b
UNIVERSITÄT
BERN

Faculty of Business, Economics
and Social Sciences

Department of Economics

Time –Varying Rational Expectations Models: Solutions, Stability, Numerical Implementation

Klaus Neusser

17-01

September 2017

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Time –Varying Rational Expectations Models: Solutions, Stability, Numerical Implementation

Klaus Neusser *

September 13, 2017

Abstract

While rational expectations models with time–varying (random) coefficients have gained some esteem, the understanding of their dynamic properties is still in its infancy. The paper adapts results from the theory of random dynamical systems to solve and analyze the stability of rational expectations models with time–varying (random) coefficients. This theory develops a "linear algebra" in terms of Lyapunov exponents defined as the asymptotic growth rates of trajectories. They replace the eigenvalue analysis used in constant coefficient models and allow the construction of solutions in the spirit of Blanchard and Kahn (1980). The usefulness of these methods and their numerical implementation is illustrated using a canonical New Keynesian model with a time–varying policy rule.

JEL classification: C02, C61
Keywords: time–varying rational expectations models, random dynamical systems, Lyapunov exponents, multiplicative ergodic theorem

*I thank Dirk Niepelt and Roger Farmer for perceptive comments.
Department of Economics, University of Bern, Schanzeneckstrasse 1, P.O. Box 8573, CH-3001 Berne, Switzerland. Email: klaus.neusser@vwi.unibe.ch

1 Introduction

Dynamical systems with time–varying (random) coefficients receive an increasing attention in the realm of theoretical macroeconomic modelling (see the pioneering papers by Farmer, Waggoner, and Zha (2009) and Farmer, Waggoner, and Zha (2011) and the recent contribution by Barthélemy and Marx (2017)) as well as in the time series literature where, starting with the groundbreaking work of Hamilton (1989, 2016) and Kim and Nelson (1999), regime–switching models have established themselves as a standard modeling tool. Reasons for the variation in the model coefficients are manifold. First, time–varying coefficient models arise from the linearization of nonlinear models along solution paths (see Elaydi, 2005, p.219–220).[1] Second, the relationships describing the economy undergo structural changes resulting in drifting coefficients as emphasized by Lucas's critique.[2] Third, policies and policy rules are subject to change. Systematic empirical evidence with regard to U.S. monetary policy has been presented by Cogley and Sargent (2005), Primiceri (2005), and Chen, Leeper, and Leith (2015) among many others. While the causes for time–varying coefficients are quite convincing, the widespread use of these models is, however, hindered by the fact that the analysis of their dynamical properties requires a higher mathematical sophistication which goes well beyond the standard tools of linear algebra. For example, it is well–known that a binary regime switching model, i.e. a model where the coefficients switch between two alternative values, may exhibit explosive behavior despite the fact that in each regime considered separately the eigenvalues of the system matrix suggest a stable behavior (see Elaydi (2005), Francq and Zakoïan (2001), and in particular Appendix A for examples and an elaboration of this point). Another peculiar feature is that the scatter plot of the simulated endogenous variables has a fractal geometry (see Berger (1993) and Diaconis and Freedman (1999)).[3]

Based on the widely acclaimed work of Oseledets (1968) and his Multiplicative Ergodic Theory (MET), the theoretical mathematical literature on random dynamical systems has developed in the last decades a "linear algebra" which allows the computation of explicit solution formulas and the analysis of stability properties.[4] The fundamental concept around which this

[1]This also called the variation equation.

[2]An interpretation of drifting parameters in terms of self–confirming equilibria is provided by Sargent (1999).

[3]This feature actually provides a probabilistic algorithm for image creation, encoding and compression (Berger, 1993, pp.157).

[4]This work builds on the analysis of products of random matrices initiated by Bellman

theory is built are the Lyapunov exponents. These exponents measure the asymptotic growth rate of the trajectories and play a similar role as the eigenvalues in models with constant coefficients. The MET thus provides the right substitute for the spectral theorem which is the basis for analyzing rational expectations models with constant coefficients. The books by Arnold (2003) and Colonius and Kliemann (2014) present a clear exposition of the relevant literature and served as a source of inspiration for this paper.

While the implications of the MET are rather straightforward, the practical implementation is not. This is due the fact that there are in general no analytical solution formulas available so that one has to resort to numerical simulations. These simulations can quickly hit the numerical capabilities of any computer because some of the trajectories may exhibit explosive behavior. Thus, one has to resort to algorithms which at each step factorizes the system matrix in such a way that the simulations of the model become numerically stable (see Dieci and Elia (2008) and Froyland et al. (2013) for details).

The goal of this paper is demonstrate the usefulness of the MET for solving and analyzing affine rational expectations models with randomly changing coefficients. As it turns out, the general solution of such models can be formulated analogously to Blanchard and Kahn (1980) and Klein (2000), the only difference being that the role of the eigenvalues will be replaced the Lyapunov exponents. In this sense the methods presented in this paper can be interpreted as generalizing the standard approach with constant coefficients to one with time–varying (random) coefficients. The reliance on Lyapunov exponents, i.e. on asymptotic growth rates of trajectories, brings the paper also close in spirit to Sims' approach (Sims, 2001).

Having presented the general theory, we illustrate the usefulness of the proposed methods by applying them to a canonical New Keynesian model where the policy rule switches (deterministically or randomly) between a simple Taylor rule and a policy rule which takes the path of the nominal interest rate as exogenously given. The consequences of these two alternative rules have been analyzed by Galí (2011). He shows that when the latter rule is adopted the model becomes indeterminate, allowing for a multiplicity of solutions. Galí also examines the possibility that the central bank reverts to the Taylor rule after having fixed the interest rate for given number of periods. However, he does not consider the possibility of systematically or randomly switching between the two policies. The methods proposed in this paper will allow a systematic treatment of these cases and pave the way for

(1954) and, in particular, Furstenberg and Kesten (1960).

a deeper understanding.

There is a related literature which deals with the existence, uniqueness and properties of solutions to linear stochastic difference equations with random changing coefficients. This literature started with the seminal paper by Brandt (1986) which established moment conditions for the existence of stationary solutions. Bougerol and Picard (1992) extended these results to multivariate stochastic processes and Goldie and Maller (2000) give an almost complete characterization. Finally, Francq and Zakoïan (2001) provided a necessary and sufficient criterion for the existence of weakly stationary and causal solution in the context of multivariate Markov-switching ARMA models. Further results and economic applications along these lines have been presented by Bhattacharya and Majumdar (2007).

The paper proceeds by first discussing as a way of introduction and attunement the case of a periodic deterministic system in Section 2. This setup can be dealt within the context of Floquet's theory and allows to gain some basic insights which will also be relevant for the stochastic case. This theory is illustrated in Sections 3 by analyzing a canonical New Keynesian model with switching policy rules. We then generalize to the stochastic case in Section 4 and examine the New Keynesian model again in Section 5. Finally, we draw some conclusions for further applications.

2 Periodic Linear Deterministic Systems

As a way of introduction, we first examine affine rational expectations model under perfect foresight. These models can be represented as nonautonomous deterministic difference equations generated by affine transformations $\psi_t : \mathbb{R}^n \to \mathbb{R}^n$ of the state space $\mathbb{R}^n$:

$$x_{t+1} = \psi_t(x_t) = A_t x_t + b_t, \qquad t \in \mathbb{Z}, A_t \in \mathbb{GL}(n), \text{ and } b_t \in \mathbb{R}^n. \tag{2.1}$$

$\mathbb{GL}(n)$ denotes the general linear group of order n, i.e. the set of invertible $n \times n$ matrices.[5] b_t denotes the vector of exogenous variables whose evolution is known and given. Moreover, we assume that $\{b_t\}$ is bounded, i.e. that $\|b_t\| < m$, for all $t \in \mathbb{Z}$, for some positive constant $m < \infty$.

We restrict the class of admissible models to those where $\{A_t\}$ is *periodic*. More specifically, we assume that the matrices $\{A_t\}$ are selected from a

[5]Allowing for singular matrices implies that for points in the range of A_t there would exist an x_t such that $x_{t+1} = A_t x_t + b_t$ with x_t being not unique. The theory could be generalized to noninvertible matrices, however, at the cost of considerable technical effort. As the noninvertibility may be the result of some redundancies, like defining equations, we stick to the case of nonsingular matrices.

finite set of matrices with cardinality $p \in \mathbb{N}^+$. Moreover, they satisfy the periodicity condition

$$A_{t+p} = A_t, \quad \text{for all } t \in \mathbb{Z} \text{ and } p \in \mathbb{N}^+.$$

We allow for the possibility that some of the matrices repeat themselves within p periods. Thus, the economy may stay within a particular regime for some time. Nevertheless, we think of p as being the smallest integer satisfying the above periodicity condition.[6] The theory of this type of difference equations is known as the Floquet theory and is well–understood. Excellent expositions can be found in ELAYDI (2005, section 3.4) or, more in line with this paper, COLONIUS and KLIEMANN (2014, section 7.1), among others. Note that equation (2.1) satisfies the *superposition principle*. This implies that the solution is given as the sum of the general solution to the linear nonautonomous equation and a particular solution of the affine equation (2.1). The linear nonautonomous equation is

$$x_{t+1} = A_t x_t, \qquad t \in \mathbb{Z}. \tag{2.2}$$

We proceed by first analyzing the linear nonautonomous equation. As a general remark note that starting with any $x \in \mathbb{R}^n$ in period 0, x_t is recursively given by

$$x_t = A_{t-1}A_{t-2}\ldots A_1 A_0 x = \Phi(t)x$$

where $\Phi(t)$ denotes the corresponding matrix product. Write this solution as $\varphi(t, x)$, then $\varphi : \mathbb{Z} \times \mathbb{R}^n \to \mathbb{R}^n$ is linear in the second argument and satisfies the *cocycle property*:

$$\varphi(t + s, x) = \varphi(s, \varphi(t, x)), \qquad \text{for all } t, s \in \mathbb{Z} \text{ and } x \in \mathbb{R}^n.$$

The cocycle property and the invertibility of the A_t's imply that

$$\Phi(t) = \begin{cases} A_{t-1}A_{t-2}\ldots A_1 A_0, & t > 0; \\ I_n, & t = 0; \\ A_0^{-1}A_1^{-1}\ldots A_{t-2}^{-1}A_{t-1}^{-1}, & t < 0. \end{cases}$$

The assumption of periodicity leads to the immediate observation that the nonautonomous equation (2.2) can be reduced to an autonomous one by

---

[6] If $p = 1$, we obtain the standard case where A_t is constant.

taking p steps at once. For any solution $\{x_t\}$ define $y_\tau = x_{\tau p}$, $\tau \in \mathbb{Z}$. Then,

$$\begin{aligned} x_{\tau p+1} &= A_{\tau p} x_{\tau p} = A_0 y_\tau, \\ x_{\tau p+2} &= A_{\tau p+1} x_{\tau p+1} = A_1 A_0 x_{\tau p} = A_1 A_0 y_\tau, \\ &\vdots \\ x_{\tau p+p} &= \left(\prod_{j=1}^{p} A_{p-j} \right) x_{\tau p} = \left(\prod_{j=1}^{p} A_{p-j} \right) y_\tau. \end{aligned}$$

This shows that $\{y_\tau\}$ is a solution of the autonomous difference equation

$$y_{\tau+1} = \Phi(p) y_\tau \qquad \text{with } \Phi(p) = \prod_{j=1}^{p} A_{p-j} = A_{p-1} \dots A_1 A_0.$$

Conversely, a solution $\{y_\tau\}$ determines a solution of equation (2.2) via the equations above: $x_{\tau p+1} = A_{\tau p} x_{\tau p} = A_0 y_\tau$, $x_{2\tau+2} = A_{\tau p+1} x_{\tau p+1} = A_1 A_0 y_\tau$, etc.[7]

The asymptotic behavior of solutions $\varphi(t, x)$ clearly depends on the eigenvalues α_j, $j = 1, \dots, p$, of $\Phi(p)$, known as the *Floquet multipliers*. The *Floquet exponents* are defined as

$$\lambda_j = \frac{1}{p} \log |\alpha_j|.$$

It turns out that they play a crucial in understanding the asymptotic behavior of solutions of equation (2.2) and the stability of the zero solution.

Define $\mathbb{Z}_p = \{0, 1, \dots, p-1\}$ and denote by $\varphi(t, \nu, x)$ the solution of $x_{t+1} = A_t x_t$ with initial condition $x_0 = x$ and $A_t = A_{\theta(t,\nu)}$, $\nu = 0, 1, \dots, p$, where $\theta(t, \nu) : \mathbb{Z} \times \mathbb{Z}_p \to \mathbb{Z}_p$ is defined as $\theta(t, \nu) = (t + \nu) \bmod p$. In the case $p = 2$, $\varphi(t, 0, x)$ becomes $x_1 = A_0 x$, $x_2 = A_1 A_0 x$, $x_3 = A_0 A_1 A_0 x$, $x_4 = A_1 A_0 A_1 A_0 x$, $\dots$, and $\varphi(t, 1, x)$ is given by $x_1 = A_1 x$, $x_2 = A_0 A_1 x$, $x_3 = A_1 A_0 A_1 x$, $x_4 = A_0 A_1 A_0 A_1 x$, $\dots$ The exponential growth rate of $\varphi(t, \nu, x)$, also known as the *Lyapunov exponent*, is defined as

$$\lambda(x, \nu) = \limsup_{t \to \infty} \frac{1}{t} \log \|\varphi(t, \nu, x)\| \qquad \text{for } (\nu, x) \in \mathbb{Z}_p \times \mathbb{R}^n.$$

Because vectors with different Lyapunov exponents are linearly independent, there are at most n Lyapunov exponents for each ν. Call the number of different Lyapunov exponents ℓ. Usually, we think of these exponents being ordered as $-\infty \leq \lambda_\ell < \lambda_{\ell-1} < \dots < \lambda_1 < \infty$ where λ_1 is known as the *top Lyapunov exponent*. These and further properties can be found f.e. in Arnold (2003, section 3.2). With these preliminaries we can quote the following Theorem from Colonius and Kliemann (2014, theorem 7.1.7).

[7] As usually, the empty product $\Pi_{j=m}^{n}$, $m > n$, is set to one by convention.

Theorem 1 (Floquet's theorem). *Consider the p-period linear (nonautonomous) difference equation* (2.2)*. The Lyapunov exponents coincide with the Floquet exponents* λ_j*,* $j = 1, \dots, \ell \leq n$*, and they exist as limits. For each* $\nu \in \mathbb{Z}_p$ *there exists a decomposition or splitting of the state space*

$$\mathbb{R}^n = L(\lambda_1, \nu) \oplus \cdots \oplus L(\lambda_\ell, \nu)$$

into linear subspaces $L(\lambda_j, \nu)$ *called the* Floquet *or* Lyapunov spaces*. These subspaces have the following properties:*

(i) The Lyapunov spaces have dimensions independent of ν*,*

$$d_j = \dim L(\lambda_j, \nu) \text{ is constant for } \nu \in \mathbb{Z}_p;$$

(ii) they are invariant under multiplication by the principal fundamental matrix $\Phi(t, \nu)$ *in the following sense:*[8]

$$\Phi(t + \nu, \nu) L(\lambda_j, \nu) = L(\lambda_j, \theta(t, \nu)) \text{ for all } t \in \mathbb{Z} \text{ and } \nu \in \mathbb{Z}_p;$$

(iii) for every $\nu \in \mathbb{Z}_p$*, the Lyapunov exponents satisfy*

$$\lambda(x, \nu) = \lim_{\pm t \to \infty} \frac{1}{t} \log \|\varphi(t, \nu, x)\| = \lambda_j \quad \text{if and only if } x \in L(\lambda_j, \nu) \text{ and } x \neq 0.$$

Remark 1. *The splitting of* $\mathbb{R}^n$ *into subspaces is given as a direct sum. The subspaces are not necessarily orthogonal to each other. Property* (*ii*) *means that the subspaces are* equivariant *or* covariant*. Note that these subspaces depend on* ν *whereas their dimensions and the Lyapunov exponents do not.*

The Lyapunov subspaces can be collected into subbundles

$$L^s(\nu) = \bigoplus_{\lambda_j < 0} L(\lambda_j, \nu), \quad L^c(\nu) = L(0, \nu), \text{ and } L^u(\nu) = \bigoplus_{\lambda_j > 0} L(\lambda_j, \nu)$$

called the stable subbundle, the center, and the unstable subbundle, respectively. Thus, the zero solution of equation (2.2) is *asymptotically stable* if and only if all Lyapunov exponents are negative. This is equivalent to $L^s(\nu) = \mathbb{R}^n$ for some (hence for all) $\nu \in \mathbb{Z}_p$. The difference equation (2.2) is called *hyperbolic* if $L^c(\nu) = \emptyset$ or, equivalently, if all Lyapunov exponent are different from zero. For a hyperbolic difference equation the zero solution is called a *saddle point* if both $L^s(\nu)$ and $L^u(\nu)$ have dimensions $d^s = \dim L^s(\nu)$, respectively $d^u = \dim L^u(\nu)$, strictly greater than zero. For the rest of this section, we make the following assumption:

[8] This matrix is defined as $\Phi(t, \nu) = \prod_{j=1}^{t} A_{\theta(t-j,\nu)}$. It is similar to the matrix product $\Phi(t)$, but starts the product with A_ν. Thus, for $\nu = 0$, we have $\Phi(t, \nu) = \Phi(t)$.

Assumption 1. *The difference equation* (2.2) *is hyperbolic.*

In economics, especially in the context of rational expectations models, we are often faced with a reversed boundary value problem: Find an initial value x_0 such that the solution $\varphi(t, \nu, x_0)$ of equation (2.2) does not explode, i.e. such that $\lambda(x_0, \nu) < 0$, subject to some initial conditions. These initial conditions can be written compactly in matrix notation as

$$\text{initial condition:} \qquad c = Rx_0, \qquad c \in \mathbb{R}^r \setminus \{0\} \text{ given}, \tag{2.3}$$

where R is a given $(r \times n)$–matrix of rank r and c a given r-vector. The nonexplosiveness condition is rationalized by the boundedness assumption of $\{b_t\}$. The restrictions implied by equation (2.3) usually fix the values of the predetermined variables. Often R takes the form $R = (I_r, 0_{r,n-r})$. In this case, the restriction sets the first r elements of x_0 to c. Depending on the rank of R several cases can be distinguished. If $r = 0$, the condition (2.3) makes effectively no restriction. In this situation $x_t = 0$ for all t is the unique nonexplosive solution if and only if $L^s(\nu) = \emptyset$. In the case $r = n$, the condition (2.3) determines a unique x_0. If this x_0 lies in $L^s(\nu)$ then $\varphi(t, \nu, x_0)$ is the unique nonexplosive solution. This is obviously the case if $L^s(\nu) = \mathbb{R}^n$. Consider finally the most interesting case $0 < r < n$. Denote by $\pi^{s,u}(\nu) : \mathbb{R}^n \to L^s(\nu)$ the projection onto $L^s(\nu)$ along $L^u(\nu)$ and by $\pi^{u,s}(\nu) : \mathbb{R}^n \to L^u(\nu)$ the projection onto $L^u(\nu)$ along $L^s(\nu)$. Taking $B(\nu)$ as a basis of $\mathbb{R}^n$ obtained from the union of a basis of $L^s(\nu)$ and $L^u(\nu)$, in that order, these projections are given by (see Meyer, 2000, chapter 5.9)

$$\pi^{s,u}(\nu) = B(\nu) \begin{pmatrix} I_{d^s} & 0 \\ 0 & 0 \end{pmatrix} B^{-1}(\nu) \quad \text{and} \quad \pi^{u,s}(\nu) = B(\nu) \begin{pmatrix} 0 & 0 \\ 0 & I_{d^u} \end{pmatrix} B^{-1}(\nu).$$

$\varphi(t, \nu, x_0)$ then determines a nonexplosive solution if and only if x_0 fulfills the simultaneous equation system consisting of the restriction (2.3) and the condition $\pi^{s,u}(\nu)x_0 = x_0$, respectively $\pi^{u,s}(\nu)x_0 = 0$. This reasoning leads to the following proposition.

Proposition 1. *The zero solution is the unique nonexplosive solution of the hyperbolic difference equation* (2.2) *if and only if*

$$\operatorname{rank} \begin{pmatrix} R \\ \begin{pmatrix} 0 & I_{d^u} \end{pmatrix} B(\nu)^{-1} \end{pmatrix} = n. \tag{2.4}$$

If condition (2.4) is satisfied, the difference equation (2.2) is said to be *determinate*. Because R has r rows and $\begin{pmatrix} 0 & I_{d^u} \end{pmatrix} B^{-1}$ has $d^u = n - d^s$ rows, we have the following corollary.

Corollary 1. *A necessary condition for a unique nonexplosive solution is that* $r = d^s$.

If $r < d^s$, there exists a whole family of nonexplosive solutions and the system (2.2) is then called *indeterminate*. If $r > d^s$, the equation system is overdetermined and no nonexplosive solution exists.

Having established the general solution to the linear equation (2.2), we then have to find a particular solution to the affine equation (2.1). The idea is to split b_t as $b_t = \pi^{s,u}(\nu)b_t + \pi^{u,s}(\nu)b_t$ and to iterate the component in the stable (unstable) subbundle backward (forward) in time. This then leads to following particular solution:

$$x_t^{(p)} = \begin{pmatrix} x_t^{b,p} \\ x_t^{f,p} \end{pmatrix} = \begin{pmatrix} \sum_{j=0}^{\infty} \left(\prod_{i=1}^{j} A_{t-i}\right) & \pi^{s,u}(\theta(t,\nu))b_{t-1-j} \\ -\sum_{j=0}^{\infty} \left(\prod_{i=0}^{j} A_{t+i}\right)^{-1} & \pi^{u,s}(\theta(t,\nu))b_{t+j} \end{pmatrix}$$

$$= \Phi(t,\nu) \begin{pmatrix} \sum_{j=0}^{\infty} \Phi(t-j,\nu)^{-1} & \pi^{s,u}(\theta(t,\nu))b_{t-1-j} \\ -\sum_{j=0}^{\infty} \Phi(t+j+1,\nu)^{-1} & \pi^{u,s}(\theta(t,\nu))b_{t+j} \end{pmatrix}. \tag{2.5}$$

These considerations, finally, lead to the following Theorem.

Theorem 2. *The hyperbolic, periodic, and affine deterministic difference equation* (2.1) *with bounded* $\{b_t\}$ *admits a unique nonexplosive solution if and only if the rank condition* (2.4) *is satisfied. The solution to the corresponding boundary value problem then is*

$$x_t = \Phi(t,\nu) \left(x_0 + \begin{pmatrix} \sum_{j=0}^{\infty} \Phi(t-j,\nu)^{-1} & \pi^{s,u}(\theta(t,\nu))b_{t-1-j} \\ -\sum_{j=0}^{\infty} \Phi(t+j+1,\nu)^{-1} & \pi^{u,s}(\theta(t,\nu))b_{t+j} \end{pmatrix} \right) \tag{2.6}$$

where x_0 *is uniquely determined by the rank condition* (2.4).

Appendix B runs a numerical example.

3 A New Keynesian Model as a Canonical Example

We elucidate the theory developed in the previous Section by applying it to a simple canonical New Keynesian macroeconomic model. This model was introduced and analyzed by Galí (2011) to illustrate the role of alternative monetary policy rules. The standard model consists of three equations:

$$\begin{aligned}
y_t &= y_{t+1} - \frac{1}{\sigma}(i_t - \pi_{t+1}), && \text{(IS-equation)}\\
\pi_t &= \beta\pi_{t+1} + \kappa y_t + u_t, && \text{(forward-looking Phillips-curve)}\\
i_t &= \phi\pi_t, && \text{(Taylor-rule)}
\end{aligned}$$

where y_t, π_t, and i_t denote income (output gap), inflation and the nominal interest rate, all measured as deviations from the steady state. u_t is an exogenous, bounded cost-push shock. Furthermore, we assume that $\sigma > 0$, $\kappa > 0$, and $0 < \beta \leq 1$. In addition, $\phi > 0$ measures the aggressiveness of the central bank against in combatting inflation. Note that in this example there is no restrictions coming from initial conditions, i.e rank $R = 0$. We only require that the solution is nonexplosive.

This model can be solved in terms of $(y_{t+1}, \pi_{t+1})'$ by inserting the Taylor-rule and the Phillips-curve into the IS-equation:

$$\begin{aligned}
x_{t+1} = \begin{pmatrix} \pi_{t+1} \\ y_{t+1} \end{pmatrix} &= \frac{1}{\beta}\begin{pmatrix} 1 & -\kappa \\ (\phi\beta - 1)/\sigma & \beta + \kappa/\sigma \end{pmatrix}\begin{pmatrix} \pi_t \\ y_t \end{pmatrix} + \begin{pmatrix} -u_t/\beta \\ u_t/(\sigma\beta) \end{pmatrix}\\
&= A_0 x_t + b_{0,t} \qquad (3.1)
\end{aligned}$$

Denote the characteristic polynomial of A_0 by $\mathcal{P}(\alpha)$ and the corresponding eigenvalues by α_1 and α_2, then we have

$$\mathcal{P}(\alpha) = (\alpha - \alpha_1)(\alpha - \alpha_2) = \alpha^2 - \mathrm{tr}(A_0)\alpha + \det A_0$$

with

$$\begin{aligned}
\mathrm{tr}A_0 &= \alpha_1 + \alpha_2 = 1 + \frac{1}{\beta} + \frac{\kappa}{\sigma\beta} > 2\\
\det A_0 &= \alpha_1\alpha_2 = \frac{1}{\beta} + \frac{\kappa\phi}{\sigma\beta} > 1\\
\Delta_0 &= (\mathrm{tr}A_0)^2 - 4\det A_0 = \left(1 - \frac{1}{\beta}\right)^2 + \frac{\kappa}{\sigma\beta}\left(\frac{\kappa}{\sigma\beta} + 2 + \frac{2}{\beta} - 4\phi\right)\\
\mathcal{P}(1) &= (1 - \alpha_1)(1 - \alpha_2) = \frac{\kappa}{\sigma\beta}(\phi - 1) > 0, \quad \text{if } \phi > 1,
\end{aligned}$$

where Δ_0 denotes the discriminant of the quadratic equation. Depending on ϕ, the roots of $\mathcal{P}(\alpha)$ may be complex. We therefore distinguish several cases:

(i) ϕ is so high such that $\Delta_0 < 0$. In this case we have two complex conjugate roots. Note that this case can only arise if $\phi > 1$. Because

$\det A_0 > 1$, they are both located outside the unit circle.[9] The model is determinate and the unique nonexplosive solution is one where $x_0 = 0$.

(ii) ϕ is small enough such that $\Delta_0 > 0$. In this case both eigenvalues are real. They must also be of the same sign because the trace and the determinant of A_0 are positive. From the expression of $\mathcal{P}(1)$, we conclude that both eigenvalues are bigger than one if $\phi > 1$.

From this discussion we conclude that the stability of the model is independent of the parameters β, κ, and σ and depends solely on ϕ.

The above results translated in terms of the Lyapunov exponents are illustrated in Figure 1 for the values $\beta = 0.985$, $\kappa = 0.8$, and $\sigma = 1$. The red curve labeled "no switching" plots the Lyapunov exponents as a function of ϕ. When ϕ is equal to zero, i.e. when the central bank does not react to inflation we have two real eigenvalues opposite of one or, equivalently, two Lyapunov exponents of opposite signs. The model is indeterminate. As the central bank progressively combats inflation more aggressively the two eigenvalues, respectively, the two Lyapunov exponents get closer to each other. When ϕ becomes greater than one, the Lyapunov exponents become positive and the model determinate. For a large enough values of ϕ the roots become complex and we have only one positive Lyapunov exponent and the model is still determinate.

When $\phi > 1$, the model is determinate with a unique solution. As both variables are not predetermined, the boundedness condition alone determines the unique solution. Thus, $x_0 = 0$ and, because $L^s(\nu) = \emptyset$, $\pi^{u,s}(\nu) = \mathrm{Id}_{R^n}$ and $\pi^{s,u}(\nu) = 0$. Applying formula (2.6) with $\Phi(t,\nu) = A_0^t$ then gives as the unique bounded solution

$$x_t = -\sum_{j=1}^{\infty} A_0^{-j} b_{0,t+j-1} = \sum_{j=1}^{\infty} Q \begin{pmatrix} \alpha_1^{-j} & 0 \\ 0 & \alpha_2^{-j} \end{pmatrix} Q^{-1} \begin{pmatrix} u_{t-1+j}/\beta \\ -u_{t-1+j}/(\sigma\beta) \end{pmatrix}$$

where the columns of Q consist of the eigenvectors corresponding to α_1 and α_2. This is usually regarded to be the standard case.

Consider next the case where the central bank is interested in the future path of income (output gap) and inflation *conditional* on an exogenously given path of the interest rate. This case is of particular concern to Galí because of his claim that central banks actually performed such thought experiments to implement theirs policies (Galí, 2011, see for details). In

[9] Another way to reach this conclusion is by observing that the real part of the roots is $\frac{\mathrm{tr} A_0}{2} > 1$.

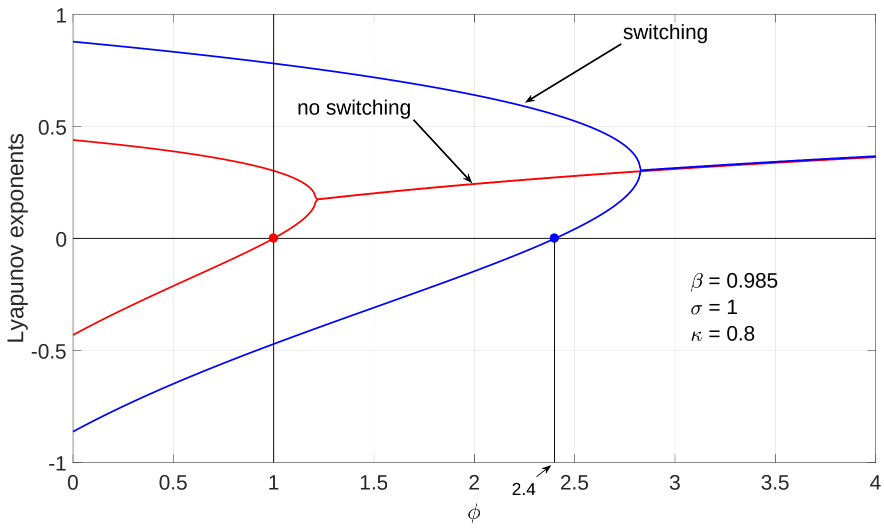

Figure 1: The Dynamics of the New Keynesian Model with and without Switching ($\beta = 0.985$, $\kappa = 0.8$, $\sigma = 1$)

this scenario the interest rate becomes an exogenous variable and the system changes to:

$$x_{t+1} = \begin{pmatrix} \pi_{t+1} \\ y_{t+1} \end{pmatrix} = \frac{1}{\beta} \begin{pmatrix} 1 & -\kappa \\ -1/\sigma & \beta + \kappa/\sigma \end{pmatrix} \begin{pmatrix} \pi_t \\ y_t \end{pmatrix} + \begin{pmatrix} -u_t/\beta \\ i_t^*/\sigma + u_t/(\sigma\beta) \end{pmatrix}$$
$$= A_1 x_t + b_{1,t} \tag{3.2}$$

where i_t^* is the exogenous (bounded) path of the interest rate. Note that A_1 is obtained from A_0 by setting ϕ equal to zero. Thus, the magnitude of $\mathrm{tr} A_1$ and $\det A_1$ are the same as for A_0, i.e. $\mathrm{tr} A_1 > 2$ and $\det A_1 > 1$. The discriminant is now unambiguously positive whereas $\mathcal{P}(1)$ becomes negative. This implies that one eigenvalue is smaller than one whereas the other one is bigger than one. Thus, the boundedness condition does not determine a unique solution so that we are faced with a situation of indeterminacy. The implications of this indeterminacy for monetary policy and possible remedies are discussed in Galí (2011). Taking the same numerical values as before, i.e. $\beta = 0.985$, $\sigma = 1$, and $\kappa = 0.8$, the eigenvalues take the values $\alpha_1 = 0.42$ and $\alpha_2 = 2.38$ and the Lyapunov exponents the corresponding values $\lambda_1 = -0.43$ and $\lambda_2 = 0.44$.

Suppose next that the central bank switches deterministically between the two policies starting with the model with Taylor rule followed by the model with interest rate fixing. This and similar settings are discussed in Galí (2011, section 4.1.1) and, in particular, Laséen and Svensson (2011). Of course any other deterministic periodic pattern is feasible as well. As shown in Figure 1 there is still a region of the parameter space for ϕ where the model becomes determinate. However, the central bank must be much more aggressive in combatting inflation when the regime with Taylor rule is in place. In the numerical example ϕ must be greater than 2.4 compared to one in the case with "no switching".

4 Random Time–Varying Rational Expectations Models

The theory of random coefficients shares some similarities to the periodic deterministic case analyzed in Section 2, but requires additional technical considerations. In analogy to equation (2.1) we analyze the following class of affine time–varying stochastic rational expectations models (TVRE models):

$$\mathbb{E}_t x_{t+1} = \psi_t(x_t) = A_t x_t + b_t, \qquad t \in \mathbb{Z}, \tag{4.1}$$

where ψ_t is a randomly chosen affine map. The sequence $\{(A_t, b_t)\}$ of random variables is defined on some probability space $(\Omega, \mathfrak{F}, \mathbf{P})$, i.e. a set Ω endowed with a σ–algebra $\mathfrak{F}$ and a probability measure $\mathbf{P}$. The sequence takes values in the state space $\mathbb{GL}(n) \times \mathbb{R}^n$ where $\mathbb{GL}(n)$ denotes the general linear group of degree n, i.e. the set of nonsingular $n \times n$ matrices.[10] Define $\mathfrak{F}_t$ as $\mathfrak{F}_t = \sigma\{(x_s, A_s, b_s) : s \leq t\}$, i.e. $\mathfrak{F}_t$ is the smallest σ–algebra such that (x_s, A_s, b_s) is measurable for all $s \leq t$. The sequence of σ–algebras $\mathfrak{F}_t$ then becomes a filtration adapted to $\{x_t\}$ and $\{(A_t, b_t)\}$ with $\mathfrak{F}_t \subseteq \mathfrak{F}$. $\mathbb{E}_t x_{t+1}$ denotes the conditional expectation with respect to $\mathfrak{F}_t$, i.e. $\mathbb{E}_t x_{t+1} = \mathbb{E}[x_{t+1} \mid \mathfrak{F}_t]$.

A solution of this TVRE model is any $\mathbb{R}^n$–valued sequence of random variables $\{x_t\}$, $t \in \mathbb{Z}$, which satisfies equation (4.1). Associated to the expectational difference equation (4.1) we investigate the following boundary value problem:

initial values: $\quad c = Rx_0, \qquad c \in \mathbb{R}^n \setminus \{0\}$ given;
boundedness: $\quad \|x_t\| < m \quad$ with probability one for some constant m.

The boundedness requirement is usually rationalized on the assumption that the sequence $\{b_t\}$ is assumed to be bounded. Thus, it makes sense to look for solution which are also bounded.

[10] For the sake of simplicity in the notation, we have omitted the dependence on $\omega \in \Omega$. Thus, we have written A_t for $A_t(\omega) = A(\theta^t \omega)$ where θ is an ergodic metric dynamical system (see Arnold, 2003, Appendix A). One can think of θ as being the time–forward shift. More details are given below.

From a conceptual point of view it is important to have a clear understanding on the randomness present in $\{A_t\}$. More precisely, we think of $\{A_t\}$ as being generated by a dynamical system in the following way. Let $\theta(t, \omega)$, $t \in \mathbb{Z}$, be a sequence of measurable maps

$$\theta(t, \omega) : \mathbb{Z} \times \Omega \to \Omega$$

such that $\theta(0, \omega) = \omega$ and such that the cocycle property $\theta(t + s, \omega) = \theta(t, \theta(s, \omega))$ is satisfied for all $t, s \in \mathbb{Z}$ and $\omega \in \Omega$. Moreover, we assume that the probability measure $\mathbf{P}$ is invariant under θ, i.e. $\theta(t, .)\mathbf{P} = \mathbf{P}$ for all $t \in \mathbb{Z}$. Because of the cocycle property, we can think of $\theta(t, \omega)$ as being generated by $\theta(1, \omega)$ and we write $\theta^t\omega$ for $\theta(t, \omega)$. In the same vein, we denote $A(\theta^t\omega)$ by $A_t(\omega)$. Sometimes we suppress the dependence on ω and write just A_t for short. Dynamical systems with these properties are called *metric dynamical systems*. Moreover, we assume θ to be ergodic. Following Arnold (2003, chapter 1), the evolution of the system on the bundle $\Omega \times \mathbb{R}^n$ can be envisioned as in Figure 2. While ω is shifted by θ to $\theta\omega$, the point x_0 in the fiber $\omega \times \mathbb{R}^n$ is shifted to $x_1 = \psi(\omega)x_0 = A(\omega)x_0 + b(\omega)$ in the fiber $\theta\omega \times \mathbb{R}^n$. In the next period $\theta\omega$ is shifted to $\theta^2\omega$ whereas x_1 is shifted to $x_2 = \psi(\theta\omega) = A(\theta\omega) + b(\theta\omega)$ and so on. The nice thing is that on each fiber the system is affine in the usual sense. We treat the random dynamical system θ as being fixed. Later in the application we specify it to follow a given Markov chain.

Suppose we are given two solutions $\{x_t^{(1)}\}$ and $\{x_t^{(2)}\}$, then $\{x_t^{(1)} - x_t^{(2)}\}$ satisfies the linear expectational equation

$$A_t x_t = \mathbb{E}_t x_{t+1}. \tag{4.2}$$

This implies that, as in the deterministic case, the superposition principle holds and that every solution is of the form:

$$x_t = x_t^{(g)} + x_t^{(p)}$$

where $\{x_t^{(g)}\}$ denotes the general solution of the linear equation (4.2) and $\{x_t^{(p)}\}$ a particular solution of equation (4.1). Thus, the solution can be found by first finding the general solution to the linear equation (4.2) and then looking for a particular solution of equation (4.1).

Define $\{\Phi(t)\} = \{\Phi(t, \omega)\}$ as the random matrix product:

$$\Phi(t) = \Phi(t, \omega) = \begin{cases} A_{t-1}(\omega) \ldots A_1(\omega) A_0(\omega), & t = 1, 2, \ldots; \\ I_n, & t = 0; \\ A_t(\omega)^{-1} \ldots A_{-1}(\omega)^{-1}, & t = -1, -2, \ldots \end{cases}$$

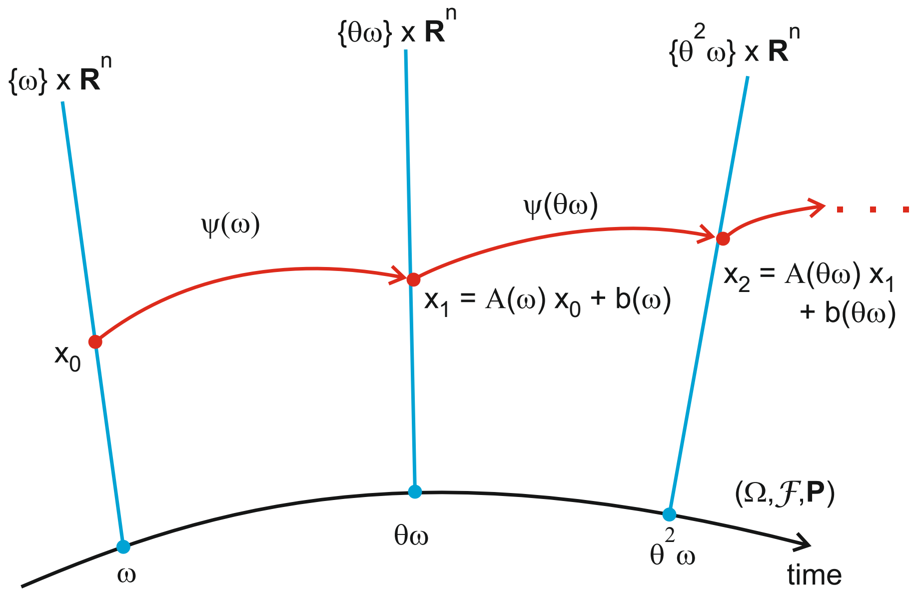

Figure 2: The Evolution of an Affine Random Dynamic System

Next define a new variable y_t as $y_t = \Phi(t)^{-1}x_t$. It is easy to see that $\{y_t\}$ is a martingale:

$$\mathbb{E}_t y_{t+1} = \mathbb{E}_t \left(\Phi(t+1)^{-1}x_{t+1}\right) = \Phi(t+1)^{-1}\mathbb{E}_t x_{t+1} = \Phi(t+1)^{-1}A_t x_t = y_t$$

This implies that the general solution of the linear equation (4.2) can be represented as

$$x_t = (A_{t-1}\dots A_1 A_0)\, m_t = \Phi(t)m_t, \tag{4.3}$$

where $\{m_t\}$ is a martingale with respect to the filtration $\{\mathfrak{F}_t\}$. Similarly, the time reversed process is also a martingale. This implies without any additional assumptions that there exists a random variable m_∞ such that $\lim_{t\to\infty} m_t = m_\infty$ a.s. and in mean (see Grimmett and Stirzaker, 2001, section 12.7). Moreover, the original martingale can be reconstructed from m_∞ by setting $m_t = E(m_\infty \mid \mathfrak{F}_t)$. Thus, the space of martingales can be continuously parameterized by the space of random variables which are measurable with respect to $\mathfrak{F}_\infty$ where $\mathfrak{F}_\infty = \sigma\left(\bigcup_{t\in\mathbb{Z}}\mathfrak{F}_t\right)$.[11] An important special case arises if m_t is constant.

The existence and the stability properties of the solutions (4.3) thus depend crucially on the convergence of the matrix products $\Phi(t)$. To study this issue, we introduce again the *Lyapunov exponent* as the asymptotic growth of solutions of the linear random dynamical system $x_{t+1} = A_t x_t$ associated to equation (4.2) (see Section 2 for comparison). Denote by $\varphi(t,\omega,x)$ the trajectory $\Phi(t,\omega)x$. The Lyapunov exponent is then defined as

$$\lambda(\omega,x) = \limsup_{t\to\infty}\frac{1}{t}\log\|\varphi(t,\omega,x)\|. \tag{4.4}$$

The Lyapunov exponent therefore describes the asymptotic exponential growth rate of the linear random dynamical system $x_{t+1} = A_t x_t$ starting with $x_0 = x$.

Throughout the paper we assume the following integrability condition.

Assumption 2 (Integrability)**.** *For every $t\in\mathbb{Z}$,*

$$\log^+\|A_t\|,\quad \log^+\|A_t^{-1}\| \quad \textit{and}\ \log^+\|b_t\| \in \mathrm{L}^1(\Omega,\mathfrak{F},\mathbf{P}).$$

Thereby $\log^+ x$ stands for $\max\{\log x, 0\}$. This assumption will be satisfied if the random variables would be essentially bounded. Here and in the following $\|.\|$ denotes the operator norm induced by the Euclidean metric, i.e. $\|A\| = \max_{\|x\|=1}\|Ax\| = \delta_1$ where δ_1 is the largest singular value of A. Because all norms are equivalent, the integrability assumption and the limits below are independent from the norm the chosen provided that it is submultiplicative.

[11] Compare this to Klein (2000, Definition 4.3 and Assumption 4.2)

A result analogous to Theorem 1 was originally proven by OSELEDETS (1968) for random dynamical systems. This theorem is also known as the Multiplicative Ergodic Theorem (MET) and provides a substitute for the spectral theorem which is key in the understanding deterministic difference equations.[12] In particular, the MET postulates the existence of invariant subspaces and constant asymptotic growth rates (Lyapunov exponents). An extensive exposition of this theorem with proofs can be found in ARNOLD (2003). Here we follow COLONIUS and KLIEMANN (2014) and present a more accessible version.

Theorem 3 (Multiplicative Ergodic Theorem (MET))**.** *Under the integrability assumption 2 and the assumptions on the random dynamical system $\theta(t, \omega)$, the linear time–varying system $x_{t+1} = A_t x_t$ induces a splitting of $\mathbb{R}^n$ into $\ell \leq n$ linear subspaces $L_j(\omega) = L(\lambda_j, \omega)$, $j = 1, \ldots, \ell$. These subspaces have the following properties:*

(i) There is a decomposition (splitting)

$$\mathbb{R}^n = L_1(\omega) \oplus \cdots \oplus L_\ell(\omega)$$

of $\mathbb{R}^n$ into ℓ invariant random linear subspaces $L_j(\omega)$, i.e. $A(\omega)L_j(\omega) = L_j(\theta\omega)$ for $j = 1, \ldots, \ell$. The linear subspaces $L_j(\omega)$ are called Lyapunov or Osoledets spaces and have constant dimensions d_j.

(ii) There are real numbers $\lambda_1 > \cdots > \lambda_\ell$ such that for each $x \in \mathbb{R}^n \setminus \{0\}$ the Lyapunov exponent $\lambda(\omega, x) \in \{\lambda_1, \ldots, \lambda_\ell\}$ exists as a limit and

$$\lambda(\omega, x) = \lim_{t \to \pm\infty} \frac{1}{t} \log \|\varphi(t, \omega, x)\| = \lambda_j \text{ if and only if } x \in L_j(\omega) \setminus \{0\}.$$

(iii) The limit

$$\Upsilon(\omega) = \lim_{t \to \infty} \left(\Phi(t, \omega)' \Phi(t, \omega)\right)^{1/2t} \tag{4.5}$$

exists as a positive definite matrix. The different eigenvalues of $\Upsilon(\omega)$ are constants and can be written as $\exp(\lambda_1) > \cdots > \exp(\lambda_\ell)$; the corresponding random eigenspaces are $L_1(\omega), \ldots, L_\ell(\omega)$.

(iv) The Lyapunov exponents are obtained as limits from the singular values δ_k of $\Phi(t, \omega)$: The set of indices $\{1, 2, \ldots, n\}$ can be decomposed into subsets S_j, $j = 1, \ldots, \ell$, such that for all $k \in S_j$,

$$\lambda_j = \lim_{t \to \infty} \frac{1}{t} \log \delta_k(\Phi(t, \omega)).$$

[12]See MEYER (2000, chapter 7) for a comprehensive account of the spectral theorem.

If the state space is one–dimensional, i.e. if $n = 1$, the MET reduces to Birkhoff's ergodic theorem.[13] To see this, define $f(\omega) = \log|a(\omega)|$ where $a(\omega)$ stands for $A(\omega)$ to emphasize that we are dealing with the scalar case. Assume that $f^+ = \max\{0, f\} \in \mathrm{L}^1(\Omega, \mathfrak{F}, \mathbf{P})$ with $\int_\Omega f \mathrm{d}\mathbf{P} = \lambda$, then the Birkoff pointwise ergodic theorem asserts

$$\frac{1}{t}\sum_{j=0}^{t-1} f(\theta^j\omega) \longrightarrow \int_\Omega f \mathrm{d}\mathbf{P} = \lambda.$$

Noting that $\log|\varphi(t,\omega,x)| = \sum_{j=0}^{t-1}\log|a(\theta^j\omega)| + \log|x|$, statement *(ii)* in the MET corresponds exactly the Birkoff's theorem because $\lim_{t\to\pm\infty}\frac{1}{t}|x|$ converges to zero.

A remark analogous to Remark 1 also applies in the context of random time-varying models. The Lyapunov subspaces are not orthogonal and depend on ω. Their dimensions and the Lyapunov exponents, however, are independent of ω. Note also the invariance (equivariance or covariance) of these subspaces that is the property $A(\omega)L_j(\omega) = L_j(\theta\omega)$ which becomes important in the numerical implementation (see Froyland et al., 2013). There exists an alternative decomposition of $\mathbb{R}^n$ into orthogonal subspaces. These subspaces are, however, no longer invariant.

We collect the Lyapunov subspaces into subbundles:

$$L^s(\omega) = \bigoplus_{\lambda_j<0} L(\lambda_j,\omega), \quad L^c(\omega) = L(0,\omega), \text{ and } L^u(\omega) = \bigoplus_{\lambda_j>0} L(\lambda_j,\omega)$$

called the stable subbundle, the center, and the unstable subbundle, respectively. Thus, the zero solution is *asymptotically stable* if and only if all Lyapunov exponents are negative. This is equivalent to $L^s(\omega) = \mathbb{R}^n$ for some (hence for all) ω. The difference equation (4.1) is called *hyperbolic* if $L^c(\omega) = \emptyset$ or, equivalently, if all Lyapunov exponent are different from zero. For a hyperbolic difference equation the zero solution is called a *saddle point* if both $L^s(\omega)$ and $L^u(\omega)$ have dimensions $d^s = \dim L^s(\omega)$, respectively $d^u = \dim L^u(\omega)$, strictly greater than zero. Analogously to the deterministic case, define by $\pi^{s,u}(\omega) : \mathbb{R}^n \to L^s(\omega)$ the projection onto $L^s(\omega)$ along $L^u(\omega)$ and by $\pi^{u,s}(\omega) : \mathbb{R}^n \to L^u(\omega)$ the projection onto $L^u(\omega)$ along $L^s(\omega)$. Note that these projections also depend on ω.

As in the deterministic case, consider the inverse boundary problem: find an initial value x_0 such that the trajectory $\varphi(t,\omega,x_0)$ remains bounded, i.e. such that $\lambda(\omega,x_0) < 0$ subject to initial conditions given by equation (2.3).

[13] An introduction to ergodic theory can be found, f.e., Grimmett and Stirzaker (2001, section 9.5) or Colonius and Kliemann (2014, section 10.1).

As before a unique solution is obtained if and only if a rank condition similar to (2.4) is satisfied:

$$\operatorname{rank}\begin{pmatrix} R \\ \begin{pmatrix} 0 & I_{d^u} \end{pmatrix} B(\omega)^{-1} \end{pmatrix} = n \tag{4.6}$$

where $B(\omega)$ is a basis of $\mathbb{R}^n$ obtained as the union of a basis of $L^s(\omega)$ and $L^u(\omega)$. Note that $\pi^{u,s}(\omega) = B(\omega)\begin{pmatrix} 0 & 0 \\ 0 & I_{d^u} \end{pmatrix} B^{-1}(\omega)$. If this equation system has a unique solution, the TVRE model (4.1) is said to be *determinate*. As R has r rows and $\begin{pmatrix} 0 & I_{d^u} \end{pmatrix} B^{-1}$ has $d^u = n - d^s$ rows, a necessary condition for a unique solution is that $r = d^s$. If $r < d^s$, there is a whole family of solutions and the model (4.1) is then *indeterminate*. If $r > d^s$, the equation system is overdetermined and no solution exists.

Having established the general solution to the linear equation, the objective is then to find a particular solution to the affine equation. Assuming hyperbolicity, the idea is to project $b_t = b(\theta^t\omega)$ on the stable and unstable subbundles $L^s(\theta^{t+1}\omega)$ and $L^u(\theta^{t+1}\omega)$. This effectively decomposes b_t into the direct sum of two components where the first component lies in $L^s(\theta^{t+1}\omega)$ and the second in $L^u(\theta^{t+1}\omega)$:

$$b_t = b(\theta^t\omega) = \pi^{s,u}(\theta^{t+1}\omega)b_t + \pi^{u,s}(\theta^{t+1}\omega)b_t.$$

Using the towering property of conditional expectations (Law of Iterated Expectations), the next step consists in the iteration of the stable (unstable) component backward (forward) in time. This then leads to following particular solution:

$$\begin{aligned} x_t^{(p)} &= \begin{pmatrix} x_t^{b,p} \\ x_t^{f,p} \end{pmatrix} = \begin{pmatrix} \sum_{j=0}^{\infty} \left(\prod_{i=1}^{j} A_{t-i}\right) \quad \pi^{s,u}(\theta^{t-j}\omega)b_{t-1-j} \\ -\mathbb{E}_t\left[\sum_{j=0}^{\infty}\left(\prod_{i=0}^{j} A_{t+i}\right)^{-1} \pi^{u,s}(\theta^{t+j+1}\omega)b_{t+j}\right] \end{pmatrix} \\ &= \Phi(t)\begin{pmatrix} \sum_{j=0}^{\infty} \Phi(t-j)^{-1} \quad \pi^{s,u}(\theta^{t-j}\omega)b_{t-1-j} \\ -\mathbb{E}_t\left[\sum_{j=0}^{\infty} \Phi(t+j+1)^{-1}\, \pi^{u,s}(\theta^{t+j+1}\omega)b_{t+j}\right] \end{pmatrix}. \end{aligned} \tag{4.7}$$

Theorem 5.6.5 in Arnold (2003) shows that the infinite sums are well–defined and that equation (4.7) determines the unique φ-invariant solution. It is straightforward to verify that this is indeed a solution to the affine expectational equation (4.1).

These considerations, finally, lead to the following Theorem which presents a solution similar in spirit to the analysis of Blanchard and Kahn (1980).

Theorem 4. *Under the integrability assumption 1, the hyperbolic affine random rational expectations model* (4.1) *admits a unique nonexplosive solution if and only if the rank condition* (4.6) *is satisfied. The solution to the corresponding boundary value problem then is*

$$x_t = \Phi(t)\left(x_0 + \begin{pmatrix} \sum_{j=0}^{\infty} \Phi(t-j)^{-1} & \pi^{s,u}(\theta^{t-j}\omega)b_{t-1-j} \\ -\mathbb{E}_t\left[\sum_{j=0}^{\infty} \Phi(t+j+1)^{-1}\right. & \left.\pi^{u,s}(\theta^{t+j+1}\omega)b_{t+j}\right] \end{pmatrix}\right) \tag{4.8}$$

where x_0 is uniquely determined by the rank condition (4.6).

While the above solution formula is appealing from a theoretical perspective, it hard, if not impossible, to come up with an explicit analytical solution in a nontrivial context. Thus, one has to resort to numerical solutions. Some of these algorithms will be introduced by means of the New Keynesian model with stochastically varying Taylor–rule.

The approach could be equally well applied when the model (4.1) is brought into the format advocated by Sims (2001). Define for this purpose $y_t = \mathbb{E}_t x_{t+1}$ and enlarge the state space by introducing the equation $x_t = y_{t-1} + \eta_t$ where η_t denotes the expectation error $x_t - \mathbb{E}_{t-1}x_t$. The augmented system then is

$$\Gamma_t^{(0)}\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \Gamma_t^{(1)}\begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} b_t \\ \eta_t \end{pmatrix} \tag{4.9}$$

where $\Gamma_t^{(0)}$ and $\Gamma_t^{(0)}$ are defined appropriately as

$$\Gamma_t^{(0)} = \begin{pmatrix} I_n & -A_t \\ 0 & I_n \end{pmatrix} \quad \text{and} \quad \Gamma_t^{(1)} = \begin{pmatrix} 0 & 0 \\ I_n & 0 \end{pmatrix}.$$

Note that $\Gamma_t^{(0)}$ is invertible so that we can bring the equation into a standard format and analyze it from there on using again the tools of the MET. While the invertible case presents no problem, it is natural to ask to what extent is it possible to generalize to the singular case. This issue is of some importance because some equation might not involve expectational variables making $\Gamma_t^{(0)}$ singular.

A possible remedy is to use a generalized inverse. The natural generalized inverse in the context of difference equations is the Drazin inverse which we will now define. Let k denote index of the matrix $\Gamma_t^{(0)}$, i.e. the smallest integer for which $\operatorname{rank}{\Gamma_t^{(0)}}^k = \operatorname{rank}{\Gamma_t^{(0)}}^{k+1}$, then there exists a core–nilpotent decomposition

$$\Gamma_t^{(0)} = P_t\begin{pmatrix} C_t & 0 \\ 0 & N_t \end{pmatrix}P_t^{-1}$$

where C_t is a nonsingular $r \times r$ matrix and N_t is nilpotent matrix of index k, meaning $N_t^k = 0$. The Drazin inverse of $\Gamma_t^{(0)}$ is then defined as

$$\Gamma_t^{(0)D} = P_t \begin{pmatrix} C_t^{-1} & 0 \\ 0 & 0 \end{pmatrix} P_t^{-1}.$$

If $\Gamma_t^{(0)}$ is nonsingular, $k = 0$ and $r = n$, so N_t is not present. A detailed exposition of the Drazin inverse can be found in Campbell and Meyer (1979, chapter 8 and the application to difference equations in section 9.3) and Meyer (2000, section 5.10). Because we think that whether or not an expectational variables appears in a equation is related to structure of the economic model, we treat the index k as being fixed. In addition, we assume that there always exists some $z_t \in \mathbb{R}$ such that $z_t\Gamma_t^{(0)} + \Gamma_t^{(1)}$ is invertible. This assumption is standard in this context (see Sims (2001, footnote 1) or Klein (2000) for details and economic examples).

With these tools and assumptions at hand, we can decouple the system in two components. To see this concentrate on the linear case first and multiply both sides by $(z_t\Gamma_t^{(0)} + \Gamma_t^{(1)})^{-1}$:

$$\begin{aligned}
(z_t\Gamma_t^{(0)} + \Gamma_t^{(1)})^{-1}\Gamma_t^{(0)}U_t &= (z_t\Gamma_t^{(0)} + \Gamma_t^{(1)})^{-1}\Gamma_t^{(1)}U_{t-1} \\
\widehat{\Gamma}_t^{(0)}U_t &= (I_{2n} - z\widehat{\Gamma}_t^{(0)})U_{t-1} \\
\begin{pmatrix} C_t & 0 \\ 0 & N_t \end{pmatrix} \widetilde{U}_t &= \left(I_{2n} - z_t \begin{pmatrix} C_t & 0 \\ 0 & N_t \end{pmatrix}\right) \widetilde{U}_{t-1}.
\end{aligned}$$

where $\widehat{\Gamma}_t^{(0)} = (z_t\Gamma_t^{(0)} + \Gamma_t^{(1)})^{-1}\Gamma_t^{(0)}$, $U_t = (y_t', x_t')'$, and $\widetilde{U}_t = P^{-1}U_t$. This leads to

$$\begin{aligned}
C_t\widetilde{U}_{1t} &= (I_{n_1} - z_tC_t)\widetilde{U}_{1,t-1} \\
N_t\widetilde{U}_{2t} &= (I_{n_2} - z_tN_t)\widetilde{U}_{2,t-1}
\end{aligned}$$

where $\widetilde{U}_t$ is partitioned conformable into $\widetilde{U}_{1t}$ and $\widetilde{U}_{2t}$. The first set of equations can then be analyzed as set out previously because C_t is invertible. It is clear that in the case $k = 1$ which is equivalent to $N_t = 0$ the only solution to the second set of equations is $\widetilde{U}_{2t} = 0$. Given the solution for $\widetilde{U}_{1t}$ and $\widetilde{U}_{2t}$ one can then work backward to determine the solution for the original variables. In this endeavor it is important to check the consistency of the initial condition (Campbell and Meyer, 1979, theorem 9.3.2).

5 Stability of the Stochastically Time–Varying New Keynesian Model

We start by exploring the properties of the linear (nonautonomous) equation (4.2). Suppose that, in contrast to Section 3, the policymaker switches randomly according to some Markov process between the two policies. Thus, the system matrices A_t are chosen randomly from the set

$$\left\{\frac{1}{\beta}\begin{pmatrix} 1 & -\kappa \\ (\phi\beta-1)/\sigma & \beta+\kappa/\sigma \end{pmatrix}, \frac{1}{\beta}\begin{pmatrix} 1 & -\kappa \\ -1/\sigma & \beta+\kappa/\sigma \end{pmatrix}\right\}$$

Both matrices are the same except for the term $(A_t)_{21}$ which is $(\phi\beta-1)/(\beta\sigma)$ in one case and $-1/(\beta\sigma)$ in the other. We can therefore think of the random process as being defined on the state space $S = \{\phi, 0\}$ where ϕ takes a specific and fixed value. The Markov process is given as a Markov chain with transition probabilities $\mathbf{P}$ where $(\mathbf{P})_{ij}$ denotes the probability of moving from state i to state j, $i, j = 1, 2$.[14] We assume that all elements of $\mathbf{P}$ are strictly positive implying that the chain is ergodic and aperiodic, i.e. regular, and admits a unique stationary distribution $\pi = (\pi_1, \pi_2)$ which fulfills $\pi = \pi\mathbf{P}$. The mean exit time from i is $1/(1-(\mathbf{P})_{ii})$. Following Shorrocks (1978), we define a mobility index $M(\mathbf{P})$ as

$$M(\mathbf{P}) = \frac{n - \mathrm{tr}\mathbf{P}}{n-1}.$$

This index is equal to the reciprocal of the harmonic mean of the mean exit times.[15] This index can be interpreted as measuring the randomness or mobility of the chain.

We specify the parameters β, κ, and σ as before as 0.985, 0.8, and 1, respectively. The value for ϕ ranges from 0 to 4 in steps of length 0.01. Finally, we specify $\mathbf{P}$ as

$$\mathbf{P} = \begin{pmatrix} p & 1-p \\ 0.5 & 0.5 \end{pmatrix}$$

where p ranges from 0.01 to 0.99 in steps of length 0.01. This implies that the mean exit time from state 1 – the state where the central bank responds to inflation – ranges from 1.01 to 100 periods and that the mobility index from 0.1.49 to 0.51.

[14] As shown in the Appendix C, every Markov chain can be represented as a i.i.d. random dynamical system. This point becomes important in the simulation exercises.

[15] The index is actually conceived by Shorrocks (1978) for stochastic matrices with quasi dominant diagonals. This aspect is, however, irrelevant for our purposes.

Next we compute the Lyapunov exponents for different values of ϕ and p, keeping the other parameters fixed. Because of the exponential growth, the numerical computation is not a straightforward task. I therefore use the iterative QR procedure as outlined in DIECI and ELIA (2008) (see Appendix E). For the purpose of this specific simulation exercise we can omit the computation of the Lyapunov spaces because in the case of a determinate model amounts $L^u(\omega) = \mathbb{R}^n$ which corresponds to the case where both Lyapunov exponents are greater than zero. In the case of indeterminacy, L^s and L^u would be one dimensional and thus nontrivial. However, there is a lack of boundary condition to achieve a unique solution. Thus, the knowledge of the Lyapunov spaces are of no help in this situation.[16] The results from this simulation exercise are plotted in Figure 3 which shows the minimal Lyapunov exponent as a function of ϕ and $p = (\mathbf{P})_{11}$. In this plot the black line shows the combinations of ϕ and p such that the minimal Lyapunov exponent equals zero. Below this line the model is indeterminate, above it the model is determinate. The picture clearly documents the negative trade–off between the aggressiveness with which the central bank reacts to inflation and the mean exit time of staying in a regime where the central bank combats inflation: if ϕ is large the mean exit time from state 1 can be low and vice versa. A similar message is conveyed by looking at the corresponding contour map in figure 4. The fuzziness of the zero line (black line in Figure 3) and the level lines in Figure 4 are due the numerical imprecisions. More iterations and higher accuracy would lead to smoother lines, however at the expense of computation time.

6 Conclusion

The purpose of this paper was to present to economists the mathematical tools which enable them to analyze rational expectations models with time-varying coefficients. The theoretical core of this methodology evolves around the concept of Lyapunov exponents which measure the asymptotic growth rates of trajectories. The Multiplicative Ergodic Theorem by Oseledets then showed that the Lyapunov exponents play a similar role in the analysis of the stability of random dynamical systems as the eigenvalues do in the standard case of constant coefficients. Based in this insight, the paper shows how to construct solutions and analyze the stability of rational expectations models with time–varying coefficients. This approach brings the paper close to the spirit of the standard Blanchard–Kahn analysis of rational expecta-

[16] Alternative numerical procedure for the computation of Lyapunov spaces have been proposed by FROYLAND et al. (2013).

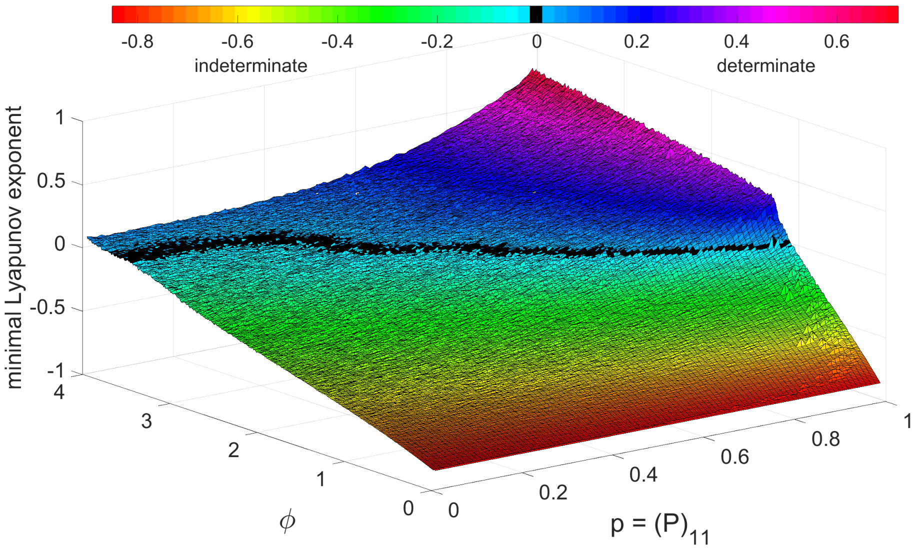

Figure 3: Trade-off between aggressiveness in reacting to inflation (ϕ) and the probability of leaving state 1 (p)

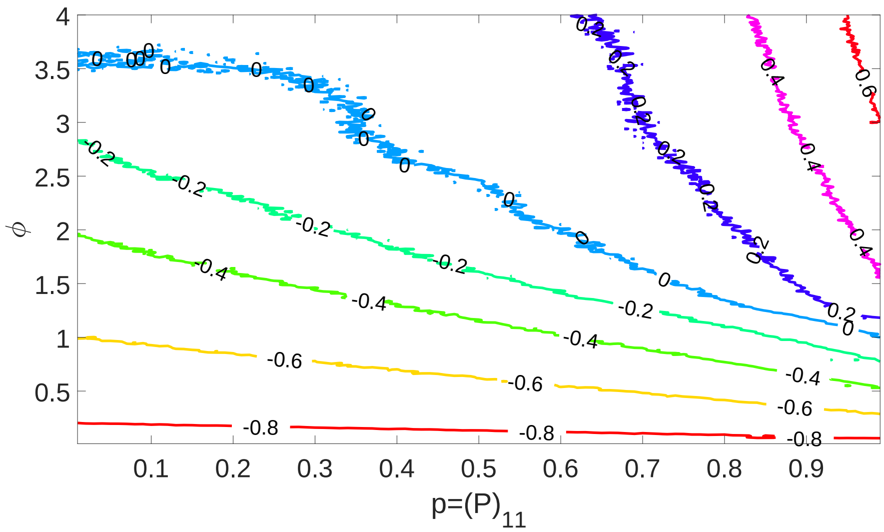

Figure 4: Contour plot of Figure 3. Each level line corresponds to a certain minimal Lyapunov exponent

tions models with constant coefficients. The methodology is also relevant for the analysis of regime–switching models á la Hamilton and complements the approach by Francq and Zakoïan (2001).

The application of these tools requires numerical methods as analytical solutions are almost never available. Fortunately, powerful procedures to estimate the Lyapunov exponents as well as the corresponding Lyapunov spaces have been developed (see Dieci and Elia (2008) and Froyland et al. (2013), f.e.). Finally, the paper runs a simple simulation exercise of a prototype New Keynesian model with Taylor rule to demonstrate the practical usefulness of the approach. From this exercise, it becomes clear that there are no conceptual obstacles to apply this methodology to more sophisticated models.

References

Arnold, Ludwig (2003), *Random Dynamical Systems*, corrected second printing edn., Berlin: Springer-Verlag.

Barthélemy, Jean and Magali Marx (2017), “Solving endogenous regime switching models”, *Journal of Economic Dynamics and Control*, 77, 1–25.

Bellman, Richard (1954), “Limit theorems of non–commutative operations”, *Duke Mathematical Journal*, 21, 491–500.

Berger, Marc A. (1993), *An Introduction to Probability and Stochastic Processes*, New York: Springer–Verlag.

Berman, A. and R. J. Plemmons (1994), *Nonnegative Matrices in the Mathematical Sciences*, vol. 9, Society for Industrial and Applied Mathematics.

Bhattacharya, Rabi and Mukul Majumdar (2007), *Random Dynamical Systems – Theory and Applications*, Cambridge, Massachusetts: Cambridge University Press.

Blanchard, Olivier J. and Charles M. Kahn (1980), “The solution of linear difference models under rational expectations”, *Econometrica*, 48, 1305–1311.

Bougerol, Philippe and Nico Picard (1992), “Stationarity of GARCH processes and some nonnegative time series”, *Journal of Econometrics*, 52, 115–127.

Brandt, Andreas (1986), "The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients", *Advances in Applied Probability*, 18, 211–220.

Campbell, Stephen L. and Carl D. Meyer (1979), *Generalized Inverses of Linear Transformations*, New York: Dover Publications, Inc.

Chen, Xiaoshan, Eric L. Leeper, and Campbell Leith (2015), "US monetary and fiscal policies – conflict or cooperation?", Working Paper 2015–16, Business School – Economics, University of Glasgow.

Cogley, Timothy and Thomas J. Sargent (2005), "Drifts and volatilities: Monetary policies and outcomes in the post WWII US", *Review of Economics Dynamics*, 8, 262–302.

Colonius, Fritz and Wolfgang Kliemann (2014), *Dynamical Systems and Linear Algebra*, vol. 158 of *Graduate Studies in Mathematics*, Providence, Rhode Island: American Mathematical Society.

Diaconis, Persi and David Freedman (1999), "Iterated random functions", *SIAM Review*, 41, 45–76.

Dieci, Luca and Cinzia Elia (2008), "SVD algorithms to approximate spectra of dynamical systems", *Mathematics of Computers in Simulation*, 79, 1235–1254.

Elaydi, Saber N. (2005), *An Introduction to Difference Equations*, 3rd edn., New York: Springer.

Farmer, Roger E. A., Daniel F. Waggoner, and Tao Zha (2009), "Understanding Markov-switching rational expectations models", *Journal of Economic Theory*, 144, 1849–1867.

Farmer, Roger E. A., Daniel F. Waggoner, and Tao Zha (2011), "Minimal state variable solutions to Markov-switching rational expectations models", *Journal of Economic Dynamics and Control*, 35, 2150–2166.

Francq, Christian and Jean-Michel Zakoïan (2001), "Stationarity of multivariate Markov–switching ARMA models", *Journal of Econometrics*, 102, 339–364.

Froyland, Gary, Thorsten Hüls, Gary Morriss, and Thomas M. Watson (2013), "Computing covariant Lyapunov vectors, Oseledets vectors, and dichotomy projectors: A comparative numerical study", *Physica D*, 247, 18–39.

FURSTENBERG, HARRY and HARRY KESTEN (1960), "Products of random matrices", *Annals of Mathematical Statistics*, 31, 457–469.

GALÍ, JORDI (2011), "Are central banks' projections meaningful?", *Journal of Monetary Economics*, 58, 537–550.

GOLDIE, CHARLES M. and ROSS A. MALLER (2000), "Stability of perpetuities", *Annals of Probability*, 28, 1195–1218.

GRIMMETT, GEOFFREY and DAVID STIRZAKER (2001), *Probability and Random Processes*, 3rd edn., Oxford University Press.

HAMILTON, JAMES D. (1989), "A new approach to the economic analysis of nonstationary time series and business cycle", *Econometrica*, 57, 357–384.

HAMILTON, JAMES D. (2016), "Macroeconomic regimes and regime shifts", in *Handbook of Macroeconomics*, John B. Taylor and Harald Uhlig, eds., vol. 2, chap. 3, pp. 163–201, Elsevier Press.

KIM, CHANG-JIN and CHARLES R. NELSON (1999), "Has the U.S. economy become more stable? A Bayesian approach based on a Markov–switching model of the business cycle", *Review of Economics and Statistics*, 81, 608–616.

KLEIN, PAUL (2000), "Using the generalized Schur form to solve a multivariate linear rational expectations model", *Journal of Economic Dynamics and Control*, 24, 1405–1423.

LASÉEN, STEFAN and LARS E. O. SVENSSON (2011), "Anticipated alternative instrument-rate paths in policy simulations", *International Journal of Central Banking*, 7, 1–35.

MEYER, CARL D. (2000), *Matrix Analysis and Applied Linear Algebra*, Philadelphia: Society for Industrial and Applied Mathematics.

OSELEDETS, VALERY I. (1968), "A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems", *Transactions of the Moskow Mathematical Society*, 19, 197–231.

PRIMICERI, GIORGIO E. (2005), "Time-varying structural vector autoregressions and monetary policy", *Review of Economic Studies*, 72, 821–852.

ROCKAFELLAR, R. TYRELL (1970), *Convex Analysis*, Princeton, New Jersey: Princeton University Press.

Sargent, Thomas J. (1999), *Conquest of American Inflation*, Princeton, New Jersey: Princeton University Press.

Shorrocks, Anthony F. (1978), "The measurement of mobility", *Econometrica*, 46, 1013–1024.

Sims, Christopher A. (2001), "Solving linear rational expectations models", *Computational Economics*, 20, 1–20.

A The Uninformativeness of Eigenvalues

This section presents a systematic way to construct examples of nonautonomous linear difference equations such that the "time frozen" or "local stability" does not imply overall or global stability. Thus, we generate examples which "*demonstrate without any doubt that eigenvalues do not generally provide any information about the stability of the nonautonomous difference systems*" (ELAYDI, 2005, p.191).[17] Consider the linear nonautonomous deterministic difference equation

$$x_{t+1} = A_t x_t \qquad \text{with } x_t \in \mathbb{R}^2$$

where $A_t = \exp(tG(\omega))B\exp(-tG(\omega))$, $\omega > 0$, with

$$G(\omega) = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \quad \text{implying } \exp(tG(\omega)) = \begin{pmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{pmatrix}.$$

Because A_t and B are similar matrices, they share the same eigenvalues. Next define $y_t = \exp(-tG(\omega))x_t$. y_t is therefore obtained from x_t by a rotation of angle ωt implying that $\|y_t\| = \|x_t\|$. Thus, x_t diverges if and only if y_t diverges. Using the defining difference equation for x_t, we see that y_t follows the autonomous difference equation

$$\begin{aligned} y_{t+1} &= \exp(-(t+1)G(\omega))x_{t+1} \\ &= \exp(-(t+1)G(\omega)) \underbrace{\exp(tG(\omega))B\exp(-tG(\omega))}_{=A_t} x_t \\ &= \exp(-G(\omega))By_t. \end{aligned}$$

The stability of y_t is determined by the product of the matrices $\exp(-G(\omega))$ and B. If we can find a matrix B and an ω such that the spectral radii $\rho(\exp(-G(\omega))B)$ and $\rho(B)$ are such that $\rho(\exp(-G(\omega))B) > 1$ whereas $\rho(B) < 1$, we have found an example where each of the "time frozen" coefficient matrices would imply stability, but where the nonautonomous system is unstable.

One such specification due to ELAYDI (2005, p. 190) is obtained by taking

$$\omega = 1 \quad \text{and } B = \begin{pmatrix} 0 & 1/2 \\ 3/2 & 0 \end{pmatrix}.$$

[17]This construction translates the continuous time approach of COLONIUS and KLIEMANN (2014, p.109–110) to a discrete time framework. FRANCQ and ZAKOÏAN (2001) provide another ad hoc example.

In this example the eigenvalues of B are $\pm\sqrt{3}/2$, thus they are both smaller than one in absolute terms, but $\exp(-G(\omega))B$ has eigenvalues $1.3836 > 1$ and -0.5421. x_t therefore diverges although every A_t has eigenvalues with modulus strictly smaller than one.

Another specification is obtained by taking

$$\omega = 2 \quad \text{and } B = \begin{pmatrix} 1/10 & 1/2 \\ -3/2 & 1/10 \end{pmatrix}.$$

In this case the eigenvalues of B are $0.1 \pm \imath\sqrt{3}/2$ whose moduli are strictly smaller than one. The eigenvalues of $\exp(-G(\omega))B$, however, are 1.3307 and -0.5711. Thus again, x_t diverges although every A_t has eigenvalues with modulus strictly smaller than one.

B Linear Nonautonomous Systems: A Numerical Example

To get a better understanding of the behavior of periodically switching linear difference equations, we consider the following numerical example for $n = 2$ and $p = 2$:

$$A_0 = \begin{pmatrix} 1 & 0.2 \\ 1 & 1 \end{pmatrix}, \qquad A_1 = \begin{pmatrix} 1 & -0.5 \\ 3 & -2 \end{pmatrix}.$$

These matrices have eigenvalues $\lambda_1^{(0)} = 1.4472$ and $\lambda_2^{(0)} = 0.5528$, respectively $\lambda_1^{(1)} = 0.3660$ and $\lambda_2^{(1)} = -1.3660$. The matrices A_1A_0 and A_0A_1 are then given by

$$A_1A_0 = \begin{pmatrix} 0.5 & -0.3 \\ 1.0 & -1.4 \end{pmatrix}, \qquad A_0A_1 = \begin{pmatrix} 1.6 & -0.9 \\ 4.0 & -2.5 \end{pmatrix}.$$

The Lyapunov, respectively the Floquet, exponents in this case can be computed from the eigenvalues α_1 and α_2 of A_1A_0, respectively A_0A_1, as $\lambda_j = \frac{1}{2}\log|\alpha_j|$, $j = 1, 2$. This gives $\lambda_1 = -0.5601$ and $\lambda_2 = 0.1020$. From the eigenvectors of A_1A_0 and A_0A_1 we can compute the stable and the unstable bundle:

$$L^s(0) = \operatorname{span}\begin{pmatrix} 0.8653 \\ 0.5013 \end{pmatrix} \qquad L^u(0) = \operatorname{span}\begin{pmatrix} 0.1712 \\ 0.9852 \end{pmatrix}$$

and

$$L^s(1) = \operatorname{span}\begin{pmatrix} 0.5770 \\ 0.8167 \end{pmatrix} \qquad L^u(1) = \operatorname{span}\begin{pmatrix} 0.3034 \\ 0.9529 \end{pmatrix}.$$

The corresponding projection matrices are

$$\pi^{s,u}(0) = \begin{pmatrix} 1.1119 & -0.1932 \\ 0.6442 & -0.1119 \end{pmatrix} \quad \text{and} \quad \pi^{u,s}(0) = \begin{pmatrix} -0.1119 & 0.1932 \\ -0.6442 & 1.1119 \end{pmatrix},$$

and

$$\pi^{s,u}(1) = \begin{pmatrix} 1.8205 & -0.5797 \\ 2.5766 & -0.8205 \end{pmatrix} \quad \text{and} \quad \pi^{u,s}(1) = \begin{pmatrix} -0.8205 & 0.5797 \\ -2.5766 & 1.8205 \end{pmatrix}.$$

Suppose $b_t = (b, (-1)^t b)'$, then we can compute the solution (2.6).

C Markov Chains as Random Dynamical Systems

Denote the state space by S which consists of the matrices $\{A_t\}$ used to move from x_t to $x_{t+1} = A_t x_t + b_t$. Assume S to be finite with cardinality $s = |S|$. Take Ω to be the set of all maps from S into itself. These maps can be represented by $\{0,1\}$–matrices with exactly one element equal to one in each row. These matrices are sometimes called *deterministic transition matrices*. According to the Krein–Milman theorem, respectively the Birkhoff–von Neumann theorem, (see Rockafellar (1970) or Berman and Plemmons (1994, p. 49)) every transition matrix $\mathbf{P}$ can be represented as a convex combination of the deterministic transition matrices:

$$\mathbf{P} = \sum_{j=1}^{s^s} \delta_j D_j, \qquad \sum_{j=1}^{s^s} \delta_j = 1 \text{ and } \delta_j \geq 0$$

where D_j are the deterministic transition matrices. Note that this representation is not unique. In the case $s = 2$ and $p+q < 1$, a possible representation of $\mathbf{P}$ is given by

$$\mathbf{P} = \begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} = p \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + q \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + (1-p-q) \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Similarly, for the case $p + q > 1$. Note that we need only three deterministic transition matrices to represent $\mathbf{P}$ in this case. For a given transition matrix $\mathbf{P}$, let Ω be the smallest set of deterministic transition matrices necessary to represent $\mathbf{P}$ as a convex combination $\mathbf{P} = \sum_{j=1}^{|\Omega|} \delta_j D_j$ with $\sum_{j=1}^{|\Omega|} \delta_j = 1$ and $\delta_j > 0$. $(\Omega, \mathcal{P}(\Omega), \Delta)$ is then a probability space where $\mathcal{P}(\Omega)$ is the σ–algebra given by all subsets of Ω and Δ is the discrete probability measure generated by the point probabilities $(\delta_1, \ldots, \delta_{|\Omega|})$. The chain is then generated by randomly drawing a deterministic transition matrix where the probabilities are given by Δ. If D_j is drawn, the system moves deterministically to the next state according to D_j. These draws are independent from one period to the next. In this way every Markov chain can be represented as a random dynamical system with i.i.d increments.

D A Digression: The Case of Triangular Matrices

To get a better understanding of the algorithm used for the estimation of the Lyapunov exponents, it is instructive to examine the case of triangular matrices. This section replicates the expositions of Berger (1993, p. 155) and Arnold (2003, p. 129–130). Let $\{A_t\}$ be a sequence of 2×2 upper triangular matrices:

$$A_t = \begin{pmatrix} a_t & c_t \\ 0 & b_t \end{pmatrix} \in \mathbb{GL}(n).$$

$\Phi(t)$ is then given by

$$\Phi(t) = A_{t-1} \dots A_1 A_0 = \begin{pmatrix} \prod_{j=0}^{t-1} a_j & \sum_{k=0}^{t-1} a_{t-1} \dots a_{k+1}\, c_k\, b_{k-1} \dots b_0 \\ 0 & \prod_{j=0}^{t-1} b_j \end{pmatrix}.$$

Note that $\mathbb{R}e_1 = \text{span}(1,0)'$ is an invariant subspace for the $\Phi(t)$'s. We assume that the sequences $\{a_t\}$, $\{b_t\}$, and $\{c_t\}$ are ergodic with $\alpha = \mathbb{E}\log|a_t|$, $\beta = \mathbb{E}\log|b_t|$, and $\gamma = \mathbb{E}\log|c_t| \in \mathbf{L}^1$. Therefore

$$\frac{1}{t}\sum_{j=0}^{\infty} \log|a_t| \to \alpha$$
$$\frac{1}{t}\sum_{j=0}^{\infty} \log|b_t| \to \beta$$

which implies that

$$\frac{1}{t}\log|\det\Phi(t)| \to \alpha + \beta.$$

$\alpha+\beta$ is twice the average Lyapunov exponent λ_Σ defined as $\lambda_\Sigma = (\lambda_1+\lambda_2)/2$. Obviously, the Lyapunov exponent of $[\Phi(t)]_{11}$ is α and that of $[\Phi(t)]_{22}$ β. Because $\lambda(x+y) \leq \max\{\lambda(x), \lambda(y)\}$ with equality if $\lambda(x) \neq \lambda(y)$ (Arnold, 2003, p. 114),

$$\lambda_1 = \max\{\alpha,\beta\} > \frac{\alpha+\beta}{2} > \lambda_2 = \min\{\alpha,\beta\} \quad \text{for } \alpha \neq \beta.$$

When $\alpha = \beta$, we have $\lambda_1 = \lambda_\Sigma = \alpha = \beta$ with multiplicity 2.

To compute the Oseledets or Lyapunov spaces, we assume without loss of generality $\alpha > \beta$. For any vector $x = (x_1, 1)'$ to grow at rate β, x must be an eigenvector with respect to eigenvalue $b(t) = [\Phi(t)]_{22}$. Thus

$$\begin{pmatrix} a(t) & c(t) \\ 0 & b(t) \end{pmatrix} \begin{pmatrix} x_1 \\ 1 \end{pmatrix} = b(t) \begin{pmatrix} x_1 \\ 1 \end{pmatrix}$$

with $a(t) = [\Phi(t)]_{11}$ and $c(t) = [\Phi(t)]_{12}$. Taking limits and recognizing that $b(t)/a(t) \to 0$, we get

$$x_1 = -\frac{c(t)}{a(t)} = -\sum_{k=0}^{\infty} \frac{c_k b_{k-1} \ldots b_1 b_0}{a_k \ldots a_1 a_0}.$$

This defines the random Lyapunov subspace $L(\lambda_2) = \text{span}(x_1, 1)'$. The other Lyapunov subspace is $L(\lambda_1) = \mathbb{R}e_1$. Note that for x to grow at rate β, x must be random. Moreover, this randomness depends on the entire sequence $\{A_t\}$.

E Computing Lyapunov Exponents and Lyapunov Spaces

Although the Lyapunov exponents and the corresponding subspaces are defined in a straightforward manner on the theoretical level, it is not a straightforward task to compute them numerically. The reason for this difficulty stems from the exponential growth of the elements in $\Phi(t)$, respectively Υ, as t becomes large. Trying to compute these matrices directly very quickly hits the numerical bounds of any computer. To avoid this problem iterative QR and SVD decompositions have been proposed (see Dieci and Elia, 2008).

In this paper, we use the QR approach which is very easy to implement. Let $X(t)$ be the principal fundamental matrix satisfying $X(t+1) = A_t X(t)$. Suppose we initialize the algorithm by taking some $X(0)$ as a starting value. Let the QR decomposition of $X(0)$ be given as $X(0) = Q_0 R_0$ where Q_0 is an orthogonal matrix and R_0 an upper triangular matrix. Then we compute $X(1) = A_0 X(0)$ and perform the QR decomposition of $X(1)Q_0 = Q_1 R_1$. Obviously, $X(1)X(0) = Q_1 R_1 R_0$. Proceeding in this way, we obtain a QR decomposition of $\Phi(t) = X(t) \ldots X(1)X(0)$:

$$\Phi(t) = Q_t R_t \ldots R_1 R_0.$$

Generalizing the arguments of appendix D from 2×2 triangular matrices to $n \times n$ triangular matrices, one gets

$$\lambda_j = \lim_{t \to \infty} \frac{1}{t} \sum_{j=k}^{t} \log [R_k]_{jj}, \quad j = 1, \ldots \ell,$$

where $[R_k]_{jj}$ is the j-th diagonal element of R_k. The algorithm stops when a sufficient precision is obtained. For further details see Dieci and Elia (2008).