# Simulating Continuous-Time Infinite Horizon Optimization Models: A Simple Method and Applications

vorgelegt von
Diplom Wirtschaftsmathematiker Timo Trimborn
aus Hannover

Hamburg, den 9. November 2007

Mitglieder der Prüfungskomission


Vorsitzender: Prof. Dr. Peter Stahlecker

Erstgutachter: Prof. Dr. Bernd Lucke

Zweitgutachter: Prof. Dr. Richard S.J. Tol


Das wissenschaftliche Gespräch fand am 29.09.2008 statt

**Abstract**

We propose the relaxation algorithm as a simple and powerful method for simulating the transition process in growth models. This method has a number of important advantages: First, it can easily deal with a wide range of dynamic systems including stiff differential equations and systems giving rise to a continuum of stationary equilibria. Second, the application of the procedure is fairly user friendly. The only input required is the dynamic system. Third, the variant of the relaxation algorithm we propose exploits the infinite time horizon, which usually underlies optimal control problems in economics, in a natural manner. Fourth, the algorithm can solve preannounced or anticipated parameter changes in an easy and intuitive way, which makes it very useful for policy analysis, e.g. in a dynamic General Equilibrium context. As an illustrative application, we compute the transition process of the Jones (1995) and the Lucas (1988) models. In addition, we solve the Ramsey-Cass-Koopmans model for a shock consisting of anticipated parameter changes. Finally, we construct a General Equilibrium model with heterogenous households to demonstrate the potential of the relaxation algorithm. The model is calibrated to Jordanian data and employed to solve different scenarios of trade liberalization between Jordan and the European Union. This model exhibits many of the attributes that makes the application of usual procedures highly inefficient or even impossible.

# Contents

Contents

iii

# List of Tables

# List of Figures

List of Figures

v

# 1  Introduction

Infinite horizon continuous-time optimization problems arise frequently in dynamic macroeconomic theory. The dynamic system representing the first order conditions of optimality is then interpreted to describe the evolution of the economy under study. Many studies have confined their analysis to the long run balanced growth rate or the balanced growth path. In this way, additional theoretical insights about the transition process and possible counteracting effects during the transition phase of the economy are neglected. Therefore, a comprehensive analysis of the model under study requires an investigation of transitional dynamics. However, in many cases it is difficult or even impossible to derive qualitative effects of the transition path analytically, which makes a numerical computation of the transition path necessary. Moreover, if a sufficiently exact quantification of transitional effects is needed, e.g. for a pareto ranking of policy measures, a numerical computation of the transition path is inevitable. For example, in the field of dynamic Computable General Equilibrium modeling, it is very often so that the only possible analysis consists of numerical solutions of transition paths.

This dissertation contributes to the literature on dynamic macroeconomic theory by proposing the relaxation algorithm as a powerful method to determine the transition process in growth models numerically. We show that the procedure is able to solve continuous-time infinite horizon optimization models, for which the numerical solution is difficult or even impossible, if standard solution algorithms are applied. For applying the relaxation algorithm we exploit that the model's dynamics can be represented by a system of differential equations potentially augmented by algebraic equations. The presence of initial conditions and final boundary conditions turn the numerical problem into a two-point boundary value problem. In the case of infinite-horizon models the latter conditions consists of the requirement for solution

trajectories to converge towards an interior point or curve. Relaxation algorithms are known to be well-suited for this kind of numerical problem.

The principle of relaxation is to construct a large set of non-linear equations, which solution represents the solution trajectory. Equations representing the slope of the trajectory are obtained by discretization of the differential equations on a mesh of points in time. The set of equations is augmented by initial and final boundary equations. Potential algebraic equations are appended for every mesh point. This set of non-linear equation is solved, employing a Gauß-Newton procedure, which exploits the sparsity of the equations. The algorithms aims at solving for the transition path simultaneously and can, therefore, meet initial and final boundary conditions easily.

The relaxation procedure and similar finite-difference procedures have already been employed in various fields of economics. Prominent examples comprise the solution of two point boundary value difference equations (e.g. Laffargue, 1990; Juillard et al., 1998), differential-difference equations (e.g. Boucekkine et al., 1997) as well as partial differential equations (e.g. Candler, 1999). However, to the best of our knowledge, the relaxation algorithm has not been exploited yet systematically to solve deterministic continuous-time two point boundary value problems in growth theory. Nonetheless, there are a few applications in the economics literature. For instance, Oulton (1993) and Robertson (1999) employ the relaxation routine provided by Press et al. (1989) to solve a continuous-time deterministic growth model.

Other solution algorithms to solve continuous-time infinite horizon optimization models comprise backward integration (Brunner and Strulik, 2002), the finite-difference method as proposed by Candler (1999), time elimination (Mulligan and Sala-i-Martin, 1991), projection methods (e.g. Judd, 1992; Judd, 1998, Chapter 11), and the method of Mercenier and Michel (1994 and 2001). The merits and shortcomings of these methods are diverse, however, many of them have problems to solve models

exhibiting a multi-dimension stable manifold. Some procedures have to be modified if the dimension of the state space increases, or computational requirements increase drastically, which could permit the numerical solution of higher dimensional models.

More precisely, the method of Candler is generic with respect to the dimension of the state space, however, computational costs grow exponentially as the number of state variables increases. The method of backward integration is not generic for models exhibiting higher-dimensional stable manifolds. If the model exhibits two state variables, a backward shooting algorithm can be applied. However, the problem of stiff differential equations might occur, which could make the solution of higher dimensional models difficult or even impossible. The same criticism applies for the method of time elimination, since its generalization to multi-dimensional stable manifolds is conceptually equal to the backward integration procedure. The method of Mercenier and Michel is generic with respect to the number of state variables. Moreover, computational costs grow acceptably if the model dimension increases. The only disadvantage of the method is that it uses first order difference schemes and, therefore, it might need a large amount of mesh points to yield accurate solutions. Projection methods are also generic with respect to the model dimension. However, computational requirements grow exponential if the dimension of the model increases. This "curse of dimensionality" can be attenuated by selecting a special basis, but, still, computational costs grow considerably.

The advantage of the relaxation algorithm is that it treats higher dimensional systems in a generic way. This means, no conceptual changes have to be made with respect to the algorithm if the dimension of the model increases. This holds for an increase in the dimension of control (or jump) variables and for an increase in the dimension of state (or predetermined) variables. Moreover, a great emphasis has been made to design the application of the relaxation algorithm as user-friendly as

possible.

We show the relaxation algorithm also to be able to handle properties of dynamic systems that arise for prominent growth models. The first property is that of stiff differential equations, i.e. a considerably difference in the real part of the stable eigenvalues. This characteristic arises, for example, in the Jones (1995a) model. We simulate this model numerically, and show that the analysis of transitional dynamics can bring about additional theoretical insights concerning the model's dynamic properties.

The second property arising in prominent growth models is that of a center manifold of stationary equilibria, i.e. long run equilibria form a curve. Nonetheless, transitional dynamics are unique, and initial values of the state variables determine to which particular steady state level the economy converges. This characteristic arises in the prominent Lucas (1988) model. It is surprising that to the best of our knowledge no numerical analysis of the model's original representation as described in Lucas (1988) exists, although this is the secondary most cited paper on the economists' platform IDEAS.[1] Moreover, theoretical analysis of the dynamic system as presented in the seminal paper is rare (e.g. Caballe and Santos, 1993, confine to a parameter restriction).

The third property arising in growth models is that of a preannounced or anticipated shock. We derive the well-known continuity principle of adjoint variables for preannounced or anticipated changes in parameters for continuous-time, infinite horizon, perfect foresight optimization models. The resulting multi-point boundary value problem can be solved numerically by employing the relaxation algorithm. By ensuring that the state variables and the adjoint variables are continuous, potential jumps in the control variables are calculated automatically. We solve a Ramsey

---

[1]See http://ideas.repec.org/top/top.item.simple.html, evaluated at Nov. 7, 2007.

model extended by an elementary Government sector numerically as an example.

Finally, we demonstrate the potential of the relaxation algorithm by solving a highly dimensional Computable General Equilibrium (CGE) model numerically. The model exhibits a six-dimensional stable manifold, a five-dimensional center manifold of stationary equilibria, and a schedule of preannounced shocks. Therefore, the model exhibits many of the attributes that make the application of usual procedures highly inefficient or even impossible. We introduce heterogenous households into an otherwise standard neoclassical dynamic CGE model. We calibrate the model to Jordanian data represented by a Social Accounting Matrix (SAM) of 2002 and a household survey. The model is employed to investigate the economic implications of trade liberalization on welfare and income distribution of heterogenous households induced by the Association Agreement of 2002 between Jordan and the European Union.

The thesis is organized as follows. In Section 2, we provide the formal mathematical framework for the theoretical analysis of the dynamics of the models presented in this thesis. Moreover, the case of stiff differential equations is presented. In Section 3, we describe the relaxation algorithm in detail and compare it to competing solution algorithms. Moreover, we present a theoretical and numerical analysis of transitional dynamics for the Jones (1995a) model and the Lucas (1988) model. In Section 4, we analyze the theoretical treatment of anticipated or preannounced shocks in continuous-time, perfect foresight optimization models. Moreover we discuss how to solve models numerically, which exhibit an anticipated shock. In Section 5, we construct a large Computable General Equilibrium model with heterogenous households. We analyze the model's dynamic behavior theoretically and numerically.

# 2  Mathematical Preliminaries

## 2.1  Introduction

In this Section, we want to set the formal, mathematical framework for the analysis of the economic models presented in this thesis. Since transitional dynamics of these models is represented by a dynamic system, we focus on the analysis of transitional dynamics and long-run behavior of dynamic systems. First, we analyze a dynamic system that exhibits an isolated, hyperbolic fixed point. Transitional dynamics are described by the Hartman-Grobman theorem, which states that the local behavior of the system equals its linearization. Note, that the conclusions only apply for the neighborhood of a fixed point. At some distance, the non-linear behavior may dominate the linear behavior. Then, we analyze transitional dynamics around a curve of stationary equilibria. The Lucas (1988) model, for example, exhibits this characteristic. In the proximate Section, we analyze transitional dynamics around a hyperbolic fixed point on a two-dimensional stable manifold. We focus on the case of two real eigenvalues which differ considerably, because the Jones (1995a) model exhibits this characteristic. Moreover, we describe the phenomenon of stiff differential equations, referring to a two-dimensional stable manifold. Again, the Jones (1995a) model exhibits this characteristic, and this exposure is intended to enhance the understanding of problems that could arise for the numerical solution of the model.

## 2.2  Qualitative long-run behavior of dynamic systems

We start by defining a flow, which is defined by a system of differential equations.[2]

---

[2]The definitions of this Subsection originate from Tu (1994) if not indicated otherwise. If corollaries or theorems are not the authors own contribution this is indicated as well.

2 Mathematical Preliminaries

**Definition 1** *(Flow) Given the system of n nonlinear ordinary differential equations*

$$\dot{x} = f(x) \qquad f : M \subset \mathbb{R}^n \to \mathbb{R}^n \tag{1}$$

*with $f$ being sufficiently smooth. Then $f$ is a vector field which generates a flow $\phi^t : M \to \mathbb{R}^n$ where $\phi^t(x) \equiv \phi(x,t)$, defined for all $x \in M$ and $t \in (a,b)$ satisfying*

$$\frac{d\phi(x,t)_{t=s}}{dt} = f[\phi(x,s)] \tag{2}$$

*for all $x \in M$ and $s \in (a,b)$.*

An orbit of the flow is a set in $\mathbb{R}^n$ that the flow passes. Referring to the system of differential equations, an orbit is the maximum trajectory that results from a specific initial point, whereas the trajectory is extended in both directions of time.

**Definition 2** *(Orbit) Given $x \in M$ the orbit of $x$, $\gamma(x)$ is defined as*

$$\gamma(x) := \{\phi^t(x) : t \in [t_{min}, t_{max}]\}$$

*where $[t_{min}, t_{max}]$ is the maximum time interval that $\phi$ is defined for. The forward orbit of $x$ is defined as*

$$\gamma^+(x) := \{\phi^t(x) : t \in [t_{min}, t_{max}]\}$$

*and the reverse orbit as*

$$\gamma^-(x) := \{\phi^t(x) : t \in [t_{min}, t_{max}]\}.$$

An invariant set is a set in $\mathbb{R}^n$ that the flow stays in forever. It could refer to an arbitrary shaped region that the flow does not leave. However, in the context of the following theorems it should better be understood as an orbit defined for $t_{max} = \infty$.

**Definition 3** *(Invariant Set) The non-empty set $V \subset M$ is said to be invariant, if $\phi^t(x)$ is defined for all $t > 0$ and*

$$\phi^t(V) \subseteq V \qquad \forall t \geq 0$$

We do not want to address existence and uniqueness of solutions, but instead focus on the long-run behavior of the system and the topology implied by the flow. At first, we focus on the system's behavior around a single fixed point.

**Definition 4** *(Fixed Point) Given the system of $n$ nonlinear ordinary differential equations*

$$\dot{x} = f(x) \qquad f : M \subset \mathbb{R}^n \to \mathbb{R}^n \tag{3}$$

*a point $x^*$ at which $f(x^*) = 0$ is called a stationary point or fixed point.*

Usually, in the literature further assumptions about the fixed point are made. At first, we assume the fixed point to be isolated, i.e. no further stationary points are 'nearby'. A more precise definition is the following.

**Definition 5** *(Simple Fixed Point) A fixed point is said to be simple if its linearized system $D_x f|_{x^*} \cdot x$ has no zero eigenvalues. i.e. if*

$$\det(D_x f|_{x^*}) \neq 0.$$

Linearization theorems like the Hartman-Grobman theorem, however, need an even stronger assumption. From a linear system of differential equations it is well-known that the magnitude of the real parts of the eigenvalue indicates the speed of convergence towards the fixed point. If the real part is zero, but the complex part is nonzero, the linear system would exhibit periodic closed orbits around the fixed

point.[3] The solutions would neither converge to the fixed point nor diverge from the fixed point. However, this result cannot be generalized to non-linear systems. In the literature, the case of a zero eigenvalue is often excluded from the analysis of qualitative long-run behavior of non-linear systems.

**Definition 6** *(Hyperbolic Fixed Point) A fixed point is said to be hyperbolic if its linearized system* $D_x f|_{x^*} \cdot x$ *has no eigenvalues with zero real part.*

Next, we want to draw the connection between a dynamics system and the corresponding flow. Since the connection can be made easily, we switch between both concepts whenever it seems convenient.

**Corollary 7** *Consider the n-dimensional dynamic system* $\dot{x} = f(x)$ *and the corresponding flow* $\phi^t$. *The Jacobian matrixes are denoted by* $Df$ *and* $D\phi^t$, *respectively. Then for a fixed point* $x^*$

$$D\phi^t(x^*) = \exp(Df(x^*) \cdot t) \tag{4}$$

*holds. Furthermore, if* $\lambda_1, \ldots, \lambda_k$, $k \leq n$ *are eigenvalues of* $Df$ *with corresponding eigenvectors* $x_1, \ldots, x_k$, *then* $\exp(\lambda_1 t), \ldots, \exp(\lambda_k t)$, $k \leq n$ *are eigenvalues of* $D\phi^t$ *with corresponding eigenvectors* $x_1, \ldots, x_k$.

**Proof.** The solution can be verified by linearization and differentiating the equation with respect to time (e.g. Königsberger, 1997, pp.152). If $\lambda$ is an eigenvalue of $Df$ with corresponding eigenvector $x$, then $Df \cdot x = \lambda x$ holds. Multiplying equation (4) with $x$ it follows that

$$
\begin{aligned}
D\phi^t x &= \exp(Dft)x \\
&= \left( I + Dft + \frac{1}{2!}(Df)^2 t^2 + \frac{1}{3!}(Df)^3 t^3 \ldots \right) x
\end{aligned}
$$

---

[3]See, for example, Barro and Sala-i-Martin (2004), p. 588.

$$= \quad x + \lambda t x + \frac{1}{2!}\lambda^2 t^2 x + \frac{1}{3!}\lambda^3 t^3 x + \dots$$

$$= \quad \exp(\lambda t)x$$

Therefore, $\exp(\lambda t)$ is an eigenvalue of $D\phi^t(x)$ with eigenvector $x$. ∎

The previous definition focusses on the eigenvalues of the linearized system $f$. The linearization of the corresponding flow $\phi^t$ provides the same information.

We want to investigate the dynamic behavior of the system around a simple, hyperbolic fixed point. The theorem of Hartman and Grobman is the standard theorem providing information about the topology.[4]

**Theorem 8** *(Linearization Theorem of Hartman and Grobman, Tu (1994))*
*Let the nonlinear dynamic system*

$$\dot{x} = f(x) \tag{5}$$

*have a simple hyperbolic fixed point $x^*$. Let the Jacobian matrix $D_x f$ evaluated at $x^*$ have $n_u$ eigenvalues with positive real part and $n_s$ eigenvalues with negative real part with the corresponding eigenspaces $N^u$ and $N^s$, respectively ($N^u \oplus N^s = \mathbb{R}^n$). Then the following claims hold*

1. *In the neighborhood $U$ of $x^* \in \mathbb{R}^n$ of this equilibrium, the phase portraits of the original system and its linearization*

$$\dot{x} = D_x f|_{x^*} \cdot x \tag{6}$$

   *are equivalent.*

2. *The corresponding flow $\phi^t$ is topologically equivalent to the linearization $D\phi^t|_{x^*}$.*

---

[4]The Hartman-Grobman theorem appears in several versions in Tu (1994), e.g. pages 135, 145, 146, 157 and 188. We collect the main conclusions in one theorem.

3. *There exists locally smooth manifolds $W^u(x^*)$ and $W^s(x^*)$, called a local unstable manifold and a local stable manifold, respectively, tangent to the linear spaces $N^u$ and $N^s$, respectively.*

4. *$W^s(x^*)$ is characterized by $||\phi(y) - \phi(x^*)|| \to 0$ exponentially as $t \to \infty$ for any $y \in W^s(x^*)$, and $W^u(x^*)$ is characterized by $||\phi(y) - \phi(x^*)|| \to 0$ exponentially as $t \to -\infty$ for any $y \in W^u(x^*)$.*

The main statement of the theorem is that in the neighborhood of a simple, hyperbolic fixed point the system behaves the same way as the linearized system. Therefore, it is possible to infer from the eigenvalues on the qualitative behavior and the speed of convergence. Moreover, it is possible to infer from the eigenvectors on the direction of convergence. The phase space can be divided into a stable and unstable manifold.

A trajectory which starts on the stable manifold converges towards the steady state. Therefore, by defining $n_s$ equations transversal to the stable manifold a unique trajectory is defined, if convergence towards the fixed point is required.

Until now, we have assumed the fixed point to be hyperbolic, since the theorem can only be applied to hyperbolic fixed points. A more general statement about the neighborhood of a fixed point that is potentially non-hyperbolic is the Center Manifold Theorem.

**Theorem 9** *(Center Manifold Theorem, Tu (1994))*
*Let the nonlinear dynamic system*

$$\dot{x} = f(x) \tag{7}$$

*be a $C^r$ vector field with a fixed point $x^*$ set at the origin for simplicity. Let the Jacobian matrix $D_x f$ evaluated at $x^*$ have $n_u$ eigenvalues with positive real part and*

$n_s$ *eigenvalues with negative real part and* $n_c$ *eigenvalues with zero real part. We denote the corresponding eigenspaces* $N^u$, $N^s$ *and* $N^c$, *respectively. Then there exists* $C^r$ *stable, unstable manifolds* $W^s$, $W^u$ *and a* $C^{r-1}$ *central manifold* $W^c$ *tangent to* $N^s$, $N^u$ *and* $N^c$ *at* $x^*$ *and invariant for the flow of* $f$.
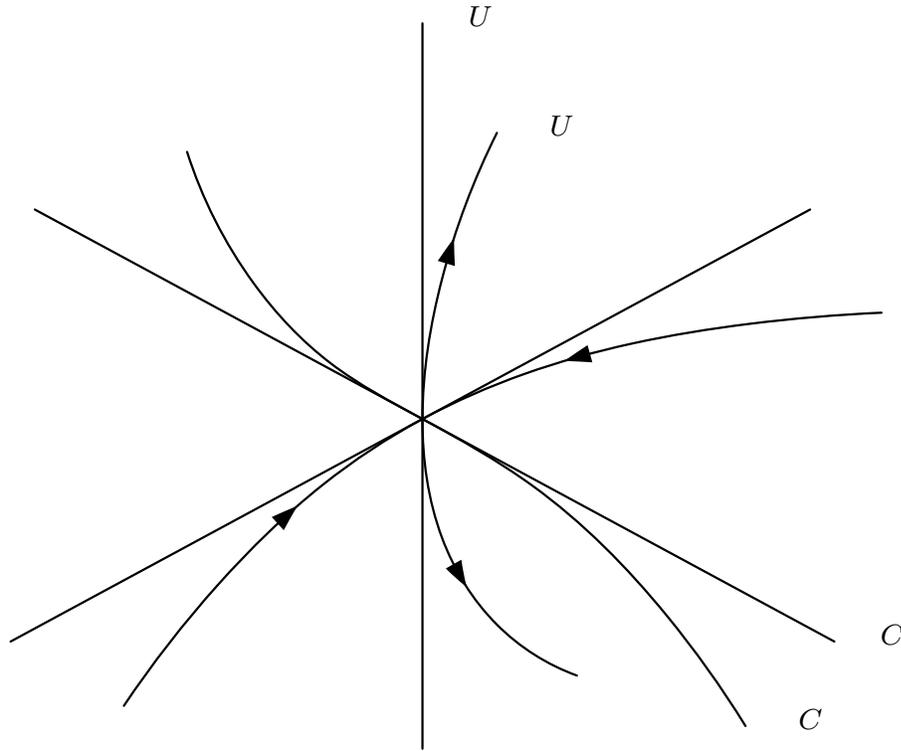


Figure 1: Fixed point with stable, unstable and center manifold

The theorem states that, besides the stable and unstable manifold, there exists a center manifold, associated to the eigenvalues with zero real part and tangent to the corresponding eigenvectors (see Figure 1). Whenever the Jacobian matrix evaluated at a fixed point exhibits a zero eigenvalue, the corresponding eigenspace is tangent to the center manifold.

However, the Hartman-Grobman theorem cannot be applied to conclude about the dynamic behavior on the center manifold in general. The reason is that the linearized system cannot give sufficient information about dynamics on a center manifold. In direction of an eigenvalue with non-zero real part, the linearization provides information about the local behavior of the system, because near the fixed point the linear part will dominate the nonlinear components. However, this conclusion cannot be drawn for systems exhibiting an eigenvalue of zero real part, since the linear behavior would indicate no movement. Therefore, any nonlinear behavior would dominate the linear behavior.

This can be seen clearly by considering the following differential equation

$$\dot{x} = x^2 \,. \tag{8}$$

The fixed point is the origin, $x^* = 0$. The Jacobian matrix evaluated at the fixed point yields

$$2x\big|_{x^*=0} = 0 \,.$$

Therefore, the neighborhood of the fixed point consists of a center manifold. No stable or unstable manifold is present. While the linearized system exhibits no movement, the nonlinear system exhibits qualitatively different behavior, depending on which side of the origin is examined. For $x < 0$ it follows that $\dot{x} > 0$ and, hence, $x$ is converging towards the origin. For $x > 0$ it follows that $\dot{x} > 0$ and $x$ is diverging.

We do not want to investigate the dynamics on a center manifold in general, but instead focus on a special case. For some growth models the phenomenon of a center manifold of stationary equilibria occurs.[5]

---

[5]The most prominent model is the Lucas (1988) model. In the seminal paper of Lucas and the analysis of Caballe and Santos (1993) the dynamic system exhibits a center manifold.

**Corollary 10** *A $n_c$-dimensional manifold of stationary equilibria is a center manifold. The Jacobian matrix evaluated at the center manifold exhibits at least $n_c$ zero eigenvalues.*

**Proof.** A manifold of stationary equilibria is a center manifold, because the solution of the problem

$$\dot{x} = 0 \quad \Leftrightarrow \quad f(x) = 0$$

is locally not unique, i.e. the system of nonlinear equations $f(x)$ is singular. Therefore, on the manifold $f(x) = 0$ the tangent bundle tangent to the manifold $f(x) = 0$ is in the null space of the Jacobian matrix $Df(x)$. ∎

In this special case, the additional information is available that there is no movement on the center manifold. A simple, linear differential equation exhibiting the same property is $\dot{x} = 0$. As a next step we want to focus on the dynamic properties of higher dimensional, nonlinear systems in the neighborhood of a center manifold of stationary equilibria. To answer that question we can apply the Fundamental theorem of normally hyperbolic invariant manifolds. The standard textbook stating the theorem is Hirsch, Pugh and Shub (1977). However, since the presentation of the theorem is very abstract, we refer to a more concrete statement of the theorem in Li et al. (2003).[6] Before we state the theorem we have to extend the definition of a hyperbolic set from a single point to a manifold.

**Definition 11** *(r-normally hyperbolic invariant manifold, Hirsch et al. (1977), Li et al. (2003)) An invariant manifold $V \subset M$ is called r-normally hyperbolic ($1 \leq r \leq \infty$) with respect to a flow $\phi^t$, if*

*1. $\phi^t$ is $C^r$*

---

[6]Li et al. (2003) apply the theorem to an economic model that exhibits similar mathematical properties, but which is not related to economic growth.

2. $T_V M$, the tangent bundle of $M$ restricted on $V$, splits into 3 continuous sub-bundles $N^u$, $TV$, and $N^s$ such that

$$T_V M = N^u \oplus TV \oplus N^s \tag{9}$$

Thereby, $TV$ is tangent to the invariant manifold $V$.

3. There exists a vector norm $|| \cdot ||$ such that for every $x \epsilon \mathbb{R}$ and $0 \leq k \leq r$

$$||D\phi^t|_{N_x^u}||_{inf} \quad > \quad ||D\phi^t|_{T_x V}||_{sup}^k \tag{10}$$

$$||D\phi^t|_{N_x^s}||_{sup} \quad < \quad ||D\phi^t|_{T_x V}||_{inf}^k. \tag{11}$$

Thereby, given a matrix $A$, $||A||_{inf}$ and $||A||_{sup}$ are defined as

$$||A||_{inf} \quad := \quad \inf_{||x||=1} \{||Ax||\} \tag{12}$$

$$||A||_{sup} \quad := \quad \sup_{||x||=1} \{||Ax||\} \tag{13}$$

Roughly speaking, a hyperbolic invariant manifold is a manifold where the flow stays in forever and where the dynamics outside the manifold are 'faster' than on the manifold itself, where the 'speed' refers to the movement transversal to the manifold. Conditions (10) and (11) can be interpreted as $N_x^u$ is expanding faster than the movement on $T_x V$, and $N_x^s$ is contracting faster than the movement on $T_x V$. A center manifold of stationary equilibria exhibits no movement and is, therefore, a hyperbolic manifold.

Since we want to state a theorem analogous to the Hartman-Grobman theorem that connects the eigenvalues to the dimension of stable and unstable manifold, we want to identify $N_x^u$ and $N_x^s$ as eigenspaces. To avoid laborious notations we focus on the case in which the Jacobian matrix is diagonalizable in $\mathbb{C}$. That is, the eigenvalues may potentially be complex, but the algebraic and geometric dimension of the eigenspace of each eigenvalue is the same. In this case, the whole space $\mathbb{C}^n$ can

be divided into $n$ eigenspaces.[7] Moreover, we assume the center manifold to exhibit the same dimension as the number of zero eigenvalues, i.e. the center manifold and the manifold $f(x) = 0$ coincide.

**Corollary 12** *Consider a $n$-dimensional dynamic system $\dot{x} = f(x)$ and the corresponding flow $\phi^t$. Let the state space exhibit a manifold of stationary equilibria of dimension $n_c$ and the Jacobian matrix $Df$ evaluated on this manifold exhibits exactly $n_c$ eigenvalues with zero real part. Moreover, let the Jacobian matrix be diagonalizable in $\mathbb{C}$. Then, the manifold of stationary equilibria is a $\infty$-normally hyperbolic manifold.*

**Proof.**  With Corollary (7) and Corollary (10) we can conclude that exactly $n_c$ of the flow's eigenvalues are of value 1. We assume, that $n_u$ are of absolute value greater than 1 and $n_s$ are of absolute value smaller than 1. First, we construct a basis $y_1, \ldots, y_n$ for $\mathbb{R}^n$. Since we assumed the matrix to be diagonalizable in $\mathbb{C}$, the space $\mathbb{C}^n$ can be divided into $n$ invariant subspaces, each represented by an eigenvector. Each subspace associated with a real eigenvalue will be represented by a real eigenvector. Therefore, if $l \leq n$ real eigenvalues are present, we found already $l$ basis elements for $\mathbb{R}^n$, since a basis of real vectors for a subspace $\mathbb{C}^l$ is also a basis for $\mathbb{R}^l$. In the case of complex eigenvalues they will appear in complex conjugate pairs $\lambda, \bar{\lambda}$. If $y$ is an eigenvector of $\lambda$, then $\bar{y}$ is an eigenvector of $\bar{\lambda}$, because $\overline{Ay} = \overline{\lambda y}$ and $\bar{A} = A$. Since $y = \Re(y) + i\Im(y)$[8], for every pair $\lambda, \bar{\lambda}$ the pair $\Re(y), \Im(y)$ represents a basis of the corresponding subspace in $\mathbb{C}^n$. Note, that for that complex pair

$$
\begin{aligned}
D\phi \cdot \Re(y) &= \Re(\lambda)\Re(y) - \Im(\lambda)\Im(y) \\
D\phi \cdot \Im(y) &= \Im(\lambda)\Re(y) - \Re(\lambda)\Im(y)
\end{aligned} \tag{14}
$$

---

[7]According to the fundamental theorem of algebra $n$ eigenvalues exists in $\mathbb{C}$, counted by their multiplicity.

[8]$\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary part of a complex element, respectively.

holds. With these additional vectors, we constructed a basis for $\mathbb{C}^n$ only consisting of real vectors. Therefore, this will also represent a basis for $\mathbb{R}^n$.

For proofing equations (10) and (11), we change coordinates and employ the euclidian norm with respect to the basis $y_1, \ldots, y_n$ that we constructed

$$||x|| := \left\{ \sqrt{\alpha_1^2 + \ldots + \alpha_n^2}, x = \alpha_1 y_1 + \ldots + \alpha_n y_n, \ \alpha_i \epsilon \mathbb{R} \right\}$$

Note, that $||D\phi^t|_{T_xV}||_{sup}^k = D\phi^t|_{T_xV}||_{inf}^k = 1$ for any $k \geq 0$, because $T_xV$ is spanned by eigenvectors with an eigenvalue of 1. This means that every $x \epsilon T_xV$ is mapped on itself by $D\phi^t$. In detail

$$||D\phi^t|_{T_xV}||_{sup}^k = \sup_{||x||=1} \left\{ \sqrt{\alpha_1^2 + \ldots + \alpha_n^2}, x = \alpha_1 y_1 + \ldots + \alpha_n y_n \right\} = 1$$

$$||D\phi^t|_{T_xV}||_{inf}^k = \inf_{||x||=1} \left\{ \sqrt{\alpha_1^2 + \ldots + \alpha_n^2}, x = \alpha_1 y_1 + \ldots + \alpha_n y_n \right\} = 1$$

W.l.o.g. we focus on the linear space $N^u$ spanned by the eigenvectors corresponding to the $n_u$ eigenvalues. We have to show that $||D\phi^t|_{N_x^u}||_{inf} > 1$. Let $\lambda_{u,1}, \ldots, \lambda_{u,k}$ denote the eigenvalues of $N^u$ and let $y_{u,1}, \ldots, y_{u,k}$ denote the real basis of the corresponding subspace. Note, that any $x \epsilon N^u$ can be decomposed as $x = \alpha_1 y_{u,1} + \ldots + \alpha_k y_{u,k}$. Then, if all eigenvalues would be real,

$$\begin{aligned} ||D\phi^t|_{N_x^u}||_{inf} &= \inf_{||x||=1, x \epsilon N_x^u} \left\{ ||D\phi^t \cdot x|| \right\} \\ &= \inf_{||x||=1} \left\{ ||D\phi^t|_x \cdot x||, x = \alpha_1 y_{u,1} + \ldots + \alpha_k y_{u,k} \right\} \\ &= \inf_{||x||=1} \left\{ ||\alpha_1 \lambda_{u,1} y_{u,1} + \ldots + \alpha_k \lambda_{u,k} y_{u,k}|| \right\} \\ &= \inf_{||x||=1} \left\{ \sqrt{\alpha_1^2 \lambda_{u,1}^2 + \ldots + \alpha_k^2 \lambda_{u,k}^2} \right\} \end{aligned}$$

However, the norm of the vector is always greater than 1, since

$$\sqrt{\alpha_1^2 \lambda_{u,1}^2 + \ldots + \alpha_k^2 \lambda_{u,k}^2} > \sqrt{\alpha_1^2 + \ldots + \alpha_k^2} = 1$$

The inequality follows because $\lambda_{u,i}^2 > 1$, $\forall i$. If e.g. the first two eigenvalues are complex, the calculation has to modified. We continue to denote the basis by $y_{u,1}, y_{u,2}, \ldots, y_{u,k}$, but keep in mind that $y_{u,1}$ and $y_{u,2}$ are real and imaginary part of the eigenvector of $\lambda_{u,1} \epsilon \mathbb{C}$. Moreover, $\lambda_{u,1} = \bar{\lambda}_{u,2}$ and $|\lambda_{u,1}| = |\lambda_{u,2}| > 1$. The matrix norm now calculates with equation (14) according to

$$
\begin{aligned}
||D\phi^t|_{N_x^u}||_{inf} &= \inf_{||x||=1, x \epsilon N_x^u} \{||D\phi^t \cdot x||\} \\
&= \inf_{||x||=1} \{||D\phi^t|_x \cdot x||, x = \alpha_1 y_{u,1} + \ldots + \alpha_k y_{u,k}\} \\
&= \inf_{||x||=1} \Big\{||\alpha_1 \Re(\lambda_{u,1}) y_{u,1} - \alpha_1 \Im(\lambda_{u,1}) y_{u,1} + \alpha_2 \Im(\lambda_{u,1}) y_{u,1} \\
&\qquad\quad -\alpha_2 \Re(\lambda_{u,1}) y_{u,1} + \ldots + \alpha_k \lambda_{u,k} y_{u,k}||\Big\} \\
&= \inf_{||x||=1} \Big\{\sqrt{\alpha_1^2 |\lambda_{u,1}|^2 + \alpha_2^2 |\lambda_{u,1}|^2 + \ldots + \alpha_k^2 \lambda_{u,k}^2}\Big\}
\end{aligned}
$$

Again, it follows that

$$
\sqrt{\alpha_1^2 |\lambda_{u,1}|^2 + \alpha_2^2 |\lambda_{u,1}|^2 + \ldots + \alpha_k^2 \lambda_{u,k}^2} > \sqrt{\alpha_1^2 + \ldots + \alpha_k^2} = 1
$$

Therefore, $||D\phi^t|_{N_x^u}||_{inf} > 1 = ||D\phi^t|_{T_x V}||_{sup}^k$ for any $k \geq 0$. The argument for $||D\phi^t|_{N_x^s}||_{sup} < ||D\phi^t|_{T_x V}||_{inf}^k$ is analogous. ∎

Now, we can apply the fundamental theorem of normally hyperbolic invariant manifolds. The theorem states conclusions about both, the dynamic behavior in the neighborhood of a single point of the manifold, and the dynamic behavior in the neighborhood of the manifold as a whole.

**Theorem 13** *(Fundamental theorem of normally hyperbolic invariant manifolds, Li et al. (2003))*

*Let $\phi^t : M \to M$ be a $C^r$ flow of a $C^\infty$ manifold $M$ with $r \geq 1$ leaving the $C^1$ submanifold $V \subset M$ invariant, where $V$ is assumed to be compact. Assume that $\phi^t$ is r-normally hyperbolic at $V$ respective to the tangent bundle splitting $T_V M =$*

$N^u \oplus TV \oplus N^s$, where $D\phi^t$ exponentially expands and contracts the vectors in $N^u$ and $N^s$, respectively. Then

1. *Existence: There exists locally $\phi^t$-invariant submanifolds $W^u(\phi^t)$ and $W^s(\phi^t)$, called a local unstable manifold and a local stable manifold at $V$, respectively, tangent at $V$ to $N^u \oplus TV$, and $TV \oplus N^s$, respectively.*

2. *Uniqueness: Any locally invariant set near $V$ lies in $W^u \cup W^s$.*

3. *Characterization: $W^s$ consists of all points whose forward $\phi^t$-orbits never strays far from $V$, and $W^u$ of all points whose reverse $\phi^t$-orbits never stray far from $V$.*

4. *Smoothness: $W^u$, $W^s$ and $V$ are class $C^r$.*

5. *Foliation: $W^u$ and $W^s$ are invariantly fibered by $C^r$ submanifolds $W_x^{uu}$, $W_x^{ss}$, $x \in V$, tangent at $x$ to $N_x^u$ and $N_x^s$, respectively. $W^{uu}$ and $W^{ss}$ are invariant in the sense that $\phi^t(W_x^{ss}) \subset W_{\phi^t(x)}^{ss}$ for $t < 0$ and $\phi^t(W_x^{uu}) \subset W_{\phi^t(x)}^{uu}$ for $t > 0$. $W_x^{ss}$ is characterized by $||\phi(y) - \phi(x)|| \to 0$ exponentially as $t \to \infty$ for any $y \in W_x^{ss}$, and $W_x^{uu}$ is characterized by $||\phi(y) - \phi(x)|| \to 0$ exponentially as $t \to -\infty$ for any $y \in W_x^{uu}$.*

6. *Continuity: The leaves of foliation $W^{uu}$ and $W^{ss}$ are continuous on parameter $x \in V$ in $C^1$-topology.*

7. *Permanence: If $\tilde{\phi}^t$ is another $C^r$ flow on $M$ and is $C^r$ close to $\phi^t$ (i.e., close in $C^r$-norm). Then $\tilde{\phi}^t$ is r-normally hyperbolic at some unique submanifold $\tilde{V}$, which is $C^r$ close to $V$. The invariant manifolds $W^u(\tilde{\phi}^t)$, $W^s(\tilde{\phi}^t)$, and the leaves $W_x^{uu}(\tilde{\phi}^t)$, $W_x^{ss}(\tilde{\phi}^t)$, are $C^r$ close to those of $\phi^t$.*

8. *Linearization:  Near $V$, $\phi^t$ is topologically conjugate (i.e. $C^0$ conjugate) to $D\phi^t|_{N^u \oplus N^s}$, the restriction of the Jacobian of the flow to the subspace $N^u \oplus N^s$.*

It has to be verified that the assumptions of the theorem are fulfilled. We already proved that a center manifold of stationary equilibria is a $\infty$-normally hyperbolic invariant manifold. The theorem further demands the manifold to be compact.[9] If we restrict the phase space to a cuboid, compact area $I \subset \mathbb{R}^n$, the truncated center manifold is also compact. Since the theorem only holds for a neighborhood $U$ of the center manifold, we can now apply the theorem to $U \cap I$. The area beyond the cuboid is not of economic interest.

The results of theorem (13) demand a detailed mathematical and economic interpretation. First, we exploit statement (1) of the theorem. It states that there exists a local stable and unstable manifold tangent to corresponding linear spaces. Since we consider the case of a center manifold of stationary equilibria, the linear space can readily be identified. The space $TV$ is spanned by the eigenvectors of eigenvalue zero, $N^s$ is spanned by the eigenvectors with eigenvalues of negative real part, and $N^u$ is spanned by the eigenvectors with eigenvalues of positive real part, referred to the dynamic system respectively.[10] Therefore, the stable and unstable manifold of the center manifold are of dimension $n_c + n_s$ and $n_c + n_u$, respectively. Analogously to the Hartman-Grobman theorem a locally unique trajectory can be identified by giving $n_c + n_s$ initial conditions, i.e. a $(n_c + n_s)$-dimensional system of equations transversal to the stable manifold. Different to the Hartman-Grobman theorem the stable manifold now refers not to a single fixed point but to the whole center manifold. This means that from theorem (13.1) it can only be concluded that

---

[9]Li et al. (2003) state alternative requirements the manifold could fulfill. To prove that these requirements are fulfilled is beyond the scope of this thesis, wherefore we simplify matters by constructing a compact manifold.

[10]Note that according to Corollary (7) the corresponding eigenvalues of the flow are of absolute value smaller or greater than one, respectively.

the trajectory is converging towards the center manifold of stationary points, but not to which of the stationary points in particular.

Further information about transitional dynamics can be drawn from conclusion (5) from theorem (13). This conclusion provides information about dynamics in the neighborhood of one fixed point. For each fixed point $x$ of the center manifold there exists a stable submanifold $W_x^{ss}$ and an unstable submanifold $W_x^{uu}$ called fibers, which are of dimension $n_s$ and $n_u$, respectively. Each fiber is tangent to the subspace spanned by the corresponding eigenvectors. That is, the unstable and stable fibers are tangent to the eigenvectors associated to eigenvalues with a positive and negative real part. According to conclusion (6), the fibers are continuous in $x$, that is, the fibers of point $\hat{x}$ close to $x$ will be close to the fibers of point $x$. The foliation is illustrated in Figure 2. It is an example of a two-dimensional system with a one-dimensional center manifold of stationary equilibria. All solutions converge towards the center manifold. Therefore, the stable manifold is two-dimensional, whereas each fiber is one-dimensional. The center manifold of stationary points is represented by a solid line. Trajectories converging towards the center manifold are dashed, and each initial point is indicated by a cross. It can be recognized that each fiber represented by a trajectory converges to a different point on the center manifold.

We want to emphasize that theorem (13) includes theorem (8), if it is applied to a single, hyperbolic fixed point. Moreover, conclusions from theorem (8) applied to one single point of a center manifold of stationary points match conclusions of theorem (13), although it is not possible to apply theorem (8) formally. The reason is that theorem (8) applies a linearization. Since the center manifold consists of stationary points there is no movement on the center manifold and, therefore, the non-linear system and its linearization are topologically equivalent. However, note that theorem (13) provides additional conclusions about the center manifold as a whole.
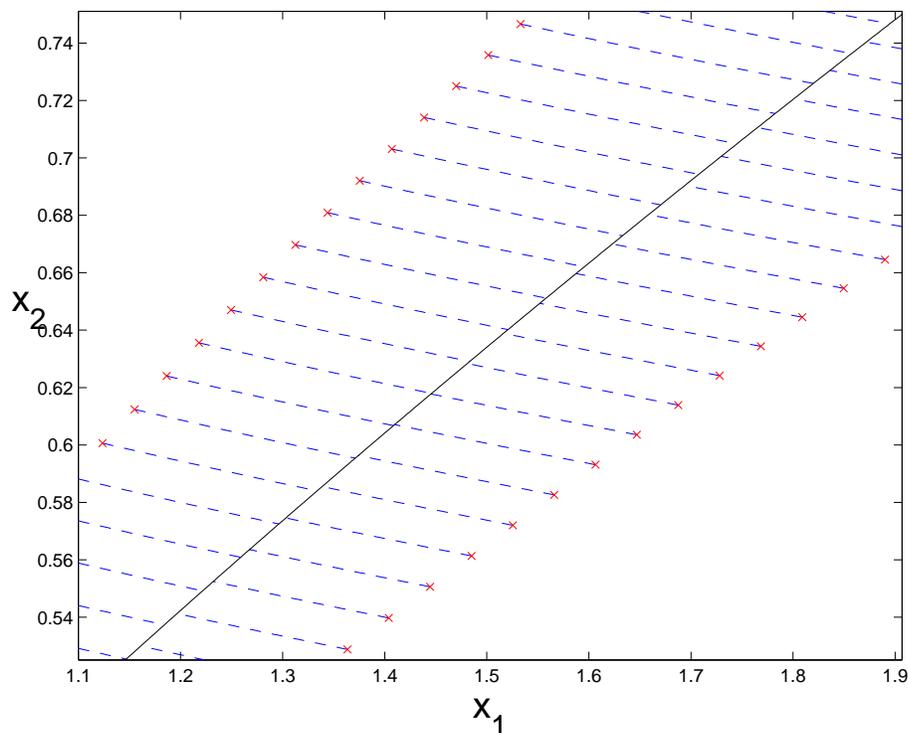
Figure 2: Center manifold of stationary points and converging trajectories

We can already conduct an economic interpretation, suitable for a macroeconomic infinite horizon optimization model. Often, the dynamic system is interpreted to describe the evolution of an economy under study. Each trajectory, then, represents an economy that differs at the beginning of the time horizon only in the initial conditions with respect to the other economies. These initial conditions often comprise the stock of physical capital, human capital or knowledge, that each economy possesses. Then, each economy converges to a different long-run equilibrium (i.e. stationary point), depending on the initial conditions. This dynamic behavior is very different to a system, which possesses an isolated fixed point. In the latter case, all economies would converge towards the same long-run equilibrium, no matter what their particular initial condition consists of.

## 2.3 Topology of two-dimensional stable manifolds

Since the method presented in this dissertation aims at solving models with multi-dimensional stable manifolds numerically, we want to illustrate potential dynamic patterns that could emerge in these type of models. However, the outline is not a complete discussion of the topology. Instead, we focus on the special case of a two-dimensional stable manifold with (distinct) real associated eigenvalues. This is most relevant for economics since many economic models with multi-dimensional stable manifolds exhibit this characteristic, e.g. the Jones (1995a) model that we analyze in this thesis.[11]

We consider a dynamic system exhibiting a hyperbolic fixed point. We focus this exposition on the stable manifold of the fixed point, since the solution of the underlying optimization problem usually requires convergence towards an interior steady state.[12] As a concise example we consider the two dimensional system

$$\dot{x}_1 = \lambda_1 x_1$$

$$\dot{x}_2 = \lambda_2 x_2 \tag{15}$$

with the eigenvalues $\lambda_1$, $\lambda_2 < 0$. The system possesses a fixed point at the origin.

This example is simple, however, it will prove to be instructive. First of all, the system matrix is diagonal and, therefore, equivalent to every diagonalizable $2 \times 2$ Matrix $A$. Hence, the system's behavior is equivalent to any two-dimensional system $\dot{x} = Ax$ with $A$ being diagonalizable in $\mathbb{R}$. Moreover the Hartman-Grobman theorem states that the behavior of any non-linear dynamic system in the neighborhood of a

---

[11]One prominent counter-example is the Lucas (1988) model, which is also analyzed in this thesis. For a special set of parameters, the models's stable manifold exhibits a pair of conjugate complex eigenvalues.

[12]We did not state the an optimization problem so far, but will do so in the following Section. However, at this point it is important that the dynamic system originates from an intertemporal optimization problem, and, therefore, optimal solutions usually are required to reach an interior steady state as $t \to \infty$.

hyperbolic fixed point is equivalent to its linearized system. Therefore, the system's behavior is equivalent to any two-dimensional dynamic system $\dot{x} = f(x)$, which Jacobian matrix $Df$ is diagonalizable in $\mathbb{R}$ at a hyperbolic fixed point $y^*$. In the light of these conclusions, we think that the study of system (15) is quite instructive.

The analytical solution of system (15) is straightforward and reads

$$x_1(t) = x_1(0)e^{\lambda_1 t} \tag{16}$$

$$x_2(t) = x_2(0)e^{\lambda_2 t} \tag{17}$$

Since $\lambda_1$, $\lambda_2 < 0$, solutions converge towards the origin independent of the initial value $(x_1(0), x_2(0))$.

The stable manifold can be divided into two submanifolds associated with the eigenvalues $\lambda_1$ and $\lambda_2$. In this case, the manifolds coincide with the sets $(\mathbb{R}, 0)$ and $(0, \mathbb{R})$, respectively, which represent the axes $x_1$ and $x_2$. Although solutions on both manifolds converge towards the fixed point, the speed differs according to equation (16) and equation (17), respectively. If $\lambda_1$ and $\lambda_2$ differ in magnitude, solutions will first converge towards the submanifold associated with the smaller eigenvalue, in absolute terms, because the other component is decreasing faster. This can be seen in Figure 3 (i) and (ii). In both Figures, system (15) is simulated with initial conditions on the unit circle, employing a fourth order Runge-Kutta procedure. In Figure (i), $\lambda_1$ and $\lambda_2$ differ by the factor 2. Solutions slightly bent towards the $x_2$ axis, the stable submanifold associated with the smaller eigenvalue $\lambda_2$, in absolute terms. In Figure (ii), $\lambda_1$ and $\lambda_2$ differ by the factor 20. In this case, solutions bent heavily towards the manifold represented by the $x_2$ axis.

The reason to focus on the dynamic behavior on two-dimensional stable manifolds is because qualitatively different behavior could possibly occur, compared to systems exhibiting a one-dimensional stable manifold. On a one-dimensional stable manifold, transitional dynamics are bound to be monotonic. That means, all variable mono-
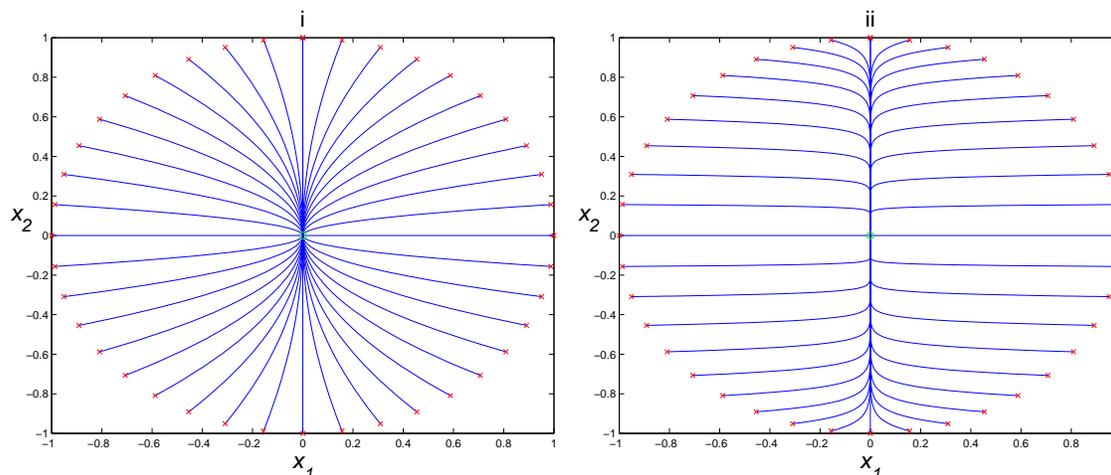
Figure 3: Phase diagram of system (15) with $\frac{\lambda_1}{\lambda_2} = 2$ and $\frac{\lambda_1}{\lambda_2} = 20$

tonically increase or decrease along their time paths towards the fixed point.[13] On a two-dimensional stable manifold, however, the variables may overshoot. By that we denote the phenomenon that variables first decrease and then increase in time along their transition path, or vice versa.

If the ratio of stable eigenvalues is one, trajectories in Figure 3 would be linearly connecting the initial point and the fixed point. Therefore, overshooting cannot occur, no matter which starting point is chosen. We argue that if the ratio of stable eigenvalues differs considerably, e.g. as shown in Figure 3 (ii), overshooting is quite likely to occur, depending on the initial conditions. This cannot be recognized from Figure 3 right away, because the axes and the submanifolds coincide. Therefore, we rotate the axes by 45° by transforming the system matrix equivalently with the

---

[13]Following the principles of dynamic system, the adjustment path could exhibit non-monotonic behavior that stems from non-linearities. However, the dynamic system results from an optimization problem. Therefore, trajectories projected into the state space cannot intersect and non-monotonic behavior of the state variable is not optimal, if the system is autonomous. However, control variables could exhibit non-monotonic adjustment.

rotating matrix

$$R = \begin{pmatrix} \cos(\pi/4) & \sin(\pi/4) \\ -\sin(\pi/4) & \cos(\pi/4) \end{pmatrix}$$

In Figure 4 (i) almost every trajectory exhibits an overshooting behavior with respect to time, either in coordinate $x_1$ or in coordinate $x_2$. This can be seen by the fact that almost every trajectory exhibits an extremum. The only exceptions are trajectories that follow the 45° angle bisector. These exactly take the shape of the two submanifolds. In Figure 4 (ii) we display the graph of a specific adjustment path of $x_1$ with respect to time. The overshooting behavior can be recognized.
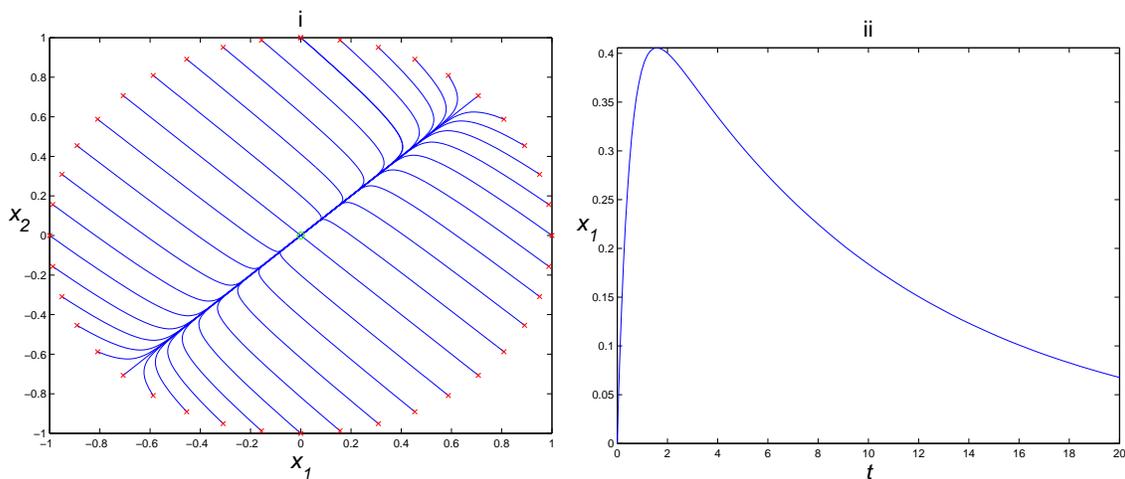


Figure 4: Rotated phase diagram of system (15) with $\frac{\lambda_1}{\lambda_2} = 20$ and one selected trajectory in time

From the analytical solution (16) and (17) we know that the convergence of the solution can be divided into two components represented by the eigenvalues $\lambda_1$ and $\lambda_2$. Both components decay with different speed. Therefore, the solution converges towards the submanifold represented by the smaller eigenvalue in absolute terms first, because the component represented by this eigenvalue decays slower. In a second phase, the solution converges along the one-dimensional submanifold towards the

fixed point.[14] This division of convergence into two phases is more pronounced, the higher the ratio of the stable eigenvalues is.

The examples above illustrates that overshooting behavior is more likely the higher the ratio of stable eigenvalues is. Overshooting behavior does not arise, if the transition takes places along one of the submanifolds, or if the submanifolds coincide with the axes. This illustration serves as a exposition about the possible pattern of transitional dynamics that can be inferred from the eigenvalues at the fixed point. Eventually, non-linearities could further enrich transitional dynamics.

## 2.4  Stiff differential equations

In this Section, we want to focus on the numerical part of multi-dimensional transitional dynamics. If the state space is multi-dimensional, a trajectory representing a transition path has to be computed that lies on a multi-dimensional stable manifold. We will exploit the insights on topology of the previous Section to describe the phenomenon of stiff differential equations. A general description of stiff differential equations can be found in Press et al. (1989, pp. 734) or Ascher and Petzold (1998, pp. 49 and pp. 214), however, we modify the illustration according to the problems that will be described later in this thesis.[15]

Consider system (15) with positive, real eigenvalues $\lambda_1$ and $\lambda_2$. Then, the origin is unstable. The case of stiff differential equations occurs, if the solutions on the submanifolds are of very different scale. More precisely, the transition speed on the submanifolds, represented by the corresponding eigenvalues, differs a lot. For

---

[14]Strictly speaking, the solution is, of course, only very close to the submanifold, because solutions cannot intersect.

[15]We also slightly deviate from the definition of stiff differential equations usually given in textbooks, because a description of stability regions of initial value problems would be beyond the scope of this thesis.

illustration, we assume the eigenvalues to take the values

$$\lambda_1 = 1$$

$$\lambda_2 = 0.1$$

and, hence, the numerical solution is

$$x_1(t) = x_1(0)e^t \tag{18}$$

$$x_2(t) = x_2(0)e^{0.1 \cdot t} \tag{19}$$

Now consider the two-point boundary value problem of system (15) together with the boundary conditions

$$\sqrt{(x_1(0))^2 + (x_2(0))^2} < 10^{-6} \tag{20}$$

$$x_1(T) = x_2(T) = 1 \tag{21}$$

with an unknown time $T$. Graphically, this is the problem of finding a trajectory that starts on or in a circle around the origin of radius $10^{-6}$ and ends at a specific point in the phase space, here $(1,1)$.[16]

The problem exhibits one degree of freedom to make it analytically tractable. If either $T$ is fixed (sufficiently high) or the initial conditions (20) are forced with equality the solution to the problem is unique. From the final boundary conditions we get

$$x_1(0)e^T = 1$$

$$x_2(0)\left(e^T\right)^{0.1} = 1$$

which yields

$$x_1(0) = x_2(0)^{10} \tag{22}$$

---

[16]The adept reader may notice the similarity to the backward integration procedure as introduced by Brunner and Strulik (2002). Indeed, we want to show problems that arise employing this procedure for multi-dimensional stable manifolds.

We set $x_2(0) = 10^{-7}$ and, therefore, a value of $x_1(0) = 10^{-70}$ would fulfill both boundary conditions (20) and (21). For $T$ we calculate $T = 70$.

The numerical problem now lies in the fact that both initial conditions are of very different scale. Note that the coordinates $x_1$ and $x_2$ also represent the eigenspaces. Therefore, this example can be generalized to the conclusion that the initial condition for meeting an arbitrary point has to be such that the eigenvector with the smaller eigenvalue has to add considerably less to the initial condition than the eigenvector with the greater eigenvalue. Roughly speaking, the ratio of both eigenvalues can be interpreted as the elasticity of the norm of both eigenvectors.[17] When the initial conditions are stored in the computer memory this can only be done with machine precision (see Press et al., 1989, pp. 28). Most machines have a machine precision of $\hat{\epsilon} \approx 10^{-15}$. Therefore, the absolute error for storing $x_2(0)$ in this example is $\epsilon \approx 10^{-22}$. If, however, some of this error spills over to the other eigenvector, $x_1(0)$ may depart from its correct value in the order of $\epsilon$. In this simple example this cannot happen, because the eigenvectors coincide with the coordinates. In a more general case it is only possible to store the vector $x_1(0)$ with precision $\epsilon$. However, $\epsilon \gg x_1(0)$, and therefore, the final boundary condition (21) cannot be met. Moreover, even if the initial condition for both eigenvectors were correct, errors would spill over to each other eigenvalue during the integration process. Therefore, the error along the eigenvector associated to the bigger eigenvalue would be very high. We can conclude in which direction the final boundary condition (21) will be missed. Since the correct starting value $(x_1(0), x_2(0))$ is biased towards $x_1(0)$, the solution will remain close to the associated submanifold of $\lambda_1$.

As a numerical experiment we simulate system (15) with rotated axes by 45° with different but positive values for $\lambda_1$ and $\lambda_2$. Again, the rotation is conducted

---

[17]This holds in this simple example, since $x_1(0)^{\frac{\lambda_1}{\lambda_2}} = x_2(0)$.

for numerical reasons. Otherwise, the submanifolds coincide with the axes and, therefore, the different scale of both solution components would not cause a numerical problem. We choose initial values equally distributed on a circle of radius $10^{-6}$ around the fixed point. Figure 5 (i) shows the trajectories for a simulation exercise with a ratio of eigenvalues equal to 1.1. It can be recognized that the trajectories are attracted by one submanifold. Unlike in Figure 4 (i), solutions are now attracted by the submanifold associated with the larger, in absolute terms, eigenvalue $\lambda_1$. In Figure 5 (ii) the ratio of eigenvalues is 10. Now, every trajectory is attracted by the $\lambda_1$-submanifold such that they are indistinguishable. While simulations for Figure 5 (i) and (ii) were conducted with a moderate number of trajectories and, therefore, a rough division of the circle, the same phenomenon maintains with a finer division of initial values on the circle. For a run of 4000 trajectories only one trajectory was distinguishable from the $\lambda_1$-submanifold. Moreover, for simulation exercises we took special measures for solving the system of stiff differential equations by employing the *ODE15s.m* MatLab procedure. This is a medium order procedure specialized in solving stiff differential equations. We constrain the local error to $10^{-13}$, close to the machine epsilon of $2.2 \cdot 10^{-16}$.

This illustration shows that it might be difficult or even impossible to meet any pre-specified point in the phase space as final boundary condition. This problem is more pronounced the higher the ratio of the eigenvalues is. In other words, if starting values close to the submanifold associated with the smaller eigenvalue are chosen, small changes of the initial values cause large changes in the resulting trajectories. The problem of hitting a pre-specified point in the phase space is ill-conditioned. The mathematical literature suggest measures to lessen these problems. If special difference schemes, mostly implicit schemes, were chosen for integration, and if a high precision is demanded from the integration procedure, the problem of stiff dif-
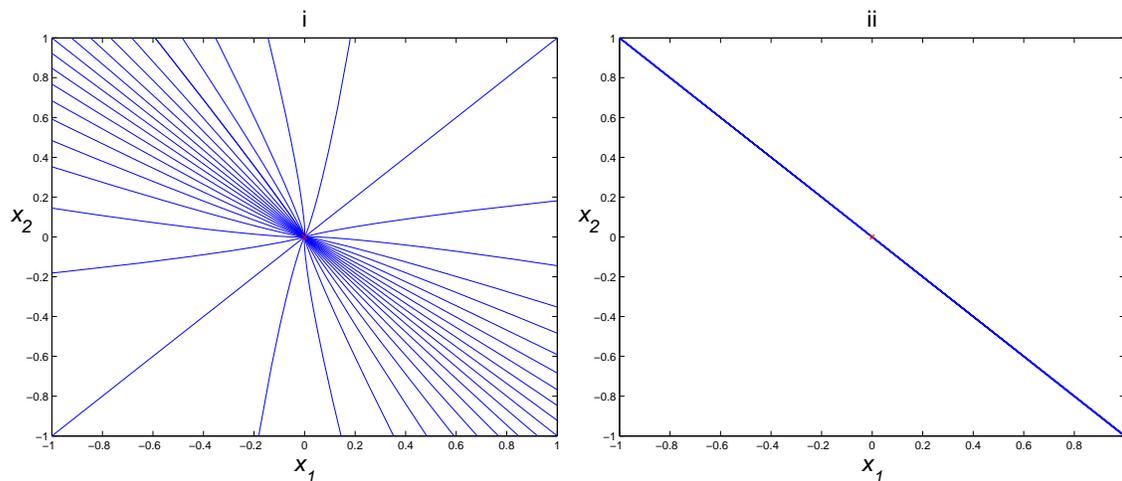
Figure 5: Phase diagram of system (15) with $\frac{\lambda_1}{\lambda_2} = 1.1$ and $\frac{\lambda_1}{\lambda_2} = 10$

ferential equations can be overcome to some extend. However, the problem remains ill-conditioned and, therefore, it is still impossible to solve very stiff systems accurately.

# 3 Multi-Dimensional Transitional Dynamics: A Simple Numerical Procedure

This Section is based on the paper of Trimborn, Koch and Steger (2007). All contributions of this Section are the author's own if not indicated otherwise.

## 3.1 Introduction

For applying the relaxation algorithm to solve infinite-horizon, continuous time optimization models we follow Pontryagin's maximum principle. First order conditions of the optimization problem comprise a set of differential equations augmented by algebraic equations that have to hold at all points of time. Initial and final boundary conditions turn the mathematical problem into a two-point boundary value problem. Then, a trajectory of the differential equations has to be found, which satisfies the algebraic constraints as well as initial and final boundary conditions. Relaxation algorithms solve for the solution trajectory simultaneously, and neglect the direction implied by time. This is in contrast to various shooting algorithms, which start at one boundary and integrate with respect to time. For these procedures, an iteration has to be applied, until the other boundary condition is satisfied.

The implementation of the relaxation algorithm follows Press et al. (1989). At first, we fix a mesh of points in time. The differential equations are discretized along this mesh grid. These difference equation represent a set of non-linear equations carrying information about the slope of the solution trajectory. The set of equations is augmented by the algebraic equations, which have to hold at all points of time, and equations representing initial and final boundary conditions. The solution to this square set of non-linear equations, then, represents the solution trajectory at the mesh of points. In principle, it can be solved by any standard routine suitable for non-linear equations. However, the number of equations may be very large,

because it increases with the number of mesh points. Note that the Jacobian matrix of the equations inherits a sparsity pattern, which is due to the time structure of the problem. Therefore, we apply a specialized Gauß-Newton procedure, which exploits the sparsity of the Jacobian matrix and allows to solve models employing a high number of mesh points.

We programmed the algorithm in Matlab. A detailed description of this work can be found in the appendix, together with the code for the main parts of the algorithm. Moreover, the code generating the simulation results presented in this thesis can be found on the accompanied CD. We want to emphasize that great effort has been made to construct the code and its application as user-friendly as possible. A researcher, who intends to simulate a particular model can use the code as a "black-box", since he does not need to know how the algorithm works. The only information that has to be provided is the dynamic system together with parameter values and a guess of the steady state of the model.

The state-of-the-art of the relaxation algorithm in the literature is described in Press et al. (1989).[18] They construct the algorithm for a finite time horizon and employ the midpoint rule for discretization of the differential equations. Koch (2003) suggests to employ the relaxation algorithm for infinite-horizon problems. He suggests to transform time from the interval $[0, \infty)$ to $[0, 1)$ to solve the system for the full horizon. This idea is borrowed for this thesis, however, the allocation of the time mesh given by Koch's transformation is fixed. Therefore, the algorithm does not converge for Koch's time transformation, wherefore we introduced the parameter $\nu$ to allow for a more flexible allocation of the mesh. Moreover, we provide the theoretical background for this transformation. Since relaxation methods were only employed for finite horizon problems the mathematics literature does not discuss how

---

[18]Ascher and Petzold (1998) do not allow for algebraic equations and describe the Gauss-Newton procedure only vaguely.

final boundary conditions can be formulated for infinite horizon problems. Koch also does not discuss this issue. Therefore, contributions to this topic are the author's own. Furthermore, we are the first to exploit symmetry in the Jacobian matrix, which enhances the speed of the algorithm considerably. Finally, great emphasis has been made to adapt the algorithm to the requirements of economic researchers. For example, the algorithm allows to simulate the impact of specific economic shocks, i.e. transition from one steady state to another steady state.

In Section 3.2 we describe the relaxation algorithm in detail. In Section 3.3, we compare it to alternative solution algorithms. In Section 3.4, we present the Jones (1995a) model and the Lucas (1988) model as two illustrative examples, and in Section 3.5 we summarize.

## 3.2 The relaxation procedure

### 3.2.1 Statement of the mathematical problem

Economic dynamic optimization problems frequently lead to a system of differential equations potentially augmented by algebraic equations:

$$\dot{x} = f(t, x, y) \tag{23}$$

$$0 = g(t, x, y) \tag{24}$$

with $x \in \mathbb{R}^{n_d}$, $y \in \mathbb{R}^{n_a}$, $f : (\mathbb{R} \times \mathbb{R}^{n_d} \times \mathbb{R}^{n_a}) \to \mathbb{R}^{n_d}$ and $g : (\mathbb{R} \times \mathbb{R}^{n_d} \times \mathbb{R}^{n_a}) \to \mathbb{R}^{n_a}$. We define the total dimension of the problem as $n := n_d + n_a$. The algebraic equations and, hence, $y$ are not required to appear in the problem. Then, $n_a = 0$ and $n_d = n$ would hold. Usually, the system has to be solved over an infinite horizon with boundary conditions

$$h_i(x(0), y(0)) = 0 \tag{25}$$

$$\lim_{t \to \infty} h_f(x(t), y(t)) = 0 \tag{26}$$

whereby $h_i : \mathbb{R}^n \rightarrow \mathbb{R}^{n_i}$ defines $n_i$ initial boundary conditions and $h_f : \mathbb{R}^n \rightarrow \mathbb{R}^{n_f}$ defines $n_f$ final boundary conditions such that $n_i + n_f = n_d$.

The system (23) and (24) together with (25) and (26) is labeled two-point boundary value problem, since a system of differential equations has to be solved subject to boundary conditions at the beginning and the end of the time horizon. The specific characteristics of the problem at hand are that equation (23) has to be solved for an infinite time horizon and that the solution is bound to a manifold defined by (24). In case of an infinite time horizon we require the final boundary conditions to enforce convergence towards a manifold of dimension $n_f$ or less. In many economic applications the desired solution is known to converge towards a single point. We do not treat this case explicitly but classify it under equation (26). Often the algebraic equations are differentiated with respect to time such that the resulting system of differential equations is square. However, we do not require this. We merely assume that the system (23) and (24) is of differential index one, that is, a one-time differentiation of (24) with respect to time would yield a square system of differential equations.[19]

### 3.2.2  Description of the relaxation procedure

Finite difference procedures like the relaxation procedure are used in various fields of numerical mathematics, e.g. the solution of partial differential equations, initial value problems, boundary value problems and delayed differential equations. While the first variant for solving initial value problems goes back to Leonard Euler, Runge and Kutta generalized finite difference procedures for one-step solution algorithms of initial value problems (see Runge, 1895, and Kutta, 1901). Solution of bound-

---

[19]For more on the differential index of differential algebraic equations see Ascher and Petzold (1998), pp. 231. Differential algebraic equations of higher differential index exhibit far more complex characteristics and are, hence, more complicated to solve numerically.

ary value problems with finite difference procedures go back to Collatz (1951), Fox (1957), and Keller (1968). Here, we employ the relaxation algorithm, a particular type of finite difference procedure, to solve boundary value problems (see e.g. Press et al., 1989, pp. 645, or Ascher and Petzold, 1998, pp. 193).

The general idea of Relaxation is to solve for the whole transition path simultaneously. The differential equations are approximated by finite difference equations on a mesh of points in time. This set of equations, augmented by equations representing initial and final boundary conditions, is a set of non-linear equations. Thereby, the variable values at the mesh points represent the unknowns. Then, this set of non-linear equation is solved for, not taking into account the time-structure the system took on from the difference equations. For solving a set of non-linear equations, a rough initial guess of the true solution is needed that is updated in a sequence of iterations. This initial guess 'relaxes' towards the true solution in every iteration along the path as a whole.

Figure 6 illustrates the adjustment of an initial guess towards the saddle path in the Ramsey-Cass-Koopmans model. The initial guess starts with a fixed initial value of the state variable $k$ and an arbitrary initial value of the control variable $c$. It consists of 30 mesh points lined up equidistantly between the starting point and the known steady state of the model.

We divide the outline of the algorithm into three parts. In the first part, we reformulate the problem consisting of equations (23), (24), (25), and (26) to match the requirements of the relaxation algorithm. By transformation of the independent time variable we can solve the infinite horizon problem conveniently. In the second part, we define a mesh of points in time and discretize the differential equations. In the third part, we solve the resulting set of non-linear equations, taking into account its special sparsity structure.
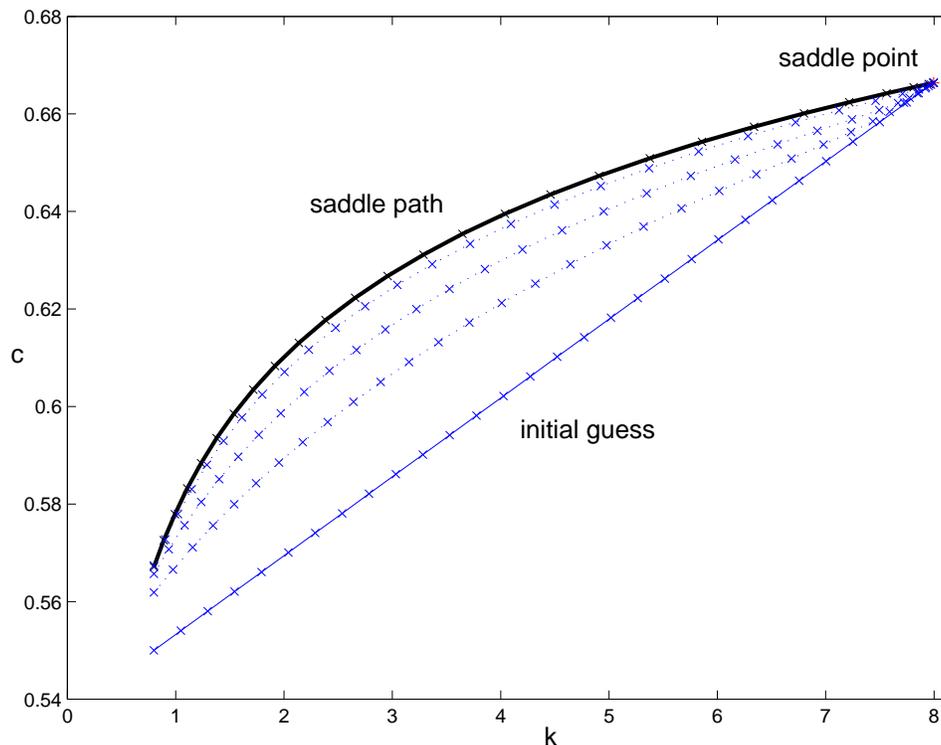
Figure 6: Relaxation in the Ramsey-Cass-Koopmans model

### The mathematical problem

We rescale the time range $\mathbb{R}^+$ by introducing a new time parameter $\tau$ running from 0 to 1

$$\tau = \frac{\nu t}{1 + \nu t} \tag{27}$$

where $\nu > 0$ is a fixed parameter.[20]  It will be needed as a degree of freedom to change the allocation of the time mesh.[21]  Note that (27) is a strictly monotonic relation for

---

[20]Prof. Koch suggested an earlier version of the time transformation function, for which it was not possible to change the allocation of the mesh.

[21]The higher the value of $\nu$, the denser the mesh is allocated at the origin of the time interval.

$t \geq 0$ and, therefore, we can transform it into $t = \frac{\tau}{\nu(1-\tau)}$. We express equations (23) and (24) with respect to $\tau$ according to

$$\frac{dx}{d\tau} = \frac{f\left(\frac{\tau}{\nu(1-\tau)}, x, y\right)}{\nu(1-\tau)^2} =: \xi(\tau, x, y) \tag{28}$$

$$0 = g\left(\frac{\tau}{\nu(1-\tau)}, x, y\right) =: \phi(\tau, x, y) \tag{29}$$

with boundary conditions (25) and (26) with respect to $\tau$

$$h_i(x(0), y(0)) = 0 \tag{30}$$

$$h_f(x(1), y(1)) = 0 \tag{31}$$

According to the original problem, the final boundary condition (31) has to be $n_f < n_d$ dimensional. Admittedly, for many economic applications it is known that $x$ and $y$ approach a stationary point as $t \to \infty$. In this case there are two possibilities for reformulation. One could construct (31) by setting all the time derivatives to zero, e.g. $\dot{x}_1^2 + \ldots + \dot{x}_n^2 = 0$, or one could omit some of the available information by demanding only some variables to be stationary in the long-run. Theoretically, the latter may involve the search of well defined final boundary conditions by trial and error. In practise, however, boundary conditions always worked for every combination of variables. We will discuss this below employing the Ramsey model as a concise example. To sum up, we have $n_d$ differential equations, $n_a$ algebraic equations, $n_i$ initial conditions and $n_f$ final conditions such that

$$n_i + n_f + n_a = n_d + n_a = n.$$

The functional form of time transformation includes some arbitrariness, since there are many possibilities to map the interval $[0, \infty)$ to $[0, 1)$ with a strictly monotonic relation. However, we have to choose a transformation such that the differential equations (28) is well defined. Consider a fixed point $(x^*, y^*)$ such that $f(\infty, x^*, y^*) =$

0. It is not possible to evaluate $\xi(1, x^*, y^*)$ because $\frac{dt}{d\tau}|_{\tau=1} = \infty$ will hold for every time transformation. We at least demand a trajectory $(x, y)$ converging to $(x^*, y^*)$ to fulfill

$$\lim_{\tau \to 1,} \xi(\tau, x(\tau), y(\tau)) = 0 .$$

or, stated differently, $f(t, x, y)$ should converge faster to zero than $\left(\frac{dt}{d\tau}\right)^{-1}$. The time transformation at hand fulfills this condition. To see this, remember that the Hartman Grobman theorem and the Fundamental theorem of hyperbolic invariant manifolds state that $(x, y)^t$ converge towards a fixed point exponentially.[22] However, this is faster than convergence of the term $\left(\frac{dt}{d\tau}\right)^{-1} = \nu(1 - \tau)^2$, since the latter converges only quadratically towards zero.

### Discretization of the system

We define a mesh of $M$ points in transformed time $\tau$ by $T = \{\tau_1, \ldots, \tau_M\}$. The mesh points $\tau_i$ are placed equidistantly on the interval $[0, 1]$. To unify notation, we subsume the variables $x$ and $y$ to $z$. Therefore, we define $z = \{z_1, \ldots z_M\}$ as a vector associated to the time mesh $T$ such that $z_i$ represents $(x, y)^t$ evaluated at time $\tau_i$. We use the midpoint of each interval $(\tau_i, \tau_{i+1})$ to discretize the differential equation according to

$$\frac{x_{i+1} - x_i}{\tau_{i+1} - \tau_i} = \xi\left(\frac{\tau_i + \tau_{i+1}}{2}, \frac{z_i + z_{i+1}}{2}\right) \qquad i = 1, \ldots, M - 1 \qquad (32)$$

Considering one interspace, i.e. one index $i$, we can construct an $n_d$ dimensional error function $H$ from equation (32), $H(T \times \mathbb{R}^n)^2 \to \mathbb{R}^{n_d}$:

$$H(\tau_i, z_i, \tau_{i+1}, z_{i+1}) = x_{i+1} - x_i - (\tau_{i+1} - \tau_i)\xi\left(\frac{\tau_i + \tau_{i+1}}{2}, \frac{z_i + z_{i+1}}{2}\right) \qquad (33)$$

---

[22]For this consideration, we assume the system to be autonomous 'in the long-run', which means that after some $\tilde{t}$ it is no longer directly time dependent.

This adds up to $(M-1) \cdot n_d$ equations. We define $n_i$ initial conditions via (30) and $n_f$ final conditions via (31), and force the algebraic equations to be fulfilled at every mesh point $\tau_i$. Therefore, the latter add $M \cdot n_a$ equations. Altogether given a mesh $T \epsilon \mathbb{R}^M$ this defines a system of $n \cdot M$ equations in $n \cdot M$ unknowns, since $z = (z_1, \ldots, z_M)^t \epsilon \mathbb{R}^{n \cdot M}$.

For convenience, we list the equations according to the unknown vector $z_i$ considering the time structure. We start with the initial conditions, which only depend on $z_1$. We continue with the difference equation involving $z_1$ and $z_2$ together with the algebraic equations for $z_1$. After proceeding through the whole mesh of points the final boundary conditions, which depend only on $z_M$, close the list together with the algebraic equation for $z_M$. We rename the equations according to

$$
\begin{aligned}
E_0(z_1) &:= h_i(z_1) \\
E_k(z_i, z_{i+1}) &:= \begin{pmatrix} g(z_i) \\ H(y_i, y_{i+1}) \end{pmatrix} \qquad i = 1, \ldots, M-1 \\
E_M(z_M) &:= \begin{pmatrix} g(z_M) \\ h_f(z_M) \end{pmatrix}
\end{aligned}
$$

Together, this yields the set of non-linear equations

$$
E(z) \equiv \begin{pmatrix} E_0(z) \\ \vdots \\ E_i(z) \\ \vdots \\ E_M(z) \end{pmatrix} = \begin{pmatrix} \left( h_i(z_1) \right) \\ \vdots \\ \begin{pmatrix} g(z_i) \\ H(z_i, z_{i+1}) \end{pmatrix} \\ \vdots \\ \begin{pmatrix} g(z_M) \\ h_f(z_M) \end{pmatrix} \end{pmatrix} \tag{34}
$$

For discretizing the differential equations we choose the midpoint rule, which is a second order discretization scheme (see e.g. Press et al., 1989, p. 710). This means that given the true solution $\tilde{z}$, equation (32) can be obtained by a Taylor expansion of $\tilde{z}$, if the summands containing the first and second derivative of $z$ are included,

but higher order derivatives are truncated. It is apparent that the more summands of Taylor's expansion are considered, i.e. the higher the order of the rule, the smaller is the truncation error, and the better is the approximation. In general, using a $k$-th order method will reduce the computational error by order $k$ if the step size of the mesh is decreased. This means that there is a constant $c$ such that the computational error $\epsilon$ follows

$$|\epsilon| < c(\tau_{i+1} - \tau_i)^k \ . \tag{35}$$

We provide evidence using a computational exercise that the midpoint rule is indeed of second order in the next Subsection.

For selecting a difference scheme there seems to be numerous possibilities, since at first glance every scheme suitable for an initial value procedure could also yield a discretization according to (32). If a higher order discretization, e.g. a fourth order Runge-Kutta procedure, would be chosen, truncation errors would be lower and, therefore, the global error would reduce faster according to equation (35) if the mesh points are placed more densely. However, the difference scheme has to fulfill a number of additional requirements. First of all, the rule has to be symmetric. This is intuitive, since the relaxation procedure is symmetric with respect to the direction of time and, therefore, any direction implied by the difference scheme would be arbitrary. Moreover, this choice has a theoretical foundation originating from stability considerations. A detailed description of stability in the context of numerical solutions of differential equations is beyond the scope of this thesis and can be found in Ascher and Petzold (1998). Roughly speaking, a difference scheme is called stable, if it yields a qualitatively correct representation of the solution, even for a moderate number of mesh points. From initial value problems it is known that systems with a negative eigenvalue should be solved with implicit, backward schemes, whereas system with a positive eigenvalue should be solved with an ex-

plicit, forward scheme.[23] For systems exhibiting positive and negative eigenvalues symmetric schemes are known to guarantee stability. Many economic models exhibit exactly this characteristic, since the resulting saddle-point structure is desirable from the perspective of dynamic optimization. Therefore, we have to rely to a symmetric difference scheme.[24]

The second requirement for a difference scheme is that it does not evaluate the function $\xi$ (as defined in equation (28)) for $\tau = 1$, since $\xi$ is not defined at this point. Because of symmetry, then, we cannot evaluate $\xi$ at any mesh point, but only in the interspace of each interval. The midpoint rule is one of the only discretization schemes that fulfills both requirements. To sum up, the midpoint rule is a dicretization scheme suitable for constructing a general purpose algorithm for the problem (28)–(31).

### Solution of the system of non-linear equations

In principle, equation (34) could be solved by any algorithm suitable for solving systems of non-linear equations. The problem is that system (34) can be quite big, because it grows not only with the dimension of model variables $n_d$ and $n_a$, but also with the number of mesh points $M$. For example, considering a small scale model with $n_d = 3$ and $n_a = 0$, for a mesh of $M = 1000$ points system (34) exhibits 3000 equations. Solving this number of equations would be very time-consuming and permit the solution of higher-dimensional models. However, relaxation algorithms exploit the time structure of system (34). Since each block of equations $E_i$ depends only on neighboring unknowns $z_i$ and $z_{i+1}$, a Gauß-Newton procedure can utilize

---

[23]For the Euler method, for example, the implicit (backward) scheme turns into the explicit (forward) scheme and vice versa if time is reversed.

[24]It is known from initial value problems that the midpoint rule exhibits desirable stability properties beyond this description, i.e. 0-stability, A-stability and AN-stability.

the sparsity of the equations. Note, that we still solve the system (34) of equations simultaneously, but exploit the sparsity to economize on computational requirements.

A Gauß-Newton procedure requires an initial guess $z^0$ of the solution.[25] The initial guess is upgraded by linearizing system (34) at $z^0$ and finding the root of this linear equation, which then represents the updated solution. The updated solution is origin of another Newton step, until the the change of the solution $\Delta z$ is sufficiently small. Denoting the Jacobian matrix of $E(z)$ by $D_z E$ we solve in each iteration $k$

$$D_z E(z^k) \cdot \Delta z^k = -E(z^k) \tag{36}$$

for $\Delta z^k$ and change $z$ according to

$$z^{k+1} := z^k + \Delta z^k \tag{37}$$

Due to the ordering of subsystems $E$ equation (36) is of following form:

$$
\begin{pmatrix}
S^{0,R} & & & & \\
S^{1,L} & S^{1,R} & & 0 & \\
 & S^{2,L} & S^{2,R} & & \\
 & & \ddots & & \\
0 & & & S^{M-1,L} & S^{M-1,R} \\
 & & & & S^{M,L}
\end{pmatrix}
\begin{pmatrix}
\Delta z_1^k \\
\vdots \\
\Delta z_M^k
\end{pmatrix}
=
\begin{pmatrix}
-E_0(z^k) \\
\vdots \\
-E_M(z^k)
\end{pmatrix}
\tag{38}
$$

with Jacobian matrices $S^{i,L}$ and $S^{i,R}$ defined by

$$S^{i,L} = \frac{\partial E_i(z)}{\partial z_i}, \quad \text{and} \quad S^{i,R} = \frac{\partial E_i(z)}{\partial z_{i+1}} \qquad i = 0, \dots, M \tag{39}$$

Note from equation (33) that the partial derivatives of $H$ with respect to $z_i$ and $z_{i+1}$ differs only in their derivatives of $x_{i+1}$ and $x_i$, respectively, and this is plus or minus the identity matrix of dimension $n_d$. Therefore, $S^{i,R}$ can be computed from $S^{i,L}$ by substraction of $2 \cdot I_{n_d}$ for $i = 1, \dots, M-1$. The upper left matrix $S^{0,R}$ has $n_i$ rows and the lower right matrix $S^{M,L}$ only $n_a + n_f$ rows, whereas all other matrices $S^{i,L}$

---

[25]The superscript indicates the iteration number.

and $S^{i,R}$ , respectively, are $n \times n$. Hence, the system is not overdetermined. The solution $\Delta z$ can be computed by a specialized Gaussian algorithm. This algorithm starts in the upper left corner of the matrix and works downward block by block to the lower right corner. The result is a system in upper triangular form with a sequence of $n \times n_i$ non-zero blocks above the diagonal. Finally the vector $\Delta z$ can be computed from bottom to top.

We transform the first $(n + n_i + n_a, 2 \cdot n)$-dimensional part of the Jacobian matrix into a $(n + n_i + n_a, n + n_i + n_a)$-Identity matrix and a $(n + n_i + n_a, n - n_i - n_a)$ matrix only with elementary transformation of rows. The same operations are conducted with the corresponding part of vector E. For transforming the rows the algorithm includes a column pivot search. The $(n - n_i, n)$-block at the lower right is going to be the upper left part of the $(n + n_i + n_a, 2 \cdot n)$-matrix that is manipulated next. Figure 7 shows an example with $n_i = n_f = 2$, $n_a = 1$, $n_d = 4$, and, hence, $n = 5$. Possible non-zero entries in the sparse matrix are indicated by $x$. The left hand side shows the original Jacobian matrix before the elimination. The right hand side shows the Jacobian matrix after the first Gaussian elimination of the upper $8 \times 10$ block matrix. The upper right $5 \times 5$ matrix of this block is saved in memory. Although this block exhibits a sparsity structure, we do not exploit it further. The lower right $3 \times 5$ part of the matrix will be the upper left part of the matrix assigned for the next Gaussian step.

Step by step the matrix is transformed into the following shape

$$
\begin{pmatrix}
I & A_1 & & & & \\
 & I & A_2 & & 0 & \\
 & & I & A_3 & & \\
 & & & \ddots & & \\
 & 0 & & & I & A_{M-1} \\
 & & & & & I
\end{pmatrix}
$$

with $(n, n)$ identity matrices $I$ and $(n, n)$ matrices $A_k$, $k = 1, \ldots, M-1$. The matrices

$$
\begin{bmatrix}
x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x
\end{bmatrix}
\quad
\begin{bmatrix}
1\,o\,o\,o\,o\,o\,o\,o\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,1\,o\,o\,o\,o\,o\,o\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,1\,o\,o\,o\,o\,o\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,1\,o\,o\,o\,o\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,1\,o\,o\,o\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,1\,o\,o\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,o\,1\,o\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,o\,o\,1\,x\,x\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,o\,o\,o\,o\,o\,o \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x \\
o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,o\,x\,x\,x\,x\,x\,x
\end{bmatrix}
$$

Figure 7: Jacobian matrix before manipulation and after the first Gaussian step

$A_k$ are stored and the solution vector $\Delta z$ can be computed from bottom to top. The last $n$ arguments of $\Delta z$ equal $E$. For each block we have to solve

$$
\begin{pmatrix} I & A_i \end{pmatrix}
\begin{pmatrix} \Delta z_i \\ \Delta z_{i+1} \end{pmatrix} = E
\qquad i = M-1, \ldots, 1
$$

with $n$ dimensional vectors $\Delta z_i$, $\Delta z_{i+1}$ and $E$ . Since $\Delta z_{i+1}$ is known, $\Delta z_i$ can be computed by

$$
\Delta z_i = E - A_i \Delta z_{i+1} \qquad i = M-1, \ldots, 1.
$$

We iterate according to equation (36) until the euclidian norm of the change $\Delta z$ is small

$$
\frac{|\Delta z|}{n \cdot M} < \epsilon.
$$

To sum up, we can solve equation (36) by exploiting the sparsity of the Jacobian matrix for storage and computational requirements. Both increase only linearly as the number of mesh points $M$ increase, thus, allowing to solve systems using a high number of mesh points. Moreover, since each block increases with the model's dimension $n$, computational requirements for a higher dimensional model grow according to $o(n^3)$. The reason is that Gaussian elimination is known to be of order $o(n^3)$ for a $n$ dimensional matrix (see Deuflhard and Hohmann, 1993, p.7).

It is worth noting that the relaxation algorithm is able to solve systems exhibiting multi-dimensional stable manifolds and center manifolds of stationary equilibria conveniently. For solving systems exhibiting one or both of these characteristics no conceptual changes for the algorithm have to be made. The first characteristic entails no difficulties, because state and control variables are handled conceptually equally by the relaxation algorithm. An increase in the dimension of the control or state space merely causes the corresponding block-matrices to increase accordingly. Moreover, the relaxation algorithm can solve systems exhibiting a center manifold of stationary equilibria conveniently, because the knowledge of the stationary equilibrium that the economy reaches is not need to be known in advance. While many competing algorithms need this information, the relaxation algorithm determines the stationary equilibrium that is actually reached through iteration.

### 3.2.3  Implementation of the algorithm

To illustrate, we describe the steps which must be taken when implementing the relaxation algorithm using the Ramsey-Cass-Koopmans model (Ramsey, 1928; Cass, 1965; Koopmans, 1965) as an example. It is important to notice, however, that this description serves as an illustration only. The researcher who intends to solve a specific model numerically using the program (provided as a supplement to this thesis) need not follow these steps.

It is well known that this simple growth model exhibits saddle-point stability and hence the determination of the solution is all but trivial.[26] The model gives rise to a system of two differential equations for consumption $c$ and capital per effective labor $k$ (Barro and Sala-i-Martin, 2004, Chapter 2):

$$\dot{c} = \frac{c}{\theta}\left(\alpha k^{\alpha-1} - (\delta + \rho + x\theta)\right) \tag{40}$$

$$\dot{k} = k^{\alpha} - c - (n + x + \delta)k, \tag{41}$$

where $\alpha$ denotes the elasticity of capital in production, $n$ the population growth rate, $\delta$ the depreciation rate, $x$ the exogenous growth rate of technology, $\rho$ the parameter for time preference and $\theta$ the inverse of the intertemporal elasticity of substitution, respectively. The steady state is $k^* = \left(\frac{\alpha}{\delta+\rho+x\theta}\right)^{\frac{1}{1-\alpha}}$ and $c^* = (k^*)^{\alpha} - (n + x + \delta)k^*$ and is saddle point stable.

As a first step, one must choose a time mesh, i.e. a set of points in time at which the solution should be calculated. We select the time mesh to be uniform in the transformed time scale (as explained in Section 3.2.2).

Second, the two differential equations have to be transformed into two non-linear equations which describe the slope between two neighboring mesh points. These equations have to be satisfied between every two mesh points. For $M$ mesh points this leads to $2 \cdot (M - 1)$ nonlinear equations.

Third, two boundary conditions have to be chosen to complete the set of equations to $2 \cdot M$. In this example the relaxation algorithm needs one initial boundary condition and one terminal boundary condition. We set the initial value of the state variable (capital) equal to 10% of its steady state value. For the terminal boundary condition there are several possibilities to formulate an equation. It would be possible to choose each of the two equations (40) or (41) and set the RHS equal to zero.

---

[26]Nonetheless, the model is comparably simple in that the stable manifold is one dimensional. We will turn to a model with a multi-dimensional stable manifold below.

Here the steady state values for consumption and capital can be computed analytically and, therefore, we can set consumption equal to its steady state value as the terminal boundary condition. It should be noted that only one terminal condition is needed. Thus the algorithm does not make use of the knowledge of the steady state value of capital. It is reached automatically. However, different choices for terminal boundary conditions do not have any noticeable impact on the performance of the algorithm for this model.

At last, an initial guess for the solution has to be made. For instance, we can choose $c$ and $k$ to be constant at their steady state values $(c_t, k_t) \equiv (c^*, k^*)$.[27] The Newton procedure always converged quickly, indicating a high degree of robustness with respect to the initial guess.

### 3.2.4  Evaluation of the procedure and error estimation

For the special parametrization $\theta = \frac{\delta+\rho}{\alpha(\delta+n+x)-x}$ the representative consumer chooses a constant saving rate $s = \frac{1}{\theta}$ and hence the solution can be expressed analytically (Barro and Sala-i-Martin, 2004, pp. 106-110).[28] This allows us to compare the computed results with the analytical solution, which has a precision close to the machine epsilon. The relative error is computed for every mesh point. Table 1 shows the maximum relative error of consumption and capital per effective labor for different numbers of mesh points. In addition, the quadratic mean error of combined $c$ and $k$ provides information about the distribution of the error.[29] Table 1 reveals that multiplying the number of mesh points by $x$ reduces the maximum error of each solution vector by the factor $\frac{1}{x^2}$, which indicates the order 2 of the difference procedure. Even

---

[27]This is in contrast to Figure 6 where the initial guess is an upward sloping line.

[28]The analytical solution is $k(t) = \left[ \frac{1}{(\delta+n+x)\theta} + \left( k_0^{1-\alpha} - \frac{1}{(\delta+n+x)\theta} \right) e^{-(1-\alpha)(\delta+n+x)t} \right]^{\frac{1}{1-\alpha}}$ and $c(t) = (1 - \frac{1}{\theta})k(t)^\alpha$.

[29]It is defined as $\varepsilon = \frac{1}{NM} \sqrt{\sum_{i=1}^{N} \varepsilon_{c_i}^2 + \sum_{i=1}^{N} \varepsilon_{k_i}^2}$ with $\varepsilon_{c_i}$ and $\varepsilon_{k_i}$ denoting the relative error of $k$ and $c$ at mesh point $i$, respectively.

with a moderate number of mesh points and therefore a short computation time, a sufficiently high degree of accuracy can be achieved. Moreover, the accuracy can be improved to a very high degree by increasing the number of mesh points.[30] The

Table 1: Accuracy of the relaxation algorithm for the Ramsey-Cass-Koopmans model

| number of mesh points | max error $c$ | max error $k$ | mean error |
|---|---|---|---|
| 10 | $< 1.3 \cdot 10^{-2}$ | $< 3.4 \cdot 10^{-2}$ | $< 3.0 \cdot 10^{-3}$ |
| 100 | $< 1.1 \cdot 10^{-4}$ | $< 8.6 \cdot 10^{-5}$ | $< 2.7 \cdot 10^{-6}$ |
| 1,000 | $< 1.1 \cdot 10^{-6}$ | $< 8.5 \cdot 10^{-7}$ | $< 8.2 \cdot 10^{-9}$ |
| 10,000 | $< 1.1 \cdot 10^{-8}$ | $< 8.5 \cdot 10^{-9}$ | $< 2.6 \cdot 10^{-11}$ |
| 100,000 | $< 1.1 \cdot 10^{-10}$ | $< 8.5 \cdot 10^{-11}$ | $< 8.2 \cdot 10^{-14}$ |

treatment of higher dimensional systems with multi-dimensional stable manifolds is largely analogous to the example described above. This is the reason why the algorithm performs similarly well for more complicated models.

For models with an unknown solution an estimate for the error is desirable. Solvers of initial value problems usually control local errors by implying a higher order difference scheme. This is lost for the relaxation algorithm, however, we can control the global error instead. The idea is to exploit information about the order of the difference scheme.[31] If we obtain a second solution, but with a finer mesh grid, the global error reduces according to equation (35). More precisely, we run the Relaxation algorithm with two different mesh grids, whereas the second mesh grid possesses additional meshpoints at the interspaces of the first mesh grid. We denote the solution obtained by the thin mesh by $\{\hat{z}_i\}_{i=1}^{M}$, the solution obtained by the dense mesh by $\{\tilde{z}_n\}_{i=1}^{2M-1}$, and the correct, usually unknown solution by $z(\tau)$. The global

---

[30]It should be mentioned that the allocation of the mesh was chosen exogenously. The accuracy of the algorithm could be improved with a self allocating time mesh as proposed by Press et al. (1989, Chapter 16.5). They suggest to automate the allocation of mesh points so that more mesh points are placed to regions in which the variables are changing rapidly.

[31]This is also known under the term extrapolation, see Ascher and Petzold (1998, pp. 207).

error of the solution obtained by the thinner mesh behaves according to

$$\epsilon_i = z(\tau_i) - \hat{z}_i = ch^2 + o(h^4) \qquad i = 1, \ldots, M \tag{42}$$

where $h$ is the step size, i.e. $h = \frac{1}{M}$, and $c$ is an unknown constant that does not depend on $h$, but may vary slowly in $t$.[32] Comparing the solution $\tilde{z}$ with the correct solution at the same mesh points yields a global error that is one quarter smaller

$$\epsilon_i = z(\tau_i) - \tilde{z}_i = \frac{1}{4}ch^2 + o(h^4) \qquad i = 1, \ldots, 2M - 1 \tag{43}$$

whereas $h$ still indicates the bigger step size $\frac{1}{M}$. Then, the sum $\frac{4\tilde{z}-\hat{z}}{3}$ is of order four, since

$$
\begin{aligned}
\epsilon_i &= z(\tau_i) - \frac{4\tilde{z}_{2i-1} - \hat{z}_i}{3} \\
&= z(\tau_i) - \frac{4 \cdot z(\tau_i) - ch^2 - 4o(h^4) - z(\tau_i) + ch^2 + o(h^4)}{3} \\
&= 0 + o(h^4) \qquad i = 1, \ldots, M
\end{aligned}
$$

This solution of higher precision can be used to estimate the global error of $\hat{z}$ by

$$\epsilon_i = z(\tau_i) - \hat{z}_i \approx \frac{4}{3}(\tilde{z}_{2i-1} - \hat{z}_i) \qquad i = 1, \ldots, M \tag{44}$$

and the global error of $\tilde{z}$ by

$$\epsilon_{2i-1} = z(\tau_i) - \tilde{z}_{2i-1} \approx \frac{1}{3}(\tilde{z}_{2i-1} - \hat{z}_i) \qquad i = 1, \ldots, M \tag{45}$$

at every second mesh point. Equations (44) and (45) justify a very intuitive estimation of the global error. If we want to know if a simulation result yields a sufficiently accurate representation of the solution, we have to run a second simulation with a different mesh size $M$. If both solutions do not differ significantly, the global error of both solutions is small.

---

[32]Note that third order terms $o(h^3)$ are canceled by the midpoint rule (see Ascher and Petzold, 1998, p. 208). However, this is not essential for the estimation to work.

We employ the error estimation procedure to compare the estimated errors with the exact errors for the Ramsey-Cass-Koopmans model. Figure 8 shows the exact error (crosses) and the estimated error (circles) estimated by equation (44) and transformed into relative errors for capital (i) and consumption (ii). Simulations were run with a mesh $M = 30$. The maximum estimated relative error is smaller than $1.0 \cdot 10^{-3}$ for capital and $1.4 \cdot 10^{-3}$ for consumption, respectively. For $c$ and $k$ the estimated error is almost indistinguishable from the analytically calculated error. Therefore, we employ equation (44) to estimate the maximum error for the simulations of the Jones (1995a) and Lucas (1988) model.
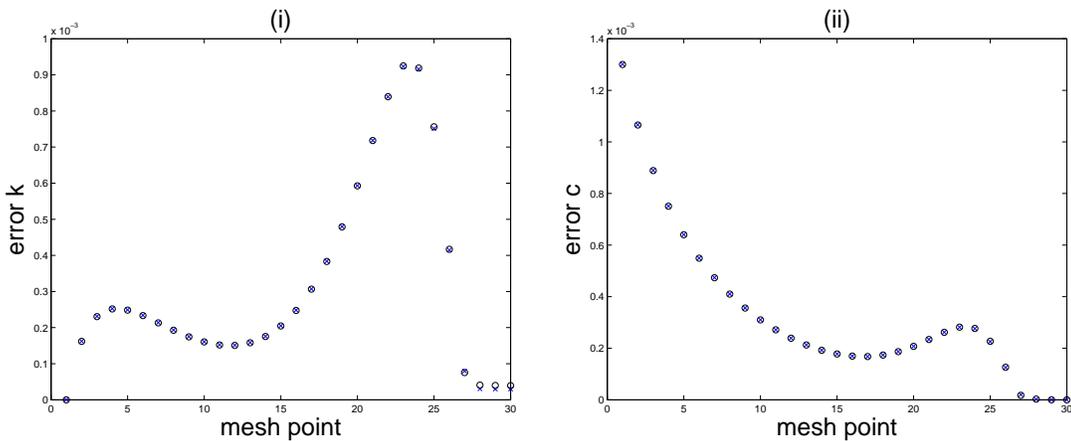


Figure 8: Exact and estimated error for the RCK model

## 3.3  Comparison to other procedures

The relaxation procedure and similar finite-difference procedures have already been employed in various fields of economics. Prominent examples comprise the solution of two point boundary value difference equations (e.g. Laffargue, 1990; Juillard et al., 1998), differential-difference equations (e.g. Boucekkine et al., 1997) as well as partial differential equations (e.g. Candler, 1999). However, to the best of our

knowledge, the relaxation algorithm has not been exploited yet systematically to solve deterministic continuous-time two point boundary value problems in growth theory. Nonetheless, there are a few applications in the economics literature. For instance, Oulton (1993) and Robertson (1999) employ the relaxation routine provided by Press et al. (1989) to solve a continuous-time deterministic growth model.

In this Section we provide a systematic comparison between the Relaxation algorithm and competing methods. As a criterion, we choose how each method performs for solving higher-dimensional models numerically. However, we do not exemplify the comparison by solving one model with all methods described. The reason for this is that we do not consider this to be informative. If we would solve a simple model (e.g. the Ramsey model) with the competing methods, computational costs would be negligible, since today's computers are fast enough to solve this kind of models in less than a second. Moreover, computational time of the simple model can not be used as an estimator for computational time of more sophisticated models. The reason is that computational costs grow with different rates for each method as the model dimension increases. Therefore, we consider a comparison of the growth rates as more meaningful. If we would try to solve a more sophisticated model (e.g. the Jones model) with competing methods, we would probably not be able to solve the model with every competing method. We could be blamed not to have tried hard enough, or to have programmed some of the methods in an inefficient way. Therefore, for an algorithm competition usually experts for each method are asked to compete. Then, incentives would ensure a fair comparison, as demonstrated in the seminal paper of Taylor and Uhlig (1990). This kind of comparison is beyond the scope of this thesis.

We compare the relaxation procedure to the most popular alternative solution methods employed in deterministic growth theory. These comprise backward inte-

gration (Brunner and Strulik, 2002), the finite-difference method as proposed by Candler (1999), time elimination (Mulligan and Sala-i-Martin, 1991), projection methods (e.g. Judd, 1992; Judd, 1998, Chapter 11), and the method of Mercenier and Michel (1994 and 2001). Most of the procedures and their comparative advantages are described in Judd (1998) and Brunner and Strulik (2002).

The numerical methods are designed to solve a continuous-time, perfect-foresight, infinite horizon, optimal control problem

$$\max_u \int_{t_0}^{\infty} f(x,u)e^{-\rho t}dt \tag{46}$$
$$\text{s.t.} \quad \dot{x} = g(x,u), \quad x(0) = x_0,$$

whereas $x$ denotes a $n_x$ dimensional state variable and $u$ a $n_u$ dimensional control variable. The functions $f$ and $g$ are sufficiently smooth. We assume that $x$ and $u$ are continuously differentiable functions of time. For an illustrative comparison of the procedures we will, for some procedures, show how the Ramsey-Cass-Koopmans model (Ramsey, 1928; Cass, 1965; Koopmans, 1965) can be solved numerically. According to this model, an infinitely-lived household solves

$$\max_c \int_0^{\infty} e^{-\rho t}u(c)dt \tag{47}$$

subject to

$$\dot{k} = f(k) - c \quad k(0) \text{ given,} \tag{48}$$

where $u(c)$ is the utility function, $f(k)$ denotes a neoclassical production function in per capita terms and $\rho$ is the household's discount rate (see Barro and Sala-i-Martin, 2004, Chapter 2).

The basic problem (46) can be transformed into different numerical problems. The first possibility is to discretize the integral and the differential equation of problem (46), and solve the discrete optimization problem by standard optimization routines.

This approach is chosen by Mercenier and Michel (1994 and 2001). The second possibility is to apply Pontryagins (1962) maximum principle and to solve the first order conditions represented by the differential equations, algebraic equations, and boundary conditions numerically. Mulligan and Sala-i- Martin (1991), Judd (1992 and 1998), and Brunner and Stulik (2002) choose different approaches to solve the system of first order conditions. The third approach is to apply Bellman's (1957) principle of dynamic programming. Candler (1999) follows this approach and solves for the value function numerically.

### 3.3.1  The method of Candler

Finite-difference methods as described by Candler (1999) employ an algorithm similar to the relaxation procedure to solve partial differential equations. Candler exemplifies the method by solving a stochastic version of the Ramsey model. An infinitely-lived household solves

$$\max_{c} E \int_{0}^{\infty} e^{-\rho t} u(c) dt \tag{49}$$

subject to

$$dk = (f(k) - c)dt + \sigma(k)dz \tag{50}$$

where the denotation of parameters is the same as in problem (47) and (48). Additional to the illustration of the model above is the stochastic term $\sigma(k)dz$. It denotes the effect of the stochastic Wiener process $dz$ on investment.

The stochastic problem can be solved by employing Bellman's (1957) principle. We define the value function $V(t, k)$ as the expected value of future utility at date $t$, given that the representative household possesses the capital stock $k$ at $t$ and given that the household chooses an optimal strategy:

$$V(t, k) = \max_{c} E \int_{t}^{\infty} e^{-\rho(s-t)} u(c) ds \,. \tag{51}$$

Bellmans principle states that an optimal solution for the time horizon $[t, \infty)$ is also optimal for each subinterval $[t, t+\Delta t)$ and $[t+\Delta t, \infty)$. Employing this rule, one can derive the general dynamic programming equation for the value function[33]

$$-V_t(t,k) + \rho V(t,k) = \max_c \left\{ u(c) + (f(k) - c)V_k(t,k) + \frac{1}{2}\sigma(k)^2 V_{kk}(t,k) \right\} \quad (52)$$

where $V_i$ denotes the partial derivative of $V$ with respect to $i$. We assume the functional forms to equal $u(c) = \frac{c^{1-\theta}}{\theta}$, $\theta > 0$, $f(k) = k^\alpha - \delta k$, and $\sigma(k) = \sigma \cdot k$. Using the first order condition $u'(c) = V_k$ and changing the sign of $V_t$ we obtain[34]

$$V_t(t,k) + \rho V(t,k) = \frac{\theta}{1-\theta}V_k(t,k)^{\frac{1-\theta}{\theta}} + V_k(t,k)(k^\alpha - \delta k) + 1/2\sigma^2 k^2 V_{kk}(t,k). \quad (53)$$

This is a partial differential equation, because the partial derivative of the unknown value function $V$ with respect to $t$ and $k$, and the second partial derivative with respect to $k$ is included. Since $V$ is time-independent, it could easily be transformed into an ordinary differential equation by substituting $V_t \equiv 0$. This procedure, however, would not be generalizable to cases with more than one state variable. If another state variable would be present in the model, equation (53) would contain partial derivatives with respect to that variable. In this case, assuming $V_t \equiv 0$ would not transform equation (53) into an ordinary differential equation. To keep the procedure general, Candler suggests to solve the partial differential equation (53) directly.

Numerical solutions of partial differential equations is a wide and complex field. We only want to give a general description of the procedure proposed by Candler. Numerical problems of partial differential equations can be divided into boundary

---

[33]A detailed derivation of the standard equation can be found in Kamien and Schwartz (1981), pp. 238.

[34]The change of the sign of $V_t$ is equivalent to a reversal of time. Candler notes that this is to turn final boundary conditions for $t = t_{final}$ into initial conditions, if the problem were time dependent. In this example (and related examples), the transformation is of no relevance, since $V_t \equiv 0$ for the correct solution.

value problems and initial value problems (see Press et al., 1989, pp. 827). In both cases the solution is a function, defined on a mesh of independent variables. The solution of boundary value problems is a 'static' function of the independent variables, satisfying conditions on the boundaries of the region of interest. Therefore, it is usually solved simultaneously for the whole mesh. By contrast, the solution of initial value problems can be seen as a function that describes how the solution propagates as one independent variable proceeds. In this case, the independent variable often can be interpreted as time. The problem is usually solved for the initial point of time and then integrated in time.

Equation (53) states how the value function is connected along $k$ and $t$. More precisely, the equation states how $V$ changes if the initial capital stock $k$ changes or if time $t$ changes. For the numerical solution of equation (53) it is essential that $V_t \equiv 0$, which is an information that is not included in the equation. Therefore, the equation can be seen as an initial value problem. In a first step, an initial guess for $V$ at $t = 0$ is made. This guess, of course, satisfies the partial differential equation only with $V_t \neq 0$, because, loosely speaking, the residuals are caught by $V_t$. Then, the equation is integrated in time until $V_t \equiv 0$.

For example, given the mesh of $(t, k)$, the partial derivative of $V_k$ can be approximated by the formula

$$V_k(t, k) = \frac{V(t, k + \Delta k) - V(t, k)}{\Delta k} + o(\Delta k)$$

which indicates that the approximation is of order one, $o(\Delta k)$. It is constructed by applying Taylor's rule and truncating terms of lower order. Analogously, the derivatives $V_t$ and $V_{kk}$ can be approximated by

$$V_t(t, k) = \frac{V(t + \Delta t, k) - V(t, k)}{\Delta t} + o(\Delta t)$$

and

$$V_{kk}(t,k) = \frac{V(t, k + \Delta k) - 2V(t,k) + V(t, k - \Delta k)}{(\Delta k)^2} + o((\Delta k)^2)$$

Inserting the approximations in equation (53) yields

$$\frac{V(t + \Delta t, k) - V(t,k)}{\Delta t} + \rho V(t,k) = \frac{\theta}{1 - \theta} \left( \frac{V(t, k + \Delta k) - V(t,k)}{\Delta k} \right)^{\frac{1-\theta}{\theta}}$$
$$+ \frac{V(t, k + \Delta k) - V(t,k)}{\Delta k}(k^\alpha - \delta k)$$
$$+ 1/2\sigma^2 k^2 \frac{V(t, k + \Delta k) - 2V(t,k) + V(t, k - \Delta k)}{(\Delta k)^2} + o(\Delta t) + o(\Delta k) \qquad (54)$$

Now, equation (54) can be solved explicitly for $V(t + \Delta t, k)$ given a solution $V(t, k)$ at $t$.

Since equation (54) is a relation for the interspace between mesh points, proper conditions must be provided for the boundary of the mesh. Candler suggests to use boundary conditions for $k = 0$ and $k = k_{max}$. The first boundary condition would cause problems, because $V_k(t, 0) = \infty$, however, due to a more special difference scheme Candler proposes, it is not necessary to evaluate $V_k$ for $k = 0$. For the second boundary condition Candler suggests to extend $k_{max}$ until $V_k(t, k_{max}) = 0$ is a good approximation.

The algorithm for finding the solution can be sketched as follows:

- Define a mesh of points on $(t, k)$

- Choose an initial guess for $V$. (Candler chooses $V(k) = \sqrt{k}$)

- Use (54) to compute $V(t + \Delta t, k) - V(t, k)$, for all $k_i$

- Compute an error norm $\epsilon := ||V(t + \Delta t, k) - V(t, k)||$

- Update the solution $V(t + \Delta t, k)$ until $\epsilon$ is sufficiently small, i.e. $V$ is no longer changing with respect to time

Candler proposes several measures to enhance the stability and the speed of the method. We want to focus the critique on the generalization of the method to more than one state variable and the associated increase in computational requirement. Consider a mesh of $(m \times n)$, i.e. $t \, \epsilon \, (t_1, \ldots, t_m)$ and $k \, \epsilon \, (k_1, \ldots, k_n)$. For each time step, equation (54) has to be evaluated for every pair $(k_i, k_i + 1)$, $i < n$. Given that the algorithm converges after $m$ time steps, a system of $n$ equations has to be solved $m$ times.[35]

Now consider a problem with two state variables. The resulting mesh will be of dimension $m \times n \times n$. For each time step, $2 \cdot n \cdot n$ difference equations have to be evaluated, since every interior mesh point has now four neighboring points instead of two. While there exists methods to leave out some of the difference equations, the computational demand will be at least $n$ times higher compared to a problem with only one state variable.[36] This exponential increase in computational expense is known as the 'curse of dimensionality'. It makes the numerical solution of higher dimensional models difficult if not impossible.

### 3.3.2  Backward integration

Backward integration as suggested by Brunner and Strulik (2002) is a method to solve the first order conditions of Pontryagin's maximum principle represented by a system of ordinary differential equations numerically.

Consider a continuous-time, perfect-foresight, infinite-horizon, optimal control problem as described by (46). Necessary conditions for an optimal solution are employing Pontryagins Maximum principle (see Pontryagin et al., 1962) with the

---

[35]Equation (54) refers only to the interspaces. Together with the boundary conditions, a square system of equations results.

[36] Press et al. (1989, p. 700) also report that computational requirements increase by a factor of at least 100 if one switches from a one-dimensional problem with 100 grid points to a two-dimensional problem with $100 \times 100$ grid points.

current-value Hamiltonian $H = f + \lambda^T g$

$$\frac{\partial f}{\partial u} + \lambda^T \frac{\partial g}{\partial u} = 0 \tag{55}$$

$$\dot{x} = g(x, u) \tag{56}$$

$$\frac{\partial f}{\partial x} + \lambda^T \frac{\partial g}{\partial x} = \rho\lambda - \dot{\lambda} \tag{57}$$

together with the initial conditions $x(0) = x_0$ and transversality condition

$$\lim_{t \to \infty} \lambda(t)x(t) = 0 \, . \tag{58}$$

We assume that the costate variable $\lambda$ can be eliminated from the system to obtain a square system of differential equations

$$\dot{u} = F_1(u, x)$$

$$\dot{x} = F_2(u, x) \tag{59}$$

that is saddlepoint stable in the neighborhood of a hyperbolic, interior steady state $(u^*, x^*)$. The system (59) exhibits $n_x$ state variables or (more generally) $n_x$ initial conditions. We assume the transversality conditions (58) to ensure convergence to the interior steady state. For the transition path towards the steady state to be unique, the system has to exhibit a $n_x$-dimensional stable manifold.

The initial values $(x(0), u(0))$ of the solution are only known for the state variable $x$. The initial value of the control variable is determined but unknown and often impossible to compute for non-linear problems. If the system (59) would be integrated with respect to time with the correct initial state but an incorrect initial control variable, the solution would diverge from the stable manifold exponentially. Therefore, small deviations in the initial value of $u$ cause large differences in $(x, u)$ over time, and the correct steady state $(x^*, u^*)$ would be missed by far. The problem of forward integration, also named as multiple-shooting, is recognized as highly

ill-conditioned.[37] Backward integration exploits this instability by reverting time. If the flow of the system is reversed, trajectories converge instead of diverge towards the solution manifold exponentially. Since the steady state is known for state and control variables, the solution manifold can be computed easily by starting near the steady state an integrating backward in time.

For reversing the flow we make use of the fact that system (59) is autonomous. Defining $\tilde{t} := -t$ and applying the chain rule yields

$$
\begin{aligned}
\frac{du}{d\tilde{t}} &= \frac{du}{dt}\frac{dt}{d\tilde{t}} = -F_1(u, x) \\
\frac{dx}{d\tilde{t}} &= \frac{dx}{dt}\frac{dt}{d\tilde{t}} = -F_2(u, x)\,.
\end{aligned}
\tag{60}
$$

Since system (60) exhibits the same shape as system (59), only multiplied by $-1$, time transformation exactly reverses the flow. Therefore, the stable manifold becomes an unstable manifold, and the unstable manifold becomes a stable manifold of the fixed point $(x^*, u^*)$. We are interested in computing the stable manifold of (59) to which we will refer to as the solution manifold. The initial boundary condition $x(0) = x_0$ of problem (46) will be relabeled as $x(\hat{t}) = x_0$. If the system (60) is integrated with respect to time $\tilde{t}$ with an initial value close to the steady state, the numerical solution converges towards the solution manifold exponentially. Therefore, errors caused by the fact that the initial point is not exactly on but only close to the solution manifold, decay exponentially. To minimize the errors, the deviation from the steady state can be determined in direction of the corresponding eigenvector of the solution manifold, which is tangent to the manifold. The outline of the backward integration algorithm for systems exhibiting a one-dimensional solution manifold is as follows

- Choose an initial value $(x(0), u(0))$ close to the steady state $(x^*, u^*)$ by deviat-

---

[37]Multiple Shooting was introduced into the economic literature by Lipton et al. (1982). However, Ascher et al. (1988) consider it to be ill-conditioned for saddlepoint problems.

ing along the corresponding eigenvector $v$:

$$(x(0), u(0)) = (x^*, u^*) + \epsilon v$$

- Integrate system (60) numerically, until the solution trajectory hits the boundary condition $x(\hat{t}) = x_0$.

- Re-transform time and the corresponding solution vector.

Note, that the numerical integration of system (60) can be conducted with any standard routine implemented in mathematical program packages.

To exemplify the algorithm on a simple, one-dimensional example, we describe how to solve the Ramsey-model with the backward integration procedure. The first order conditions of problem (47) and (48), employing the production function $f(k) = k^\alpha$ and the utility function $u(c) = \frac{c^{1-\theta}}{1-\theta}$, are

$$
\begin{aligned}
\dot{c} &= c\frac{\alpha k^{\alpha-1} - \rho}{\theta} \\
\dot{k} &= k^\alpha - c\,.
\end{aligned}
$$

The steady state can be calculated analytically as $(k^*, c^*) = \left( \left(\frac{\alpha}{\rho}\right)^{\frac{1}{1-\alpha}}, \left(\frac{\alpha}{\rho}\right)^{\frac{\alpha}{1-\alpha}} \right)$, and the time-transformed system reads

$$
\begin{aligned}
\dot{c} &= -c\frac{\alpha k^{\alpha-1} - \rho}{\theta} \\
\dot{k} &= -(k^\alpha - c)\,.
\end{aligned}
$$

For choosing the starting values of backward integration we can avail the fact that the shape of the phase diagram is known. The one-dimensional stable manifold passes the steady state from south-west to north-east. Along the adjustment path, capital and consumption increase or decrease monotonically, respectively. We fix the initial value of capital $k_0$ to 10% of its steady state value and, therefore, want to calculate

the south-west part of the stable manifold. In this case, we can deviate from the steady state by reducing $(c^*, k^*)$ by a small $\epsilon$ in both components. The backward looking trajectory can be computed easily by employing a fourth order Runge-Kutta ODE solver. We integrate with respect to time until $k(\tilde{t}) = k_0$. The ODE solver delivers a list $(\tilde{t}, c, k)$ that has to be rearranged to re-transform time.

The method of backward integration may run into two different problems. The first one refers to an increase in the dimension of the state space. Then, the backward looking trajectory has to fulfill more than one boundary condition. In this case, Brunner and Strulik (2002) suggest to generate starting values on an orbit around the steady state. To pass through a pre-specified point (determined by the specific shock under study), it is necessary to iterate until the trajectory hits this point. However, if the real parts of the stable eigenvalues differ substantially, the problem of stiff differential equations occurs as described in Section 2.3. For small deviations in the starting values the resulting trajectories depart strongly from each other. For large differences between the stable eigenvalues, it is impossible to meet the pre-specified point, because the backward directed trajectories will be attracted by the submanifold, which is associated with the eigenvalue with the largest, in absolute terms, real part. The resulting trajectories, hence, cannot represent a specified shock and potentially have no economic meaning.

The second problem of the method refers to economic models, which do not exhibit an isolated steady state but a continuum of steady states represented by a hyperbolic center manifold of stationary equilibria as described in Section 2.1. In this case, the specific steady state to which the economy converges depends on the initial boundary conditions.[38] The backward integration procedure as described above, however, requires the knowledge of the final steady state as an initial condition for backward

---

[38]For instance, in the Lucas (1988) model presented below the actual steady state to which the economy converges depends on the initial level of human and physical capital $h_0$ and $k_0$.

integration. If one particular steady state is chosen for backward integration, only one initial condition can be satisfied. To find a trajectory, which fulfills all initial conditions, an iteration process has to be applied. This procedure typically could give rise to problems of convergence.

### 3.3.3  The method of Mercenier and Michel

Mercenier and Michel (1994 and 2001) propose to transform the continuous time, infinite horizon problem (46) into a finite horizon maximization problem in discrete time. This transformation requires the choice of an exogenous final time $t_N$ and an exogenous mesh grid $T = \{t_0, \ldots, t_N\}$. We denote the time steps as $\Delta_n := t_{n+1} - t_n$, $n = 0, \ldots, N - 1$. Then, the transformed problem is

$$\max \sum_{n=0}^{N-1} \alpha_n \Delta_n f(x(t_n), u(t_n)) + \beta_N F(x(t_N)) \tag{61}$$

$$\text{s.t.} \quad x(t_{n+1}) - x(t_n) = \Delta_n g(x(t_n), u(t_n)), \quad n = 0, \ldots, N - 1,$$

$$x(t_0) = x_0$$

where $\alpha_n$ denotes discount factors that remain to be determined and may depend on the chosen mesh $T$. The function $F$ and the associated discount factor $\beta_N$ are responsible for a proper treatment of the post terminal phase $[t_N, \infty)$, and remain to be chosen, as well. The advantage of solving problem (61) instead of the original problem (46) is that it can be solved straightforward with a standard optimization routine.[39] The discretetized problem can be seen as a problem of maximizing a function subject to static constraints.

Mercenier and Michel (1994 and 2001) show that the discount factors $\alpha_n$, $\beta_N$ and the function $F$ cannot be chosen arbitrarily to make the solution of optimization

---

[39]Mercenier and Michel (1994) annotate, that problem (61) could also originate from the transformation of another discrete time model. For example, it might be favorable to transform a discrete time model with respect to the time mesh to economize on computational requirements.

problem (61) a good approximation of the solution of optimization problem (46). They introduce the property of steady state invariance and steady state growth invariance, respectively. The idea is that the transformed problem (61) should exhibit the same stationary solution as the original optimization problem (46), if the model exhibits exogenous growth. Analogously, if the model exhibits endogenous growth, the transformed problem (61) should exhibit the same long-run growth rate as problem (46). Mercenier and Michel state necessary and sufficient conditions for both properties, which significantly enhances numerical accuracy given a certain number of mesh points.

Considering the case of steady state invariance it can be proven that the discount factors $\alpha_n$ have to fulfill the recurrence relation

$$\alpha_{n+1} = \frac{\alpha_n}{1 + \rho \Delta_{n+1}}, \quad n > 0 \tag{62}$$

with $\alpha_0$ given. Moreover, the function $F$ should be chosen according to

$$F(x) = \frac{1}{\rho} f(x^*, u^*) \tag{63}$$

with the stationary solution $(x^*, u^*)^t$ of problem (46).[40] In this case (46) and (61) possess the same stationary solution.

Mercenier and Michel (2001) generalize this result for a class of endogenous growth models by extending equation (62). Then the balanced growth rate of problem (46) is the same as for the transformed problem. Again, the property of steady growth invariance yields a considerable improvement in the numerical accuracy of the discrete time model. Moreover, Alemdar et al. (2006) show that the overall optimization performance can be improved substantially if an optimal allocation of the time mesh is chosen for the transition.

---

[40]Mercenier and Michel (1994) derive a more general condition for $F$ that is necessary for steady state invariance, however, they present this functional form as very intuitive.

The approach of Mercenier and Michel is very user-friendly, since the discretized model can be solved by any standard routine implemented in mathematical programming packages. Therefore, it is to some extent difficult to compare its performance to other algorithms, since the method relies partly on the performance of an exchangeable subroutine. As an advantage of the method, it has to be mentioned that this method is suitable for higher dimensional models. The discrete optimization model (61) could exhibit any number of state or control variables, since an increase in the number of variables does not cause any conceptual differences. Moreover, usually computational costs grow of order three for optimization routines if the number of variables increase. Therefore, computational costs grow only moderately with the model's dimension.

However, the transformation of the optimization problem has also its drawbacks, besides the fact that the researcher has to conduct laborious transformations by converting the maximization problem. First, the choice of a terminal time is to some extent arbitrary. In the worst case, the user has to find a terminal time by trial end error, since the time span in which the solution has almost approached its steady state is usually not known in advance. Also, the treatment of a post terminal stationary phase causes additional effort. Second, the method of Mercenier and Michel does not leave room for selecting different discretization rules, also of higher order. The discretization rule of the method is a first order rule, which means that the error only decreases linearly with an increase of the number of mesh points.[41] Therefore, it might be difficult to decrease the maximum error below some predetermined value when the model is high-dimensional. Moreover, computational costs grow of order three if the number of mesh-points is increased, whereas the respective increase in computational costs for the relaxation algorithm grow only quadratically.

---

[41]The discretization of the differential equations is a forward Euler discretization scheme.

### 3.3.4 Projection methods

Projection methods, introduced in Judd (1992) and Judd (1998, Chapter 11), cover a wide range of algorithms.[42]   Therefore, they can be applied to a large number of numerical problems. We first state the general approach how to apply projection methods and how it is implemented numerically. In a second step, we describe how to solve the standard neoclassical model as an example. The method can solve problems, for which a function $f$ has to be computed. Since this function is potentially multi-dimensional, a wide range of problems is covered. For example, $f$ could represent the Bellman-equation or policy function of an intertemporal optimization problem.

Consider an operator $\mathcal{N}$ that maps the function space $B_1$ into another function space $B_2$:

$$\mathcal{N} : B_1 \rightarrow B_2$$

Then, the problem is to find a function $f \, \epsilon \, B_1$ that is a zero of the operator:

$$\mathcal{N}(f) = 0$$

For numerical computations a computable approximation $\hat{\mathcal{N}}$ of $\mathcal{N}$ has to be chosen.[43] Moreover, bases $\Phi_j = \{\phi_i^j\}_{i=1}^\infty$ and inner products, $< \cdot, \cdot >_j$, for $B_j$, $j = 1, 2$ have to be chosen. In numerical terms the problem is now to find $a_i \, \epsilon \, \mathbb{R}$, $1 \leq i \leq n$, such that the function

$$\hat{f} := \sum_{i=1}^{n} a_i \phi_i^1$$

with $n$ large enough satisfies

$$\hat{\mathcal{N}}(\hat{f}) \doteq 0$$

---

[42]A rich literature describes several variants of projection methods, see e.g. Taylor and Uhlig (1990) and the papers cited therein, McGratton (1996) or McGrattan (1999). An example for a recent application is Judd (2002).

[43]In the example below, $\hat{\mathcal{N}} = \mathcal{N}$. E.g., the difference between $\hat{\mathcal{N}}$ and $\mathcal{N}$ could be that $\hat{\mathcal{N}}$ maps $C^\infty$ functions while $\mathcal{N}$ maps $C^1$ functions.

Coefficients $a_i$ are found through iteration. An initial guess of $a = (a_i)_{i=1}^n$ has to be made and the residual of $\hat{\mathcal{N}}(\hat{f})$, measured by the inner product, is improved in every iteration.

The generality of this description clarifies that projection methods cover a wide range of algorithms, since for each step different approaches can be used to implement the algorithm. Therefore, we want to address common implementations.

The bases $\Phi_j$ should be chosen such that they are simple to compute and each element 'differs a lot' from the other elements. The latter requirement means that given the basis elements $i = 1, \ldots n$ the $(n+1)$st element adds a lot of new information. Therefore, often an orthogonal basis with respect to the inner product is chosen. One example of a basis together with an inner product that are used frequently are Chebyshev Polynomials. The are defined over $[-1, 1]$ by

$$T_n(x) = \cos(n \ \arccos(x)) \qquad n \geq 0$$

and generated by the three-term recursion

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_n(x) &= 2xT_{n-1}(x) - T_{n-2}(x), \qquad \forall n \geq 2\,. \end{aligned}$$

Chebyshev Polynomials are orthogonal with respect to the inner product

$$< f(x), g(x) > := \int_{-1}^{1} \frac{f(x)g(x)}{\sqrt{(1-x^2)}} dx\,. \tag{64}$$

Besides the orthogonality, Chebyshev Polynomials have the advantage to be easy to compute and they exhibit the same scale on $[-1, 1]$.[44] Although they are considered to be very efficient, other bases could be used such as basis which have only small

[44]Of course, they can be transformed to any interval of interest.

support.   E.g. sometimes tent function are used, which take the shape of a tent on some subinterval and are zero elsewhere.  If the basis elements are zero almost everywhere, the method is classified as a finite-element technique, otherwise it is called spectral method.

There are also several possibilities to measure how much the residual

$$R(x; a) := \hat{\mathcal{N}}(\hat{f})(x)$$

deviates from zero and how the initial guess of $a$ should be improved.  One possibility is to minimize the sum of squared residuals

$$\min_a < R(x; a), R(x; a) >_2 .$$

To achieve this, a standard minimization algorithm can be employed.  Another possibility is to choose $n$ projections $p_1, \dots, p_n \in B_2$ and solve

$$P_i(a) \equiv < R(x; a), p_i(x) >= 0, \qquad i = 1, \dots, n.$$

The latter approach has the advantage that projections $P_i$ could be defined such that the minimization of residuals emphasizes certain aspects.  For example, the average error in an Euler equation could be minimized. The sum of squared residuals approach would not put special emphasis on that error and could, therefore, yield poor approximations of the Euler equation. Depending on the chosen projections, the method is classified as a Galerkin method, method of moments, subdomain method, or collocation method. Galerkin methods use the first $n$ elements of the basis $\Phi$ for projection.[45] For the method of moments the first $n$ monomials $x^{i-1}$, $i = 1, \dots, n$ are chosen for projection. On the other hand, subdomain methods minimize the residual on a collection of subdomains that cover the whole region of interest. Therefore, the projections are chosen as $I_{D_i}$, $i = 1, \dots, n$ with the indicator functions $I_{D_i}$ of the

---

[45]This requires the same basis and inner products for $B_1$ and $B_2$.

subinterval $D_i$. At last, collocation methods choose $a$ to solve $R(x_i, a) = 0$ at $n$ different points. Formally, this corresponds to choosing $n$ Dirac delta functions for projections, i.e. the projection $\delta(x - x_1), \ldots, \delta(x - x_n)$, $i = 1, \ldots, n$.

It is necessary to mention the advantages of collocation methods, for they are widely used and known to be very efficient.[46] Given the number of projections $n$, collocation methods have less computational requirement, since the residuum $R(x; a)$ only has to be evaluated at $n$ points instead of integrated $n$ times. One would suspect that collocation methods yield poor approximations of $f$, because the residuum is only controlled at certain points and not in between those points. However, if the points $x_i$ are chosen such that they exploit the characteristics of the basis, it is guaranteed that convergence occurs very quickly. For example, if Chebyshev polynomials are used, the projection points should be placed at the zeros of the Chebyshev polynomials, defined as $x_i = \cos\left(\frac{2i-1}{2n}\pi\right)$, $i = 1, \ldots, n$ on the interval $[-1, 1]$. The Chebyshev interpolation theorem states that interpolating at these points is optimal and that $\hat{f}$ converges to $f$ quickly as $n \to \infty$.

For implementing the continuous-time version of the Ramsey model (problem (47) and (48)) we solve for the policy function $c(k)$.[47] To derive a useable operator that maps the policy function we employ Pontryagin's maximum principle and derive the first order conditions with utility function $u(c) = \frac{c^{1-\theta}}{1-\theta}$ and production function $f(k) = k^\alpha$. After eliminating the adjoint variable, we derive two differential equations, the Keynes-Ramsey rule and the capital accumulation equation

$$\dot{c} = c\frac{\alpha k^{\alpha-1} - \rho}{\theta}$$
$$\dot{k} = k^\alpha - c.$$

---

[46]For example, the standard routine implemented in MatLab to solve two-point boundary value problems is a collocation method.

[47]Judd (1992) solves for the policy function in a stochastic, discrete-time version of the Ramsey model. Since we focus on the numerical solution of deterministic, continuous-time models, we deviate from the original illustration.

## 3 Multi-Dimensional Transitional Dynamics

We can derive the policy function $c'(k)$ by dividing both equations

$$c'(k) = \frac{dc/dt}{dk/dt} = \frac{c\frac{\alpha k^{\alpha-1}-\rho}{\theta}}{k^\alpha - c} \qquad \dot{k} \neq 0\,. \tag{65}$$

By rearranging we get the following differential equation for $c'(k)$

$$c'(k)(k^\alpha - c) - \left(c\frac{k^{\alpha-1}-\rho}{\theta}\right) = 0\,. \tag{66}$$

The task is now to find a policy function $c(k)$ that is the zero of operator $\hat{\mathcal{N}}$ represented by equation (66). As an example, we further specify the bases and the projections that could be used for solving the problem. The domain of approximation will be an interval $[k_{min}, k_{max}]$ that includes the steady state. For a suitable basis we choose Chebyshev polynomials transformed to the interval $[k_{min}, k_{max}]$ by the linear transformation $k \rightarrow 2\frac{k-k_{min}}{k_{max}-k_{min}} - 1$. The inner product is defined by

$$< f(k), g(k) > \quad := \quad \int_{-1}^1 \frac{f(k)g(k)}{w(k)}dk, \tag{67}$$

$$w(k) \quad := \quad \sqrt{\left(1 - \left(2\frac{k-k_{min}}{k_{max}-k_{min}} - 1\right)^2\right)}$$

and the projections are the Dirac delta functions evaluated at

$$x_i = 2\frac{\left(\cos\left(\frac{2i-1}{2n}\pi\right)\right) - k_{min}}{k_{max} - k_{min}} - 1, \; i = 1,\dots,n \tag{68}$$

To summerize, we search $a = (a_i)_{i=1}^n$ such that the function

$$\hat{c}(k; a) = \sum_{i=1}^n a_i\phi_i(k)$$

with transformed Chebyshev polynomials $\phi_i$ satisfies equation (66) evaluated at points (68). These are $n$ non-linear equations which could be solved using a Newton algorithm or any proper standard routine.

The method may run into two different problems. The first one refers to the derivation of the policy function. Projection methods are often employed to approximate the policy function, because the dimensionality of the problem is smaller. In the simple example of the Ramsey-model, the first order conditions derived from Pontryagrin's maximum principle are two differential equations on the time interval $[0, \infty)$. The derivation of the policy function allows to reduce the system to a one dimensional differential equation with respect to $k$ on an interval $[k_{min}, k_{max}]$. Therefore, the problem was reduced with respect to the dimension and the interval of integration. However, if policy function approximation are used, the policy function may not be defined at interior points. From equation (65) it becomes apparent that the policy function is not defined for $\dot{k} = 0$. At the steady state the additional information that $\dot{c} = 0$ and $c'(k) \neq \infty$ is available. Therefore, equation (66) could still be evaluated. However, if the stable manifold is multi-dimensional, trajectories in the state space may exhibit extremal values. In this case, the slope of the policy function may be infinity as it is illustrated in Section 2.2. Then, it is not possible to evaluate equation (66).

The second detriment of projection methods also refers to the solution of higher dimensional models. To represent a higher-dimensional function a multi-dimensional basis has to be constructed. Employing Tensor methods, multi-dimensional basis functions can be constructed from a simple one-dimensional basis. For example, for constructing a two-dimensional basis from the one-dimensional basis $(\phi_i(x))_{i=1}^{\infty}$ the set of pairwise products $(\phi_i(x)\phi_j(x))_{i,j=1}^{\infty}$ represents the Tensor product basis. The number of basis elements will be $n$ basis elements for the one-dimensional case, $n^2$ basis elements for the two-dimensional basis and so forth. Therefore, the number of basis elements will increase exponentially as the dimension increases, and computational costs will also grow exponentially. To avoid this "curse of dimensionality",

a special complete polynomial basis is chosen but still the computation costs grow polynomially.[48]

### 3.3.5  Time elimination

The method of time elimination introduced by Mulligan and Sala-i-Martin (1991, 1993) is a method, similar to backward integration, to turn the two-point-boundary value problem of first order conditions into an initial value problem. Different to backward integration, the system of differential equations is manipulated such that it is independent of time. Therefore, the solution method aims at solving for the policy function. We demonstrate the method on the simple example of optimization problem (47) and (48) first, and then discuss the generalization to higher dimensional systems.

The first order conditions of problem (47) and (48), employing the production function $f(k) = k^\alpha$ and the utility function $u(c) = \frac{c^{1-\theta}}{1-\theta}$, are

$$
\begin{aligned}
\dot{c} &= c\frac{\alpha k^{\alpha-1} - \rho}{\theta} \\
\dot{k} &= k^\alpha - c.
\end{aligned}
$$

(69)

The transversality condition ensures convergence towards the interior steady state, represented by

$$
(k^*, c^*) = \left( \left( \frac{\alpha}{\rho} \right)^{\frac{1}{1-\alpha}}, \left( \frac{\alpha}{\rho} \right)^{\frac{\alpha}{1-\alpha}} \right).
$$

(70)

For applying the method of time elimination we derive the policy function $c'(k)$ as

$$
c'(k) = \frac{\dot{c}}{\dot{k}} = \frac{c\frac{\alpha k^{\alpha-1}-\rho}{\theta}}{k^\alpha - c}
$$

(71)

---

[48]That means an denoting the number of state variables by $n$ the procedure is of order $o(c^n)$ with some constant $c$.

for $\dot{k} \neq 0$. For deriving the boundary condition for this differential equation in $k$, we exploit the fact that the economy converges towards the steady state (70). Boundary condition is, therefore, $c(k^*) = c^*$. The numerical problem now turns from a two-point boundary value problem into a problem with boundary conditions for the final boundary $c(k^*) = c^*$. To transform the problem into an initial value problem that can be solved with standard numerical routines easily we transform it via $\tilde{k} := k^* - k$. The transformed equation reads

$$c'(k) = -\frac{c\frac{\alpha(k^*-k)^{\alpha-1}-\rho}{\theta}}{(k^*-k)^\alpha - c} . \tag{72}$$

For numerical calculation of the policy function we have to integrate equation (72) on the interval $[k^*, k_0]$, employing a standard ODE solver for initial value problems.[49] However, one problem remains to be solved. We have exchanged the asymptotic transversality condition for a boundary condition that can be satisfied for a final value of $k$, namely $k^*$. However, equation (72) cannot be evaluated for $k^*$, since $\dot{k^*} = 0$. $c'(k^*)$ can be calculated using L'Hopital's rule or linearization of system (69) at the steady state. In the second case, the eigenvector of the stable eigenvalue displays the slope of the policy function at the steady state. The ratio of both components of the eigenvector in the $(c, k)$ space can be interpreted as the first derivative $c'(k^*)$. We apply L'Hopital's rule and find the slope of the policy function at the steady state to equal

$$c'(k^*) = \frac{1}{2}\left(\rho + \sqrt{\rho^2 + \frac{4}{\theta}c^*\alpha(1-\alpha)(k^*)^{\alpha-2}}\right) \tag{73}$$

Finally, we should mention that the time paths of $c$ and $k$ can easily be calculated by inserting the numerical solution of $c(k)$ into system (69) and integrating with respect to time. As initial conditions $(c, k) = (c(k_0), k_0)$ must be chosen.

---

[49]For example, we could use a fourth order Runge-Kutta procedure with step-size control.

The generalization of the method to models, which exhibit more than one control variable, is straightforward. Instead of deriving a one-dimensional policy function, the policy function $p$ will be mapped into $n_c$ dimensions, if we consider the case of $n_c$ control variables, $p : \mathbb{R} \to \mathbb{R}^{n_c}$. Therefore, system (71) will be an $n_c$ dimensional system of ordinary differential equations. The crucial point is that integration still takes place with respect to one variable, the only state variable.

For systems exhibiting more than one state variable, it is in general not possible to represent the policy function through a system of ordinary differential equations analogous to system (71). If time were eliminated, the system would contain derivatives with respect to the different state variables. Therefore, we would have to solve a partial differential equation, whose numerical solution is far more complicated. Instead, Mulligan and Sala-i-Martin suggest to shoot backward, employing the original, time-dependent system of ordinary differential equations. However, this is conceptually equal to backward integration as suggested by Brunner and Strulik (2002), and, therefore, the same criticism applies.

### 3.3.6 Concluding comparison

The exceptional attribute of the relaxation algorithm is that it can solve systems exhibiting a multi-dimensional stable manifold and a center manifold of stationary equilibria conveniently. We want to summarize the merits and disadvantages of each method. Since every method can solve simple models like the Ramsey model in less than a second, we concentrate on how the methods perform for models exhibiting a multi-dimensional stable manifold. More precisely, a crucial point of comparison is how the methods' computational demand increases if the dimension of the state space increases. Though, it is not possible to prove that for a method it is impossible to solve a certain kind of model. For the comparison we will not consider that for some methods the model has to be transformed or that other forms of human capital

is needed as an input.[50]

The essential disadvantage of the method of Candler is the treatment of higher-dimensional models. The method provides a generic treatment of higher-dimensional stable manifolds, and is, therefore, able to simulate these kind of models. However, computational requirements grow exponentially as the number of state variables increases. Therefore, the method loses its merits for models exhibiting a multi-dimensional stable manifold. The method of backward integration does not allow for a generic treatment of higher-dimensional models. If the model exhibits two state variables, a backward shooting algorithm can be applied. However, the problem of stiff differential equations might occur. Hence, the numerical solution of higher-dimensional models might be difficult or even impossible. The method of Mercenier and Michel is generic with respect to the number of state variables. Moreover, computational costs grow only of order three if the model dimension increases. The only disadvantage of the method is that it uses first order difference schemes and, therefore, it might need a large amount of mesh points to yield accurate solutions. Projection methods are also generic with respect to the model dimension. However, computational requirements grow exponentially if the dimension of the model increases. This "curse of dimensionality" can be attenuated by selecting a special basis, but, still, computational costs grow polynomially. The method of time elimination cannot be generalized easily to more than one state variable. In fact, the generalization is conceptually equal to the backward integration procedure, wherefore the same criticism applies.

---

[50]In the following we will refer to models exhibiting a multi-dimensional stable manifold as higher-dimensional models.

## 3.4 Two illustrative applications

The relaxation procedure is employed to investigate the transition process of two prominent growth models. As a first example, we consider the Jones (1995a) model. For usual calibrations this model gives rise to a system of stiff differential equations. The four-dimensional transition towards the unique steady state appears to be non-monotonic. The second example, the Lucas (1988) model, implies a saddle-point stable center manifold. The different points on this curve reflect level effects of transition towards long-run growth. It should be noticed that the transition process of these popular growth models has hardly been investigated in detail so far, which is probably due to the conceptual problems mentioned above.

### 3.4.1 The Jones (1995a) model

The presentation of the Jones (1995a) model basically follows Eicher and Turnovsky (1999) who have formulated the social planner's solution of the general non-scale R&D-based growth model. For a detailed derivation of the decentralized solution see Steger (2005). As in Jones (1995a), the focus here is on the market solution. The final-output technology is given by

$$Y = \alpha_F(\phi L)^{\sigma_L} \int_0^A x(i)^{1-\sigma_L} di \qquad (74)$$

where $Y$ denotes final output, $\phi$ the share of labor allocated to final-output production, $x(i)$ the amount of differentiated capital goods of type $i$, $A$ the number of differentiated capital goods, $\alpha_F$ a constant overall productivity parameter and $\sigma_L$ the elasticity of labor in final-output production. Each intermediate good $x_i$ is produced by firm $i$, which owns a patent on producing this good exclusively. Final output producing firms maximize profits, which yields that wages equal the marginal product $w = \sigma_L \frac{Y}{\phi L}$ and the price of one intermediate is equal to $p_i = (1 - \sigma_L)(\phi L)^{\sigma_L} x_i^{-\sigma_L}$.

Noting the general symmetry among $x(i)$ and using the definition of aggregate capital $K := Ax$, the final-output technology can be written as

$$Y = \alpha_F (A\phi L)^{\sigma_L} K^{1-\sigma_L} \tag{75}$$

Intermediate firms produce one unit of intermediate from one unit of foregone consumption. Therefore, profit maximization of intermediate firms yield

$$p = p_i = \frac{r}{1 - \sigma_L}, \tag{76}$$

$$\pi = \pi_i = \sigma_L (1 - \sigma_L) \frac{Y}{A}. \tag{77}$$

Patents for new intermediate goods are produced according to the R&D technology

$$\dot{A} = J = \alpha_J A^{\eta_A} [(1 - \phi)L]^{\eta_L} \tag{78}$$

with $\eta_L := \eta_L^p + \eta_L^e$, $\eta_L^p = 1$, $-1 < \eta_L^e < 0$, $\eta_A < 1$ where $\alpha_J$ denotes a constant overall productivity parameter, $\eta_A$ the elasticity of technology in R&D and $\eta_L$ the elasticity of labor in R&D. For the market solution we distinguish between the elasticity of labor observed by private firms, $\eta_L^p$, and a negative external effect of labor in research, $\eta_L^e$, caused by the 'stepping on toe' effect introduced by Jones (1995a). The former elasticity equals one because of perfect competition in the research sector. Different to Romer (1990), we assume $\eta_A < 1$, which reflects the fact that spill-overs known as the 'standing on shoulders' effect do not fulfill the knife-edge condition to be linear. Workers are allowed to enter the R&D sector freely, wherefore they are paid according their marginal product

$$w = V_a \alpha_J A^{\eta_A} [(1 - \phi)L]^{\eta_L - 1} \tag{79}$$

with the value of one blueprint $V_a$. A no arbitrage condition forces the discounted stream of profits to equal the patent's value

$$V_a(t) = \int_t^\infty \pi(\tau) e^{-\int_t^\tau r(s)ds} d\tau \tag{80}$$

Differentiating with respect to time yields

$$\dot{V}_a = rV_a - \pi \qquad (81)$$

Finally, households maximize intertemporal utility according to

$$\max_c \int_0^\infty \frac{(C/L)^{1-\theta}}{1-\theta} e^{-\rho t} dt \qquad (82)$$

subject to

$$\dot{K} = rK + wL - V_a\dot{A} + A\pi - C \qquad (83)$$

with consumption $C$ and the relative risk aversion equal to $\frac{1}{\theta}$. First order condition of the consumer's maximization problem is

$$\dot{C} = \frac{C}{\theta}(r - \rho - n) + nC \qquad (84)$$

We transform the variables into stationary ones by expressing the system in scale adjusted variables, which are defined by $y := Y/L^{\beta_K}$, $k := K/L^{\beta_K}$, $c := C/L^{\beta_K}$, $a := A/L^{\beta_A}$, $j := J/L^{\beta_A}$ and $v_a := v/L$ with $\beta_K = \frac{1-\eta_A+\eta_L}{1-\eta_A}$, $\beta_A = \frac{\eta_L}{1-\eta_A}$. The dynamic system which governs the evolution of the economy under study, can be summarized as follows:

$$\dot{k} = y - c - \delta k - \beta_K n k \qquad (85)$$

$$\dot{a} = j - \beta_A n a \qquad (86)$$

$$\dot{c} = \frac{c}{\theta}[r - \delta - \rho - (1-\gamma)n] - \beta_K n c \qquad (87)$$

$$\dot{v}_a = v_a[r - n] - \pi \qquad (88)$$

$$\frac{\sigma_L y}{\phi} = v_a \frac{\eta_L^p j}{1 - \phi} \qquad (89)$$

where $y = \alpha_F(a\phi)^{\sigma_L} k^{1-\sigma_L}$, $j = \alpha_J a^{\eta_A}(1-\phi)^{\eta_L}$, $r = \frac{(1-\sigma_L)^2 y}{k}$, $\pi = \frac{\sigma_L(1-\sigma_L)y}{a}$. The (unique) stationary solution of this dynamic system corresponds to the (unique) BGP of the economy expressed in original variables.

Equations (85) and (86) are the equations of motion of (scale-adjusted) capital and technology, (87) is the Keynes-Ramsey rule of optimal consumption $c$, (88) shows capital market equilibrium with $v_a$ denoting the (scale-adjusted) price of blueprints and (89) determines the privately efficient allocation of labor across final-output production and R&D.[51]

To determine the set of feasible parameter values is particulary difficult for this model. To our knowledge, it is not possible to calculate the steady state, nor the eigenvalues or eigenvectors of the linearized system at the steady state analytically. Arnold (2006) investigates the model by employing the parameter restriction $\eta_L^e = 0$. For the steady state to be unique he finds necessary and sufficient conditions equal to

$$\frac{\theta - 1}{1 - \eta_A}n + \rho > 0\,. \tag{90}$$

Moreover, Arnold proves that for this parameter restriction adjustment dynamics are locally unique. Since $\eta_A < 1$ holds in any case and $\theta \geq 1$ is empirically more plausible we will consider restriction (90) for the simulations.

The objective is to solve the four-dimensional system of differential equations (85) - (88), taking into account the static equation (89), which must hold at all points in time. Since the steady state can be determined numerically only, the algorithm computes the steady state of the system first by applying a standard algorithm for solving non-linear equations. The choice of $k(0) = k_0$ and $a(0) = a_0$ as initial boundary conditions is obvious since $k$ and $a$ are the state variables. Again, there is some freedom when it comes to the determination of boundary conditions. We have set the RHS of equations (87) and (88) equal to zero. However, if we choose different final boundary conditions the performance of the algorithm is not affected.

---

[51] The presence of the static efficiency condition (eq. (89)) is due to the fact that labor does neither enter final output nor R&D linearly. Hence, it is in general not possible to solve for the optimal amount of labor explicitly.

Moreover, we choose once more, as an initial guess, all variables to be constant at their steady state values. This always leads to quick convergence, indicating that the procedure is relatively robust with respect to the initial guess. We control the relative error of the solution to be below $10^{-4}$ at every mesh point by employing the error estimation equation (44). This is already the case for a mesh size of $M = 100$.

We want to demonstrate that the dynamic system (85) - (89) represents a system of stiff differential equations. The algebraic equation (89) does not contribute to the dynamics. Therefore, we project dynamics on the four-dimensional manifold defined by (89). For the adjustment path to be unique the system must exhibit a two-dimensional stable manifold. Employing the Hartman-Grobman theorem for the unique, hyperbolic steady state, the linearized system has to exhibit two eigenvalues with negative and two with positive real part. Moreover, for the system to be stiff, the eigenvalues associated to the stable manifold have to differ significantly. Since the steady state and the linearized system can only be calculated numerically, we employ numerical experiments to verify this result. We first consider a benchmark economy, which calibration agrees with those generally invoked in the economic literature (see Prescott, 1986, Lucas, 1988, Jones, 1995b, Ortingueiry and Santos, 1997, and Eicher and Turnovsky, 2001). The set of parameters used for the benchmark economy is: $\sigma_L = \sigma_A = 0.6$, $\sigma_K = 0.4$, $\delta = 0.05$, $n = 0.015$, $\eta_A = 0.6$, $\eta_L = 0.5$, $\eta_L^p = 1$, $\rho = 0.04$, $\theta = 1.5$, and $\alpha_J = \alpha_F = 1$. In this case, the ratio of the stable eigenvalues amounts to 11.9. We deviate each of the parameters by 20% and calculate the ratio of the stable eigenvalues again. The results are shown in Table 2. Note that a deviation in $\sigma_L$ causes also a deviation in $\sigma_A$ and $\sigma_K$, since $\sigma_A = \sigma_L$ and $\sigma_A = 1 - \sigma_K$.[52] It can be seen from Table 2 that the ratio of the stable eigenvalues differs substantially

---

[52]We adopt this setting from Jones (1995a). Eicher and Turnovsky (1999 and 2001) allow for a more general parametrization of the production function. However, assuming $\sigma_L \neq 1 - \sigma_K$ would need the final production sector to exhibit an external effect for maintaining the assumption of perfect competition for this sector.

for a broad range of parameter values, i.e. the ratio amounts to more than 10 for most of the chosen parameter sets. This has two important implications. First, from an economic perspective, overshooting behavior is very likely to occur when transitional dynamics resulting from a specific shock is investigated as described in Section 2.2. Second from a numerical perspective, this model is difficult to solve numerically as described in Section 2.3, since it exhibits the characteristics of stiff differential equations. For example, for the backward shooting algorithm it would be very difficult or even impossible to meet given initial conditions.

Table 2: Sensitivity analysis for the Jones model

| parameter | direction of change | ratio of stable eigenvalues |
|-----------|---------------------|------------------------------|
| $\sigma_A$ | $+20\%$ | 18.7 |
| $\sigma_A$ | $-20\%$ | 7.9 |
| $\eta_A$ | $+20\%$ | 14.0 |
| $\eta_A$ | $-20\%$ | 10.7 |
| $\eta_L$ | $+20\%$ | 10.3 |
| $\eta_L$ | $-20\%$ | 14.5 |
| $\delta$ | $+20\%$ | 12.8 |
| $\delta$ | $-20\%$ | 11.0 |
| $\rho$ | $+20\%$ | 12.1 |
| $\rho$ | $-20\%$ | 11.7 |
| $\theta$ | $+20\%$ | 11.4 |
| $\theta$ | $-20\%$ | 12.6 |
| $n$ | $+20\%$ | 10.9 |
| $n$ | $-20\%$ | 13.2 |

Before investigating a specific shock, we want to illustrate the dynamics on the two-dimensional stable manifold of the model. We select initial values for 28 trajectories on a circle around the steady state. Simulation results projected into the $(k, a)$-plane can be seen in Figure 9. The initial value of each trajectory is indicated by a cross, while the steady state is indicated by a circle. Transitional behavior proceeds

similar to Figure 4 in Section 2.2. Accordingly, convergence can be divided into two phases for most trajectories. In the first phase, trajectories converge towards the submanifold associated with the eigenvalue of smaller magnitude, in absolute terms. In the second phase, trajectories move along this submanifold. As can be seen from Figure 9, convergence in the first phase mainly takes place along the $k$ coordinate, while in the second phase, $k$ and $a$ increase or decrease jointly.
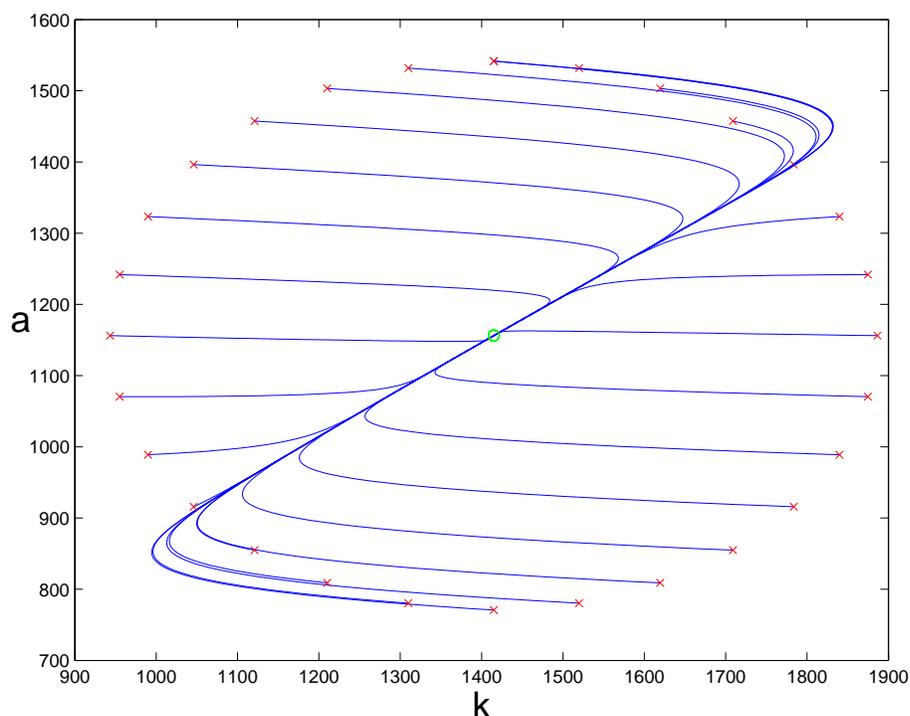


Figure 9: Phase diagram of the Jones (1995a) model

This result entails an economic interpretation. Note first that in the Jones model, both $a$ and $k$ serve as an asset, which an individual household can save in. While capital yields future payments in the form of interest, patents yield future payments in the form of monopoly rents. Now consider an economy that possesses its steady state value of patents $a$, but deviates from its steady state value of capital $k$. Transitional

dynamics for this economy will roughly equal that of a Ramsey economy. Capital will be accumulated or deaccumulated until the steady state value is reached, but the (scale-adjusted) number of patents $a$ will roughly remain on its steady state value during transition. Therefore, transitional dynamics are almost monotonic. This can be seen by considering trajectories in Figure 9 that start horizontally from the steady state. By contrast, an economy that possesses its steady state capital stock but deviates for its number of patents from the steady state value will exhibit an overshooting pattern for its stock of capital. This is shown in Figure 10. In this simulation, a shock is investigated, which consists in a reduction of the stock of patents only. In a first phase the economy accumulates capital until the ratio $\frac{k}{a}$ roughly equals its steady state value. In this phase, both types of assets are interchanged in the sense that the accumulation of capital is accompanied by a deaccumulation of patents.[53] In the second phase, the economy deaccumulates both assets until the steady state is reached. This can be seen in Figure 10 (ii) and (iii). Figure 10 (iv) displays the phase diagram of the $(k, a)$-plane and Figure 10 (i) shows that consumption follows the overshooting pattern of capital.

To summarize, adjustment dynamics can be divided into two stages. In the first stage it is dominated by accumulation (or deaccumulation) of capital, whereas in the second stage, accumulation (or deaccumulation) of patents or knowledge drives economic development. This is perfectly in line with the literature on economic growth, which assigns the accumulation of knowledge to be the source of long-run growth, whereas the accumulation of capital is responsible for growth in the medium run. (see e.g. Funke and Strulik, 2000a, for a model, which explicitly distinguishes between different phases of economic development).

Finally, we present simulation results originating from a change in the model's

---

[53]In a growing economy, the scale adjusted number of patents $a$ can be deaccumulated by, for instance, keeping $A$ constant.
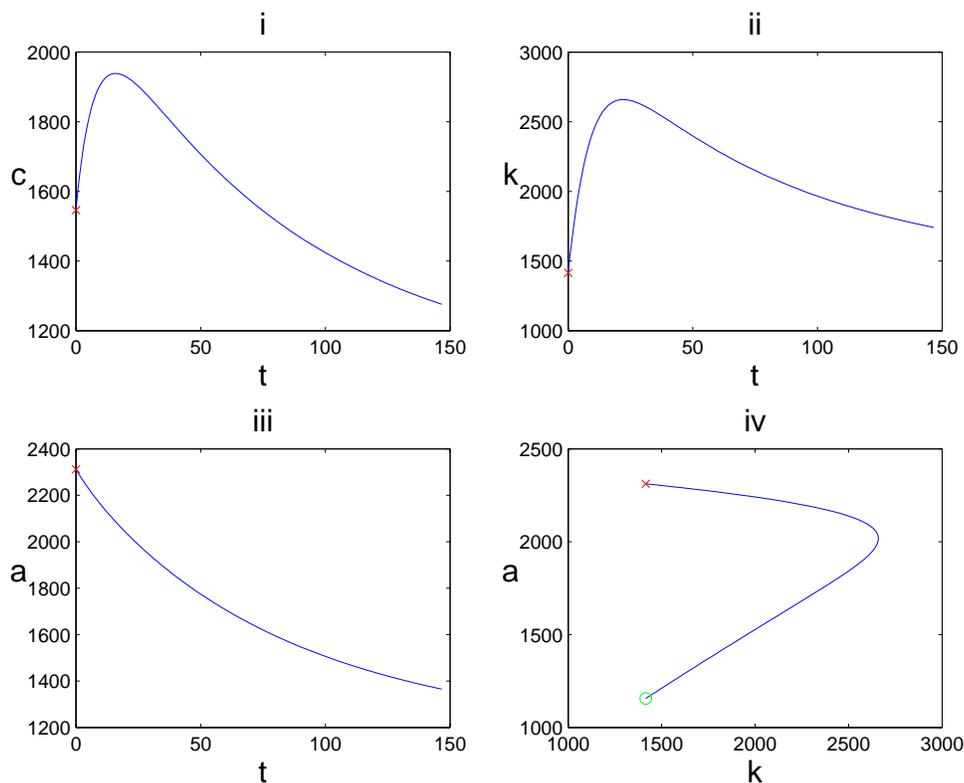
Figure 10: Transition of the Jones (1995a) model

parameters. The transition process considered below results from a combination of two simultaneous shocks. Specifically, it is assumed that the overall productivity parameter in the production function for final output $\alpha_F$ increases from 1.0 to 1.3, while the overall productivity parameter in the production function for new ideas $\alpha_J$ decreases from 1.0 to 0.9. Figure 11 gives a summary of the adjustment process. The plots (i) to (iii) show the time path of the jump variables $c$, $\phi$, $v_a$, plots (iv) and (v) display the time path of the state variables $k$ and $a$, while plot (vi) gives the projection of the adjustment trajectory into the $(k, a)$-plane. For plot (iv) the initial value of the trajectory is indicated by a cross, and the final value by a circle.
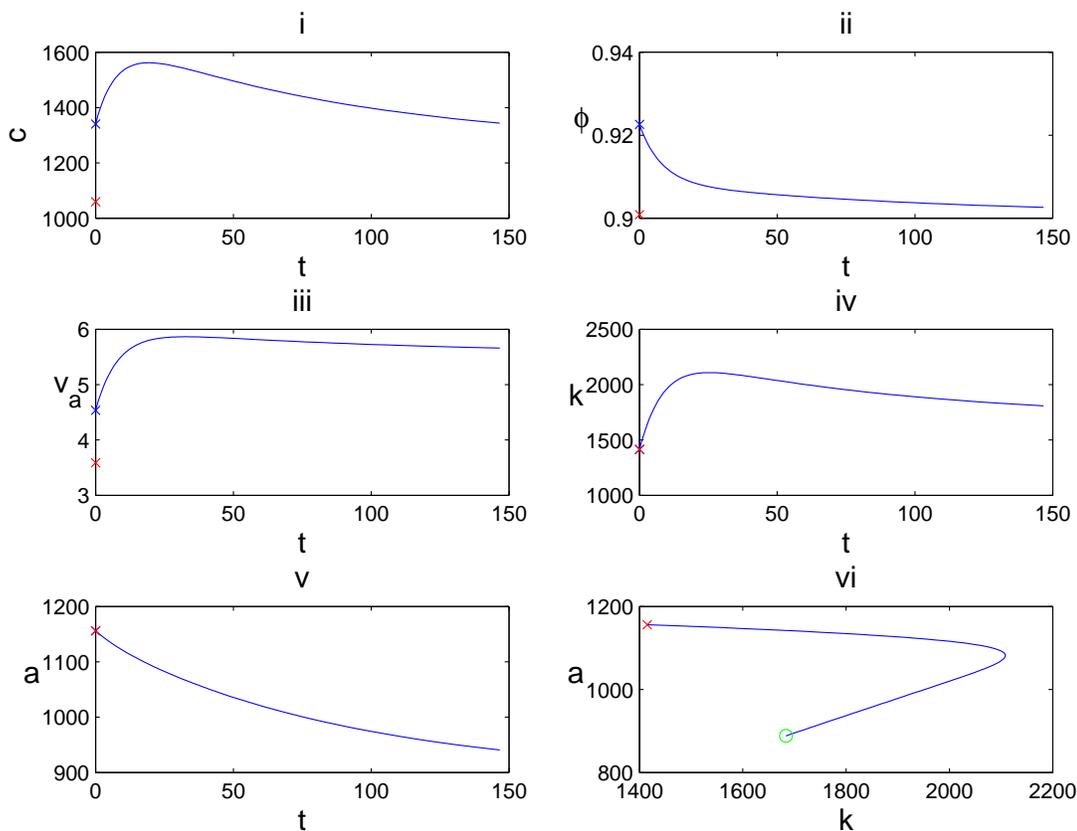
Figure 11: Summary of the transition of the Jones (1995a) model

### 3.4.2 The Lucas (1988) model

As a second example we want to simulate the Lucas (1988) model, which is also discussed in Mulligan and Sala-i-Martin (1993), Caballe and Santos (1993), and Benhabib and Perli (1994). The analysis of the model by these authors differs in one technical aspect, namely how stationary variables are constructed for analyzing the dynamic system. While Mulligan and Sala-i-Martin (1993) and Benhabib and Perli (1994) construct stationary variables by creating ratios of endogenous variables that exhibit the same balanced growth rate, Lucas (1988) and Caballe and Santos (1993)

apply scale-adjustment. The latter method consists of slowing down the motion of variables according to their respective balanced growth rates. Scale-adjustment possesses the advantage that the time paths of variables are obtained right away and, for example, utility integrals can easily be computed. Moreover, by employing scale adjustment an important characteristic of the model becomes apparent. The long-run equilibria in the scale adjusted system are not represented by an isolated fixed point, but form a center manifold of stationary equilibria. Therefore, following Section 2.1 theorem 13, the specific steady state to which the economy converges depends on the initial conditions, i.e. the initial endowment of physical and human capital.

Final output is produced from physical capital $k$ and human capital $h$. The share $u$ of human capital is used for final output production

$$y = A k^\alpha (uh)^{1-\alpha} h^\gamma \tag{91}$$

with the elasticity $\alpha$ of physical capital in output production, the overall productivity parameter $A$, and the external effect $\gamma$ of human capital in final output production. Due to human capital spill over effects there are increasing returns to scale in the production sector. The remainder $1 - u$ of human capital is employed to increase human capital according to

$$\dot{h} = \delta(1 - u)h \tag{92}$$

with overall productivity parameter of human capital accumulation $\delta$. A representative household maximizes intertemporal utility of consumption $c$

$$\max_c \int_0^\infty \frac{c^{1-\theta}}{1-\theta} e^{-\rho t} dt \tag{93}$$

with constant elasticity of intertemporal substitution $\sigma^{-1}$ and discount rate $\rho$. The first order conditions for optimal solutions in terms of a system of differential equa-

tions read (see Benhabib and Perli, 1994, for the derivation)

$$\dot{k} = A k^{\alpha} h^{1-\alpha+\gamma} u^{1-\alpha} - c \tag{94}$$

$$\dot{h} = \delta(1-u)h \tag{95}$$

$$\dot{c} = \sigma^{-1} c (\alpha A k^{\alpha-1} h^{1-\alpha+\gamma} u^{1-\alpha} - \rho) \tag{96}$$

$$\dot{u} = u \left( \frac{(\gamma-\alpha)\delta}{\alpha}(1-u) + \frac{\delta}{\alpha} - \frac{c}{k} \right) . \tag{97}$$

Balanced growth requires $u$, $c/k$ as well as $k^{\alpha-1} h^{1-\alpha+\gamma}$ to be constant. The latter requirement in turn demands $(1-\alpha)\frac{\dot{k}}{k} = (1-\alpha+\gamma)\frac{\dot{h}}{h}$.

In this simple case the common balanced growth rate $\mu$ of $k$ and $c$ can be computed by solving the system under balanced growth assumptions:

$$\mu = \frac{1-\alpha+\gamma}{(1-\alpha+\gamma)\sigma - \gamma} (\delta - \rho)$$

Growth is balanced if the four variables of the system satisfy three equations:

$$1-u = \frac{1-\alpha}{(1-\alpha+\gamma)\sigma - \gamma}(1 - \rho/\delta) \tag{98}$$

$$c/k = ((\gamma-\alpha)\psi\mu + \delta)/\alpha \tag{99}$$

$$k^{\alpha-1} h^{1-\alpha+\gamma} = \frac{\sigma\mu + \rho}{\alpha A} u^{\alpha-1} \tag{100}$$

where $\psi := (1-\alpha)/(1-\alpha+\gamma)$. We construct a stationary system by employing scale-adjustment. The transformed variables are

$$ke^{-\mu t}, \quad he^{-\psi\mu t}, \quad ce^{-\mu t} \quad \text{and} \quad u.$$

To avoid extra notation we continue to use the old designations of variables. The new, adjusted growth rates are reduced by the constants of adjustment, $\mu$ and $\psi\mu$, respectively. The growth rate of $u$ remains unchanged. Therefore, the transformed

system reads

$$\dot{k} = A\,k^{\alpha}h^{1-\alpha+\gamma}u^{1-\alpha} - c - \mu k \tag{101}$$

$$\dot{h} = \delta(1-u)h - \psi\mu h \tag{102}$$

$$\dot{c} = \sigma^{-1}c(\alpha A\,k^{\alpha-1}h^{1-\alpha+\gamma}u^{1-\alpha} - \rho) - \mu c \tag{103}$$

$$\dot{u} = u\left(\frac{(\gamma-\alpha)\delta}{\alpha}(1-u) + \frac{\delta}{\alpha} - \frac{c}{k}\right). \tag{104}$$

Due to scale adjustment, balanced growth solutions represented by equations (98), (99) and (100) now turn into stationary points of system (101) - (104).[54] Therefore, according to corollary 10, the system exhibits a one dimensional center manifold of stationary equilibria. The model exhibits two state variables, $k$ and $h$. Hence, according to theorem 13 adjustment dynamics are locally unique, if the Jacobian matrix of system (101) - (104) evaluated at the center manifold of stationary equilibria exhibits one zero eigenvalue, one eigenvalue with negative real part, and two eigenvalues with positive real part.

The set of valid parameter values has been investigated by Benhabib and Perli (1994). They define the subsets

$$\Theta_1 = \left\{(A,\alpha,\gamma,\delta,\rho,\sigma) > 0 \,\middle|\, 0 < \rho < \delta \,\wedge\, \sigma > 1 + \frac{\rho(\alpha-1)}{\delta(1-\alpha+\gamma)}\right\} \tag{105}$$

$$\Theta_2 = \left\{(A,\alpha,\gamma,\delta,\rho,\sigma) > 0 \,\middle|\, \delta < \rho < \frac{\delta(1-\alpha+\gamma)}{\alpha-1} \,\wedge\, \right.$$
$$\left. \sigma < 1 + \frac{\rho(\alpha-1)}{\delta(1-\alpha+\gamma)}\right\}. \tag{106}$$

For parameter sets originating from $\Theta_1$ the balanced growth path and trajectories converging towards the balanced growth path are unique. For parameter sets originating from $\Theta_2$ the balanced growth path is unique, but for local transitional dynamics two different cases occur. $\Theta_2$ can further be divided into $\Theta_2^A$ and $\Theta_2^B$. If

---

[54]It is straightforward to verify that the right hand sides of system (101) - (104) are linearly dependent, and that equations (98), (99) and (100) are the only solution.

parameters from $\Theta_2^B$ are chosen, the balanced growth path is unstable.[55] For parameters from $\Theta_2^A$ the case of indeterminacy occurs, i.e. given initial endowments of physical and human capital, there are infinitely many trajectories converging towards the balanced growth path.[56] [57] The distinction between $\Theta_2^A$ and $\Theta_2^B$ is rather complicated to state analytically, which is why we refer to the appendix of Benhabib and Perli (1994) for this issue. For both sets of parameters $\Theta_1$ and $\Theta_2$, stationary equilibria exhibit an interior share $u$, i.e. $0 < u < 1$, and trajectories converging towards this equilibrium satisfy the transversality conditions. Note that the fact that $u$ is interior also implies that the denominator of equation (98) is not zero. We select the benchmark set of parameters $A = 1$, $\alpha = 0.3$, $\delta = 0.1$, $\gamma = 0.3$, $\sigma = 1.5$ and $\rho = 0.05$ that lies in $\Theta_1$ and, hence, guaranties determinate adjustment dynamics. The parameters are chosen according to Prescott (1986), Lucas (1988), Benhabib and Perli (1994), and Ortingueira and Santos (1997).

Numerical computation employing the relaxation algorithm requires the solution of the system of differential equations (101) - (104) with two initial conditions and two final conditions. The initial conditions are given by the initial values of state variables $k(0) = k_0$, $h(0) = h_0$. Final conditions should ensure convergence towards the center manifold given by equations (98), (99) and (100). However, the relaxation algorithm can only incorporate two final boundary conditions.[58] Therefore, we incorporate stationary conditions for the state variables, implicitly defined by $\dot{k}(\infty) = 0$ and

---

[55]The Jacobian matrix numerically evaluated at the center manifold of system (101) - (104) exhibits a zero eigenvalue and three eigenvalues with positive real part.

[56]This means that the Jacobian matrix numerically evaluated at the center manifold of system (101) - (104) exhibits a zero eigenvalue, one eigenvalue with positive real part, and two eigenvalues with negative real part.

[57]We did not calculate the eigenvalues analytically, but computed them numerically employing a representative parametrization from $\Theta_1$, $\Theta_2^A$, and $\Theta_2^B$, respectively.

[58]Of course, equations (98), (99) and (100) could be incorporated together by constructing an equation which is the norm of equations (98), (99) and (100). This was not necessary for solving this model numerically.

$\dot{h}(\infty) = 0$. We estimate the global error of simulation results employing equation (44). We restrict the global error below $10^{-4}$, which was fulfilled with a mesh of $M = 100$.

To demonstrate adjustment dynamics towards the center manifold of stationary equilibria we simulate transitional dynamics of several economies with different initial endowments of physical and human capital. We select initial values for 20 trajectories on a circle around one point of the center manifold. Simulation results projected into the $(k, h)$-plane can be seen in Figure 12.[59] Trajectories originating from unbalanced initial conditions are indicated by dashed lines, while the center manifold is indicated by a solid line. The initial value of each trajectory is indicated by a cross, while the original stationary point on the center manifold is indicated by a circle.

Transitional dynamics can be interpreted in the context of theorem 13 of Section 2.1. Numerical evaluation of the Jacobian matrix at the center manifold confirms that the matrix possesses a zero eigenvalue, one negative (real) eigenvalue and two positive (real) eigenvalues. Therefore, the stable manifold of the center manifold is two-dimensional, and transitional dynamics are determinate. The stable manifold can be divided into one-dimensional fibres. Each fibre is associated with one particular stationary point on the center manifold in the sense that all points of each fibre converge towards the same stationary point on the center manifold. Therefore, economies converge towards different points on the center manifold, depending on their initial endowments of physical and human capital, unless they belong to the same fiber.

For economic interpretation note first that transitional dynamics differ with respect to the accumulation of physical capital and human capital. We exemplify this

---

[59]Notice that Figure 12 is very similar to Figure 1 in the seminal paper of Lucas (1988). The difference is that Lucas constructed transitional dynamics due to what he conjectured it to be. However, Figure 12 in this thesis is due to numerical simulations. Moreover, transitional dynamics correspond to the theoretical results presented in Section 2.1.
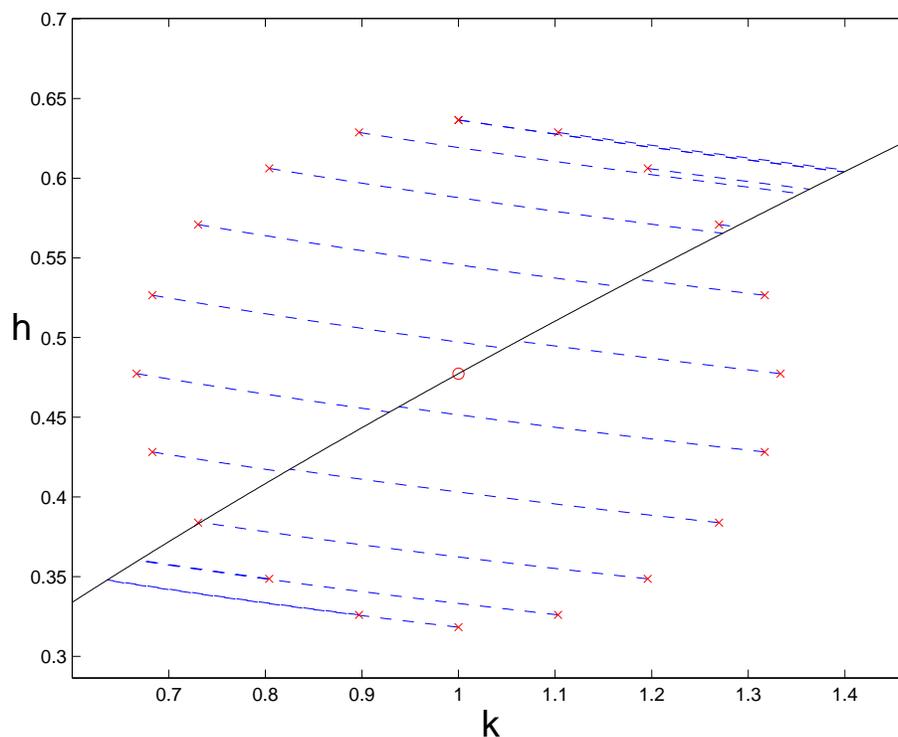
Figure 12: Center manifold of stationary points and trajectories starting on a circle

by considering two different economic shocks. In Figure 13 we show simulation results of two economies, which initial endowments of physical and human capital are reduced with respect to a benchmark economy that is in balanced growth. While the economy represented by solid trajectories exhibits one third less of human capital than the benchmark economy as initial endowment, the economy represented by dashed trajectories exhibit one third less of physical capital than the benchmark economy as initial endowment. This can be seen in Figure 13 (iv) that shows the $(k, h)$-phase diagram of transitional dynamics. The benchmark economy is indicated by a circle, while initial points of both unbalanced economies are indicated by crosses. Figure 13 (i), (ii), and (iii) display the time paths of physical capital, human capital and consumption, respectively. Steady state levels for the benchmark economy are

indicated by dotted lines.
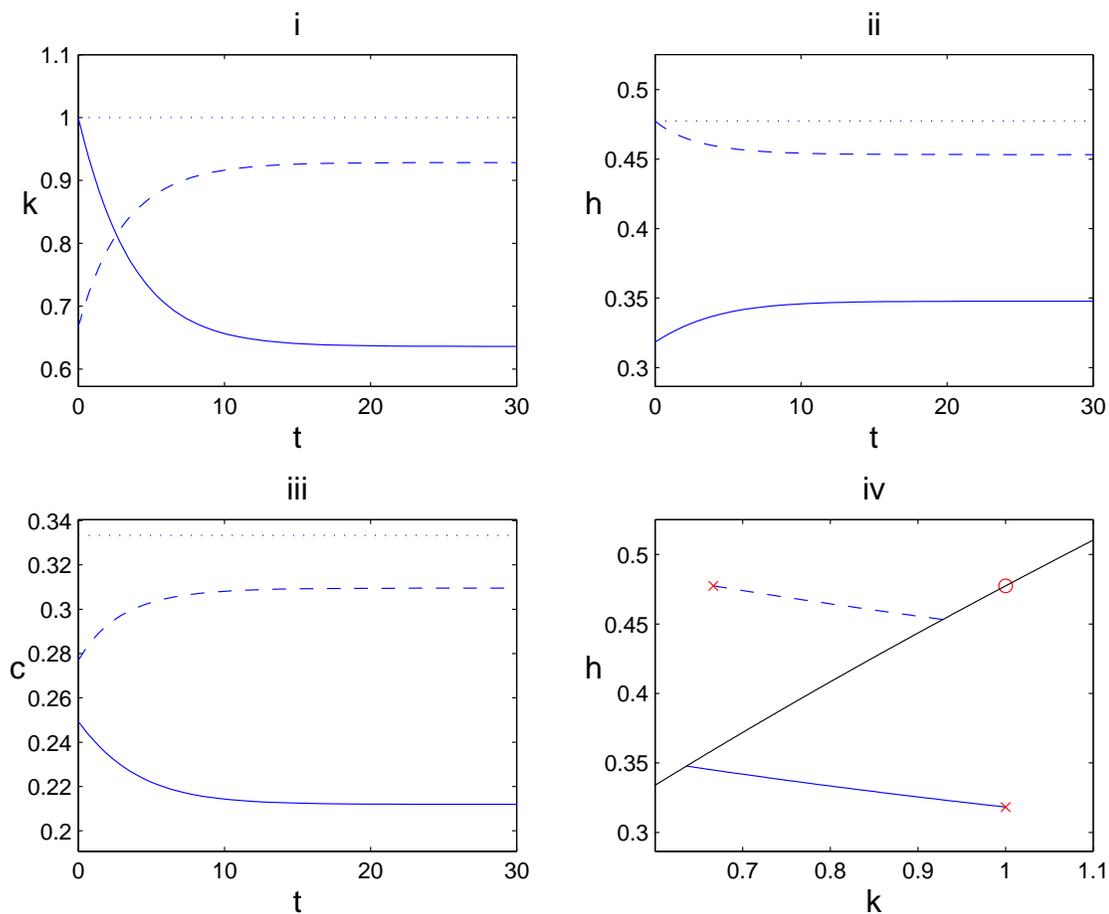


Figure 13: Comparison of two different economies

It can be recognized, that the economies do not belong to the same fibre and, hence, converge towards different long-run equilibria. The economy, whose initial endowment of human capital is reduced (solid line), converges towards a lower stationary equilibrium than the economy, whose initial endowment of physical capital is reduced (dashed line). This reflects the fact that human capital is the engine of growth in the Lucas model, and that human capital serves as the only input for hu-

man capital accumulation. Therefore, reducing the stock of human capital hampers economic growth considerably, and a catch-up towards the benchmark economy is relatively small. During the transition process, human capital is accumulated and physical capital is deaccumulated. By contrast, reducing the stock of physical capital hampers economic growth considerably less. During the transition process capital is accumulated and human capital is deaccumulated. The long-run equilibrium that the economy approaches possesses a considerably higher level of (scale-adjusted) consumption than the economy represented by solid trajectories. Though, both levels of consumption are still lower than that of the benchmark economy.

First it must be noted that the properties of the center manifold of stationary equilibria cause dependence of long-run equilibria from initial conditions. This is, therefore, not an implication of the special parametrization, but a general attribute of the Lucas (1988) model, also emphasized by Caballe and Santos (1993). The model contributes a pessimistic view to the debate of convergence discussed by Baumol (1986), DeLong (1988), Barro (1991), Barro and Sala-i-Martin (1992), Sala-i-Martin (1997), and Sala-i-Martin et al. (2004) among others, because it does not give rise to absolute $\beta$-convergence for economies exhibiting identical parameters. Rather, it implies that the particular steady state an economy reaches is higher, the higher the initial level of human capital. The same influence holds for the initial stock of physical capital, but this connection is much smaller in magnitude as discussed earlier. A positive influence of human capital on growth is frequently found in empirical studies (e.g. Barro, 1991, and Sala-i-Martin et al., 2004), whereas the stock of human capital is estimated by school enrolment rates at the base year of the regression.

Moreover, the model implies club convergence. Countries of the same fibre converge towards the same steady state, but might exhibit very different initial endow-

ments of physical and human capital. Therefore, countries of the same fibre (or fibres close by) form a convergence club. E.g. Quah (1996 and 1997), Jones (1997), Pritchet (1997), and Sala-i-Martin (2006) found club-convergence empirically. However, the model does not imply convergence towards only two different steady states as labeled as the emergence of "twin-peaks", and empirically found by Quah (1996 and 1997), and Jones (1997). According to the model's dynamics, each economy converges towards its own steady state with probability 1, if the distribution of initial endowments is random. A more comprehensive analysis of the convergence pattern implied by this kind of model would include the possibility of human capital spill-over effects between different countries as suggested by Lucas (1993), and is, therefore, beyond the scope of this thesis.

We want to emphasize that the numerical solution of the scale-adjusted system instead of the numerical solution of the reduced system (as in Benhabib and Perli, 1994) has a major advantage. Time paths of variables are obtained right away. Therefore, different economies can easily be compared with respect to initial values of consumption or their long-run equilibria. Moreover, utility integrals can be computed right away. In Figure 14, we compare two economies, which differ in their initial endowments of physical and human capital. Figure 14, (i) displays the $(k, h)$-phase diagram of unscaled variables, Figure 14, (ii) displays the $(k, h)$-phase diagram of scaled variables, and Figure 14, (iii) displays the time paths of consumption. The special characteristic of transitional dynamics for this example is that one economy (dashed line) starts with a higher initial value of consumption than the comparison economy (solid line). However, eventually, consumption of the economy represented by a dashed line is overtaken by consumption of the comparison economy, and converges towards a steady state represented by lower values of consumption and output. Since it is not straightforward to decide, which economy enjoys a higher value of the

sum of discounted utility, we compute utility integrals. The value of the utility integrals amount to -4.81 for the solid path and -9.22 for the dashed one.
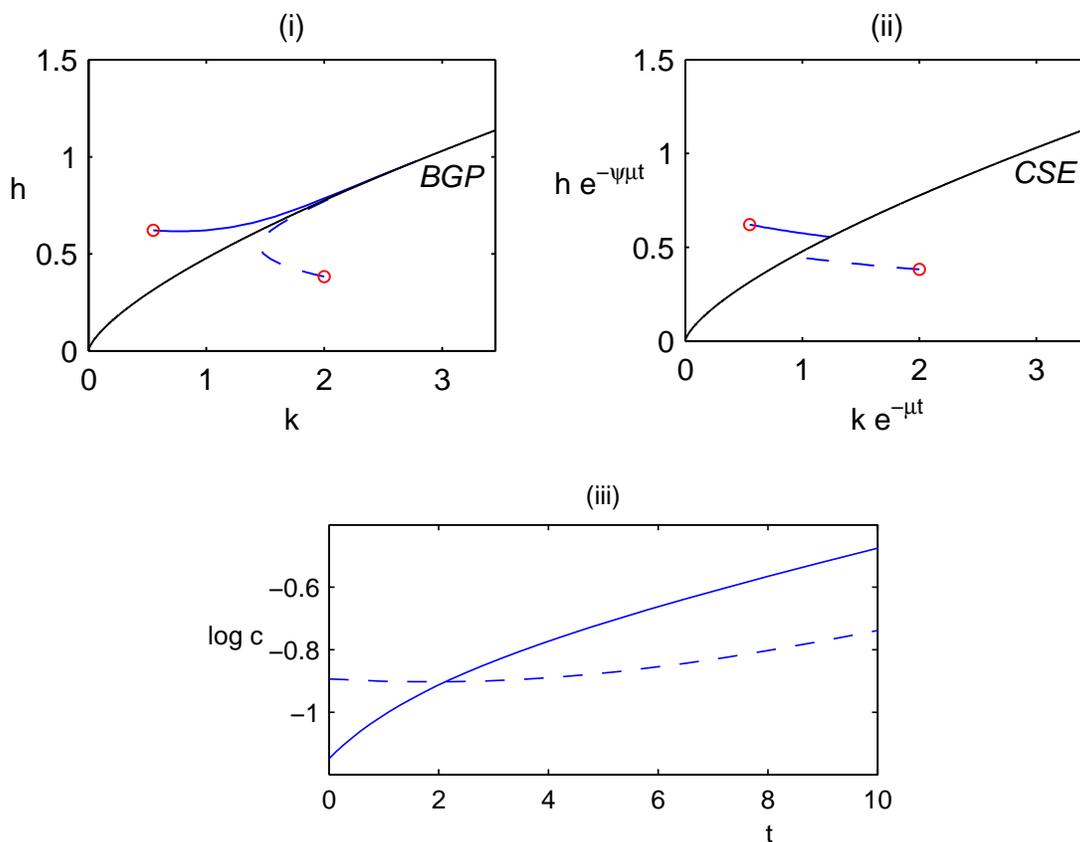


Figure 14: Summary of the transition of the Lucas (1988) model

Finally, we want to present simulation results with a set of parameters originating from $\Theta_2^A$, implying an indeterminate adjustment path with given initial endowments of physical and human capital. The set of parameters we choose is $\sigma = 0.15$, $\rho = 0.0505$, $\beta = 0.65$, $\gamma = 0.1$, $\delta = 0.05$, and $A = 1$. Employing this set of parameters yields a zero eigenvalue, a positive eigenvalue, and a pair of complex eigenvalues with negative real part for the Jacobian matrix evaluated at the center manifold. According to theorem 13 of Section 2.1 unique adjustment dynamics require three

initial conditions. Therefore, given initial endowments of physical and human capital households have one degree of freedom to choose the initial value of consumption $c(0)$ and the initial share of human capital for final output production $u(0)$, such that the resulting trajectory converges towards the center manifold.

For numerical computation employing the relaxation algorithm trajectories have to be unique. Otherwise, the Jacobian matrix of equation (36) is singular. Therefore we add an additional initial condition for $u$,

$$u(0) = u^* \tag{107}$$

with the unique steady state value of $u$, $u^*$, for each trajectory we compute. Simulation of system (101) - (104) with the relaxation algorithm, however, yield slow convergence or no convergence of the iteration process. Obviously, the combination of the center manifold together with oscillating behavior renders difficulties, or the initial condition (107) is not transversal to the set of solution trajectories.[60] Therefore, we transform the center manifold of stationary equilibria of system (101) - (104) into a saddle-point following Benhabib and Perli (1994). We intoduce $x := kh^{\frac{1-\alpha+\gamma}{\alpha-1}}$ and $q := \frac{c}{k}$. The transformed system reads

$$\dot{x} = Ax^{\alpha}u^{1-\alpha} - \frac{\delta(1-\alpha+\gamma)}{1-\alpha}(1-u)x - qx \tag{108}$$

$$\dot{u} = u\left(\frac{(\gamma-\alpha)\delta}{\alpha}(1-u) + \frac{\delta}{\alpha} - q\right) \tag{109}$$

$$\dot{q} = q^2 + A\left(\frac{\alpha}{\sigma}-1\right)x^{\alpha-1}u^{1-\alpha}q - \frac{\rho}{\sigma}q \tag{110}$$

Now, equations (98), (99) and (100) define a unique saddle-point of system (108) - (110).

Figure 15 displays six different trajectories converging towards the steady state in the $(u, x, q)$-phase diagram. For initial conditions, we fix the value of $u$ according to

---

[60]We tried different initial conditions, but were not successful in finding one which yields fast convergence.

equation (107), and the initial value of $x$ to six different values lower than its steady state value. For final boundary conditions we fix the value of $x$ to its steady state value. Convergence proceeded quickly, and the relative global error was controlled to remain smaller than $10^{-4}$. Simulation results coincide with that of Brunner and Strulik (2002), although they applied the backward integration method for numerical calculation.
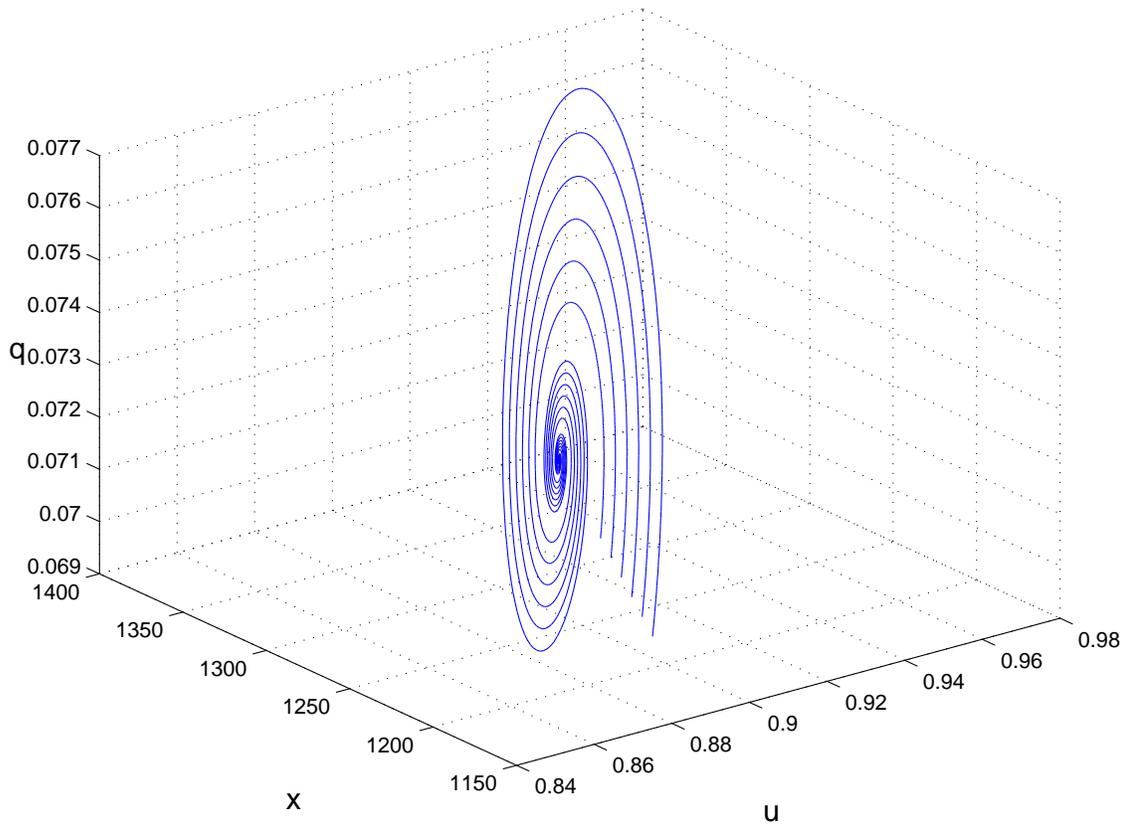


Figure 15: Phase diagram of the Lucas model with indeterminate adjustment paths

## 3.5  Summary

We propose the relaxation algorithm as a powerful and efficient procedure to investigate the transition process of infinite-horizon continuous-time optimization models numerically. The procedure has the advantages that it can easily simulate transitional dynamics of multi-dimensional models. In contrast to many other procedures it is generic with respect to the model dimension, in particular with respect to the dimension of the state space. Moreover, computational costs increase only moderately with the model dimension. We demonstrated the potential of the algorithm by simulating the transition process of the Jones (1995a) and Lucas (1988) model numerically. Both models' dynamic systems exhibit characteristics, which make the application of standard procedures difficult. Moreover, we show that additional theoretical insights can be obtained by numeric simulation, which, for these examples, cannot be derived analytically. Although we demonstrate the algorithm's potential by simulating infinite-horizon growth models, the application of the method is neither restricted to growth models nor to infinite horizon problems.

# 4 Anticipated Shocks in Continuous-time Optimization Models: Theoretical Investigation and Numerical Solution

This Section is based on the working paper of Trimborn (2007).

## 4.1 Introduction

Dynamic macroeconomic theory very often assumes the agents in the model to experience perfect foresight. For example, a representative, infinite horizon consumer maximizes discounted utility subject to his or her budget constraint. Then, it is assumed that at the point of time when the maximization takes place, the maximizing agent is aware of the whole set of information. More precisely, the agent knows the future time path of variables that are exogenous to him. Then, following Bellman's principle, the decisions are time-consistent, that is, at a later point of time the same solution is optimal if the set of information did not change.

Changes in the underlying parameters, e.g. tax rates or preference parameters are frequently analyzed by assuming a sudden, immediate shock. More precisely, at a specific point of time, say $t_0$, new parameter values are applied and the optimizing agent experiences this information at exactly the same point of time. Then, he or she can optimize over the remaining time horizon. Since the point of time of the shock and the point of time of the information propagation coincide, i.e. the shock is not preannounced or anticipated, control variables can jump at $t_0$. State variables cannot jump at $t_0$ by construction.[61]

This simplifying assumption is very useful to analyze e.g. policy measures in a stylized way, however, in some cases it may be oversimplifying. The reason is that

---

[61]For the analysis of anticipated shocks with jumps in the state variables see Vind (1967) and Auernheimer and Lozada (1990).

usually the time of the parameter changes and the information propagation to the agent do not coincide. For example, policy measures are usually announced some time before they take place. Therefore, the stylized analysis of unexpected shocks cannot take account for anticipatory actions conducted by the agents. In addition, the shock might comprise a schedule of policy measures, which do not enter into force simultaneously.

It is straightforward to state necessary conditions in the context of optimal control, which have to hold at points of anticipated shocks. This is the familiar requirement of continuity for adjoint variables, also known as the first Weierstraß-Erdmann corner condition. However, they do not imply that control variables are continuous. Since control variables are only required to be piecewise continuous, they can potentially jump twice, first at the point of the announcement and secondly at the point where the parameter change actually takes place. Requirement of continuity for adjoint variables determines the height of the jump under weak regularity assumptions.

In a second step, we describe how numerical computation can be conducted efficiently. We propose the relaxation algorithm as described in Section 2 to simulate continuous-time, infinite horizon optimization models with expected shocks. The algorithm allows to solve expected shocks without any a priori information about the behavior of the model at the time of the parameter change. Only the dynamic system has to be provided together with the underlying parameters and their changes along time. Moreover, the size of the jump in the control variables is calculated automatically. This is illustrated by employing the Ramsey-Cass-Koopmans model as a concise example.

The analysis of preannounced policy measures in the context of perfect foresight optimizing agents has a long tradition. The continuity of adjoint variables (or market prices like asset prices) has already been exploited in many cases (e.g. Wilson, 1979,

Judd, 1985, Howitt and Sinn, 1989, Turnovsky, 1996, Ch. 11). All the more, it is surprising that many standard textbooks of dynamic optimization in economics do not cover how to treat anticipated policy changes in the context of optimal control.[62] Moreover, to our knowledge no formal analysis of this problem has been made in the economic literature. Vind (1967) and Auernheimer and Lozada (1990) analyze how to treat jumps in the state variable theoretically, which is a related problem.

On the other hand, no attempt has been made so far to construct algorithms which can solve these kind of problems in a generic way. Buiter (1984) describes how to solve the linearized problem numerically. Perhaps this explains, why, to our knowledge, there exists hardly any numerical solutions to models with anticipated policy changes in the literature (see e.g. Funke and Strulik, 2000a and 2000b, for simulation results with a technically equivalent problem).

In Section 4.2, we investigate the treatment of expected shocks in the context of optimal control theoretically. In Section 4.3, we describe the numerical implementation of expected shocks. In Section 4.4, we present the Ramsey-Cass-Koopmans model as a concise example. In Section 4.5 we conclude.

## 4.2  Theoretical investigation of expected shocks

Consider an agent who solves a finite horizon or infinite-horizon, perfect-foresight, continuous-time, optimal control problem

$$\max_u \int_{t_0}^{T} f(t,x,u)dt \tag{111}$$
$$\text{s.t.} \qquad \dot{x} = g(t,x,u), \qquad x(0) = x_0$$

---

[62]Kamien and Schwartz (1991) and Intriligator (1971) state necessary conditions for corners for problems of calculus of variation, but not for problems of optimal control. Seierstad and Sydsæter (1987) state conditions for both problems, but without proof. Chiang (1992) does not cover this subject. As we will show, these conditions even have to be modified for the problem at hand.

whereas $x$ denotes a $n_x$ dimensional state variable and $u$ a $n_u$ dimensional control variable. The functions $f$ and $g$ are, for the time being, assumed to be continuously differentiable functions in all three arguments. We assume that $x$ is a continuous function of time, i.e. $x$ cannot jump. On the other hand, $u$ is only assumed to be a piecewise continuous function of time, i.e. $u$ can exhibit interior jumps. The determinate terminal time $T$ can also equal infinity.

To start with, we want to state necessary conditions for corners in $x$ and jumps in $u$. These are based on the Weierstraß-Erdman corner conditions, though these conditions are stated for problems of the calculus of variation.[63] However, it is straightforward to extend these conditions to problems of optimal control (see Bryson and Ho, 1975, pp. 125, for a statement of the conditions without proof). We transform problem (111) into a problem of the calculus of variation. We introduce the piecewise continuous, time-dependent adjoint variable $\lambda : \mathbb{R} \rightarrow \mathbb{R}^{n_x}$, and include the differential equation in the integral according to

$$\max_{x,u,\lambda} \int_{t_0}^{T} f(t,x,u) + \lambda^t(\dot{x} - g(t,x,u))dt \tag{112}$$
$$\text{s.t.} \qquad x(0) = x_0 \,.$$

Since $u$ and $\lambda$ are assumed to be only piecewise continuous, we cannot apply first order conditions of the calculus of variation. Therefore, we introduce continuous and piecewise differentiable functions $U$ and $\Lambda$ according to

$$U(t) \; := \; \int_{t_0}^{t} u(s)ds \qquad U(t_0) = 0 \tag{113}$$

$$\Lambda(t) \; := \; \int_{t_0}^{t} \lambda(s)ds \qquad \Lambda(t_0) = 0 \tag{114}$$

and the problem can be written as

$$\max_{x,U,\Lambda} \int_{t_0}^{T} f(t,x,\dot{U}) + \dot{\Lambda}^t(\dot{x} - g(t,x,\dot{U}))dt \tag{115}$$

---

[63]Corners of a function are points, where the function is continuous, but not differentiable, whereas jumps are points, where the function is not continuous, either.

$$\text{s.t.} \qquad x(0) = x_0, \quad \Lambda(0) = 0, \quad U(0) = 0$$

To simplify notation we introduce $\phi(t, x, U, \Lambda, \dot{x}, \dot{U}, \dot{\lambda}) := f(t, x, \dot{U}) + \dot{\Lambda}^t(\dot{x} - g(t, x, \dot{U}))$, $y := (x, U, \Lambda)^t$, and $y_0 := (x_0, 0, 0)^t$. Then, the problem can be written as

$$\max_{y} \int_{t_0}^{T} \phi(t, y, \dot{y}) dt \tag{116}$$
$$\text{s.t.} \qquad y(0) = y_0, \ .$$

First order conditions of optimal control of problem (111) could be calculated by applying the principles of calculus of variation to problem (116), and retransforming the first order conditions in notation. E.g. applying the Euler-Lagrange equations to problem (116) yields the familiar first order conditions $H_u = 0$, $H_\lambda = \dot{x}$, and $H_x = -\dot{\lambda}$ with the Hamiltonian function $H \equiv f + \lambda^t g$. However, we want to focus on conditions that must be satisfied, if $y$ exhibits a corner, i.e. $x$ exhibits a corner and $u$ and $\lambda$ exhibit jumps. Therefore, we state the theorem of the Weierstraß-Erdmann corner conditions according to Kamien and Schwartz (1991, pp. 79), Bryson and Ho (1975, pp. 125) and Intriligator (1971, pp. 312) and give a proof that is based on the exposition in Kamien and Schwartz (1991, pp. 53-82). A complete and more formal proof can be found in Bryson and Ho (1975, pp. 87 and pp. 101).

**Theorem 14** *(Weierstraß-Erdmann corner conditions for calculus of variation)*
*Consider a problem of calculus of variation according to (116), with continuous and piecewise differentiable function $y$. If the optimal solution $y^*$ exhibits a corner at time $\tilde{t}$, the following conditions must hold in addition to the usual first order conditions of optimality*

$$\phi_{\dot{y}}|_{\tilde{t}^-} = \phi_{\dot{y}}|_{\tilde{t}^+} \tag{117}$$
$$(\phi - (\dot{y})^t \phi_{\dot{y}})|_{\tilde{t}^-} = (\phi - (\dot{y})^t \phi_{\dot{y}})|_{\tilde{t}^+} \tag{118}$$

*We refer to (117) as the first Weierstraß-Erdmann corner condition, and to (118) as the second Weierstraß-Erdmann corner condition.*

**Proof.** We define

$$I(y, \tilde{t}) := \int_{t_0}^{\tilde{t}} \phi(t, y, \dot{y}) dt + \int_{\tilde{t}}^{T} \phi(t, y, \dot{y}) dt \qquad (119)$$

and

$$J(\varepsilon) := I(y^* + \varepsilon v, \tilde{t} + \varepsilon \cdot \delta t) \qquad (120)$$

with the optimal solution $y^*$, $\varepsilon \in \mathbb{R}$, and the time-dependent deviation $v$ of the optimal path. We require $v$ to be differentiable, but since $y^*$ may exhibit a corner at $\tilde{t}$, we allow for corners of $v$ at $\tilde{t}$. In addition, we vary the point of time, at which the corner occurs by $\varepsilon \cdot \delta t$. Note that we have to distinguish between $y^*(t + \varepsilon \cdot \delta t) + v(t + \varepsilon \cdot \delta t)$ as a variation with respect to the interval $[t_0, \tilde{t}]$, or as a variation with respect to the interval $[\tilde{t}, T]$, because $y^*$ may exhibit a corner at $\tilde{t}$. Therefore we indicate the first by $t^-$, and the second by $t^+$. Differentiation $J$ with respect to $\varepsilon$ and inserting $\varepsilon = 0$ yields

$$
\begin{aligned}
\delta I(y^*; v) &= J'(0) \qquad &(121)\\
&= \int_{t_0}^{\tilde{t}} ((\phi_y)^t v + (\phi_{\dot{y}})^t \dot{v}) dt + \phi(\tilde{t}^-, y^*(\tilde{t}^-), \dot{y}^*(\tilde{t}^-))\delta t \qquad &(122)\\
&+ \int_{\tilde{t}}^{T} ((\phi_y)^t v + (\phi_{\dot{y}})^t \dot{v}) dt - \phi(\tilde{t}^+, y^*(\tilde{t}^+), \dot{y}^*(\tilde{t}^+))\delta t \qquad &(123)
\end{aligned}
$$

with the Gâteaux-differential or first variation $\delta I$.[64] We transform the second summand from each integral with partial integration, which yields

$$
\begin{aligned}
\delta I(y^*; v) &= ((\phi_{\dot{y}})^t v)\Big|_{t_0}^{\tilde{t}} + \int_{t_0}^{\tilde{t}} \left( \left( \phi_y - \frac{d}{dt}\phi_{\dot{y}} \right)^t v \right) dt \\
&+ ((\phi_{\dot{y}})^t v)\Big|_{\tilde{t}}^{T} + \int_{\tilde{t}}^{T} \left( \left( \phi_y - \frac{d}{dt}\phi_{\dot{y}} \right)^t v \right) dt \\
&+ \phi(\tilde{t}^-, y^*(\tilde{t}^-), \dot{y}^*(\tilde{t}^-))\delta t - \phi(\tilde{t}^+, y^*(\tilde{t}^+), \dot{y}^*(\tilde{t}^+))\delta t \qquad (124)
\end{aligned}
$$

---

[64]To keep the notation clear we omit arguments of $\phi$ wherever it does not cause confusion.

Extracting and rearranging yields

$$
\begin{aligned}
\delta I(y^*; v) \;=\;& \int_{t_0}^{\tilde{t}} \left( \left( \phi_y - \frac{d}{dt}\phi_{\dot{y}} \right)^t v \right) dt + \int_{\tilde{t}}^{T} \left( \left( \phi_y - \frac{d}{dt}\phi_{\dot{y}} \right)^t v \right) dt \\
&+ \phi_{\dot{y}}(T, y^*(T), y^*(T))^t v(T) - \phi_{\dot{y}}(t_0, y^*(t_0), y^*(t_0))^t v(t_0) \\
&+ \phi_{\dot{y}}(\tilde{t}^-, y^*(\tilde{t}^-), y^*(\tilde{t}^-))^t v(\tilde{t}^-) - \phi_{\dot{y}}(\tilde{t}^+, y^*(\tilde{t}^+), y^*(\tilde{t}^+))^t v(\tilde{t}^+) \\
&+ \phi(\tilde{t}^-, y^*(\tilde{t}^-), \dot{y}^*(\tilde{t}^-))\delta t - \phi(\tilde{t}^+, y^*(\tilde{t}^+), \dot{y}^*(\tilde{t}^+))\delta t \qquad (125)
\end{aligned}
$$

It has to be considered that the function $y \equiv y^* + v$ changes at $\tilde{t}$ if $v$ changes or if $\tilde{t}$ changes. Therefore, we approximate the variation $\delta y(\tilde{t})$ by linearization according to

$$
\delta y(\tilde{t}) \approx y(\tilde{t}) - y^*(\tilde{t}) + \dot{y}^*(\tilde{t}) \cdot \delta t = v(\tilde{t}) + \dot{y}^*(\tilde{t}) \cdot \delta t
$$

We substitute for $v(\tilde{t})$ in equation (125) and get by rearranging terms

$$
\begin{aligned}
\delta I(y^*; v) \;=\;& \int_{t_0}^{\tilde{t}} \left( \left( \phi_y - \frac{d}{dt}\phi_{\dot{y}} \right)^t v \right) dt + \int_{\tilde{t}}^{T} \left( \left( \phi_y - \frac{d}{dt}\phi_{\dot{y}} \right)^t v \right) dt \\
&+ \phi_{\dot{y}}(T, y^*(T), y^*(T)) v(T) - \phi_{\dot{y}}(t_0, y^*(t_0), y^*(t_0)) v(t_0) \\
&+ \phi_{\dot{y}}(\tilde{t}^-, y^*(\tilde{t}^-), y^*(\tilde{t}^-))^t \delta y - \phi_{\dot{y}}(\tilde{t}^+, y^*(\tilde{t}^+), y^*(\tilde{t}^+))^t \delta y \\
&+ [\phi(\tilde{t}^-, y^*(\tilde{t}^-), \dot{y}^*(\tilde{t}^-)) - \dot{y}^*(\tilde{t})^t \phi_{\dot{y}}(\tilde{t}^-, y^*(\tilde{t}^-), y^*(\tilde{t}^-))]\delta t \\
&- [\phi(\tilde{t}^+, y^*(\tilde{t}^+), \dot{y}^*(\tilde{t}^+)) - \dot{y}^*(\tilde{t})^t \phi_{\dot{y}}(\tilde{t}^+, y^*(\tilde{t}^+), y^*(\tilde{t}^+))]\delta t \qquad (126)
\end{aligned}
$$

For an optimal solution $y^*$, $\delta I = 0$ must hold. In a first stage, we allow for functions $v$ with $v(t_0) = 0$, $\delta t = 0$ and $v \equiv 0$ on $[\tilde{t}, T]$. Then, equation (126) reduces to

$$
\delta I(y^*; v) \;=\; \int_{t_0}^{\tilde{t}} \left( \left( \phi_y - \frac{d}{dt}\phi_{\dot{y}} \right)^t v \right) dt \qquad (127)
$$

Since $\delta I = 0$ has to hold for every possible deviation $v$, we get as a necessary condition for optimal solutions the Euler-Lagrange equation

$$
\phi_y - \frac{d}{dt}\phi_{\dot{y}} = 0 \qquad (128)
$$

for the interval $[t_0, \tilde{t})$. Analogously, the Euler-Lagrange equation has to hold at $(\tilde{t}, T]$. Moreover, by allowing for functions $v$ that satisfy $v(\tilde{t}) = 0$ and restricting $\delta\tilde{t} = 0$, proper boundary conditions for $t_0$ and $T$ can be derived. We focus on $\tilde{t}$ and, therefore, assume that equation (126) has reduced according to the considerations above to

$$
\begin{aligned}
0 \;=\; & \phi_{\dot{y}}(\tilde{t}^-, y^*(\tilde{t}^-), y^*(\tilde{t}^-))^t \delta y - \phi_{\dot{y}}(\tilde{t}^+, y^*(\tilde{t}^+), y^*(\tilde{t}^+))^t \delta y \\
& + [\phi(\tilde{t}^-, y^*(\tilde{t}^-), \dot{y}^*(\tilde{t}^-)) - \dot{y}^*(\tilde{t})^t \phi_{\dot{y}}(\tilde{t}^-, y^*(\tilde{t}^-), y^*(\tilde{t}^-))]\delta t \\
& - [\phi(\tilde{t}^+, y^*(\tilde{t}^+), \dot{y}^*(\tilde{t}^+)) - \dot{y}^*(\tilde{t})^t \phi_{\dot{y}}(\tilde{t}^+, y^*(\tilde{t}^+), y^*(\tilde{t}^+))]\delta t \,.
\end{aligned}
\tag{129}
$$

Allowing only for variations satisfying $\delta t = 0$ in a first stage yields equation (117), and allowing only for variations satisfying $\delta y = 0$ in a second stage yields equation (118). ∎

The corner conditions can be extended to any finite number of corners. However, solutions with an infinite number of corners are not admissible. The next step is to retransform equations (117) and (118) into notation of optimal control. This yields the following theorem

**Theorem 15** *(Weierstraß-Erdmann corner conditions for problems of Optimal Control)*

*Consider a problem of Optimal Control according to (111), with continuous and piecewise differentiable function $x$, and piecewise continuous functions $\lambda$ and $u$. We define the associated Hamilton function as $H \equiv f + \lambda^t g$. If the optimal solution $x$ exhibits a corner at time $\tilde{t}$, and/or the optimal solution $u$ exhibits a jump at time $\tilde{t}$ the following conditions must hold in addition to the usual first order conditions of optimality*

$$
\lambda(t)|_{\tilde{t}^-} = \lambda(t)|_{\tilde{t}^+} \tag{130}
$$

$$
H(t)|_{\tilde{t}^-} = H(t)|_{\tilde{t}^+} \tag{131}
$$

*We refer to (130) as the first Weierstraß-Erdmann corner condition of optimal control, and to (131) as the second Weierstraß-Erdmann corner condition of optimal control.*

**Proof.** We apply theorem (14) to the problem of optimal control. Evaluating the first Weierstraß-Erdmann corner condition for $x$, $U$, and $\Lambda$, respectively, yields

$$\lambda|_{\tilde{t}^-} = \lambda|_{\tilde{t}^+} \tag{132}$$

$$H_u|_{\tilde{t}^-} = H_u|_{\tilde{t}^+} \tag{133}$$

$$(\dot{x} - g(t, x, u))|_{\tilde{t}^-} = (\dot{x} - g(t, x, u))|_{\tilde{t}^+} \tag{134}$$

The first equation already represents the first Weierstraß-Erdmann corner condition of optimal control. The second and third equation do not yield additional information, since $H_u = 0$ and $\dot{x} = g(t, x, u)$ are already first order conditions. Evaluating the second Weierstraß-Erdmann corner condition yields, omitting the arguments

$$
\begin{aligned}
& \left[ f + \lambda^t(\dot{x} - g) - (\dot{x}^t\lambda + u^t(f_u - \lambda^t g_u) + \lambda^t\dot{x} - \lambda^t g) \right]\Big|_{\tilde{t}^-} \\
= & \left[ f + \lambda^t(\dot{x} - g) - (\dot{x}^t\lambda + u^t(f_u - \lambda^t g_u) + \lambda^t\dot{x} - \lambda^t g) \right]\Big|_{\tilde{t}^+}
\end{aligned}
$$

and exploiting that $H_u = 0$, yields the second Weierstraß-Erdmann corner condition of optimal control. ∎

Consider a continuous change in a parameter or a variable exogenous to the agent. In problem (111) these parameters and variables are caught by the time arguments of $f$ and $g$. That means a change in a parameter or exogenous variable along time causes a change of $f$ and $g$ along time. Due to perfect foresight the agent is aware of this time dependence.[65]  Then, an anticipated shock in terms of a jump in a parameter after $t_0$ yields discontinuities in $f$ and $g$.

---

[65]However, the agent is not aware that potentially his decisions influence variables exogenous to him.

For the analysis of an expected shock we follow Bryson and Ho (1975, pp. 101). First of all, the set of differential equations switches at, for the time being, an unspecified point of time $\tilde{t}$,

$$
\begin{aligned}
\dot{x} &= g^{(1)}(t, x, u) \qquad t < \tilde{t} \\
\dot{x} &= g^{(2)}(t, x, u) \qquad t \geq \tilde{t}\,.
\end{aligned}
$$

In addition, $f$ changes its functional form at $\tilde{t}$ from $f^{(1)}$ to $f^{(2)}$. We fix the point of time where the jump in the parameter occurs with an interior boundary condition:

$$
\psi[\tilde{t}] = \tilde{t} - t_{jump} = 0
$$

Now, the transformed optimization problem is

$$
\max_{u,\tilde{t}} \left[ \int_{t_0}^{\tilde{t}} f^{(1)}(t, x, u)dt + \int_{\tilde{t}}^{T} f^{(2)}(t, x, u)dt \right] \tag{135}
$$

$$
\text{s.t.} \qquad \dot{x} = g^{(1)}(t, x, u) \qquad t < \tilde{t} \tag{136}
$$

$$
\dot{x} = g^{(2)}(t, x, u) \qquad t \geq \tilde{t}
$$

$$
x(0) = x_0 \tag{137}
$$

$$
\tilde{t} - t_{jump} = 0 \tag{138}
$$

with $t_{jump}$ denoting the (numerical) time value where the jump occurs. Note, that now $\tilde{t}$ has become part of the optimization problem. First order conditions specifying the behavior of $x$ and $u$ are stated in the following theorem

**Theorem 16** *(corner conditions for problems of optimal control with anticipated parameter changes)*
*Consider problem (135) - (138). We define the associated Hamiltonian function piecewise for the intervals $[0, \tilde{t})$ and $[\tilde{t}, T]$ as $H \equiv f + \lambda^T g$. First order conditions*

*for an optimal solution are*

$$H_u = 0 \tag{139}$$

$$H_\lambda = \dot{x} \tag{140}$$

$$H_x = -\dot{\lambda} \tag{141}$$

*and initial boundary condition (137) and final boundary condition $\lambda(T) = 0$ for $T\epsilon\mathbb{R}$,*

*or*

$$\lim_{t \to \infty} \lambda(t)x(t) = 0 \tag{142}$$

*for $T = \infty$.[66] Since $g$ and $f$ are defined piecewise before and after $\tilde{t}$, these conditions have to be applied piecewise. Moreover, at $\tilde{t}$ the following conditions have to hold*

$$\lambda(\tilde{t}^-) = \lambda(\tilde{t}^+) \tag{143}$$

$$H(\tilde{t}^-) = H(\tilde{t}^+) - \nu \tag{144}$$

*whereas $\nu$ is a constant Lagrangian multiplier associated with the interior boundary condition.*

**Proof.**  Again, we transform the problem of optimal control into a problem of the calculus of variation. We append the internal boundary condition with a constant multiplier $\nu$. Then, the transformed problem reads

$$\max_y \int_{t_0}^T \phi(t, y, \dot{y}) + \nu(\tilde{t} - t_{jump})dt \tag{145}$$
$$\text{s.t.} \qquad y(0) = y_0,$$

The proof is now analogous to the proof of theorem (14). We define

$$I(y, \tilde{t}) := \int_{t_0}^{\tilde{t}} \phi(t, y, \dot{y}) + \nu_1(\tilde{t} - t_{jump})dt + \int_{\tilde{t}}^T \phi(t, y, \dot{y}) + \nu_2(\tilde{t} - t_{jump})dt \tag{146}$$

---

[66]There is disagreement if this equation is indeed a necessary condition (see e.g. Chiang, 1992). We do not want to adress this point.

and

$$J(\varepsilon) := I(y^* + \varepsilon v, \tilde{t} + \varepsilon \cdot \delta t) \tag{147}$$

with the optimal solution $y^*$, $\varepsilon \epsilon \mathbb{R}$, and the time-dependent deviation $v$ of the optimal path. For differentiating $J$ with respect to $\epsilon$ we have to consider the internal boundary condition. Therefore, we get

$$
\begin{aligned}
\delta I(y^*; v) &= J'(0) \\
&= \int_{t_0}^{\tilde{t}} ((\phi_y)^t v + (\phi_{\dot{y}})^t \dot{v}) dt + \phi(\tilde{t}^-, y^*(\tilde{t}^-), \dot{y}^*(\tilde{t}^-)) \delta t \\
&\quad + \int_{\tilde{t}}^{T} ((\phi_y)^t v + (\phi_{\dot{y}})^t \dot{v}) dt - \phi(\tilde{t}^+, y^*(\tilde{t}^+), \dot{y}^*(\tilde{t}^+)) \delta t \\
&\quad + \nu_1 \delta t + \nu_2 \delta t
\end{aligned}
$$

Following the proof, equation (126) has to be extended according to

$$
\begin{aligned}
\delta I(y^*; v) &= \int_{t_0}^{\tilde{t}} \left( \left( \phi_y - \frac{d}{dt} \phi_{\dot{y}} \right)^t v \right) dt + \int_{\tilde{t}}^{T} \left( \left( \phi_y - \frac{d}{dt} \phi_{\dot{y}} \right)^t v \right) dt \\
&\quad + \phi_{\dot{y}}(T, y^*(T), y^*(T)) v(T) - \phi_{\dot{y}}(t_0, y^*(t_0), y^*(t_0)) v(t_0) \\
&\quad + \phi_{\dot{y}}(\tilde{t}^-, y^*(\tilde{t}^-), y^*(\tilde{t}^-))^t \delta y - \phi_{\dot{y}}(\tilde{t}^+, y^*(\tilde{t}^+), y^*(\tilde{t}^+))^t \delta y \\
&\quad + [\phi(\tilde{t}^-, y^*(\tilde{t}^-), \dot{y}^*(\tilde{t}^-)) - \dot{y}^*(\tilde{t})^t \phi_{\dot{y}}(\tilde{t}^-, y^*(\tilde{t}^-), y^*(\tilde{t}^-))] \delta t \\
&\quad - [\phi(\tilde{t}^+, y^*(\tilde{t}^+), \dot{y}^*(\tilde{t}^+)) - \dot{y}^*(\tilde{t})^t \phi_{\dot{y}}(\tilde{t}^+, y^*(\tilde{t}^+), y^*(\tilde{t}^+))] \delta t \\
&\quad + \nu_1 \delta t + \nu_2 \delta t \tag{148}
\end{aligned}
$$

After exploiting the Euler-Lagrange equations and proper boundary conditions for $t_0$ and $T$ we get

$$\phi_{\dot{y}}|_{\tilde{t}^-} = \phi_{\dot{y}}|_{\tilde{t}^+} \tag{149}$$

by setting $\delta t = 0$. These equations yield the first Weierstraß-Erdmann corner condition of optimal control. Different to the previous proof we get the equation

$$0 = +[\phi(\tilde{t}^-, y^*(\tilde{t}^-), \dot{y}^*(\tilde{t}^-)) - \dot{y}^*(\tilde{t})^t \phi_{\dot{y}}(\tilde{t}^-, y^*(\tilde{t}^-), y^*(\tilde{t}^-))] \delta t$$

$$-[\phi(\tilde{t}^+, y^*(\tilde{t}^+), \dot{y}^*(\tilde{t}^+)) - \dot{y}^*(\tilde{t})^t\phi_{\dot{y}}(\tilde{t}^+, y^*(\tilde{t}^+), y^*(\tilde{t}^+))]\delta t$$

$$+\nu_1\delta t + \nu_2\delta t \tag{150}$$

by seting $\delta y = 0$. Now, optimal solutions have to fulfill

$$(\phi - (\dot{y})^t\phi_{\dot{y}})|_{\tilde{t}^-} = (\phi - (\dot{y})^t\phi_{\dot{y}})|_{\tilde{t}^+} + \nu$$

with $\nu := \nu_1 + \nu_2$. Therefore, the second Weierstraß-Erdmann corner condition does not hold. Since we derived the continuity of the Hamilton function of optimal control from this condition, the second Weierstraß-Erdmann corner condition of optimal control now turns into

$$H(t)|_{\tilde{t}^-} = H(t)|_{\tilde{t}^+} + \nu$$

∎

Theorem (16) can be derived from the general interior boundary condition as described in Bryson and Ho (1975, pp. 101).[67] To summarize, theorem (16) states that $\lambda$ is continuous, whereas $H$ jumps at the particular point of time when the jump in the parameter takes place. Therefore, the first Weierstraß-Erdmann corner condition of optimal control holds, whereas the second does not. The reason is that the Weierstraß-Erdmann corner condition are stated for possible corners at an unknown point of time $\tilde{t}$. We, however, assumed the anticipated parameter change to take place at a predetermined point of time. Reducing the degrees of freedom for the solution by one, renders ones condition unnecessary. Note that the first Weierstraß-Erdmann corner condition is $n_x$ dimensional, while the second is one dimensional.

Equation (144) is the only equation where the shadow price $\nu$ appears. It equals the jump in the Hamiltonian at time $\tilde{t}$. Therefore, the set of equations and variables can be solved without equation (144), after which (144) trivially determines $\nu$.

---

[67]Bryson and Ho (1975) explicitly address discontinuities in $g$ but not in $f$. However, their proof also allows for discontinuities in $f$.

The problem at hand comprises $2 \cdot n_x$ differential equations and $n_u$ algebraic equations, together with $n_x$ initial conditions and $n_x$ transversality conditions as well as $2 \cdot n_x$ interior boundary conditions. The $2 \cdot n_x$ interior boundary conditions comprise the requirement of continuity for $n_x$ state and $n_x$ adjoint variables. Note that the local dynamic behavior of the system can implicitly be described by $2 \cdot n_x$ differential equations in coordinates $(\lambda, x)$ if $H_u$ exhibits no singularity with respect to $u$.[68] For the system (111) to possess a unique solution it has to exhibit a $n_x$ dimensional stable manifold around the stationary equilibrium or the stationary equilibria.[69] Applying Bellman's principle, we know that at time $\tilde{t}$ the solution has to be on this stable manifold. This results in a well defined two point boundary value problem for the time interval $[t_0, \tilde{t}]$ and for the time interval $[\tilde{t}, \infty)$. For both intervals, the $2 \cdot n_x$ differential equations together with $n_u$ algebraic equations have to hold. The $n_x$ initial boundary conditions for the first two point boundary value problem are (137) and the $n_x$ final boundary conditions are represented by the requirement to be on the $n_x$ dimensional stable manifold at $\tilde{t}$. For the second two point boundary value problem $n_x$ initial boundary conditions are the requirement to start on the stable manifold, whereas $n_x$ final boundary conditions are represented by the transversality conditions (142). Since initial and final boundary conditions of both problems are linked, this is labeled as a three point boundary value problem in the mathematic literature.

Statements about necessary or sufficient conditions for existence and uniqueness of solutions of two-point-boundary value problems are in no case as advanced as the theory of initial value problems (see Ascher and Petzold, 1998, Ch. III, 6). If the

---

[68]This follows from the implicit function theorem. We assume the differential algebraic system to be of differential index one. Higher order differential algebraic systems exhibit a far more complex behavior (e.g. Ascher and Petzold, 1998, Ch. IV, 9). If $\frac{\partial H_u}{\partial u}$ exhibits full rank, the system is of index one.

[69]Some economic models exhibit a center manifold of stationary equilibria (e.g. Lucas, 1988), as previously described in this thesis.

linearized system exhibits a unique solution it is at least possible to conclude that the original system exhibits a unique solution locally. However, it is beyond the scope of this dissertation to prove uniqueness of the linearized system. Necessary conditions for the existence of a unique solution comprise the dimension of the boundary conditions to add up to the dimension of the dynamic system. This is fulfilled here for both intervals, $[t_0, \tilde{t}]$ and $[\tilde{t}, \infty)$.

If a finite sequence of expected shocks is given, the same arguments as above can be applied. The resulting problem can be decomposed in a sequence of two point boundary value problems, each of which exhibits $n_x$ initial and $n_x$ final boundary conditions. Then the problem turns into a multi point boundary value problem.

## 4.3  Numerical implementation

There exists a rich economic literature how to solve infinite horizon maximization problems numerically (see Section 3.3 of this thesis, Judd, 1998, or Brunner and Strulik, 2002, for a survey). Many approaches employ the first order conditions derived by Pontryagin's maximum principle for finding the solution, together with the initial conditions and transversality conditions (e.g. Judd, 1992, Brunner and Strulik, 2002, Mulligan and Sala-i-Martin, 1991 or Section 3 of this thesis).[70] Then, the approach comprises the solution of a two point boundary value problem (e.g. Judd, 1992, and Section 3 of this thesis), or the problem is transformed into a (stable) initial value problem (e.g. Brunner and Strulik, 2002 and Mulligan and Sala-i-Martin, 1991). However, the simulation of anticipated shocks transforms the two point boundary value problem into a multi point boundary value problem. Therefore, it is not straightforward to employ the existing solution algorithms. In the former case, the algorithms at hand cannot incorporate arbitrary internal boundary conditions with-

---

[70]Exceptions are e.g. Mercenier and Michel (1994 and 2001), and Candler (1999).

out undergoing fundamental modifications. In the latter case, it is no longer possible
to transform the problem into a single stable initial value problem.

In the mathematic literature it is suggested to reformulate the multi point bound-
ary value problem into a two point boundary value problem to make it accessible for
standard algorithms (see Ascher et al., 1985, and Ascher and Russell, 1981). How-
ever, we suggest to solve the multi point boundary value problem directly, which
is straightforward if the relaxation algorithm as proposed in this thesis is applied.
The advantage is that this proceeding is easy and intuitive. The disadvantage is
merely that the efficiency in terms of calculation requirement of the algorithm re-
duces slightly.

The special property of the multi point boundary value problem at hand is that
the internal boundary conditions are fulfilled, if the variables $\lambda$ and $x$ are forced
to be continuous along the solution whereas no restrictions regarding $u$ are made.
The relaxation algorithm perfectly exploits this fact, since the algorithm exactly de-
mands this property from differential and algebraic variables.[71] More precisely, the
principle of relaxation is to replace the differential equations by approximate finite
difference equations on a mesh of points in time. The residuum of the algebraic
equations is minimized at every mesh point separately. Therefore, the relaxation
algorithm treats differential and algebraic variables conceptually different. Whereas
no connection along time is made for algebraic variables, differential variables are
connected along time through the difference equations. Therefore, algebraic vari-
ables can exhibit jumps, whereas differential variables have to be continuous but
potentially can exhibit corners. In case the solution indeed exhibits corners along
the time path of the differential variables, the employed discretization rule decreases

---

[71]We refer to differential variables as variables for which differential equations are present (i.e. $\lambda$
and $x$) whereas the time derivative of algebraic variables does not appear in the set of equations
(i.e. $u$).

from second to first order. This means that by increasing the number of mesh points by factor $x$ the global error reduces by $x$, while in case the solution does not exhibit corners the global error reduces by $x^2$.

In the economic literature, very often the set of algebraic equations is differentiated with respect to time and the adjoint variables $\lambda$ are eliminated (see e.g. the analysis of the Lucas (1988) model by Caballe and Santos (1993) or Benhabib and Perli (1994)). In many cases, this eases economic interpretation, as well as numerical computation. Many algorithms cannot solve differential algebraic systems, and by eliminating adjoint variables the dimensionality of the problem is reduced. However, this procedure is not advisable here, since differential equations for the control variables can only be applied piecewise. Since jumps can occur at internal points it would then be necessary to specify the height of each jump. This is, however, not possible without information about the adjoint variables. Therefore we suggest to simulate system (139), (140), and (141) together with the boundary conditions and ensure that $x$ and $\lambda$ are continuous.

So far, we assumed the differential algebraic system to be scaled, i.e. the variables approach constant (finite) values in the long-run. This may not be the case, for example if the model at hand is an endogenous growth model. For numerical computation the variables have to be scaled such that they approach constants in the long-run. If ratios are created it has to be ensured that the time path of the differential variables remains continuous. That is, only combinations of different $x$ and $\lambda$ can be chosen for creating artificial differential variables. By contrast, arbitrary ratios can be created as algebraic variables. We recommend the scale adjustment as proposed by Lucas (1988), since continuous variables conserve their properties after scaling.

To summarize, the relaxation algorithm can solve infinite horizon problems ex-

hibiting anticipated shocks conveniently, since it solves the multi point boundary value problem directly. This is done by implementing the continuous variables as differential variables and the possible jumping variables as algebraic ones. The only input the user has to provide are the time dependent parameter values.[72]

## 4.4 A simple example

### 4.4.1 Description of the Model

To illustrate the numerical solution of infinite-time optimization models with expected shocks, we employ the Ramsey-Cass-Koopmans model (Ramsey, 1928; Cass, 1965; Koopmans, 1965) as an example. For reasons of clarity, we keep the model as simple as possible. We omit technological progress and assume the population to grow with constant rate $n$. We follow Barro and Sala-i-Martin (2004, Chapter 3) and introduce proportional taxes on wage income, $\tau_w$, private asset income, $\tau_r$, and consumption, $\tau_c$. Thus, the representative household's maximization problem is

$$\max_c \int_0^\infty \frac{c^{1-\sigma}-1}{1-\sigma} e^{(n-\rho)t} dt \tag{151}$$
$$\text{s.t.} \quad \dot{k} = (1-\tau_w)w + (1-\tau_r)rk - (1+\tau_c)c - nk, \quad k(0) = k_0$$

whereas $c$ denotes consumption per capita, $k$ the capital stock per capita, $w$ the wage rate, $r$ the interest rate, $\sigma$ the inverse of intertemporal elasticity of substitution, and $\rho$ the discount factor, respectively.

The government is assumed to run a balanced budget. Therefore, government revenues equal total outlays. However, we assume that government spending appears nowhere else in the economy.[73] Firms produce according to a Cobb-Douglas

---

[72]Using the supplemented software this can be done by stating an *if*-clause.

[73]Alternatively, it could be assumed that government spending increases consumer utility, whereas household exhibit an additively separable utility function, or government revenues are spend on transfers to households. Since both alternative assumptions do not contribute anything to the topic discussed in this Section we chose the simplest assumption.

production function

$$Y = K^\alpha L^{1-\alpha}$$

whereas $Y$ denotes the output, $K$ the capital stock, $L$ the amount of labor employed in production, and $\alpha$ the elasticity of capital in final-output production, respectively.

Since perfect competition in factor markets is assumed, firms pay the factors according to their marginal product,

$$r = \alpha k^{\alpha-1} - \delta$$
$$w = (1-\alpha)k^\alpha .$$

Turning back to the representative household optimization problem we now consider an anticipated change in the tax rates at time $\tilde{t} > t_0$. This means that at $\tilde{t}$ the consumer's budget constraint changes its functional form and potentially exhibits a jump on the right hand side,

$$\dot{k} = \begin{cases} (1-\tau_{w,1})w + (1-\tau_{r,1})rk - (1+\tau_{c,1})c - nk & t < \tilde{t} \\ (1-\tau_{w,2})w + (1-\tau_{r,2})rk - (1+\tau_{c,2})c - nk & t \geq \tilde{t} . \end{cases}$$

For simplicity, we do not distinguish between $\tau_{i,1}$ and $\tau_{i,2}$ below, but denote the tax rates with a time index to indicate that different values have to be applied before and after $\tilde{t}$. Then, the Hamiltonian is[74]

$$H = \frac{c^{1-\sigma} - 1}{1 - \sigma} + \lambda((1-\tau_{w,t})w + (1-\tau_{r,t})rk - (1+\tau_{c,t})c - nk) . \qquad (152)$$

Necessary conditions for an optimal solution are

$$c^{-\sigma} = \lambda(1+\tau_{c,t}) \qquad (153)$$
$$\dot{\lambda} = (\rho - n)\lambda - \lambda((1-\tau_{r,t})r - n) \qquad (154)$$
$$\dot{k} = (1-\tau_{w,t})w + (1-\tau_{r,t})rk - (1+\tau_{c,t})c - nk \qquad (155)$$

[74]In contrast to the general derivation above we use the current-value Hamiltonian here, since then the shadow price $\lambda$ is already a stationary variable. This does not affect conclusions about continuity, since $H$ and $\lambda$ only differ by the factor $e^{(n-\rho)t}$.

together with

$$\lambda(\tilde{t}^-) = \lambda(\tilde{t}^+) \tag{156}$$

$$H(\tilde{t}^-) = H(\tilde{t}^+) - \nu \,. \tag{157}$$

From the firm's sector we additionally obtain two algebraic equations for the factor prices, which we can substitute into the capital accumulation equation. Note, that for the derivation of the familiar Euler-Equation $\frac{\dot{c}}{c} = \frac{(1-\tau_{r,t})r-\rho}{\sigma}$ the time derivative of $c$ must be taken. However, $c$ is not necessarily differentiable at time $\tilde{t}$. Therefore, the Euler equation can only be applied piecewise.

For numerical computation of expected shocks we exploit the information that $k$ and $\lambda$ have to be continuous. Therefore, we implement them in the form of differential equations whereas we introduce $c$ as an algebraic variable. To summarize, the differential algebraic system is

$$\dot{k} = (1-\tau_{w,t})(1-\alpha)k^\alpha + (1-\tau_{r,t})(\alpha k^\alpha - \delta k) - (1+\tau_{c,t})c - nk \tag{158}$$

$$\dot{\lambda} = \lambda(\rho - (1-\tau_{r,t})r) \tag{159}$$

$$c^{-\sigma} = \lambda(1+\tau_{c,t}) \,. \tag{160}$$

### 4.4.2  Simulation of expected shocks

Equation (160) displays that consumption will exhibit an interior jump if and only if the consumption tax $\tau_c$ evolves discontinuously, since $\lambda$ must be continuous at an expected shock. As an example, we will focus on an expected increase of $\tau_c$ from 10% to 20% at time $\tilde{t} = 20$. We assume the economy to be in steady state prior to the shock with a tax rate $\tau_c = 10\%$. At time zero the household experience that at time $\tilde{t}$ the consumption tax will increase. Therefore, it will re-optimize its consumption plan such that consumption potentially jumps immediately as well as at time $\tilde{t}$. We conduct three simulations with the inverse of intertemporal elasticity of substitution,

$\sigma$, equal to $\frac{1}{2}$, 1, and 2, respectively.[75] For the simulation no a priori information about the time path of the variables or the shape of the flow was given. We choose, as an initial guess, all variables to be constant at their steady state values. This always leads to quick convergence.

Note, that the steady state value of $k$ is not affected by a change in $\tau_c$ while the steady state value of $c$ is reduced by $\frac{1+\tau_c}{1+\tau_c+\Delta\tau_c}$. If the tax on consumption increases, households consume less in the long-run while spending the same amount for consumption.[76] Therefore, it is straightforward to analyze an unexpected, immediate shock at $t_0$. In this case, consumption would jump to its lower, new steady state value at $t_0$ without any dynamics in capital. This analysis holds for any feasible set of parameters.

For the case of an anticipated shock we can reason that consumption jumps down at the preannounced time of the tax increase, $\tilde{t}$. It is not optimal for the households to smooth consumption such that this jump vanishes, which again can be seen from equation (160).

If $\sigma$ is high, i.e. $\sigma = 2$, consumers have a strong preference for smoothing consumption over time. Figure 16, (i), (ii), and (iii) display the time path of $\lambda$, $c$, and $k$, respectively, referring to this simulation. Variables are normalized to unity at the new steady state, whereas green crosses designate old steady state values. Households try to soften the drop in consumption at $\tilde{t}$. They do so by abandoning consumption in the period between $t_0$ and $\tilde{t}$. Figure 16, (ii) shows that consumption drops down immediately at $t_0$ and is decreasing in the subsequent period until $\tilde{t}$. Then, it jumps down again and is approaching the new steady state from above. Households increase savings until $\tilde{t}$, and dissave after $\tilde{t}$, which can be seen in Figure

---

[75]The remaining parameters and tax rates are set to $\alpha = .3$, $\delta = .03$, $\rho = .02$, $n = .01$, $\tau_w = .4$, and $\tau_k = .3$, respectively.
[76]Households spend $c(1 + \tau_c)$ for consuming the amount $c$.

16, (iii). Figure 16, (iv) shows the $(c, k)$-phase diagram. The economy starts initially at the green cross and moves along the blue line to the new steady state indicated by the red cross. Dotted lines designate jumps, whereas solid lines designate continuous dynamics along time. Note that $\lambda$ is continuously differentiable whereas $k$ is continuous but experiences a corner at $\tilde{t}$ (Figure 16, (i) and (iii)).
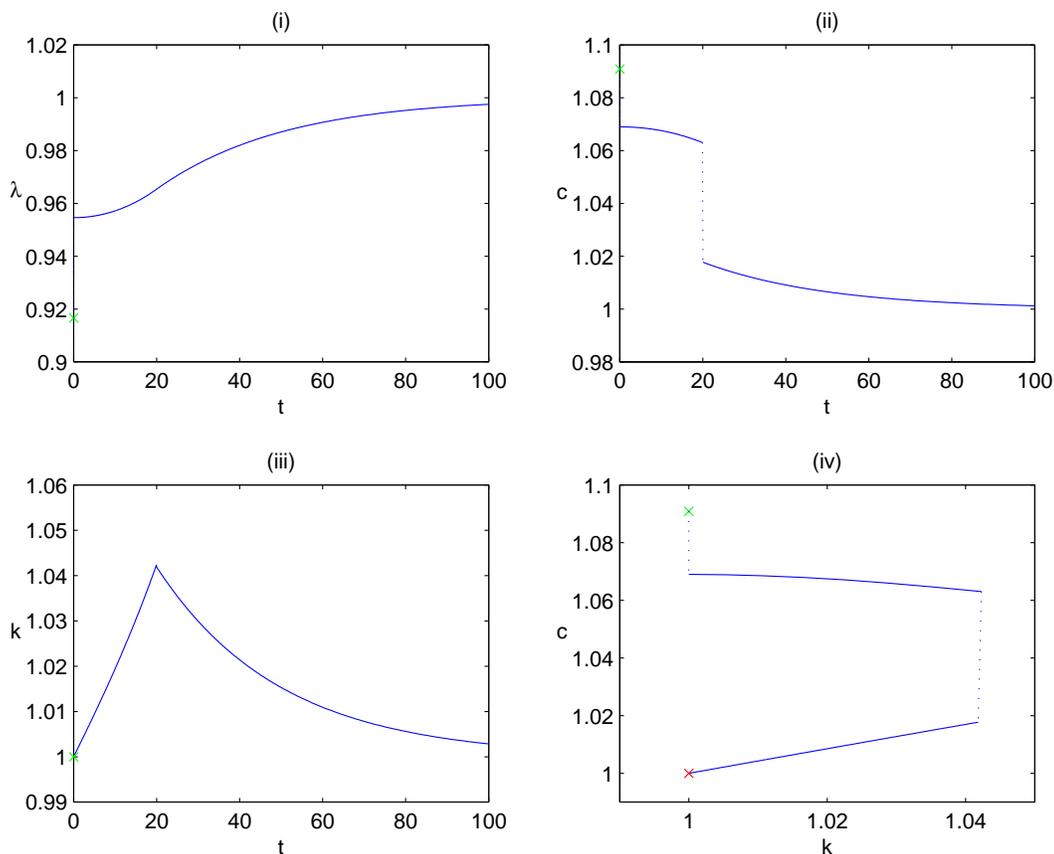


Figure 16: Anticipated increase in $\tau_c$ with $\sigma = 2$

Conversely, if $\sigma$ is low, i.e. $\sigma = 0.5$, households do not put much emphasis on consumption smoothing. They are willing to accept sharp kinks in consumption if this on the other hand yields initially higher consumption. Figure 17, (i), (ii), and (iii) display the time path of $\lambda$, $c$, and $k$, respectively, referring to the sec-

ond simulation. Again, variables are normalized to unity at the new steady state, whereas green crosses designate old steady state values. In Figure 17, (ii) it can be seen that consumption jumps up at $t_0$ and increases in the subsequent period until $\tilde{t}$. Then, consumption jumps down and is approaching the new steady state from below. Households, therefore, exploit the fact that for the same amount spent for consumption they receive a higher consumption before $\tilde{t}$ than after. They do so by dissaving before $\tilde{t}$ (Figure 17, (iii)) and regain their former level of assets by saving after $\tilde{t}$. Figure 17, (iv) shows the $(c, k)$-phase diagram. The economy starts initially at the green cross and moves along the blue line to the new steady state indicated by the red cross. Again, dotted lines designate jumps, whereas solid lines designate continuous dynamics along time. As in the first simulation, $\lambda$ is continuously differentiable whereas $k$ is continuous but experiences a corner at $\tilde{t}$ (Figure 17, (i) and (iii)).

For $\sigma = 1$, which implies a logarithmic utility function, a third pattern emerges. Then, consumption does not jump at $t_0$ and remains in its old steady state value. At $\tilde{t}$ consumption jumps down by $\frac{1+\tau_c}{1+\tau_c+\Delta\tau_c}$ and the economy is immediately in its new steady state.[77] The model's behavior is as if households would experience an unexpected shock, since the above mentioned counteracting effects cancel each other.

### 4.4.3  Verification of simulation results

For deriving the error estimation equation (44), we assumed the time path of variables to be at least three times continuously differentiable. However, this assumption does not hold for the simulations at hand, wherefor we cannot estimate the error by equation (44). For verifying the simulation results, we exploit that for a smooth variation of parameter values theorem (16) does not have to be applied. Therefore,

---

[77]We do not show simulation results for this parameter setting since all variables are constant, despite from very small numerical errors.
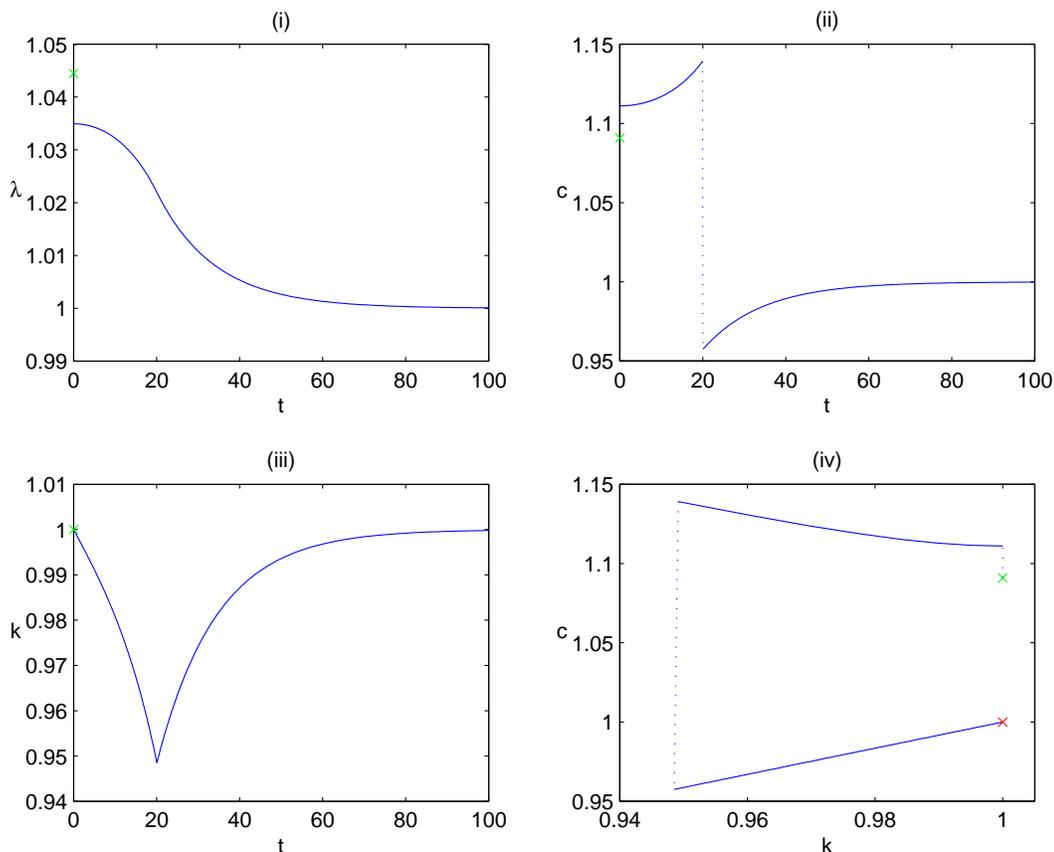
Figure 17: Anticipated increase in $\tau_c$ with $\sigma = .5$

if $\tau_c$ is a three times differentiable function of time, $c$, $\lambda$ and $k$ are also at least three times continuously differentiable with respect to time. Hence, we can estimate the error by error estimation equation (44). The idea is to choose a sufficiently smooth parameterized time path for $\tau_c$ on an interval $[t_{min}, t_{max}]$ with $t_{min} < \tilde{t} < t_{max}$. Then, changing the parameter such that it approaches infinity should make the time path of $\tau_c$ converge towards the discontinuous step function assumed in the previous paragraph. Selecting a sequence of parameters converging towards infinity should result in a sequence of simulation results converging towards simulation results presented in the previous Section. The crucial point is that we can control the error

for each simulation of this sequence, unless $\tau_c$ follows indeed a step function.

Since $\tilde{t} = 20$, we select an interval of $[10, 30]$, and $\tau_c$ to equal

$$\tau_c(t; \zeta) = \begin{cases} 0.1 & t \leq 10 \\ 0.15 + 0.05 \cdot \tanh\left(\zeta \frac{\frac{t-20}{10}}{1 - \left(\frac{t-20}{10}\right)^2}\right) & 10 < t < 30 \\ 0.2 & t \geq 30 \end{cases} \tag{161}$$

Clearly, $\tau_c(t; \zeta)$ converges towards the step function as $\zeta \to \infty$. The tangens hyperbolicus ensures that the time path of $\tau_c$ is infinitely smooth on $[0, \infty)$. Figure 18 displays the time path of $\tau_c$ for $\zeta \in \{1.6, 2.5, 4, 10, 30, 100, \infty\}$.
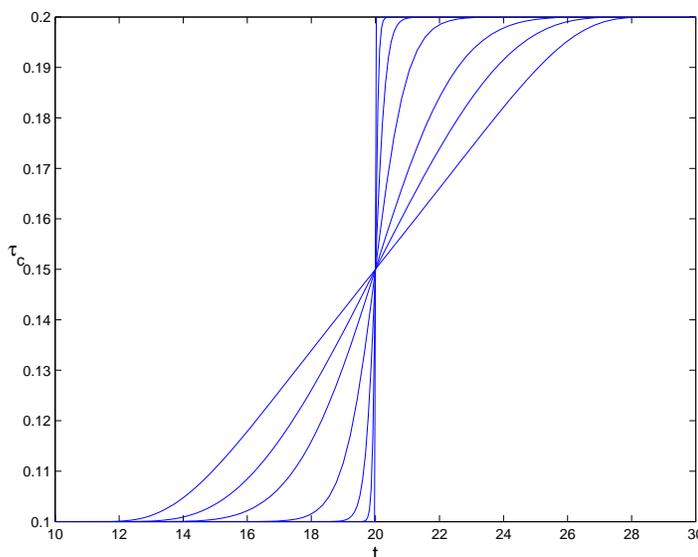


Figure 18: Time path of $\tau_c$ for $\zeta \in \{1.6, 2.5, 4, 10, 30, 100, \infty\}$

We repeat simulation exercises with $\zeta \in \{1.6, 2.5, 4, 10, 30, 100, \infty\}$ and $\sigma = .5$, and control the global relative error to be below $10^{-5}$ for $\zeta \neq \infty$. Results of the $(k, c, )$-phase diagram are displayed in Figure 19. Solid lines represent simulation results for $\zeta \neq \infty$, and the dotted line represents simulation results for $\zeta = \infty$. Crosses indicate starting points of trajectories and circles the final steady state. It can be recognized that solutions in the phase diagram converge to simulation results presented in the

previous section as $\zeta \to \infty$. However, there is one qualitative difference between paths for $\zeta \neq \infty$ and $\zeta = \infty$. Paths of the former parametrization of $\tau_c$ do not exhibit jumps at $\tilde{t}$, whereas the path for $\zeta = \infty$ exhibit a jump in $c$ at $\tilde{t}$. The speed of $c$ at $\tilde{t}$ is the higher, the higher $\zeta$ is and approaches infinity as $\zeta \to \infty$. We conclude that trajectories approach that of $\zeta = \infty$, wherefore we assume simulation results of the previous Section to be valid.
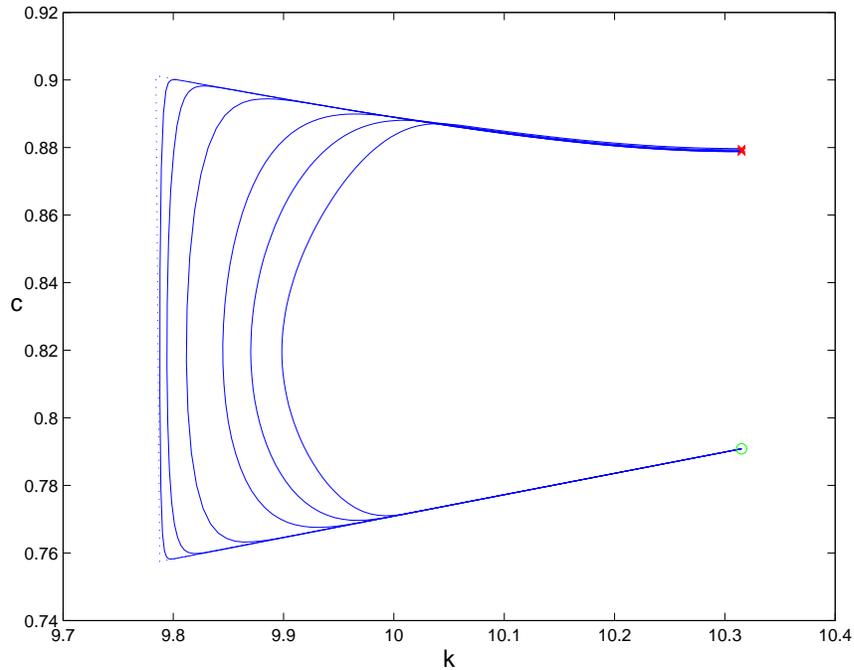


Figure 19: Phase-diagram of the RCK-model with different time paths of $\tau_c$

## 4.5  Summary

We derived the well-known continuity principle for adjoint variables for preannounced or anticipated changes in parameters for a broad class of infinite-horizon, perfect foresight, optimization models. For easy and intuitive numerical computation of the resulting multi point boundary value problem we suggested to simulate the resulting

differential algebraic system representing the first order conditions. By ensuring that the state variable $x$ and the adjoint variable $\lambda$ are continuous, potential jumps in the control variable $u$ are calculated automatically. This can easily be conducted employing the relaxation algorithm as proposed in this thesis. This algorithm treats differential and algebraic variables conceptually different such that the requirements for simulating the multi point boundary value problem at hand are automatically met.

The proposed algorithm was employed to solve a Ramsey model extended by an elementary Government sector. Simulations of a preannounced increase in the consumption tax showed a qualitative different pattern depending on the intertemporal elasticity of substitution.

Potential applications of this method emerge throughout in economic fields where the reaction on preannounced policy measures is of special interest in the context of perfect foresight optimizing agents.

# 5   Trade Liberalization and Income Distribution: A CGE Model for Jordan

This Section is based on the working paper of Feraboli and Trimborn (2006). All contributions of this Section are the author's own if not indicated otherwise.

## 5.1   Introduction

In this Section of the thesis we want to demonstrate the potential of the relaxation algorithm by simulating a large Computable General Equilibrium model. On the computational level, the model exhibits many of the difficulties described in this thesis. First, it exhibits six state variables and, therefore, a six-dimensional stable manifold. Second, the steady state is not unique, but long-run equilibria form a five-dimensional center manifold of stationary equilibria. Third, transitional dynamics are not result of a single, unanticipated shock, but due to a schedule of preannounced policy measures. Therefore, time paths of variables may exhibit interior corners or jumps.

The model at hand builds on previous work done by Feraboli et al. (2003), who implement a dynamic CGE model calibrated to the Jordanian economy. The model is characterized by the assumption of one representative consumer as used by Ramsey, Cass and Koopmans (see Ramsey, 1928, Cass, 1965 and Koopmans, 1965). We augment their dynamic CGE model by introducing heterogeneous households. The aim of this extension is to study the distribution of income and wealth in the context of a large CGE model. In detail, we disaggregate households into six different groups ranked by their disposable income. Within each group a representative consumer maximizes the sum of discounted utility according to his or her own budget constraint. Household groups' individual tax rate, wage rate, initial endowment of assets, transfers from government and abroad, as well as preferences concerning the

126

consumption basket are calibrated by data from a household survey undertaken in Jordan. Moreover, different households' time preferences are also calibrated from survey data.

In the context of General Equilibrium modelling several studies have been conducted to assess aspects of income distribution and poverty (see Reimer, 2002, and Winters et al., 2004, for a survey). We build on the strand of the literature, which embeds the disaggregated household groups within the CGE model (e.g. Bourguignon et al., 1992, Gibson, 2002).[78] This approach guarantees that the model is internally consistent, i.e. behavioral changes at the household level can transmit back into the macroeconomic solution. Moreover, these models exhibit additional channels, which can potentially influence income distribution, e.g. inflation, human capital accumulation, or labor market distortions. We extend the existing studies by relaxing the assumption of an exogenous saving rate. To our knowledge, this study is the first approach analyzing income distribution in a dynamic General Equilibrium framework with utility maximizing agents as used by Ramsey, Cass and Koopmans (see Ramsey, 1928, Cass, 1965 and Koopmans, 1965). This fact, perhaps, illustrates that the relaxation algorithm extends the range of numerically tractable models.

On the other hand, theoretical contributions analyze the effects of implementing heterogeneous consumers into a neoclassical framework (see e.g. Chatterjee, 1994 or Caselli and Ventura, 2000). By imposing restrictions on the utility maximizing agents it is guaranteed that the sum of all households behave as if it were a single household. This is of analytical convenience, since it is possible to analyze a model with one representative consumer in a first step and calculate the effects on heterogenous households in a second step. However, the restrictions on the utility maximizing agents imposed by this strand of the literature are not fulfilled in our model and

---

[78]Other studies of this strand are e.g. Löfgren (1999), Decaluwé et al. (1999), Cogneau and Robilliard (2000), Cockburn (2001), and Harrison et al. (2002).

would be neglected by the available survey data. Therefore, in our approach the behavior of disaggregated variables influences aggregate variables. This is described in detail in Section 5.3.

The model is implemented by means of the mathematical software Gauss, and calibrated to the Jordan Social Accounting Matrix of 2002. As a simulation exercise, we simulate the process of trade liberalization, which the Jordanian economy undergoes since 2002. The concrete institutional framework is reported in Feraboli and Trimborn (2006), and Feraboli (2007). Our simulation results indicate changes in per-capita level of welfare in Jordan between -0.03% and 0.19%, providing evidence that trade liberalization has indeed a different impact across heterogeneous households. Remarkably, the behavior of aggregate variables is qualitatively consistent with previous work done by Feraboli et al. (2003).[79]

In Section 5.2, we explain the model briefly, in Section 5.3, we describe dynamic properties of the model, in Section 5.4, we present the calibration process and explain the numerical solution method, in Section 5.5, we briefly analyze and discuss the simulations, and in Section 5.6 we draw the main conclusions.

## 5.2 The Model

We model a dynamic small open economy, building on the model of Feraboli et al. (2003). For each of six different household groups, a representative consumer maximizes discounted intertemporal utility subject to a budget constraint. In the domestic economy there are nine production sectors, eight of which are producing goods and one produces services. Aggregate private consumption, government consumption, and aggregate investment are Cobb-Douglas composites of nine different sectoral outputs, which, in turn, are Armington (1969) composites of domestically

---

[79]Previous work by Hosoe (2001) on Jordan's trade liberalization implements a static model with one representative household. Simulation results suggest average welfare gains of 0.44%.

produced and imported goods. Firms produce nine different commodities using a Leontief production technology between sectoral goods and a value-added factor, which is a CES composite of capital and different kinds of labor. Total output can be sold domestically or exported according to a CET specification. The Government raises taxes and collects import tariffs. Government revenues are spent for a fixed amount of government consumption as well as for transfers to households.

The domestic economy accepts the world price as given in international markets. Perfect competition and full employment are assumed in all sectors. Production factors are perfectly mobile across sectors.

In the following, we focus on the main mathematical equations. The remainder of the equations used in the model is relegated to the appendix.

### *Households*

The problem of each representative infinitely-lived household, $i$, is to maximize discounted intertemporal utility

$$\int_0^\infty \log\left(C_i\right)e^{-\rho_i t}dt \qquad i = 1, \ldots, 6 \tag{162}$$

subject to

$$\dot{K}_i = \frac{YD_i - P_C C_i}{P_I} - \delta K_i \tag{163}$$

$$K_i\left(0\right) = K_{i,0}$$

where $C_i$, $YD_i$, $K_i$ are consumption, disposable income, and capital of household $i$, respectively. Each representative household discounts future utility with discount rate $\rho_i$, which is specific to each household group. We discuss this modeling choice in the subsequent Section.

Disposable income of each household group is given by

$$YD_i = (1 - \tau_i)(w_i L_i + r K_i + GT_i + FT_i) \tag{164}$$

whereby $w_i$, $L_i$, $K_i$, $GT_i$ and $FT_i$ denote the individual wage rate, labor endowment, and capital endowment of household $i$, as well as government and foreign transfers to household $i$, respectively. The interest rate $r$ is identical for each household since capital is a homogenous good. Each household pays a different income tax $\tau_i$ depending on its household group.

### *Firms*

Sectoral output in the domestic economy is determined by a two-stage production process, which exhibits at the top tier a Leontief (or fixed-proportions) specification between intermediate input and value-added output. Each representative firm producing commodity $j$ generates total output according to the following production technology

$$Q_j = min\left\{\frac{VA_j}{a_{VA,j}}, \frac{q_{1,j}}{a_{1,j}}, \dots, \frac{q_{9,j}}{a_{9,j}}\right\} \qquad j = 1, \dots, 9 \tag{165}$$

where $Q_j$ and $VA_j$ are sectoral output and value-added output, respectively. $q_{k,j}$ is intermediate input produced by sector $k$ and used in the production of activity $j$. Leontief coefficients are denoted by $a_{k,j}$, and productivity of value added in producing commodity $j$ is $1/a_{VA,j}$.

At the second tier, intermediate input $q_{i,j}$ is a Cobb-Douglas composite of domestic and foreign intermediate consumption goods.

Value-added production is determined by a technology characterized by a constant elasticity of substitution between the primary inputs, capital $(KD_j)$ and six different

types of labor $LD_{i,j}$, pertaining to each household group $i$

$$VA_j = A_j \left[ \sum_{i=1}^{6} \alpha_{i,j} LD_{i,j}^{\frac{\sigma_j-1}{\sigma_j}} + \left(1 - \sum_{i=1}^{6} \alpha_{i,j}\right) KD_j^{\frac{\sigma_j-1}{\sigma_j}} \right]^{\frac{\sigma_j}{\sigma_j-1}}$$

$$\alpha_{i,j} > 0, \ 0 < \sum_{j=1}^{6} \alpha_{i,j} < 1, \ \sigma_j > 0, \ \sigma_j \neq 1$$

(166)

where $A_j$ is the time-invariant technological parameter, $\alpha_{i,j}$ is the share of labor of household $i$, and $\sigma_j$ denotes the constant elasticity of substitution between primary inputs. At the value-added production stage, firms minimize production costs subject to the above technology constraint.

### Government

The government consumes an exogenous amount of goods, raises taxes and tariffs, and provides transfers to consumers. We assume the government to run a balanced budget. Government consumption is determined by a CES Armington specification between domestically-produced goods and imports. Government revenue is generated from the Value-Added Tax, that applies with different rates to domestic and imported goods ($VAT^D$ and $VAT^M$), the income tax ($TY$) and import duties ($TM$), which apply with different rates to the EU and the rest of the world, and exogenous and fixed foreign grants, ($FRG$). The expenditure is given by an aggregate transfer to households ($TR$) and an aggregate fixed consumption of goods and services ($\bar{G}$).

The government budget is, therefore, given by

$$VAT^D + VAT^M + TY + TM + FRG = TR + \bar{G}.$$

(167)

### Market clearing

The equilibrium in the factors markets requires for each type of labor, aggregate endowment of labor to be equal to aggregate labor demand and aggregate capital

stock to be equal to aggregate demand for capital

$$L_i = \sum_{j=1}^{9} LD_{i,j} \qquad i = 1, \ldots, 6 \tag{168}$$

$$\sum_{i=1}^{6} K_i = \sum_{j=1}^{9} KS_j \tag{169}$$

where $L_i$ and $K_i$ are, respectively, labor and capital supplied by household $i$.

The equilibrium condition on the domestic goods markets is

$$X_j = \sum_{k=1}^{9} q_{k,j} + C_j + I_j + G_j \qquad j = 1, \ldots, 9 \tag{170}$$

where $I_j$ and $G_j$ are investment demand and government consumption, respectively.

The equilibrium in the balance of payments is given by

$$\sum_{j=1}^{9} PWM_j M_j = \sum_{j=1}^{9} PWE_j E_j + \sum_{i=1}^{6} FT_i + FGR \tag{171}$$

where $M_j$ and $E_j$ are, respectively, imports and exports of sector $j$, $PWM_j$ and $PWE_j$ are the exogenous world prices of, respectively, imports and exports of sector $j$, and $FGR$ are foreign grants donated to the Jordanian government.

## 5.3 Discussion of the model's dynamic properties

In a first step we discuss the way we extended the model of Feraboli et al. (2003) by different household groups, and which alternative modeling choices would be available. As we will argue, the way we introduce the household groups determines in a crucial way the model's dynamic behavior. In a second step, we discuss the model's long-run dynamic behavior with a special emphasis on the evolution of the distribution of wealth.

It is important to focus on the taxation of capital income for different household groups. In many countries the average tax paid on capital income is increasing with

the total amount of capital income. This holds also for Jordan, where a distinction between capital and labor income is not made, and marginal income taxation is increasing in total income. If we assume that the income tax $\tau_i$ does not depend on total income (see equation (164)),[80] the Euler equation for household group $i$ can be derived as

$$\frac{\dot{C}_i}{C_i} = \frac{(1 - \tau_i)r}{P^I} - \rho_i - \delta \qquad (172)$$

with the price index of aggregate investment, $P^I$. Since capital is assumed to be homogenous, the interest rate $r$, and the rate of depreciation $\delta$, is the same for all household groups. Therefore, assuming $\rho_i = \rho$ for all $i$ would cause the growth rate of consumption to be higher for household groups facing a lower tax burden $\tau_i$, and vice versa. The reason is that low income households face higher incentives for saving, since their net return on savings is higher. This, on the other hand, would cause the household, which faces the lowest tax rate, to own all the assets in the long-run. By assumption, the distribution of wealth would become extreme, even if distributional policy measures in the form of transfers are undertaken.[81]

Since this outcome is at odds with the assumption of a tax rate independent of income, a possible cure of this problem would be to assume an income-dependent tax rate $\tau_i(Y_i)$. Then, net return on savings would be higher than average for households with an income below average, while net return on savings would be lower than average for households with an income above average. Then, by assumption, the distribution of wealth would become more equal in the long-run.

By contrast, we did not want to impose any long-run distribution of wealth by assumption. The only possibility, then, is to assume that long-run growth rates of consumption are equal for every household group. This holds if $\rho_i$ is calibrated

---

[80]Strictly speaking, this requires the nominal income of household groups to be constant.
[81]Becker (1980) and Becker and Tsyganov (2002) confirm this results in a similar setting.

according to

$$\rho_i = \frac{(1 - \tau_i)r^*}{(P^I)^*} - \delta \tag{173}$$

with $r^*$ and $(P^I)^*$ representing the steady state values of the interest rate and the price index of aggregate investment, respectively. Note that we assume $\tau_i$ to be independent of the development of individual income. It will become clear that this assumption is justifiable under this setting. Moreover, simulation results for the specific policy measures analyzed in this Section confirm that the relative income position of households does not change. Nonetheless, this assumption would not be feasible any more, if huge distributional policy measures would be undertaken.

The assumption of equal long-run growth rates of consumption has strong implications for the dynamic behavior. To start with, the steady state can no longer be unique. The reason is that the standard model with one representative consumer has to fulfill

$$\frac{(1 - \tau)r^*}{(P^I)^*} - \rho - \delta = 0 \tag{174}$$

at the steady state. This condition determines a unique ratio of $\frac{r^*}{(P^I)^*}$ at the steady state for given parameter values. If the model is extended to six household groups, the equation becomes six dimensional according to

$$\frac{(1 - \tau_i)r^*}{(P^I)^*} - \rho_i - \delta = 0 \qquad i = 1, \dots, 6. \tag{175}$$

However, all the six equations determine the same ratio $\frac{r^*}{(P^I)^*}$ at the steady state. The real interest rate is considered to be given if (175) is solved for the individual discount rates. Since the number of variables, which refer to the household groups (i.e. $C_i$, $K_i$) has increased sixfold, there is not enough information to determine the steady state uniquely. The system of equations has five degrees of freedom. Therefore, according to corollary 10 of Section 2.1 there exists a five-dimensional center manifold

of stationary equilibria. This result is verified by numerical computation of the Jacobian matrix at the steady state. Since the matrix exhibits five zero eigenvalues, the model exhibits a five-dimensional center manifold. Nonetheless, according to theorem 13 of Section 2.1 adjustment dynamics are unique in a neighborhood of the center manifold.[82] Note that formally theorem 13 does only hold from a time $\tilde{t}$ onwards, for which no further parameter changes occur. The reason is that the theorem can only be applied to autonomous systems. Therefore, the theorem cannot be applied to investigate the dynamics during anticipated parameter changes, but only for the time period after these changes. Hence, no connection between the distribution of initial wealth and the distribution of steady state wealth can be made.

Chatterjee (1994) as well as Caselli and Ventura (2000) investigate in a neoclassical framework under which conditions the sum of all households behave as if it were a single household. They state restrictions for the utility function of heterogeneous households to hold. This is of analytical convenience, since then it is possible to analyze a model with one representative consumer in a first step and calculate the effects on heterogenous households in a second step. However, these restrictions on the utility functions are not fulfilled in the model at hand since individual households' discount rates differ.

## 5.4 Calibration procedure and numerical solution technique

We calibrate the model to match the Jordanian economy of 2002 as a simulation exercise to demonstrate the potential of the relaxation algorithm on the one hand, and the explanatory potential of the model on the other hand. As a concrete policy example we select the process of trade liberalization, which the Jordanian economy undergoes since 2002. The concrete institutional framework is reported in Feraboli

---

[82]The Jacobian matrix numerically evaluated at the center manifold exhibits one eigenvalue with negative real part.

and Trimborn (2006), and Feraboli (2007). For simulation exercises we reduce the tariff rates according to Table 3, which corresponds to the actual tariff reduction that Jordan has agreed on due to the EU-Jordan Association Agreement signed in 1997.

Table 3: Tariff reduction schedule of the AA

| Year | Agric. | Mining | Food | Text. | Paper | Chemic. | Miner. | Other |
|---|---|---|---|---|---|---|---|---|
| Pre-AA | 17.0% | 9.4% | 29.2% | 14.1% | 13.2% | 2.8% | 12.2% | 12.2% |
| 2002 | 17.0% | 5.6% | 29.2% | 8.5% | 7.9% | 1.7% | 7.3% | 7.3% |
| 2003 | 17.0% | 5.0% | 29.2% | 7.5% | 7.0% | 1.5% | 6.5% | 6.5% |
| 2004 | 17.0% | 4.4% | 29.2% | 6.6% | 6.2% | 1.3% | 5.7% | 5.7% |
| 2005 | 17.0% | 3.8% | 29.2% | 5.7% | 5.3% | 1.1% | 4.9% | 4.9% |
| 2006 | 15.3% | 2.8% | 26.3% | 4.2% | 4.0% | 0.8% | 3.7% | 3.7% |
| 2007 | 13.6% | 2.5% | 23.4% | 3.8% | 3.5% | 0.8% | 3.3% | 3.3% |
| 2008 | 11.9% | 2.2% | 20.4% | 3.3% | 3.1% | 0.7% | 2.9% | 2.9% |
| 2009 | 10.2% | 1.9% | 17.5% | 2.8% | 2.6% | 0.6% | 2.4% | 2.4% |
| 2010 | 8.5% | 1.6% | 14.6% | 2.4% | 2.2% | 0.5% | 2.0% | 2.0% |
| 2011 | 8.5% | 1.3% | 14.6% | 1.9% | 1.8% | 0.4% | 1.6% | 1.6% |
| 2012 | 8.5% | 0.9% | 14.6% | 1.4% | 1.3% | 0.3% | 1.2% | 1.2% |
| 2013 | 8.5% | 0.6% | 14.6% | 0.9% | 0.9% | 0.2% | 0.8% | 0.8% |
| 2014 | 8.5% | 0.0% | 14.6% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |

The calibration procedure is based on the Social Accounting Matrix (SAM) for Jordan constructed for the year 2002.[83] The model's parameters are calibrated such that the SAM represents a solution of the model where all variables are stationary except asset accumulation of individual households. The reason for this is that the fractions of savings and assets are not the same across households, and, therefore, the assumption of a stationary individual capital accumulation would violate the SAM. This can easily be seen by considering individual household group's capital accumulation equation (163). To be stationary, this equation is required to satisfy

$$\frac{YD_i - P_C C_i}{P^I} = \delta \qquad i = 1, \ldots, 6 \tag{176}$$

[83]The SAM was constructed by Feraboli and Kolev. We thank the latter for very helpful research assistance.

Given the prices $P_C$ and $P^I$, $\delta$ can be calibrated such that it satisfies equation (176) for one household. However, data for $YD_i$ and $C_i$ does not allow to satisfy equation (163) for six households because of missing degrees of freedom. Therefore, simulating the model without any parameter changes does not result in a stationary solution. However, running the simulation without a parameter shock yields transitional dynamics of small magnitude. This indicates that a stationary solution is very close to the initial state of the economy.

Household survey data allows disaggregation into six different groups of households. Each group differs with respect to labor income, capital income, transfers from government and from abroad, income-tax payments, and savings, as well as total consumption and the composition of total consumption. Within the calibration process, these differences result in varying exogenous variables for each group of households as well as diverse parameters. We rank heterogeneous households by their income in 2002, i.e. household group one earns the lowest income and household group six the highest.[84] This brings about an almost monotonous ranking in labor income, capital income and, reversely, transfers received. Also, the share of capital income (transfers) on total income is almost monotonously increasing (decreasing) while the share of labor income on total income is hump-shaped (see Figure 20).

Elasticities of substitution are obtained from the existing literature (see Devarajan et al., 1999, Devarajan and Go, 1998, and Lucke, 2001). The domestic interest rate is set to 10%. Once these parameters have been fixed, the remaining parameters are calibrated from the SAM.

The model is programmed in Gauss and solved with the relaxation procedure as proposed in this thesis. Since the model exhibits a continuum of stationary equilibria, we explicitly utilize the fact that this numerical procedure does not require

---

[84]For convenience we will denote household group one the poorest and household group six the richest household.
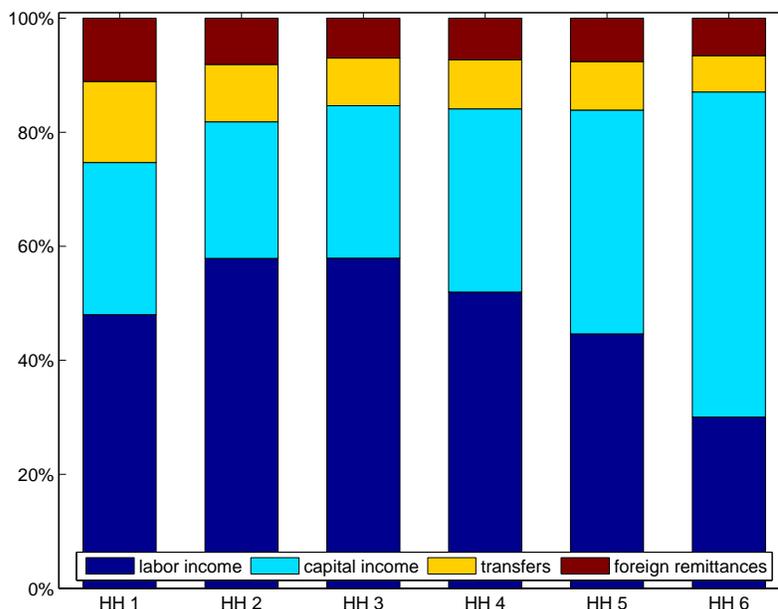
Figure 20: Income composition of households

information regarding an achieved stationary equilibrium in advance. The particular stationary equilibrium is determined within the iteration process. In addition, the relaxation procedure can simulate transitional dynamics on multi-dimensional stable manifolds. This means that an increase in the dimension of the model, especially in the state space, does not cause any conceptual problems. Moreover, households experience a preannounced schedule of changes in tariffs. Therefore, the model exhibits anticipated changes in parameters as investigated in Section 4.

## 5.5  Simulations

As a simulation exercise, we simulate a gradual reduction of tariff rates due to table (3). Since the data available for the calibration procedure represents the Jordan economy of the year 2002, this is our benchmark year. In our simulation, tariff rates are gradually reduced in the subsequent years. We assume the government budget

138

to be balanced. The endogenous variable to achieve this is total transfers to households. Precisely, this means that total transfers from the government, granted to households, are endogenous, whereas the share each household receives is fixed. This assumption guarantees that the reduction of distortionary tariffs is not accompanied by distortionary side-effects as additional taxation. Since poor household groups rely heavily on transfer income, the cut in transfers implies an effect on income distribution according to the discussion in Section 5.3. Since the impact of tariff reduction may be different for household groups income, consumption, and welfare, we report simulation results for all of them. In a second scenario we investigate how an additional ten-percent increase in all VAT rates affects the economy.

Our simulation results indicate an increase in aggregate capital, output, disposable income, and consumption in the long-run. Therefore, simulation results are in line with previous findings by Feraboli et al. (2003). We find an increase in welfare for most household groups, and welfare gains of the poor households are slightly higher than gains of the rich households. However, trade liberalization is not pareto improving since some households (group five) are even worse off. Figure 21 represents welfare changes of both scenarios. The blue line summarizes the impact on welfare for each household group and its absolute size in the baseline scenario, whereas the green line refers to the second scenario.

Since welfare gains are roughly higher for poor households, one may expect inequality to decrease. However, the opposite is the case. We measure inequality with the Gini index of income (see Gini, 1912), which increases immediately with trade liberalization and over time, which is measured in years (see Figure 22, (i)).[85] The reason for this can be seen in Figure 22 (ii), which indicates that the initial response of income to trade liberalization is positive for household groups 3, 4, 5, and 6 and

---

[85]An alternative measure of inequality, Theils entropy of income (see Theil, 1967), yields qualitatively the same result.
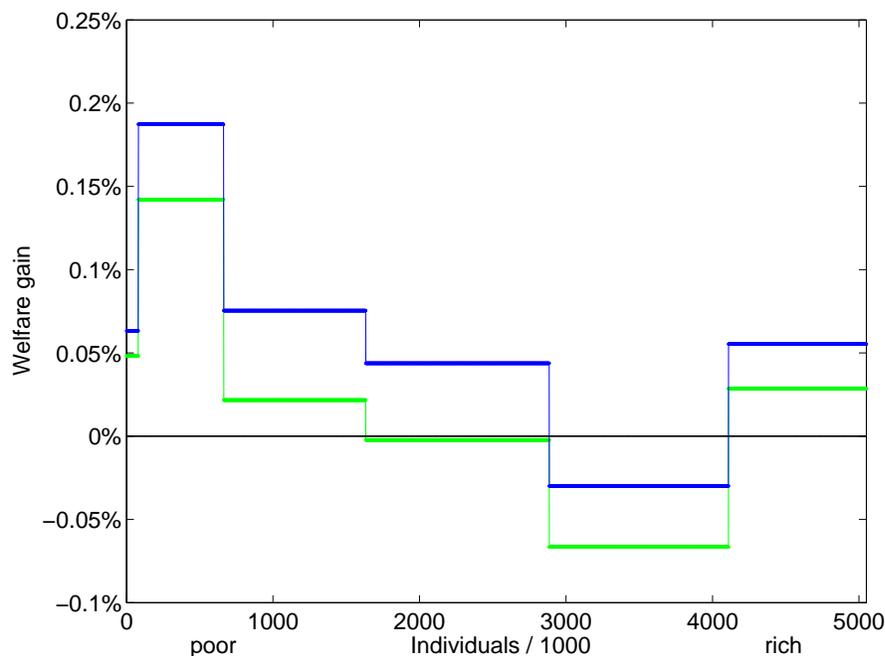
Figure 21: Welfare effects of both simulations

negative for household groups 1 and 2. In addition, income increases more drasti-
cally over time, the richer the household group is.[86] That means, the gap in income
increases over time, as well. The reason for this is that households rely differently
on various kinds of income. First, transfers are cut immediately when trade liberal-
ization starts and are even decreasing in the subsequent years because the tax base
and, therefore, government revenue increase sluggishly. This affects poor households
relatively severely. Secondly, since the aggregate capital stock grows, wage income
increases over time. Poor households benefit more from this because they rely heavily
on labor income. Due to the fact that a large part of their income is labor income,

---

[86]Whereas time is continuous, the import tariff reduction takes place at specific points in time.
Therefore, government transfers to households drop sharply at the beginning of each year and re-
cover smoothly during the remainder of the year. Hence, the income flow follows a discontinuous
path.

poor households can offset the negative effect of reduced transfers after some periods. Finally, households own different amounts of capital. Higher incentives for investments condense in a higher interest rate. Therefore, capital income for the four richest groups of households is increasing instantaneously and over time, due to capital accumulation. This capital accumulation can be seen in Figure 22 (iii). Poor households use their already tiny amount of assets to smooth consumption, since they have to overcome temporary losses in income (see Figure 22 (iv)). Therefore, poor households even deaccumulate capital, and this deaccumulation is insignificant for the economy as a whole.
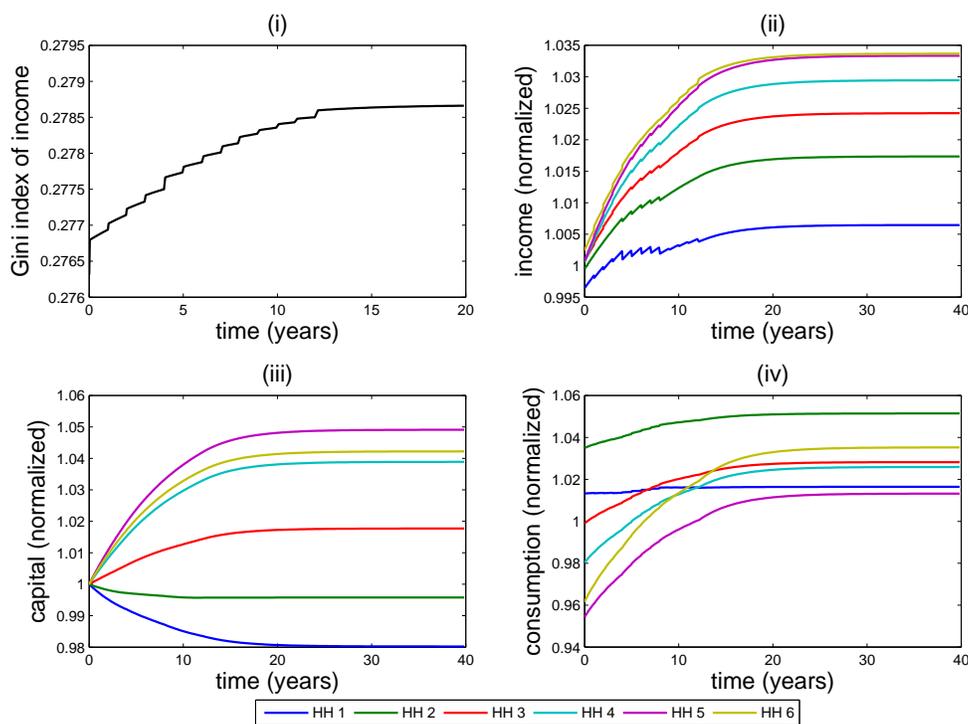


Figure 22: Effects of AA on heterogeneous households (baseline simulation)

In the second scenario we assume the government to undertake the additional fiscal

measure of a 10% increase of all VAT rates to overcome losses in revenues.[87]  This
10% increase has two effects on the economy.  Since this allows the government to
limit the cut in transfers, we expect poor households to benefit more in relation to
rich households from this additional fiscal measure.

Simulation results indicate that the effect of trade liberalization dominates the rise
in VAT rates.  Aggregate variables behave qualitatively the same.  However, welfare
gains are reduced for every group of households, and for one household group (group
four), the welfare gain turns into a loss.  This is illustrated in Figure 21 where
the green line indicates welfare changes of the second scenario.  Although transfers
remain even higher than in the benchmark year, every households' welfare is lower
compared to the previous simulation.  The reason is that the rise in the VAT rates has
a negative impact on investment and, therefore, reduces the aggregate accumulation
of capital compared to the baseline scenario.  This determines steady-state values
for private consumption and capital which are below the steady-state levels in the
previous simulation.

## 5.6  Summary

We constructed a large CGE model incorporating heterogenous household groups to
demonstrate the potential of the relaxation algorithm.  More precisely, the model
exhibits a six-dimensional stable manifold, a five-dimensional center manifold of
stationary equilibria, and a preannounced schedule of shocks.  Therefore, the model
exhibits many of the characteristics, which complicate the numerical analysis of
transitional dynamics.  The presence of a center manifold of stationary equilibria is
due to the fact that we assumed growth rates of consumption to be equal for different
household groups in the long-run.

---

[87]However, total government transfers to households remain the endogenous variable to balance
the government's budget.

The model was calibrated to the Jordanian economy of the year 2002 as an computational exercise. Moreover, household specific data was calibrated according to a Jordanian household survey. As a baseline simulation the trade liberalization between Jordan and the EU was simulated. We found that welfare effects were different across household groups, but predominantly positive. In addition, we found effects to be contrarian concerning welfare and income distribution. While on the one hand welfare gains are slightly higher for low income households, on the other hand the gap in income will increase, especially in the long-run.

To our knowledge, this is the first large CGE model, which addresses the issue of wealth and income distribution. Perhaps this is due to the numerical difficulties, which can be overcome by employing the relaxation algorithm as the numerical solution method. Therefore, we hope that this analysis paves the way for future analysis in this field.

# 6  Conclusion

We proposed the relaxation algorithm as a simple and powerful method for simulating the transition process in growth models. This method has a number of important advantages: First, the application of the procedure is fairly user friendly. Second, it is generic for a wide range of dynamic systems including stiff differential equations and systems giving rise to a continuum of stationary equilibria. Third, the variant of the relaxation algorithm we propose exploits the infinite time horizon, which usually underlies optimal control problems in economics, in a natural manner. And fourth, the algorithm can solve preannounced or anticipated parameter changes in an easy and intuitive way. As an illustrative application, we computed the transition process of the Jones (1995a) and the Lucas (1988) models. In addition, we solved the Ramsey-Cass-Koopmans model, extended by a simple government sector. Preannounced or anticipated changes in the consumption tax induced qualitatively different results, depending on the intertemporal elasticity of substitution. Lastly, we constructed a General Equilibrium model with heterogenous households to demonstrate the potential of the relaxation algorithm. The model was calibrated to Jordanian data and employed to solve different scenarios of trade liberalization between Jordan and the European Union. This model exhibits a six-dimensional stable manifold, a five-dimensional center manifold of stationary equilibria, and a schedule of preannounced policy measures. These characteristics make the application of usual procedures highly inefficient or even impossible.

Future research could go in different directions. On the computational level, the relaxation algorithm could possibly be extended to solve stochastic continuous-time models. Another extension is that of Trimborn (2006), which allows for solving growth models, for which the balanced growth rates cannot be calculated analytically. On the economics level, availableness of a more powerful tool for investigating

transitional dynamics could allow for richer dynamic models. For example, growth models often suffer from the fact that the steady state and transitional dynamics towards it cannot be analyzed analytically. This especially holds for models, which have several sources of dynamics, e.g. different engines of growth. Analyzing transitional dynamics by numerical simulation offers a way to extend the dynamics of existing models, without loosing track of the model's dynamic behavior. On the other hand, many applied Computable General Equilibrium models suffer from the fact that numerical solution of high-dimensional dynamic models is difficult or even impossible, if the model exhibits a saddlepoint with a multi-dimensional stable manifold or a center manifold of stationary equilibria. This is perhaps why many models in this area do not contain perfect-foresight optimizing households. We showed that these models are numerically tractable, if the relaxation algorithm is employed.

# 7 Appendix

## 7.1 Appendix on Section 3

In Section 7.1.1, we describe how the relaxation algorithm is implemented in Matlab. Special emphasis is given on the description of how to modify the files if a researcher wishes to simulate a specific model. In Section 7.1.2, 7.1.3, 7.1.4, and 7.1.5, respectively, we print the code of the most important files of the relaxation algorithm, *main.m*, *initrelax.m*, *relax.m*, and *rrefmod.m*. The actual algorithm can be found in Section 7.1.4 and 7.1.5, in which the code of the files *relax.m* and *rrefmod.m* is printed.

### 7.1.1 Implementation of the Relaxation Algorithm

We want to describe in more detail how the algorithm is implemented in MatLab, and which steps a researcher has to perform to simulate a model numerically. At this point, we want to emphasize the user-friendlyness of the algorithm. It is possible to simulate the transition process of a wide range of infinite horizon optimization models without the knowledge how the algorithm works. The MatLab files can be used as a "black box". The researcher who intends to simulate a specific model only has to provide the dynamic system and the set of underlying parameters.

### *Installation*

There are two different types of files, the system files and the user files. The system files are general to every model and should be copied into a separate folder, which has to be assigned as a MatLab search path (menu *File → Set Path*). These files should not be modified. The user files carry the information due to a particular model and should be copied into a model-related folder, which has to be assigned as

the MatLab *Current Directory.*

### *Structure of the Code*

The aim of the structure of the code is not only to make standard applications comfortable, but also to allow for advanced applications, e.g. multiple simulations of the same model with different initial conditions or parameter values. Therefore, in this section the structure of the code and the usage of the two most important functions *initrelax.m* and *relax.m* will be described in more detail. Note, that for convenience some files are double among the user files and the system files. The reason is that files in the current directory have priority when Matlab is executing them. Therefore it is possible to place empty files in the system folder to keep the number of files in the user directory small and concise.

The central file of the code is the file *main.m*, which calls the different scripts and functions in the following order:

| | |
|---|---|
| globalpar | initializes the global parameters |
| parini | loads the values of the parameters |
| relaxsetting | loads the settings for the relaxation procedure |
| initrelax | converts the settings into a form suitable for *relax.m* |
| relax | executes the Relaxation algorithm |
| varex | disentangles the variables and stores them in the memory |

The actual relaxation algorithm and the code to prepare the model for the relaxation algorithm are exported to functions and are not integrated into *main.m* as a script. The reason for this is that as a script they would be able to manipulate any variables in the memory. If they are implemented as functions their interactions with model specific variables in the memory is kept as small as possible. Detailed information about the functions *initrelax.m* and *relax.m* can be listed by calling *help initrelax* and *help relax* at the MatLab command window. They are used according

to

```
[guess, start, errorcode] = initrelax(@funcODE, @funcSTAT, n,
 n1, n3, nu, y, M, statev, t)
```

```
[t, x] = relax(@funcODE, @funcSTAT, @funcINI, @funcfinal, n,
 n1, n3, nu, y, M, start, Endcond, maxit, tol, damp, dampfac)
```

Output from *initrelax.m* are the variables *guess*, *start*, and *errorcode*. The first two are again input arguments for *relax.m*, the last is a code indicating $error = 0$ for the case of no error and $error > 0$ for different errors that the code detected in the model set-up. Output of *relax.m* is a row vector $t$ and a matrix $x$. Each column of $x$ is the list of the variable values at the point of time of the corresponding entry of $t$. The file *varex.m* disentangles the variables, such that each variable is a row vector with corresponding time vector $t$.

Input for both functions are different parameters, which will be described in more detail below. Further, input arguments for both functions are different function handles, which pass the information about a specific function name. These functions have the following shape:

| function | input arguments | output arguments |
|---|---|---|
| @funcODE | time, vector of variables (dim $N$) | vector of RHS of ODE (dim $n$) |
| @funcSTAT | time, vector of variables (dim $N$) | vector of residuals of static equations (dim $n_3$) |
| @funcINI | time, vector of variables (dim $N$) | vector of residuals of initial boundary conditions (dim $n_1$) |
| @funcfinal | time, vector of variables (dim $N$) | vector of residuals of final boundary conditions (dim $n - n_1$) |

These functions are among the system files with the above names but may well be replaced by more user specific files.

### *Preparation of the model*

First of all, the dynamic system under study must be transformed into stationary variables, i.e. in variables that approach a finite constant asymptotically as time goes to infinity. It is assumed that now the model consists of $n$ ordinary differential equations, $n_1 (\leq n)$ initial conditions and (optional) $n_3$ static equations that have to be satisfied at all times. Notice that you do not have to modify these equations. They are plugged into the algorithm as they are written in continuous time. Only the static equations (if present in the model) must be rewritten as a residual $(g(y) = 0)$.[88] You should go through the following steps carefully.

### *Simulation of a model*

The model dependent files are the following:

globalpar.m
varex.m
parini.m
relaxsetting.m
ODE.m

with the optional files

StatEq.m
shock.m
initbound.m
finalbound.m

Go through the following steps carefully.[89]

- **File globalpar.m**

  List all the names of the parameters with the command *global*.

---

[88]In some algorithms it is necessary to differentiate the static equations with respect to time. This is not necessary for the application of the relaxation procedure.

[89]Note, that you cannot name any variable or parameter with $x$.

- **File varex.m**

  List all the variables like in the original with an increasing index. You have to bring the variables in a special ordering. First write down all the state variables then all the control (or adjoint) variables and finally all variables, for which the static equations should hold. If you have an auxiliary variable that does not appear in the static equations you have to place it at the very end.

- **File parini.m**

  Assign all parameter values which are valid at the new steady state.

- **relaxsetting.m**

  – Specify the dimensions:

    n1   number of initial boundary conditions
         (equals the number of state variables, for instance,
         n1=1 in the case of the RCK model)

    n    number of differential equations

    n3   number of static equations to be solved simultaniously

    M    number of mesh points

  – *normal* specifies, which variables you want to normalize to 1 at the steady state. The vector [1   3   6] for example will normalize variables 1, 3 and 6. If you do not want to normalize any variable return an empty vector [   ].

  – You have to provide a rough guess for the steady state. Here the same ordering for the variables should be maintained as for the file *varex.m*.

  – You can induce a transition process by reducing one of the state variables compared to its new steady state value. E.g. $statev(1) = .5$ halves the value of the first state variable. If you do not enter anything, the initial values of the state variables will be calculated due to a shock in parameter values as specified in *shock.m*. If you assign $statev = 0$, you can specify initial conditions in the file *initbound.m* (see below).

– Change the last section of the file only if problems occur.

*Endcond:* Specifies the final conditions. The vector [1  3  6] for example will set the RHS of differential equations 1, 3 and 6 to zero. Default setting is to use the differential equations of the control (or adjoint) variables. You should mind the dimensions when changing this. If you set $Endcond = 0$ you can enter the final boundary conditions in the file *finalbound.m* (see below).

*tol:* This declares the tolerance for the Newton procedure.

*maxit:* Maximum number of iterations

*nu:* This denotes the parameter for time transformation. Higher values allocate the mesh more towards the beginning of time, while lower values increase the density around the steady state. A good value for $nu$ depends crucially on the speed of convergence towards the steady state. It should be chosen such that it minimizes the maximum change of the variables between the mesh points. For models with a half-life of about $50 - 80$ time units, a value of $nu = 0.04$ is advisable.

*damp:* You can dampen the Newton procedure by the dampening factor *damp*. It will be multiplied with the factor *dampfac* in every iteration until it equals 1. A value of 1 does not damp, values between 0 and 1 increase the convergence radius of the Newton procedure but also lead to slower convergence and therefore more iterations.

• **File ODE.m**

List the right-hand-sides of your differential equations here. Assign them in vector form to *funcODE* as in the original. You should mind the ordering, since the first differential equation should belong to the first variable and so on. In the non-autonomous case time is denoted as $t$.

The following files are optional. If they are needed, create them in the model-related current directory.

- **File StatEq.m**

  List the residuals of your static equations here. Assign them in vector form to *funcSTAT* as in the original. Time is again denoted as $t$.

- **File shock.m**

  In this file you can specify a specific shock in the parameter values. Assign (e.g. in an *if* loop) the old steady state values to the time $t = -1$ and the new steady state values to all other $t$. It is also possible to investigate continuous parameter changes or interior (expected) jumps in the parameters.

- **File initbound.m**

  If you want to employ variable initial boundary conditions, enter them here and set $statev = 0$ in the file *relaxsetting.m*. Assign the residuals to the vector *funcINI*. Mind that you need $n1$ equations.

- **File finalbound.m**

  If you want to employ final boundary conditions different to the RHS of the differential equations, enter them here and set $Endcond = 0$ in the file *relaxsetting.m*. Assign the residuals to the vector *funcfinal*. Mind that you need $n - n1$ equations.

After running the code by entering *main*, the variables are stored in the memory with their assigned names. For plotting the results you should consider the time vector $t$. You can view the results quickly by calling *plotrelax(t,x,n1)*, whereas $t$ and $x$ are output from *relax.m*. For further information call *help plotrelax*. Eigenvalues and eigenvectors of the linearized system at the steady state can be evaluated by calling *eigDASg.m*. This routine can also handle differential algebraic systems.

### 7.1.2 Code for main.m

```
% Version 3.0
% RELAXATION algorithm to solve infinite-horizon continuous time models.
%
% For further information contact Timo Trimborn, University of Hamburg
% or visit http://www.rrz.uni-hamburg.de/IWK/trimborn/relaxate.htm


clear all
disp(['Initialize Relaxation algorithm']);disp([' ']);

tic
% Initialization of global parameter
globalpar
% Loading the parameter values
parini
% Loading the settings for the Relaxation algorithm, i.e. dimensions,
% boundary conditions etc.
relaxsetting

% Converting settings into a form suitable for relax.m
[guess, start, errorcode]=initrelax(@funcODE, @funcSTAT, n, n1, n3,
nu, y, M, statev);

% Execution of relaxation algorithm if no error
% occurred during initilization
if errorcode==0

    [t, x]=relax(@funcODE, @funcSTAT, @funcINI, @funcfinal, n, n1, ...
        n3, nu, guess, M, start, Endcond, maxit, tol, damp2, dampfac);

    %Normalization of specified variables
    for i=1:M
        x(normal,i)=x(normal,i)./x(normal,end);
    end;
    % Extracting of variables and storing them
    varex

end

%Calculation time
time=toc; timesec=mod(time,60); timemin=floor(time/60);
disp(['Calculation time: ',num2str(time),' seconds
(',num2str(timemin),' min ',num2str(timesec),' sec)']);
```

### 7.1.3 Code for initrelax.m

```
function [guess, start, errorcode] = initrelax(funcODE, funcSTAT, n,
```

```
n1, n3, nu, y, M, statev, t)
% Version 3.0
% Copyright by Trimborn, Koch, Steger
%
% COMMAND:
% [guess, start, errorcode] = initrelax(@funcODE, @funcSTAT, n,
%                                    n1, n3, nu, y, M, statev, t)
%
% INPUT ARGUMENTS:
% funcODE: function which returns the right hand side of the
%     differential equations whereas time and vector of variales are inputs
% funcSTAT: function which returns the residual of the static equations
%     whereas time and vector of variables are inputs
% n: number of differential equations
% n1: number of initial boundary conditions
% n3: (optional, default value is 0) number of static equations
% nu: (optional, default value is 0.1) parameter for time transformation
% y: (optional, default value is vector 1) initial guess. If no time
%     dependent guess is made,
%     this should be a column vector, representing a guess for the steady
%     state values of the variables (dimension N). If a time dependent
%     guess is made, this should be a matrix maching to t, such that
%     each column i is a guess for the variables at time t(i)
% M: (optional, default value is 50) number of mesh points
% statev: (optional, default value is 0) if statev=0, external file
%     will be used for initial boundary conditions. if
%     statev=ones(N,1), the initial steady state will be determined
%     via shock.m. if this is a vector of positiv numbers, the state
%     variables will be multiplied by the correspondent entry
% t: (optional) if a time vector is supplied, y should match for a
%     time dependent guess (see above)
%
% OUTPUT ARGUMENTS:
% guess: initial guess, suitable for relax.m
% start: vector which indicates treatment of initial boundary conditions
% errorcode: 0 no error
%            1 not enough input arguments
%            2 funcODE or funcSTAT return non real values or vectors in
%              wrong dimension
%            3 fsolve does not converge
%
% For further information contact Timo Trimborn, University
% of Hamburg
% or visit http://www.rrz.uni-hamburg.de/IWK/trimborn/relaxate.htm

N=n+n3; errorcode=0;
if nargin < 10,          %if vector t is supplied together with a
    preciseguess=0;      %matrix y, a precise, time dependent guess
```

```
else                        %will be constructed
    preciseguess=1;
end

if nargin < 9,
    statev=0;
    start=0;
    stateshock=0;
elseif statev'*statev==0;
    start=0;
    stateshock=0;
elseif sum((statev-1).^2)~=0
    stateshock=1;
    start=1;
else
    stateshock=0;
    start=1;
end

if nargin < 8, M=50; end if nargin < 7, y=ones(N,1); end if nargin <
6, nu=0.1; end if nargin < 5, n3=0; end if nargin < 4, errorcode=1;
end

%determination of guess for steady state values
if preciseguess==0
    finalguess=y;
else
    finalguess=y(:,M);
end

ode=feval(funcODE,inf,finalguess); if n3>0
    stat=feval(funcSTAT,inf,finalguess);
    res=[ode;stat]'*[ode;stat];
else
    res=ode'*ode;
    stat=[];
end

%Test, if funcODE and funcSTAT return real values
if ~(res<inf) | ~isreal(res)
    disp('ERROR: funcODE or funcSTAT return non real entries!');
    errorcode=2;
end

%Test, if funcODE and funcSTAT return vectors of right dimension
[test1 test2]=size(ode); [test3 test4]=size(stat); if
~((test1==n)&(test2==1)&(test3==n3)&(test4==1)) & n3>0
    disp('ERROR: funcODE or funcSTAT returns vector in wrong dimension!');
```

```
    errorcode=2;
elseif ~((test1==n)&(test2==1))
    disp('ERROR: funcODE returns vectors in wrong dimension!');
    errorcode=2;
end

if (sqrt(res) < 10^(-10)) & errorcode==0
    disp('Guess for steady state is sufficiently good.');
    disp(' ');
    if (stateshock==0) & (start==1)
        disp(['Calculation of initial steady state values...']);
        [start,fval,exitflag]=fsolve(@attach,finalguess,optimset('Display'...
         ,'off','MaxFunEvals',1000000,'MaxIter',100000),-1,funcODE,funcSTAT);
        if exitflag>0
            y=real(y);
            disp(['Convergence achieved.']);
            disp([' ']);
        else
            disp(['No convergence for initial steady state values!']);
            errorcode=3;
        end
    elseif stateshock==1 & start==1
        start=finalguess.*statev;
    end
elseif errorcode==0
    if ~(preciseguess==1 & start==0)
        disp(['Calculation of final steady state values...']);
        [finalguess,fval,exitflag]=fsolve(@attach,finalguess,optimset(...
            'Display','off','MaxFunEvals',1000000,'MaxIter',100000),inf,...
            funcODE,funcSTAT);
        if exitflag>0
            y=real(y);
            disp(['Convergence achieved.']);
            disp([' ']);
        else
            disp(['No convergence for final steady state values!']);
            errorcode=3;
        end
        if (stateshock==0) & (start==1) & (errorcode==0)
            disp(['Calculation of initial steady state values...']);
            [start,fval,exitflag]=fsolve(@attach,finalguess,optimset(...
                'Display','off','MaxFunEvals',1000000,'MaxIter',100000)...
                ,-1,funcODE,funcSTAT);
            if exitflag>0
                y=real(y);
                disp(['Convergence achieved.']);
                disp([' ']);
            else
```

```
                    disp(['No convergence for initial steady state values!']);
                    errorcode=3;
                end
            elseif stateshock==1  & start==1
                start=finalguess.*statev;
            end
        end
    end
end

guess=ones(M*N,1); if (preciseguess==0) & (errorcode==0)
    for i=1:M
        for ii=1:N
            guess((i-1)*N+ii)=finalguess(ii);
        end;
    end;
elseif errorcode==0
    tt=zeros(1,M);
    tau=[0:1/(M-1):1];
    for i=1:M
        if tau(i)~=1
            tt(i)=tau(i)/nu/(1-tau(i));
        else
            tt(i)=inf;
        end;
    end;
    y=releval(tt,t,y);
    for i=1:M
        guess((i-1)*N+1:i*N)=y(:,i);
    end
end

function attach=attach(x,t,funcODE,funcSTAT);
attach=[feval(funcODE,t,x);feval(funcSTAT,t,x)];
```

### 7.1.4 Code for relax.m

```
function [t, x] = relax(funcODE, funcSTAT, funcINI, funcfinal,...
n,n1, n3, nu, y, M, start, Endcond, maxit, tol, damp, dampfac)
% Version 3.0
% RELAXATION algorithm to solve infinite-horizon continuous time models.
%
% Copyright by Trimborn, Koch, Steger
%
% For further information contact Timo Trimborn, University of Hamburg
% or visit http://www.rrz.uni-hamburg.de/IWK/trimborn/relaxate.htm
%
% COMMAND:
% [t, x] = relax(@funcODE, @funcSTAT, @funcINI, @funcfinal, n, n1, n3, nu,
```

```
%              y, M, start, Endcond, maxit, tol, damp, dampfac)
%
% INPUT ARGUMENTS
% funcODE: function which returns the right hand side of the
%     differential equations whereas time and vector of variales are inputs
% funcSTAT: function which returns the residual of the static equations
%     whereas time and vector of variables are inputs
% funcINI: function which returns the residual of the initial
%     boundary conditions (In case you want to determine them
%     otherwise, nevertheless input @funcINI)
% funcfinal: function which returns the residual of the final
%     boundary conditions (In case you want to determine
%     them otherwise, nevertheless input @funcfinal)
% n: total number of differential equations
% n1: number of initial boundary conditions
% n3: (optional, default value is 0) number of static equations
% nu: (optional, default value is 0.1) parameter for time transformation
% y: (optional, default value is vector 1) vector of dimension M*Nx1,
%     guess of the time path of the variables. the first N
%     variables are the guess at time 0 and so on whereas the time
%     vector is determined through M and nu
% M: (optional, default value is 50) number of mesh points
% start: (optional, default is start=0) vector with the values of the
%     state variables at time 0. If input is start=0, external file
%     funcINI will be used to determine the initial boundary conditions
% Endcond: (optional, default is Endcond=0) vector with determines,
%     which differential equations will be used for constructing the
%     final boundary conditions. If input is Endcond=0, external file
%     funcfinal will be used to determine the final boundary conditions
% maxit: (optional, default is maxit=50) maximum number of iterations
% tol: (optional, default is tol=10^-9) tolerance for Newton-procedure
% damp: (optional, default is damp=1) Dampening factor of the Newton
%     procedure. The dampening factor will be multiplied by the factor
%     dampfac in every iteration until it equals 1. Therefore damp=1
%     means no dampening.
% dampfac: (optional, default is dampfac=2)
%
% OUTPUT ARGUMENTS:
% t: time vector
% x: solution vector of variables through time: column i are
%    the values at time t(i)

%****************************************************************************
%---------------------------------------------------------------------------
N=n+n3;                     %calculation of total dimension
if nargin < 16, dampfac=2; end              %default values
if nargin < 15, damp=1; end if nargin < 14, tol=10^-9; end if nargin
< 13, maxit=50; end if nargin < 12, Endcond=0; end if nargin < 11,
```

```
start=0; end if nargin < 10, M=50; end if nargin < 9, y=ones(N*M,1);
end if nargin < 8, nu=0.1; end if nargin < 7, n3=0; end
%The convergence criteria will be (deltay'.*deltay) normalized by the
%final steady state vector, therefore for components which are zero
%an increment has to be added to the final vector
crit=(y(end-N+1:end).*y(end-N+1:end)<eps*10)*10^-5;
crit=y(end-N+1:end)+crit;
%reset number of iterations
it=0;
%This vector contains the residuals from the nonlinear equations
E=zeros(N*M,1);
%This is a storing vector for the numerical calculation of the jacobian
fac=[];fac2=[];fac3=[];
%Later on, in this vector the sections of the jacobian matrix will be
%stored
Jac=zeros(N*(M-1),N); KonvKrit=tol+1;
%start of iteration++++++++++++++++++++++++++++++++++++++++++++++++++++++
while KonvKrit>tol & it<maxit
    if it==0; disp(['Start of main loop:']); end;
    it=it+1;

    %Input of initial conditions and first static equation in E
    if start==0
        E(1:n1)=feval(funcINI,0,y(1:N));
    else
        E(1:n1)=y(1:n1)-start(1:n1);
    end
    if n3>0
        E(n1+1:n1+n3)=feval(funcSTAT,0,y(1:N));
        %creation of the first part of the jacobian due to the initial
        %conditions and the first static equation
        [Jacstat,fac2]=numjac(funcSTAT,0,y(1:N),feval(funcSTAT,0,...
            y(1:N)),sqrt(eps)*ones(N,1),fac2,0);
    else
        Jacstat=[];
    end
    if start==0
        [Jacini,fac3]=numjac(funcINI,0,y(1:N),feval(funcINI,0,y(1:N)),...
            sqrt(eps)*ones(N,1),fac3,0);
        Jacini=Jacini(1:n1,:);
    else
        Jacini=[eye(n1) zeros(n1,N-n1)];
    end
    Jacsml=[Jacini ; Jacstat];
    for i=1:M-1
        %numerical evaluation of H, SL and SR
        %First, function H is evaluated
        E(n1+n3+(i-1)*N+1:n1+n3+i*N)=feval(@funcH,i/(M-1),...
```

```
            y(i*N+1:i*N+N),y((i-1)*N+1:(i-1)*N+N),(i-1)/(M-1),n,N,nu,...
            funcODE,funcSTAT);
    %Jacobian block SR is calculated
    [SRnum,fac]=numjac(@funcH,i/(M-1),y(i*N+1:i*N+N),feval(...
        @funcH,i/(M-1),y(i*N+1:i*N+N),y((i-1)*N+1:(i-1)*N+N),...
        (i-1)/(M-1),n,N,nu,funcODE,funcSTAT),...
        sqrt(eps)*ones(N,1),fac,0,[],[],y((i-1)*N+1:(i-1)*N+N),...
        (i-1)/(M-1),n,N,nu,funcODE,funcSTAT);
    %Jacobian block SL can be computed out of SR
    SLnum=SRnum-2*eye(size(SRnum));
    SLnum(N-n3+1:N,:)=0;
    %The block which has to be modified is composed
    Jactmp=[Jacsml zeros(size(Jacsml)); SLnum SRnum];
    %Transformation of actual part of the jacobian
    [Jactmp E((i-1)*N+1:i*N+n1+n3)]=feval(@rrefmod,Jactmp , ...
        E((i-1)*N+1:i*N+n1+n3));
    %relevant areas are stored
    Jac((i-1)*N+1:i*N,:)=Jactmp(1:N , N+1:2*N);
    %One part is stored for the next passage of the loop
    Jacsml=Jactmp(N+1:N+n1+n3,N+1:2*N);
end

%input of final boundary conditions in E
End=feval(funcODE,inf,y(N*M-N+1:N*M));
if n>n1 & Endcond~=0
    E(N*M-n+n1+1:N*M)=End(Endcond);
elseif n>n1
    E(N*M-n+n1+1:N*M)=feval(funcfinal,inf,y(N*M-N+1:N*M));
end

%Calculation of part of the jacobian due to final boundary conditions
if Endcond~=0 | n==n1
    [Jacend,fac3]=numjac(funcODE,inf,y(N*(M-1)+1:N*M),...
        feval(funcODE,inf,y(N*(M-1)+1:N*M)),sqrt(eps)*ones(N,1)...
        ,fac3,0);
else
    [Jacend,fac3]=numjac(funcfinal,inf,y(N*(M-1)+1:N*M),...
        feval(funcfinal,inf,y(N*(M-1)+1:N*M)),sqrt(eps)*ones(N,1)...
        ,fac3,0);
end

%Transforming final part of the jacobian
if n>n1 & Endcond~=0
    Jactmp=[Jacsml ; Jacend(Endcond,:)];
elseif n>n1
    Jactmp=[Jacsml ; Jacend];
else
    Jactmp=Jacsml;
```

```
    end
    [Jactmp E(N*(M-1)+1:N*M)]=feval(@rrefmod,Jactmp , E(N*(M-1)+1:N*M));

    %Calculation of correction vector and correction of last block
    deltay=E(N*M-N+1:N*M);
    y(N*M-N+1:N*M) = y(N*M-N+1:N*M) - damp*deltay;
    %Calculation of convergence criteria
    KonvKrit=sqrt((deltay./(crit))'*(deltay./(crit)));
    %Calculation of correction vector and correction of all other blocks
    for i=M-1:-1:1
        deltay=E((i-1)*N+1:i*N)-Jac((i-1)*N+1:i*N,:)*deltay;
        y((i-1)*N+1:i*N) = y((i-1)*N+1:i*N) - damp*deltay;
         %convergence criteria
        KonvKrit=KonvKrit+sqrt((deltay./(crit))'*(deltay./(crit)));
    end
    %Calculation of convergence criteria
    KonvKrit=KonvKrit/(M*N);
    if damp<1
        disp(['Iteration number: ',num2str(it),...
                ' convergence criteria: ',num2str(KonvKrit),...
                ' dampening factor: ',num2str(damp)]);
    else
        disp(['Iteration number: ',num2str(it),...
                ' convergence criteria: ',num2str(KonvKrit)]);
    end
    damp=min(dampfac*damp,1);
    y=real(y);
end;
%Test of convergence
if KonvKrit > tol | not(isfinite(KonvKrit))
    disp(['No convergence!']);
elseif KonvKrit~=0
    disp(['Convergence achieved.']);disp([' ']);
end
%End of iteration+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++

%Storing of variables in x and creation of actual time t
x=zeros(N,M); t=zeros(1,M); tau=[0:1/(M-1):1]; for i=1:M
    x(:,i)=y((i-1)*N+1:i*N);
    if tau(i)~=1
        t(i)=tau(i)/nu/(1-tau(i));
    else
        t(i)=inf;
    end;
end;


%***********************************************************************
%-----------------------------------------------------------------------
```

```
function funcH = funcH(taua,ya,yb,taub,n,N,nu,funcODE,funcSTAT)
%ya: variables at knot k+1, yb: variables at knot k
%taua: time at k+1, taub: time at k
%The residuals of the differential equations and the static equations
%are composed to one vector. The differential equation is tranformed
%into a difference equation (midpoint rule)

time=((taua+taub)/2); if taua<1
    funcH=[ya(1:n)-yb(1:n)-(taua-taub)*feval(funcODE,time/(nu*(1-time)),...
        (ya+yb)/2)*1/nu/(1-time)^2; ...
      feval(funcSTAT,taua/(nu*(1-taua)),ya);];
else
    funcH=[ya(1:n)-yb(1:n)-(taua-taub)*feval(funcODE,time/(nu*(1-time)),...
        (ya+yb)/2)*1/nu/(1-time)^2; ...
      feval(funcSTAT,inf,ya);];
end
```

## 7.1.5 Code for rrefmod.m

```
function [A,b] = rrefmod(A,b)
% RREFMOD    Reduced row echelon form.
%   form: [B , x] = RREFMOD(A,b)
%   The modified algorithm transforms the rows of the full rank
%   (m x n) matrix A (m<=n) such that B has the following shape:
%   B = IR with Identity matrix I and a (m x n-m) matrix R. In addition
%   the algorithm performs the same operations with the vector b. Output
%   is the transformed matrix B and corresponding vector x. For
%   transforming the rows the algorithm includes a column pivot search.
%
%   See RREF.

[m,n] = size(A); q=length(b); if q~=m
    disp('Warning: Wrong dimension of vector!');
end

% Compute the default tolerance if none was provided.
 tol = max(m,n)*eps*norm(A,'inf');

% Loop over the entire matrix.
i = 1; j = 1; while (i <= m) & (j <= n)
   % Find value and index of largest element in the remainder of column j.
   [p,k] = max(abs(A(i:m,j))); k = k+i-1;
   if (p <= tol)
      % The column is negligible, zero it out.
      A(i:m,j) = zeros(m-i+1,1);
      j = j + 1;
   else
```

```
        % Swap i-th and k-th rows.
        A([i k],j:n) = A([k i],j:n);
        b([i k])=b([k i]);
        % Divide the pivot row by the pivot element.
        b(i)=b(i)/A(i,j);
        A(i,j:n) = A(i,j:n)/A(i,j);
        % Subtract multiples of the pivot row from all the other rows.
        for k = [1:i-1 i+1:m]
            b(k)=b(k)-A(k,j)*b(i);
            A(k,j:n) = A(k,j:n) - A(k,j)*A(i,j:n);
        end
        i = i + 1;
        j = j + 1;
    end
end
```

## 7.2  Appendix on Section 5

**A. List of equations**

(Note: The time index has been dropped for simplicity.)

Euler equation

$$\frac{\dot{C}_i}{C_i} = \frac{(1 - \tau_i)\, r}{P^I} - \rho_i - \delta,\ i = 1, 2, .., 6$$

Composite private consumption

$$C_i = \Omega_i \prod_{j=1}^{9} c_{i,j}^{\theta_{i,j}},\ \Omega_i > 0,\ 0 < \theta_{i,j} < 1$$

$$P_i^C C_i = \sum_{j=1}^{9} pc_j c_{i,j}$$

$$\frac{c_{h,i}}{c_{h,j}} = \frac{\theta_{h,i} pc_j}{\theta_{h,j} pc_i},\ i, j = 1, 2, .., 9 \text{ and } h = 1, 2, ,.., 6$$

Consumption prices

$$P_i^C = \frac{1}{\Omega_i} \prod_{j=1}^{9} \left( \frac{pc_j}{\theta_{i,j}} \right)^{\theta_{ij}}$$

Private consumption demand functions

$$c_{i,j} = \theta_{i,j} \frac{P_i^C C_i}{pc_j}$$

The same consumption demand function applies to government consumption $G$ and investment $I$.

Labor demand functions

$$L_{i,j} = (A_j)^{(\sigma_j - 1)} V A_j \left( \frac{\alpha_{i,j} P_j^{VA}}{w_i} \right)^{\sigma_j}$$

Capital demand

$$K_j = (A_j)^{(\sigma_j - 1)} V A_j \left[ \frac{\left( 1 - \sum\limits_{i=1}^{6} \alpha_{i,j} \right) P_j^{VA}}{r} \right]^{\sigma_j}$$

Value-added price

$$P_j^{VA} = \frac{1}{A_j} \left[ \sum_{i=1}^{6} (w_i)^{(1-\sigma_j)} (\alpha_{i,j})^{\sigma_j} + r^{(1-\sigma_j)} \left( 1 - \sum_{i=1}^{6} \alpha_{ij} \right)^{\sigma_j} \right]^{\frac{1}{1-\sigma_j}}$$

CES Armington function

$$X_i = \Phi_i \left[ \varepsilon_i (M_i)^{\frac{\gamma_i - 1}{\gamma_i}} + (1 - \varepsilon_i)(D_i)^{\frac{\gamma_i - 1}{\gamma_i}} \right]^{\frac{\gamma_i}{\gamma_i - 1}}$$
$$\Phi_i > 0, \ 0 < \varepsilon_i < 1, \ \gamma_i > 0, \ \gamma_i \neq 1, \ i = 1, 2, ..., 9$$

$$P_i^X X_i = P_i^M M_i + \left( 1 + vat_i^D \right) P_i^D D_i$$

$$\frac{D_i}{M_i} = \left[ \frac{(1 - \varepsilon_i) P_i^M}{\varepsilon_i \left( 1 + vat_i^D \right) P_i^D} \right]^{\gamma_i}$$

Import demand function

$$M_i = (\Phi_i)^{(\gamma_i - 1)} X_i \left( \frac{\varepsilon_i P_i^X}{P_i^M} \right)^{\gamma_i}$$

Domestic good demand function

$$D_i = (\Phi_i)^{(\gamma_i - 1)} X_i \left[ \frac{(1 - \varepsilon_i) P_i^X}{\left( 1 + vat_i^D \right) P_i^D} \right]^{\gamma_i}$$

Composite CES Armington price

$$P_i^X = \frac{1}{\Phi_i} \left\{ \left( P_i^M \right)^{(1-\gamma_i)} (\varepsilon_i)^{\gamma_i} + \left[ \left( 1 + vat_i^D \right) P_i^D \right]^{(1-\gamma_i)} (1 - \varepsilon_i)^{\gamma_i} \right\}^{\frac{1}{1-\gamma_i}}$$

Cobb-Douglas total imports

$$M_i = \Phi_i^M \left(M_i^{EU}\right)^{\varepsilon_i^{EU}} \left(M_i^{RW}\right)^{\varepsilon_i^{RW}}$$
$$\Phi_i^M > 0,\ 0 < \varepsilon_i^{EU}, \varepsilon_i^{RW} < 1,\ \varepsilon_i^{EU} + \varepsilon_i^{RW} = 1,\ i = 1, 2, .., 9$$

$$P_i^M M_i = PM_i^{EU} M_i^{EU} + PM_i^{RW} M_i^{RW}$$

$$\frac{M_i^{EU}}{M_i^{RW}} = \frac{\varepsilon_i^{EU} PM_i^{RW}}{\varepsilon_i^{RW} PM_i^{EU}}$$

Regional import demand functions

$$M_i^j = \varepsilon_i^j \frac{P_i^M M_i}{PM_i^j},\ i = 1, 2, .., 9,\ j = EU, RW$$

Import composite price

$$P_i^M = \frac{1}{\Phi_i^M} \left(\frac{PM_i^{EU}}{\varepsilon_i^{EU}}\right)^{\varepsilon_i^{EU}} \left(\frac{PM_i^{RW}}{\varepsilon_i^{RW}}\right)^{\varepsilon_i^{RW}}$$

Import prices

$$PM_i^j = PWM_i \left(1 + tm_i^j\right) \left(1 + vat_i^M\right),\ j = EU, RW$$

CET function

$$Q_i = \chi_i \left[\lambda_i \left(E_i\right)^{\frac{1+\Psi_i}{\Psi_i}} + \left(1 - \lambda_i\right) \left(D_i\right)^{\frac{1+\Psi_i}{\Psi_i}}\right]^{\frac{\Psi_i}{1+\Psi_i}}$$
$$\chi_i > 0,\ 0 < \lambda_i < 1,\ \Psi_i > 0,\ i = 1, 2, ..., 9$$

$$P_i^Q Q_i = P_i^E E_i + P_i^D D_i$$

$$\frac{D_i}{E_i} = \left[\frac{\lambda_i P_i^D}{\left(1 - \lambda_i\right) P_i^E}\right]^{\Psi_i}$$

Export supply function

$$E_i = \frac{Q_i}{(\chi_i)^{(1+\Psi_i)} \left(P_i^Q\right)^{\Psi_i}} \left(\frac{P_i^E}{\lambda_i}\right)^{\Psi_i}$$

Domestic good supply function

$$D_i = \frac{Q_i}{(\chi_i)^{(1+\Psi_i)} \left(P_i^Q\right)^{\Psi_i}} \left(\frac{P_i^D}{1-\lambda_i}\right)^{\Psi_i}$$

Composite output price

$$P_i^Q = \frac{1}{\chi_i} \left[\frac{\left(P_i^E\right)^{(1+\Psi_i)}}{(\lambda_i)^{\Psi_i}} + \frac{\left(P_i^D\right)^{(1+\Psi_i)}}{(1-\lambda_i)^{\Psi_i}}\right]^{\frac{1}{1+\Psi_i}}$$

CET composite exports

$$E_i = \chi_i^E \left[\lambda_i^{EU} \left(E_i^{EU}\right)^{\frac{1+\Psi_i^E}{\Psi_i^E}} + \lambda_i^{RW} \left(E_i^{RW}\right)^{\frac{1+\Psi_i^E}{\Psi_i^E}}\right]^{\frac{\Psi_i^E}{1+\Psi_i^E}}$$
$$\chi_i^E > 0, \ 0 < \lambda_i^{EU}, \lambda_i^{RW} < 1, \ \lambda_i^{EU} + \lambda_i^{RW} = 1, \ > 0, \ i = 1, 2, , ..., 9$$

$$P_i^E E_i = PE_i^{EU} E_i^{EU} + PE_i^{RW} E_i^{RW}$$

$$\frac{E_i^{EU}}{E_i^{RW}} = \left(\frac{\lambda_i^{RW} PE_i^{EU}}{\lambda_i^{EU} PE_i^{RW}}\right)^{\Psi_i^E}$$

Export supply functions

$$E_i^j = \frac{E_i}{\left(P_i^E\right)^{\Psi_i^E} \left(\chi_i^E\right)^{\left(1+\Psi_i^E\right)}} \left(\frac{P_i^E}{\lambda_i^j}\right)^{\Psi_E}$$

Export composite price

$$P_i^E = \frac{1}{\chi_i^E} \left[\frac{\left(P_i^E\right)^{1+\Psi_i^E}}{\left(\chi_i^E\right)^{\Psi_i^E}} + \frac{\left(P_i^D\right)^{1+\Psi_i^E}}{\left(\chi_i^D\right)^{\Psi_i^E}}\right]^{\frac{1}{1+\Psi_i^E}}$$

Export prices

$$PE_i^j = PWE_i, \; j = EU, RW$$

VAT on domestic goods

$$VAT^D = \sum_{i=1}^{9} vat_i^D P_i^D D_i$$

VAT on imports

$$VAT^M = \sum_{i=1}^{9} \sum_{j=EU,RW} vat_i^M \left(1 + tm_i^j\right) PWM_i M_i^j$$

Import duties

$$TM = \sum_{i=1}^{9} \sum_{j=EU,RW} tm_i^j PWM_i M_i^j$$

Income tax

$$TY = \sum_{i=1}^{6} \tau_i (w_i L_i + rK_i + GT_i + FT_i)$$

Government transfer to households

$$TR = \sum_{i=1}^{6} GT_i$$

## B. Glossary

$C_i$:      total consumption of household $i$

$P_i^C$:      consumption price (index) of household $i$

$\tau_i$:      income tax rate applying to household $i$

$\rho_i$:      household $i$'s discount rate

$\Omega_i$:      shift parameter in the Cobb-Douglas consumption function of household $i$

$c_{i,j}$:      household $i$'s consumption of good $j$

$pc_j$:      price of consumption good $j$

$\theta_{i,j}$:      share parameter in the Cobb-Douglas function of household $i$ for good $j$

$I$:      aggregate investment

$P^I$:      price index of aggregate investment

$G$:      aggregate government consumption

$P^G$:      price index of aggregate government consumption

$\delta$:      deprecation rate of capital

$L_{i,j}$:      sector $j$'s demand for labor of type $i$

$K_j$:      sector $j$'s demand for capital

$A_j$:      shift parameter of the value-added production function in sector $j$

$\sigma_j$:      elasticity of substitution between primary inputs in sector $j$

$\alpha_{i,j}$:      share parameter of labor of type $i$ used in sector $j$

$VA_j$:      sector $j$'s value-added production

$P_j^{VA}$:      sector $j$'s value-added price

$w_i$:      nominal wage rate of labor of type $i$

$r$:      nominal return to capital

$X_i$:      domestic absorption of sector $i$

$M_i$:      total imports of sector $i$

$D_i$:      domestic production of sector $i$

$\Phi_i$:      shift parameter in the CES Armington function of sector $i$

$\varepsilon_i$:          imports share parameter in the CES Armington
                  function of sector $i$

$\gamma_i$:          sector $i$'s elasticity of substitution between imports
                  and domestically-produced output

$P_i^X$:          composite price of domestic absorption of sector $i$

$P_i^M$:          import price of sector $i$

$P_i^D$:          price of sector $i$'s domestically-produced good

$vat_i^D$:          VAT rate on sector $i$'s domestically-produced good

$M_i^j$:          imports of sector $i$ from region $j$

$\Phi_i^M$:          shift parameter in the imports CES function of sector $i$

$\varepsilon_i^j$:          region $j$'s share parameter in the imports
                  CES function of sector $i$

$tm_i^j$:          import tax rate applying to sector $i$'s imports from region $j$

$vat_i^M$:          VAT rate on sector $i$'s imported goods

$PWM_i$:          sector $i$'s world price of imports

$Q_i$:          total output of sector $i$

$P_i^Q$:          composite output price of sector $i$

$E_i$:          total exports of sector $i$

$P_i^E$:          export price of sector $i$

$\chi_i$:          shift parameter in the CET function of sector $i$

$\lambda_i$:          export share parameter of sector $i$

$\Psi_i$:          elasticity of transformation between exports and
                  domestically-sold output of sector $i$

$E_i^j$:          exports of sector $i$ to region $j$

$\chi_i^E$:          shift parameter in the CET exports function of sector $i$

$\lambda_i^j$:          share parameter of exports to region $j$ in sector $i$

$\Psi_i^E$:          elasticity of transformation between exports to different
                  regions of sector $i$

$PE_i^j$:          price of exports to region $j$ of sector $i$

$PWE_i$:          world price of exports of sector $i$

# References

ABED, G. T. (1998): "Trade Liberalization and Tax Reform in the Southern Mediterranean Countries," *IMF Working Paper*, 49.

ALEMDAR, N. M., S. SIRAKAYA, AND F. HÜSSEINOV (2006): "Optimal Time Aggregation of Infinite Horizon Control Problems," *Journal of Economic Dynamics and Control*, 30, 569–593.

ARMINGTON, P. S. (1969): "A Theory of Demand for Products Distinguished by Place of Production," *IMF Staff Papers*, 16, 159–176.

ARNOLD, L. G. (2006): "The Dynamics of the Jones R&D Growth Model," *Review of Economic Dynamics*, 9, 143–152.

ASCHER, U. M., R. M. M. MATTHEIJ, AND R. D. RUSSELL (1985): *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*. Englewood Cliffs, New York: Prentice Hall.

ASCHER, U. M., AND L. R. PETZOLD (1998): *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*. Philadelphia : Society for Industrial and Applied Mathematics.

ASCHER, U. M., AND R. D. RUSSELL (1981): "Reformulation of boundary value problems into 'standard' form," *SIAM Review*, 23, 238–254.

AUERNHEIMER, L., AND G. A. LOZADA (1990): "On the Treatment of Anticipated Shocks in Models of Optimal Control with Rational Expectations: An Economic Interpretation," *American Economic Review*, 80(1), 157–169.

BARRO, R. J. (1974): "Are Government Bonds Net Wealth?," *Journal of Political Economy*, 82(6), 1095–1117.

——— (1991): "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 106, 407–443.

BARRO, R. J., AND X. SALA-I-MARTÍN (1992): "Convergence," *Journal of Political Economy*, 100, 223–251.

——— (2004): *Economic Growth*. MIT-Press, Cambridge, Massachusetts.

BAUMOL, W. J. (1986): "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show," *American Economic Review*, 76(5), 1072–1085.

BECKER, R. A. (1980): "On the Long-Run Steady State in a Simple Dynamic Model of Equilibrium with Heterogeneous Households," *Quarterly Journal of Economics*, 95(2), 375–382.

BECKER, R. A., AND E. N. TSYGANOV (2002): "Ramsey Equilibrium in a Two-Sector Model with Heterogenous Households," *Journal of Economic Theory*, 105, 188–225.

BELLMAN, R. (1957): *Dynamic Programming*. Princeton University Press, Princeton.

BENHABIB, J., AND R. PERLI (1994): "Uniqueness and Indeterminacy: On the Dynamics of Endogenous Growth," *Journal of Economic Theory*, 63, 113–142.

BOUCEKKINE, R., O. LICANDRO, AND C. PAUL (1997): "Differential-Difference Equations in Economics: On the Numerical Solution of Vintage Capital Growth Models," *Journal of Economic Dynamics and Control*, 21, 347–362.

BOURGUIGNON, F., W. H. BRANSON, AND J. DE MELO (1992): "Adjustment and Income Distribution: A Micro-Macro Model for Counterfactual Analysis," *Journal of Development Economics*, 38, 17–39.

BRUNNER, M., AND H. STRULIK (2002): "Solution of Perfect Foresight Saddlepoint Problems: A Simple Method and Applications," *Journal of Economic Dynamics and Control*, 26, 737–753.

BRYSON, A. E., AND Y.-C. HO (1975): *Applied Optimal Control: Optimization, Estimation, and Control.* Hemisphere Publishing Corporation, Washington, DC.

BUITER, W. H. (1984): "Saddlepoint Problems in Continuous Time Rational Expectation Models: A General Method and Some Macroeconomic Examples," *Econometrica*, 52(3), 665–680.

CABALLE, J., AND M. S. SANTOS (1993): "On Endogenous Growth with Physical and Human Capital," *The Journal of Political Economy*, 101(6), 1042–1067.

CANDLER, G. V. (1999): "Finite-Difference Methods for Continuous-time Dynamic Programming," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott. Oxford University Press, Oxford.

CASELLI, F., AND J. VENTURA (2000): "A Representative Consumer Theory of Distribution," *American Economic Review*, 90, 909–926.

CASS, D. (1965): "Optimum Growth in an Aggregate Model of Capital Accumulation," *Review of Economic Studies*, 32, 233–240.

CHATTERJEE, S. (1994): "Transitional Dynamics and the Distribution of Wealth in a Neoclassical Growth Model," *Journal of Public Economics*, 54, 97–119.

CHIANG, A. C. (1992): *Elements of Dynamic Optimization.* McGraw-Hill, New York.

COCKBURN, J. (2001): "Trade Liberalization and Poverty in Nepal: A Computable General Equilibrium Micro Simulation Analysis," *CREFA Working Paper*, 01-18.

COGNEAU, D., AND A.-S. ROBILLIARD (2000): "Growth Distribution and Poverty in Madagascar: Learning from a Microsimulation Model in a General Equilibrium Framework," *IFPRI Working Paper, Trade and Macroeconomic Division.*

COLLATZ, L. (1951): *Numerische Behandlung von Differentialgleichungen.* Springer.

DE LONG, J. B. (1988): "Productivity Growth, Convergence, and Welfare: Comment," *American Economic Review*, 78(5), 1138–1154.

DECALUWÉ, B., A. PATRY, L. SAVARD, AND E. THORBECKE (1999): "Poverty Analysis within a General Equilibrium Framework," *CREFA Working Paper*, 9909.

DEUFLHARD, P., AND A. HOHMANN (1993): *Numerische Mathematik I.* Walter de Gruyter, Berlin.

DEVARAJAN, S., AND D. GO (1998): "The Simplest Dynamic General-Equilibrium Model of an Open Economy," *Journal of Policy Modelling*, 20(6), 677–714.

DEVARAJAN, S., J. D. LEWIS, AND S. ROBINSON (1990): "Policy Lessons from Trade-Focused, Two-Sector Models," *Journal of Policy Modelling*, 12(4), 625–657.

EICHER, T. S., AND S. J. TURNOVSKY (1999): "Convergence in a Two-Sector Nonscale Growth Model," *Journal of Economic Growth*, 4, 413–428.

———— (2001): "Transitional Dynamics in a Two Sector Model," *Journal of Economic Dynamics and Control*, 25, 85–113.

FERABOLI, O. (2007): *A Dynamic General Equilibrium Analysis of Jordan's Trade Liberalisation.* PhD. dissertation, University of Hamburg.

FERABOLI, O., B. LUCKE, AND B. GAITAN SOTO (2003): "Trade Liberalisation and the Euro-Med Partnership: A Dynamic Model for Jordan," *Discussion Paper, University of Hamburg.*

# References

FERABOLI, O., AND T. TRIMBORN (2006): "Trade Liberalization and Income Distribution: A CGE Model for Jordan," *Discussion Paper, University of Hamburg.*

FOX, L. (1957): *Numerical Solution of Two-Point Boundary Value Problems in Ordinary Differential Equations.* Oxford University Press, Oxford, UK.

FUNKE, M., AND H. STRULIK (2000a): "On Endogenous Growth with Physical Capital, Human Capital and Product Variety," *European Economic Review*, 44, 491–515.

———— (2000b): "Transitional Dynamics in Funke and Strulik (2000): A Qualification," *www.rrz.uni-hamburg.de/wst/strulik/fsextra.pdf.*

GIBSON, B. (2005): "The Transition to a Globalized Economy: Poverty, Human Capital and the Informal Sector in a Structuralist CGE Model," *Journal of Development Economics*, 78(1), 60–94.

GINI, C. (1912): *Variabilità e mutabilità.* Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T), Rome: Libreria Eredi Virgilio Veschi (1955).

HARRISON, G. W., T. F. RUTHERFORD, AND D. G. TARR (2002): "Trade Policy Options for Chile: The Importance of Market Access," *World Bank Economic Review*, 16(1), 49–79.

HIRSCH, M. W., C. C. PUGH, AND M. SHUB (1977): *Invariant Manifolds.* Springer-Verlag, New York.

HOSOE, N. (2001): "A General Equilibrium Analysis of Jordan's Trade Liberalization," *Journal of Policy Modeling*, 23, 595–600.

References

HOWITT, P., AND H.-W. SINN (1989): "Gradual Reforms of Capital Income Taxation," *American Economic Review*, 79(1), 106–124.

INTRILIGATOR, M. D. (1971): *Mathematical Optimization and Economic Theory*. Prentice-Hall, Inc., Englewood Cliffs, New York.

JONES, C. I. (1995a): "R&D-Based Models of Economic Growth," *Journal of Political Economy*, 103(3), 759–784.

——— (1995b): "Time series tests of endogenous growth models," *Quarterly Journal of Economics*, 110, 495–527.

——— (1997): "On the Evolution of the World Income Distribution," *Journal of Economic Perspectives*, 11(3), 19–36.

JUDD, K. L. (1985): "Short-Run Analysis of Fiscal Policy in a Simple Perfect Foresight Model," *Journal of Political Economy*, 93, 298–315.

——— (1992): "Projection Methods for Solving Aggregate Growth Models," *Journal of Economic Theory*, 58, 410–452.

——— (1998): *Numerical Methods in Economics*. MIT Press.

——— (2002): "The Parametric Path Method: An Alternative to Fair–Taylor and L–B–J for Solving Perfect Foresight Models," *Journal of Economic Dynamics and Control*, 26, 1557–1583.

JUILLARD, M., D. LAXTON, P. MCADAM, AND H. PIORO (1998): "An algorithm competition: First-order iterations versus Newton-based techniques," *Journal of Economic Dynamics and Control*, 22(8/9), 1291–1319.

KAMIEN, M. I., AND N. L. SCHWARTZ (1991): *Dynamic Optimization.* 2nd edition, North Holland.

KELLER, H. (1968): *Numerical Methods for Two-Point Boundary Value Problems.* Blaisdell Publishing Company, Waltbam, Massachusets.

KOCH, K.-J. (2003): "Relaxation," *Mimeo.*

KOOPMANS, T. C. (1965): "On the Concept of Optimal Economic Growth," In: *The Econometric Approach to Development Planning,* Amsterdam: North Holland.

KUTTA, M. W. (1901): "Beitrag zur näherungsweisen Integration totaler Differentialgleichungen," *Zeitschrift für Mathematik und Physik*, 46, 435–453.

KÖNIGSBERGER, K. (1997): *Analysis 2.* Springer Verlag, Berlin.

LAFFARGUE, J.-P. (1990): "Résolution d'un modèle économique avec anticipations rationelles," *Annales d'Economie et Statistique*, 17, 97–119.

LI, W., M. RYCHLIK, F. SZIDAROVSKY, AND C. CHIARELLA (2003): "On the Attractivity of a Class of Homogeneous Dynamic Economic Systems," *Nonlinear Analysis*, 52, 1617–1636.

LIPTON, D., J. POTERBA, J. SACHS, AND L. SUMMERS (1982): "Multiple shooting in rational expectation models," *Econometrica*, 50, 1329–1333.

LÖFGREN, H. (1999): "Trade Reform and the Poor in Morocco: A Rural-Urban General Equilibrium Analysis of Reduced Protection," *IFPRI Working Paper, Trade and Macroeconomic Division*, 38.

LUCAS, R. E. JR.. (1988): "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3–42.

——— (1993): "Making a Miracle," *Econometrica*, 61(2), 251–272.

LUCKE, D. (2001): "Fiscal Impact of Trade Liberalization: The Case of Jordan," *FEMISE Research Programme Final Report*.

MCGRATTAN, E. R. (1996): "Solving the Stochastic Growth Model with Finite Element Method," *Journal of Economic Dynamics and Control*, 20, 19–42.

——— (1999): "Application of Weighted Residual Methods to Dynamic Economic Models," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott. Oxford University Press, Oxford.

MERCENIER, J., AND P. MICHEL (1994): "Discrete-Time Finite Horizon Approximation of Infinite Horizon Optimization Problems with Steady-State Invariance," *Econometrica*, 62(3), 635–656.

——— (2001): "Temporal Aggregation in a Multi-Sector Economy with Endogenous Growth," *Journal of Economic Dynamics and Control*, 25, 1179–1191.

MULLIGAN, C., AND X. SALA-I-MARTIN (1991): "A Note on the Time Elimination Method for Solving Recursive Growth Models," *Technical Working Paper No. 116*.

——— (1993): "Transitional Dynamics in Two-Sector Models of Endogenous Growth," *Quarterly Journal of Economics*, 108(3), 737–773.

ORTINGUEIRA, S., AND M. S. SANTOS (1997): "On the speed of convergence in endogenous growth models," *American Economic Review*, 87, 383–399.

OULTON, N. (1993): "Widening the Human Stomach: The Effect of New Consumer Goods on Economic Growth and Leisure," *Oxford Economic Papers*, 51, 577–594.

PONTRYAGIN, L. S., V. G. BOLTYANSKII, R. V. GAMKRELIDZE, AND E. F. MISHCHENKO (1962): *The Mathematical Theory of Optimal Processes.* Wiley and Sons, New York.

PRESCOTT, E. C. (1986): "Theory ahead of business-cycle measurement," *Federal Reserve Bank of Minneapolis Quarterly Review*, 10, 9–22.

PRESS, W., B. FLANNERY, S. TEUKOLSKY, AND W. VETTERLING (1989): *Numerical Recipes in Pascal.* Cambridge University Press.

PRITCHET, L. (1997): "Divergence, Big Time," *Journal of Economic Perspectives*, 11(3), 3–17.

QUAH, D. T. (1996): "Twin Peaks: Growth and Convergence in Models of Distribution Dynamics," *Economic Journal*, 106, 1045–1055.

——— (1997): "Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs," *Journal of Economic Growth*, 2, 27–59.

RAMSEY, F. (1928): "A Mathematical Theory of Saving," *Economic Journal*, 38, 543–559.

REIMER, J. J. (2002): "Estimating the Poverty Impacts of Trade Liberalization," *Policy Research Working Paper Series 2790, The World Bank.*

ROBERTSON, P. E. (1999): "Economic Growth and the Return to Capital in Developing Economies," *Oxford Economic Papers*, 51, 269–283.

ROMER, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, S71–S102.

RUNGE, C. D. (1895): "Über die numerische Auflösung von Differentialgleichungen," *Mathematische Annalen*, 46, 167–178.

SALA-I-MARTÍN, X. (1997): "I Just Ran Two Million Regressions," *American Economic Review*, 87, 178–183.

————— (2006): "The World Distribution of Income: Falling Poverty and . . . Convergence, Period," *Quarterly Journal of Economics*, CXXI(2), 351–397.

SALA-I-MARTÍN, X., G. DOPPELHOFER, AND R. I. MILLER (2004): "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review*, 94(4), 813–835.

SEIERSTAD, A., AND K. SYDSÆTER (1987): *Optimal Control Theory with Economic Applications*. North-Holland, Amsterdam.

STEGER, T. M. (2005): "Welfare Implications of Non-Scale R&D-based Growth Models," *Scandinavian Journal of Economics*, 107(4), 737–757.

STRULIK, H. (2004): "Solving Rational Expectations Models Using Excel," *Journal of Economic Education*, 35, 269–283.

TAYLOR, J. B., AND H. UHLIG (1990): "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods," *Journal of Business and Economic Statistics*, 8, 1–18.

TEMPLE, J. (2003): "The Long-Run Implications of Growth Theories," *Journal of Economic Surveys*, 17(3), 497–510.

THEIL, H. (1967): *Economic and Information Theory*. Amsterdam: North Holland.

TRIMBORN, T. (2006): "A Numerical Procedure for Simulating Models of Endogenous Growth," *Discussion Paper, University of Hamburg.*

——— (2007): "Anticipated Shocks in Continuous-time Optimization Models: Theoretical Investigation and Numerical Solution," *Discussion Paper, University of Hannover*, 363.

TRIMBORN, T., K.-J. KOCH, AND T. M. STEGER (2008): "Multi-Dimensional Transitional Dynamics: A Simple Numerical Procedure," *Macroeconomic Dynamics*, 12(3), 1–19.

TU, P. N. (1994): *Dynamical Systems, An Introduction with Applications in Economics and Biology.* Springer-Verlag, Berlin.

TURNOVSKY, S. J. (1996): *Methods of Macroeconomic Dynamics.* MIT Press, Second Edition, Camebridge, Massachusetts.

VIND, K. (1967): "Control Systems with Jumps in the State Variables," *Econometrica*, 35(2), 273–277.

WILSON, C. A. (1979): "Anticipated Shocks and Exchange Rate Dynamics," *Journal of Political Economy*, 87, 639–647.

WINTERS, L. A., N. MCCULLOCH, AND A. MCKAY (2004): "Trade Liberalization and Poverty: The Evidence so far," *Journal of Economic Literature*, XLII, 72–115.