\sigma_i(T) - \tau_i^2(T) = e^{-\lambda_i T}(\mu_i\sigma_i(0) - \tau_i^2(0)) + \lambda_i e^{-\lambda_i T}\int_0^T e^{\lambda_i t}\tau_i^2(t)dt. \tag{3.2}$$

As we are mainly going to concentrate on (3.2), we find it convenient to simplify the notation. The expression for the normal stress $\sigma_i(T)$ is too bulky to manipulate and, since we can only control the integral part of it, for the sake of simplicity we introduce a new function $S_i(T)$ to represent just the integral part of the expression:

$$S_i(T) = \mu_i\sigma_i(T) - \tau_i^2(T) - e^{-\lambda_i T}(\mu_i\sigma_i(0) - \tau_i^2(0)) = \lambda_i\int_0^T e^{-\lambda_i(T-t)}\tau_i^2(t)\,dt. \tag{3.3}$$

Notice that the quantity $\tau_i^2(T)$ is in general affected by the control $\dot{\gamma}$ and combining the effects of the control on two variables into one quantity is not always appropriate. However, we point out that because we earlier established the controllability of the shear stresses $\tau_i(t)$ to any desirable state we may as well assume that the final states $\tau_i(T)$ are constant for our case and, consequently, we reduce the problem of controllability of the normal stresses $\sigma_i(T)$ to the problem of controllability of the quantities $S_i(T)$. We summarize everything in the definition of the modified problem we are to concentrate on for the rest of this work:

Definition 3.1 (*Problem Statement*)**.** We denote by $\Omega_S = \{S_1(T), S_2(T), \dots, S_n(T)\}$ the set of all attainable values S_i at the final time T given the initial and final shear stresses $\tau_i(0)$ and $\tau_i(T)$. The control problem of interest here is to find the complete description of the set Ω_S under these conditions.

Once we obtain the characterization of the set Ω_S we can then vary the final values of $\tau_i(T)$ to observe all possible reachable states.

We note that we are only interested in finite $S_i(T)$ values. Thus we require the integral expression $\int_0^T e^{\lambda_i t}\tau_i^2(t)dt$ to be finite. This forces each $\tau_i(t)$ to be in $L^2[0,T]$. This requirement, however, puts the validity of the boundary values of τ_i in question. We shall later circumvent this difficulty by introducing the H^1 function $q(t)$ and then reformulate the problem in terms of this new variable.

One immediate observation is the fact that S_i are not fully controllable: they must all be positive (unless $\tau_i \equiv 0$). Moreover if we choose one particular i it is possible to tune $\tau_i(t)$ (and consequently $\dot{\gamma}$) so that $S_i(T)$ is ε-close to zero. In case of only one relaxation mode our quest for controllability would end here. However, it is natural to expect that in case when several relaxation modes are present the boundary of Ω_S does not consist of a union of the sets $S_i(T) = 0$.

3.1 Two relaxation modes

We begin our analysis with just two relaxation modes. This case is simple enough to yield a comprehensive analysis method. However it is complex enough to reveal the effects that, after suitable generalization, might be extrapolated to cases of more relaxation modes.

The analysis below will be separated into several steps:

- Analyze what values of S_1 we can achieve, disregarding the effect of the control on S_2;
- Analyze what values of S_2 we can achieve, keeping S_1 controlled at the desired level;
- Analyze the interconnection between S_1 and S_2 and attempt to describe the behavior in the extreme cases.

We will achieve these goals by employing the methods of calculus of variations. By definition, the quantities S_i depend only on the behavior of τ_i, which are interconnected to each other through $\dot{\gamma}$. Thus it seems natural to formulate the problem in terms of variations of $\dot{\gamma}$. However, we pointed out earlier that the boundary conditions for these functions do not make sense. Thus we perform a change of variables to overcome this issue. The new variable will belong to H^1 instead of L^2, and this will be of great convenience. The derivation is simple. We combine the equations for τ_i eliminating $\mu_i\dot{\gamma}$ terms and arrive at:

$$\mu_2\dot{\tau}_1 + \mu_2\lambda_1\tau_i = \mu_1\dot{\tau}_2 + \mu_1\lambda_2\tau_2.$$

This, consequently, can be rewritten as

$$\partial_t(\mu_2\tau_1 - \mu_1\tau_2) + (\mu_2\lambda_1\tau_1 - \mu_1\lambda_2\tau_2) = 0. \tag{3.4}$$

Now, by calling

$$\begin{array}{rcl} q &=& \mu_2\tau_1 - \mu_1\tau_2, \\ \text{and} && \\ q_1 &=& -(\mu_2\lambda_1\tau_1 - \mu_1\lambda_2\tau_2), \end{array} \tag{3.5}$$

the equation (3.4) finally turns into

$$\dot{q} = q_1.$$

The variable q we introduced here has a derivative in L^2. Therefore it is a member of H^1 and continuous.

Expressing τ_1 and τ_2 in terms of these quantites we get

$$\begin{array}{rcl} \tau_1 &=& \frac{\lambda_2 q + q_1}{(\lambda_2-\lambda_1)\mu_2} = \frac{\lambda_2 q + \dot{q}}{(\lambda_2-\lambda_1)\mu_2}, \\ \tau_2 &=& \frac{\lambda_1 q + q_1}{(\lambda_2-\lambda_1)\mu_1} = \frac{\lambda_1 q + \dot{q}}{(\lambda_2-\lambda_1)\mu_1}. \end{array} \tag{3.6}$$

We can see that the boundary values of τ_i (in the rare case when they make sense) put restrictions on boundary values of both q and $\dot{q}$. These restrictions are too tight to yield a solution in most cases, thus we drop the restrictions on $\dot{q}|_{t=0,T}$ since $\dot{q}$ is a member of L^2. We'll return to this matter later; for now we restrict ourselves to:

$$\begin{array}{l} q(0) = \mu_1\tau_1(0) - \mu_2\tau_2(0) = Q_0 \\ q(T) = \mu_1\tau_1(T) - \mu_2\tau_2(T) = Q_T \end{array} \tag{3.7}$$

These restrictions for q define a hyperplane in H^1, which we reference later as the **q-plane**. The functions $\tau_i(t)$ corresponding to each q would also form a hyperplane in L^2, which will be referred to as the $\boldsymbol{\tau_i}$**-plane**.

In terms of this new variable, the functionals we are going to analyze transform into

$$\begin{array}{rcl} S_1(T) &=& \lambda_1 e^{-\lambda_1 T}\int_0^T e^{\lambda_1 t}\left(\frac{\lambda_2 q + \dot{q}}{(\lambda_2-\lambda_1)\mu_2}\right)^2 dt, \\ S_2(T) &=& \lambda_2 e^{-\lambda_2 T}\int_0^T e^{\lambda_2 t}\left(\frac{\lambda_1 q + \dot{q}}{(\lambda_2-\lambda_1)\mu_1}\right)^2 dt. \end{array} \tag{3.8}$$

We note that the quantities $S_i(T)$ depend not only on the final time T but also on the function $q(t)$. When we need to emphasize this fact we write $S_i(T, q)$.

3.1.1 S_1 analysis

The first step, as explained above, is to analyze S_1. It is easy to see that $S_1(T)$ is convex with respect to q and $\dot{q}$ (the integrand must be "strongly convex", which is true since the integrand is a quadratic function). Hence the minimizing function will be a solution of the Euler-Lagrange equation

$$\partial_t f_{\dot{q}} = f_q,$$

where $f(t, q, \dot{q}) = \frac{e^{-\lambda_1(T-t)}(\lambda_2 q + \dot{q})^2}{(\lambda_2 - \lambda_1)^2 \mu_2^2}$. This leads to the following boundary value problem:

$$\begin{aligned} \ddot{q} + \lambda_1 \dot{q} + \lambda_2(\lambda_1 - \lambda_2) q &= 0, \\ q(0) &= Q_0, \\ q(T) &= Q_T. \end{aligned} \tag{3.9}$$

This yields a solution of the form

$$q(t) = C_1 e^{(\lambda_2 - \lambda_1)t} + C_2 e^{-\lambda_2 t}, \tag{3.10}$$

where the constants C_1, C_2 are chosen so that the solution satisfies the boundary conditions. After we have obtained $q(t)$ we easily obtain the control $\dot{\gamma}$. We can see here that if we had opted to keep the boundary conditions for $\dot{q}$, then we wouldn't have a continuous solution unless the conditions for $\dot{q}$ were compatible (i.e. redundant) with the conditions on q. This means that only in the latter case can the minimum be achieved with a certain control. In all other cases the minimum is not achievable. However, if we alter q on an ε small interval we would not change the integral expression in S_1 much. Therefore it is possible to produce a smooth control by altering q near 0 and T, so that the final state is within prescribed distance from the minimum. We will give complete consideration to this matter when we consider both relaxation modes. For later references we refer to the solution (3.10) as $q_1^\star(t)$.

By doing analysis similar to one above, we can find a minimum value for S_2. However, we must understand that a minimizer for S_2 is usually not a minimizer for S_1 and vice versa. Therefore to fully understand their interaction we must consider both relaxation modes at the same time.

3.1.2 Intermediate results

Before we start to investigate the second relaxation mode normal stress, we state and prove two theoretical results that will help us with our investigation. In case of general minimum inquiries, knowing whether the functional is convex or not is of great help. The following two lemmas provide theoretical ground for our coming analysis. For a short time we put aside our problem and work with somewhat more general integral functionals I_i.

Lemma 3.2. *Let*

$$\pi_f(u,v) = \int_0^T f(t)(\dot{u} + \lambda_2 u)(\dot{v} + \lambda_2 v)\, dt$$

$$\pi_g(u,v) = \int_0^T g(t)(\dot{u} + \lambda_1 u)(\dot{v} + \lambda_1 v)\, dt$$

$$I_1(v) = \pi_f(v,v) = \int_0^T f(t)(\dot{v} + \lambda_2 v)^2\, dt,$$

$$I_2(v) = \pi_g(v,v) = \int_0^T g(t)(\dot{v} + \lambda_1 v)^2\, dt,$$

with given positive continuous functions $f(t)$ and $g(t)$. Then there is a constant $\bar{C}_1$ and $\bar{C}_2$ such that $I_2(v) \geqslant \bar{C}_1 I_1(v)$ and $I_1(v) \geqslant \bar{C}_2 I_2(v)$ for every $v \in H_0^1(0,T)$.

Proof. We argue by contradiction. Assume there is a sequence v_n such that $\forall n$ we have $I_2(v_n) = 1$, but $I_1(v_n) \to \infty$. Simple algebraic manipulation leads to:

$$I_1(v_n) \leqslant \int_0^T 2f(t)[(\dot{v}_n + \lambda_1 v_n)^2 + (\lambda_2 - \lambda_1)^2 v_n^2]\, dt.$$

Since the first term has to be bounded, it follows that the L_2-norm of v_n tends to infinity.

On the other hand, let $w_n = \dot{v}_n + \lambda_1 v_n$. By assumption, the L^2-norm of w_n is bounded. This differential equation has the following solution:

$$v_n(t) = \int_0^t \exp(\lambda_1(\tau - t)) w_n(\tau)\, d\tau,$$

which implies that the L^2 norm of v_n can be bounded in terms of the L^2-norm of w_n. Hence, there exists $\bar{C}_1 = \sup(1/M)$ where $M \in \{M > 0 : I_1(v_n) \leqslant M\}$.

The existence of $\bar{C}_2$ is proved by interchanging the roles of I_1, I_2 in the proof above. □

Remarks:

1. It is not difficult to show that for any $C < \bar{C}$ the inequalities above are strict $\forall v \neq 0$.

2. The argument in this lemma also proves that each $I_i(v)$ is coercive in H_0^1 norm. E.g. setting $\lambda_1 = 0$ gives us

   $$I_1(v) \geqslant K\|\dot{v}\|^2 \geqslant M\|v\|_{1,2}^2.$$

   The last step is done using the Poincaré's inequality.

3. By normalizing the coefficients for I_1, I_2 we can rewrite the inequality in the following way:
$$\alpha_2 I_2(v) + \alpha_1 I_1(v) \geqslant 0, \tag{3.11}$$
where $\alpha_1^2 + \alpha_2^2 = 1$, $\alpha_2 \geqslant \frac{-\bar{C}_2}{\sqrt{1+\bar{C}_2^2}}$, $\alpha_1 \geqslant \frac{-\bar{C}_1}{\sqrt{1+\bar{C}_1^2}}$

Lemma 3.2 works only for the functions in H_0^1. However, we need a similar result for the function $q \in H^1[0,T]$ which does not always have vanishing boundary conditions. We note that the proof of lemma 3.2 is very symmetric in regard that the roles of the functionals I_1 and I_2 can be interchanged. For now we will consider only "one sided" cases, assuming that the reader understands that the similar result holds for the case I_1, I_2 are switched. The following lemma shows that a result similar to the one in lemma 3.2 holds for functions in H^1:

Lemma 3.3. *Let $q(t) \in H^1[0,T]$ satisfy the boundary conditions:*
$$q(0) = Q_0 \qquad q(T) = Q_T.$$
Then $\exists C_1, D_1$ so that $I_2(q) \geqslant D_1 + C_1 I_1(q)$.

Proof. Proving the claim is equivalent to proving that there is a finite minimum of $I_2(q) - C_1 I_1(q)$, i.e.
$$I_2(q) - C_1 I_1(q) \geqslant D_1.$$
Consider a functional $J_{C_1}(q)$
$$J_{C_1}(q) = I_2(q) - C_1 I_1(q).$$
This functional is convex iff
$$J_{C_1}(q+v) - J_{C_1}(q) - \delta J_{C_1}(q,v) \geqslant 0 \qquad \forall v \in H_0^1,$$
where $\delta J_{C_1}(q,v)$ is a Gateaux variation. By working out the details, we can show that
$$J_{C_1}(q+v) - J_{C_1}(q) - \delta J_{C_1}(q,v) = I_2(v) - C_1 I_1(v) \geqslant (\bar{C}_1 - C_1) I_1(v). \tag{3.12}$$
Thus, the question of convexity of J_{C_1} reduces to the result of lemma 3.2. Hence, for any $C_1 < \bar{C}_1$ the functional J_{C_1} is strictly convex, also, the modified functional
$$\tilde{J}_{C_1}(v) = J_{C_1}(q+v) - J_{C_1}(q) = J_{C_1}(v) + \delta J_{C_1}(q,v), \quad v \in H_0^1,\, q \,\text{is fixed} \tag{3.13}$$
is coercive and strictly convex. Since $\tilde{J}_{C_1}$ has clearly the same minimizer as J_{C_1} (after an addition of fixed element q), we conclude that the minimizer q_m satisfying $J_{C_1}(q_m) = \inf_q J_{C_1}(q) = D_1$ has to exist and be unique (see [19]). Consequently, we have
$$I_2(q) \geqslant D_1 + C_1 I_1(q) \quad \forall q \in H^1,\, q(0) = Q_0,\, q(T) = Q_T. \quad \square$$

There are two questions that we need to answer before we proceed: is the mapping $C_1 \to q_m$ continuous and what happens if $C_1 \to \bar{C}_1$? To investigate the continuity we analyze the minimizers for the family of functionals $J_{C_1+\varepsilon}$, when $|\varepsilon|$ is sufficiently small, so that $C_1+\varepsilon < \bar{C}_1$. Let q_m be the minimizer of the functional when $\varepsilon = 0$, i.e.

$$J_{C_1}(q_m) = \inf J_{C_1}(q)$$

Consider the modified functional

$$\tilde{J}_{C_1+\varepsilon}(q) = J_{C_1+\varepsilon}(q) - J_{C_1+\varepsilon}(q_m). \tag{3.14}$$

This expression can be rewritten as (using (3.12))

$$\tilde{J}_{C_1+\varepsilon}(q) = J_{C_1}(v) - \varepsilon I_1(v) - \varepsilon\delta I_1(q_m, v) = J_{C_1+\varepsilon}(v) - \varepsilon\delta I_1(q_m, v), \tag{3.15}$$

where $v = q - q_m$. This functional has a minimizer, say v_ε, that satisfies (see [19])

$$\pi_g(v_\varepsilon, u - v_\varepsilon) - (C_1+\varepsilon)\pi_f(v_\varepsilon, u - v_\varepsilon) \geqslant \frac{\varepsilon}{2}\delta I_1(q_m, u - v_\varepsilon) \quad \forall u \in H_0^1. \tag{3.16}$$

By letting $u = 0$ we arrive at

$$-\big(\pi_g(v_\varepsilon, v_\varepsilon) - (C_1+\varepsilon)\pi_f(v_\varepsilon, v_\varepsilon)\big) = -\tilde{J}_{C_1+\varepsilon}(v_\varepsilon) \geqslant -\frac{\varepsilon}{2}\delta I_1(q_m, v_\varepsilon). \tag{3.17}$$

As long as $C_1 + \varepsilon < \bar{C}$ we have $\tilde{J}_{C_1+\varepsilon}$ coercive, thus $\exists c_1,\, c_2 > 0$ so that

$$c_1\|v_\varepsilon\|_{1,2}^2 \leqslant \tilde{J}_{C_1+\varepsilon}(v_\varepsilon) \leqslant \frac{\varepsilon}{2}\delta I_1(q_m, v_\varepsilon) \leqslant \varepsilon c_2\|v_\varepsilon\|_{1,2}.$$

Therefore, if we let $\varepsilon \to 0$ we will have $\|v_\varepsilon\|_{1,2} \to 0$. Hence q_m depends on C_1 continuously.

The second question, however, doesn't have an easy answer. Nevertheless, we can analyze the situation numerically. For now, we will consider several scenarios, that might take place under different circumstances.

Let c_i be an increasing sequence so that $c_i \to \bar{C}_1$ in lemma 3.2, then for each c_i there exists a function $q_{m,i}$ that minimizes J_{c_i}, and hence

$$I_2(q) - c_i I_1(q) \geqslant D_i = I_2(q_{m,i}) - c_i I_1(q_{m,i}).$$

What happens in the limit is dependent on whether $\lim D_i$ exists. The sequence D_i must be a decreasing one, since $c_i > c_j \Rightarrow \forall q\ J_{c_i}(q) < J_{c_j}(q)$.

At this moment, there is no direct evidence that any of the following scenarios will never happen, thus we list all the possibilities:

A) $\lim \mathbf{D_i} = -\infty$
In this case we clearly have $I_1, I_2 \to \infty$, and therefore $||q_i||_2 \to \infty$. Thus, while trying to control the medium we will never reach the extreme case and stay always in the convexity region, having just one unique minimizer.

B) $\lim \mathbf{D_i} = \mathbf{\bar{D}} > -\infty$
In this case we can have the following sub-scenarios happen:

- **I)** $\lim I_1 = \infty$ and $\lim I_2 = \infty$. The conclusion is the same as in previous case.
- **II)** $\lim I_{1,2} < \infty$. In this case, the sequence q_i has a weakly convergent subsequence (since the sequence has bounded L_2 norm), that converges to a minimizer of $J_{\bar{C}_1}$. If this minimizer is unique, then we have just one pair of numbers I_1, I_2 belonging to a line $I_2 - \bar{C}_1 I_1 = \bar{D}_1$, and for all greater values of I_1 we can arrange I_2 to be as close to this line as we want. This is a direct consequence of the fact that $J_{\bar{C}_1}$ is convex, but not strictly. On the contrary, we can have infinitely many minimizers, with the only difference that there will be infinitely many pairs I_1, I_2 that lie on the line $I_2 - \bar{C}_1 I_1 = \bar{D}_1$ for all I_1 bigger than $\lim I_1(q_i)$.

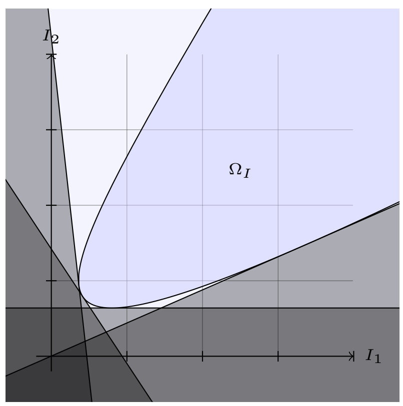


Figure 3.1: The graphic interpretation of the inequalities (3.18)

After similar manipulations as in (3.11), we can conclude that for the pairs α_1, α_2 that satisfy the inequalities in Remark 3 there exists a number D_α so that

$$J_\alpha(q) = \alpha_2 I_2(q) + \alpha_1 I_1(q) \geqslant D_\alpha \quad \forall q \in H^1,\ q(0) = Q_0,\ q(T) = Q_T. \tag{3.18}$$

By virtue of the above discussion the minimizer q_m of the functional $J_\alpha(q)$ depends continuously on the vector $\langle \alpha_1, \alpha_2 \rangle$ (assuming a Euclidian metric on $\mathbb{R}^2$).

Since this one-vector-parameter approach is more flexible and general than the approach we took in the proofs of lemmas (the one that uses two symmetric "one-sided" inequalities), we will use it for all subsequent calculations.

We can interpret the results of Lemma 3.3 (and of the family of inequalities (3.18)) in the following way: there is a family of lines that cut the plane into two halves: one contains the set Ω_I of all possible points $(I_1(q), I_2(q))$ and another has none (see Fig. 3.1).

3.1.3 S_2 analysis

Now we have all the tools necessary, and it is time to turn our attention to S_2. Since we know what possible S_1 values we can achieve, it seems natural to try to investigate the possible S_2 states when certain S_1 level is required to be reached. Using standard COV approach, this task is accomplished by minimizing (or maximizing)

$$S_2(T) + \alpha S_1(T) \tag{3.19}$$

for some suitable choice of a Lagrange multiplier α. This would give us a characterization of the extreme values of S_1, S_2.

However, since most of the analysis we do here fits into the framework of lemmas 3.2 and 3.3, we find it more convenient to consider the minimization of the linear combination of the following type:

$$\alpha_2 S_2(T) + \alpha_1 S_1(T). \tag{3.20}$$

Certainly (3.20) is more general than (3.19); it also has a number of advantages, one of which is easier numerical implementation. The redundancy of the coefficients α_1, α_2 is removed if we impose the requirements similar to those in (3.11). Also, this restrictions guarantee that the functional (3.20) is coercive and the minimization problem has a solution. The detailed discussion on the bounds for α_i is postponed until later. To the moment we assume that the values of α_i are within the coercivity region.

Applying the methods of Calculus of Variations we reduce the problem to the boundary value problem of the following form:

$$\begin{aligned} P(t)\ddot{q} + Q(t)\dot{q} + R(t)q &= 0, \\ q(0) &= Q_0, \\ q(T) &= Q_T, \end{aligned} \tag{3.21}$$

where:

$$\begin{aligned} P(t) &= \alpha_2 \frac{\lambda_2 e^{\lambda_2(t-T)}}{\mu_1^2} + \alpha_1 \frac{\lambda_1 e^{\lambda_1(t-T)}}{\mu_2^2}, \\ Q(t) &= \alpha_2 \frac{\lambda_2^2 e^{\lambda_2(t-T)}}{\mu_1^2} + \alpha_1 \frac{\lambda_1^2 e^{\lambda_1(t-T)}}{\mu_2^2}, \\ R(t) &= \alpha_2 \frac{\lambda_2 e^{\lambda_2(t-T)}(\lambda_1\lambda_2 - \lambda_1^2)}{\mu_1^2} + \alpha_1 \frac{\lambda_1 e^{\lambda_1(t-T)}(\lambda_1\lambda_2 - \lambda_2^2)}{\mu_2^2}. \end{aligned}$$

The existence and uniqueness of the solution here is guaranteed (see the previous subsection), as long as α_i satisfies the coercivity requirements. However, we can deduce by just looking at the equations, that in order to avoid singularities in the solution, we must have $P(t)$ to be bounded away from zero. Thus we must require either $P(t) > 0$ or $P(t) < 0$. However, we can clearly see that without loss of generality, we can restrict ourselves to the requirement of $P(t) > 0$ only. This puts restrictions on the values α_i can take; namely, the set of permissible α values can be characterized by (assuming $\lambda_2 > \lambda_1$, otherwise the inequalities have a slightly different form)

$$\alpha_1 = \cos\theta,\ \alpha_2 = \sin\theta,$$
$$-\cos^{-1}\left(\frac{\lambda_2\mu_2^2}{\sqrt{\lambda_1^2\mu_1^4+\lambda_2^2\mu_2^4}}\right) < \theta < \cos^{-1}\left(\frac{-e^{\lambda_2 T}\lambda_2\mu_2^2}{\sqrt{e^{2\lambda_1 T}\lambda_1^2\mu_1^4+e^{2\lambda_2 T}\lambda_2^2\mu_2^4}}\right). \tag{3.22}$$

The restrictions we have put on the α_i suggest that they might have something to do with coercivity restrictions in lemmas 3.2 and 3.3. However, these constants are different in nature. These bounds represent the region where the Weierstrass necessary condition for a minimum holds, thus the coercivity bounds have to lie within these bounds. Consult the example in Appendix B on the page 52 for detailed discussion on this topic. We also take a short look at this issue in the numerical experiments.

The equations above, as we might expect, incorporate the behavior of the equations (3.9). If we set $(\alpha_1, \alpha_2) = (1, 0)$ the system (3.21) becomes identical to (3.9), and therefore the solution we get is also identical to the one of (3.9). In other words, the analysis we performed two subsections above was just a special case of the current one.

Although we managed to "trace" the outside boundary of the set Ω_S , we still know nothing about its interior. Does it have any "bubbles", or is it "solid"? In answer to this question we prove that for any finite fixed S_1 value, there are the functions $q(t)$ that would continuously connect the maximum S_2 value with its minimum. In other words, show that the straight line that connects the minimum S_2 value with maximum S_2 value for any given fixed S_1 lies entirely within the set Ω_S.

Let's fix the value of $S_1 = A$, where A is a constant bigger than $\min S_1$. There will be two q functions on of which minimizes S_2 and another maximizes S_2. Let's denote these two functions by $q_m(t)$ and $q_M(t)$ respectively. The functional S_1 is a seminorm for $q(t)$, thus the set $B_A = \{q(t) : S_1 \leqslant A\}$ must be a convex set. If we restrict this convex set to the hyperplane

$$Q_{plane} = \{q(t) : q(0) = Q_0,\ q(T) = Q_T\}$$

we will end up with a convex set again. Let $\bar{q}(t)$ be any function in this hyperplane, i.e. any function that satisfies the given boundary condition. The functions $q_0(t) = q(t) - \bar{q}(t)$ will have zero boundary conditions for all $q \in Q_{plane}$. Thus they make a vector space inside the hyperplane. The set

$$V = \mathrm{Span}(\{q_{0,M},\ q_{0,m}\}), \quad q_{0,M} = q_M - \bar{q},\ q_{0,m} = q_m - \bar{q},$$

will make a two dimensional subspace (with proper choice of $\bar{q}$), that contains the segment:

$$q_0^\beta(t) = \beta\, q_{0,m}(t) + (1-\beta) q_{0,M}(t), \quad \beta \in [0, 1].$$

Because of convexity of B_A the segment q_0^β lies entirely within it, as well as within the intersection of V and B_A. Thus, if we manage to find a continuous transformation

$$\gamma(\beta, t) = \bar{q}(t) + C_\beta\, q_0^\beta, \quad \beta \in [0, 1]$$

that does the following:

$$S_1(T)|_{q=\gamma(\beta,t)} = A,$$

then we are done. Therefore, if we can find a function $\bar{q}$ that makes $q_{0,m}$ and $q_{0,M}$ linearly independent, we will turn the problem into a two dimensional one, which will always have a solution for C_β (see Fig. 3.2).

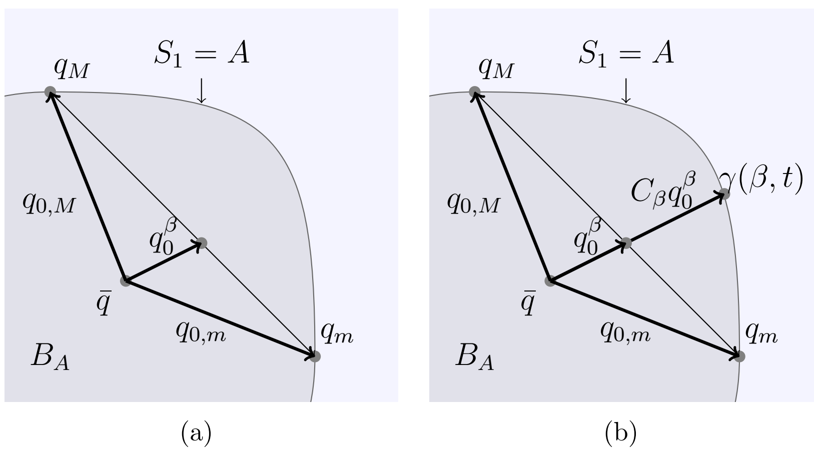


Figure 3.2: Projection of the segment onto the convex boundary

Can we find such a $\bar{q}$? The answer is yes, since the set of all the continuous functions that satisfy the boundary conditions $q(0) = Q_0$ and $q(T) = Q_T$ is clearly infinitely dimensional.

Since we can use this reasoning for any fixed S_1, we conclude that the interior of the set of attainable stresses in (S_1, S_2) plane is bounded only by the curves corresponding to extremal values of S_2.

Now, we almost ready to perform the numerical experiments. However the picture wouldn't be full unless we investigate what happens if the boundary conditions on τ_1, τ_2 that put boundary conditions on $\dot{q}$ in (3.5) are incompatible with the solution of (3.9) or (3.21).

So far, we used only the first equation of (3.5) for our analysis. Naturally, for almost all cases the solutions of (3.21) would produce (see (3.6)) functions τ_1, τ_2 that do not satisfy the boundary conditions with which we started. To resolve this issue, we must alter the solution $q(t)$. This will inevitably drive $S_1(T)$ and $S_2(T)$ away from the achieved minimum. The question is whether we can or cannot keep this drifting away controlled. The answer is yes.

We start in the usual way: let ε be a small positive number. The goal we need to achieve is the following (which we assume has not been reached yet):

$$\dot{q}(t)|_{t=0,T} = -\left(\mu_2\lambda_1\tau_1(t) - \mu_1\lambda_2\tau_2(t)\right)|_{t=0,T}.$$

Thus all we need to do here is to alter the initial and final values of $\dot{q}$. The solution to our problem is to add a correcting function v to q that is zero at the boundary (and for most of the time interval), but whose derivative is exactly the difference

$$\dot{v}(t) = -\left(\mu_2\lambda_1\tau_1(t) - \mu_1\lambda_2\tau_2(t)\right) - \dot{q}(t) \qquad \text{for}\, t = 0, T.$$

This inevitably will increase the values of S_1, S_2 since the new, corrected, q-function will not be the minimizer anymore, but if we keep the deviation small, then such a difference wouldn't matter much. Let $v \in C^\infty([0,T])$ be a function that satisfies the following conditions:

- $v(0) = v(T) = 0$,
- $v(t) = 0$ for $t \in [\varepsilon, T-\varepsilon]$,
- $\dot{v}(0) = -\left(\mu_2\lambda_1\tau_1(0) - \mu_1\lambda_2\tau_2(0)\right) - \dot{q}(0)$ and
  $\dot{v}(T) = -\left(\mu_2\lambda_1\tau_1(T) - \mu_1\lambda_2\tau_2(T)\right) - \dot{q}(T)$,
- $\sup_t |v(t)| \leqslant \varepsilon$ and $\sup_t |\dot{v}(t)| \leqslant \max(|\dot{v}(0)|, |\dot{v}(T)|)$.

In other words, we define v as a zero function with two bumps at the boundaries added. There are, of course, infinitely many ways to define such a function. However, any function that satisfies the restrictions will clearly do job. I.e. it will bring the boundary conditions of τ_i to desired levels. It will also provide the mean (by making ε smaller) to control the drift of S_1, S_2 from their theoretical extremal values. This means that the set achievable values of S_1, S_2 in this case is almost the same as the set Ω_S we obtained before. The only difference is that the boundary $\partial\Omega_S$ is no longer reachable. This result is another proof of the usefulness of the function q. It makes it easy to see that the initial and final conditions put on $\tau_i(t)$ only affect the reachable set of normal stresses through their difference, but not through their actual values.

3.2 Three relaxation modes.

In the case of three relaxation modes things are not as simple as in case of just two relaxation modes. However, the methods we employ here are closely related to those used previously. First of all, we will have to deal with more than one stress difference functions $q_{i,j}$. Here the subscripts are introduced to indicate what stress functions are chosen for making the stress difference function. These stress functions are not completely independent of each other, there is a set of differential equations that interconnects them together. Nevertheless, we shall see that calculus of variations still provides enough analysis tools even for this more complicated case. We shall also see that the analysis performed in this section is general enough to be easily applied to cases with more than three relaxation modes.

The notation and conventions, that are going to be used throughout this section are not much different from the notation in previous section. However to make things clear we list here the main changes:

A) We denote the shear stress' difference by $q_{i,j} = \mu_j \tau_i - \mu_i \tau_j$, thus the function q we used in the previous section becomes $q_{1,2}$.

B) The quantities S_i we introduced before have the same meaning here. We just add S_3 to this collection.

C) The initial and final values of stress difference functions are denoted $Q^0_{i,j}$ and $Q^T_{i,j}$ correspondingly.

As was done in the previous section, we concentrate on controlling the normal stresses while regarding the final shear stresses as given.

Since we have more than two relaxation modes, we need more than one shear stress difference function. Out of many options, we choose to tie all difference functions to the first relaxation mode, i.e. we are going to work only with functions $q_{1,j}$. These difference functions are not independent of each other, we have to account for the relation between them:

$$\frac{\dot{q}_{1,i} + \lambda_i q_{1,i}}{(\lambda_i - \lambda_1)\mu_i} = \tau_1 = \frac{\dot{q}_{1,j} + \lambda_j q_{1,j}}{(\lambda_j - \lambda_1)\mu_j}. \tag{3.23}$$

From this point, we have two options:

1. Solve the differential equation for $q_{1,j}$ and express the solution in terms of $q_{1,i}$. Then use the boundary conditions on $q_{1,j}$ as an integral constraint on the minimization problem later on.

2. Use τ_1 as an L_2 control in the system of two differential equations. Then use the ODE system (3.23) as a constraint on the problem.

The latter method proved to be more elegant and more suitable for subsequent generalization in cases of arbitrary relaxation modes. Thus we implement only the second one.

The analysis we performed in the two relaxation modes case was performed in several steps. The analysis of the three relaxation modes case is not very different in its outline. First, we shall derive the Euler-Lagrange equations that characterize the boundary of the set of achievable values $\{S_1, S_2, S_3\}$, and find what are the conditions on the parameters α for these equations to make sense. I.e. we will investigate the convexity of the functional we are going to minimize and the regions of nonsingularity of the solutions. Finally, we shall address the question whether the set of achievable normal stresses is solid (i.e. has nonempty interior with no "holes").

As was said before, we begin our analysis with consideration of all three relaxation modes included. In mathematical terms, we look for minimum values of linear combinations (this is similar to (3.19)):

$$\alpha_1 S_1 + \alpha_2 S_2 + \alpha_3 S_3. \tag{3.24}$$

It is easy to see that setting two of the α_i parameters to zero would yield to minimization of just one of the S_j quantities and setting just one of the α_i to zero would produce a minimizer of a linear combination of only two of the S_j. However the situation is different from the two modes case, since we didn't have any restrictions of the form (3.23). Also, we can notice that there is a great deal of redundancy when we introduce three α parameters instead of two. For example multiplying all of them by any nonzero constant would not change the problem at all. We can get around the redundancy of three parameters by restricting the α_i triples to be vectors of unit norm and whose dot product with a certain *a priori* chosen vector is nonnegative (the latter requirement is here to rule out the triples that are negatives of each other). We shall see later how to generalize the coercivity constrains introduced in (3.11) and the Weierstrass condition in (3.22).

The quantities S_i are defined as follows:

$$\begin{aligned} S_1(T) &= \lambda_1 e^{-\lambda_1 T} \int_0^T e^{\lambda_1 t} \left(\frac{\lambda_2 q_{1,2} + \dot{q}_{1,2}}{(\lambda_2 - \lambda_1)\mu_2} \right)^2 dt, \\ S_2(T) &= \lambda_2 e^{-\lambda_2 T} \int_0^T e^{\lambda_2 t} \left(\frac{\lambda_1 q_{1,2} + \dot{q}_{1,2}}{(\lambda_2 - \lambda_1)\mu_1} \right)^2 dt, \\ S_3(T) &= \lambda_3 e^{-\lambda_3 T} \int_0^T e^{\lambda_3 t} \left(\frac{\lambda_1 q_{1,3} + \dot{q}_{1,3}}{(\lambda_3 - \lambda_1)\mu_1} \right)^2 dt. \end{aligned} \tag{3.25}$$

The functions $q_{1,2}$ and $q_{1,3}$ depend on each other through the following differential equations:

$$\begin{aligned} \dot{q}_{1,2} &= -\lambda_2 q_{1,2} + \mu_2(\lambda_2 - \lambda_1)\tau_1, \\ \dot{q}_{1,3} &= -\lambda_3 q_{1,3} + \mu_3(\lambda_3 - \lambda_1)\tau_1. \end{aligned} \tag{3.26}$$

They act as a constraint on our minimization problem. Therefore we include them into the functional to be minimized with an appropriate Lagrange multiplier vector function $\boldsymbol{p}(t) = (p_{1,2}(t), p_{1,3}(t))$. Here we must use multiplier function, since the constraint in the form of ODE is different from a point constraint or integral constraint.

As was mentioned earlier, the Euler-Lagrange equations for the minimization problem will have a more elegant and compact form if we first substitute $\dot{q}_{1,j}$ in (3.25) by the right hand side of (3.26). The resulting equations now look like

$$\begin{aligned} S_1(T) &= \lambda_1 e^{-\lambda_1 T} \int_0^T e^{\lambda_1 t} (\tau_1)^2 \, dt, \\ S_2(T) &= \lambda_2 e^{-\lambda_2 T} \int_0^T e^{\lambda_2 t} \left(\frac{\mu_2 \tau_1 - q_{1,2}}{\mu_1} \right)^2 dt, \\ S_3(T) &= \lambda_3 e^{-\lambda_3 T} \int_0^T e^{\lambda_3 t} \left(\frac{\mu_3 \tau_1 - q_{1,3}}{\mu_1} \right)^2 dt. \end{aligned} \tag{3.27}$$

Bringing all these things together, we arrive at the following minimization problem:

$$\text{Minimize} \quad J(q) = \int_0^T f(t, q_{1,2}, q_{1,3}, \dot{q}_{1,2}, \dot{q}_{1,3}, \tau_1) \, dt \quad \text{where :}$$

$$\begin{aligned} f(t, \ldots, \tau_1) &= \alpha_1 F_1 + \alpha_2 F_2 + \alpha_3 F_3 + p_{1,2}(t) E_1 + p_{1,3}(t) E_2, \\ F_1 &= \lambda_1 e^{\lambda_1 (t-T)} (\tau_1)^2, \\ F_2 &= \lambda_2 e^{\lambda_2 (t-T)} \left(\tfrac{\mu_2 \tau_1 - q_{1,2}}{\mu_1} \right)^2, \\ F_3 &= \lambda_3 e^{\lambda_3 (t-T)} \left(\tfrac{\mu_3 \tau_1 - q_{1,3}}{\mu_1} \right)^2, \\ E_1 &= -\dot{q}_{1,2} - \lambda_2 q_{1,2} + \mu_2 (\lambda_2 - \lambda_1) \tau_1, \\ E_2 &= -\dot{q}_{1,3} - \lambda_3 q_{1,3} + \mu_3 (\lambda_3 - \lambda_1) \tau_1, \end{aligned} \tag{3.28}$$

subject to the following boundary conditions :

$$\begin{aligned} q_{1,i}(0) &= Q^0_{1,i} \quad i = 2, 3, \\ q_{1,i}(T) &= Q^T_{1,i}. \end{aligned}$$

The minimizer for this problem will have to satisfy the Euler-Lagrange equations that can be expressed in brief as:

$$\begin{aligned} \frac{d}{dt} D_{\dot{q}_{12}} f &= D_{q_{12}} f, \\ \frac{d}{dt} D_{\dot{q}_{13}} f &= D_{q_{13}} f, \\ \frac{d}{dt} D_{\dot{\tau}_1} f &= D_{\tau_1} f. \end{aligned}$$

We couple the equations above with the equations from (3.26) and arrive at

$$\begin{array}{rcl}\dot{p}_{1,2} & = & \lambda_2\left(\frac{2\alpha_2 e^{\lambda_2(t-T)}(-q_{1,2}+\mu_2\tau_1)}{\mu_1^2}\right),\\ \dot{p}_{1,3} & = & \lambda_3\left(\frac{2\alpha_3 e^{\lambda_3(t-T)}(-q_{1,3}+\mu_3\tau_1)}{\mu_1^2}\right),\\ 0 & = & -2\tau_1(\alpha_1\lambda_1\mu_1^2 e^{\lambda_1(t-T)}+\alpha_2\lambda_2\mu_2^2 e^{\lambda_2(t-T)}+\alpha_3\lambda_3\mu_3^2 e^{\lambda_3(t-T)})+\\ & & +2(\alpha_2\lambda_2\mu_2 e^{\lambda_2(t-T)}q_{1,2}+\alpha_3\lambda_3\mu_3 e^{\lambda_3(t-T)}q_{1,3})+\\ & & +(\lambda_1-\lambda_2)\mu_1^2\mu_2 p_{1,2}+(\lambda_1-\lambda_3)\mu_1^2\mu_3 p_{1,3},\\ \dot{q}_{1,2} & = & -\lambda_2 q_{1,2}+\mu_2(\lambda_2-\lambda_1)\tau_1,\\ \dot{q}_{1,3} & = & -\lambda_3 q_{1,3}+\mu_3(\lambda_3-\lambda_1)\tau_1.\end{array} \tag{3.29}$$

Add the boundary conditions from (3.28) and we get the closed system of equations. Notice that the third equation in (3.29) is not differential because there was no derivative of τ_1 involved. Thus we can directly solve for τ_1 from this equation unless

$$\sum_{i=1}^{3}\alpha_i\lambda_i\mu_i^2 e^{\lambda_i(t-T)}=0. \tag{3.30}$$

This is the Weierstrass necessary condition for a minimum similar to (3.22). We will therefore require this sum to be nonzero. Without loss of generality we may only asuume

$$\sum_{i=1}^{3}\alpha_i\lambda_i\mu_i^2 e^{\lambda_i(t-T)}>0. \tag{3.31}$$

Of course, we cannot expect the solution of (3.29) to be a minimizer of (3.24) unless we have (3.24) coercive. And of course, similarly to the case for two relaxation modes, the equation (3.31) does only provide the inexact bounds on the region of coercivity of (3.24). Therefore, we must make sure that we are well within the region of coercivity before drawing any conclusions from (3.29).

The investigation of the coercivity of the functional (3.24) follows exactly the same lines as in lemma 3.3, and since the difference is marginal we skip the discussion here. For rigorous derivation of the result for the case of arbitrary finite number of relaxation modes the reader might check the next section.

We'll also save till later the discussion on how to adjust the ends of $q_{1,\bullet}$ to satisfy the boundary conditions on $\tau_i(t)$. The result and the procedure are identical to the ones we have for two relaxation modes.

The last unresolved issue is the question of the interior of Ω_S. In the previous case, we used the fact that the points that have the same value of S_1 form a connected boundary of a convex set. To use the same idea in the current setting we would need the connectedness of the set of points that share the same $S_1,\ S_2$ values. However since this set is an intersection of the boundaries of two convex sets its connectedness is questionable. Therefore, we cannot use here the argument that worked in two relaxation mode case. At the moment, we can neither prove nor disprove any statements about the topology of the set Ω_S and its boundary.

3.3 n relaxation modes, $n > 2$

In this section we summarize and generalize all the results we have obtained for controllability of the quantities S_i.

Let the point $\boldsymbol{s} = (s_1, s_2, \ldots, s_n) \in \Omega_S$, then the functional

$$J_\alpha(\boldsymbol{s}) = \alpha_1 s_1 + \alpha_2 s_2 + \ldots + \alpha_n s_n \tag{3.32}$$

is a continuous linear mapping in variables α_i and $\boldsymbol{s}$. The functional J_α does have a minimum if Ω_S is bounded in the direction opposite to the vector $(\alpha_1, \alpha_2, \ldots, \alpha_n)$. Therefore the hyperplane defined by the equation for a fixed collection of α_i:

$$J_\alpha(\boldsymbol{s}) = \inf_{\bar{\boldsymbol{s}} \in \Omega_S} J_\alpha(\bar{\boldsymbol{s}}) \tag{3.33}$$

separates the state space $\{s_i\}$ into two halves, only one of which contains the set Ω_S. By varying the parameters α_i we would trace the boundary of a convex hull that contains Ω_S, possibly having multiple minimizers for some combinations of α_i and no minimum at all if there is no such separating plane can be found.

The state space we are dealing with consists of quantities

$$S_i = \lambda_i \int_0^T e^{-\lambda_i (T-t)} \tau_i^2(t)\, dt.$$

Since the shear stresses τ_i are not assumed to be continuous, they are only in L^2, we reformulate the problem in terms of the H^1 functions

$$q_{1,i} = \mu_i \tau_1 - \mu_1 \tau_i,$$

each of which satisfies

$$\tau_1 = \frac{\dot{q}_{1,i} + \lambda_i q_{1,i}}{(\lambda_i - \lambda_1)\mu_i}, \quad i = 2, 3, \ldots, n. \tag{3.34}$$

By virtue of the definition of $q_{1,i}$ each has to satisfy some boundary conditions. Following the previous section, we designate the quantities $Q^0_{1,i}$ and $Q^T_{1,i}$ to denote the boundary values at $t = 0, T$. With these functions we have now:

$$\begin{aligned} S_1(T, q_{1,2}) &= \lambda_1 \int_0^T e^{-\lambda_1 (T-t)} \left(\frac{\lambda_2 q_{1,2} + \dot{q}_{1,2}}{(\lambda_2 - \lambda_1)\mu_2} \right)^2 dt, \\ S_i(T, q_{1,i}) &= \lambda_i \int_0^T e^{-\lambda_i (T-t)} \left(\frac{\lambda_1 q_{1,i} + \dot{q}_{1,i}}{(\lambda_i - \lambda_1)\mu_1} \right)^2 dt, \quad i = 2, 3, \ldots, n. \end{aligned} \tag{3.35}$$

To trace the boundary $\partial\Omega_S$ of the achievable states of S_i we consider the problem of minimizing the quantity

$$J_\alpha(T, \boldsymbol{q}) = \alpha_1 S_1(T, q_{1,2}) + \alpha_2 S_2(T, q_{1,2}) + \alpha_3 S_3(T, q_{1,3}) + \ldots + \alpha_n S_n(T, q_{1,n}) \tag{3.36}$$

subject to the constraint (3.34). Again, to avoid redundancy in the coefficients α_i we require the vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ to have unit norm.

Naturally, we need to know for which values of α_i we have existence of the minimizers. The following theorem addresses this issue

Theorem 3.4. *The set of n-tuples $\boldsymbol{\alpha}$ (not necessarily of unit norm) for which the solution to the minimizing problem exists and is unique forms a convex cone in $\mathbb{R}^n$*

Proof. First, we need to make a change of variables. Let $\bar{q}_{1,i}$ be a collection (we denote this collection as $\bar{\boldsymbol{q}}$) of functions satisfying the boundary conditions and the restriction (3.34) for some L^2 function τ_1, any other function q can be expressed as $q_{1,i}(t) = \bar{q}_{1,i}(t) + v_i(t)$, where $v_i \in H_0^1$. However, since $\bar{q}_{1,i}$ are constrained, the functions v_i cannot be chosen arbitrary, as the following should hold:

$$\begin{gathered} \tfrac{\dot{q}_{1,i}+\lambda_i q_{1,i}}{(\lambda_i-\lambda_1)\mu_i} = \tfrac{\dot{q}_{1,j}+\lambda_i q_{1,j}}{(\lambda_j-\lambda_1)\mu_j} \Rightarrow \tfrac{\dot{v}_i+\lambda_i v_i}{(\lambda_i-\lambda_1)\mu_i} = \tfrac{\dot{v}_j+\lambda_i v_j}{(\lambda_j-\lambda_1)\mu_j} \Rightarrow \\ \Rightarrow v_j(t) = \int_0^t e^{-\lambda_j(t-s)} \tfrac{(\lambda_j-\lambda_1)\mu_j}{(\lambda_i-\lambda_1)\mu_i} (\dot{v}_i + \lambda_i v_i)\, ds. \end{gathered} \tag{3.37}$$

The constraint $v_i(0) = 0$ is satisfied automatically, but the other boundary condition is not. Thus we must enforce

$$0 = \int_0^T e^{-\lambda_j(T-s)} \frac{(\lambda_j-\lambda_1)\mu_j}{(\lambda_i-\lambda_1)\mu_i} (\dot{v}_i + \lambda_i v_i)\, ds, \quad j \neq i. \tag{3.38}$$

Since this is an intersection of finitely many kernels of linear functionals on H_0^1 the set of functions satisfying this constraint is clearly infinitely dimensional.

The functional J_α can be now expressed as

$$J_\alpha(T, \boldsymbol{q}) = J_\alpha(T, \bar{\boldsymbol{q}}) + \delta J_\alpha(T, \bar{\boldsymbol{q}}, \boldsymbol{v}) + J_\alpha(T, \boldsymbol{v}). \tag{3.39}$$

In the equation above $J_\alpha(T, \bar{\boldsymbol{q}})$ is fixed, and the Gâteaux variation $\delta J_\alpha(T, \bar{\boldsymbol{q}}, \boldsymbol{v})$ is linear with respect to $\boldsymbol{v}$. Thus the question of uniqueness and existence of minimizers of $J_\alpha(T, \boldsymbol{q})$ is reduced down to that of the coercivity of $J_\alpha(T, \boldsymbol{v})$.

The components of the vector function $\boldsymbol{v} = (v_2, \ldots, v_n)$ are not independent (see (3.37)). Therefore we might tie all v_j to v_2. Moreover, since the linear mapping (3.37) is one to one on H_0^1 we only need to show that

$$J_\alpha(T, \boldsymbol{v}) > C\|v_2\|_{1,2}. \tag{3.40}$$

This means that we can actually redefine the functional as $\bar{J}_\alpha(T, v_2) = J_\alpha(T, \boldsymbol{v})$, where $\boldsymbol{v}$ is constructed using (3.37) and v_2 satisfies (3.38).

We have that $S_1(T, v_2)$ and all of $S_i(T, v_i)$, $i > 1$ are coercive in terms of the respective functions v_i (c.f. Lemma 3.2). Also we have just proven the fact that for all v_i there exists a constant C_i so that $\|v_i\|_{1,2} \geqslant C_i \|v_2\|_{1,2}$. These two facts together allow us to conclude that if all of the $\alpha_i > 0$ then $\bar{J}_\alpha(T, v_2)$ is coercive. Since for each S_i the coercivity constants are positive, and we have only finitely many of them, for each coercive α_i combination there should exist a small neighborhood of it so that the coercivity still holds. Therefore we may have coercivity even if some of the α_i are negative. The coercivity of $\bar{J}_\alpha(T, v_2)$ necessitates that the functional (3.36) has a unique minimizer.

Finally, if we have two α_i combinations, say $\boldsymbol{\alpha}$ and $\bar{\boldsymbol{\alpha}}$ for which $\bar{J}_\alpha(T, v_2)$ is coercive, then any convex combination of the form $\theta \bar{J}_\alpha(T, v_2) + (1-\theta)\bar{J}_{\bar{\alpha}}(T, v_2)$ is also coercive. Therefore the set of allowable α_i combination does indeed form a convex cone.

□

With this theorem we can now state that for α_i within the coercivity region the functional $J_\alpha(T, \boldsymbol{q})$ has a unique minimizing vector function $(q^\alpha_{1,2}, \dots, q^\alpha_{1,n})$ and the dependence of $q^\alpha_{1,i}$ on $\boldsymbol{\alpha}$ is continuous. This in turn means that the minimizing states S_i also depend continuously on $\boldsymbol{\alpha}$ since they are continuous functionals on H^1.

Now we are ready to derive the Euler-Lagrange equations that describe the boundary $\partial\Omega_S$. We solve that equations (3.34) for derivatives $\dot{q}_{1,i}$ then substitute the solutions into each S_i:

$$\begin{aligned} S_1(T, q_{1,2}) &= \lambda_1 \int_0^T e^{-\lambda_1(T-t)} \left(\tau_1\right)^2 \, dt, \\ S_i(T, q_{1,i}) &= \lambda_i \int_0^T e^{-\lambda_i(T-t)} \left(\frac{\mu_i\tau_1 - q_{1,i}}{\mu_1}\right)^2 \, dt, \quad i = 2, 3, \dots, n. \end{aligned} \tag{3.41}$$

We also include them into J_α as a constraint on the minimization problem:

$$\text{Minimize} \quad J_\alpha(T, \boldsymbol{q}) = \int_0^T f(t, q_{1,2}, \dot{q}_{1,2}, \dots, q_{1,n}, \dot{q}_{1,n}, \tau_1) \, dt \quad \text{where :}$$

$$\begin{aligned} f(t, \dots, \tau_1) &= \textstyle\sum_{i=1}^n \alpha_i F_i + \sum_{i=2}^n p_{1,i}(t) E_i, \\ F_1 &= \lambda_1 e^{\lambda_1(t-T)} \left(\tau_1\right)^2, \\ F_i &= \lambda_i e^{\lambda_i(t-T)} \left(\frac{\mu_i\tau_1 - q_{1,i}}{\mu_1}\right)^2, \\ E_i &= -\dot{q}_{1,i} - \lambda_i q_{1,i} + \mu_i(\lambda_i - \lambda_1)\tau_1, \end{aligned} \tag{3.42}$$

subject to the following boundary conditions :

$$\begin{aligned} q_{1,i}(0) &= Q^0_{1,i}, \quad i = 2, 3, \dots, n, \\ q_{1,i}(T) &= Q^T_{1,i}. \end{aligned}$$

This yields to the following system whose solutions describe the the boundary $\partial\Omega_S$:

$$\begin{aligned} \dot{p}_{1,i} &= 2\alpha_i\lambda_i\mu_1^{-2} e^{-\lambda_i(T-t)}(-q_{1,i} + \mu_i\tau_1), \\ \dot{q}_{1,i} &= -\lambda_i q_{1,i} + \mu_i(\lambda_i - \lambda_1)\tau_1, \\ 0 &= -2\tau_1 \left(\textstyle\sum_{i=1}^n \alpha_i\lambda_i\mu_i^2 e^{-\lambda_i(T-t)}\right) + \\ &\quad + \textstyle\sum_{i=2}^n \left(2\alpha_i\lambda_i\mu_i e^{-\lambda_i(T-t)} q_{1,i} + (\lambda_1 - \lambda_i)\mu_1^2\mu_i p_{1,i}\right), \\ q_{1,i}(0) &= Q^0_{1,i}, \\ q_{1,i}(T) &= Q^T_{1,i}. \end{aligned} \tag{3.43}$$

Again, as was done in earlier sections, we must require $\sum_{i=1}^{n} \alpha_i \lambda_i \mu_i^2 e^{-\lambda_i (T-t)} > 0$; otherwise the minimizing solution doesn't exist. This requirement represents a special case of the Wierstrass necessary condition for a minimum. We provide the formal derivation below.

The condition requires a function $\mathcal{E}(t, \boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{z})$ to be nonnegative $\forall \boldsymbol{z} \in \mathbb{R}^{n-1}$. The function $\mathcal{E}$ is defined as

$$\mathcal{E}(t, \boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{z}) = f(t, \boldsymbol{q}, \dot{\boldsymbol{q}}) - f(t, \boldsymbol{q}, \boldsymbol{z}) - \partial_{\boldsymbol{z}} f(t, \boldsymbol{q}, \boldsymbol{z}) \circ (\dot{\boldsymbol{q}} - \boldsymbol{z}), \tag{3.44}$$

where f is[1]

$$f(t, \boldsymbol{q}, \boldsymbol{\xi}) = \lambda_1 e^{-\lambda_1 (T-t)} \left(\frac{\lambda_2 q_{1,2} + \xi_2}{(\lambda_2 - \lambda_1)\mu_2} \right)^2 + \sum_{i=2}^{n} \lambda_i e^{-\lambda_i (T-t)} \left(\frac{\lambda_1 q_{1,i} + \xi_i}{(\lambda_i - \lambda_1)\mu_1} \right)^2. \tag{3.45}$$

The condition therefore reads

$$0 \leqslant \mathcal{E}(t, \boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{z}) = \alpha_1 \lambda_1 \mu_1^2 e^{-\lambda_1 (T-t)} (\dot{q}_{1,2} - z_2)^2 + \sum_{i=2}^{n} \alpha_i \lambda_i \mu_i^2 e^{-\lambda_i (T-t)} (\dot{q}_{1,i} - z_i)^2. \tag{3.46}$$

As we have seen before, the functions $\dot{q}_{1,i}$ and z_i are not independent of each other. Thus the condition does not need to be enforced on all $\boldsymbol{z} \in \mathbb{R}^{n-1}$, but only on a special subset. In one of these special cases we have a condition:

$$0 \leqslant \sum_{i=1}^{n} \alpha_i \lambda_i \mu_i^2 e^{-\lambda_i (T-t)} \tag{3.47}$$

The last issue we need to address to complete the picture is the issue of meeting the boundary conditions for $\tau_i(0)$ and $\tau_i(T)$. For this purpose we only alter $q_{1,2}$ in the exactly the same way we did it in the case of two relaxation modes. The only requirement is that the correcting function v must also satisfy (3.38). Now we have the boundary conditions for τ_1 and τ_2 corrected. For any other τ_i we have the boundary conditions for $q_{1,i}$ preserved and since the time derivative $\dot{q}_{1,i}$ depends on $q_{1,i}$ and τ_1 we have that the desired boundary conditions for any other τ_i are met automatically.

[1] just for convenience we use shifted numbering of components of $\boldsymbol{\xi} = (\xi_2, \xi_3, \ldots, \xi_n)$

Chapter 4

Numerical results

In this subsection we perform numerical experiments to investigate the dependence of the set Ω_S on model parameters and initial conditions for stresses. We present the numerical results only for cases of two and three relaxation modes. The cases of more than three relaxation modes are not very different from the case of three relaxation modes, but very cumbersome to visualize.

4.1 Two relaxation modes

We employ two different approaches here. One of them is the use of finite differences methods. In other words, we discretize the Euler-Lagrange equations (3.21), then run the simulation for different values of $\alpha_{1,2}$ to trace the boundary of the set Ω_S. We repeat the procedure several times for different parameters. The advantages of this method is robustness, short running time and easy implementation. The biggest disadvantage the fact that the Euler-Lagrange equations do not allow for easy analysis of coercivity bounds on $\alpha_{1,2}$.

To overcome this issue, we turn to Rayleigh-Ritz analysis. Instead of discretizing the Euler-Lagange equations, we discretize the minimization problem (3.18)[1]. We do so by minimizing (3.18) over a finite dimensional subset of H^1. In our case this subset will be the set of piecewise linear functions with fixed step size. This transforms the the problem (3.18) into a finite dimensional one, that can be rewritten as the minimization of a matrix problem. We employ the decomposition of (3.18) similar to that in (3.13) to get

$$\tilde{J}_\alpha(v) = J_\alpha(q+v) - J_\alpha(q) = J_\alpha(v) + \delta J_\alpha(q,v). \tag{4.1}$$

The loss of coercivity will be indicated by the presence of negative eigenvalues in the matrix representation of $J_\alpha(v)$ in (4.1). We shall investigate this matter later in this section.

[1]By $J_\alpha(v)$ we now denote the the functional (3.20), replacing quantities I_i with S_i

At first let's look at the minimizing functions $q(t)$ for different values of α_i.

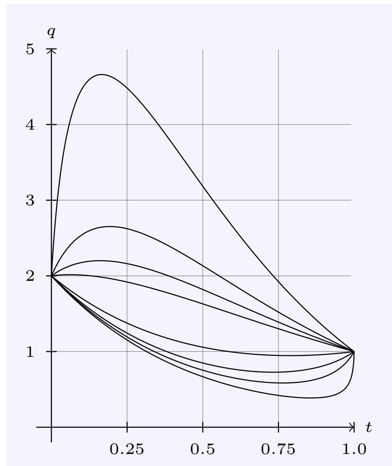


Figure 4.1: q-functions for different α parameters

On the Figure 4.1 we can see that as α gets closer to the extreme values, the minimizer exhibits close to singular behavior. This makes the normal stresses very large, which is consistent with the well known fact that the UCM model predicts the unlimited growth of the stresses under high strain rates. Since the parameters $\alpha_{1,2}$ are not the part of the model (they are only complementary Lagrangian multipliers) we postpone the discussion of the coercivity bounds till later. For now we will look at how the model parameters affect the shape of the set of attainable stresses Ω_S.

Table 4.1:

Var	Value
T	1
$q(0)$	2
$q(T)$	1
λ_1	1
λ_2	2
μ_1	1
μ_2	2

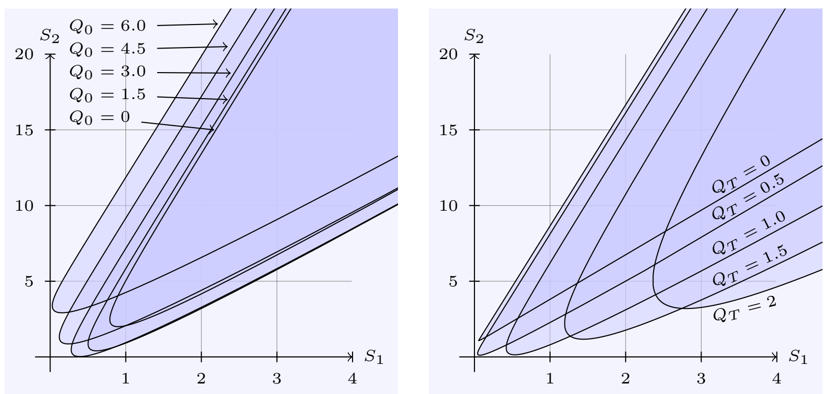
The parameters that affect the shape of Ω_S are λ_i, μ_i, the time T alloted for control and the boundary conditions Q_0, Q_T. As we shall see, the physical parameters have influence on the slopes of the bounding cone[2], and the initial conditions determine the location of the vertex of this cone and how close the attainable set boundary is to the boundary of the cone.

In the first series of experiments we investigate the relation between the form of the attainable set Ω_S and initial and final conditions. Since the physical parameters do not interfere qualitatively in this process, we will not vary them throughout the experiments. In the Table 4.1 the values that we used as the base values are listed. First, we are going to vary the initial condition on q. We set the values of Q_0 to range from 0 to 6 with increments of 1.5. The sets in the Fig. 4.2(a) tend to change location only, but since we require the relatively big shear stress difference at the end, the variations do not help much with "sharpening" the set Ω_S.

[2]We saw that the set Ω_S is bounded by a series of lines (hyperplanes in the case of more than 2 dimensions). The lines (hyperplanes) that correspond to the extreme values of α_i cut the cone in the S-space. Clearly this cone contains Ω_S, and the boundary of Ω_S converges to the cone boundary as we let $S_i \to \infty$

One curious thing that happened is the fact that the set Ω_S for the zero initial stress difference (the rightmost set on the figure) is contained completely in the next two sets for bigger initial difference. Thus, if we start from the relatively calm flow with low shear stress difference that needs to be driven to relatively big shear stress difference, we will have to pay the price of overall increased normal stresses.

We have quite the different situation when varying the final value Q_T of q. Here (Fig. 4.2(b)) we can see that both the location and "sharpness" of the set Ω_S vary greatly. The values Q_T go from 0 to 2 with increments of 0.5. Here, the effect of lowering the final shear stress difference is clearly seen. We can see the considerable sharpening of the set Ω_S and also we see that the closer to the origin states become achievable.

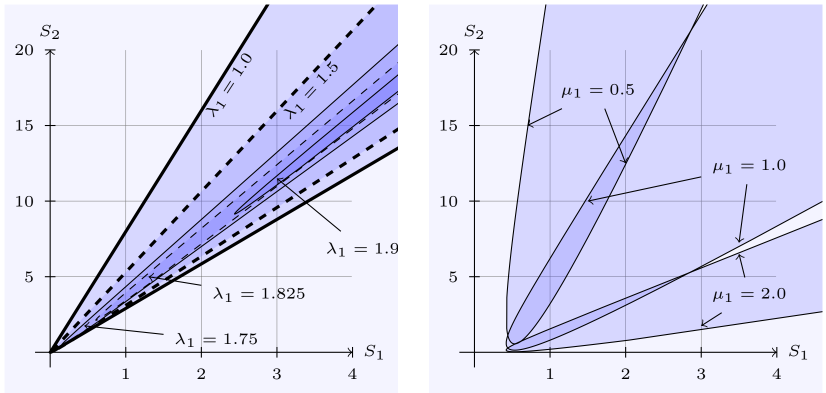


(a) Q_0 variations

(b) Q_T variations

Figure 4.2: The sets Ω_S for various Q_0, Q_T

In all these experiments we have seen that the slopes of the sets Ω_S at infinity are the same. This is a direct consequence of the fact that the convexity of the minimized functional depends only on values λ_i, μ_i and T. Having considered the effect of boundary conditions variations, we next turn our attention to the variations of physical parameters. Since the problem is symmetrical in terms of dependence on λ_i and μ_i, I chose to vary only λ_1 and μ_1 in this set of experiments. The variation of λ_1, especially when λ_1 gets close to λ_2, leads to results that are off the scale when the initial and final stress difference are relatively big. Therefore we opted to temporarily make them very small ($Q_0 = 0.2$ and $Q_T = 0.1$). This can be explained in the following way: when λ_i are close, by virtue of the equations (3.1) that govern τ_i, the respective τ-solutions start to behave in very similar fashion. The closer the constants are in their value, the harder it is to make the solutions to deviate from each other and thus control the shear stress difference to the desired level. Therefore, if we keep the stress difference big, it will require more aggressive control, consequently making the normal stresses big.



(a) λ_1 variations

(b) μ_1 variations

Figure 4.3: The sets Ω_S for various λ_1, μ_1

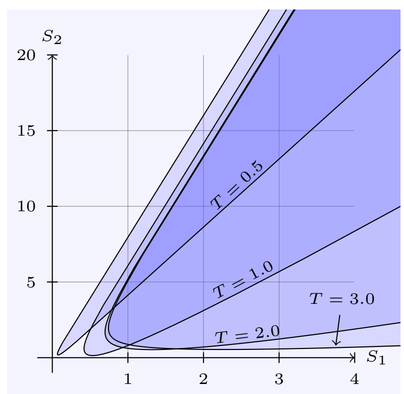


Figure 4.4: The sets Ω_S for various time for control T

The slopes of the boundary curve at infinity are defined by the coercivity bounds which in turn are required to lay within the region defined by the Weierstrass necessary condition (3.22). Thus we can predict the results of the subsequent experiments by analyzing the Weierstrass condition. As was expected, getting λ_1 close to or far from λ_2 has the effect of contracting or broadening the set Ω_S, and the variations of μ_1 cause the rotations (or skewing effect) of the attainable set.

We observe a different effect when we vary the time for control T. In the Figure 4.4 we see that only one of the limiting slopes varies. Also the location of the set changes slightly. This effect is predictable since in the Weierstrass conditions only one end of the admissible θ region depends on T, cf. (3.22). This gives a ground for speculation that the actual coercivity bounds might also depend on T at only one end.

Now, it is time to investigate the convexity of the functional and what happens if we go close to the boundary the region described by (3.22). As was mentioned before, the inequalities (3.22) only provide the region where the Weierstrass necessary condition for the minimum holds. The counter example in Appendix B shows that it might be possible that if $\lambda_i > 2$ the functional (3.20) would become nonconvex before we reach the boundary of the region described by (3.22).

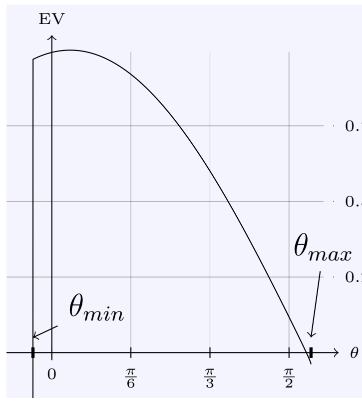


Figure 4.5: The eigenvalues vs. θ

To investigate such possibility, we employ the second method mentioned in the beginning of this subsection. Namely we approximate the H_0^1 space by piecewise linear functions. This will convert our functional $J_\alpha(v)$ in (4.1) into a matrix bilinear form. Nonconvexity of the functional on this approximate finite dimensional space will be indicated by presence of negative eigenvalues of the matrix in the approximate bilinear form. In the Fig. 4.5 we can see that in the case $\lambda_1 = 4.0$, $\lambda_2 = 8$ the minimum eigenvalue drops below zero well before the boundary θ_{max} of (3.22) is achieved. The numerical solution in this case tends to infinity as α approaches the coercivity boundary, just as it did in the case $\lambda_i < 2$ when α was approaching the boundary of (3.22), so the overall shape of the set Ω_S is preserved. In other words, this case fall into category A) or B.I) (see pg. 21). We note that the apparent drop of the eigenvalues as θ approaches θ_{min} is a result of eigenvalue crossing and the graph is continuous despite its deceptive appearance.

4.2 Three relaxation modes

The numerical procedure changes slightly in this case since the equations we are solving have qualitative changes. First, the system (3.29) has boundary condition on only two equations for $q_{1,i}$ functions; the equations for auxiliary functions $p_{1,i}$ assume free boundary conditions. The standard "shooting" procedure is used here, with a slight modification. Since the final values $q_{1,i}(T)$, $p_{1,i}(T)$ depend linearly on the initial conditions, we only need to solve the equations 5 times for each α_i combination (4 to investigate the mapping and 1 for final run). The standard Runge-Kutta solver performs rather well in this setting. The numerical procedure that is used for the simulation generalizes easily to the cases of more than three relaxation modes.

Another issue is how to select the values for α_i. In a two dimensional setting we could simply tie everything to one angular parameter θ that is bounded at two ends. In three or more dimensions we can use the fact that the set of all values α_i that yield the minimizer is a convex cone, and therefore is an intersection of series of half spaces. We use this fact to cut the unit sphere in $\mathbb{R}^3$ to describe the set of $\boldsymbol{\alpha}$ that are used in the numerical experiment. A graphical illustration of how this process works for Weierstrass necessary condition is on the Figure 4.6. The figure also explains how to implement the automatic filtering of the α_i

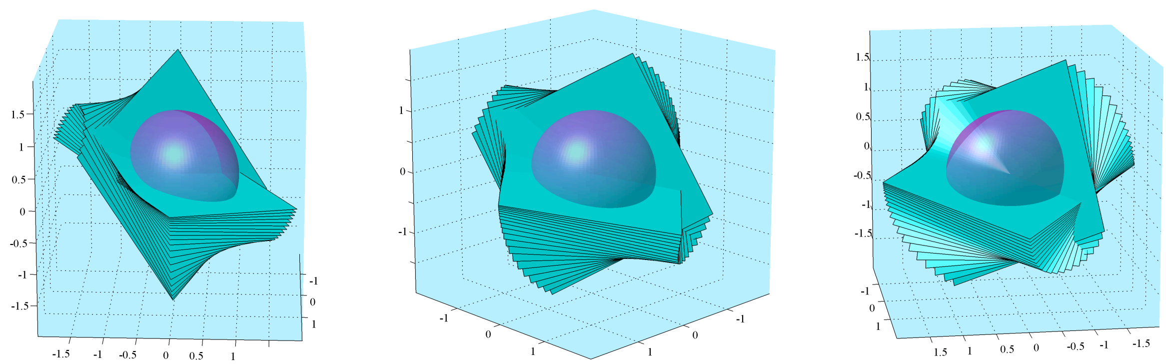


Figure 4.6: Weierstrass necessary condition

values: one simply needs to evaluate a series of dot products of $\boldsymbol{\alpha}$ vectors with the normal vectors that describe the cut off planes. The simulation is then performed only if all the dot products are positive.

Table 4.2:

Var	Start	End		Var	1	2	3
t	0	$T = 8$		λ_i	1	2	3
$q_{1,2}$	2	5		μ_i	1	2	3
$q_{1,3}$	2	6					

We perform the series of simulations similar to those in the previous section. Except for the presence of another dimension, the sensitivity of the achievable set Ω_S to changes of model parameters follows (qualitatively) the results obtained before. The parameters that are taken as base values are listed in the Table 4.2. One can notice several changes (apart from the presence of one more set of physical constants): the final time T and $q_{1,i}(T)$ has been increased for illustrative purposes. The explanation is simple: in the case of three relaxation modes, the additional restriction that has been put on the function $q_{1,2}$ made the set Ω_S extremely thin. Thus these measures has been taken to widen it.

We start, as before, with the variations of the boundary conditions. However, this series of experiments brought a surprise. The variations of the initial condition in the previous section resulted in the changes of the location of the set Ω_S, in these experiments. Since the scale is much larger and the time for control is bigger, the changes of location are indistinguishable for very large range of initial conditions. For the record, we did run this experiment for smaller time for control T and the changes of location were much more pronounced. However since the resulting pictures turned out to be of very poor quality we omit them here.

The variations of the final conditions yield predictable results. For smaller final values the set Ω_S becomes sharper and closer to the origin, and does the opposite if we increase the final values. The values used for the simulation are listed below:

Experiment No. :	1	2	3
$q_{1,2}(T) =$	5	8	12
$q_{1,3}(T) =$	6	10	16

The results of the simulation are shown in the Fig. 4.7, where the base case is shaded red while the others are blue. The four parts of the figure correspond to four different points of view on the sets. Just as before, the slopes of the set Ω_S at infinity are not changed. This is predictable since the coercivity of the minimized functional does not depend on the boundary conditions.

The variations of the model parameters also change the set in the predictable way. This time the slopes at infinity do change, since they affect the coercivity, thus affect the admissible values of α_i. Since we do not assign any real physical sense to the values of the constants, there is no necessity of varying all of the model parameters and trying to get the representative of all possible combinations. For variations of relaxation parameter λ_i I chose to vary only λ_3. The reason for this is simple: as there will be one maximum λ and one minimum λ only the relative position of the λ in the middle matters. Thus we can opt to vary just one of them because there will be no qualitative differences. There were four experiments with the following values of λ_i:

Experiment No. :	1	2	3	4
$\lambda_1 =$	1	1	1	1
$\lambda_2 =$	2	2	2	2
$\lambda_3 =$	2.5	3	4.5	6

The resulting sets Ω_S are shown in the Fig. 4.8. The behavior is similar to that of the previous section. When the values of λ_2 and λ_3 are close (in the first experiment) the set Ω_S shrinks and pulls further away from the origin. It widens and gets closer to the origin when the relaxation parameters are far from each other.

The variations of the viscosity parameter μ_i also do not bring any surprises. The sets Ω_S do change their slopes in a way similar to the two relaxation modes case. Instead of shrinking or expanding we observed when we varied λ_3 the set undergoes a skewing transformation. This is in agreement of the results obtained in the previous section. Since the variations of different μ constants are very predictable, for the sake of savings space and not introducing chaos I opted to vary just two of them, namely μ_2 and μ_3. The list of used μ_i values is shown

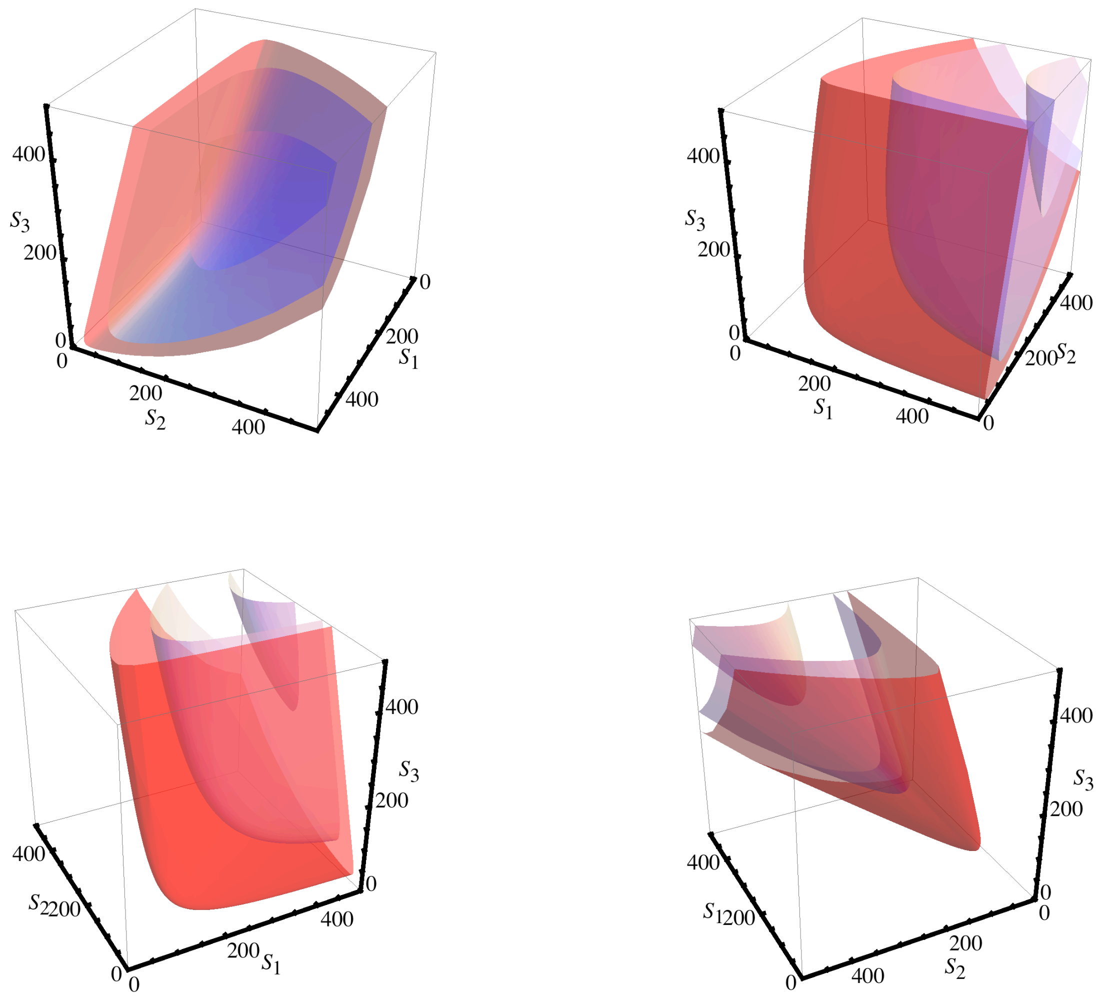


Figure 4.7: Variations of the final values $q_{1,i}(T)$

below:

Experiment No. :	1	2	3
$\mu_1 =$	1	1	1
$\mu_2 =$	2	2	2.5
$\mu_3 =$	2.5	3	3

Finally, we conduct the experiments when the time for control varies. As we might expect, some of the slopes of the set Ω_S will change, some will not. Since the set doesn't change much after time $T = 8$, but instead brings the numerical stability problems, we only run the

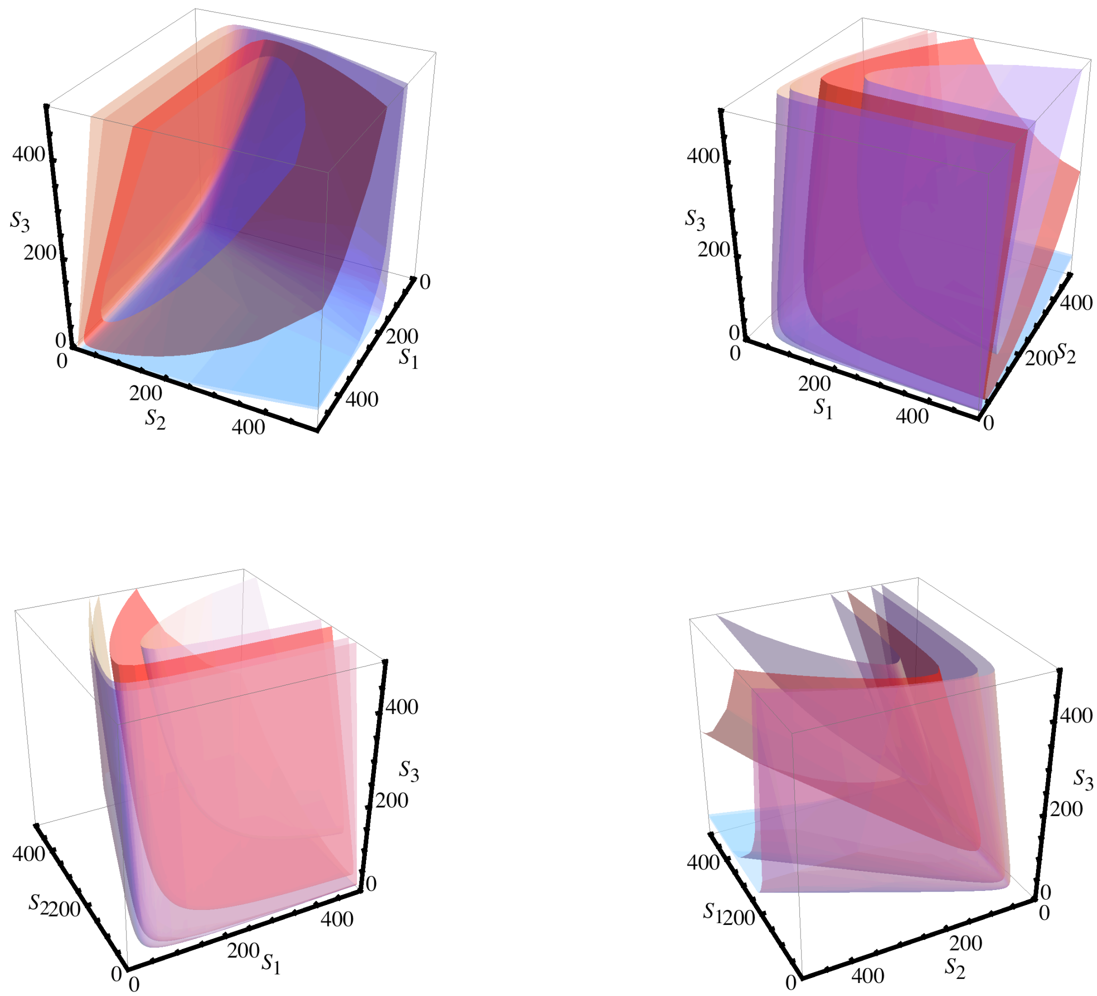


Figure 4.8: Variations of the relaxation parameter λ_3

simulation for smaller times, namely $T = 2, 3, 8$. Such strange choice of times is dictated by the dependence of the set on the time for control. For values $T > 4$ the set doesn't change much, however, for smaller times the set changes rapidly, shrinking to a very thin plate far away from the origin at $T = 1$. The calculated sets Ω_S are shown on the Figure 4.10, this picture doesn't have the base case set colored red for clarity purposes, the base set in the figure is the widest one.

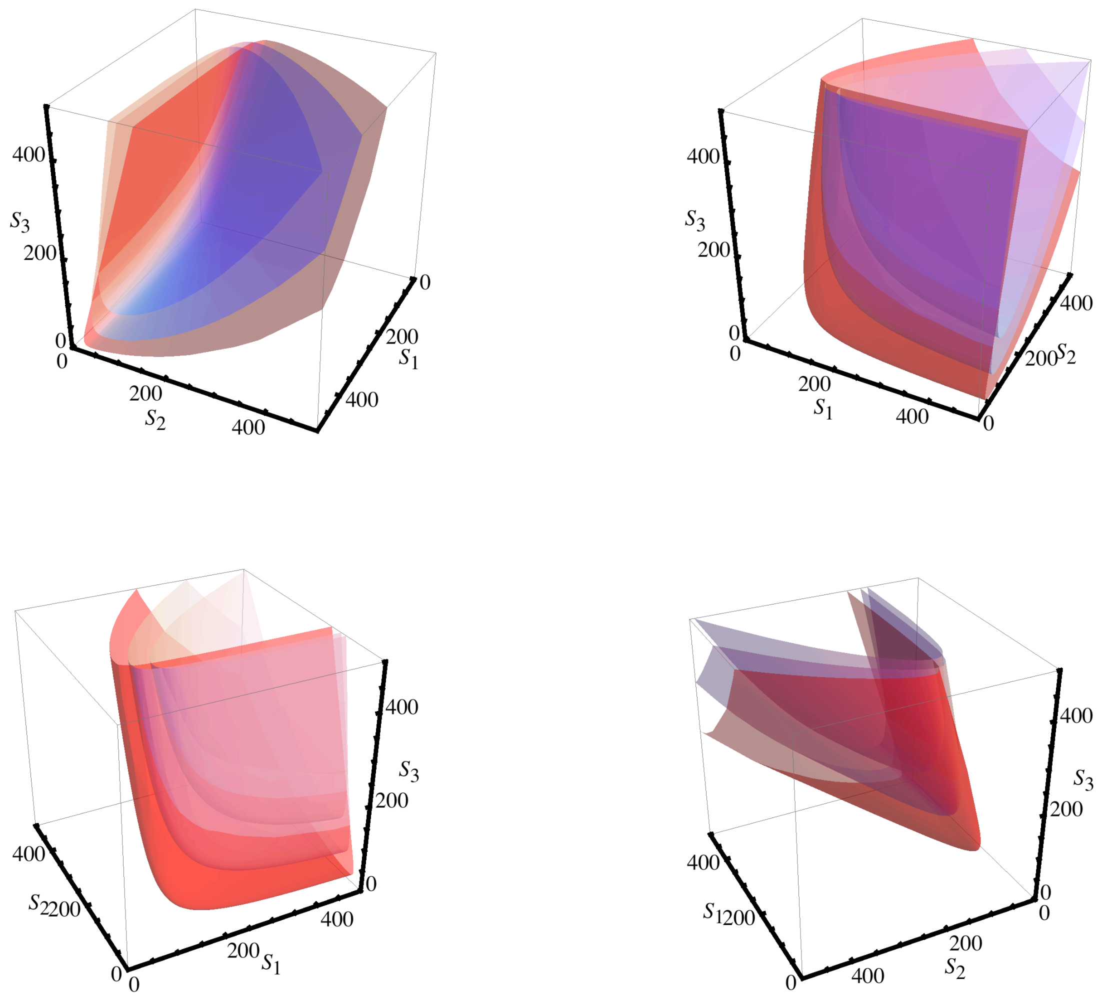


Figure 4.9: Variations of the viscosity parameter μ_i

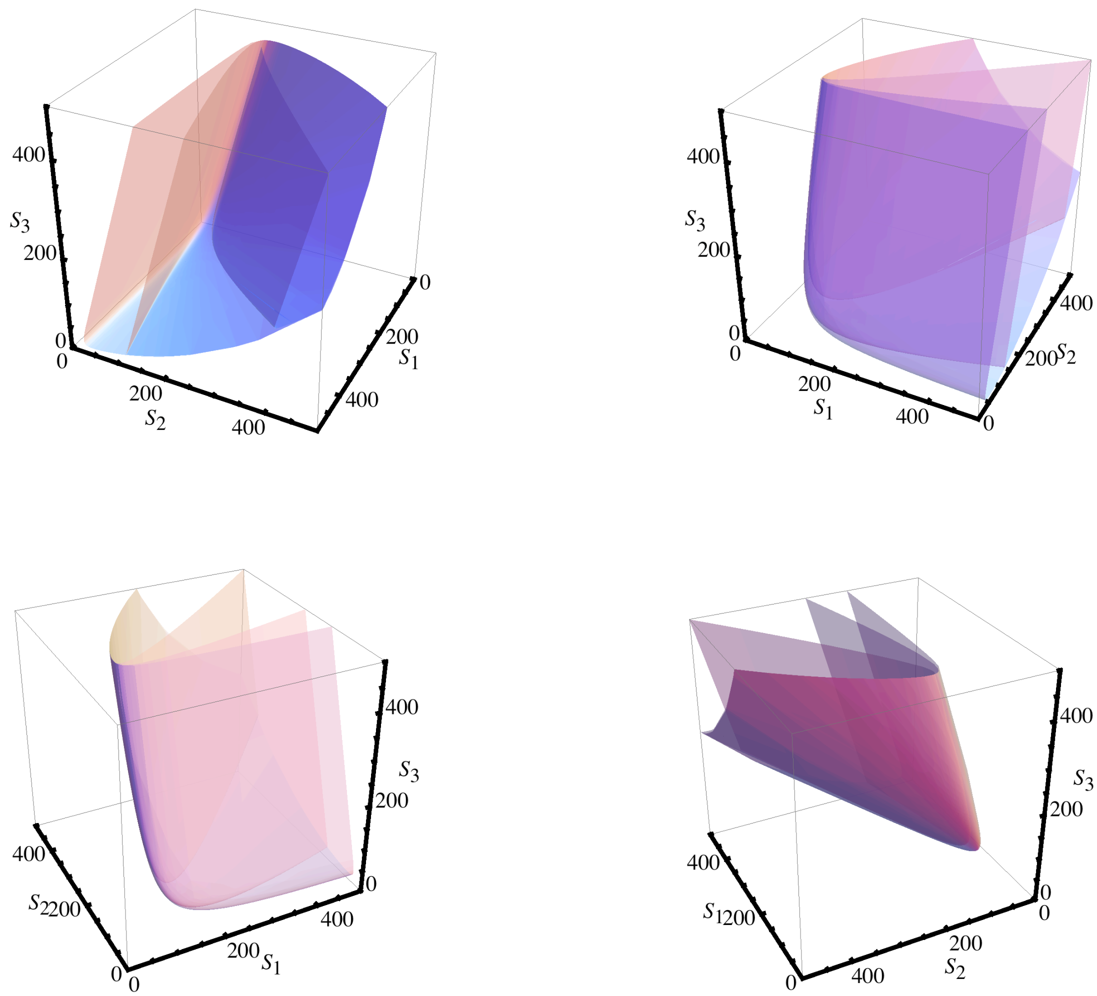


Figure 4.10: Variations of the time for control T

Bibliography

[1] S.A. Avdonin and S.A. Ivanov. *Families of exponentials: the method of moments in controllability problems for distributed parameter systems*. Cambridge University Press, 1995.

[2] Ben M. Chen, Zongli Lin, and Yakov Shamash. *Linear Systems Theory. A structural decomposition approach*. Birkhäuser, 2004.

[3] G. Chen, W.H. Mills Jr, and G. Crosta. Exact controllability theorems and numerical simulations for some nonlinear differential equations. *SIAM Journal on Control and Optimization*, 19:765, 1981.

[4] Earl A. Coddington and Robert Carlson. *Linear Ordinary Differential Equations*. Society for Industrial and Applied Mathematics, 1997.

[5] S. Dolecki and D.L. Russell. A general theory of observation and control. *SIAM Journal on Control and Optimization*, 15:185, 1977.

[6] A. Doubova, E. Fernandez-Cara, and M. Gonzalez-Burgos. Controllability results for linear viscoelastic fluids of the Maxwell and Jeffreys kinds. *Comptes Rendus de l'Academie des Sciences Series I Mathematics*, 331(7):537–542, 2000.

[7] A.V. Fursikov and O.Y. Imanuvilov. Exact controllability of the Navier-Stokes and Boussinesq equations. *Russian Mathematical Surveys*, 54(3):565–618, 1999.

[8] G.W. Haynes and H. Hermes. Nonlinear controllability via Lie theory. *SIAM Journal on Control*, 8:450, 1970.

[9] R. Hermann and A. Krener. Nonlinear controllability and observability. *IEEE Transactions on Automatic Control*, 22(5):728–740, 1977.

[10] L.F. Ho. Observabilite frontiere de l'équation des ondes [Boundary observability of the wave equation]. *CR Acad. Sci. Paris Ser. I Math*, 302:443–446, 1986.

[11] V. Jurdjevic and HJ Sussmann. Controllability of nonlinear systems. *J. Differential Equations*, 12:95–116, 1972.

[12] V. Komornik. *Exact controllability and stabilization*. Masson Paris, 1994.

[13] V. Komornik and P. Loreti. *Fourier series in control theory*. Springer, 2005.

[14] K. Kunisch and X. Marduel. Optimal control of non-isothermal viscoelastic fluid flow. *Journal of Non-Newtonian Fluid Mechanics*, 88(3):261–301, 2000.

[15] I. Lasiecka. *Mathematical control theory of coupled PDEs*. Society for Industrial Mathematics, 2002.

[16] G. Leugering. Exact controllability in viscoelasticity of fading memory type. *Applicable Analysis*, 18(3):221–243, 1984.

[17] G. Leugering. Exact boundary controllability of an integrodifferential equation. *Applied Mathematics and Optimization*, 15(1):223–250, 1987.

[18] G. Leugering. Time optimal boundary controllability of a simple linear viscoelastic liquid. *Math. Meth. Appl. Sci*, 9:413–430, 1987.

[19] Jacques-Louis Lions. *Optimal control of systems governed by partial differential equations*. Springer-Verlag, Berlin, New York,, 1971.

[20] JL Lions. Exact controllability, stabilization and perturbations for distributed systems. *SIAM REVIEW*, 30:1, 1988.

[21] JL Lions. Are there connections between turbulence and controllability. In *9th INRIA International Conference, Antibes*, 1990.

[22] W.J. Liu and G.H. Williams. Partial exact controllability for the linear thermo-viscoelastic model. *Electronic Journal of Differential Equations*, 1998(17):1–11, 1998.

[23] X. Marduel and K. Kunisch. Suboptimal control of transient nonisothermal viscoelastic fluid flows. *Physics of Fluids*, 13:2478, 2001.

[24] JG Oldroyd. On the Formulation of Rheological Equations of State. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences (1934-1990)*, 200(1063):523–541, 1950.

[25] M. Renardy. Are viscoelastic flows under control or out of control? *Systems & Control Letters*, 54(12):1183–1193, 2005.

[26] M. Renardy. Shear flow of viscoelastic fluids as a control problem. *Journal of Non-Newtonian Fluid Mechanics*, 131(1-3):59–63, 2005.

[27] M. Renardy. On control of shear flow of an upper convected Maxwell fluid. *ZAMM*, 87(3), 2007.

[28] M. Renardy. Controllability of viscoelastic stresses for nonlinear Maxwell models. *Journal of Non-Newtonian Fluid Mechanics*, 156(1-2):70–74, 2009.

[29] Michael Renardy. *Mathematical analysis of viscoelastic flows*, volume 73. Society for Industrial and Applied Mathematics, Philadelphia, 2000.

[30] RS Rivlin and JL Ericksen. Stress-deformation relations for isotropic materials. *Collected Papers of RS Rivlin*, 1997.

[31] D.L. Russell. A unified boundary controllability theory for hyperbolic and parabolic partial differential equations. *Studies appl. Math.*, 52:189–211, 1973.

[32] D.L. Russell. Exact boundary value controllability theorems for wave and heat processes in star-complemented regions, in" Differential Games and Control Theory. *Liu and Sternberg (Marcel Dekker Inc., New York, 1974)*, 1974.

[33] D.L. Russell. Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions. *SIAM REVIEW*, 20:639, 1978.

[34] E.D. Sontag. *Mathematical control theory: deterministic finite dimensional systems.* Springer, 1998.

[35] C. Truesdell and W. Noll. *The Non-Linear Field Theories of Mechanics.* Springer-Verlag, 3rd edition, 2004.

Appendix A

Index of used notation

In this appendix we list and explain the notation used throughout the work.

We denote the set of all real numbers by $\mathbb{R}$. Let X be any set, we denote as X^n the set of ordered n-tuples $(x_1, x_2, \ldots, x_n)$, where $x_i \in X$ and n is a known positive integer. For the sake of saving space we designate the bold letter to denote the vector quantities or n-tuples, for example $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. For two sets X and Y we denote as $X \times Y$ the set of pairs (x, y), where $x \in X$ and $y \in Y$. In this dissertation we only work with real valued functions.

A.1 Function spaces

Let X be an open or closed subset of $\mathbb{R}^n$. Below we provide the notation for different function spaces used in throughout the work. We

$C(X)$ — The space of continuous functions $f(x)$. The space is equipped with the norm

$$\|f(x)\|_{C(X)} \overset{def}{=} \sup_{x \in X} |f(x)|.$$

In case X is not compact, the space of functions with the finite norm is sometimes denoted $C_b(X)$.

$C^m(X)$ — The space of continuous functions with m continuous derivatives. The number m is a nonnegative integer. The space is equipped with the norm

$$\|f(x)\|_{C^m(X)} \overset{def}{=} \sup_{x \in X, 0 \leqslant k \leqslant m} |f^{(k)}(x)|.$$

$L^p(X)$ The space of Lebesgue measurable functions $f(x)$ which satisfy $\int_X |f(x)|^p \, dx < \infty$. The number p is a real number that satisfies: $p \geqslant 1$. The space is equipped with the norm

$$\|f(x)\|_{L^p(X)} \stackrel{def}{=} \left(\int_X |f(x)|^p \, dx \right)^{1/p} \stackrel{def}{=} \|f(x)\|_p.$$

There are two special cases:

- If $p = 2$, then the space is a Hilbert space equipped with the inner product $\langle f, g \rangle_2 = \int_X f \cdot g \, dx$, where $f,\ g \in L^2(X)$;
- If $p = \infty$ then the norm definition changes to $\|f(x)\|_{L^\infty(X)} \stackrel{def}{=} \operatorname*{ess\,sup}_{x \in X} |f(x)| \stackrel{def}{=} \|f(x)\|_\infty$. Since the requirement for finite integral above does not make sense in this case, we require the norm to be finite.

$W^{m,p}(X)$ The Sobolev space of measurable functions $f(x)$, such that $f(x)$ and its distributional derivatives $D^\alpha f(x)$ of order less than or equal to m belong to L^p. The number m is a nonnegative integer and p is a real number satisfying $p \geqslant 1$. The space is equipped with the norm

$$\|f(x)\|_{W^{m,p}(X)} \stackrel{def}{=} \left(\sum_{|\alpha| \leqslant m} \|D^\alpha f(x)\|_p^p \right)^{1/p} \stackrel{def}{=} \|f(x)\|_p.$$

As above, there are two special cases:

- If $p = 2$, the space is a Hilbert space and denoted as $H^m(X)$. The inner product is defined as $\langle f, g \rangle_{H^m(X)} \stackrel{def}{=} \sum_{|\alpha| \leqslant m} \langle D^\alpha f, D^\alpha g \rangle_2$, where $f,\ g \in H^m(X)$;
- If $p = \infty$ then the norm definition changes to $\|f(x)\|_{W^{m,\infty}(X)} \stackrel{def}{=} \max_{|\alpha| \leqslant m} \|D^\alpha f(x)\|_\infty \stackrel{def}{=} \|f(x)\|_{m,\infty}$.

For the following functional spaces we assume that $I \subset \mathbb{R}$ is an open or closed interval (possibly infinite), X is a real Banach space with the norm $\| \bullet \|_X$.

$C(I;X)$ The space of continuous mappings $f(t)$ so that $\forall t \in I$ we have $f(t) \in X$. The space is equipped with the norm

$$\|f(x)\|_{C(I;X)} \overset{def}{=} \sup_{t\in I} \|f(t)\|_X.$$

In case I is not compact, the space of functions with the finite norm is sometimes denoted $C_b(I;X)$.

$C^m(I;X)$ The space of continuous mappings $f(t)$ with m continuous derivatives $\frac{\partial^k f}{\partial t^k} \in X \ \forall 0 \leqslant k \leqslant m$. The number m is a nonnegative integer. The space is equipped with the norm

$$\|f(x)\|_{C^m(X)} \overset{def}{=} \sup_{x\in X, 0\leqslant k\leqslant m} \|f^{(k)}(x)\|_X.$$

$L^p(I;X)$ The space of Lebesgue measurable X-valued functions $f(t)$ which satisfy $\int_I \|f(t)\|_X^p \, dt < \infty$. The number p is a real number that satisfies: $p \geqslant 1$. The space is equipped with the norm

$$\|f(x)\|_{L^p(I;X)} \overset{def}{=} \left(\int_I \|f(t)\|_X^p \, dt \right)^{1/p}.$$

Again, we distinguish two special cases:

- If $p = 2$, and X is a Hilber space then the space $L^2(I;X)$ is a Hilbert space equipped with the inner product $\langle f, g \rangle_{L^2(I;X)} = \int_I \langle f(t), g(t) \rangle_X, dt$, where f, $g \in L^2(I;X)$;

- If $p = \infty$ then the norm definition changes to $\|f(x)\|_{L^\infty(I;X)} \overset{def}{=} \operatorname{ess\,sup}_{t\in I} \|f(t)\|_X$. Since the requirement for finite integral above does not make sense in this case as well, we require the norm to be finite.

$W^{m,p}(I;X)$ The Sobolev space of measurable X-valued functions $f(t)$, such that $f(t)$ and its distributional derivatives $D^k f(x)$ belong to $L^p(I;X)$ $\forall 0 \leqslant k \leqslant m$. The number m is a nonnegative integer and p is a real number satisfying $p \geqslant 1$. The space is equipped with the norm

$$\|f(x)\|_{W^{m,p}(I;X)} \stackrel{def}{=} \left(\sum_{0 \leqslant k \leqslant m} \|D^k f(t)\|_X^p \right)^{1/p}.$$

We also separate the following two special cases:

- When $p = 2$ and X is a Hilbert space, the space $W^{m,2}(I;X)$ is a Hilbert space itself and denoted as $H^m(I;X)$. The inner product is defined as $\langle f, g \rangle_{H^m(I;X)} \stackrel{def}{=} \sum_{0 \leqslant k \leqslant m} \langle D^k f, D^k g \rangle_{L^2(I;X)}$, where f, $g \in H^m(I;X)$;

- If $p = \infty$ then the norm is defined as $\|f(x)\|_{W^{m,\infty}(I;X)} \stackrel{def}{=} \underset{0 \leqslant k \leqslant m}{max} \|D^\alpha f(x)\|_{L^\infty(I;X)}$.

Appendix B

Analysis of coercivity booundaries

Here we will investigate whether the α value needed for equations (3.21) to be nonsingular and the value of α needed for the functional (3.19) to be convex have anything in common.

The functional that we investigate for convexity is:

$$J(q) = S_2(T) + \alpha S_1(T) = \int_0^T f(t, q, \dot{q}) dt$$

where $f(t, q, \dot{q})$ is:

$$f(t, q, \dot{q}) = \frac{\lambda_2 e^{\lambda_2 (t-T)} (\lambda_1 q + \dot{q})^2}{(\lambda_2 - \lambda_1)^2 \mu_1^2} + \alpha \frac{\lambda_1 e^{\lambda 1 (t-T)} (\lambda_2 q + \dot{q})^2}{(\lambda_2 - \lambda_1)^2 \mu_2^2}$$

For this thing to be convex, we need the folowing to be true:

$$J(q+v) - J(q) \geqslant \delta J(q, v) \qquad\qquad v(0) = v(T) = 0$$

This can be rearranged into:

$$\int_0^T \lambda_2 e^{\lambda_2 (t-T)} (\dot{v} + \lambda_1 v)^2 \mu_2^2 + \alpha \lambda_1 e^{\lambda_1 (t-T)} (\dot{v} + \lambda_2 v)^2 \mu_1^2 dt \geqslant 0$$

The critical value for α when the Euler-Lagrange equations are on the edge of being singular is $\alpha = e^{-T(\lambda_2 - \lambda_1)} \frac{\lambda_2 \mu_2^2}{\lambda_1 \mu_1^2}$ (provided $\lambda_2 > \lambda_1$). To find out whether the functional can become nonconvex at this value of α we simply plug it in, thus arrive at:

$$\int_0^T (e^{\lambda_2 t} - e^{\lambda_1 t}) \dot{v}^2 - (\lambda_2 - \lambda_1)(\lambda_1 e^{\lambda_2 t} + \lambda_2 e^{\lambda_1 t}) v^2 \, dt \geqslant 0$$

This integral is not so easy to analyze, therefore, we make use of a little trick. If λ_2 and λ_1 are close ($\lambda_1 \approx \lambda \approx \lambda_2$), we can say that approximately:

$$\begin{aligned} &\tfrac{1}{\lambda_2 - \lambda_1} \textstyle\int_0^T (e^{\lambda_2 t} - e^{\lambda_1 t}) \dot{v}^2 - (\lambda_2 - \lambda_1)(\lambda_1 e^{\lambda_2 t} + \lambda_2 e^{\lambda_1 t}) v^2 \, dt \approx \\ &\approx \textstyle\int_0^T t \, e^{\lambda t} \dot{v}^2 - 2 e^{\lambda t} v^2 \, dt \end{aligned}$$

For simplicity, we put $T = 1$. To see if any function makes this functional negative, we minimize it over the set of functions that satisfy $\int_0^1 e^{\lambda t} v^2 \, dt = 1$.

Therefore, the problem converts into the minimization of the following integral:

$$\int_0^T t\, e^{\lambda t} \dot{v}^2 + \Lambda e^{\lambda t} v^2 \, dt$$

where Λ is a Lagrange multiplier.

For this minimization problem, the Euler-Lagrange equations can be solved in closed form, yielding:

$$v(t) = e^{-\lambda t}(C_1 U(-n, 1, \lambda t) + C_2 L_n(\lambda t)) \quad n = \frac{\Lambda - \lambda}{\lambda}$$

Here U is a confluent hypergeometric function and L_n is a Laguerre function.

Since we are looking for a fuunction with zero boundary condition, we have to drop U function, because it is singular at zero, thus we canot find a proper minimizer here. Fortunately, we don't need the minimizer, we only need the function that makes the integral negative, and there is a definite way to find such a function.

Let's start with $v(t) = e^{-\lambda t} L_n(\lambda t)$. After some experimenting we will find that for $\lambda \geqslant 2$ we can find a value for Λ that makes $v(1) = 0$. Unfortunately $v(0) = 1$ for whatever the coefficient Λ we use. To circumvent this issue we make some correction:

$$\tilde{v}(t) = v(t_2)\, w(t) + I_{[t_2,1]}\, v(t) \quad t_2 \in (0, 1)$$

where $w(t)$ is a piecewise continuous function that has the following properties:

$$\begin{array}{ll} w(t) = 0, & t \in [0, t_1] \cup (t_2, 1], \\ \dot{w}(t) = \frac{\delta}{t}, & t \in [t_1, t_2] \\ w(t) = 1 & t = t_2 \end{array}$$

The goal here is to tweak three parameters t_1, t_2 and δ so that: $\|v\|_2$ is small, the function is continuous on $[0, t_2]$ and $\int_0^{t_2} t\, \dot{w}^2(t)\, dt$ is small too.
Let δ be small, we choose t_2 to achieve:

$$\int_0^{t_2} v^2(t_2)\, dt \leqslant \int_0^{t_2} 1\, dt < \delta\,,$$

then we choose t_1 to maintain continuity, i.e.:

$$\int_{t_1}^{t_2} \dot{w}(t)\, dt = \delta\, (\ln(t_2) - \ln(t_1)) = 1$$

Now we have:

$$\int_0^{t_2} t\, \dot{w}^2(t)\, dt = \delta^2\, (\ln(t_2) - \ln(t_1)) = \delta$$

Therefore:

$$\int_0^1 t\, e^{\lambda t}\dot{\tilde{v}}^2 - 2e^{\lambda t}\tilde{v}^2\, dt \leqslant$$
$$\leqslant 3\delta + \int_{t_2}^1 t\, e^{\lambda t}\dot{v}^2 - 2e^{\lambda t}v^2\, dt$$

If we choose δ sufficiently small, we will get that this integral is indeed negative. This means that the boundary of the forbidden region $\alpha = e^{-T(\lambda_2-\lambda_1)}\frac{\lambda_2\,\mu_2^2}{\lambda_1\,\mu_1^2}$ doesn't provide us convexity. For some combinations of λ_1 and λ_2 (especially if they are close to each other) we can make the functional to be nonconvex.

The numerical experiment confirms that for λ_1, λ_2 sufficiently big, the functional (3.19) does become nonconvex.