Exponential Dichotomy for Asymptotically Hyperbolic Two-Dimensional Linear Systems *

Weishi Liu† and Erik S. Van Vleck ‡

April 6, 2007

Abstract

We consider the problem of determing the existence of exponential dichotomy for a class of linear nonautonomous ODEs. An approach is introduced that combines numerical techniques with rigorous perturbation theory. It is applicable to a given problem within the class we consider and for practical purposes we develop a continuation technique. Numerical results illustrate the utility of the approach.

Keywords: Exponential dichotomy, Sacker-Sell spectrum, QR methods.
AMS Subject classifications: 35A, 65L.

1 Introduction

Exponential dichotomy (ED) is a characterization of hyperbolicity for solutions of linear non-autonomous systems. An important class of linear non-autonomous systems are related to linearization of non-linear problems along one distinguished solution or an invariant set. For example, to detect the linear stability of traveling waves, one studies the spectrum property of linear non-autonomous systems associated to the linearization along the solution. The notion of normal hyperbolicity of invariant manifolds of nonlinear flows is also closely related to exponential dichotomy (a more accurate notion is exponential separation). Theories have been developed to the study of exponential dichotomy from works of Lyapunov [14], Millionshchikov [15, 16], Bylov and Izobov [3], Coppel [7], Palmer [17, 18, 19, 20], Sacker and Sell [21], and many others. In terms of exponential dichotomy or some generalizations, various spectral theories have been established. Main results of the theory include: (1) Invariant splitting according to asymptotic rates defined by spectrum; (2) Roughness or perturbation results; (3) generic property of ED. All have said, it is a highly non-trivial task to determine whether or not a given system has an ED. The only available non-perturbation result is the Lyapunov sufficient condition and its variations.

A purpose of this work is to introduce an approach of combining analytical and numerical tools for the study of exponential dichotomy. This approach depends heavily on the analysis of QR methods developed in [9] and [10]. These works were initially developed for approximation of Lyapunov exponents, but they first established a backward error analysis and then a forward error analysis for the approximation of fundamental matrix solutions using QR methods. The basic idea of this work can be explained as follows. The numerical approximation provides in fact true

*This work was supported in part under NSF Grants DMS-0139824, DMS-0406998, and DMS-0513438.
†Department of Mathematics, University of Kansas, Lawrence, Kansas 66045 (wliu@math.ku.edu).
‡Department of Mathematics, University of Kansas, Lawrence, Kansas 66045 (evanvleck@math.ku.edu).

solutions for a nearby system. With the help of error analysis, one can rigorously estimate the departure of the computed system from the original one. In the case that the computed system has ED with enough hyperbolicity (relative to the error), one can continue the ED to the original system for a fixed problem parameter (see Theorems 3.3 and 4.3). In addition, if for a given problem parameter value the system has strong enough ED with respect to the hyberbolicity and error, then the ED may be continued to a nearby problem (see Theorem 5.1). For a fixed problem parameter we refine our error estimates starting from crude a priori estimates for that parameter value or from refined estimates already obtained for a nearby problem. Likewise, the approach can also provide quantitative information on how "close" the original system is to one without ED (see Example 6.1).

The results obtained here are similar in flavor to the Roughness Theorem for Exponential Dichotomy (see Coppel [7]), but with important differences. The basic idea is to compare a discrete time system (corresponding to an approximate solution of (2.8)) with ED to a nearby continuous time system to determine ED for the continous time system. However, the results are not just of perturbation flavor, but iterative as well in that error estimates on the solution of nonlinear scalar differential equation (2.8) are refined. We emphasize that obtaining these error estimates does not rely on having ED. We are able to resolve rigorously small neighborhoods of parameter space where there exists a problem with no ED. All quantities needed to determine these neighborhoods are computable for the specific class of problems we focus on.

To be definite in this paper we consider the following differential equation on the real line:

$$y'' - q(t)y \;=\; 0\;, \quad t \in I\!R\,, \tag{1.1}$$

where the function q is asymptotically constant; that is,

$$\lim_{t\to\pm\infty} q(t) \;=\; q_\pm > 0\;.$$

It is a general and important problem because the model (1.1) is the form in which many second order differential system can be recast (see e.g. [5, 13]). Upon rewriting this 2nd order problem as the linear system

$$\dot{x} = A(t)x, \quad \text{where} \quad A(t) = \begin{pmatrix} 0 & 1 \\ q(t) & 0 \end{pmatrix}, \tag{1.2}$$

we will investigate if (1.2) admits ED.

Recall that (1.2) has ED, if for a fundamental solution $X(t)$, there exist a projection P and constants $K \geq 1$ and $\alpha > 0$ for which

$$\begin{aligned} \|X(t)PX^{-1}(s)\| \;&\leq\; Ke^{-\alpha(t-s)}\;,\;\; t \geq s\;, \\ \|X(t)(I-P)X^{-1}(s)\| \;&\leq\; Ke^{\alpha(t-s)}\;,\;\; t \leq s\;. \end{aligned} \tag{1.3}$$

The Sacker-Sell spectrum, $\Sigma_{\rm ED}$, is defined to be those values of $\lambda \in I\!R$ for which the shifted system $\dot{x} = (A(t) - \lambda I)x$ does not have ED. As a consequence, $0 \notin \Sigma_{\rm ED}$ if and only if (1.2) does have ED.

For the case under examination here, let $A_\pm = \lim_{t\to\pm\infty} A(t)$. Then the eigenvalues of A_- are $\pm\sqrt{q_-}$ and the eigenvalues of A_+ are $\pm\sqrt{q_+}$. If we let $a_+ = \min\{\sqrt{q_+}, \sqrt{q_-}\}$, $b_+ = \max\{\sqrt{q_+}, \sqrt{q_-}\}$, $a_- = \min\{-\sqrt{q_+}, -\sqrt{q_-}\}$, and $b_- = \max\{-\sqrt{q_+}, -\sqrt{q_-}\}$, then the possibilities for $\Sigma_{\rm ED}$ of (1.2) are

1. $\Sigma_{\rm ED} \;=\; [a_-, b_-] \cup [a_+, b_+]$, in which case the system (1.2) has ED, and

2. $\Sigma_{\mathrm{ED}} = [a_-, b_+]$, in which case (1.2) does not have ED.

In spite of the apparent simplicity of the problem, determining whether or not there is ED for (1.2) is a classical, general, and important problem, as well as a formidable task. Ascertaining ED for (1.2) is also a classical problem, strongly related to the oscillatory nature of solutions of (1.1): The celebrated Lyapunov theorem (see [13]) can in fact be rephrased as an exclusion of ED for (1.2). Such results are limited in scope as they are applicable to functions q that are in small in some sense and do not allow for resolving a neighborhood in parameter space where there exists a problem with no ED. Similarly, many direct results on having ED for (1.2) are of perturbation nature; see Coppel [7]. So, in practice, these results become applicable only in very special situations, such as when q is a small perturbation of a constant value or more generally for problems for which ED for the unperturbed problem is known.

Our approach is very different from either of the above mentioned points of view: We are going to provide sufficient conditions based on numerical QR approximation for having (and not excluding) ED for (1.2) directly, and we do not rely on a perturbation argument. Rather, our approach is based upon determining whether an approximate problem has ED and then giving explicit conditions that imply the original problem has ED. Our method may be viewed as a combined numerical and analytical technique for detecting or excluding ED. The main relevance of our method is that it is applicable to very general functions q, which can also be very oscillatory in nature. The caveat is that our ability to infer that there is ED is limited by the practical limitations one has to compute with arbitrary precision. However, in the ever so important continuation context for problems with parameters, we will be able to detect explicit (and in principle arbitrarily small) range of parameters inside which there is a parameter value for which there is no ED; see Example 6.1.

Without loss of generality, we will consider the case in which $\lim_{t\to\pm\infty} q(t) = 1$, since this will simplify notation considerably. Different limiting values of q can be handled through a rescaling of time. To make the presentation clear we first consider the case where $q(t)$ is continous and $q(t) \equiv 1$ for $|t| \geq T$, and then extend the study to the case where $q(t)$ is continous and $q(t) \to 1$ as $t \to \pm\infty$.

This paper is organized as follows. In section 2 we outline the basic idea we exploit, a change of variables that transforms (1.2) to upper triangular form. For simplicity we consider the forward Euler method to approximate the orthogonal change of variables and recall the a priori global error analysis. Section 3 contains our main results that show how to obtain improved bounds on the error in approximating the orthogonal change of variables. In section 4 motivated by integral separation we show how information obtained during the numerical computations may be used to quantify the degree to which there is integral separation. We iteratively improve the bounds on the integral separation and the error in the approximate orthgonal change of variables and define a process which, under very reasonable conditions, we show converges. Section 5 contains a perturbation result which allows for efficiently obtaining bounds on the error in the orthgonal change of variables as a problem parameter changes without the need for a priori global error bound. Thus, if refined bounds have been obtained, ED may be continued to a nearby problem. In section 6 we illustrate the efficacy of our results with an example, while in section 7 we extend our results from the case with constant tails to the case with asymptotically constant tails.

2 An Equivalent Formulation of the Problem

2.1 QR Decomposition of the Fundamental Matrix Solution

We make use of an orthogonal change of variables of the form

$$Q(t) = \begin{pmatrix} \cos(\theta(t)) & \sin(\theta(t)) \\ -\sin(\theta(t)) & \cos(\theta(t)) \end{pmatrix} \tag{2.1}$$

to convert (1.2) into a system with an upper triangular coefficient matrix function. In order to do so, the change of variables Q satisfies the differential equation

$$\dot{Q} = QS(Q,A), \;\; S(Q,A)_{ij} = \begin{cases} (Q^T A Q)_{ij}, & i > j, \\ 0, & i = j, \\ -(Q^T A Q)_{ji}, & i < j, \end{cases} \tag{2.2}$$

and results in a transformed upper triangular coefficient matrix function of the form

$$B(t) := Q^T(t)A(t)Q(t) - S(Q,A)\,, \tag{2.3}$$

with upper triangular fundamental matrix solution satisfying

$$\dot{R} = B(t)R. \tag{2.4}$$

The existence of ED for (1.2) is reduced to the existence of ED for (2.4). Numerically, we will approximate Q and compare the strength of the ED for (2.4) with the error in approximating Q.

It is useful to briefly consider here the first problem: $q(t) \equiv 1$ for $|t| \geq T$, for some $T > 0$. Let A_1 denote the matrix $A(t)$ in (1.2) with $q(t) \equiv 1$, and consider

$$Q_+ = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \;\; Q_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{2.5}$$

Observe that for $Q = \pm Q_+$,

$$Q^T A_1 Q = \mathrm{diag}(1,-1)\ , \tag{2.6}$$

and for $Q = \pm Q_-$,

$$Q^T A_1 Q = \mathrm{diag}(-1,1)\ . \tag{2.7}$$

Define $Q(t) = Q_+$ for $t \leq -T$ and $\theta(-T) = -\pi/4$. Observe that for two cases of $q(t)$ defined for $t \in (-T,T)$ the equation (2.8) is extremely simple: For $q(t) = 1$, $\dot{\theta}(t) = -\cos(2\theta(t))$ and for $q(t) = -1$, $\dot{\theta}(t) = 1$. Now, if $q(t) \equiv 1$, $\theta = -\pi/4 + k\pi$, $k \in \mathbb{Z}$, are stable equilibria that correspond to Q_+ and $\theta = \pi/4 + k\pi$, $k \in \mathbb{Z}$, are unstable equilibria that correspond to Q_-. In the case we are considering, $q_\pm = 1$, the possibilities for Σ_{ED} are either $\{-1,1\}$ or $[-1,1]$. If $q(t) = -1$ for $-T < t < T$ and $q(t) = 1$ for $|t| \geq T$, then Σ_{ED} is $[-1,1]$ if and only if $\theta(T) = \pi/4 + k\pi$ for some $k \in \mathbb{Z}$. Since $\theta(-T) = -\pi/4$ and $\theta(T) = \theta(-T) + 2T$, if $q(t) = -1$ for $-T < t < T$, the spectrum is $[-1,1]$ if and only if $2T = \pi/2 + k\pi$ for some $k \in \mathbb{Z}$.

2.2 Numerical Integration of Q and A Priori Error Estimate

Using the form of Q in (2.1) we have

$$\dot{\theta}(t) = \sin^2(\theta(t)) - q(t)\cos^2(\theta(t)) \equiv f(\theta(t), q(t)) \tag{2.8}$$

We consider for simplicity the Forward Euler method applied to (2.8) and employ the classical global error result to obtain an initial bound $\rho^{(0)}$ on the global error in $\theta(t)$ from $t=-T$ to $t=T$. Integration of the equation (2.8) for θ using the forward Euler method gives

$$\theta_{j+1}=\theta_j+h_j[\sin^2(\theta_j)-q(t_j)\cos^2(\theta_j)],\ \ \theta_0=-\pi/4,\ \ h_j=t_{j+1}-t_j \tag{2.9}$$

for $j=0,...,N-1$ with $t_0=-T$ and $t_N=T$. Then

$$\theta_N=-\pi/4+\sum_{j=0}^{N-1}h_j[\sin^2(\theta_j)-q(t_j)\cos^2(\theta_j)] \tag{2.10}$$

The local error in integration of θ has the form

$$|e_j|:=|\theta_{j+1}-\theta(t_{j+1};\theta_j)|\le\frac{h_j^2}{2}\sup_{t_j\le t\le t_{j=1}}|\frac{d}{dt}f(\theta(t),q(t))| \tag{2.11}$$

where $\theta(t_{j+1};\theta_j)$ is the solution at $t=t_{j+1}$ of (2.8) with the initial condition $\theta(t_j;\theta_j)=\theta_j$. Note that

$$\frac{d}{dt}f(\theta(t),q(t))=2(1+q(t))\cos(\theta(t))\sin(\theta(t))[\sin^2(\theta(t))-q(t)\cos^2(\theta(t))]-q'(t)\cos^2(\theta(t)) \tag{2.12}$$

which can be bounded as

$$|\frac{d}{dt}f(\theta(t),q(t))|\le|1+q(t)|\cdot\max\{1,|q(t)|\}+|q'(t)|=:\Omega(t). \tag{2.13}$$

We next summarize the global error result for Forward Euler specialized to the equation considered here. We have

$$\theta(t_{j+1})-\theta_{j+1}=\theta(t_j)-\theta_j+h_j[f(t_j,\theta(t_j))-f(t_j,\theta_j)]+\frac{h_j^2}{2}\ddot{\theta}(\xi_j),\ \ \xi_j\in(t_j,t_{j+1}). \tag{2.14}$$

Then for $\Omega_j=\sup_{t\in(t_j,t_{j+1})}\Omega(t)$ for $\Omega(t)$ given in (2.13) and

$$L_j=\sup_{t\in(t_j,t_{j+1})}L(t),\quad L(t):=|1+q(t)|\ge|\frac{\partial}{\partial\theta}f(t,\theta)| \tag{2.15}$$

we have

$$|e_{j+1}|\le|e_j|+h_jL_j|\theta(t_j)-\theta_j|+\frac{h_j^2}{2}\ddot{\theta}(\xi_j)\le(1+h_jL_j)|e_j|+\frac{h_j^2}{2}\Omega_j. \tag{2.16}$$

Then in a standard way, one obtains

Proposition 2.1. *The global error in approximating θ may be bounded by*

$$\rho^{(0)}\le|e_N|\le M_N+M_{N-1}(1+h_NL_N)+M_{N-2}(1+h_NL_N)(1+h_{N-1}L_{N-1})+\cdots+M_0(1+h_NL_N)\cdots(1+h_1L_1) \tag{2.17}$$

where $M_j=\frac{h_j^2}{2}\Omega_j$.

We will employ this global error bound $\rho^{(0)}$ to obtain refined error bounds as outlined in sections 3, 4, 5. A reasonably good error bound $\rho^{(0)}$ is necessary to obtain refined error bounds. In theory, one can obtain good $\rho^{(0)}$ by employing small enough time steps, but this is often inpractical becauses it increases the computational complexity. As an integral part of our approach we employ a continuation strategy: we begin from a "simple" problem for which $\rho^{(0)}$ in Proposition 2.1 can be determined accurately and efficiently and then continue to the problem we wish to solve. The procedure can be carried out even if a member of the family of problems along the continuation path does not have ED. We remark that the error bounds ultimately obtained are generally smaller than the a priori bounds for a given set of time steps $\{h_j\}_{j=1}^N$ and much smaller as the Lipschitz constant for (2.8) becomes larger.

3 Exponential Dichotomy for Constant Tails

In this section we consider the existence of exponential dichotomy for the problem with constant tails ($q(t) = 1$ for $|t| \geq T$). The main result of this section relies on the backward error analysis of [9], arguments similar to those in [10], and the uniqueness of the QR factorization.

In exact arithmetic the solution of (2.2) may be alternatively obtained using the so-called discrete QR technique whereby the solution Q is found at discrete points by forming $X(t_{j+1}, t_j)Q(t_j)$, the product of the transition fundamental matrix solution and the solution Q at the previous time, and then forming the QR factorization,

$$Q(t_{j+1})R(t_{j+1}, t_j) = X(t_{j+1}, t_j)Q(t_j), \;\; j = 0, 1, ...$$

The exact fundamental matrix solution at $t = t_N$ is then obtained in QR form as

$$X(t_N) = Q(t_N)R(t_N, t_{N-1}) \dots R(t_2, t_1)R(t_1, t_0)R(t_0). \tag{3.1}$$

Typically one is not able to form the transition fundamental matrix solutions exactly and $X(t_{j+1}, t_j)$ is approximated by $X_{j+1,j}$ and what is found numerically is

$$Q_{j+1}R_{j+1} = X_{j+1,j}Q_j, \;\; j = 0, 1, ...$$

In approximating $Q(t)$ we form $Q_j \approx Q(t_j)$ and an approximation to the fundamental matrix solution X_N at t_N where

$$X_N = Q_N R_N R_{N-1} \dots R_2 R_1 R(t_0)\,. \tag{3.2}$$

This can be written in terms of the exact Q and R factors and the **local** error matrices, see [9] Theorem 3.1, as

$$X_N = Q(t_N)[R(t_N, t_{N-1}) + E_N] \dots [R(t_2, t_1) + E_2][R(t_1, t_0) + E_1]R(t_0)\,, \tag{3.3}$$

where $E_j = Q^T(t_j)N_jQ(t_{j-1})$ for $N_j = X_{j,j-1} - X(t_j, t_{j-1})$, the local error in approximating the transition fundamental matrix solution. The backward error result obtained in [9] is summarized next. For precise statements we refer to the original work. See, in particular, Theorem 3.12 of [9] for the result for the discrete QR method and Theorem 3.16 of [9] for the continuous QR method in which the local error in approximation of (2.2) is used to bound the local error in the transition fundamental matrix solution.

Summary 3.1. *With a numerical realization of the QR methods, we are not obtaining the triangular system (2.4-2.3), but rather the perturbed triangular system*

$$\dot{R} = \big(B(t) + E(t)\big)R\;, \tag{3.4}$$

where B is given in (2.3), and E is bounded as

$$\|E\| \leq c\eta + O(\eta^2) \leq \omega\;, \tag{3.5}$$

where $\eta := \sup_j ||E_j||$ with the main contribution to the magnification factor c being the departure from normality of the exact triangular factor R.

We will obtain in Lemma 3.4 explicit, computable bounds on $\sup_t ||E(t)||$ for the problem, (1.2), considered here.

Now, consider the 2×2 upper triangular coefficient matrix functions

$$B \text{ and } B+E, \;\; B = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}, \;\; \sup_t ||E(t)|| \leq \omega. \tag{3.6}$$

We next prove the existence of an orthogonal change of variables

$$\hat{Q}(t) = \begin{pmatrix} \cos(\varphi(t)) & \sin(\varphi(t)) \\ -\sin(\varphi(t)) & \cos(\varphi(t)) \end{pmatrix} \tag{3.7}$$

that brings the perturbed upper triangular system (3.4) to upper triangular and derive a bound on $\sup_t ||\hat{Q}(t) - I||$. Note that from (2.2) and (3.7),

$$\frac{d}{dt}\sin(\varphi(t)) = -\cos(\varphi(t))(S(\hat{Q}, B+E))_{21}, \tag{3.8}$$

where $(S(\hat{Q}, B+E))_{21} = \cos(\varphi(t))\sin(\varphi(t))(B_{11} - B_{22}) - \sin^2(\varphi(t))B_{12} + (\hat{Q}^T E \hat{Q})_{21}$.

Lemma 3.2. *There exists an orthogonal change of variables $\hat{Q}$ that brings (3.4) to $\dot{R} = \tilde{B}R$, $\tilde{B}$ upper triangular. Moreover, assume for some $K > 0$,*

$$\sup_{t \geq -T} \int_{-T}^{t} e^{-\int_\tau^t (B_{11}(r) - B_{22}(r))dr} d\tau \leq K, \tag{3.9}$$

$\beta = \alpha K$ for some $\alpha > 1$, and $\kappa = \sup_t |1 + q(t)|$. Set $a_1 = \kappa\beta^2$ and $a_2 = \kappa\beta^3$, and

$$\omega_+(\alpha, K, \kappa) := (\sqrt{a_1^2 + 4(\alpha - 1)a_2} - a_1)/(2a_2). \tag{3.10}$$

If $\omega := \sup_t ||E(t)|| < \omega_+(\alpha, K, \kappa)$, then

$$|\sin(\varphi(t))| \leq \rho := \beta \cdot \omega \equiv \alpha K \omega \;\text{ for all }\; t \in [-T, +T]. \tag{3.11}$$

Proof. From (3.8) we have for $s(t) \equiv \sin(\varphi(t))$ and $c(t) \equiv \cos(\varphi(t))$,

$$\begin{aligned} \frac{d}{dt}s &= -c[cs(B_{11} - B_{22}) - s^2 B_{12} + (\hat{Q}^T E \hat{Q})_{21}] \\ &= -s(B_{11} - B_{22}) + (s - c^2 s)(B_{11} - B_{22}) + c[s^2 B_{12} - (\hat{Q}^T E \hat{Q})_{21}] \\ &=: -s(B_{11} - B_{22}) + p(t, s, \omega). \end{aligned} \tag{3.12}$$

The proof follows from Theorem IV.2.1 in [12]. Using the nonlinear variation of constants formula we have for $s(0) = 0$,

$$\sin(\varphi(t)) = \int_{-T}^{t} e^{-\int_\tau^t (B_{11}(r) - B_{22}(r))dr} p(\tau, s(\tau), \omega) d\tau \tag{3.13}$$

Thus, $\sup_t |\sin(\varphi(t))| \leq K \sup_t |p(t, \sin(\varphi(t)), \omega)|$. We have

$$|p(t, s, \omega)| \leq |p(t, s, \omega) - p(t, 0, \omega)| + |p(t, 0, \omega)| \leq \eta(\rho, \omega)|s| + \omega \tag{3.14}$$

where

$$\eta(\rho, \omega) \leq \kappa\rho^2 + \kappa\rho, \tag{3.15}$$

using $|c - 1| \leq s^2$ for $0 \leq c \leq 1$ and $\sup_t |B_{11}(t) - B_{22}(t)| \leq \kappa, \sup_t |B_{12}(t)| \leq \kappa$. Theorem IV.2.1 of [12] may be applied if $K[\eta(\rho, \omega)\rho + \omega] < \rho$. Using the bound on $\eta(\rho, \omega)$ in (3.15) and the form for $\rho = \beta\omega$, this condition is equivalent to $a_2\omega^2 + a_1\omega + (1 - \alpha) < 0$ or $\omega < \omega_+(\alpha, K, \kappa)$ with $\omega_+(\alpha, K, \kappa)$ given in (3.10).

□

Remark 3.1. *The value α is adjustable, but a reasonable choice is $\alpha = 2$. Note that this shows that $|\sin(\varphi(t))| \le \rho$ which implies that $|\cos(\varphi(t)) - 1| \le \rho^2/(2-\rho^2)$ provided ρ is such that $0 \le \cos(\varphi(t)) \le 1$. Recall that $|\varphi(t)| \le \rho$ implies $|\sin(\varphi(t))| \le \rho$ which implies $||\hat{Q}(t) - I|| \le \sqrt{2}\rho$ where $||\cdot||$ is the two norm or the Frobenius norm.*

Remark 3.2. *An alternative approach to obtaining $\rho^{(0)}$ is to employ Lemma 3.2. However, an a priori bound on K is required, e.g. $K = 2Te^{\gamma+1}$ where $\gamma = \int_{-T}^{T} |1 + q(t)|dt$, and this tends to unduly restrict the stepsizes h_j/increase the number of steps N which adds greatly to the computational complexity.*

The following is our main result and establishes a uniform error in the approximation of the orthogonal change of variables Q.

Theorem 3.3. *Assume that $\{Q_k\}_{k=0}^N$ with Q_k an approximation to $Q(t_k)$, $-T = t_0 < t_1 < \cdots < t_N = T$, and $Q_0 = Q(-T) = Q_+$ gives a backward error result with $\|E\| \le \omega$. Suppose $\omega < \omega_+$ as defined in Lemma 3.2. For ρ as defined in (3.11) in Lemma 3.2, we have*

$$\|Q(t_k) - Q_k\| \le \sqrt{2}\rho \tag{3.16}$$

for all k.

Proof. The idea of the proof is to use the uniqueness of the QR factorization to equate terms in (3.2) and (3.3). By Lemma 3.2 there exists $\rho > 0$ such that in (3.2),

$$\hat{Q}\hat{R} = [R(t_k, t_{k-1}) + E_k] \dots [R(t_2, t_1) + E_2][R(t_1, t_0) + E_1] \tag{3.17}$$

with $\|\hat{Q} - I\| \le \sqrt{2}\rho$, $\hat{Q}$ orthogonal, and $\hat{R}$ is upper triangular with positive diagonal elements. Thus, from (3.3) we have

$$Q(t_k)\hat{Q}\hat{R} = Q_k R_k R_{k-1} \dots R_2 R_1 R(t_0). \tag{3.18}$$

So, by the uniqueness of the QR factorization, $Q(t_k)\hat{Q} = Q_k$ and

$$\|Q(t_k) - Q_k\| = \|\hat{Q} - I\| \le \sqrt{2}\rho. \tag{3.19}$$

□

In summary we compute Q_N which approximates $Q(T)$ with $Q(-T) = Q_0 = Q_+$. If $\|Q_N - Q_-\| > \sqrt{2}\rho$, then original system has ED. The difference $\|Q_N - Q_-\|$ measures the strength of the hyperbolicity while $\sqrt{2}\rho$ bounds the error in approximating $Q(T)$ (see Theorem 3.3).

For the example we are considering we have

$$B(t) = (1 + q(t)) \begin{pmatrix} -\cos(\theta(t))\sin(\theta(t)) & \cos^2(\theta(t)) - \sin^2(\theta(t)) \\ 0 & \cos(\theta(t))\sin(\theta(t)) \end{pmatrix}, \quad \dot{R}(t) = B(t)R(t) \tag{3.20}$$

where $\theta(t)$ satisfies (2.8). Observe that $B_{11}(t) = -B_{22}(t)$ and $4B_{11}^2(t) + B_{12}^2(t) = (1 + q(t))^2$. We have $\sup_t |B_{11}(t) - B_{22}(t)| \le \sup_t |1 + q(t)|$ and $\sup_t |B_{12}(t)| \le \sup_t |1 + q(t)|$.

The local error matrix E_j for Q over the interval $[t_j, t_{j+1}]$ has the form

$$E_j := Q^T(t_{j+1}; t_j, Q_j)Q_{j+1} - I = \begin{pmatrix} \cos(e_j) - 1 & \sin(e_j) \\ -\sin(e_j) & \cos(e_j) - 1 \end{pmatrix}, \quad \texttt{TOL} := \max_j \|E_j\| \tag{3.21}$$

where $Q(t_{j+1};t_j,Q_j)$ denotes the exact Q at $t=t_{j+1}$ such that $Q(t_j;t_j,Q_j)=Q_j$ and

$$Q_j=\begin{pmatrix}\cos(\theta_j) & \sin(\theta_j)\\ -\sin(\theta_j) & \cos(\theta_j)\end{pmatrix}. \tag{3.22}$$

To determine bounds on the perturbation E, from [9] (page 633) we have, see (3.5), that $\omega\le\max_j\omega_j$ where a bound for ω_j is given in the following Lemma.

Lemma 3.4. *If the local error in approximating θ on the jth time interval $[t_j,t_{j+1}]$ is bounded by $\frac{h_j^2}{2}\Omega_j$ where $h_j=t_{j+1}-t_j$, then for $t\in[t_j,t_{j+1}]$, the error $||E_j||$ in the perturbation of the upper triangular coefficient matrix function is bounded by*

$$\omega_j\le 2(1+\gamma_j e^{\gamma_j})^2\mathtt{TOL}_j e^{\gamma_j/2}+\frac{\alpha_j^2}{1-\alpha_j}e^{\gamma_j/2}(1+\gamma_j e^{\gamma_j}) \tag{3.23}$$

for $\gamma_j=\int_{t_j}^{t_{j+1}}|1+q(t)|dt$, $\mathtt{TOL}_j=h_j^2\Omega_j e^{\gamma_j/2}[1+(1-e^{h_j^2\Omega_j\gamma_j})(1+1/(h_j^2\Omega_j))]$, provided $\alpha_j=\mathtt{TOL}_j e^{\gamma_j/2}(1+\gamma_j e^{\gamma_j})<1$.

Proof. The proof follows by adapting the bounds obtained in Theorems 3.12 and 3.16 of [9] to the form of the equation (1.2) and the parameterization of the equation for Q by the function θ (2.1). By Theorem 3.12 of [9] we have that the norm of the difference of the coefficient matrix functions is bounded as

$$||E_j||\le||h_j\hat{B}_j-h_jB_j||\le||L_j||+\alpha_j^2\delta_j(1+\delta_j\nu_j)/(1-\alpha_j),\ \ \alpha_j<1 \tag{3.24}$$

where $\delta_j=\min_p[\min(1,\exp(\int_{t_j}^{t_{j+1}}B_{pp}(t)dt)]^{-1}=1$ and $\nu_j\le\gamma_j e^{\gamma_j/2}$ in our case. Again by Theorem 3.12 of [9],

$$||L_j||\le 2(1+\delta_j\nu_j)^2\mathtt{TOL}_j e^{\gamma_j/2}.$$

All that remains is to bound $\mathtt{TOL}_j$ which we do using Theorem 3.16 of [9]. We have

$$\mathtt{TOL}_j\le||\hat{N}_j||e^{\gamma_j/2}+e^{\gamma_j/2}[1-e^{-||\hat{N}_j||\gamma_j}]+||\hat{N}_j||\cdot e^{\gamma_j/2}[1-e^{-||\hat{N}_j||\gamma_j}].$$

where $||\hat{N}_j||\le h_j^2\Omega_j$ which completes the proof.

□

4 Improved Bounds

In this section we show how the initial bounds, specifically the $\rho^{(0)}$ in Proposition 2.1, obtained on the error in the orthogonal change of variables can be used as a starting point from which improved bounds may be obtained. Our strategy is, given some potentially weak control on the error in approximating θ, to bound the constant K so that using Lemma 3.2 and in particular (3.11) we derive a stronger control on the error in approximating θ. The following two lemmas establish a bound on K by perturbing from the K obtained with the approximate solution.

If $a>0$ and $d\ge 0$ are such that for $T\ge t\ge s\ge -T$,

$$\int_s^t(B_{11}(\tau)-B_{22}(\tau))d\tau\ge a(t-s)-d, \tag{4.1}$$

then

$$\sup_{t\in[-T,T]}\int_{-T}^{t} e^{-\int_\tau^t (B_{11}(r)-B_{22}(r))dr}d\tau \le \sup_{t\in[-T,T]}\int_{-T}^{t} e^{-a(t-\tau)+d}d\tau \le \frac{e^d}{a} := K. \tag{4.2}$$

The next two Lemmas provide a way to refine the error bound ρ. The idea is to use the solution and an error bound, initially $\rho^{(0)}$, to obtain an improved error bound, $\rho^{(1)}$, via an improved K and (3.11) and then use the solution and $\rho^{(1)}$ to obtain an improved $\rho^{(2)}$, etc..

Lemma 4.1. *If $\Psi(t)$ is the piecewise linear function that interpolates θ_j, $j = 0, ..., N$, and $\theta(t)$ is the exact solution such that $\theta(-T) = \theta_0$ and $|\theta(t_j) - \theta_j| \le \rho$, $j = 0, ..., N$, then for $p(t) = -2\cos(\theta(t))\sin(\theta(t))(1 + q(t))$, $t \in [t_j, t_{j+1}]$, there exists $a_j > 0$ and $d_j \ge 0$ such that for $t_j \le s \le t_{j+1}$,*

$$\int_s^{t_{j+1}} p(r)dr \ge a_j(t_{j+1} - s) - d_j. \tag{4.3}$$

where for $h_j = t_{j+1} - t_j$ and for

$$Y_j = \min_{t_j\le s\le t_{j+1}} \frac{1}{t_{j+1}-s}\int_s^{t_{j+1}} \tilde{p}(r)dr, \ \ \tilde{p}(r) \text{ defined in (4.8)}, \tag{4.4}$$

$$a_j = \begin{cases} h_j^{-1}, & h_jY_j < 1, \\ Y_j, & h_jY_j \ge 1, \end{cases} \quad d_j = \begin{cases} 1 - h_jY_j, & h_jY_j < 1, \\ 0, & h_jY_j \ge 1. \end{cases} \tag{4.5}$$

Proof. Let $\theta_j(t)$ denote the local exact solution for $t \in [t_j, t_{j+1}]$ such that $\theta_j(t_j) = \theta_j$. Then by a standard Lipschitz/Gronwall argument since $|\theta(t_j) - \theta_j| \le \rho$,

$$|\theta(t) - \theta_j(t)| \le e^{L_j t}\rho, \ \ t_j \le t \le t_{j+1}. \tag{4.6}$$

where $L_j \le \sup_{t\in[t_j,t_{j+1}]} |1+q(t)|$ is the local Lipschitz contant for (2.8) on $[t_j, t_{j+1}]$. For $t_j \le t \le t_{j+1}$, $\Psi(t)$ is the Forward Euler method starting from θ_j with stepsize $t - t_j$, so for $t_j \le t \le t_{j+1}$,

$$|\theta_j(t) - \Psi(t)| \le \frac{h_j^2}{2}(\frac{t-t_j}{t_{j+1}-t_j})^2\Omega_j = \frac{(t-t_j)^2}{2}\Omega_j \tag{4.7}$$

where $\Omega_j = \sup_{t\in[t_j,t_{j+1}]} |1 + q(t)| \cdot \max\{1, |q(t)|\} + |q'(t)|$. Thus, for $s \in [t_j, t_{j+1}]$,

$$\begin{aligned} \int_s^{t_{j+1}} p(r)dr &= 2\int_s^{t_{j+1}} (B_{11}^{\Psi}(r) + B_{11}^{\theta_j}(r) - B_{11}^{\Psi}(r) + B_{11}^{\theta}(r) - B_{11}^{\theta_j}(r))dr \\ &\ge \int_s^{t_{j+1}} [2B_{11}^{\Psi}(r) - (2e^{L_j(r-t_j)}\rho + (r-t_j)^2\Omega_j)L_j]dr =: \int_s^{t_{j+1}} \tilde{p}(r)dr. \end{aligned} \tag{4.8}$$

If $h_jY_j < 1$, then for $s \in [t_j, t_{j+1}]$, $a_j = 1/h_j$, and $d_j = 1 - h_jY_j \ge 0$,

$$\int_s^{t_{j+1}} p(r)dr \ge \int_s^{t_{j+1}} \tilde{p}(r)dr \ge (t_{j+1}-s)Y_j = (t_{j+1}-s)\left[\frac{1}{h_j} - \frac{(1-h_jY_j)}{h_j}\right] \ge \frac{1}{h_j}(t_{j+1}-s) - (1-h_jY_j). \tag{4.9}$$

If $h_jY_j \ge 1$, then for $s \in [t_j, t_{j+1}]$, $a_j = Y_j$, and $d_j = 0$,

$$\int_s^{t_{j+1}} p(r)dr \ge \int_s^{t_{j+1}} \tilde{p}(r)dr \ge (t_{j+1} - s)Y_j = a_j(t_{j+1} - s). \tag{4.10}$$

□

Lemma 4.2. *If $\Psi(t)$ is the piecewise linear function that interpolates θ_j, $j = 0, ..., N$, and $\theta(t)$ is the exact solution such that $\theta(-T) = \theta_0$ and $|\theta(t_j) - \theta_j| \leq \rho$, $j = 0, ..., N$, then for $p(t) = -2\cos(\theta(t))\sin(\theta(t))(1 + q(t))$ and $n = 1, ..., N$,*

$$\int_{-T}^{t_n} e^{-\int_s^{t_n} p(r)dr} ds \leq \sum_{j=0, d_j=0}^{n-1} \frac{(1 - e^{-Y_j h_j})}{Y_j} + (e-1) \cdot \sum_{j=0, a_j = h_j^{-1}}^{n-1} h_j e^{-h_j Y_j} =: K_n, \tag{4.11}$$

where Y_j, a_j, and d_j are given in Lemma 4.1.

Proof. The proof is a direct consequence of the following estimate

$$\int_{-T}^{t_n} e^{-\int_s^{t_n} p(r)dr} ds = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} e^{-\int_s^{t_{j+1}} p(r)dr} ds \leq \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} e^{-a_j(t_{j+1}-s)+d_j} ds = \sum_{j=0}^{n-1} \frac{e^{d_j}}{a_j}(1-e^{-a_j h_j}). \tag{4.12}$$

□

4.1 Convergence

We next show that under reasonable conditions we can produce a convergent sequence $\{\rho^{(j)}\}$ of bounds on the global error in the approximation of θ. Recall from (3.11) in Lemma 3.2 that $\rho = \alpha K \omega$. We set $\alpha = 2$, keep ω fixed, and thus define for $j = 0, 1, ...$

$$\rho^{(j+1)} = 2K^{(j+1)}\omega, \quad K^{(j+1)} := \min\{K^{(j)}, K_N^{(j+1)}\} \tag{4.13}$$

and $K_N^{(j+1)}$ is the K_N obtained from Lemmas 4.1 and 4.2 using the previous bound $\rho^{(j)}$.

Theorem 4.3. *If $\rho^{(1)} \leq \rho^{(0)}$, then there exists $\rho := \lim_{j\to\infty} \rho^{(j)}$ and $|\sin(\theta(t_j) - \theta_j)| \leq \rho$.*

Proof. The $\rho^{(j)}$ are defined using (4.13). We show that $\lim_{j\to\infty} \rho^{(j)}$ exists by constructing a monotone sequence, the proof of monotonicity is by induction. If $\rho^{(1)} \leq \rho^{(0)}$, then we must show that $\rho^{(j+1)} \leq \rho^{(j)}$ for $j = 1, 2, ...$ This clearly holds if $K^{(j+1)} \leq K^{(j)}$ which follows from (4.13).

□

Corollary 4.4. *Let $K^{(1)}$ denote the value of K_N obtained in Lemma 4.1 and 4.2 using $\rho := \rho^{(0)}$. If $\rho^{(1)} \leq \rho^{(0)}$ and for $j = 0, ..., N-1$,*

$$\overline{Y}_j = \min_{t_j \leq s \leq t_{j+1}} \frac{1}{t_{j+1} - s} \int_s^{t_{j+1}} [2B_{11}^{\Psi}(r) - (r - t_j)^2 \Omega_j L_j] dr < h_j^{-1}, \tag{4.14}$$

then $K^{(n+1)} < K^{(n)}$ and $\rho^{(n+1)} < \rho^{(n)}$ for $n = 1, 2, ...$

Proof. This follows since K_N in Lemma 4.2 decreases as ρ decreases provided $h_j Y_j < 1$ which follows from (4.4) and (4.14).

□

4.2 Minimization

To determine Y_j and hence a_j and d_j we must determine a lower bound on $\frac{1}{t_{j+1}-s}\int_s^{t_{j+1}} \tilde{p}(r)dr$ as in (4.8):

$$\begin{aligned} & \frac{1}{t_{j+1}-s}\int_s^{t_{j+1}} \tilde{p}(r)dr = \frac{1}{t_{j+1}-s}\int_s^{t_{j+1}} [2B_{11}^{\Psi}(r) - (2e^{L_j(r-t_j)}\rho + (r-t_j)^2\Omega_j)L_j]dr \\ = & \frac{1}{t_{j+1}-s}\left[\int_s^{t_{j+1}} 2B_{11}^{\Psi}(r)dr - 2(e^{L_j h_j} - e^{L_j(s-t_j)})\rho - \frac{\Omega_j L_j}{3}(h_j^3 - (s-t_j)^3)\right]. \end{aligned} \tag{4.15}$$

Using the trapezoidal rule to approximate $\int_s^{t_{j+1}} 2B_{11}^{\Psi}(r)dr$ and employing its error formula we have that

$$Y_j(\rho) \geq \min\left\{B_{11}^{\Psi}(t_j) + B_{11}^{\Psi}(t_{j+1}) - \frac{1}{6}h_j^2 \sup_{t_j\leq t\leq t_{j+1}} |\frac{d^2}{dt^2}B_{11}^{\Psi}(t)|, B_{11}^{\Psi}(t_{j+1})\right\} - 2L_j e^{L_j h_j}\rho - h_j^2\Omega_j L_j \tag{4.16}$$

and

$$\frac{d^2}{dt^2}B_{11}^{\Psi}(t) = -4\cos(2\Psi(t))\left(\frac{\theta_{j+1}-\theta_j}{h_j}\right)q'(t) - \sin(2\Psi(t))\left[q''(t) - 4\left(\frac{\theta_{j+1}-\theta_j}{h_j}\right)^2(1+q(t))\right]. \tag{4.17}$$

If (4.14) is satisfied, then

$$K_N(\rho) := (e-1)\cdot\sum_{j=0}^{N-1} h_j e^{-h_j Y_j(\rho)} \leq (e-1)\sup_{0\leq j\leq N-1}\left\{e^{2\rho L_j h_j e^{L_j h_j}}\right\}\sum_{j=0}^{N-1} h_j e^{-h_j Z_j} \tag{4.18}$$

where $Z_j = Y_j - 2\rho L_j h_j e^{L_j h_j}$.

Finally, we note that what is needed is to bound the K not for $B(t)$ but for $\tilde{B}(t)$ where $\tilde{B}(t)$ is the piecewise constant upper triangular matrix function where for $i = 1, 2$,

$$\tilde{B}_{ii}(t) = \frac{1}{h_j}\int_{t_j}^{t_{j+1}} B_{ii}(s)ds,\ t_j \leq t < t_{j+1}.$$

We have

$$\sup_{t\geq -T}\int_{-T}^{t} e^{-\int_\tau^t(\tilde{B}_{11}(r)-\tilde{B}_{22}(r))dr}d\tau \leq e^{4\kappa h}\sup_{t\geq -T}\int_{-T}^{t} e^{-\int_\tau^t(B_{11}(r)-B_{22}(r))dr}d\tau \tag{4.19}$$

where $h = \max_j h_j$, since for all j and $i = 1, 2$,

$$\int_{t_j}^{t_{j+1}} B_{ii}(r)dr = \int_{t_j}^{t_{j+1}} \tilde{B}_{ii}(r)dr, \tag{4.20}$$

$$|\int_{t_j}^{t}(B_{11}(r)-B_{22}(r))dr| \leq \kappa\cdot h_j,\ |\int_{t_j}^{t}(\tilde{B}_{11}(r)-\tilde{B}_{22}(r))dr| \leq \kappa\cdot h_j, \tag{4.21}$$

for $t_j < t < t_{j+1}$, and

$$|\int_{\tau}^{t_k}(B_{11}(r)-B_{22}(r))dr| \leq \kappa\cdot h_j,\ |\int_{\tau}^{t_k}\tilde{}(B_{11}(r)-\tilde{B}_{22}(r))dr| \leq \kappa\cdot h_j, \tag{4.22}$$

for $t_{k-1} < \tau < t_k$.

Thus, to determine $\rho := \lim_{j\to\infty}\rho^{(j)}$ we have for $h = \max_j h_j$,

$$\rho^{(j+1)} = 2\omega(e-1)e^{4\kappa h}\sup_{0\leq j\leq N-1}\left\{e^{2\rho^{(j)}L_j h_j e^{L_j h_j}}\right\}\sum_{j=0}^{N-1} h_j e^{-h_j Z_j} \equiv Ce^{\chi\rho^{(j)}}. \tag{4.23}$$

5 A Perturbation Result

In this section we prove a perturbation result that allows for continuation in a problem parameter. The idea is to avoid the need for a priori forward error estimates and instead use the information (bound on the error in the computed θ and bound on the integral separation constants) we have previously obtained to continue to a nearby problem. Generally we expect this to give a better error bound than the a priori bounds while still having the ability to improve the error bound for the new parameter value.

Consider two problems with coefficient matrix functions $A_0(t)$ and $A_1(t)$ corresponding to functions $q_0(t)$ and $q_1(t)$, respectively. We assume that for the A_0 problem we have a uniform bound $\overline{\rho}$ on the approximation of $\theta(t)$ that brings the problems to upper triangular. In addition, we assume that the information used to approximate the constant K is available.

Theorem 5.1. *Assume for the A_0 problem that for all j, $|\theta(t_j) - \theta_j| \le \overline{\rho}$ and*

$$\int_s^{t_{j+1}} p(r)dr \ge a_j(t_{j+1} - s) - d_j, \;\; p(t) = -2\cos(\theta(t))\sin(\theta(t))(1 + q_0(t)), \;\; t_j \le s \le t_{j+1} \quad (5.1)$$

with a_j, d_j given in (4.5) and $\overline{K} := K_N$ as in Lemma 4.2. Let $\kappa_i = \sup_t |1 + q_i(t)|$ for $i = 0, 1$ and assume $\sup_t |q_1(t) - q_0(t)| \le \delta$. Suppose we approximate θ for the A_1 problem with $\{\phi_j\}$ such that the perturbation matrix function is uniformly bounded by ω. If there exists $\alpha_0, \alpha_1 > 1$ such that

$$\delta < \omega_+(\alpha_0, \overline{K}, \kappa_0), \;\; \text{and} \;\; \omega < \omega_+(\alpha_1, \tilde{K}, \kappa_1) \quad (5.2)$$

where

$$\tilde{K} = \sum_{j=0, \tilde{d}_j = 0}^{N-1} \frac{(1 - e^{-\tilde{Y}_j h_j})}{\tilde{Y}_j} + (e-1) \cdot \sum_{j=0, \tilde{a}_j = h_j^{-1}}^{N-1} h_j e^{-h_j \tilde{Y}_j} \quad (5.3)$$

($\tilde{Y}_j, \tilde{a}_j, \tilde{d}_j$ are defined in the proof analogously to Y_j, a_j, d_j, respectively, in Lemma 4.1), then $|\sin(\phi(t_j) - \phi_j)| \le \rho$ where $\rho = \alpha_1 \tilde{K} \omega$.

Proof. Let $\phi(t)$ denote the exact θ that brings the A_1 problem to upper triangular. Since $\theta(t)$ brings A_0 to the upper triangular B_0, we can use $\theta(t)$ to construct a perturbed upper triangular problem $B_0(t) + E(t)$ for the A_1 problem with $\sup_t ||E(t)|| \le \delta$. If the first inequality in (5.2) holds, then by Lemma 3.2, $|\sin(\theta(t) - \phi(t))| \le \rho^* := \alpha_0 \overline{K} \delta$.

Next we proceed as in Lemma 4.1 to obtain a bound of the form

$$-2\int_s^{t_{j+1}} \cos(\phi(r))\sin(\phi(r))(1 + q_1(r))dr \ge \tilde{a}_j(t_{j+1} - s) - \tilde{d}_j, \;\; s \in [t_j, t_{j+1}]. \quad (5.4)$$

By Lemma 4.1, we have

$$-2\int_s^{t_{j+1}} \cos(\theta(r))\sin(\theta(r))(1 + q_0(r))dr \ge a_j(t_{j+1} - s) - d_j \quad (5.5)$$

and since

$$-2\cos(\phi(t))\sin(\phi(t))(1 + q_1(t)) + 2\cos(\theta(t))\sin(\theta(t))(1 + q_0(t)) \ge -2[\delta + \kappa_0 \rho^*] \quad (5.6)$$

we set $\tilde{Y}_j = Y_j - 2[\delta + \kappa_0 \rho^*]$ where Y_j is defined in Lemma 4.1.

Thus, by Lemma 4.2 with $\tilde{Y}_j, \tilde{a}_j, \tilde{d}_j$ replacing Y_j, a_j, d_j, respectively, $\tilde{K}$ is given as in (5.3), and if the second inequality in (5.2) holds, then an application of Lemma 3.2 completes the proof. □

The perturbation result allows for the continuation to a new parameter value while avoiding the need for a crude a priori estimate of $\rho^{(0)}$.

6 Example

To illustrate our results for problems with constant tails, we consider the example where $q(t)$ is given by

$$q(t) = \begin{cases} 1 - a(t+T), & -T \le t \le -T + 2/a, \\ -1 + b\sin(c(t+T-2/a)), & -T + 2/a \le t \le T - 2/a, \\ 1 + a(t-T), & T - 2/a \le t \le T, \end{cases} \tag{6.1}$$

for $a > 0$, $2T - 4/a > 0$, $b \in \mathbb{R}$, and $c \in \mathbb{R}$ such that $c(2T - a/4) = n\pi$ for some $n \in \mathbb{Z}$. Then $\Omega = \max\{2 + a, b^2 + |b|(1 + |c|)\}$ in (2.13), $\kappa = \max\{2, |b|\}$ in Lemma 3.2. This problem allows us to vary the amplitude and frequency of the potential $q(t)$ that is far from the case of pure rotation ($q(t) \equiv -1$) and the case of no rotation ($q(t) = 1$ and $\theta(-T) = -\pi/4$). The examples we consider illustrate that the combined analytical/numerical result developed here can capture exponential dichotomy or lack of exponential dichtomy to a resolution proportional to the error in the approximation of θ.

To employ (4.16) and (4.17) we need

$$q'(t) = \begin{cases} -a, & -T \le t \le -T + 2/a, \\ bc\cos(c(t+T-2/a)), & -T + 2/a \le t \le T - 2/a, \\ a, & T - 2/a \le t \le T, \end{cases} \tag{6.2}$$

and

$$q''(t) = \begin{cases} 0, & -T \le t \le -T + 2/a, \\ -bc^2\sin(c(t+T-2/a)), & -T + 2/a \le t \le T - 2/a, \\ 0, & T - 2/a \le t \le T. \end{cases} \tag{6.3}$$

Our strategy is to continue in b from $b = 0$ where we can obtain reasonable a priori bounds since for $b = 0$, $q(t) = -1$ for $-T + 2/a \le t \le T - 2/a$ over which there is no error in numerical integration. As b increases for the a priori estimate to be useful requires small timesteps which may force the complexity to become unacceptable. We fix $a = 10^2$ and consider two values of T, $T = 9\pi/4$ and $T = 19\pi/4$. Besides continuing in the parameter b, we vary the value of c that controls the frequency of the *sine* function in (6.1) by varying the integer n where $n\pi = c(2T - a/4)$. When continuing we choose the value of δ in (5.2) to be $\omega_+(2, \overline{K}, \kappa_0)/2$ and take $\alpha_1 = 2$ in the second inequality in (5.2).

For $T = 9\pi/4$, we exhibit in Figure 1 a plot of b versus θ_N for various values of n. For the values of b for which $\theta(T) = \pi/4 + k\pi$ for some integer k there is no exponential dichotomy since in this case there exists a non-trivial bounded solution, e.g. $x(0) = e_1$. The values of K, δ, and ρ obtained are fairly uniform with $K \approx 24.3$, $\delta \approx 8.9E{-}3$, and $\rho \approx 1.05E{-}4$. In Figure 1 the values of θ_N were plotted in increments of b of 10^{-1}. The figure illustrates the robust, non-monotone behavior in the θ_N value as a function of both b and n.

To determine error bars we examine Table 1. We report on three different values of $\rho^{(0)}$ that were used only for $b = 0$ to begin the calculations. We then continued in b for the different values of n considered in Figure 1. The values of K, ω_+, and δ, the size of the perturbation in b allowed in Theorem 5.1, are fairly uniform in the values of n and b considered here. In Table 1 we see that by decreasing the value of $\rho^{(0)}$ employed the error bar (of size ρ) about the computed θ value is decreased and one can restrict to a narrow range the value of b for which there is not exponential dichotomy. We are able to resolve to a resolution ρ obtained in Table 1 values of b and n in Figure 1 for which there is not exponential dichtomy.

For $T = 19\pi/4$ and $a = 10^2$ we consider a more rapidly oscillating problem and restrict attention to smaller values of the amplitude b. In Figure 2 we plot the values of θ_N versus b for

$n = 80$ for increments in b of size 10^{-4}. For these parameters values there are many values of b in the plot that correspond to no exponential dichotomy and the range of b in which there is a b^* such that the problem does not have exponential dichotomy is determined for different values of $\rho^{(0)}$ (again only employed for $b = 0$) from the value ρ in Table 2.

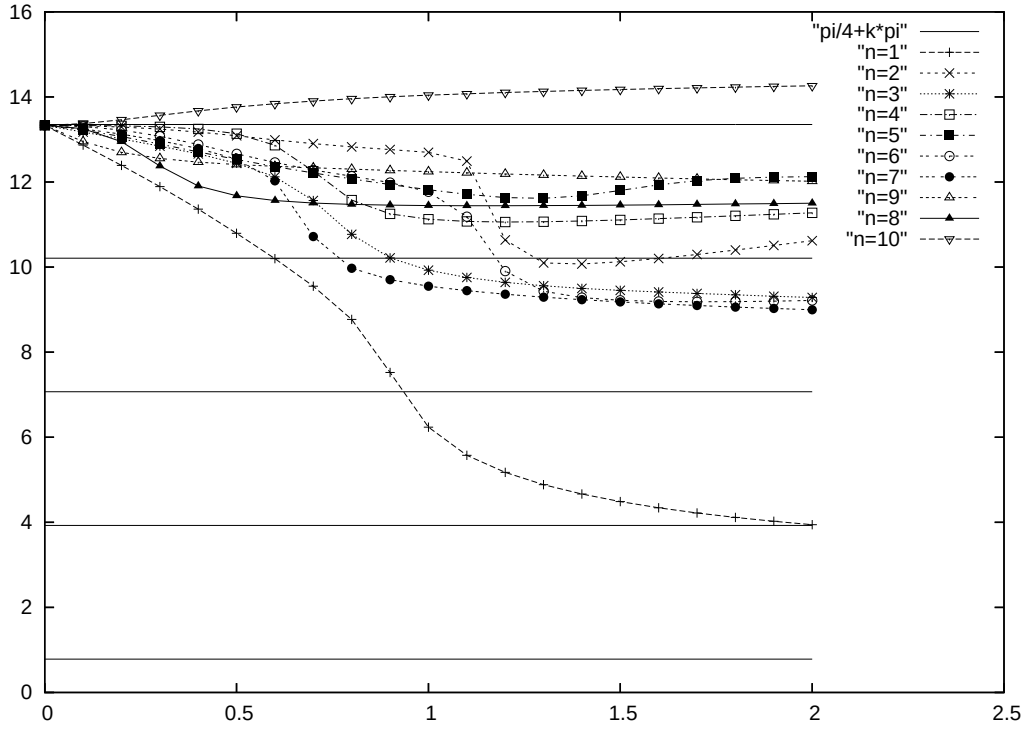

Figure 1: *Plot of (b, θ_N) for different values of n for the example with $a = 10^2$ and $T = 9\pi/4$ which shows that there is exponential dichotomy or lack of exponential dichotomy for both fixed b and varying n as well as fixed n and varying b.*

7 Exponential Dichotomy for Asymptotically Constant Tails

In this section we show how to extend these results to the case of asymptotically constant tails. Recall that $\dot{\theta}(t) = \sin^2(\theta(t)) - q(t)\cos^2(\theta(t))$.

Lemma 7.1. *Suppose that $q(t) \to 1$ uniformly as $t \to \infty$, and for $t \geq T$, $|1-q(t)| < \epsilon_{+T}$ for some $0 < \epsilon_{+T} \ll 1$. Let $\delta_{+T} = 1/2 \arcsin(\epsilon_{+T}/(2-\epsilon_{+T}))$. Then, for any $\theta_0 \in (-3\pi/4+\delta_{+T}, \pi/4-\delta_{+T})$, the solution $\theta(t)$ with $\theta(T) = \theta_0$ satisfies $\theta(t) \in (-3\pi/4 + \delta_{+T}, \pi/4 - \delta_{+T})$ for $t \geq T$ and $\theta(t) \to -\pi/4$ as $t \to \infty$.*

Proof. First of all, by [24], the assumption that $q(t) \to 1$ uniformly as $t \to \infty$ implies that either $\theta(t) \to -\pi/4$ or $\theta(t) \to \pi/4 \pmod{\pi}$ as $t \to \infty$. We will show that, if $\theta_0 \in (-3\pi/4 + \delta_{+T}, \pi/4 -$

Values of N, ω, and ρ obtained with $K \approx 24.3, \omega_+ \approx 2.1E-4$, and $\delta \approx 1.05E-4$.

$\rho^{(0)}$	N	ω	ρ
$1.E-3$	$30,025$	$4.5E-5$	$2.2E-3$
$1.E-4$	$300,249$	$4.5E-7$	$2.2E-5$
$1.E-5$	$3,002,481$	$4.5E-9$	$2.2E-7$

Table 1: Values were essentially uniform for b and n considered in Figure 1.

Values of N, ω, and ρ obtained with $K \approx 51.3, \omega_+ \approx 4.7E-5$, and $\delta \approx 2.4E-5$.

$\rho^{(0)}$	N	ω	ρ
$1.E-3$	$63,386$	$4.5E-5$	$4.6E-3$
$1.E-4$	$633,857$	$4.5E-7$	$4.6E-5$
$1.E-5$	$6,338,570$	$4.5E-9$	$4.6E-7$

Table 2: Values were essentially uniform for b considered in Figure 2.

δ_{+T}), then the solution $\theta(t)$ with $\theta(T) = \theta_0$ satisfies $\theta(t) \in (-3\pi/4 + \delta_{+T}, \pi/4 - \delta_{+T})$ for $t \geq T$, and hence, $\theta(t) \to -\pi/4$ as $t \to \infty$.

Note that, if $\theta(t_0) = \pi/4 - \delta_{+T}$ for some $t_0 > T$, then

$$\begin{aligned} 2\theta'(t_0) &= 2[\sin^2(\theta(t_0)) - q(t_0)\cos^2(\theta(t_0))] \\ &= 1 - \cos(2\theta(t_0)) - q(t_0)(1 + \cos(2\theta(t_0))) \\ &= (1 - q(t_0)) - (1 + q(t_0))\cos(\pi/2 - 2\delta_{+T}) \\ &< \epsilon_{+T} - (2 - \epsilon_{+T})\sin(2\delta_{+T}) = 0. \end{aligned}$$

Similarly, if $\theta(t_0) = -3\pi/4 + \delta_{+T}$ for some $t_0 > T$, then

$$2\theta'(t_0) = (1 - q(t_0)) - (1 + q(t_0))\cos(-3\pi/2 + 2\delta_{+T}) > -\epsilon_{+T} + (2 - \epsilon_{+T})\sin(2\delta_{+T}) = 0.$$

Therefore, the interval $(-3\pi/4 + \delta_{+T}, \pi/4 - \delta_{+T})$ is positively invariant for $t \geq T$; that is, if $\theta_0 \in (-3\pi/4 + \delta_{+T}, \pi/4 - \delta_{+T})$, then the solution $\theta(t)$ with $\theta(T) = \theta_0$ satisfies $\theta(t) \in (-3\pi/4 + \delta_{+T}, \pi/4 - \delta_{+T})$ for $t \geq T$, and hence, $\theta(t) \to -\pi/4$ as $t \to \infty$. □

Lemma 7.2. *Suppose that $q(t) \to 1$ uniformly as $t \to -\infty$, and for $t \leq -T$, $|1 - q(t)| < \epsilon_{-T}$ for some $0 < \epsilon_{-T} \ll 1$. Let $\delta_{-T} = 1/2 \arcsin(\epsilon_{-T}/(2-\epsilon_{-T}))$. Then, for any $\theta_0 \in (-\pi/4+\delta_{-T}, 3\pi/4-\delta_{-T})$, the solution $\theta(t)$ with $\theta(-T) = \theta_0$ satisfies $\theta(t) \to \pi/4$ as $t \to -\infty$. In particular, there exists $\theta_{-T} \in (-\pi/4 - \delta_{-T}, -\pi/4 + \delta_{-T})$ such that the solution $\theta(t)$ with $\theta(-T) = \theta_{-T}$ satisfies $\theta(t) \to -\pi/4$ as $t \to -\infty$.*

Proof. Using the same proof as above, one can show that the interval $(-\pi/4 + \delta_{-T}, 3\pi/4 - \delta_{-T})$ is negatively invariant for $t \leq -T$; that is, if $\theta_0 \in (-\pi/4 + \delta_{-T}, 3\pi/4 - \delta_{-T})$, then the solution $\theta(t)$ with $\theta(-T) = \theta_0$ satisfies $\theta(t) \in (-\pi/4 + \delta_{-T}, 3\pi/4 - \delta_{-T})$ for $t \leq -T$.

In view of the fact that solutions $\theta(t)$ of the equation are symmetric with respect to $\theta = \pi$, the interval $(3\pi/4 + \delta_{-T}, 7\pi/4 - \delta_{-T})$ is also negatively invariant for $t \leq -T$. The assumption that $q(t) \to 1$ uniformly as $t \to -\infty$ implies that there does exist θ_{-T} so that the solution $\theta(t)$

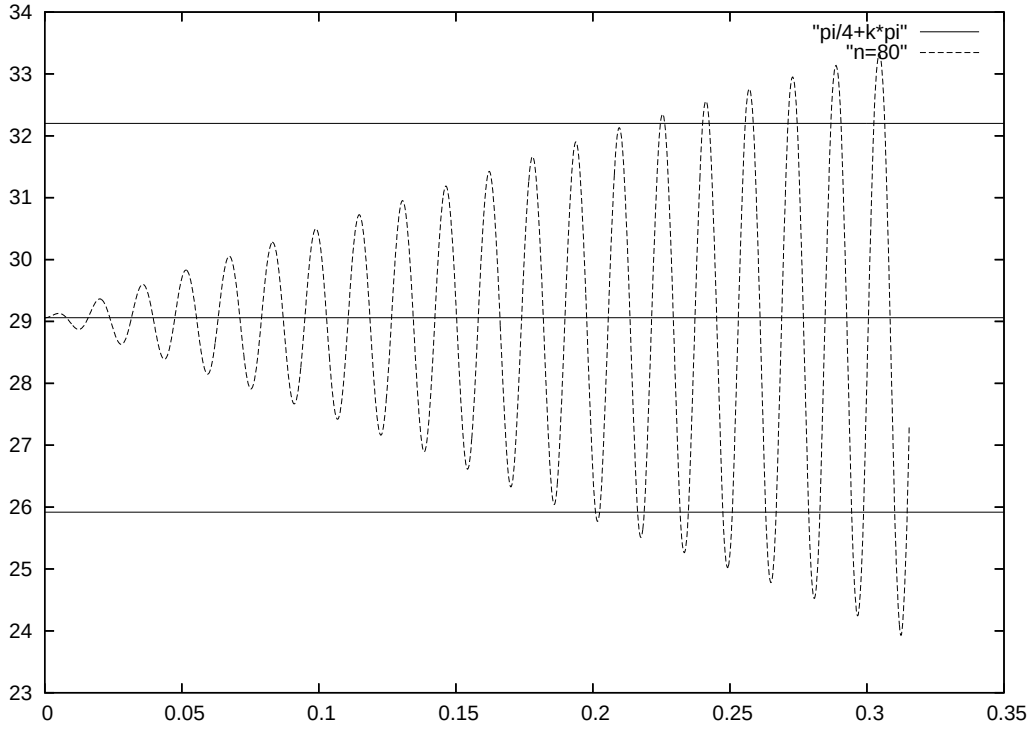

Figure 2: *Plot of* (b, θ_N) *for* $n = 80$ *for the example with* $a = 10^2$ *and* $T = 19\pi/4$. *May resolve to within* ρ *as in Table 2 the values of* b *for which there is not exponential dichotomy.*

with $\theta(-T) = \theta_{-T}$ satisfies $\theta(t) \to -\pi/4$ as $t \to -\infty$. We then claim that there exists a such $\theta_{-T} \in (-\pi/4 - \delta_{-T}, -\pi/4 + \delta_{-T})$ (and one in $(3\pi/4 - \delta_{-T}, 3\pi/4 + \delta_{-T})$). □

Theorem 7.3. *Assume there exists* $\delta_{-T} > 0$ *such that for some* θ_{-T} *with* $|\theta_{-T} + \pi/4| < \delta_{-T}$ *the solution of the initial value problem (2.8),* $\theta(-T) = \theta_{-T}$ *satisfies* $\theta(t) \to -\pi/4$ *as* $t \to -\infty$. *Assume further that there exists* $\delta_{+T} > 0$ *such that for any* θ_{+T} *with* $|\theta_{+T} - \pi/4| > \delta_{+T}$ *the solution of the initial value problem (2.8),* $\theta(+T) = \theta_{+T}$ *satisfies* $\theta(t) \not\to \pi/4$ *as* $t \to \infty$. *Then there exists* $\rho > 0$ *such that if for* $\overline{\theta}_N = \theta_N \bmod \pi$, $|\overline{\theta}_N - \pi/4| > \rho + \delta_{+T}$, *then the system (1.2) where* $q(t) \to 1$ *as* $t \to \pm\infty$ *has exponential dichotomy.*

Proof. Overall logic: Show the existence of $\theta(t)$ defined for $t \in \mathbb{R}$ such that mod π, $\theta(t) \to -\pi/4$ as $t \to \pm\infty$.

Modifications to the ω and K needed to determine ρ:

- For ω defined in (3.23), $\texttt{TOL} \mapsto \texttt{TOL} + \delta_{-T} e^{L_0 h_0}$ for $j = 0$.
- The calculation of Y_0 and hence K_1 and hence $K \equiv K_N$ becomes

$$Y_0 = \min_{t_0 \le s \le t_1} \frac{1}{t_1 - s} \int_s^{t_0} 2B_{11}^{\Psi}(r) - (2e^{L_0(r-t_0)}\delta_{-T} + (r - t_0)^2\Omega_0)L_0 dr \tag{7.1}$$

Thus, if (3.23) with $\texttt{TOL} \mapsto \texttt{TOL} + \delta_{-T} e^{L_0 h_0}$ only for $j = 0$, gives the same value ω as in the case of constant tails and $\delta_{-T} \leq \rho$ for ρ the value obtained in the case of the constant tails, then the value ρ obtained in the case of constant tails may be employed in the case of asymptotically constant tails. Otherwise, a new, potentially larger, value of ρ may be obtained.

□

8 Conclusions

In this paper we have developed techniques for determining whether a class of linear nonautonomous systems has exponential dichotomy. The technique is a combined analytical and numerical approach and relies on an error analysis for an orthogonal change of variables. If the computed system has enough hyperbolicity relative to the numerical error, then the exponential dichtomy may be continued to the original system. Techniques are developed based on using crude bounds and then obtaining refined bounds for a fixed problem parameter and also by continuing in a problem parameter using refined bounds obtained for the previous problem parameter. In particular, for the equation (1.2) considered here we are able to compute in a practical way all quantities necessary to obtain rigorous error bound on the solution to (2.8). In turn we are able to resolve rigorously small neighborhoods in parameter space where there exists a problem with no ED. The bounds obtained are much sharper than those that one might obtain using existing classical a priori error techniques. In addition, we are able to determine much more than is possible with Lyapunov type theorems that in our context perclude no ED by restricting the solution (2.8) so that a connection between $\theta(-T) = -\pi/4$ and $\theta(T) = \pi/4 + j\pi$, $j \in \mathbb{Z}$ is not possible.

We view this as a starting point and hope to develop further the ideas that form the basis of this paper. In particular, the basic ideas developed here are applicable to higher dimensional problems and the results have a wide range of applications.

Acknowledgment. We thank Luca Dieci for several helpful discussions and his input on a previous version of this paper.

References

[1] L. Ya. Adrianova, *Introduction to Linear Systems of Differential Equations*, Translations of Mathematical Monographs Vol. 146, AMS, Providence, R.I. (1995).

[2] W.J. Beyn and J. Lorenz, "Stability of traveling waves: dichotomies and eigenvalue conditions on finite intervals," *Numer. Funct. Anal. Opt.* **20** (1999), pp. 201–244.

[3] B.F. Bylov and N.A. Izobov, "Necessary and sufficient conditions for stability of characteristic exponents of a linear system," *Differentsial'nye Uravneniya* **5** (1969), pp. 1794–1903.

[4] B.F. Bylov, R.E. Vinograd, D. M. Grobman, and V.V. Nemyckii, *The theory of Lyapunov exponents and its applications to problems of stability*, Nauka Pub., Moscow (1966).

[5] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York (1955).

[6] B. A. Coomes, H. Koak, and K. J. Palmer, Homoclinic shadowing, J. Dyn. Diff. Eqn. 17 (2005), no. 1, 175–215.

[7] W.A. Coppel, *Dichotomies in Stability Theory*, Lecture Notes in Mathematics 629, Springer-Verlag, Berlin (1978).

[8] L. Dieci and E.S. Van Vleck, "Lyapunov Spectral Intervals: Theory and Computation," (2003) *SIAM J. Numer. Anal.* **40** pp. 516–542.

[9] L. Dieci and E.S. Van Vleck, "On the Error in Computing Lyapunov Exponents by QR Methods," (2005) *Numer. Math.* **101** pp. 619–642.

[10] L. Dieci and E.S. Van Vleck, "Perturbation Theory for Approximation of Lyapunov Exponents by QR Methods," (2006) *J. Dyn. Diff. Eqn.* **18** pp. 815–840.

[11] L. Dieci and E.S. Van Vleck, "Lyapunov and Sacker-Sell Spectral Intervals," (2007) *to appear in J. Dyn. Diff. Eqn.* **19** pp. XYZ–ZYX.

[12] J. K. Hale, *Ordinary Differential Equations,* (1980) Krieger.

[13] P. Hartman, *Ordinary Differential Equations,* (1982) Birkhauser.

[14] A. Lyapunov, "Problém géneral de la stabilité du mouvement," *Int. J. Control* **53** (1992), pp. 531–773.

[15] V.M. Millionshchikov, "Systems with integral division are everywhere dense in the set of all linear systems of differential equations," *Differentsial'nye Uravneniya* **5** (1969), pp. 1167–1170.

[16] V.M. Millionshchikov, "Structurally stable properties of linear systems of differential equations," *Differentsial'nye Uravneniya* **5** (1969), pp. 1775–1784.

[17] K. J. Palmer, "Exponential dichotomy, integral separation and diagonalizability of linear systems of ordinary differential equations," *J. Diff. Eqn.* **43** (1982), pp. 184–203.

[18] K. J. Palmer, "Exponential separation, exponential dichotomy and spectral theory for linear systems of ordinary differential equations," *J. Diff. Eqn.* **43** (1982), pp. 184–203.

[19] K. J. Palmer, "Exponential Dichotomies and transversal homoclinic points," *J. Diff Eqn.* **55** (1984), pp. 225–256.

[20] K. J. Palmer, "Exponential Dichotomies and Fredholm Operators," *Proc. Amer. Math. Soc.* **104** (1988), pp. 149–156.

[21] R. J. Sacker and G. R. Sell, "A spectral theory for linear differential systems," *J. Diff. Eqn.* **7** (1978), pp. 320–358.

[22] B. Sandstede, "Stability of Traveling Waves", *Handbook of Dynamical Systems* **2** (2002), pp. 983–1055.

[23] B. Sandstede and A. Scheel, "Absolute and convective instabilities of waves on unbounded and large bounded domains," *Physica D* **145** (2000) pp. 233–277.

[24] H. R. Thieme, Asymptotically autonomous differential equations in the plane. Rocky Mount. J. Math. 24 (1994), no. 1, 351–380.