

CEP 18-11

**Saddle Cycles: Solving Rational Expectations
Models Featuring Limit Cycles (or Chaos) Using
Perturbation Methods**

Dana Galizia
Carleton University

September 10th, 2018

CARLETON ECONOMIC PAPERS



Department of Economics

1125 Colonel By Drive
Ottawa, Ontario, Canada
K1S 5B6

Saddle Cycles: Solving Rational Expectations Models Featuring Limit Cycles (or Chaos) Using Perturbation Methods

Dana Galizia*[†]

July 2018

Version 5.0

Abstract

Unlike their linear counterparts, non-linear models of the business cycle can generate sustained economic fluctuations even in the absence of shocks (e.g., via limit cycles or chaos). A popular approach to solving non-linear models is the use of perturbation methods. I show that, as typically implemented, these methods are generally incapable of finding solutions that feature limit cycles or chaos, a fact that does not appear to be recognized in the existing literature. Standard algorithms only seek solutions that feature convergence to the steady state, which is stronger than the standard definitional requirement that a solution simply cannot explode. Because of this, in estimation exercises any parameterization that involves limit cycles would typically (and incorrectly) be discarded. I propose a modification to standard algorithms that does not impose the overly strong requirement that solutions involve convergence.

Key Words: Dynamic equilibrium economies; Computational methods; Non-linear solution methods; Limit cycles; Chaos; JEL Class.: C63; C68; E37

*Department of Economics, Carleton University

[†]Thanks to Paul Beaudry and Franck Portier for helpful comments. All errors are mine.

1 Introduction

Rational expectations models where non-linearities play an important role have become increasingly used in macroeconomics—and, in particular, in business cycle research—in recent years, both in theoretical and quantitative settings. Unlike their linear counterparts, non-linear models are in some cases capable of generating sustained fluctuations in economic aggregates even in the absence of stochastic forces, such as through limit cycle dynamics.¹ As argued in Beaudry, Galizia, and Portier [2016] and Beaudry, Galizia, and Portier [2018], models that combine limit cycles with traditional stochastic forces are capable of generating strong endogenous propagation, with the stochastic forces contributing only modestly to the persistence and volatility of the endogenous variables, while still producing business cycles that are as unpredictable and irregular as those found in the data.² Given that a common criticism of many existing models of the business cycle is their perceived over-reliance on unobserved (and often empirically unjustified) exogenous stochastic forces in matching the persistence and volatility found in the data, models featuring limit cycles may be a promising avenue for business cycle research.

For a non-linear rational expectations model featuring a limit cycle (or chaos) to be useful for quantitative applications, one must be able to (at least approximately) obtain a solution to it. There is a substantial body of literature devoted to the development of approximate solution methods for non-linear rational expectations models. These methods vary in the complexity of their implementation, their computational burden, and in the accuracy of the resulting approximation. As a number of studies have noted (e.g., Gaspar and Judd [1997], Aruoba, Fernández-Villaverde, and Rubio-Ramírez [2006]), one of these methods—the perturbation (i.e., Taylor approximation) method—is relatively simple to implement, can in many cases produce a reasonable degree of accuracy, and generates a substantial savings in computation time relative to many popular alternatives, making it a good candidate for

¹Another possible type of deterministic fluctuations is chaotic dynamics. This paper focuses on limit cycles, which appear to be the more empirically plausible case, but the proposed solution method will apply without modification to models featuring chaos.

²These papers were by no means the first to explore the relevance of limit cycles and/or chaos to macroeconomic fluctuations, and indeed this literature has a long history (see the references contained in Beaudry, Galizia, and Portier [2018]). However, combining such forces with more conventional stochastic ones appears to have gone largely unexplored in that earlier literature.

solving larger models and/or for use in estimation exercises. Given these properties, perturbation methods would seem like an attractive choice for solving models that may feature limit cycles. However, as discussed in detail in this paper, existing perturbation algorithms³ are incapable of dealing with these situations. This fact—which, to my knowledge, this paper is the first to point out—has at least two important consequences. First, the potentially fruitful research avenue noted above may have gone largely unexplored thus far at least in part because of the unavailability of methods to solve models featuring limit cycles. Second, and perhaps more deleteriously, there could in fact be any number of existing models in the literature that are capable of generating limit cycles for certain parameterizations, but for which (a) such parameterizations have been discarded because the researcher (or the software package used by the researcher) incorrectly believes that there is no solution to the model, or (b) in addition to the solution(s) found by the researcher, there are additional solutions featuring limit cycles that have been inadvertently ignored.

The reason standard perturbation methods are generally incapable of solving models featuring attractive limit cycles—that is, limit cycles to which the system tends to converge in the absence of any shocks (as long as it does not begin exactly at the steady state)⁴—is straightforward. First, note that a solution that features an attractive limit cycle will generally feature a steady state that is locally unstable, since arbitrarily small initial deviations from the steady state will result in convergence to the limit cycle, rather than back to the steady state. Note also that, by definition, a solution is typically required both to satisfy a set of model equations, and also to result in trajectories that do not explode, i.e., that satisfy

³This paper deals specifically with perturbation methods that yield general (approximate) solutions to DSGE models in the form of the coefficients of a Taylor expansion around a single point. See, for example, the perturbation methods discussed in Judd [1996], Judd and Guu [1997], Schmitt-Grohé and Uribe [2004], Aruoba, Fernández-Villaverde, and Rubio-Ramírez [2006], Gomme and Klein [2011], Kollman, Maliar, Malin, and Pichler [2011], and Caldara, Fernández-Villaverde, Rubio-Ramírez, and Yao [2012]. See Fernández-Villaverde, Rubio-Ramírez, and Schorfheide [2016] (section 4) for a recent comprehensive overview of these methods. In contrast, this paper does not discuss methods that, for example, involve taking perturbations around the present location of the system in the phase space, and which are thus suited primarily to solving for a single transition path rather than a general solution to the model (e.g., Mennuni and Stepanchuk [2018]).

⁴Models with repellant limit cycles can also exhibit interesting dynamics in which trajectories beginning near the steady state converge to it, those beginning far enough away produce relatively sudden collapses or explosions, and the border between these two regions is an unstable limit cycle. Given the complicated nature of these dynamics, perturbation methods may not be well suited to solving such models.

a transversality condition (TVC). Next, as is well known,⁵ to obtain a non-linear perturbation approximation, one must in practice obtain as a first step the linear (i.e, first-order) approximation to the solution, before sequentially obtaining the second-order approximation, then third-order, etc., up to the desired order. Further, given the linear approximation, the higher-order approximations are typically pinned down uniquely, with each set of k -th-order derivatives ($k \geq 2$) given by the solution to a square linear system whose coefficients are functions of the $(k - 1)$ -th- and lower-order derivatives. Thus, the problem of computing a non-linear approximation to a solution fundamentally boils down to choosing the first-order coefficients, i.e., to finding the correct linear approximation (or set of linear approximations in the case of multiple solutions). Standard solution algorithms typically equate this set of linear approximations to the set of solutions to the *linearized* model. However, the latter set typically consists of solutions that not only satisfy the TVC, but that in fact meet the stronger requirement of converging to the steady state:⁶ in a linear environment, trajectories that do not converge typically explode, which would violate the TVC for the linearized model and therefore not be a solution to it.⁷ As noted above, however, this generally excludes solutions featuring attractive limit cycles, since they typically do not feature convergence to the steady state. Thus, by searching only for linear approximations to the solution that are also solutions to the linearized model, one is fundamentally being overly restrictive, the upshot of which is that one cannot generally discover solutions featuring limit cycles. In practice, when the only solutions to a model involve limit cycles, typical software packages (e.g., the widely used MATLAB add-in Dynare) would usually stop and report an error for a single solution attempt, while in estimation exercises the associated parameterization would typically be discarded (or ruled out in advance by parameter restrictions). This discarding

⁵See, e.g., Judd [1998].

⁶Essentially, this is the content of the well known Blanchard and Kahn [1980] condition, which states that the number of (endogenous) eigenvalues outside the complex unit circle must be no greater than the number of jump variables in order for a solution to a linear(ized) model to exist.

⁷Under special circumstances—i.e., where the linearized model has one or more (endogenous) eigenvalues lying exactly on the complex unit circle—there may also exist trajectories of the linearized model that neither converge nor explode. In these cases, there may exist solutions featuring limit cycles that could be discovered by standard perturbation algorithms. For the sake of simplifying the presentation, I henceforth ignore such “endogenous unit root” cases since they would typically correspond to a zero measure of the parameter space, and thus be largely irrelevant in practice. Nonetheless, the method proposed in this paper will deal automatically with such cases without modification.

of potential limit cycle parameterizations is in turn the basis for the assertion above that there may very well be existing models of the business cycle that are capable of generating limit cycles, but that the relevant parameterizations have been inadvertently discarded.

As described above, the fundamental source of the problem with standard perturbation methods is their implicit assumption that the following two sets are equal: (i) the set of linear approximations to the (non-linear) solutions of the model; and (ii) the set of solutions to the linearized model. In fact, while solutions in (ii) are also solutions in (i), the reverse is not generally true, and indeed these two sets are only equal if all the (non-linear) solutions to the model result in convergence. If, on the other hand, there is a solution to the model that features a limit cycle, its linear approximation will *not* be a solution to the linearized model. This fact does not appear to be fully understood in the literature. For example, in their recent Handbook of Macroeconomics chapter, Fernández-Villaverde, Rubio-Ramírez, and Schorfheide [2016] state, “For many DSGE models, we will have exactly \tilde{n} stable generalized eigenvalues and the stable solution would also be unique. If we have too few stable generalized eigenvalues, the equilibrium dynamics will be inherently unstable. If we have too many, we can have sunspots... Suffice it to note here that all these issues would depend only on the first-order approximation and that going to higher-order approximations would not change the issues at hand. If we have uniqueness of equilibrium in the first-order approximation, we will also have uniqueness in the second-order approximation.” As the present paper makes clear, in fact there may in general be solutions to the model that require taking higher-order approximations in order to discover.

Geometrically, the goal of any perturbation method is to find a surface (i.e., a manifold) in the phase space that (a) is of dimension at least as great as the number of pre-determined variables; and (b) is such that the system remains bounded when restricted to this manifold (so that it satisfies the TVC), but explodes in all other areas of the phase space. A typical solution is then expressed as a function that projects the system onto this manifold by choice of the jump variables. As this paper shows, there is in principle no reason why one cannot employ perturbation methods to this problem without imposing the stronger condition that trajectories on this manifold converge to the steady state; that is, it proposes a method that does not assume that the sets (i) and (ii) described above are equal. This allows for the

possibility of finding “saddle cycles”; that is, a configuration where trajectories on this manifold feature limit cycles, but trajectories off of it explode. Note that the proposed method is a strict generalization of standard methods, and in particular it can be applied without modification to models that do not feature limit cycles. Thus, one need not determine in advance whether limit cycles are relevant in order to apply the method. This is useful, since it can, through an estimation exercise, allow the data to determine whether and to what extent limit cycles may be relevant.

As noted above, perturbation methods are not the only possible way to solve models featuring limit cycles. In practice, there are at least two classes of solutions that could in principle be used. The first class involves approximately solving a system of equations made up of the first-order conditions, constraints, and equilibrium conditions of the model, where the solution gives the jump variables as functions of the pre-determined and exogenous variables. Within this class, there are two principal methods that have been used to obtain such an approximate solution: perturbation methods, which are used in this paper, and the projection methods formally introduced to economics by Judd [1992] and Gaspar and Judd [1997] (among others). The main drawback to this class of approaches is that, even when a unique solution to the model exists, there will generally be multiple solutions to this system of equations. The particular solution of interest would be the one that also satisfies the TVC. Thus, in principle one needs to find every solution to the system of equations and check them for conformity with the TVC. Finding every solution to the system is relatively simple in the case of perturbation methods—and indeed, this is an important part of the method proposed here—but more difficult for projection methods, which generally rely on computationally expensive numerical solvers that would typically have to be initialized and tuned appropriately in order to ensure convergence to a given desired solution.⁸

The second class of solution methods are based on solving for the value functions using the Bellman equations of the model. The principal drawback of these methods are their computational burden: they suffer acutely from the well known curse of dimensionality. This can be especially problematic in cases where the first welfare theorem fails to hold, since

⁸One approach would be to first obtain an approximate solution using perturbation methods, and then use it as an initial guess for a (generally more accurate) projection method. Obviously, this approach still requires an algorithm for finding a perturbation solution.

one must typically solve for private value and policy functions taking the aggregate state as given, and then search numerically for the aggregate state that produces an equilibrium (where for each different aggregate state checked, the private value and policy functions would have to be found anew). Since limit cycles may be especially relevant in cases where there are strategic complementarities⁹ (and therefore the welfare theorems would fail), this is precisely an application in which value function methods would be especially computationally expensive. Furthermore, complementarities may in fact cause the search problem to become non-convex, in which case numerical algorithms are not guaranteed to converge. For these reasons, the perturbation methods discussed in this paper may be the most tractable available approach for solving such models.

The remainder of the paper is organized as follows.¹⁰ Section 2 outlines the main elements of the solution algorithm as it applies to deterministic models. While modern business cycle analysis is conducted mainly using stochastic models, focusing first on the deterministic case allows the basic ideas underlying the method to be conveyed in the simplest possible setting. As part of the presentation, I show how the proposed method is fundamentally a strict generalization of existing methods. I also illustrate how to apply the method to a non-linear rational expectations model based on Beaudry, Galizia, and Portier [2018]. The model is simple enough that it can be solved analytically using our method, and is capable of generating limit cycles as the unique rational expectations equilibrium, so that standard perturbation methods would erroneously conclude that there is no solution. Section 3 then discusses how one can implement the method in practice. Section 4 extends the discussion to the more general stochastic case, and finally Section 5 concludes.

2 The Deterministic Case

While most modern business cycle models feature stochastic elements, to make the exposition as simple as possible we begin with the case of a deterministic set-up. Once we have established the results and intuition for this case, it will be straightforward to extend them

⁹See Beaudry, Galizia, and Portier [2018].

¹⁰Note that, for considerations of length, it will generally be assumed that the reader has a basic familiarity with standard non-linear perturbation algorithms (both theory and implementation). See Fernández-Villaverde, Rubio-Ramírez, and Schorfheide [2016], section 4, for an overview of these standard methods.

to the stochastic case, a task undertaken in Section 4.

Note that, throughout the paper, a number of assumptions are made. In most cases, these assumptions fall into one of two categories: (1) unnecessary (and easily relaxed) assumptions made only for ease of presentation, and (2) assumptions that are necessary for all perturbation algorithms, not just the one proposed here. For the latter category, it is often the case in the literature that these assumptions are not explicitly stated. I state them here for completeness, but the reader should be aware that these do not represent extra restrictions imposed by the current method over and above those imposed in other perturbation methods.

2.1 Set-Up

Suppose the economy evolves deterministically according to the potentially non-linear dynamic relation

$$\Gamma(x_{t+1}, x_t) = 0, \tag{1}$$

where $x_t \in \mathbb{R}^n$ is a vector of endogenous variables and $\Gamma: \mathbb{R}^{2n} \rightarrow \mathbb{R}^n$ is a locally C^∞ function.¹¹ We will take Γ as primitive (and known to the researcher). We impose the following assumption.

Assumption (A1). $\Gamma(x, x) = 0$ has a unique solution given by \bar{x} , and for any x_t there is a unique value of x_{t+1} that solves (1).

The first part of Assumption (A1) ensures that the system has a unique steady state.¹² This steady state will serve as the natural point around which to perturb the system. The second part of Assumption (A1), meanwhile, requires both that, given x_t , a solution for x_{t+1} exists (an unrestrictive assumption in most settings) and that this solution is unique. The latter in turn requires as a first step that all “static” relationships (i.e., those involving

¹¹In practice, we only require that Γ be C^k , where k is at least as big as the desired perturbation order.

¹²We can replace this part of (A1) with the less stringent assumption that there is only one *relevant* steady state; that is, that there is only one steady state of practical interest. This will allow the method to address, for example, models in which capital is a necessary input to production, which typically feature as an uninteresting steady state the one associated with no capital. Note also that, since Γ does not depend on t , the existence of a (non-trivial) steady state also implies that there is no trend growth in this economy. As is well known, an economy exhibiting balanced growth can be put into the form (1) by an appropriate change of variables. We assume throughout that this has been done already.

only date- t variables) be eliminated from the system.¹³ Without loss of generality we also henceforth make the normalization that $\Gamma(0,0) = 0$, so that $\bar{x} = 0$.

The above properties imply that we can write

$$x_{t+1} = h(x_t) , \tag{2}$$

for some function h with $h(0) = 0$. While h may not initially be known to the researcher (and indeed may not have a closed-form solution), its derivatives at $x = 0$ —and thus its Taylor series representation at that point—can be computed in a straightforward manner from the (known) derivatives of Γ .¹⁴ Since we will ultimately solve the model using perturbation methods, these Taylor series coefficients will be sufficient for our purposes. We thus henceforth take (2) as our primitive system.

Next, let $m_0 \subset \mathbb{R}^n$ be the set of all x 's such that if $x_0 = x$ then the sequence $\{x_t\}$ obtained from (2) satisfies the transversality condition (TVC) $\limsup_{t \rightarrow \infty} \|x_t\| < \infty$ (i.e., the system does not explode).¹⁵ We make the basic regularity assumption that h is such that m_0 is a locally analytic manifold,¹⁶ and let n_0 denote its dimension. Note that m_0 is clearly h -invariant,¹⁷ so that it can be equivalently characterized as the set of all trajectories that do not explode. We will thus refer to m_0 in what follows as the “non-explosive manifold” (NEM), and note that all trajectories not beginning on the NEM necessarily explode.

As is often the case, it will be useful to partition the state vector as $x = (y, z)$, where $y \in \mathbb{R}^{n_y}$ corresponds to the set of variables that are pre-determined at date t , $z \in \mathbb{R}^{n_z}$ corresponds to the set of variables free to change at date t (i.e., the jump variables), and $n_y + n_z = n$. We are now in a position to define a solution to the model given in (2) as

¹³The assumption of a unique solution for x_{t+1} given x_t was present in the linear setting of Blanchard and Kahn [1980]. It is nonetheless straightforward to extend the analysis to the more general case (e.g., using the techniques of King and Watson [1998], Klein [2000], or Sims [2002]), though at the cost of complicating the presentation and, in practice, increasing the computational burden (potentially significantly).

¹⁴This is a straightforward application of the implicit function theorem. Note that we are assuming here that a Taylor series representation of h at $x = 0$ exists (i.e., h is locally analytic).

¹⁵This TVC is slightly stronger than those often encountered in the literature, which generally allow $\|x_t\|$ to possibly become unbounded as long as it does not grow too quickly (e.g., as long as it grows less than exponentially). Since there is no long-run trend in our set-up, in most cases such an assumption would be sufficient to ensure that indeed $\|x_t\|$ remains bounded. We therefore adopt the latter as our TVC for simplicity.

¹⁶A q -dimensional manifold is (locally) analytic if it can be expressed as the image of \mathbb{R}^q through a (locally) analytic function. Note that this assumption is also implicitly made by all standard perturbation algorithms.

¹⁷A manifold m is said to be h -invariant if $x \in m$ implies $h(x) \in m$.

follows: a solution is a locally analytic function ϕ such that $(y, \phi(y)) \in m_0$ for any y . In other words, given the initial value of the pre-determined variables y_0 , a solution is a value of z_0 that “projects” the system onto the NEM.¹⁸ Next, for a function ψ that maps a q -element subvector \tilde{y} of x into the complementary $(n - q)$ -element subvector \tilde{z} ,¹⁹ we will say that ψ is an *invariant function* if, for any \tilde{y}_t , setting $\tilde{z}_t = \psi(\tilde{y}_t)$ implies $\tilde{z}_{t+1} = \psi(\tilde{y}_{t+1})$, where \tilde{y}_{t+1} and \tilde{z}_{t+1} are the corresponding elements of x_{t+1} obtained from (2). We then have the following proposition, the content of which will be familiar to most macroeconomists.

Proposition 1.

- (i) If $n_0 < n_y$, then there is no solution.
- (ii) If $n_0 = n_y$, then there is a unique solution ϕ ,²⁰ and that solution is invariant.
- (iii) If $n_0 > n_y$, then there are an infinity of solutions.

Proof. All proofs in Appendix A. □

Lastly, let $A \equiv D_x^h$,²¹ so that the linearized version of (2) is given by

$$x_{t+1} = Ax_t, \tag{3}$$

As is the case for standard perturbation algorithms, this linearized system will play an important role here. For ease of exposition, we will assume in what follows that none of the eigenvalues of A lie on the complex unit circle.²²

2.2 Description of the Method

We propose here a method to (approximately) obtain the solution ϕ when it exists uniquely (i.e., when $n_0 = n_y$), and to determine when this is in fact the case. The method is based on

¹⁸We will make the regularity assumption that for any value of y_j (where y_j is the j -th element of y), there is an $x \in m_0$ with $x_j = y_j$. If this were not the case, then no solution could ever exist.

¹⁹Note that we do not require the elements of \tilde{y} or \tilde{z} to be consecutive elements of x .

²⁰Without placing further restrictions on the model, it is technically possible for there to be multiple solutions in this case. However, these solutions, which are by definition locally analytic, would all have the same Taylor series representation. Since we will only be able to solve for the Taylor series representation, these solutions would for our purposes be identical.

²¹For a function $F(\cdot)$ that takes arguments x_1, x_2, \dots, x_k , we use the notation $D_{x_j}^F$ to denote the Jacobian of F with respect to x_j , evaluated at the point $(x_1, x_2, \dots, x_k) = 0$.

²²It is straightforward—though somewhat tedious—to relax this assumption.

the standard perturbation algorithms typically used in the literature, which are themselves fundamentally based on the method introduced by Blanchard and Kahn [1980] for linear models²³ (i.e., models for which h is linear). The method works for linear models as follows. First, under our assumption that A has no unit-modulus eigenvalues, the stable manifold of A —that is, the A -invariant linear subspace²⁴ made up of the set of all trajectories that converge linearly to the steady state—is precisely the set of trajectories that remain bounded; that is, in the linear case with no unit roots, the NEM is given by the stable manifold. Further, using standard eigenvalue-eigenvector methods, one can show that the dimension of the stable manifold (i.e., n_0) is given by the number of stable eigenvalues of A , and if this is equal to n_y , one can easily obtain an analytical expression for the unique solution ϕ .

Every linear dynamic model possesses a stable manifold, and, essentially, Blanchard and Kahn’s [1980] method is a way to solve for it. It turns out that every (differentiable) non-linear dynamic model also possesses a stable manifold,²⁵ which is itself generally non-linear. In fact, when one uses the Blanchard and Kahn [1980] method (or one based on it) to obtain a unique solution to a linear *approximation* to a non-linear model, one is effectively solving for the linear approximation to the non-linear stable manifold; that is, one is solving for the linear subspace that is tangent to the stable manifold at the steady state. Assuming a unique solution to the linearized model exists, standard higher-order perturbation methods²⁶ begin with that linear (i.e., first-order perturbation) approximation and use it to sequentially obtain higher-order approximations to the stable manifold.

Clearly, then, standard perturbation methods are only capable of discovering solutions that lie on the (non-linear) stable manifold. However, in models featuring limit cycles (or chaos), it is possible for a solution to the non-linear model to exist without the same being true of the linearized model. In particular, the NEM necessarily contains the stable manifold,

²³A number of papers (e.g., King and Watson [1998], Klein [2000], Sims [2002]) have proposed alternative methods for solving linear rational expectations models. These alternatives make a number of important improvements in generality, ease of implementation, and computational speed and reliability over Blanchard and Kahn [1980], and as such have largely supplanted Blanchard and Kahn [1980] in most practical situations. Under the assumptions made in Section 2.1, however, it can be verified that these methods are mathematically equivalent to one another, though in practice one may encounter some numerical discrepancies due to imperfect computational precision.

²⁴By A -invariant we mean invariant with respect to the function $x \mapsto Ax$.

²⁵See, for example, Kuznetsov [1998], Theorem 2.3.

²⁶For example, those described in Fernández-Villaverde, Rubio-Ramírez, and Schorfheide [2016].

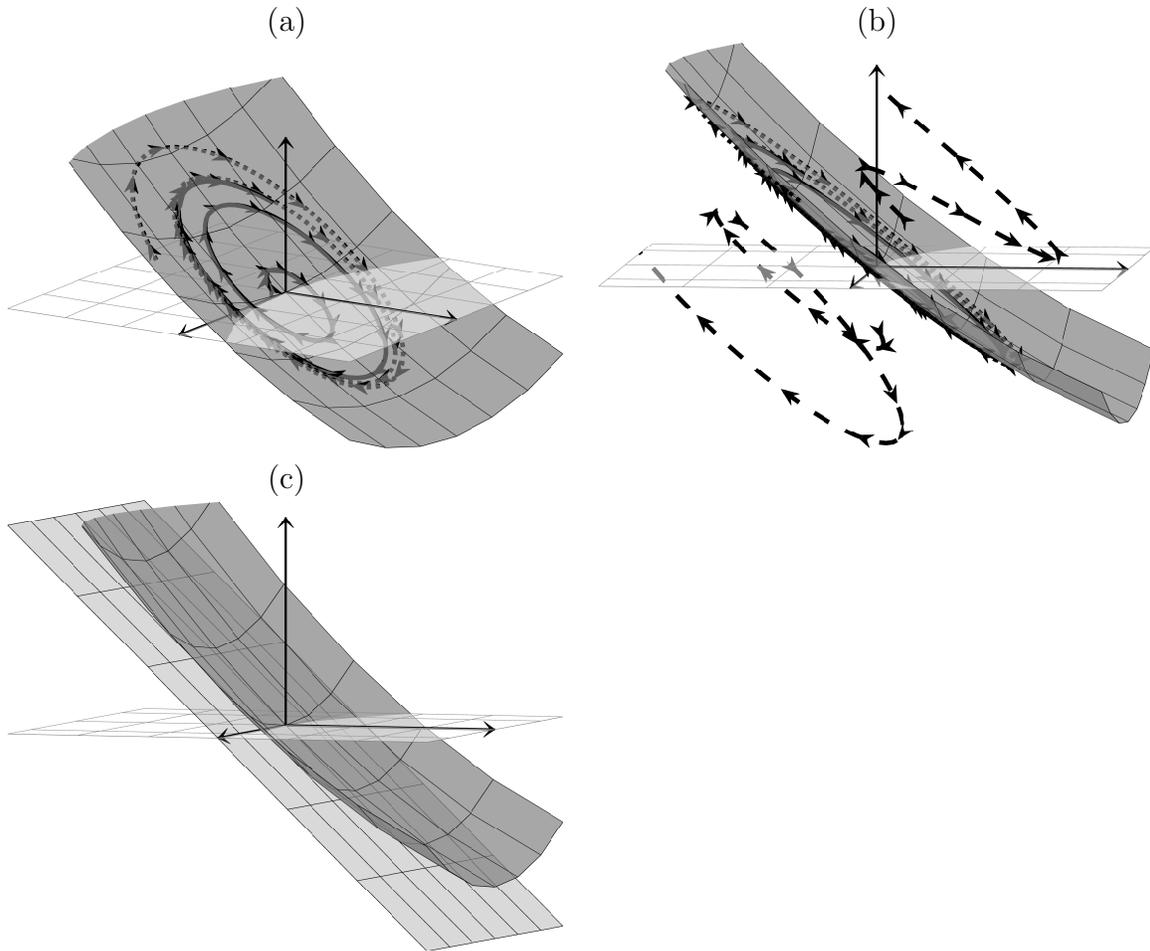
but it may contain additional trajectories as well. As noted in Proposition 1, if the NEM has dimension n_y , then there is a unique solution. If this is the case, but the steady state does not attract all nearby trajectories on the NEM (e.g., some trajectories converge to a limit cycle instead), this would imply that the stable manifold must be of dimension strictly less than n_y , in which case there could not be a solution to the linearized model, and standard perturbation methods would fail.

Figure 1 presents an example of such a situation for a case with two pre-determined variables ($n_y = 2$) and one jump variable ($n_z = 1$). Each panel of the Figure shows the same phase diagram from a slightly different angle. The dark gray two-dimensional surface shown in each panel is the NEM. Trajectories beginning on this NEM, such as the ones illustrated with the black solid and dotted lines in panels (a) and (b), neither converge to the steady state nor become unbounded, instead converging to a limit cycle (which also lies on the NEM). On the other hand, trajectories not beginning on the NEM, such as those illustrated by the dashed black lines in panel (b) of the Figure, diverge and become unbounded. Thus, in obtaining a solution, the TVC would force the system to jump (by choice of the only jump variable) onto the NEM. The solution to this non-linear system is thus simply the function that, for given values of the two pre-determined variables, returns the value of the jump variable that places the system on the NEM. Note however that, since trajectories on the NEM do not converge to the steady state, the stable manifold evidently has dimension zero. As a result, no solution will exist to the linearized system, and thus, even though a solution to the non-linear system exists, standard perturbation methods would be unable to discover it.

The method proposed in this paper is, conceptually, a simple modification of standard perturbation algorithms that is designed to address the possibility of such a scenario. In particular, rather than solving for the stable manifold and assuming it and the NEM are one and the same, we look for the NEM directly. In practice, the method is based on two observations. The first, which is stated formally below in Proposition 3, is that tangent to any h -invariant manifold²⁷ at the steady state is a unique A -invariant linear subspace of equal

²⁷In what follows, we only consider h -invariant manifolds that are locally analytic. For the sake of brevity, we will leave the “locally analytic” part as implied unless confusion will arise.

Figure 1: A Saddle Limit Cycle



Notes: The dark gray surface in each panel is the NEM. In panel (a), the black solid and dotted lines are two paths that converge to a limit cycle (which is located on the NEM), one from the inside and one from the outside. In panel (b), which shows the same phase space (from a slightly different angle), the dashed black lines are two paths for which the jump variable has not placed the system onto the NEM, and which therefore violate the TVC. In panel (c), the light gray plane, w , is tangent to the NEM at zero.

dimension, and vice versa. Further, given an A -invariant linear subspace, it is straightforward to compute the tangent h -invariant manifold of equal dimension to any desired order of approximation. The second observation is that, in most cases, there are a finite number of A -invariant linear subspaces of dimension at least n_y , and there are established numerical methods to find these subspaces. Thus, the solution method involves checking, for each A -invariant linear subspace of dimension at least n_y , whether the tangent h -invariant manifold of the same dimension is “bounded”, in the sense that trajectories on it satisfy the TVC. If there exists a unique solution to the model, then (a) there will be only one such bounded invariant manifold (i.e., the NEM), and (b) its dimension will be exactly n_y . On the other hand, if there is a bounded invariant manifold with dimension strictly greater than n_y , then there will be a multiplicity of solutions.²⁸ Finally, if there are no bounded invariant manifolds with dimension at least n_y , then there will be no solution.

To illustrate this method, consider again the example of Figure 1. We would like to solve for the NEM, which, as noted above, is given by the dark gray two-dimensional surface in each panel. In this particular example, it can be verified that there are exactly two A -invariant linear subspaces of dimension at least $n_y = 2$: the two-dimensional plane that is tangent to the NEM at zero, which we may denote w , and the full phase space \mathbb{R}^3 .²⁹ The three-dimensional h -invariant manifold tangent to the latter is (trivially) also \mathbb{R}^3 , and, as shown in panel (b), this manifold contains explosive trajectories and therefore clearly cannot be the NEM. The subspace w , meanwhile, which is shown as the light gray plane in panel (c) of the Figure, is precisely the linear subspace that is tangent to the h -invariant NEM. Thus, with w in hand, one can solve for the NEM to any desired order of approximation (and verify that trajectories on it do indeed remain bounded).

²⁸There will also be a multiplicity of solutions if there are multiple bounded invariant manifolds with dimension at least n_y . Under our assumption that m_0 is a manifold, however, this will necessarily imply the existence of a bounded invariant manifold with dimension strictly greater than n_y .

²⁹In this example, A has one real eigenvalue/eigenvector and a pair of complex eigenvalues/eigenvectors. w is given by the real elements of the complex space spanned by the two complex eigenvectors. Further, no other two-dimensional subspace is A -invariant, leaving \mathbb{R}^3 as the only other A -invariant subspace with dimension at least 2.

2.3 The Algorithm

This section formally presents the main elements of the proposed solution method. We focus here on establishing the theoretical underpinnings of the method, for now leaving aside discussion of how to implement the method in practice. Such practical issues are discussed below in Section 3.

Let M be the set of all h -invariant manifolds of non-zero dimension, and note that $m_0 \in M$. We then have the following proposition, which emphasizes that there is a duality between invariant manifolds and invariant functions (as defined earlier). As we shall see in Section 3 below, this will be quite useful in implementing our method.

Proposition 2. *If $m \in M$ is q -dimensional, then there exists an invariant function $\psi: \mathbb{R}^q \rightarrow \mathbb{R}^{n-q}$ such that (with appropriate re-orderings of the elements of x) $m = \mu_\psi \equiv \{(\tilde{y}, \psi(\tilde{y})) : \tilde{y} \in \mathbb{R}^q\}$. Further, for any such invariant function ψ , $\mu_\psi \in M$.*

Next, let W be the set of all real A -invariant linear subspaces of any non-zero dimension. We have the following proposition relating the non-linear system (2) to the linearized system (3).

Proposition 3.

- (i) *Suppose $m \in M$ is q -dimensional, and let $w \subset \mathbb{R}^n$ be the (unique) q -dimensional linear subspace that is tangent to m at zero. Then $w \in W$.*
- (ii) *Suppose $w \in W$ is q -dimensional. Then there exists a unique q -dimensional $m \in M$ that is tangent to w at zero.*

Proposition 3 establishes a tight relationship between the non-linear h -invariant manifolds contained in M and the linear A -invariant manifolds contained in W .³⁰ To see how this is useful, recall that we can generally decompose an invariant subspace $w \in W$ by locating a set of disjoint (except for the zero vector) lower-dimensional elements of W , and then writing w as the direct sum³¹ of these component subspaces. It can be verified further that a “maximal” such decomposition of any $w \in W$ exists; that is, there is a decomposition of

³⁰Given an $m \in M$, it is also straightforward in practice to compute the corresponding $w \in W$, and vice versa. We discuss this further below in Section 3

³¹The direct sum of two sets X and Y is given by $X \oplus Y \equiv \{x + y : x \in X, y \in Y\}$.

w in this way into elements of W that cannot themselves be so decomposed.³² In this sense, letting W^* be the set of elements of W that cannot be so decomposed, we can view the elements of W^* as “building blocks” for the elements of W .

As an example, suppose A has n distinct real eigenvalues. Then, as is well known, each of the n one-dimensional eigenspaces of A are A -invariant (i.e., elements of W), each q -dimensional $w \in W$ can be decomposed as the direct sum of q of the eigenspaces, and those n eigenspaces, being one-dimensional, cannot themselves be decomposed. Thus, in this example W^* is given by the set of n one-dimensional eigenspaces of A .

Given the preceding discussion, Propositions 1 through 3 combined suggest the following solution procedure. First, since $m_0 \in M$, by Proposition 3(i), the tangent linear subspace of the same dimension, which we denote w_0 , is an element of W . By Proposition 1, if a solution is to exist, then the dimension of m_0 —and therefore of w_0 as well—must be at least n_y . Thus, letting \widetilde{W} be the subset of elements of W of dimension at least n_y , we must have $w_0 \in \widetilde{W}$. Further, the elements of \widetilde{W} can be generated by selecting all combinations of the elements of W^* whose dimensions sum to at least n_y and taking the corresponding direct sum.³³ Each element of \widetilde{W} is then associated (by Proposition 3(ii)) to a unique element of M that is tangent and of the same dimension. We denote the set of such elements of M by \widetilde{M} .

The solution method is thus as follows: (1) generate elements of \widetilde{W} in the above fashion; (2) for each element $w \in \widetilde{W}$, get the corresponding element $m \in \widetilde{M}$; (3) for this m , compute the corresponding invariant function ψ using Proposition 2; (4) using ψ , simulate (2) numerically to check whether trajectories on m remain bounded. By definition, if a solution exists then m_0 will be the m identified in this way (i.e., among the elements of \widetilde{M} on which trajectories remain bounded) that has the largest dimension,³⁴ and w_0 the $w \in \widetilde{W}$ from which it was obtained. By Proposition 1, if the dimension of m_0 equals n_y , then there is a unique solution, while if it is greater than n_y then there is indeterminacy. Finally, if no such element of m could be identified, then this implies that the dimension of m_0 is less than n_y ,

³²Note that this decomposition may not be unique.

³³Technically, it is possible for the dimension of that direct sum to be less than n_y if A has any eigenvalues with a geometric multiplicity smaller than its algebraic multiplicity, in which case we would simply discard that particular combination.

³⁴One can verify in particular that, given our assumptions, all such identified m 's will be contained in m_0 .

and therefore there is no solution.

2.3.1 Relationship to Standard Perturbation Methods

We discuss in detail how to implement this solution method below in Section 3, but first, it is worth re-emphasizing that this is a generalization of the typical perturbation fundamentally methods based on the Blanchard and Kahn's [1980] algorithm. In particular, typical methods implicitly make the following additional assumption.

Assumption (A2). All bounded trajectories converge linearly to the steady state.

Under Assumption (A2) (which the method proposed in this paper does not make), one need only consider the linear part of the system in order to find m_0 . In particular, letting Λ denote the set of eigenvalues in A , for an eigenvalue $\lambda \in \Lambda$, let $e(\lambda)$ denote the corresponding generalized eigenspace, and let

$$r(\lambda) \equiv \{x \in [e(\lambda) \oplus e(\bar{\lambda})] : \text{Im}(x) = 0\}, \quad (4)$$

where $\bar{\lambda}$ denotes the complex conjugate of λ , $\text{Im}(x)$ denotes the imaginary part of x , and \oplus denotes the direct sum. We henceforth refer to $r(\lambda)$ as the real generalized eigenspace (RGE) associated with λ . The RGEs will play a key role in the implementation algorithm discussed below in Section 3. Note that $r(\lambda) \in W$, $r(\lambda) = r(\bar{\lambda})$, and if λ is real then the dimension of $r(\lambda)$ is the same as $e(\lambda)$, while if it is complex the dimension is twice that of $e(\lambda)$. Under Assumption (A2), the subspace w_0 is then given by the direct sum of the RGE's associated with the stable eigenvalues of A . Thus, there is no need to consider any non-linear effects in the construction of w_0 in this case. Once one has found w_0 in this way, it is then straightforward to obtain m_0 using Proposition 3(ii). Further, the dimension of w_0 (and therefore m_0) is given simply by the number of stable eigenvalues of A , and thus conditions governing the existence and uniqueness of solutions can be determined in a straightforward way. Thus, Assumption (A2) makes the solution method simpler, but at the cost of limiting its generality. In particular, it will not be possible to find solutions that feature limit cycles or chaos, since such solutions necessarily involve trajectories that are bounded but do not converge to the steady state. In contrast, the method proposed here does not require the

extra imposition of Assumption (A2), and as a result will generally be able to find these solutions.

2.4 Some Examples

Example 1: An Example Without Limit Cycles

Figure 2 illustrates the procedure discussed in Section 2.3 graphically for a simple example that does not feature limit cycles or chaos.³⁵ Each panel of the Figure shows a phase diagram for a two-dimensional system $x_{t+1} = h(x_t)$ of the type often encountered in macroeconomics, with $n_y = n_z = 1$. In all cases, A is assumed to have two distinct real positive eigenvalues. Thus, as noted above, W^* is the set made up of the two eigenspaces w_1^* and w_2^* , drawn as the gray lines in the Figure. Tangent to these eigenspaces are two elements of M , drawn as the black curves and labeled m_1^* and m_2^* . Further, it is trivial to obtain that $\widetilde{W} = \{w_1^*, w_2^*, \mathbb{R}^2\}$, with associated tangent m 's given by $\widetilde{M} = \{m_1^*, m_2^*, \mathbb{R}^2\}$.

While w_1^* , w_2^* , m_1^* , and m_2^* are the same in all panels of the Figure, the dynamics nonetheless differ, as indicated by the arrows of motion. For the case of panel (a), we see that trajectories beginning on m_1^* remain bounded, while those beginning anywhere else become unbounded. Thus, $m_0 = m_1^*$ and $w_0 = w_1^*$. Since m_0 has dimension $1 = n_y$, there thus exists a unique solution, which is given by the function that projects the system onto m_1^* by choice of z .

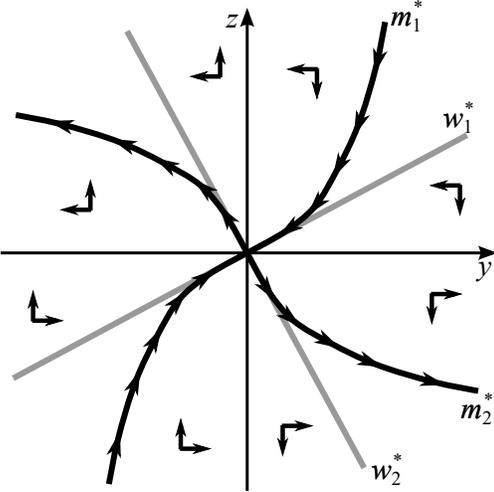
Panel (b) is similar to panel (a), except the arrows of motion in the northwest and southeast quadrants are reversed. In this case, trajectories on every element of \widetilde{M} remain bounded, so that m_0 is given by the element of \widetilde{M} with the largest dimension, i.e., $m_0 = \mathbb{R}^2$. Since m_0 has dimension $2 > n_y$ in this case, we have indeterminacy.

Finally, panel (c) is again similar to panel (a), except the arrows of motion in the northeast and southwest quadrants are reversed. In this case, trajectories on every element of \widetilde{M} become unbounded, which implies that the dimension of m_0 is less than n_y (i.e., it is zero), and thus there is no solution.

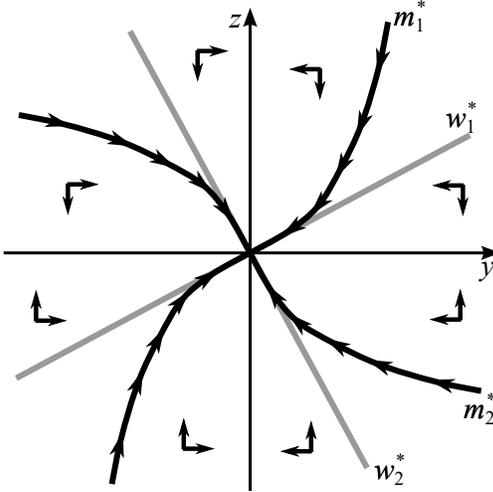
³⁵Note that, as such, standard perturbation methods could be used in this case. We nonetheless apply our proposed method here in order to illustrate it in an environment that most macroeconomists are quite familiar with. We present a second example below for which standard methods would *not* work.

Figure 2: An Example

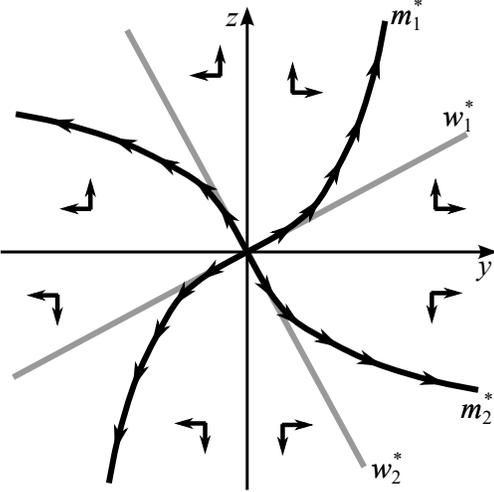
(a) Unique Solution: $m_0 = m_1^*$



(b) Multiple Solutions: $m_0 = \mathbb{R}^2$



(c) No Solution: $m_0 = \emptyset$



Example 2: An Example Featuring Limit Cycles

We now present a simple model for which a rational-expectations solution exists and can be found by the proposed method but not by standard methods. Consider the simple forward-looking model³⁶

$$\begin{aligned} K_{t+1} &= (1 - \delta) K_t + I_t \\ G(I_t) &= I_{t+1} - \alpha K_t, \end{aligned}$$

where K_t (“capital”) is pre-determined at date t , I_t (“investment”) is free to jump, $\delta \in (0, 1)$ and $\alpha > 0$ are parameters, and G is an invertible function satisfying $G(0) = 0$, $G'(I) > 0$, and $G'''(I) > 0$. To rule out the possibility of multiple steady states we assume further that $\alpha > \delta$. To keep things as simple as possible, we also assume that $G''(0) = 0$.

Letting $x_t \equiv (K_t, I_t)'$, we can write the evolution of this economy as

$$x_{t+1} = \begin{pmatrix} (1 - \delta) K_t + I_t \\ \alpha K_t + G(I_t) \end{pmatrix} \equiv h(K_t, I_t). \quad (5)$$

We then have that³⁷

$$A = \begin{pmatrix} 1 - \delta & 1 \\ \alpha & G' \end{pmatrix},$$

which has eigenvalues

$$\begin{aligned} \lambda_1 &= \frac{1 - \delta + G' - \sqrt{(1 - \delta - G')^2 + 4\alpha}}{2}, \\ \lambda_2 &= \frac{1 - \delta + G' + \sqrt{(1 - \delta - G')^2 + 4\alpha}}{2}. \end{aligned}$$

Suppose $G' < \alpha + \delta - 2$. It can then be verified that $\lambda_1 < -1$ and $\lambda_2 > 1$. Thus, if one were to apply standard perturbations methods, one would note that the stable manifold is empty, and thereby conclude that the NEM is as well, i.e., that there are no solutions to the model.³⁸ As we show now, this conclusion is erroneous here: the NEM is non-empty, and in particular there exists (at least to a third-order approximation) a unique solution to the non-linear model.

³⁶This model is a simplified version of the one presented in detail in Beaudry, Galizia, and Portier [2018].

³⁷For compactness, the notation G' means $G'(0)$, G'' means $G''(0)$, etc.

³⁸In estimation applications, this parameterization would then typically be discarded.

Since the system is two-dimensional with two distinct real eigenvalues, as in Example 1 above the set W^* has two elements, which are given by the eigenspaces w_j^* associated with each λ_j , $j = 1, 2$, so that $\widetilde{W} = \{w_1^*, w_2^*, \mathbb{R}^2\}$ and $\widetilde{M} = \{m_1^*, m_2^*, \mathbb{R}^2\}$, where m_j^* is the one-dimensional element of M that is tangent to w_j^* . As noted in Proposition 1, if a unique solution exists then it must be invariant and analytic. We check to see if such a solution exists. In particular, suppose ψ is an invariant function and $I_t = \psi(K_t)$. From (5), we then have $K_{t+1} = \pi(K_t) \equiv (1 - \delta)K_t + \psi(K_t)$, and thus in order to be invariant ψ must satisfy

$$\psi((1 - \delta)K + \psi(K)) = \alpha K + G(\psi^j(K)). \quad (6)$$

Since condition (6) must hold for all K , sequentially taking derivatives three times with respect to K and evaluating at $K = 0$, we obtain the conditions

$$\psi'^2 + (1 - \delta - G')\psi' - \alpha = 0, \quad (7)$$

$$\left[(1 - \delta + \psi')^2 + \psi' - G'\right]\psi'' = 0, \quad (8)$$

$$\left[(1 - \delta + \psi')^3 + \psi' - G'\right]\psi''' = G'''\psi'^3 - 3(1 - \delta + \psi')\psi''^2. \quad (9)$$

Equation (7) is a quadratic in ψ' whose solutions it can be verified are given by $\psi'_1 \equiv \lambda_1 - (1 - \delta)$ and $\psi'_2 \equiv \lambda_2 - (1 - \delta)$. Note that these two possible values for ψ' index two different invariant functions of our system, which we denote ψ_1 and ψ_2 , respectively, where the corresponding second- and third-order Taylor coefficients can be obtained uniquely from (8) and (9).³⁹ Note also from Proposition 2 that the images of \mathbb{R} through these two invariant functions are precisely m_1^* and m_2^* . It can be further verified that $w_j^* = \{(K, \psi'_j K) : K \in \mathbb{R}\}$, which confirms that m_j^* is in fact tangent to w_j^* at zero. We have thus effectively found a method to solve for the two non-trivial invariant manifolds that are of dimension at least $n_y = 1$. It remains to check whether either (or both) of the associated invariant functions are in fact solutions.

Substituting $\psi'_j = \lambda_j - (1 - \delta)$ into equation (8), we obtain

$$\left[\lambda_j^2 + \lambda_j - (1 - \delta + G')\right]\psi''_j = 0.$$

³⁹In principle, one can continue taking derivatives of (6) to any desired order to obtain the corresponding higher-order Taylor coefficients.

Except in knife-edge circumstances, the term in square brackets will be non-zero, and thus we must have $\psi_j'' = 0$.⁴⁰ Next, substituting $\psi_j'' = 0$ and the expression for ψ_j' into (9) yields

$$\psi_j''' = \frac{[\lambda_j - (1 - \delta)]^3}{\lambda_j^3 - \lambda_{3-j}} G''''.$$

Since $\lambda_1 < -1$ and $\lambda_2 > 1$, the numerator and denominator of the fraction on the right-hand side of this expression are of the same sign (negative when $j = 1$, positive when $j = 2$), and thus, since $G'''' > 0$, we also have $\psi_j''' > 0$. Given the properties just established, it can then be verified that, to a third-order approximation, the evolution of K_t on m_j^* is given by

$$K_{t+1} = \pi_j(K_t) \equiv \lambda_j K_t + \frac{1}{6} \psi_j''' K_t^3. \quad (10)$$

Since $\lambda_2 > 1$ and $\psi_2''' > 0$, the system given by (10) for $j = 2$ is globally unstable: all paths for K beginning on m_2^* (except those beginning exactly at the steady state) will explode. Thus, ψ_2 cannot be a solution. On the other hand, even though $\lambda_1 < -1$, so that the system given by (10) for $j = 1$ will not converge to the steady state, since $\psi_1''' > 0$ all paths beginning on m_1^* will remain bounded (as long as $|K_0|$ is not too large). Thus, there is in fact a unique solution to (5) given by $\phi = \psi_1$,⁴¹ and the resulting dynamics of the system are governed (to a third-order approximation) by (10). These dynamics will generically feature limit cycles or chaos, so that the system will typically neither explode nor converge to a single point.

The above analysis is illustrated graphically in Figures 3 and 4. Figure 3 shows the eigenspaces w_1^* and w_2^* (i.e., the elements of W^*), and the associated one-dimensional invariant subspaces m_1^* and m_2^* , the configuration of which is similar to Figure 2.⁴² A unique solution to (5), if it exists, is the projection of the system onto one of m_1^* or m_2^* by choice of I . To determine which, if either, of these invariant manifolds represents a solution, one needs to consider the dynamics of the projected systems, which are given fundamentally by the univariate systems $K_{t+1} = \pi_j(K_t)$, $j = 1, 2$. Panels (a) and (b) of Figure 4 illustrate the evolution of K on m_1^* and m_2^* , respectively. In both cases, the solid black line plots the

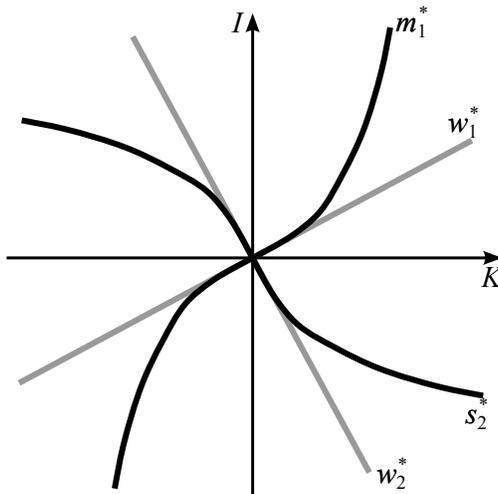
⁴⁰This property is inherited directly from our earlier simplifying assumption that $G'' = 0$.

⁴¹Note that we technically need to check that there are trajectories on $\mathbb{R}^2 \in \tilde{M}$ that become unbounded in order to rule out the possibility that $m_0 = \mathbb{R}^2$. However, since $m_2^* \subset \mathbb{R}^2$ and trajectories on m_2^* become unbounded as argued above, that result follows immediately.

⁴²Unlike in Figure 2, Figure 3 does not contain arrows of motion since the system exhibits oscillations, which are difficult to illustrate using such arrows.

function π_j , while the gray lines illustrate a typical path for K . In panel (b), we clearly see that such paths explode on m_2^* : $|K|$ grows each period without bound. As shown in panel (a), however, the behavior of K on m_1^* is more subtle. Beginning from a point near the steady state, we again observe that $|K|$ grows each period, in this case with K alternating signs. However, as the system moves away from the steady state, the growth in $|K|$ eventually tapers off, and the system settles into a 2-cycle given by $K_t = (-1)^t \bar{K}$ for some $\bar{K} > 0$. Thus, the system explodes on m_2^* but not on m_1^* , and thus $m_0 = m_1^*$. Since $n_0 = n_y = 1$, we have verified that a solution exists and is unique, and indeed is given by $\phi = \psi_1$.

Figure 3: Example Featuring Limit Cycles



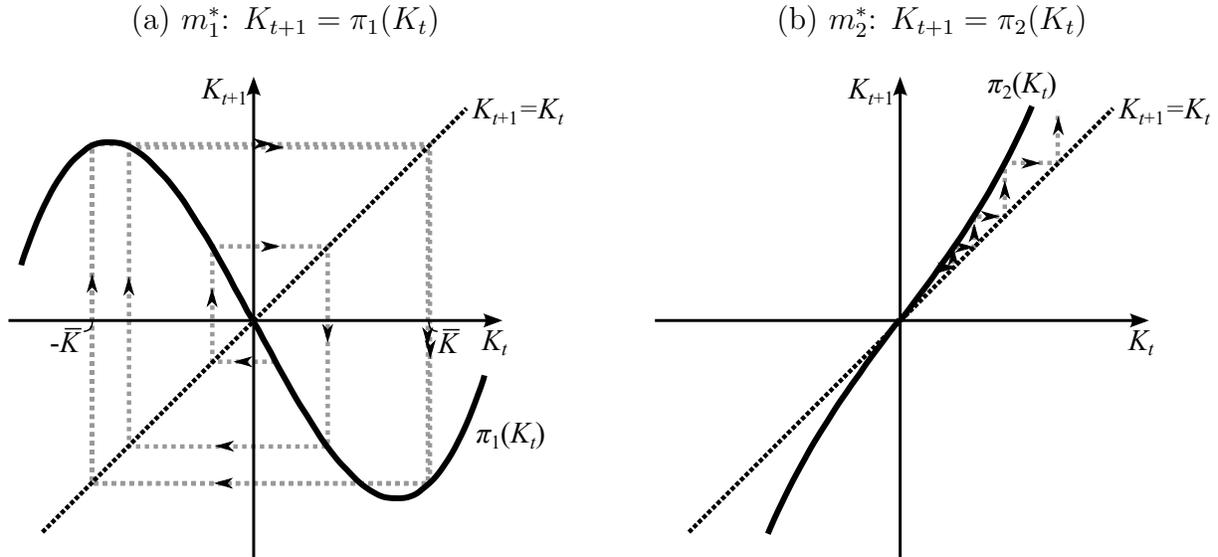
3 Implementation

3.1 Finding m_0

In this subsection we discuss an algorithm for finding the NEM m_0 whenever a solution exists.⁴³ We proceed in this subsection under the assumption that if a solution exists then it is unique; that is, we assume that $n_0 \leq n_y$. In Section 3.3 below, we discuss issues related to indeterminacy. The algorithm we propose is based on real Schur decompositions of the

⁴³Obviously, if a solution does not exist then there is no need to compute the NEM.

Figure 4: Evolution of K on m_1^* and m_2^*



matrix A .⁴⁴ A real Schur decomposition of A is given by $A = UTU'$,⁴⁵ where U is a real orthogonal matrix (and in particular, $UU' = I_n$), and T is a real block-upper-triangular matrix with the following properties: (a) the diagonal blocks are all either 1×1 or 2×2 ; (b) the real eigenvalues of A appear in the 1×1 blocks; and (c) the eigenvalues of the 2×2 blocks correspond to the pairs of complex conjugate eigenvalues of A .⁴⁶ Note that the real Schur decomposition is not unique.

Given such a Schur decomposition of A , consider the first $q > 0$ columns of U , denoted U_1 , and the upper-left $q \times q$ block of T , denoted T_{11} (where we require only that this partition does not split one of the 2×2 diagonal blocks of T described in property (c) above). As is well known, the columns of U_1 form a basis for a q -dimensional invariant subspace of A ; that

⁴⁴Under assumption (A1), the function h —and therefore the matrix A —exists, and thus real Schur decompositions will be sufficient for the task at hand. One can nonetheless relax the invertibility part of (A1) (e.g., one can allow “static” equations in (1)), and modify the algorithm in this section to use the generalized real Schur decomposition (see, for example, the discussion in Klein [2000]).

⁴⁵See, for example, Golub and Van Loan [1996], chapter 7. Algorithms to compute a real Schur decomposition are widely available. In MATLAB and Julia it can be done using the function `schur` which, for a real matrix, computes the real Schur decomposition by default. For implementation in Fortran or C++, various routines are available, including the LAPACK routine `dgees`.

⁴⁶Note that the real Schur decomposition generally differs from a regular Schur decomposition in that, in the latter, T is upper-triangular (rather than block-upper-triangular) and generally complex, with the eigenvalues of A on the main diagonal.

is, the column space w_1 of U_1 is an element of W . Further, the convergence properties of the linearized system on w_1 are governed by the eigenvalues of T_{11} . Finally, for an element $w \in W$ with dimension q , we can always find a Schur decomposition of A such that w is equal to the column space of U_1 . Thus, in principle we can generate any element of W through an appropriate choice (and partition) of the Schur decomposition.

A given Schur decomposition can also be “reordered” so as to obtain an alternative Schur decomposition of A in which the first q columns of U are a basis for a different element of W .⁴⁷ It would be useful if, by considering every possible reordering of a given Schur decomposition, we could so generate every possible element of W . Unfortunately, this is not generally the case. In particular, when A has at least one eigenvalue with a geometric multiplicity greater than one, W will have an uncountable number of elements.⁴⁸ Since there are only a finite number of possible reorderings of the Schur decomposition, it is clearly not possible to find every element of W in this way, and indeed it will not be feasible in practice to find every one of the infinite number of elements of W using *any* algorithm. To address this problem, we make the following additional assumption about w_0 .

Assumption (A3). If w_0 and $r(\lambda)$ have any non-zero element in common, then $r(\lambda) \subset w_0$.⁴⁹

Assumption (A3) states that, aside from the zero vector (which is always an element of both w_0 and $r(\lambda)$), w_0 cannot contain only part of a given RGE: either the whole RGE is contained in w_0 , or none of it is.⁵⁰ Let \mathcal{R} be the set of RGE’s of A . Under Assumption (A3), w_0 can be constructed as the direct sum of a subset of the elements of \mathcal{R} . Since there are at most n RGE’s, computing direct sums of each possible combination of the elements of \mathcal{R} is feasible in practice (though in general, as we argue in Section 3.2, it will not be necessary to do so).

⁴⁷Re-ordering of the real Schur decomposition can be done using the function `ordschur` in MATLAB and Julia, or the LAPACK routine `dtrsen` in Fortran or C++.

⁴⁸For example, for two linearly independent eigenvectors associated with the same real eigenvalue, every linear combination of these eigenvectors is also an eigenvector, and the space spanned by that new eigenvector is a one-dimensional A -invariant subspace. We can clearly generate an uncountably infinite number of invariant subspaces in this way.

⁴⁹Recall that $r(\lambda)$ is the RGE of A associated with the eigenvalue λ ; see equation (4).

⁵⁰As with several of our other assumptions, this assumption is also implicitly made by standard perturbation algorithms. More accurately, Assumption (A2), which is made by standard algorithms but which we do not make here, implies Assumption (A3). Thus, standard algorithms have no need to separately impose Assumption (A3), though they nonetheless rely on its content.

This suggests a simple way to find a solution in practice if it exists. First, obtain a Schur decomposition $A = UTU'$. Next, let $\{r(\lambda_1), \dots, r(\lambda_k)\}$ be a subset of k distinct elements of \mathcal{R} whose dimensions sum to n_y .⁵¹ Re-order the Schur decomposition so that the upper-left $n_y \times n_y$ block of T has eigenvalues $\lambda_1, \dots, \lambda_k$ with the same algebraic multiplicities as they appear in A , as well as the corresponding conjugates of any complex λ_j 's. The first n_y columns of U are then a basis for $w \equiv r(\lambda_1) \oplus \dots \oplus r(\lambda_k)$. Thus, $w \in \widetilde{W}$ (and is thus a candidate value of w_0), and is associated with an element $m \in \widetilde{M}$ by Proposition 3 and an invariant function ψ by Proposition 2. We may obtain a first-order approximation to ψ as follows. Partition U conformably with $x = (y, z)$ as

$$U = \begin{pmatrix} U_{yy} & U_{yz} \\ U_{zy} & U_{zz} \end{pmatrix}.$$

In order for x to be in the space spanned by the first n_y columns of U , the final n_z rows of $U^{-1}x (= U'x)$ must equal zero. That is, we must have⁵²

$$z_t = -U'_{zz}{}^{-1}U'_{yz}y_t. \quad (11)$$

Thus, (11) gives the A -invariant function that projects the system onto w by choice of z_t , and is therefore also the *first-order approximation* to the h -invariant function ψ that projects the system onto m by choice of z_t ; that is,

$$D_y^\psi = -U'_{zz}{}^{-1}U'_{yz}.$$

This can then be used in a straightforward manner to sequentially obtain any desired number of higher-order derivatives of ψ . One can then check numerically whether the sequence generated by $y_{t+1} = \pi(y_t) \equiv f(y_t, \psi(y_t))$ satisfies the TVC. If it does, then we have found a solution (which is unique by assumption), and we are done. If not, then one can choose an alternative subset of \mathcal{R} whose dimensions sum to n_y and repeat the above process, continuing in this way until a solution is found. If after checking every such subset of \mathcal{R} no solution has been found, then we may conclude that there is no solution.

⁵¹Recall that $r(\lambda) = r(\bar{\lambda})$, so that $r(\lambda)$ and $r(\bar{\lambda})$ are not distinct elements of \mathcal{R} .

⁵²The form of (11) will be familiar to most macroeconomists, being precisely the form that standard solutions to linear rational expectations models take.

3.2 Narrowing Down The Search

The algorithm discussed in Section 3.1 involves checking as many as n -choose- n_y combinations of the elements of \mathcal{R} to see whether the system generated by $y_{t+1} = \pi(y_t)$ satisfies the TVC. In practice, we can use theory to rule out some of these combinations, as established in the following proposition.

Proposition 4. *If λ is an eigenvalue of A with $|\lambda| < 1$, then $r(\lambda) \subset w_0$. If λ is a real eigenvalue of A with $\lambda > 1$, then $r(\lambda) \not\subset w_0$.*

The intuition for Proposition 4 is straightforward. To see the first part, suppose $|\lambda| < 1$ and let m be the h -invariant manifold tangent to and of the same dimension as $r(\lambda)$. For $x_0 \in m$ sufficiently close to zero, the dynamics of the system are dominated by the linear part of h , which are in turn governed by the eigenvalue λ . Since $|\lambda| < 1$, the resulting trajectory converges to the steady state, and thus $m \subset m_0$, so that $r(\lambda) \subset w_0$.

To see the second part of Proposition 4, let v be a real eigenvector of A associated with $\lambda > 1$, $w \equiv \{\alpha v : \alpha \in \mathbb{R}\}$ the one-dimensional space spanned by this eigenvector (which is A -invariant), and m the one-dimensional h -invariant manifold tangent to w at zero. For any smooth invertible function $\xi: m \rightarrow \mathbb{R}$ with $\xi(0) = 0$, the sequence $\alpha_t \equiv \xi(x_t)$ evolves according to

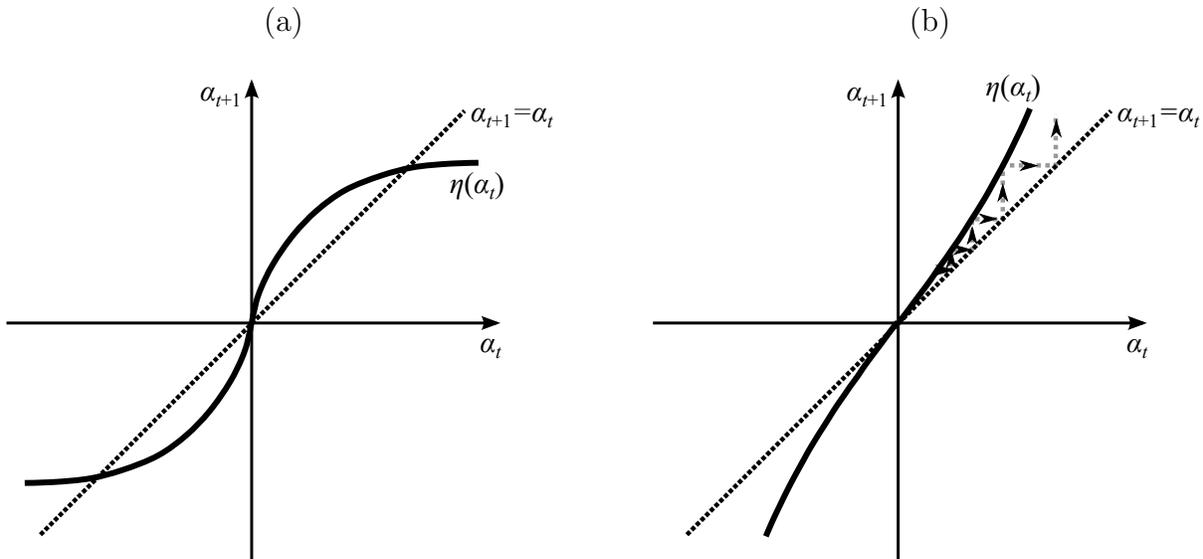
$$\alpha_{t+1} = \eta(\alpha_t), \tag{12}$$

for a function η that satisfies $\eta(0) = 0$ and $\eta'(0) = \lambda$.⁵³ System (12) thus has a steady state at $\alpha = 0$, and the dynamics near that steady state are given to a first-order approximation by $\alpha_{t+1} = \lambda\alpha_t$. Since $\lambda > 1$, there are two possible configurations for the function η , displayed separately in panels (a) and (b) of Figure 5. In the first configuration, illustrated in panel (a), η crosses the 45-degree line at at least one non-zero point (in addition to the crossing at zero). This additional crossing, however, would represent an additional steady state of (12), and therefore an additional steady state of (2), a possibility that is ruled out by Assumption (A1). Thus, this configuration for η cannot occur. The remaining possible configuration, wherein η crosses the 45-degree line at only one point (zero), is illustrated in panel (b). In this case, if $\alpha_0 \neq 0$ then $|\alpha_t|$ will grow without bound (as illustrated by the gray line

⁵³Specifically, the function η here is given by $\eta(\alpha) \equiv \xi(h(\xi^{-1}(\alpha)))$.

in the figure for the case where $\alpha_0 > 0$); that is, α_t becomes unbounded, and therefore so does x_t . Thus, if $r(\lambda) \subset w_0$, then there would be trajectories beginning on m_0 that become unbounded, which cannot happen by definition of m_0 .

Figure 5: Configurations For $\eta(\alpha_t)$ When $\lambda > 1$



Proposition 4 implies that we need only consider re-orderings of the Schur decomposition that result in every stable eigenvalue appearing in the upper-left $n_y \times n_y$ block of T , and no real eigenvalues that are greater than one appearing in that block. This can significantly reduce the number of re-orderings that one is required to check. For example, in a modestly sized system with $n_y = n_z = 5$, there are in principle as many as $10\text{-choose-}5 = 252$ possible re-orderings to check. If, however, 3 of the eigenvalues of this system were stable, and a further 3 were real and greater than one, then using Proposition 4 one may narrow down the set of re-orderings to check to at most $4\text{-choose-}2 = 6$. For larger systems, the potential computational savings are even greater.

3.3 Indeterminacy

In Section 3.1, we assumed that if a solution exists then it is unique (i.e., that $n_0 \leq n_y$). In such cases, the algorithm described in that section will be capable of finding the solution if

it exists, and will be unable to find any solution if no solution exists. In general, of course, we could have $n_0 > n_y$, i.e., multiple solutions/indeterminacy. In such cases, the algorithm presented in Section 3.1 could erroneously lead to one to conclude that an identified solution is the only solution, or even that there are no solutions at all. In this section we describe several techniques one can use to screen for indeterminacy when it may be of potential concern.

The first and simplest method to screen for indeterminacy is to check for indeterminacy of the linearized system. In particular, Proposition 4 implies that the stable manifold must be contained in m_0 . If the dimension of that stable manifold is greater than n_y (i.e., if there are at least n_y stable eigenvalues), then m_0 will necessarily also be of dimension greater than n_y . Fundamentally, this is the Blanchard-Kahn method for detecting indeterminacy, and in some cases this will be sufficient to identify indeterminacy if it exists. That is, one may be willing to assume that this “linear indeterminacy” is the only potential source of indeterminacy, and thus if there are no more than n_y stable eigenvalues then indeterminacy of the non-linear system is not a concern. This is the assumption implicitly made by standard perturbation algorithms.

If the system has no more than n_y stable eigenvalues, but one is unwilling to assume that linear indeterminacy is the only potential source of indeterminacy, one can instead modify the algorithm of Section 3.1 to explicitly implement the method discussed in Section 2.3. In particular, under Assumption (A3), one can view \mathcal{R} as an alternative set of linear building blocks (i.e., an alternative to W^*) with which to construct the elements of \widetilde{W} , and then one can check the $m \in \widetilde{M}$ corresponding to each such element to see whether trajectories on it remain bounded. In practice, this involves modifying the procedure from Section 3.1 by checking every subset $\{r(\lambda_1), \dots, r(\lambda_k)\}$ of elements of \mathcal{R} whose dimensions sum to *at least* (rather than exactly) n_y , re-ordering the Schur decomposition so that the upper-left block of T has the eigenvalues $\lambda_1, \dots, \lambda_k$, etc. While this procedure can determine conclusively whether there is indeterminacy, it does make the implementation somewhat more complicated, since the associated invariant functions ψ can no longer all typically be chosen such that $\tilde{y} = y$ and $\tilde{z} = z$ (so that $z = \psi(y)$). This necessitates taking additional sets

of analytic derivatives of the invariance relation,⁵⁴ and also writing computer code flexibly enough to accommodate different combinations of \tilde{y} and \tilde{z} . Having said that, while writing the code may be more complicated, aside from the one-time up-front cost of taking the additional analytic derivatives, actual computation times should not be made impractically long by choosing this method for addressing indeterminacy.

A third screening method, which is in the middle ground between the first and second, is to modify the algorithm discussed in Section 3.1 to check every possible combination of RGE's in \mathcal{R} whose dimensions sum to exactly n_y , instead of stopping at the first one that yields a solution. If more than one solution is found in this way, then one may conclude that there is indeterminacy. This method will detect more instances of indeterminacy than the first method, and is simpler to implement than the second method since it only requires obtaining derivatives of ψ for the case of $\tilde{y} = y$ and $\tilde{z} = z$. Further, if there are no unstable complex eigenvalues, and if the RGE's corresponding to the unstable real negative eigenvalues are all one-dimensional, this method will in fact be sufficient to determine conclusively whether there is indeterminacy. When this is not the case, however, certain forms of indeterminacy could be missed.

4 Extension to the Stochastic Case

4.1 Set-Up

Extending the above solution method to the stochastic case is straightforward, though the details depend on the precise way in which the stochastic terms enter. To keep the presentation simple, we assume that the evolution of the n_θ -vector of exogenous stochastic variables θ_t evolves according to the stationary vector autoregressive process

$$\theta_{t+1} = B\theta_t + \zeta\epsilon_{t+1}, \quad (13)$$

⁵⁴Given $\tilde{z} = \psi(\tilde{y})$, the invariance relation is given by $\psi(h^{\tilde{y}}(\tilde{y}, \psi(\tilde{y}))) = h^{\tilde{z}}(\tilde{y}, \psi(\tilde{y}))$, where $h^{\tilde{y}}$ denotes the first q elements of h and $h^{\tilde{z}}$ the remaining $n - q$ elements. Sequentially taking derivatives of this expression with respect to \tilde{y} and evaluating them at the steady state, one obtains a system in the unknown derivatives of ψ which can then be solved.

for some $n_\theta \times n_\theta$ matrix B whose eigenvalues are all strictly inside the unit circle.⁵⁵ Here, $\zeta \in \mathbb{R}$ is the “perturbation parameter,” and ϵ_t is a n_ϵ -vector of mean-zero i.i.d. innovations. In particular, we will seek a solution for the case where $\zeta = 1$. To do so, we take an approximation around the non-stochastic steady state (i.e., the steady state for the case where $\zeta = 0$). Without loss of generality, we henceforth assume that all variables are expressed in deviations from that non-stochastic steady state.

Given the above stochastic process, we assume that the n model equations governing the endogenous variables can be written

$$\begin{pmatrix} y_{t+1} - f(x_t, \theta_t) \\ \mathbb{E}_t [g(x_{t+1}, x_t, \theta_t, \zeta, \epsilon_{t+1})] \end{pmatrix} = 0, \quad (14)$$

for some $f: \mathbb{R}^{n+n_\theta} \rightarrow \mathbb{R}^{n_y}$ and $g: \mathbb{R}^{2n+n_\theta+n_\epsilon+1} \rightarrow \mathbb{R}^{n_z}$, where \mathbb{E}_t denotes expectation conditional on the information set at date t . The function g is assumed to satisfy the key restriction that, wherever ϵ_{t+1} and ζ appear, they do so as part of the term $B\theta_t + \zeta\epsilon_{t+1}$.⁵⁶

For convenience, define

$$\Gamma(x_{t+1}, x_t, \theta_t, \zeta, \epsilon_{t+1}) \equiv \begin{pmatrix} y_{t+1} - f(x_t, \theta_t) \\ g(x_{t+1}, x_t, \theta_t, \zeta, \epsilon_{t+1}) \end{pmatrix}. \quad (15)$$

Of interest to us will be the first-order approximation to (14) around the non-stochastic steady state, which can be written in the form⁵⁷

$$D_{x_{t+1}}^\Gamma \mathbb{E}_t [x_{t+1}] + D_{x_t}^\Gamma x_t + D_{\theta_t}^\Gamma \theta_t = 0. \quad (16)$$

In anticipation of the fact that the non-linear solution will have ζ as a state variable, we can augment this system with the trivial equation $\zeta = \zeta$ and the stochastic relationship $\mathbb{E}_t[\theta_{t+1}] = B\theta_t$, and then (assuming $D_{x_{t+1}}^\Gamma$ is invertible) write the full linear system in the form

$$\mathbb{E}_t [X_{t+1}] = AX_t, \quad (17)$$

where

$$A \equiv \begin{pmatrix} A_{xx} & A_{x\theta} & 0 \\ 0 & B & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

⁵⁵At the cost of complicating the presentation somewhat, this assumption can be easily relaxed to include non-linear processes.

⁵⁶That is, we assume the underlying model equations are functions of x_t , x_{t+1} , θ_t , and θ_{t+1} , so that ζ and ϵ_{t+1} only appear after using (13) to replace θ_{t+1} .

⁵⁷The fact that ζ and ϵ_{t+1} do not appear in (16) follows from our earlier restriction on g and the fact that $\mathbb{E}_t[\epsilon_{t+1}] = 0$.

$X_t \equiv (x_t, \theta_t, \zeta)$, $A_{xx} = -(D_{x_{t+1}}^\Gamma)^{-1} D_{x_t}^\Gamma$, and $A_{x\theta} = -(D_{x_{t+1}}^\Gamma)^{-1} D_{\theta_t}^\Gamma$. As we did for A in the deterministic case, we assume for ease of exposition that none of the eigenvalues of A_{xx} lie on the complex unit circle. Note that, since A is block-triangular and thus its eigenvalues are given by the eigenvalues of the diagonal blocks, and since the eigenvalues of B are strictly inside the unit circle by assumption, there is exactly one eigenvalue of A equal to one (i.e., the eigenvalue corresponding to the trivial equation $\zeta = \zeta$).

In the deterministic environment, invariant manifolds and invariant functions each had natural definitions, and we established a duality between the two concepts in Proposition 2. In the stochastic environment, we proceed slightly differently by defining invariant functions first, and then defining invariant manifolds so as to have a relationship analogous to the one in Proposition 2. The definition of an invariant function we employ here is more nuanced than in the deterministic case, since it must account for the fact that the relationships governing the jump variables in (14) (i.e., the last n_z equations) need only hold in expectation. In particular, let Ω_t denote the information set at time t (which includes the entire history of all variables up to and including date t), and let ψ be a function that maps ζ , θ , and a subset of $q > 0$ of the elements of x , which we denote \tilde{y} , into the remaining $n - q$ elements of x , which we denote \tilde{z} . We will say that ψ is invariant if there exists a function $\gamma(\Omega_t; \psi)$ such that, for any $(\tilde{y}_t, \theta_t, \zeta)$, the values $\tilde{z}_t = \psi(\tilde{y}_t, \theta_t, \zeta)$, $\tilde{z}_{t+1} = \psi(\tilde{y}_{t+1}, B\theta_t + \zeta\epsilon_{t+1}, \zeta)$, and $\tilde{y}_{t+1} = \gamma(\Omega_{t+1}; \psi)$ together satisfy (14). Note that, for variables in \tilde{y}_t that are pre-determined, the corresponding elements of $\mathbb{E}_t[\gamma(\Omega_{t+1}; \psi)]$ are uniquely pinned down by the appropriate elements of $f(y_t, \psi(y_t, \theta_t, \zeta), \theta_t)$, and therefore do not depend in any way on the realization of ϵ_{t+1} . For non-predetermined variables, however, this will in general not be the case.

Given an invariant function ψ , we may then define a corresponding $(n_\theta + q + 1)$ -dimensional manifold $\mu_\psi \equiv \{(x, \theta, \zeta) \in \mathbb{R}^{n+n_\theta+1} : \tilde{z} = \psi(\tilde{y}, \theta, \zeta)\}$. We refer to such a manifold as being invariant, and let M be the set of all such invariant manifolds. If an invariant function ψ is such that the stochastic sequence of endogenous variables generated by

$$\begin{aligned}\tilde{z}_t &= \psi(\tilde{y}_t, \theta_t, \zeta), \\ \tilde{y}_t &= \gamma(\Omega_t; \psi),\end{aligned}$$

satisfies the stochastic TVC, $\limsup_{t \rightarrow \infty} \mathbb{E}_0 \|x_t\| < \infty$, then we will refer to the associated invariant manifold μ_ψ as being “bounded”. We may then define a solution to (14) as an

invariant function ϕ that gives the value of z_t conditional on (ζ, θ_t, y_t) and is such that μ_ϕ is bounded. We may further define the NEM m_0 as the union of all such bounded manifolds.

Analogous to Proposition 3, there is a tight relationship between the manifolds in M and the real invariant subspaces of the matrix A in (17). In particular, let \widehat{W} be the set of all A -invariant linear subspaces of non-zero dimension, and let $w_{\text{ex}} \equiv \{(x, \theta, \zeta) \in \mathbb{R}^{n+n_\theta+1} : x = 0\}$ denote the subspace corresponding to the exogenous variables θ and ζ . Next, let

$$W \equiv \left\{ w \in \widehat{W} : w_{\text{ex}} \subset w \right\}. \quad (18)$$

That is, W is the set of real invariant subspaces of A that contain the “exogenous subspace” w_{ex} . Then, in a straightforward extension of Proposition 3, we can easily verify that (i) tangent to every $m \in M$ at zero is a unique element of W , and (ii) tangent to every $w \in W$ at zero is a unique element of M .

From this point, the remainder of the solution algorithm is essentially identical to the non-stochastic case. First, obtain the subset of elements W^* of W that cannot be written as the direct sum of two or more distinct elements of W . Note that, whereas in the non-stochastic case the elements of W^* could have dimension as small as one, here they must have dimension at least $n_\theta + 2$.⁵⁸ Second, obtain the set \widetilde{W} of elements of W of dimension at least $N \equiv n_y + n_\theta + 1$ by taking direct sums of the elements of W^* . Third, generate the associated tangent elements to make up \widetilde{M} and their associated invariant functions, and check each one for compliance with the TVC. The NEM m_0 is then the largest-dimensional $m \in \widetilde{M}$ on which the TVC is satisfied, and w_0 is the associated element of \widetilde{W} . If the dimension of this m equals N , then there is a unique solution, while if it is greater than N then there is indeterminacy. If no such m could be found, then no solution exists.

4.2 Implementation

With some minor modifications, the implementation method discussed in Section 3 for the non-stochastic case using the real Schur decomposition of A may be applied to the stochastic

⁵⁸This implicitly assumes that $w_{\text{ex}} \notin W$, so that the dimension of any $w \in W$ is strictly greater than $n_\theta + 1$. That is, it assumes that, to a first-order approximation, $\mathbb{E}_t[x_{t+1}]$ depends in some way on θ_t . Instances where this assumption fails would be rare in practice, but could nonetheless be easily accommodated with minor modifications.

case as well. The principal modification is that, whereas in the non-stochastic case we re-ordered and partitioned the Schur decomposition so that the first n_y columns of U formed a basis for an n_y -dimensional A -invariant subspace without any further restriction, in the stochastic case (a) the relevant A -invariant subspaces are N -dimensional, and (b) they must always contain the invariant subspace corresponding to the exogenous variables (i.e., w_{ex}). As in the deterministic case, for simplicity we proceed here assuming that if a solution exists, then it is unique. Nonetheless, it is straightforward to extend the discussion of Section 3.3 to the stochastic case.

In particular, let $z_t = \psi(y_t, \theta_t, \zeta)$ be a candidate solution. Since $\tilde{y}_t = y_t$ in this case, we have

$$\gamma(\Omega_{t+1}; \psi) = f(y_t, \psi(y_t, \theta_t, \zeta), \theta_t) \equiv \pi(y_t, \theta_t, \zeta).$$

so that $y_{t+1} = \pi(y_t, \theta_t, \zeta)$. Note that one may express the derivatives of π in a straightforward way as functions of the derivatives of ψ using the above expression. Next, substituting ψ and π into the last n_z equations in (14) for the endogenous variables, we may obtain the relationship

$$\mathbb{E} \left[g \left(\pi(y, \theta, \zeta), \psi(\pi(y, \theta, \zeta), \theta, \zeta), y, \psi(y, \theta, \zeta), \theta, \zeta, \epsilon) \right) \right] = 0, \quad (19)$$

where for notational simplicity we have dropped the time subscripts. All variables in (19) are dated t , except for ϵ which is dated $t + 1$, and with respect to which the expectation is taken. Sequentially taking derivatives of this expression and evaluating the results at the non-stochastic steady state (using the relationship between the derivatives of π and ψ alluded to above), one can solve for any desired order of the derivatives of ψ and π as functions of the first-order derivatives.

To obtain candidate first-order derivatives, let $A_{xx} = U_x T_x U_x'$ and $B = U_B T_B U_B'$ be Schur decompositions of the relevant submatrices of A . Then it can be verified that $A = \widehat{U} \widehat{T} \widehat{U}'$ is a Schur decomposition of A , where

$$\widehat{U} \equiv \begin{pmatrix} U_x & 0 & 0 \\ 0 & U_B & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\widehat{T} \equiv \begin{pmatrix} T_x & U_x^{-1} A_{x\theta} U_B'^{-1} & 0 \\ 0 & T_B & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Let \mathcal{R}_x denote the set of RGEs of A_{xx} . To obtain an A -invariant subspace satisfying (a) and (b) above, first choose a set $\{r(\lambda_1), \dots, r(\lambda_k)\}$ of k distinct RGEs from \mathcal{R}_x whose dimensions sum to n_y . Next, re-order the above Schur decomposition of A to obtain a new Schur decomposition $A = UTU'$ such that the upper-left $N \times N$ block of T contains the eigenvalues in T_B , as well as the eigenvalue 1 and the eigenvalues in $\{\lambda_1, \dots, \lambda_k\}$. Then the first N columns of U span an A -invariant subspace that contains w_{ex} .⁵⁹ To obtain the first-order approximations to the associated functions, partition the rows of U conformably with (y, z, θ, ζ) , and the columns into the first N and the remaining n_z as⁶⁰

$$U = \begin{pmatrix} U_{y1} & U_{y2} \\ U_{z1} & U_{z2} \\ U_{\theta1} & U_{\theta2} \\ U_{\zeta1} & 0 \end{pmatrix}.$$

In order for x to be in the space spanned by the first N columns of U , the final n_z rows of $U'x$ must equal zero. That is, we must have

$$z_t = -U'_{z2}{}^{-1}U'_{y2}y_t - U'_{z2}{}^{-1}U'_{\theta2}\theta_t.$$

Thus,

$$D_y^\psi = -U'_{z2}{}^{-1}U'_{y2}$$

$$D_\theta^\psi = -U'_{z2}{}^{-1}U'_{\theta2}$$

$$D_\zeta^\psi = 0$$

for this candidate solution, which can then be used in a straightforward manner to obtain $D_{(y,\theta,\zeta)}^\pi$ and, sequentially, any desired number of higher-order derivatives of ψ and π . Using π , one can then check numerically whether the stochastic sequence generated by $y_{t+1} = \pi(y_t, \theta_t, 1)$ satisfies the TVC. If it does, then we have found a solution. If not, then one

⁵⁹We have assumed here that if A_{xx} and B share any eigenvalues, then these are contained in $\{\lambda_1, \dots, \lambda_k\}$ (or in the set of complex conjugates of these eigenvalues). If this were not the case, it would not be guaranteed that U contains w_{ex} . If indeterminacy is not a concern, so that one need only in practice generate invariant subspaces that contain all of the stable eigenvalues, then this assumption about shared eigenvalues will always hold in practice. Even when this is not the case, A_{xx} and B sharing eigenvalues would typically be a knife-edge outcome in most applications, in which case this possibility can generally be ignored.

⁶⁰That the bottom-right element of this matrix is 0 follows from the facts that (a) the span of the first N columns contains w_0 , and (b) all entries in the last row and column of A are zero except for the bottom-right one.

can choose an alternative subset of \mathcal{R}_x whose dimensions sum to n_y and repeat the above process, continuing in this way until a solution is found. If after checking every such subset of \mathcal{R}_x no solution is found, then we may conclude that there is no solution.

Since it is the fundamental difference between the method in this paper and standard perturbation methods, we provide MATLAB code to implement the above process for obtaining the set of linear approximations to all the candidate solutions.⁶¹ Given any such candidate solution, one can then compute a higher-order approximation in the usual fashion. The general process for doing this is well known (see, e.g., Fernández-Villaverde, Rubio-Ramírez, and Schorfheide [2016] and the many references therein), and typically involves using software capable of obtaining analytic derivatives (such as Mathematica or the symbolic toolbox in MATLAB) to sequentially differentiate the appropriate model equations and then evaluate them at the steady state. From these, one can then create functions that take as inputs the quantitative solutions for all k -th-order and lower derivatives and use them to solve for derivatives of order $k+1$. This paper does not have anything to add to the general literature on this topic.⁶²

4.3 Stochastic Limit Cycle Example

In this section, we illustrate how to apply the methodology in the stochastic case by applying it to a stochastic version of the limit cycle example (Example 2) from Section 2.4. In particular, suppose we now include an investment demand shock θ_t . We also make an additional change in order to increase the realism of the fluctuations produced by the model. In particular, as seen in the deterministic version of the example in Section 2.4, the existing endogenous mechanisms are capable of producing limit cycles that have period 2. While such cycles may have some theoretical interest, in cases where the length of the time period is a quarter they are far too short to be empirically relevant. As shown by Beaudry, Galizia, and Portier [2018], in order to make more sustained cycles possible, one simply needs to introduce sluggishness into the investment demand equation by making current investment

⁶¹See the file `InvSub.m`, which is available for download on the author's website at <http://www.danagalizia.com/research>.

⁶²In the next subsection, we implement the proposed solution method for a particular example. The MATLAB code to do this is also provided, including code to obtain and use the relevant higher-order derivatives for that example.

also depend positively on lagged investment. We make this change here as well, so that our investment demand equation is now

$$G(I_t) = \mathbb{E}_t [I_{t+1}] - \alpha K_t + \gamma I_{t-1} + \theta_t,$$

where we assume the shock process is $\theta_t = \rho\theta_{t-1} + \zeta\epsilon_t$, and we take $\epsilon_t \sim N(0, \sigma^2)$. The capital stock K continues to evolve according to $K_{t+1} = (1 - \delta) K_t + I_t$.

In both the deterministic and stochastic cases, as part of the process to obtain a third-order approximation of the solution, one must obtain all derivatives up to third order of the relevant model equations with respect to the variables of which the solution ϕ will be a function. In the deterministic version of this example from Section 2.4, the solution was a function of only one variable (K_t), so that there were only three such derivatives to be obtained (see equations (7)-(9)), thus allowing for a relatively simple analytic characterization of the solution. In the present stochastic example, on the other hand, the solution is a function of four variables (K_t , I_{t-1} , θ , and ζ), so that there are now 34 derivatives up to the third order. Characterizing the solution analytically in this case would therefore be quite cumbersome. We thus instead illustrate the method numerically.

Figure 6 shows results for a parameterization for which, in the absence of the arrival of any investment demand shocks, the system would converge to a limit cycle.⁶³ Panel (a) illustrates the evolution of the system in $y_t = (K_t, I_{t-1})$ -space when the realized shock process is always $\theta_t = 0$,⁶⁴ beginning from two different initial points, y_0 (dotted trajectory) and y'_0 (dashed trajectory).⁶⁵ As one can see, in the absence of any shocks, the system converges to the limit cycle, which is drawn as the solid black closed curve. Panel (b) shows the evolution of I_t over time for the non-stochastic simulation beginning from y_0 (i.e., corresponding to the dotted trajectory in panel (a)). As one can see, by about $t = 75$ investment has more or less converged to the cycle, after which it repeats itself every 38 periods or so. Finally, panel (c) of Figure 6 shows a typical path for investment when we feed in a random draw for the θ_t process. Comparing panels (b) and (c), one can see that the presence of the shock

⁶³The parameters are $\delta = 0.05$, $\alpha = 0.04$, $\gamma = 2.2$, $G' = 3$, $G'' = 0$, $G''' = 0.01$, $\rho = 0.2$, and $\sigma = 2$.

⁶⁴Note that this corresponds only to the case where the *realizations* of the shocks are all zero. In particular, agents in the economy nonetheless continue to believe at date t that $\theta_{t+1} \neq 0$ is possible, and this affects their choice of I_t .

⁶⁵Here, $y_0 = (10, 1)$ and $y'_0 = (30, 8)$.

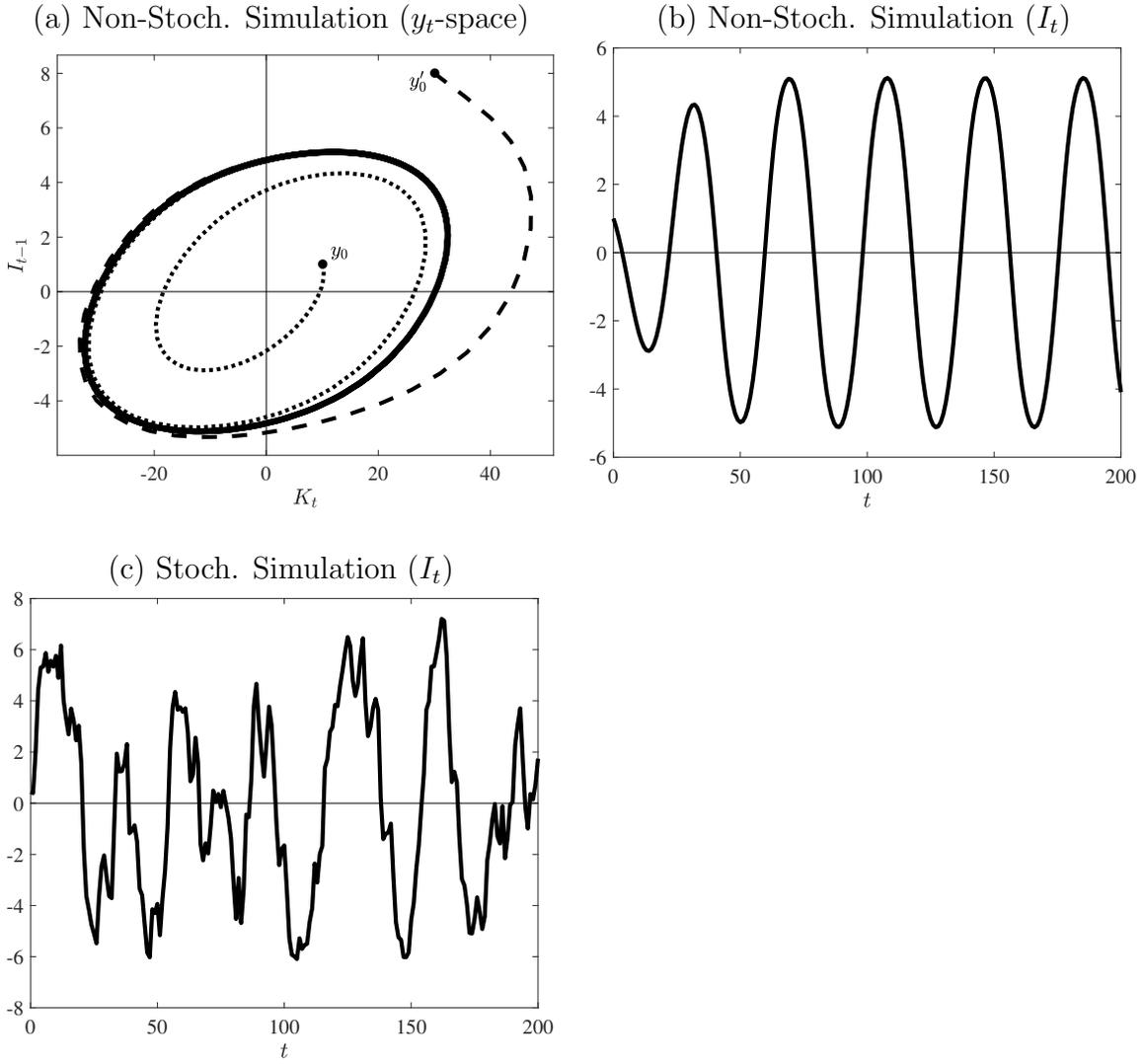
not only produces significant fluctuations in the amplitude of the cycle, but also makes the cycle quite irregular.

5 Conclusion

Non-linear rational expectations models are becoming increasingly important in the macroeconomic literature. While there are a variety of algorithms that can be used to solve these models, perturbation methods are relatively simple to implement, fast to execute, and can produce a reasonable degree of accuracy. Indeed, outside of the simplest economic models, the perturbation approach is the only one that is sufficiently fast that it can feasibly be used for estimation purposes.

Existing perturbation algorithms in the macroeconomic literature fundamentally require the linear approximation of a solution to be linearly stable; that is, they require that the eigenvalues of the solution all lie inside (or on) the complex unit circle. As a result of this constraint, these methods cannot be applied to models that feature attractive limit cycles or chaos—which generally have linearly unstable solutions—effectively excluding a large class of interesting and potentially empirically relevant economic models. As argued in this paper, however, this limitation is not inherent to perturbation methods in general, only to the precise way in which existing algorithms implement them. In particular, I generalize existing perturbation algorithms and show that the approach can be used to solve rational expectations models that may (or may not) feature limit cycles or chaos.

Figure 6: Stochastic Limit Cycle Example



Notes: The dotted and dashed curves in panel (a) show the evolution of the system beginning from $y_0 = (10, 1)$ and $y'_0 = (30, 8)$, respectively, when we feed in $\theta_t = 0$ every period. Panel (b) shows the path for I_t over time beginning from y_0 when we feed in $\theta_t = 0$ every period. Panel (c) shows a typical path for I_t when we feed in a random draw for θ_t . Model parameters for these simulations are $\delta = 0.05$, $\alpha = 0.04$, $\gamma = 2.2$, $G' = 3$, $G'' = 0$, $G''' = 0.01$, $\rho = 0.2$, and $\sigma = 2$.

References

- ARUOBA, S. B., J. FERNÁNDEZ-VILLAVERDE, AND J. F. RUBIO-RAMÍREZ (2006): “Comparing solution methods for dynamic equilibrium economies,” *Journal of Economic Dynamics and Control*, 30(12), 2477–2508.
- BEAUDRY, P., D. GALIZIA, AND F. PORTIER (2016): “Is the Macroeconomy Locally Unstable and Why Should We Care?,” in *NBER Macroeconomics Annual*, ed. by J. A. Parker, and M. Woodford, vol. 31, pp. 479–530. The University of Chicago Press.
- (2018): “Putting the Cycle Back into Business Cycle Analysis,” Working Paper.
- BLANCHARD, O. J., AND C. M. KAHN (1980): “The Solution of Linear Difference Models under Rational Expectations,” *Econometrica*, 48(5), 1305–1311.
- CALDARA, D., J. FERNÁNDEZ-VILLAVERDE, J. F. RUBIO-RAMÍREZ, AND W. YAO (2012): “Computing DSGE models with recursive preferences and stochastic volatility,” *Review of Economic Dynamics*, 15, 188–206.
- FERNÁNDEZ-VILLAVERDE, J., J. F. RUBIO-RAMÍREZ, AND F. SCHORFHEIDE (2016): “Solution and estimation methods for DSGE models,” *Handbook of Macroeconomics*, 2, 527–724.
- GASPAR, J., AND K. L. JUDD (1997): “Solving Large-Scale Rational-Expectations Models,” *Macroeconomic Dynamics*, 1(1), 45–75.
- GOLUB, G. H., AND C. F. VAN LOAN (1996): *Matrix Computations*. The Johns Hopkins University Press, Baltimore, Maryland, 3rd edn.
- GOMME, P., AND P. KLEIN (2011): “Second-order approximation of dynamic models without the use of tensors,” *Journal of Economic Dynamics and Control*, 35, 604–615.
- JUDD, K. L. (1992): “Projection Methods for Solving Aggregate Growth Models,” *Journal of Economic Theory*, 58(2), 410–452.
- (1996): “Approximation, Perturbation, and Projection Methods in Economic Analysis,” *Handbook of Macroeconomics*, 1, 509–585.
- (1998): *Numerical Methods in Economics*. MIT press.

- JUDD, K. L., AND S.-M. GUU (1997): “Asymptotic methods for aggregate growth models,” *Journal of Economic Dynamics and Control*, 21(6), 1025–1042.
- KING, R. G., AND M. W. WATSON (1998): “The Solution of Singular Linear Difference Systems Under Rational Expectations,” *International Economic Review*, 39(4), 1015–1026.
- KLEIN, P. (2000): “Using the generalized Schur form to solve a multivariate linear rational expectations model,” *Journal of Economic Dynamics & Control*, 24(10), 1405–1423.
- KOLLMAN, R., S. MALIAR, B. A. MALIN, AND P. PICHLER (2011): “Comparison of solutions to the multi-country Real Business Cycle model,” *Journal of Economic Dynamics and Control*, 35, 186–202.
- KUZNETSOV, Y. A. (1998): *Elements of Applied Bifurcation Theory*, vol. 112 of *Applied Mathematical Sciences*. Springer, New-York, 2 edn.
- MENNUNI, A., AND S. STEPANCHUK (2018): “Dynamic Perturbation,” Working Paper.
- SCHMITT-GROHÉ, S., AND M. URIBE (2004): “Solving dynamic general equilibrium models using a second-order approximation to the policy function,” *Journal of Economic Dynamics and Control*, 28, 755–775.
- SIMS, C. A. (2002): “Solving Linear Rational Expectations Models,” *Computational Economics*, 20(1-2), 1–20.

Appendix

A Proofs of Propositions

Proposition 1

For a solution $\phi: \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_z}$, let $\mu_\phi \equiv \{(y, \phi(y)) : y \in \mathbb{R}^{n_y}\}$. To see (i), suppose to the contrary that ϕ is a solution but $n_0 < n_y$. By definition, we must have $(y, \phi(y)) \in m_0$ for any y ; that is, $\mu_\phi \subset m_0$. But μ_ϕ is n_y -dimensional, and since $n_y > n_0$, μ_ϕ cannot be contained in m_0 , which establishes a contradiction.

To see (ii) and (iii), suppose $n_0 \geq n_y$. Since m_0 is a locally analytic n_0 -dimensional manifold, there exists a subset \tilde{y} of n_0 elements of x , and a locally analytic function $\psi: \mathbb{R}^{n_0} \rightarrow \mathbb{R}^{n-n_0}$ yielding the remaining $n - n_0$ elements \tilde{z} of x , such that m_0 is locally given by $\{x \in \mathbb{R}^n : \tilde{z} = \psi(\tilde{y})\}$. Note that this ψ and \tilde{y} will not necessarily be unique. Choose ψ such that \tilde{y} contains all the elements in y ,⁶⁶ with z^* denoting the remaining elements of \tilde{y} (which are jump variables). Without loss of generality, write $x = (y, z^*, \tilde{z})$.⁶⁷ Then any function $\phi(y) \in \mathbb{R}^{n_z}$ whose last $n - n_0$ elements equal $\psi(y)$ is by construction a solution. If $n_0 = n_y$, then there is exactly one such solution given by $\phi(y) = \psi(y)$. If $n_0 > n_y$, then there are an infinity of such solutions indexed by the set of all functions $\phi^* : \mathbb{R}^{n_y} \rightarrow \mathbb{R}^{n_0-n_y}$, where we then set $z^* = \phi^*(y)$. Lastly, note that since m_0 is invariant, by Proposition 2 so is ψ . Thus, if $n_0 = n_y$, then by the above reasoning $\phi = \psi$, so that ϕ must be invariant. \square

Proposition 2

To see the first part, note that by definition of a locally analytic manifold, m can be expressed as $\{\xi(a) : a \in \mathbb{R}^q\}$ for some analytic function $\xi: \mathbb{R}^q \rightarrow \mathbb{R}^n$. Without loss of generality, choose ξ so that $\xi(0) = 0$. Let τ_y be an $q \times n$ matrix “row-selection” matrix⁶⁸ such that $\xi_y(a) \equiv \tau_y \xi(a)$ is an invertible function. Let τ_z be the $(n - q) \times n$ row-selection matrix that selects the remaining elements of $\xi(a)$, and let $\xi_z(a) \equiv \tau_z \xi(a)$. Finally, set $\tilde{y} = \tau_y x$, $\tilde{z} = \tau_z x$, and $\psi(\tilde{y}) = \xi_z(\xi_y^{-1}(\tilde{y}))$. It is straightforward to verify that $m = \mu_\psi$. To see that ψ is an invariant function, note that if $x_0 \in m$ then (with appropriate re-orderings) $x_0 = (\tilde{y}_0, \psi(\tilde{y}_0))$ for some \tilde{y}_0 . Further, $x_1 \equiv h(x_0) \in m$ by invariance of m , and thus $x_1 = (\tilde{y}_1, \psi(\tilde{y}_1))$ for some \tilde{y}_1 . Invariance of ψ then follows immediately by definition.

The second part follows directly from the definition of an invariant manifold and the fact that, since h is locally analytic, so must be ψ .

Proposition 3

To see part (i), let $\psi: \mathbb{R}^q \rightarrow \mathbb{R}^{n-q}$ be the invariant function in Proposition 2 associated with m . Since w is the linear subspace tangent to m by hypothesis, it immediately follows that setting $\tilde{z} = D_{\tilde{y}}^\psi \tilde{y}$ projects x onto w for a given \tilde{y} . Next, without loss of generality, assume the elements of x are ordered such that $x = (\tilde{y}, \tilde{z})$, and partition h correspondingly as

$$h(\tilde{y}, \tilde{z}) = \begin{pmatrix} f(\tilde{y}, \tilde{z}) \\ g(\tilde{y}, \tilde{z}) \end{pmatrix}.$$

⁶⁶By the regularity assumption noted in footnote 18, this is always possible.

⁶⁷This form can always be obtained by an appropriate re-ordering of the elements of z .

⁶⁸That is, exactly one element in each row of τ_y is equal to one, and the remaining elements are zero.

We may then write

$$A = \begin{pmatrix} D_{\tilde{y}}^f & D_{\tilde{z}}^f \\ D_{\tilde{y}}^g & D_{\tilde{z}}^g \end{pmatrix}.$$

Next, note that $w \in W$ if and only if $x \in w$ implies $Ax \in w$. Suppose then that $x \in w$, which, as noted above, implies that $x = (\tilde{y}, D_{\tilde{y}}^\psi \tilde{y})$ for some \tilde{y} . We may then obtain

$$Ax = \begin{pmatrix} D_{\tilde{y}}^f + D_{\tilde{z}}^f D_{\tilde{y}}^\psi \\ D_{\tilde{y}}^g + D_{\tilde{z}}^g D_{\tilde{y}}^\psi \end{pmatrix} \tilde{y}. \quad (\text{A.1})$$

From this expression, we see that $Ax \in w$ if

$$D_{\tilde{y}}^\psi \left(D_{\tilde{y}}^f + D_{\tilde{z}}^f D_{\tilde{y}}^\psi \right) = D_{\tilde{y}}^g + D_{\tilde{z}}^g D_{\tilde{y}}^\psi. \quad (\text{A.2})$$

Now, since ψ is invariant, it follows that $g(\tilde{y}, \psi(\tilde{y})) = \psi(f(\tilde{y}, \psi(\tilde{y})))$. Totally differentiating this expression and evaluating at the steady state, we obtain (A.2), which confirms that $Ax \in w$, and thus $w \in W$.

To see (ii), fix w , and let $B \in \mathbb{R}^{(n-q) \times q}$ be the matrix that projects x onto w by choice of \tilde{z} ; i.e., $(\tilde{y}, B\tilde{y}) \in w$ for any \tilde{z} . Construct the Taylor series expression for $\psi(\tilde{y})$ around the steady state by setting $\psi(0) = 0$, $D_{\tilde{y}}^\psi = B$, and, for $k \geq 2$, given all $(k-1)$ -th order and lower derivatives of ψ , the k -th order derivatives are obtained by totally differentiating the ‘‘invariance’’ expression $g(\tilde{y}, \psi(\tilde{y})) = \psi(f(\tilde{y}, \psi(\tilde{y})))$ k times and solving.⁶⁹ Clearly, ψ is an invariant function. Define $m \equiv \mu_\psi$, where μ_ψ is as in Proposition 2. Clearly m is tangent to w at zero, and by Proposition 2 $m \in M$. Further, the ψ as constructed above is the only locally analytic function satisfying the required invariance condition, and thus m so constructed is the only element of M tangent to w , which completes the proof. \square

Proposition 4

See the discussion immediately following the statement of the Proposition.

⁶⁹As is well known, the expression to be solved will be linear in the k -th order derivatives, which can then be easily solved for in terms of known $(k-1)$ -th order and lower derivatives.