Making Sense of Unstable Convergence in the Problem of Adaptive Observer Design: Case Study

Ivan Y. Tyukin ∗ David R. Fairhurst ∗∗

∗ University of Leicester, Leicester, LE1 7RH, UK (e-mail: I.Tyukin@le.ac.uk)
∗∗ University of Leicester, Leicester, LE1 7RH, UK (e-mail: drf9@le.ac.uk)

Abstract: We consider the problem of adaptive observer design for systems in which the uncertainties are modelled by ill-conditioned regressors or/and unknown parameters of the model entering the equations nonlinearly. It is shown that both of these critical issues can be overcome in a unified yet simple way, provided that the usual requirement of global asymptotic stability of solutions of model-observer system is replaced with a weaker constraint of mere set-attractivity. The claim is illustrated with a particular adaptive observer design problem.

Keywords: unstable convergence, weak attractors, ill-posed problems, adaptive observers, nonlinear parametrization

1. INTRODUCTION

Accessing variables of the system’s model which are not available for direct observation is a well-known and important problem in the context of control. Large literature on the subject exists, covering cases of systems modeled by linear ordinary equations (ODEs) (Luenberger 1979), nonlinear systems (Nijmeijer and van der Schaft 1990), and also systems of ODEs of which the right-hand side depends on unknown parameters (Marino and Tomei 1995).

Whereas models in which the uncertainty is limited to unavailability of some of the state variables for direct observation are relevant in a wide range of engineering problems, it is the models of the latter kind, with unknown parameters in the right-hand side, which are often encountered in areas related to control and identification of given physical systems. Control and identification of bioreactors (Bastin et al. 1992) and mathematical modeling of cells (Prinz et al. 2003) are standard examples of such areas.

In the context of control, the framework of adaptive observer design is perhaps the most natural concept for setting and solving the original reconstruction problem in a meaningful and mathematically tractable way. Largely, this is due to substantial theoretical progress in the field which allows to develop adaptive observers for a broad range of systems of nonlinear ODEs with linear and nonlinear parametrization (see e.g. Besançon (2000) and Farza et al. (2009)).

Yet, despite obvious advances in the theory of adaptive observers design there are certain issues of practical nature preventing successful applications of the developed tools in a range of real-life applications. One of these issues is related to the phenomenon which can be termed as occurrence of slow motions in the model-observer dynamics. These slow motions deteriorate performance of the observer and affect accuracy of the estimation in the presence of disturbance. Let us illustrate the latter claim with an example in a relatively standard setting of linearly parameterized models.

Consider the class of models which, subject to an appropriate coordinate transformation, can be described by the following systems:

\[ \dot{x} = A_0 x + B \phi^T(t) \theta \]
\[ y = C^T x, \quad C \in \mathbb{R}^n, \quad B \in \mathbb{R}^n \]

where \( A_0, C \) are in the canonical observable form, \( B = \text{col}(b_1, \ldots, b_m) \) is such that \( b_1 s^{n-1} + \cdots + b_m \) is a Hurwitz polynomial with \( b_1 \neq 0 \), \( x \in \mathbb{R}^n \) is the state vector, \( \theta \in \mathbb{R}^m \) is the vector of unknown parameters, and \( \phi : \mathbb{R} \to \mathbb{R}^m \) is the regressor.

We shall suppose that \( \phi \) is continuous, bounded, and differentiable. Consider \( A_1 = A_0 + LC^T, \quad L \in \mathbb{R}^n \) and let’s pick \( L \) such that \( C^T (A_1 - I_s)^{-1} B \) is strictly positive real. Such choice is clearly always possible. Finally, let \( \phi \) be persistently exciting. That is there exist \( T, \delta, \Delta \in \mathbb{R}_{>0} \) such that

\[ \Delta I > R(t) = \int_0^{t+T} \phi(\tau)\phi^T(\tau)d\tau > \delta I \quad \forall \ t. \]

Then

\[ \dot{x} = A_0 \dot{x} + LC^T (\dot{x} - x) + B \phi^T \theta \]
\[ \dot{\theta} = -\phi(t)C^T (\dot{x} - x) \]

is the standard adaptive observer for (1), and the error vector \( e = (\dot{x} - x, \theta - \dot{\theta}) \) evolves according to:

\[ \dot{e} = A(t)e, \quad A(t) = \begin{pmatrix} A_1 & B \phi^T(t) \\ -\phi(t)C^T & 0 \end{pmatrix}. \]

\footnote{There is a large class of systems which can be transformed into (1), see e.g. Marino (1990), Marino and Tomei (1993).}
Uniform asymptotic stability of the error dynamics (4) in this case follows from the classical work Morgan and Narendra (1977) (see also Panteley et al. (2001)).

In theory, an observer described by (3), (4), should function as long as the requirements specified so far are met. In practice, however, performance of the observer may be severely affected by presence of noise, especially when the determinant of the matrix $R(t)$ is close to zero. The latter property is illustrated with the example below.

Motivating example Let us consider the case when $A_0 = -1$, $L = 0$, $C = 1$, $B = 1$, and $\phi(t) = (\phi_1(t), \phi_2(t)) = (\sin(t), \sin(t) + \varepsilon_0 \cos(t))$, $\varepsilon_0 \in \mathbb{R}_{>0}$. In addition, suppose that there is a disturbance $d(t) = k_d \sin(t + 0.5 + \pi)$ in the right-hand side of (1):

$$\dot{x} = -x + \phi_1(t) \theta_1 + \phi_2(t) \theta_2 + d(t). \quad (5)$$

Clearly, the function $\phi(t)$ is persistently exciting for $\varepsilon_0 \in \mathbb{R}_{>0}$. Choosing the observer according to (3)

$$\dot{\hat{x}} = -\hat{x} + \phi_1(t) \hat{\theta}_1 + \phi_2(t) \hat{\theta}_2,$$

$$\dot{\hat{\theta}}_1 = -\phi_1(t)(\hat{x} - x); \quad (6)$$

$$\dot{\hat{\theta}}_2 = -\phi_2(t)(\hat{x} - x)$$

we would expect that the estimates $\hat{x}(t)$, $\hat{\theta}_1(t)$, $\hat{\theta}_2(t)$ will converge asymptotically to a neighborhood of $x(t)$, $\theta_1$, and $\theta_2$. This is certainly true, however, the size of the neighborhood may become unacceptably large, depending on how far is the value of $\det(R(t))$ from zero for some fixed $T$. This is illustrated in the Fig. 1 below. As the figure suggests the presence of a relatively small perturbation about 10 percent of the value of the relevant part of the model, $\phi_1(t) \theta_1 + \phi_2(t) \theta_2$, can result in substantial deviations (hundreds of percent) of the estimates from the true values of $\theta_1$, $\theta_2$.

The purpose of this work is to propose a simple remedy to this and similar cases which may eventually occur in practice. It is intuitively clear that observed high sensitivity of the standard observer to unmodelled perturbations is mainly due to the fact that the matrix $R(t)$ (and hence $\phi(t) = (\phi_1(t), \ldots, \phi_N(t))$) is ill-conditioned. Therefore, if the ill-conditioning of $R(t)$ can be removed then the problem will be eliminated. There are certain obstacles, however, preventing removal of the ill-condition in this problem via an appropriately chosen coordinate transformation, or by projecting the estimates onto a lower-dimensional subspace (these are discussed in the next section). Thus an alternative approach is needed.

In the present manuscript we suggest that trading exponential stability of the observer for improved robustness to perturbations may be a plausible option. In particular, we propose an observer which is capable of delivering an acceptable level of performance when the matrices $R(t)$ are ill-conditioned. The price, however, is that exponential stability of the extended system is lost. Yet, asymptotic convergence of the estimates to the set of values corresponding to indistinguishable input-output behavior is guaranteed.

The paper is organized as follows. Section 2 contains preliminaries and statement of the problem, Section 3 presents the main result, and Section 4 concludes the paper.

---

2 In fact, the linearity constraint can be removed from the approach, as has been done e.g. in Besançon (2000). However, it can be shown that for the case of non-linear error dynamics the sources of ill-conditioning in the problem will largely remain the same.
will have to satisfy:
\[
\limsup_{t \to \infty} \|e(t)\| \leq D \sup_{t \geq t_0} \|d(t)\| \tag{8}
\]
If the matrix \(A(t)\) would be known for all \(t\) then setting the value of \(\alpha\) equal to the Lyapunov exponent of system (4) is a natural option.

\textbf{Definition 1.} Let \(x : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n\) be a solution of the form \(\dot{x} = f(x,t)\) passing through \(x_0 \in \mathbb{R}^n\) at \(t = t_0\). Suppose that \(x(t, x_0)\) is defined for all \(t \geq t_0\). Then the Lyapunov exponent of \(x(t, x_0)\) is
\[
\chi(x) = \limsup_{t \to \infty} \frac{1}{t} \ln \|x(t, x_0)\|. \tag{9}
\]
\textbf{Definition 2.} Consider system \(\Sigma: \dot{x} = f(x, t), f: \mathbb{R}^n \times \mathbb{R}\), and let \(X\) be the set of all solutions of the system defined on the open intervals \([t_0, \infty)\). The Lyapunov exponent of \(\Sigma\) is
\[
\chi(\Sigma) = \max_{x \in X} \chi(x). \tag{10}
\]
If the right-hand side of \(\Sigma\) is Lipschitz in \(x\) and continuous in \(t\) then Laypunov exponent (10) always exists. This is certainly true for (4) as well.

When the matrix \(A(t)\) is not known but constants \(T, \delta, \) and \(\Delta\) are available then the value of \(\alpha\) can be estimated from above as a function \(T, \delta, \Delta\) (see e.g. Panteley et al. (2001), Tyukin (2011) for details):

\textbf{Theorem 3.} Consider system (4), and suppose that the function \(B\phi(t)^T\) in (1) be persistently exciting. Furthermore let
\[
\max \{\|\phi(t)\|, \|\phi(t)\|\} \leq B_1.
\]
Let \(\Phi(t, t_0), \Phi(t_0, t_0) = I\) be the fundamental system of solutions of (4), and \(p\) be a vector from \(\mathbb{R}^{r+m}\). Then
\[
\|\Phi(t_2, t_1)p\| \leq e^{-\rho(t_2-t_1)}\|p\|D, \forall t_2 \geq t_1 \geq t_0
\]
where parameters \(\rho\) and \(D\) do not depend on \(t_0, p,\) and can be expressed explicitly as functions of \(B_1, \delta, T,\) and matrices \(A, B, C, L\).

Notice that \(\rho\) is an upper bound for the Lyapunov exponent of (4). A straightforward derivation based on the results of Panteley et al. (2001) shows that the value of \(\rho\) is a monotone function of \(\delta: \) the smaller the value of \(\delta,\) the smaller the \(\rho,\) and if \(\delta \to 0_+\) then \(\rho \to 0_+\). Notwithstanding that this does not necessarily imply that the Lyapunov exponent of (4) should have the same quantitative behavior as the bound for \(\rho,\) qualitative asymptotic of these quantities should obviously coincide.

Since, according to (8), the disturbances propagate with the gain \(1/\alpha,\) a viable solution to the problem would be to increase the value of \(\alpha\) in the model-observer system. In particular, by increasing the value of \(\delta\) in the persistency of excitation requirement we can attempt to reduce sensitivity of the system to unmodeled disturbances.

There are, however, limitations constraining the extent of our abilities to improve the conditioning number of the problem. First of all, notice that the Lyapunov exponent is invariant with regards to any Lyapunov transformation\(^3\) Bylov et al. (1966). This rules out regularization approaches based on finding a suitable coordinate transformation. Projecting the estimates on a hyperplane which is close to the singular direction, or approaches based on model reduction, may not be a desirable substitute either. This is because they will inevitably induce an additional unmodeled dynamics, even if no perturbations are present. Hence, if no compromise on the estimation accuracy is allowed developing alternatives is needed.

Suppose that the ill-conditioning of the matrix \(R(t)\) is due to one element, \(\phi_m(t),\) of the function \(\phi(t) = (\phi_1(t), \ldots, \phi_m(t)).\) And the matrix \(\tilde{R}(t)\)
\[
\tilde{R}(t) = \int_{t}^{t+T} \tilde{\phi}(\tau)\tilde{\phi}^T(\tau)d\tau,
\]
where \(\tilde{\phi}(t) = (\phi_1(t), \ldots, \phi_{m-1}(t)),\) is supposed to be well-conditioned. Then one way to produce an alternative to the standard adaptive observer is to use conventional structure (3) for estimating the values of \(x, \theta_1, t = 1, \ldots, m-1,\) and design a new algorithm for estimating the value of \(\theta_m.\)

Formally, we can state the problem as follows. Let the model equations be given as
\[
\dot{x} = A_0x + B\phi^T(t)\theta + Bf(y, p, t) + d(t),
\]
\[
y = C^T x
\]
where matrices \(A_0, B, C,\) and the function \(\phi\) are as in (1), vector \(\theta \in \mathbb{R}^m\) is the vector of unknown parameters, and \(d : \mathbb{R} \to \mathbb{R}^n, d \in C^1, \|d\|_\infty < D_4\) is the vector corresponding to unmodeled dynamics; \(f : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}, f \in C\) is Lipschitz wrt. the unknown parameter \(p \in [p_{\min}, p_{\max}]:\)
\[
|f(y, p, t) - f(y, p', t)| < D_3|p - p'|.
\]
\(D_3\) is well conditioned. There is, however, an additional unknown parameter \(p\) which e.g. if the function \(f\) were linearly parameterized, \(f(y, p, t) = \phi_{m+1}(t)p\) can make the extended regressor \(\phi_1(t), \ldots, \phi_{m+1}(t)\) ill-conditioned.

Let us suppose that the function \(\phi\) is identically zero then (3) is an adaptive observer for (11), and that the fundamental (normalized) matrix of solutions of the error dynamics (4) satisfies
\[
\|\Phi(t, t_0)x_0\| \leq D_4e^{-\lambda(t-t_0)}\|x_0\| \forall x_0 \in \mathbb{R}^{r+m}. \tag{12}
\]
\(D_4\) is the aim to derive an adaptive observer for (11) in the form
\[
\dot{x} = A_0x + LC^T(\dot{x} - x) + B\phi^T(\tilde{\phi}) + Bf(y, \tilde{p}, t)
\]
\[
\dot{\tilde{p}} = -\phi(t)C^T(\dot{x} - x)
\]
\(\tilde{p} = W(y, \tilde{p}, t), W \in C^0\)
such that
\[
\limsup_{t \to \infty} \|e(t)\| \leq \frac{1}{\lambda}r_1\left(\sup_{t} \|d(t)\|\right), \tag{14}
\]
\[
\limsup_{t \to \infty} \|p - \tilde{p}(t)\| \leq \frac{1}{\lambda}r_2\left(\sup_{t} \|d(t)\|\right),
\]
where \(r_1, r_2 \in \mathcal{K}\) are some known non-negative monotone functions satisfying \(r_1(0) = 0, r_2(0) = 0\) which are determined exclusively by input-output properties of the perturbed system (7), and \(e(t) = (\dot{x}(t) - x(t), \theta - \bar{\theta}(t)).\) In the next section we will show that such an observer indeed can be defined.

\(^3\) That is to transformations \(z = T(t)x\) with \(\|T(t)\|, \|T^{-1}(t)\|,\)
\(\|T(t)\| < K, K \in \mathbb{R}_{>0}.)\)
3. MAIN RESULTS

We start with the following assumption on the dynamics of (4) \(^4\)

Assumption 1. Consider system (7) and suppose that \(d(t) = Bd_0(t), \|d_0\| < D_d\) for all \(t\). Let \(e_y = C_d^T e, C_e = (C, 0)\), then there exists a function \(r \in \mathcal{K}\) such that
\[
\|e_y(t)\|_{\infty, [t_0, \infty)} \leq \varepsilon \Rightarrow \|d(t)\|_{\infty, [t_1(t_0, \varepsilon), \infty)} \leq r(\varepsilon).
\]

Now we are ready to formulate the following theorem.

Theorem 4. Consider the model-observer system (11)–(13), were the function \(W\) is defined as follows:
\[
W(y, \hat{p}, t) = p_{\min} + \frac{p_{\max} - p_{\min}}{2} (\xi_1 + 1)
\]
\[
\hat{\xi}_1 = \gamma \sigma (\xi_1 - \xi_2 - \xi_1 (\xi_1^2 + \xi_2^2))
\]
\[
\hat{\xi}_2 = \gamma \sigma (\xi_1 + \xi_2 - \xi_2 (\xi_1^2 + \xi_2^2))
\]
\[
\sigma = |C_d^T e|,\]
where \(\|s\|_e = \|s\| - \varepsilon \) if \(\|s\| \geq \varepsilon\), and \(\|s\|_e = 0\) if \(\|s\| < \varepsilon\). Let us, in addition, suppose that
1) Assumption 1 holds
2) the function \(f(y, p, t)\) be nonlinearly persistently exciting in \(p\):
\[
\exists T, \beta \in \mathbb{R}_{>0}: \forall t \geq t_0, p, p' \exists t' \in [t, t + T]: |f(y, p, t) - f(y, p', t')| \geq \beta |p - p'|,
\]
3) the derivative of the function \(f(y, p, t)\) with respect to \(t\) is bounded.

Then there exist numbers \(\varepsilon(D_d) > 0, \gamma > 0\), and functions \(r_1, r_2 \in \mathcal{K}\) such that for all \(\gamma \in (0, \gamma^*]\) the following hold along the trajectories of (11)–(13):
\[
\limsup_{t \to \infty} \|e(t)\| = r_1(D_d);
\]
\[
\limsup_{t \to \infty} |p(t) - p| = r_2(D_d),
\]
provided that \(D_d\) is sufficiently small.

Sketch of the proof. The proof of the theorem is based on the ideas of a universal adaptive control scheme presented in Tyukin et al. (2008). First we notice that
\[
|e(t)| \leq D_d e^{-\lambda (t - t_0)} |e_{0}|| + \frac{D_d D_p}{\lambda} |p - \hat{p}(\tau)|_{\infty, [t_0, \infty)} + \frac{D_d D_d}{\lambda}
\]
Let for the moment consider the case when
\[
|\hat{p}(t_0) - \hat{p}(t)| \geq \Delta - \int_{t_0}^{t} |\gamma C_d^T e(\tau)|_e dr.
\]

The following result can now be used to assess asymptotic properties of (18), (19) \(^5\):

Corollary 5. Consider an interconnection of (18), (19). Then trajectories \(e(t), \hat{p}(\tau)\) passing through \(e(t_0) = e_0, \hat{p}(t_0) = p_0\) at \(t = t_0\) are bounded in forward time provided that
\[
\gamma \leq \frac{\kappa - 1}{\kappa D_d} \left[ \ln \left( \frac{\kappa}{\varepsilon} \right) \right]^{-1} \theta_0 + \frac{\epsilon_0}{1 + \frac{D_d}{\kappa}} \Delta,
\]
and
\[
\epsilon \geq \left( \frac{\kappa - 1}{\kappa D_d} \right)^{-1} (1 + \Delta),
\]
where
\[
e = \frac{D_d D_p}{\lambda}, \Delta = \frac{D_d D_d}{\lambda}, \kappa > 1, d \in (0, 1).
\]

Proof of the corollary is provided in Appendix.

Now, invoking similar argument as in the proof of Corollary 4.2 from Tyukin et al. (2008), using properties (15), (20) and Assumption 1 one can immediately establish that conclusion of the theorem follows \(\Box\).

Remark 6. The main advantage of the proposed observer, as follows from Theorem 4, is that the following holds along the solutions of the extended model-observer system:
\[
\limsup_{t \to \infty} \|e(t)\| \leq \frac{D_d}{\lambda} \left( \sup_{t} \|d(t)\| + D_d r_2 \left( \sup_{t} \|d(t)\| \right) \right),
\]
provided that \(\|d(t)\|\) is small enough. This means that if the value of \(\lambda\) is large (that is the regressor \(\phi^T(t)\theta\) is well-conditioned) then ill-conditioning of the original problem due to the presence of one extra parameter \(p\) can be removed by reformulating the problem as (11)–(13).

Remark 7. Functions \(r_1, r_2\) in (16), (17) can, in principle, be derived explicitly. The derivations, however, are technical and thus are omitted from the manuscript. Furthermore, the smallness condition on \(d(t)\) can be removed too. We left it here for consistency with the explanatory footnote regarding Assumption 1.

Remark 8. Despite observer (13) is guaranteed to achieve desired asymptotic performance, it is necessary to stress, that solutions of the model-observer system are not globally asymptotically stable. This is due to the explorative nature of the evolution of variable \(\hat{p}\). Indeed, there will always exist a perturbation which, if added to the right-hand side of the observer, will prevent it from convergence to the required neighborhood of \(x, \theta, p\). Yet, robustness of the observer can be ensured by increasing the value of \(\varepsilon\) when necessary. As the example below demonstrates, the observer offers an advantage too. The advantage is that it eliminates high sensitivity of the outcomes of the estimation for systems with ill-conditioned regressors.

Motivating example revisited. Let us illustrate application of the theorem to problem (5), (6) discussed in the Introduction and Fig. 1. Let us now choose \(\phi(t) = \phi_1(t)\). It is clear that the function \(\phi(t)\) defined in this way is well-conditioned. The function \(f(y, p, t)\) thus can be chosen as \(f(y, p, t) = \phi_2(t)p\) (where the value of \(p\) is supposed to be

\(^5\) This is a corollary from the main theorem in Tyukin et al. (2008), and its proof is provided in Appendix.
The framework is based on the idea of trading stability for accuracy and, rather counterintuitive, robustness of estimation in some sense.

The results provided so far should be considered as the case study, and albeit they illustrate the point we do not wish to suggest that this is the only tool to tackle ill-conditioning/nonlinear parametrization issues in the problem of observer design. Yet, as our present results and examples illustrate, the approach may indeed be relevant in applications.

The results themselves can be extended to deal with ill-conditioning/nonlinear parametrization in more than one variable, and these are the subjects of our future study.

4. CONCLUSION

In this paper we proposed a framework for solving the problem of adaptive observer design for systems with ill-conditioned regressors or/and nonlinear parametrization.

REFERENCES


Appendix A. PROOF OF COROLLARY 5

Let us introduce a strictly decreasing sequence \( \{\sigma_i\} \), \( i = 0, 1, \ldots \), such that \( \sigma_0 = 1 \), and \( \sigma_i \) asymptotically converge to zero. Let \( \{t_i\} \), \( i = 1, \ldots \) be an ordered sequence of time instances such that \( \hat{p}(t_i) = \sigma_i \hat{p}(t_0) \).

We wish to show that the amount of time needed to reach the set specified by \( |x(t)| = 0 \) from the given initial condition is infinite.

Consider time differences \( T_i = t_i - t_{i-1} \). It is clear that:

\[
T_i \geq \frac{\hat{p}_0(\sigma_{i-1} - \sigma_i)}{\gamma} \left( \frac{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon}{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon} - 1 \right), \quad (A.1)
\]

Consider the case when \( \|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon > 0 \) for all \( i \), and introduce the following sequence \( \{\tau_i\} \), \( \tau_i = \tau^* \), \( \tau^* \in \mathbb{R}_{>0}, i = 1, \ldots \). Sequence \( \{\tau_i = \tau^*\} \) gives rise to the series with divergent partial sums \( \sum \tau_i = \infty \). Hence proving that

\[
T_i \geq \tau^* \Rightarrow T_{i+1} \geq \tau^* \quad \forall \ i
\]

will constitute the proof that \( x(t), y(t) \) are bounded for all \( t \geq t_0 \). Let us denote \( \beta(t) = D e^{-Mt} \), let \( T_j \geq \tau^* \) for all \( 1 \leq j \leq i - 1 \), and consider

\[
\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} \leq \beta(t) \|x(t_{i-1})\| + \hat{p}_0 \sigma_{i-1} + \Delta
\leq \beta(t) \|x(T_{i-1})\| + \hat{p}_0 \sigma_{i-2} + \hat{p}_0 \sigma_{i-1} + \beta(t)(\Delta + \Delta) \leq \beta(t) \|x(t_{i-3})\| + P_2,
\]

where

\[
P_2 = \beta(t) (\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} + \hat{p}_0 \sigma_{i-2} + \hat{p}_0 \sigma_{i-1}) + \beta(t)(\Delta + \Delta) + \Delta.
\]

Repeating this iteration with respect to \( i \) leads to

\[
\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} \leq \beta(t) \|x(t_{i-1})\| + P_{i-1}
\]

\[
P_{i-1} = \hat{p}_0 \beta(t) \left[ \sum_{j=0}^{i-2} \beta(t) \sigma_{i-j} \right] + \hat{p}_0 \beta(t) \left[ \sum_{j=0}^{i-2} \beta(t) \sigma_{i-j} \right] + \Delta \beta(t) \left[ \sum_{j=0}^{i-2} \beta(t) \sigma_{i-j} \right] + \Delta,
\]

and after \( i - 1 \) steps we obtain

\[
\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} \leq \beta(t) \|x(t_{i-1})\| + P_{i-1}
\]

\[
P_{i-1} = \hat{p}_0 \beta(t) \left[ \sum_{j=0}^{i-2} \beta(t) \sigma_{i-j} \right] + \hat{p}_0 \beta(t) \left[ \sum_{j=0}^{i-2} \beta(t) \sigma_{i-j} \right] + \Delta \beta(t) \left[ \sum_{j=0}^{i-2} \beta(t) \sigma_{i-j} \right] + \Delta, \quad (A.2)
\]

From (A.1) it follows that

\[
T_i \geq \frac{\sigma_{i-1} - \sigma_i \hat{p}_0}{\sigma_{i-1} - \sigma_i \hat{p}_0} \frac{1}{\gamma} \left( \frac{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon}{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon} - 1 \right) \quad (A.3)
\]

Hence, if we can show that there exist \( x_0 \) such that such that for some \( \Delta_0 \in \mathbb{R}_{>0} \) and \( \varepsilon \):

\[
\frac{\sigma_{i-1} - \sigma_i \hat{p}_0}{\sigma_{i-1} - \sigma_i \hat{p}_0} \geq \Delta_0
\]

the following holds

\[
\gamma \sigma_{i-1} \left( \frac{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon}{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon} \right) \leq \gamma B(x_0) \leq \Delta_0 \forall i, \quad (A.4)
\]

where \( B(\cdot) \) is a function of \( x_0 \), then boundedness of trajectories will follow. Consider the term \( \sigma_{i-1} \|x(\tau)\|_{\infty, [t_{i-1}, t_i]} \), and let

\[
\sigma_i = \frac{1}{\kappa^i}, \quad \kappa > 1
\]

According to (A.2) we have:

\[
\sigma_{i-1} \left( \frac{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon}{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon} \right) \leq \beta(t) (\|x(t_0)\| + \hat{p}_0 \beta(t) \left[ \sum_{j=0}^{i-2} \beta(t) \sigma_{i-j} \right] + \Delta - \varepsilon).
\]

Hence choosing the value of \( \tau^* \) as

\[
\kappa \beta(t)(\tau^*) \leq d, \quad d \in (0, 1)
\]

results in the following estimate:

\[
\sigma_{i-1} \left( \frac{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon}{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon} \right) \leq B(x_0) = \beta(t) (\|x(t_0)\| + \hat{p}_0 \beta(t) \left[ \sum_{j=0}^{i-2} \beta(t) \sigma_{i-j} \right] + \Delta - \varepsilon).
\]

Condition (20) implies that \( \kappa - 1 \left[ \frac{\beta(t)(\tau^*)}{1 - \varepsilon} + 1 \right] \leq 0 \).

Hence

\[
\sigma_{i-1} \left( \frac{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon}{\|x(\tau)\|_{\infty, [t_{i-1}, t_i]} - \varepsilon} \right) \leq B(x_0) = \beta(t) (\|x(t_0)\| + \hat{p}_0 (1 + \frac{\beta(t)(\tau^*)}{\kappa - 1})).
\]

Solving (A.5), (A.3) with respect to \( \Delta_0 \) results in

\[
\Delta_0 = \frac{\kappa - 1}{\kappa} \left[ \beta(t)(\tau^*) \right] \hat{p}_0.
\]

This in turn implies that for all \( x_0, \hat{p}_0 \) such that:

\[
\gamma \leq \frac{\kappa - 1}{\kappa} \left[ \beta(t)(\tau^*) \right] \hat{p}_0
\]

the following implication must hold:

\[
\Delta_0 \geq \tau^* \Rightarrow T_{i+1} \geq \tau^*.
\]

Therefore, trajectories \( x(t), y(t) \) passing through \( x(t_0) = x_0, y(t_0) = \hat{p}_0 \) at \( t = t_0 \) are bounded in forward time. □