Undershoot and Settling Time Trade-offs for Nonlinear Non-minimum Phase Systems

K. Lau*, R.H. Middleton * and J.H. Braslavsky †

Abstract

It has been known for some time that real non-minimum phase zeros imply undershoot in the step response of linear systems. Bounds on such undershoot depend on the settling time demanded and the zero locations. In this paper we review such constraints for linear time invariant systems and provide new stronger bounds that consider simultaneously the effect of two real NMP zeros. Using the concept of zero dynamics, we extend these results to a class of nonlinear systems. In particular, using concepts of constrained reachability, we show that scalar separable unstable zero dynamics imply undershoot in the step response. Furthermore, this undershoot cannot be small if a rapid settling time is required and the zero dynamics are slow.

1 Introduction

The study of performance trade-offs for feedback control systems has been studied for many years (see for example the monographs: [1] and [2]). This area of research aims to examine and expound fundamental compromises in the achievable performance of a feedback control system. Such studies have used both frequency domain (sensitivity functions, interpolation constraints, achievable H_∞ performance) and time domain (cheap control, undershoot-overshoot, settling time and rise time, L_∞ and L_2 performance) analyses. The study of performance trade-offs is most clearly developed for linear systems from a frequency domain perspective.

In attempting to extend these results to nonlinear systems, one natural avenue to explore is the extension of time domain constraints. Such an approach has been taken in [3] to extend the cheap control results of [4] to a class of nonlinear systems.

In this paper, we consider the extension of the results on undershoot and settling time constraints for non-minimum phase linear systems obtained in [5] using frequency domain tools. As in [3], we use the concept of zero dynamics [6] for the characterisation of non-minimum phase nonlinear systems. In the context of the zero dynamics formulation, we first rederive the results of [5] for linear systems, and provide new, stronger undershoot bounds that consider simultaneously the effect of two real non-minimum phase zeros.

We then present an extension of these results for a class of nonlinear systems with unstable zero dynamics. Using concepts of constrained reachability, we show that the step response of systems with scalar separable unstable zero dynamics must necessarily undershoot. Furthermore, as is the case for linear systems, the undershoot cannot be small if a short settling time is required when

*Centre for Integrated Dynamics and Control, The University of Newcastle, Callaghan, 2308 NSW, AUSTRALIA. email: {eekl,rick}@ee.newcastle.edu.au

†Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Calchaquí 5800, F. Varela (1888), ARGENTINA. email: jbrasla@unq.edu.ar

the zero dynamics are comparatively slow. We expound these results on two nonlinear system case studies, which qualitatively illustrate the extent of the performance trade-off.

1.1 Background and Definitions

Consider the general problem of trying to control a single input single output nonlinear time invariant plant,

$$y = \mathcal{P} * u \tag{1}$$

where $\mathcal{P}$ denotes, in the most general setting, a dynamic nonlinear operator mapping input signals $u(t)$ to output signals $y(t)$. In this paper, we are primarily interested in step response tracking examples, where we wish to move from an equilibrium[1] ($y = 0$) to a new equilibrium ($y = \bar{y}$). We are particularly interested in the *undershoot* that may occur during such a transition. Given an output signal $y(t)$ we define the *relative undershoot* r_{us} as

$$r_{us}(y(\cdot)) = -\inf_{t\in(0,\infty)} \left\{ \frac{y(t)}{\bar{y}} \right\}.$$

Undershoot is an important time domain characteristic of a control system. There are several reasons why it may be undesirable to permit excessive undershoot in a plant response. Firstly, large undershoot may cause state constraints to be violated. Secondly, large undershoot may be, and in the linear case definitely is, an indication of poor sensitivity robustness properties. Thirdly, large undershoot may deceive a supervisor or operator into believing the control system is faulty, and therefore manually intervening in a control system.

We are also interested in the *settling time* of a feedback control system. The settling time has a variety of definitions by different authors. In this paper, for simplicity, we use the following definition of an exact settling time, T_{es} :

$$T_{es}(y(\cdot)) = \inf_{\tau\in(0,\infty)} \{\tau : t > \tau \Rightarrow y(t) = \bar{y}\}.$$

Note that for many real control systems (for example, linear time invariant systems), the output y may not be able to settle exactly in finite time. However, the output may be able approximate, to an arbitrary degree of accuracy, an 'ideal' output signal which does have finite exact settling time. This approximation significantly simplifies the analysis in this paper.

2 Linear Time Invariant Plants

In the case where the plant can be described by a linear, causal, finite dimensional operator, we replace the description of Equation (1) by the following rational transfer function description:

$$y = P(s) * u,$$

where the notation '$P(s) * u$' represents convolution of the impulse response of $P(s)$ with u.

In this case, we say that the plant transfer function $P(s)$ is *minimum phase* if all of its zeros have non-positive real parts. If any of the zeros of $P(s)$ have positive real part, then we say that $P(s)$ is *non-minimum phase* (NMP). The study of NMP zeros, and their effects on time domain control properties have been studied previously by several authors [5], [2], [7, Chap. 4], [8], [9]. We first review the situation for a single real NMP zero.

[1] We assume, without loss of generality, that the initial equilibrium is at the origin.

2.1 Single real NMP Zero

Consider the case where we have a single real NMP zero at $s = \lambda$. We then have the following result:

Proposition 2.1 *[5] Suppose that $P(\lambda) = 0$ where λ is a positive real number. Consider any input-output signals such that $u(t)$ is bounded and the output $y(t)$ settles exactly to $\bar{y}$ in time T. Then the relative undershoot must satisfy*

$$r_{us}(y) \geq \frac{1}{e^{\lambda T} - 1}.$$

Proof

[2]Since $u(t)$ is bounded, λ is in the region of convergence of $U(s) = \mathscr{L}\{u(t)\}$ (the Laplace transform of $u(t)$). Therefore, since $Y(\lambda) = P(\lambda)U(\lambda) = 0$,

$$\int_0^\infty y(t)e^{-\lambda t}dt = 0. \tag{2}$$

Then by splitting the interval of integration in Equation (2), and using the definitions of exact settling, we obtain

$$\begin{aligned}\left(\frac{e^{-\lambda T}}{\lambda}\right) &= \int_T^\infty e^{-\lambda t}dt \\ &= \frac{1}{\bar{y}}\int_T^\infty y(t)e^{-\lambda t}dt \\ &= \int_0^T -\frac{y(t)}{\bar{y}}e^{-\lambda t}dt \\ &\leq r_{us}(y)\left(\frac{1-e^{-\lambda T}}{\lambda}\right)\end{aligned}$$

from which the result follows. □

We now wish to expand on this result to consider the case where we have two real NMP zeros.

2.2 Two Real NMP Zeros

Suppose that we have a LTI plant with two real NMP zeros at $s = \lambda_1$ and $s = \lambda_2$. Then using the same arguments as in the previous section, we obtain two interpolation constraints of the form of Equation (2). Of course, Proposition 2.1 applies (at least as a lower bound) individually to each of these two NMP zeros. However, the interaction of the two interpolation constraints gives stronger results as we now show:

Proposition 2.2 *Suppose that $P(\lambda_1) = P(\lambda_2) = 0$ where $\lambda_1 > \lambda_2 > 0$ are real numbers. Consider any input-output signals such that $u(t)$ is bounded and the output $y(t)$ settles exactly to $\bar{y}$ in time T. Then the relative undershoot must satisfy*

$$r_{us}(y) \geq \frac{\lambda_1 e^{-\lambda_2 T} - \lambda_2 e^{-\lambda_1 T}}{\lambda_1\left(1-e^{-\lambda_2 T}\right) - \lambda_2\left(1-e^{-\lambda_1 T}\right)}. \tag{3}$$

[2]We include the proof here since it is brief, and is instructive for the remaining development of the paper.

Proof

In a similar fashion to Proposition 2.1 we obtain the following interpolation conditions:

$$\int_0^\infty y(t)e^{-\lambda_1 t}dt = 0$$
$$\int_0^\infty y(t)e^{-\lambda_2 t}dt = 0.$$

Because of the exact settling time assumption, and defining $r_1 = \frac{1}{e^{\lambda_1 T}-1}$, $r_2 = \frac{1}{e^{\lambda_2 T}-1}$, $\gamma(t) = \left(-\frac{y(t)}{\bar{y}} - r_2\right)$ these conditions may be transformed to the equations:

$$\int_0^T \gamma(t)e^{-\lambda_1 t}dt = \left(\frac{e^{-\lambda_1 T}}{\lambda_1}\right) - r_2\left(\frac{1-e^{-\lambda_1 T}}{\lambda_1}\right), \tag{4}$$

$$\int_0^T \gamma(t)e^{-\lambda_2 t}dt = 0. \tag{5}$$

We now split $\gamma(t) = \gamma^+(t) - \gamma^-(t)$ where both $\gamma^+(t)$ and $\gamma^-(t)$ are non-negative functions of time. Note that these definitions with Equation (5) imply that

$$\int_0^T \gamma^+(t)e^{-\lambda_2 t}dt = \int_0^T \gamma^-(t)e^{-\lambda_2 t}dt.$$

From Equation (4) we have

$$\begin{aligned}
r_2\left(\frac{1-e^{-\lambda_1 T}}{\lambda_1}\right) - \left(\frac{e^{-\lambda_1 T}}{\lambda_1}\right) &= -\int_0^T \gamma(t)e^{-\lambda_1 t}dt \\
&= \int_0^T \gamma^-(t)e^{-\lambda_1 t}dt - \int_0^T \gamma^+(t)e^{-\lambda_1 t}dt \\
&\leq \int_0^T \gamma^-(t)e^{-\lambda_2 t}dt - \int_0^T \gamma^+(t)e^{-\lambda_1 t}dt \\
&= \int_0^T \gamma^+(t)e^{-\lambda_2 t}dt - \int_0^T \gamma^+(t)e^{-\lambda_1 t}dt \\
&= \int_0^T \gamma^+(t)\left(e^{-\lambda_2 t} - e^{-\lambda_1 t}\right)dt \\
&\leq \int_0^T (r_{us}(y) - r_2)\left(e^{-\lambda_2 t} - e^{-\lambda_1 t}\right)dt \\
&= (r_{us}(y) - r_2)\left(\left(\frac{1-e^{-\lambda_2 T}}{\lambda_2}\right) - \left(\frac{1-e^{-\lambda_1 T}}{\lambda_1}\right)\right)
\end{aligned}$$

from which the desired result follows. □

This result may be illustrated as shown in Figure 1. Note that by taking each constraint individually, we would only get the results for $\frac{\lambda_1}{\lambda_2} \to \infty$. Clearly from the figure, if the zeros are not widely spaced, the results when the two constraints are considered together may be many times worse than for the individual constraints. For example, consider $\lambda_2 T = 1$. As $\lambda_1 \to \infty$ the lower bound on the relative undershoot is 0.6 but when λ_1 is close to λ_2 the bound on $r_{us}(y)$ is approximately 2.78.

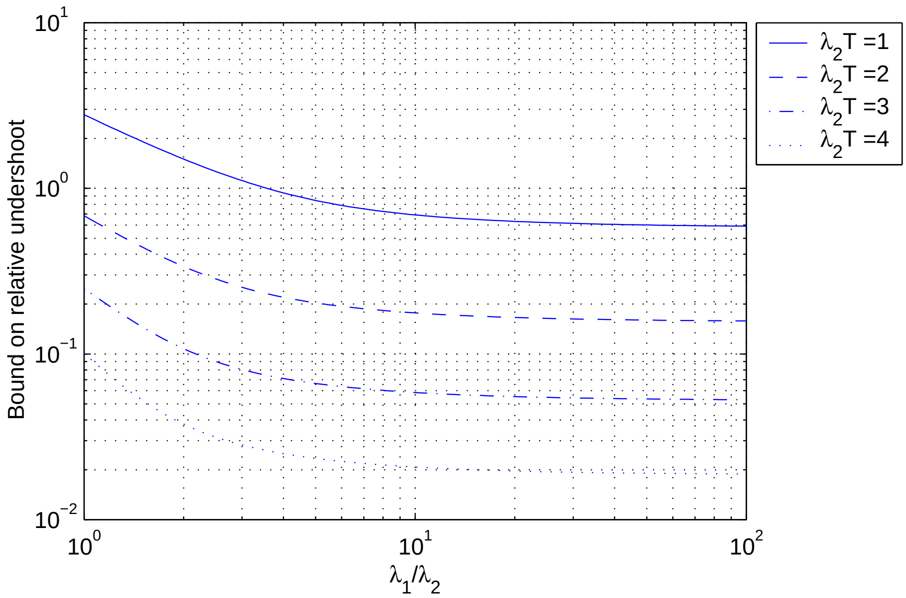

Figure 1: Graph showing the bound on $r_{us}(y)$ given in Inequality (3) versus λ_1/λ_2 for $\lambda_1 > \lambda_2$ and $\lambda_2 T = 1, 2, 3$ and 4.

The results have been derived based on Laplace transform analysis of an LTI plant. However, when considering nonlinear plants, such analysis may be inappropriate and difficult to generalise. To facilitate analysis of nonlinear plants, we now consider the same problems, using a zero dynamics formulation.

2.3 Linear Zero Dynamics Formulation

The zero dynamics formulation (see for example [10, pp538]) for a linear system performs a state transformation from the generic state space form:

$$\begin{aligned}\dot{x} &= Ax + Bu\\ y &= Cx\end{aligned}$$

to the zero dynamics form:

$$\begin{aligned}\dot{\xi} &= A_c\xi + B_cC_0z + B_cu\\ \dot{z} &= A_0z + B_0y \qquad (6)\\ y &= C_c\xi.\end{aligned}$$

In this case, the zero dynamics refers to the dynamics of the equation $\dot{z} = A_0 z$ and in particular, the eigenvalues of A_0 are the zeros of the transfer function, $P(s)$. Hence a NMP system has unstable zero dynamics.

With the system in zero dynamics form, one way of considering the relationship of undershoot to settling time and NMP zeros is as follows: Take the target output $\bar{y}$ and form the target final state, $\bar{z} = -A_0^{-1}B_0\bar{y}$. Without loss of generality, let $\bar{y} > 0$. For each $\alpha > 0$, let $\mathcal{Y}_\alpha$ denote the set of functions y which satisfy

$$y(t) \geq -\alpha \;\forall\, t \geq 0.$$

Given a permissable level of relative undershoot ρ we seek an answer to the constrained reachability question: *Does there exist an output* $y \in \mathcal{Y}_{\rho\bar{y}}$ *which takes* $z(0) = 0$ *to* $z(t) = \bar{z}$?

2.3.1 First Order Linear Zero Dynamics

In the case of unstable first order zero dynamics, the constrained reachability question has a straightforward answer as we will now show. Without loss of generality, we take $A_0 = \lambda > 0$ and $B = 1$, $\bar{y} > 0$. In this case, clearly $\bar{z} = -\frac{\bar{y}}{\lambda}$ and the general solution to the linear zero dynamics equation is given by

$$z(t) = \int_0^t e^{\lambda(t-\tau)} y(\tau)\, d\tau. \tag{7}$$

From Equation (7), and for $y \in \mathcal{Y}_\alpha$, $\alpha > 0$, we obtain

$$\begin{aligned} -z(T) &\le \int_0^T e^{\lambda(T-\tau)} \alpha\, d\tau \\ &= \alpha \frac{e^{\lambda T} - 1}{\lambda}. \end{aligned} \tag{8}$$

Clearly therefore, unless $\alpha\left(e^{\lambda T} - 1\right) \ge \bar{y}$ Inequality (8) contradicts $z(t) = \bar{z}$. It follows that $r_{us}(y)\left(e^{\lambda T} - 1\right) \ge 1$ which is equivalent to the lower bound on the undershoot given in Proposition 2.1. Thus we have rederived the result of Section 2.1 without using Laplace transforms.

2.3.2 Second Order Linear Zero Dynamics

Suppose that $\bar{y} > 0$ and we have two distinct real NMP zeros $\lambda_1 > \lambda_2 > 0$. Without loss of generality, we can take A_0 to be diagonal, that is, $A_0 = diag\{\lambda_1, \lambda_2\}$ and $B_0 = [1,\ 1]^T$. In this case $\bar{z} = -A_0^{-1} B_0 \bar{y} = -\bar{y}\left[\frac{1}{\lambda_1},\ \frac{1}{\lambda_2}\right]^T$. We also use the notation $B_0^\perp = [1,\ -1]^T$ and the definition

$$z_\alpha(T) \triangleq -\alpha \int_0^T e^{A_0(T-\tau)} B_0 d\tau \tag{9}$$

$$= -\alpha A_0^{-1}\left(e^{A_0 T} - I\right) B_0 \tag{10}$$

$$= -\alpha \begin{bmatrix} \frac{e^{\lambda_1 T} - 1}{\lambda_1} \\ \frac{e^{\lambda_2 T} - 1}{\lambda_2} \end{bmatrix}.$$

With this background, we state our first proposition:

Proposition 2.3 *For the second order zero dynamics system defined above, any state,* $z(T)$ *reachable in time* T *with* $y \in \mathcal{Y}_\alpha$, $\alpha > 0$ *must satisfy both*

$$B_0^\perp\left(z(T) - z_\alpha(T)\right) \ge 0 \tag{11}$$

and $$B_0^\perp e^{-A_0 T}\left(z(T) - z_\alpha(T)\right) \le 0. \tag{12}$$

Proof

Firstly we note that using the Cayley Hamilton Theorem (see for example [11, pp167]), $e^{A_0 t} = I\phi_0(t) + A_0\phi_1(t)$ where $\phi_0(t) = \frac{\left(\lambda_1 e^{\lambda_2 t} - \lambda_2 e^{\lambda_1 t}\right)}{(\lambda_1 - \lambda_2)}$ and $\phi_1(t) = \frac{\left(e^{\lambda_1 t} - e^{\lambda_2 t}\right)}{(\lambda_1 - \lambda_2)}$.

We now consider Inequality (11). From the definition of z_α, in Equation (9), and the linear zero dynamics equation (6), we obtain

$$\begin{aligned}
B_0^\perp \left(z(T) - z_\alpha(T)\right) &= B_0^\perp \int_0^T e^{A_0(T-\tau)} B_0 \left(y(\tau) + \alpha\right) d\tau \\
&= \int_0^T B_0^\perp \left(I\phi_0(T-\tau) + A_0\phi_1(T-\tau)\right) B_0 \left(y(\tau) + \alpha\right) d\tau \\
&= B_0^\perp A_0 B_0 \int_0^T \phi_1(T-\tau) \left(y(\tau) + \alpha\right) d\tau \\
&\geq 0
\end{aligned}$$

because $B_0^\perp A_0 B_0$ and $(y(\tau) + \alpha)$ are both non-negative, and $\phi_1(T - \tau) \geq 0$ for $T - \tau \geq 0$.

Next we consider Inequality (12). In this case we note that

$$\begin{aligned}
B_0^\perp e^{-A_0 T} \left(z(T) - z_\alpha(T)\right) &= B_0^\perp e^{-A_0 T} \int_0^T e^{A_0(T-\tau)} B_0 (y(\tau) + \alpha)\, d\tau \\
&= \int_0^T B_0^\perp \left(I\phi_0(-\tau) + A_0\phi_1(-\tau)\right) B_0 (y(\tau) + \alpha)\, d\tau \\
&= B_0^\perp A_0 B_0 \int_0^T \phi_1(-\tau)(y(\tau) + \alpha)\, d\tau \\
&\leq 0
\end{aligned}$$

because both $B_0^\perp A_0 B_0$ and $(y(\tau) + \alpha)$ are non-negative, and $\phi_1(-\tau)$ is non-positive for $\tau > 0$. □

At first sight, there might not appear to be any correspondence between this result and the earlier result based on Laplace transforms (Inequality (3)). However, as we now show, the two results give identical bounds on the undershoot.

Corollary 2.4 *For any plant with two real distinct NMP zeros, Inequality (12) implies Inequality (3).*

Proof

Suppose that $r_{us}(y) = \rho$. Inequality (12) ,together with $z(T) = \bar{z} = -\bar{y}A_0^{-1}B_0$ and Equation (10), gives

$$\begin{aligned}
0 &\geq B_0^\perp e^{-A_0 T} \left(z(T) - z_{\rho\bar{y}}(T)\right) \\
&= B_0^\perp e^{-A_0 T} \left(-\bar{y}A_0^{-1}B_0 + \rho\bar{y}A_0^{-1}\left(e^{A_0 T} - I\right) B_0\right).
\end{aligned}$$

Since $\bar{y} \geq 0$, this is equivalent to

$$\begin{aligned}
0 &\geq B_0^\perp A_0^{-1} e^{-A_0 T} \left(-I + \rho\left(e^{A_0 T} - I\right)\right) B_0 \\
&= -B_0^\perp A_0^{-1} e^{-A_0 T} B_0 + \rho B_0^\perp A_0^{-1}(1 - e^{-A_0 T}) B_0.
\end{aligned}$$

This implies, therefore, that since $B_0^\perp A_0^{-1}\left(I - e^{-A_0 T}\right) B_0 < 0$,

$$
\begin{aligned}
r_{us}(y) &\geq \frac{B_0^\perp A_0^{-1} e^{-A_0 T} B_0}{B_0^\perp A_0^{-1}\left(I - e^{-A_0 T}\right) B_0} \\
&= \frac{\frac{e^{-\lambda_1 T}}{\lambda_1} - \frac{e^{-\lambda_2 T}}{\lambda_2}}{\frac{1-e^{-\lambda_1 T}}{\lambda_1} - \frac{1-e^{-\lambda_2 T}}{\lambda_2}}
\end{aligned}
$$

which is equivalent to Inequality (3). □

We therefore see that in the case of a linear time invariant plant, the zero dynamics form allows the same results as those obtained using Laplace transforms for relating zeros, undershoot and settling time.

3 Performance Limitations for Systems with Unstable Zero Dynamics

In the previous section the zero dynamics form of a linear system was used to derive a bound on the undershoot for a given settling time. This was achieved by finding a relationship between the relative undershoot at the output and the reachable states of the zero dynamics. In this section we will show that these ideas can be extended to nonlinear systems with unstable zero dynamics.

Suppose that a nonlinear system has the form

$$
\begin{aligned}
\dot{\xi} &= F(\xi, z, u), \quad \xi \in \mathbf{R}^l, \\
\dot{z} &= F_0(z, y), \quad z \in \mathbf{R}^m, \\
y &= H(\xi),
\end{aligned}
$$

where u is the input and y is the output. We focus on the zero dynamics equation,

$$\dot{z} = F_0(z, y), \quad z \in \mathbf{R}^m, \tag{13}$$

which represents the internal dynamics of the system. Note that although we assume that Equation (13) represents the full zero dynamics, for the purpose of this paper, it is actually sufficient for this equation to describe part of the internal dynamics.

The following assumptions will be made:

Assumptions

- $\forall\, \bar{y} \in \mathbf{R}$, $\dot{z} = F_0(z, \bar{y})$ has a unique equilibrium point $\bar{z}$ which implies $0 = F_0(\bar{z}, \bar{y})$
- the Jacobian matrix at $\bar{z}$ is nonsingular
- $F_0(0, 0) = 0$

We shall be concerned with the problem of taking the system from rest to the equilibrium at $y(t) = \bar{y} \geq 0$. This is equivalent to finding y which satisfies the following constraints:

$$\lim_{t\to\infty} y(t) = \bar{y} \geq 0, \tag{14}$$

$$\lim_{t\to\infty} z(t) = \bar{z}, \tag{15}$$

where $z(t)$ is the solution to Equation (13) with $z(0) = 0$. If, in addition, $y(t) = \bar{y} \ \forall \ t > T$, then we shall say that y has *finite (exact) settling time.*

Definition 1 (Stability Definitions)
The equilibrium point $\bar{z}$ defined above is *unstable* if it is not (locally) asymptotically stable. It is *anti-stable* if $\dot{z} = -F_0(z, \bar{y})$ is (locally) asymptotically stable. The internal dynamics are unstable (anti-stable) if $\forall \bar{y}$, the corresponding equilibrium point is unstable (anti-stable). If $\bar{z}$ is unstable then the *stable manifold*, $\mathcal{M}_{\bar{z}}$, corresponding to $\bar{z}$ is given by

$$\mathcal{M}_{\bar{z}} = \left\{ z_0 \in \mathbf{R}^m : \dot{z} = F_0(z, \bar{y}) \text{ and } z(0) = z_0 \implies \lim_{t \to \infty} z(t) = \bar{z} \right\}.$$

□

Note that in the case where $\bar{z}$ is anti-stable, $\mathcal{M}_{\bar{z}} = \{\bar{z}\}$. Also, if $\bar{z}$ is globally asymptotically stable, then $\mathcal{M}_{\bar{z}} = \mathbf{R}^m$.

Recall that, for each $\alpha > 0$, $\mathcal{Y}_\alpha$ is the set of functions y which satisfy

$$y(t) \geq -\alpha \quad \forall \ t \geq 0.$$

Definition 2 (Reachability Definitions)
Consider the system described by Equation (13). For each triple (z_0, α, T) the *reachable set*, $\mathcal{R}_{z_0,\alpha,T}$ is the set given by

$$\mathcal{R}_{z_0,\alpha,T} = \{z^* \in \mathbf{R}^m : \ \exists \ y \in \mathcal{Y}_\alpha \text{ s.t. } z(T) = z^*, \ \dot{z} = F_0(z, y), \text{ and } z(0) = z_0\}.$$

Similarly, the set $\mathcal{S}_u \subseteq \mathbf{R}^m$ is *unreachable* if $\mathcal{R}_{z_0,\alpha,T} \subseteq \mathcal{S}_u^c$, where $\mathcal{S}_u^c = \mathbf{R} \setminus \mathcal{S}_u$. □

Definition 3 (Positive Invariance)
Suppose that $\dot{z}$ satisfies Equation (13). The set $\mathcal{S} \subset \mathbf{R}^m$ is *positively invariant* with respect to $\mathcal{Y}_\alpha$ if $z(0) \in \mathcal{S} \implies z(t) \in \mathcal{S} \ \forall \ y \in \mathcal{Y}_\alpha, \ t > 0$. □

Note that if $\mathcal{S}$ is positively invariant for $y \in \mathcal{Y}_\alpha$ then $\mathcal{S}^c$ is unreachable from any $z(0) \in \mathcal{S}$ (for all t).

The reachability definitions given above are quite general. In this paper, we consider only reachability from the origin, and the value of α is usually clear from the context. Hence, we use the simplified notation $\mathcal{R}_T$ for the reachable set at $t = T$.

Remark 1 For any given $\alpha > 0$, $\mathcal{R}_t$ has the additional property that $\mathcal{R}_{t_1} \subseteq \mathcal{R}_{t_2}$ if $t_1 < t_2$. This can be seen by noting that, if $z_1 \in \mathcal{R}_{t_1}$, then z_1 can be reached in time t_2 by letting $y(t) = 0$ for $t < t_2 - t_1$. It follows that a set is reachable at $t = T$ if and only if it is reachable for $t \leq T$. □

If y satisfies constraints (14) and (15), and $\bar{z}$ is unstable, then y must stabilise the internal dynamics by driving z to $\mathcal{M}_{\bar{z}}$. In the special case in which $\bar{z}$ is anti-stable, y must move z to $\bar{z}$. This leads to the following Lemma:

Lemma 3.1 *Consider the system described by Equation* (13). *Suppose that the assumptions in Section 3 are satisfied and that y satisfies constraints* (14) *and* (15). *Then the following statements hold:*

(a) If the open set $\mathcal{S}_u$ is unreachable $\forall \ y \in \mathcal{Y}_0$ and $\bar{z} \in \mathcal{S}_u$, then y must undershoot.

(b) If, for a given $\bar{y}$, $\mathcal{M}_{\bar{z}}$ is unreachable at $t = T$ $\forall$ $y \in \mathcal{Y}_\alpha$, then $r_{us}(y) \geq \alpha/\bar{y}$.

Proof

The proof is immediate, in both cases, by contradiction. □

4 Case Studies

In this section we analyse, in detail, two examples of unstable zero dynamics. In the first example we assume that the zero dynamics are scalar and have a particular structure. The general case and a particular system are both considered. In the second example we study the zero dynamics of a magnetic suspension system. We use the ideas of Section 3 and Lemma 3.1 to show that in each of these examples, the instability of the zero dynamics implies that the output y must undershoot. Bounds on the undershoot for a given exact settling time are also found.

The following definitions will be used in this section:

Let $\mathcal{Y}_{constr}$ be the set of functions y which satisfy constraints (14) and (15). For a given $\bar{y}$, the minimum relative undershoot for a specified exact settling time T can be defined by

$$r^*_{us}(T, \bar{y}) = \inf \{r_{us}(y) : y \in \mathcal{Y}_{constr},\ T_{es}(y) = T\}.$$

Similarly, if $\bar{y}$ and the allowable relative undershoot ρ are specified, then the minimum exact settling time can be defined as

$$T^*_{es}(\rho, \bar{y}) = \inf \{T_{es}(y) : y \in \mathcal{Y}_{constr},\ r_{us}(y) \leq \rho\}.$$

Remark 2 It can be seen that, if one of these two functions is known, then so is the other because

$$r^*_{us}(T^*_{es}(\rho, \bar{y}), \bar{y}) = \rho.$$

□

4.1 Scalar Zero Dynamics

4.1.1 General Case

Suppose the internal dynamics satisfy

$$\dot{z} = F_0(z, y) = f_0(z) + g_0(z)y, \quad z(0) = 0, \tag{16}$$

where

$z \in R$,

$f_0(z)$ is monotonically increasing and continuous,

$f_0(0) = 0$,

and $g_0(z)$ has constant sign $\forall$ z.

Without loss of generality, we take $g_0(z) > 0$. Note that the conditions on f_0 ensure that the system satisfies the assumptions of Section 3.

Suppose that y is required to track a step of height $\bar{y} > 0$. Let the corresponding equilibrium point be $\bar{z}$. $\bar{z} < 0$ because $f(\bar{z}) = -g(\bar{z})\bar{y} < 0$. $\bar{z}$ is also anti-stable because of the monotonicity of f_0. It follows that y must drive z to $\bar{z}$.

For this system the following proposition holds [12]:

Proposition 4.1 *Consider the system above. Suppose that* $y \in \mathcal{Y}_\alpha$ *and let* $z_\alpha(t)$ *be the solution to initial value problem* (16) *with* $y(t) = -\alpha$.

$$z(t) \geq z_\alpha(t).$$

Proof

The proof [12, Prop. 2] is a direct application of the comparison principle. □

Suppose that $y \in \mathcal{Y}_0$. When $\alpha = 0$, $z_\alpha(t) = 0$. Thus from the proposition, $z(t) \geq 0 \;\forall\; t$. But then $\bar{z}$ is unreachable, and so y must undershoot.

We can also quantify the required undershoot for a given exact settling time. In [12], it was shown that if the relative undershoot is $\leq \rho$, then $\bar{z}$ is unreachable for $t \leq T^*_{es}(\rho, \bar{y})$, where

$$T^*_{es}(\rho, \bar{y}) = \int_0^{\bar{z}} \frac{1}{f_0(z) - \rho \bar{y} g_0(z)} \, dz.$$

Thus, for a given settling time T, $r_{us} \geq \rho_T$, where ρ_T is the solution of

$$T^*_{es}(\rho_T, \bar{y}) = T.$$

4.1.2 A Particular Example

Consider the particular example in which $f_0(z) = z^3$ and $g_0(z) = 1$, i.e.

$$\dot{z} = z^3 + y.$$

Then

$$\bar{z} = -\bar{y}^{1/3}$$

and

$$T^*_{es}(\rho, \bar{y}) = \int_0^{\bar{z}} \frac{1}{z^3 - \rho \bar{y}} \, dz.$$

Let $b = \rho^{1/3}$. An expression for T^*_{es} can be derived as follows:

$$\begin{aligned}
T^*_{es}(\rho, \bar{y}) &= \int_0^{\bar{z}} \frac{1}{z^3 + (b\bar{z})^3} \, dz \\
&= \frac{1}{3(b\bar{z})^2} \int_0^{\bar{z}} \frac{1}{z + b\bar{z}} + \frac{2b\bar{z} - z}{z^2 - b\bar{z}z + (b\bar{z})^2} \, dz \\
&= \frac{1}{3(b\bar{z})^2} \left[\ln|z + b\bar{z}| - \frac{1}{2} \ln\left|z^2 - b\bar{z}z + (b\bar{z})^2\right| \right]_0^{\bar{z}} + \frac{1}{2b\bar{z}} \int_0^{\bar{z}} \frac{1}{(z - \frac{b\bar{z}}{2})^2 + \frac{3(b\bar{z})^2}{4}} \, dz \\
&= \frac{1}{6(b\bar{z})^2} \left[\ln \frac{(1+b)^2}{1 - b + b^2} \right] + \frac{1}{\sqrt{3}(b\bar{z})^2} \left[\tan^{-1} \frac{2z - b\bar{z}}{\sqrt{3} b\bar{z}} \right]_0^{\bar{z}} \\
&= \frac{1}{6(\bar{y})^{2/3}} \left[\ln \frac{(1+b)^2}{1 - b + b^2} + 2\sqrt{3} \left(\tan^{-1} \frac{2 - b}{\sqrt{3} b} + \frac{\pi}{6} \right) \right].
\end{aligned}$$

From Remark 2, a plot of $r^*_{us}(T, \bar{y})$ as a function of T may be obtained by plotting ρ against T^*_{es}. Several of these plots are shown in Figure 2. Note that the bound on the relative undershoot increases for fast settling times and smaller $\bar{y}$ (slower zero dynamics). This is qualitatively similar to the linear case where the bound $\frac{1}{e^{\lambda T} - 1}$ is worse for fast settling and slow zero dynamics.

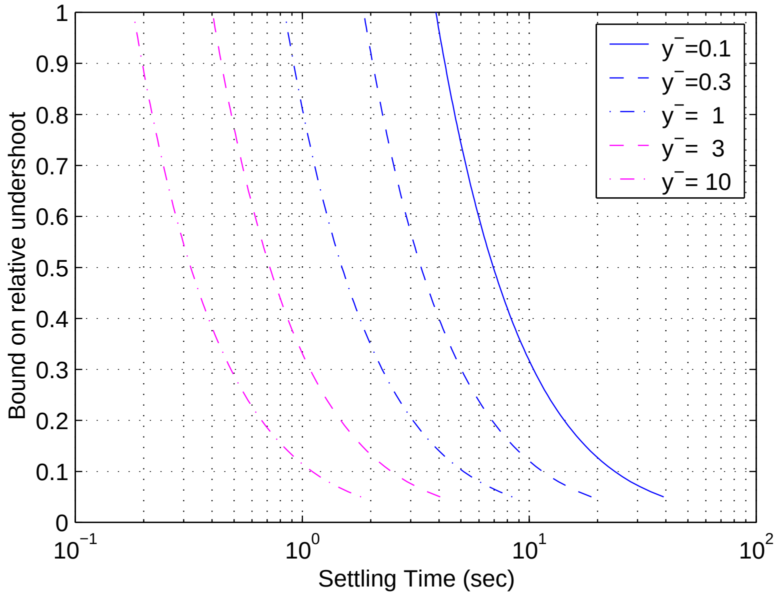

Figure 2: Scalar example - bound on relative undershoot versus settling time for several values of $\bar{y}$.

4.2 Magnetic Suspension System

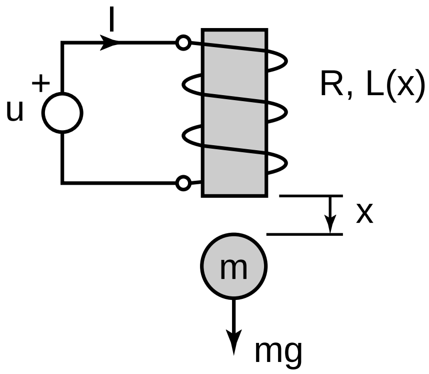

Figure 3: Diagram of magnetic suspension system.

We consider the magnetic suspension system illustrated in Figure 3. If we neglect viscous damping (air resistance), then the model is given by [10], [13]

$$
\begin{aligned}
\dot{x} &= v \\
\dot{v} &= g - \frac{L_0 a}{2m(a+x)^2} I^2 \\
\dot{I} &= \frac{1}{L(x)} \left[-Rx + \frac{L_0 a}{(a+x)^2} v I + u \right],
\end{aligned}
$$

where

$x > 0$ is the vertical distance of the ball below the electromagnet,

$m, \; v$ are the mass and velocity of the ball,

g is the acceleration due to gravity,

R is the resistance of the circuit,

$L_0, \; L_1, \; a$ are positive constants,

$L(x) = L_1 + L_0/(1 + x/a)$,

I is the current through the electromagnet and

u is the applied voltage.

The ball is stationary when the force from the electromagnet balances the gravitational force.

If we make $E = I^2$ the output of the system instead of x, as was done in [13], then the first two equations above are the internal dynamics of the system.

We assume that the system is initially at equilibrium and that $x(0) = x_0$ and $E(0) = E_0$. Note that $E_0 = 2gm(a + x_0)^2/L_0 a$. The change of variables

$$z_1 = x - x_0, \qquad z_2 = v, \qquad \text{and} \quad y = E - E_0$$

yields

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= g - \frac{k_1}{(\gamma + z_1)^2}(y + E_0), \end{aligned} \tag{17}$$

where $z \in \mathcal{D} = \left\{ z \in \mathbf{R}^2 : z_1 > -x_0 \right\}$, $k_1 = L_0 a/2m > 0$ and $\gamma = a + x_0 > 0$. Note that $y(t) \geq -E_0$. This is in the zero dynamics form of Equation (13) and satisfies the assumptions in Section 3. In the discussion that follows, the subscripts 1 and 2 will be used to denote the components of an element of $\mathbf{R}^2$. Thus z is equivalent to $[z_1, z_2]^T$.

Suppose that y is required to track a step of height $\bar{y} > 0$. The equilibrium point, $\bar{z}$, of system (17) with $y(t) = \bar{y}$ is given by

$$\bar{z}_1 = \sqrt{\frac{k_1}{g}(\bar{y} + E_0)} - \gamma, \tag{18}$$

$$\bar{z}_2 = 0. \tag{19}$$

The stability matrix of the linearisation of system (17) at the equilibrium point is given by

$$A_{\bar{z}} = \begin{bmatrix} 0 & 1 \\ \frac{2g}{(\gamma + \bar{z}_1)} & 0 \end{bmatrix}.$$

This has eigenvalues at $\pm\sqrt{\frac{2g}{(\gamma + \bar{z}_1)}}$, and so $\bar{z}$ is saddle point. It follows that the equilibrium is unstable but not anti-stable.

It is useful to note that

$$g\gamma^2 = k_1 E_0 \tag{20}$$

and

$$g(\gamma + \bar{z}_1)^2 = k_1(E_0 + \bar{y}). \tag{21}$$

These two relations hold because $y = 0$ and $y = \bar{y}$ correspond to equilibrium points at $z_1 = 0$ and $z_1 = \bar{z}_1$, respectively.

The ideas of Section 3 can be used to show that y must undershoot. A lower bound on the undershoot for a given exact settling time will also be found. We first prove two propositions.

Proposition 4.2 *For the above system, the stable manifold is given by*

$$\mathcal{M}_{\bar{z}} = \{z \in \mathcal{D} : z_2 = q_1(z_1)\},$$

where $q_1(z_1) = \sqrt{\frac{2g}{\gamma + z_1}}(\bar{z}_1 - z_1)$.

Proof

Since $\bar{z}$ is a saddle point, there will be exactly two trajectories approaching $\bar{z}$ along opposite directions [14], [15]. The trajectories will be unique because the vector field of the system is continuously differentiable on $\mathcal{D}$. In order to find the trajectories the internal dynamics equations can be combined to yield

$$\left[g - \frac{k_1(\bar{y} + E_0)}{(\gamma + z_1)^2}\right]\dot{z}_1 = z_2\dot{z}_2.$$

This can be integrated with respect to t to give

$$gz_1 + \frac{k_1(\bar{y} + E_0)}{(\gamma + z_1)} = \frac{z_2^2}{2} + c, \tag{22}$$

where c is the constant of integration and is determined by the initial condition of the system.

The trajectories which converge to, or diverge from $\bar{z}$ are referred to as separatrices. By substituting $z = \bar{z}$ into the above equation we obtain the value of c for the separatrices as

$$c \;=\; g\bar{z}_1 + \frac{k_1(\bar{y} + E_0)}{(\gamma + \bar{z}_1)}$$

With this value of c Equation (22) becomes

$$\frac{z_2^2}{2} = g(z_1 - \bar{z}_1) + k_1(\bar{y} + E_0)\left[\frac{1}{(\gamma + z_1)} - \frac{1}{(\gamma + \bar{z}_1)}\right].$$

By using relation (21) we get

$$\begin{aligned}\frac{z_2^2}{2} &= g(z_1 - \bar{z}_1) + g(\gamma + \bar{z}_1)\frac{(\bar{z}_1 - z_1)}{(\gamma + z_1)} \\ &= \frac{g}{(\gamma + z_1)}\left[(\gamma + z_1)(z_1 - \bar{z}_1) + (\gamma + \bar{z}_1)(\bar{z}_1 - z_1)\right].\end{aligned}$$

It follows that the separatrices are given by

$$z_2 = \pm q_1(z_1), \quad \text{where} \quad q_1(z_1) = \sqrt{\frac{2g}{\gamma + z_1}}(\bar{z}_1 - z_1).$$

Now suppose that $z_2(t) = q_1(z_1(t))$. Then

$$\operatorname{sgn}(\dot{z}_1(t)) = \operatorname{sgn}(z_2(t)) = -\operatorname{sgn}(z_1(t) - \bar{z}_1),$$

and so, as $t \to \infty$, $z_1(t) \to \bar{z}_1$ and $z_2(t) = q_1(z_1(t)) \to q_1(\bar{z}_1) = \bar{z}_2$. Thus we have shown that

$$\mathcal{M}_{\bar{z}} = \{z \in \mathcal{D} : z_2 = q_1(z_1)\}.$$

□

Proposition 4.3 *Suppose $\alpha \in [0, E_0]$. Let $z_\alpha(t)$ be the solution to (17) with $y(t) = -\alpha$ and $z(0) = 0$. Let $z(t)$ be the solution for another element of $\mathcal{Y}_\alpha$. Then $\tilde{z}(t) = z(t) - z_\alpha(t)$ satisfies*

$$\tilde{z}(t) \in Q_3 = \left\{ \tilde{z} \in \mathbf{R}^2 : \tilde{z}_1 \leq 0, \tilde{z}_2 \leq 0 \right\}.$$

Proof

Suppose that $\tilde{z}(t_0) \in Q_3$ at some time t_0. Then, at $t = t_0$,

$$\dot{\tilde{z}}_1 = \tilde{z}_2 \leq 0$$

and

$$\dot{\tilde{z}}_2 = k_1(E_0 - \alpha) \left[\frac{1}{(\gamma + z_{\alpha 1})^2} - \frac{1}{(\gamma + z_1)^2} \right] - \frac{k_1(y + \alpha)}{(\gamma + z_1)^2}.$$

Since $E_0 - \alpha \geq 0$, and $\tilde{z}_1 \leq 0$ implies that $(\gamma + z_1)^2 - (\gamma + z_{\alpha 1})^2 \leq 0$,

$$\dot{\tilde{z}}_2 \leq -\frac{k_1(y + \alpha)}{(\gamma + z_1)^2} \leq 0.$$

Thus, at every point in Q_3, $\tilde{z}_1$ and $\tilde{z}_2$ are decreasing. Hence Q_3 is a positively invariant set for $\tilde{z}$ and the proposition follows from the fact that $\tilde{z}(0) = 0$.

□

Suppose that $y \in \mathcal{Y}_0$. Then Proposition 4.3 implies that $z(t) \in Q_3$ (because $z_\alpha(t) = 0$ when $\alpha = 0$). However, it is clear from Equations (18) and (19) that $\bar{z} \notin Q_3$. Thus, $\bar{z}$ is unreachable $\forall\, t$. It follows that y must undershoot.

Now suppose that $\alpha \in [0, E_0]$ and $y \in \mathcal{Y}_\alpha$. The derivation of Equation (22) can be used to show that the $z_\alpha(t)$ lies on the curve defined by

$$\begin{aligned} \frac{z_2^2}{2} &= g z_1 + \frac{k_1(E_0 - \alpha)}{(\gamma + z_1)} - \frac{k_1(E_0 - \alpha)}{\gamma} \\ &= g z_1 - \frac{k_1 z_1 (E_0 - \alpha)}{\gamma(\gamma + z_1)}. \end{aligned}$$

Equation (20) can be used to simplify this to

$$\frac{z_2^2}{2} = \frac{z_1(g\gamma z_1 + k_1 \alpha)}{\gamma(\gamma + z_1)}.$$

By noting the direction of the vector field along this curve we get,

$$z_{\alpha 2}(t) = q_2(z_{\alpha 1}(t)), \quad \text{where} \quad q_2(z_1) = \sqrt{\frac{2 z_1 (g\gamma z_1 + k_1 \alpha)}{\gamma(\gamma + z_1)}}.$$

It follows that, along the trajectory, $\dot{z}_{\alpha 1} = q_2(z_{\alpha 1}) \geq 0$.

Recall that the stable manifold is given by $z_2 = q_1(z_1)$. We note that $q_2(z_1)$ increases monotonically and passes through the origin whilst $q_1(z_1)$ decreases monotonically and crosses the z_1 axis at $z_2 = \bar{z}_2 > 0$. The point of intersection of these curves occurs when $q_1(z_1) = q_2(z_1)$. This equation can be solved as follows:

$$\begin{aligned} \frac{g}{\gamma+z_1}(\bar{z}_1-z_1)^2 &= \frac{z_1(g\gamma z_1+k_1\alpha)}{\gamma(\gamma+z_1)} \\ g(z_1^2-2z_1\bar{z}_1+\bar{z}_1^2) &= gz_1^2+\frac{k_1\alpha z_1}{\gamma} \\ z_1 &= \frac{g\gamma\bar{z}_1^2}{k_1\alpha+2g\gamma\bar{z}_1} \\ &=: p(\alpha). \end{aligned}$$

It can be seen that $z_\alpha(t)$ meets $\mathcal{M}_{\bar{z}}$ when $z_{\alpha 1}(t) = p(\alpha)$. Suppose the time at which this occurs is t_α. Since $\dot{z}_{\alpha 1} = q_2(z_{\alpha 1})$,

$$t_\alpha = \frac{1}{\sqrt{2}}\int_0^{p(\alpha)} \left(\frac{\gamma(\gamma+z_1)}{z_1(g\gamma z_1+k_1\alpha)}\right)^{\frac{1}{2}} dz_1.$$

The diagram in Figure 4 shows $z_\alpha(t)$, $\mathcal{M}_{\bar{z}}$ and $p(\alpha)$. The region labelled $\mathcal{S}_{t_1}$ is $\mathcal{Q}_3$ shifted by $z_\alpha(t_1)$ for $t = t_1 < t_\alpha$ and contains $z(t_1)$ $\forall\ y \in \mathcal{Y}_\alpha$. From the monotonic nature of $q_1(z_1)$ and $q_2(z_1)$ (and from the figure), it is clear that $\mathcal{M}_{\bar{z}}$ is unreachable for $t < t_\alpha$. Hence, $T^*_{es}(\rho,\bar{y}) = t_{\rho\bar{y}}$. It follows from Remark 2 that, if y is required to settle at time T, then

$$r_{us}(y) \geq \rho_0,$$

where ρ_0 is the solution to $T^*_{es}(\rho_0,\bar{y}) = T$.

Suppose that $m = 0.01\,\text{kg}$, $k = 0.001\,\text{N/m/s}$, $g = 9.81\,\text{m/s}^2$, $a = 0.05\,\text{m}$, $L_0 = 0.01\,\text{H}$, $L_1 = 0.02\,\text{H}$, $R = 1\,\Omega$ and $x = 0.1\,\text{m}$. Note that this gives $k_1 = 0.025$, $E_0 = 8.829$ and $\gamma = 0.15$. For this particular case, plots of $r^*_{us}(T,\bar{y})$, for $\bar{y} = 1,\ 3,\ 10$, and 100, are given in Figure 5. It can be seen that, as for the scalar example, the bound on the undershoot increases for smaller $\bar{y}$ and faster settling times. We note that in this system there is a physical limit to the achievable settling time because $y(t) + E_0 = I^2 > 0$.

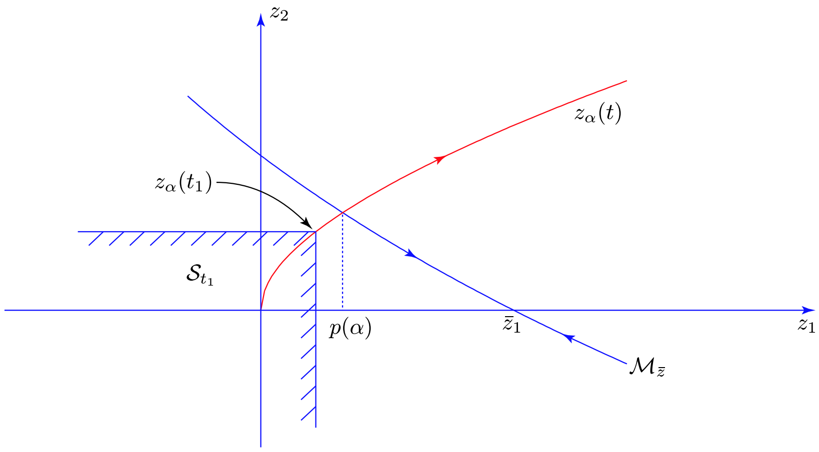

Figure 4: Magnetic suspension system - diagram showing the region ($\mathcal{S}_{t_1}$) which contains $z(t_1)$ $\forall\ y \in \mathcal{Y}_\alpha$.

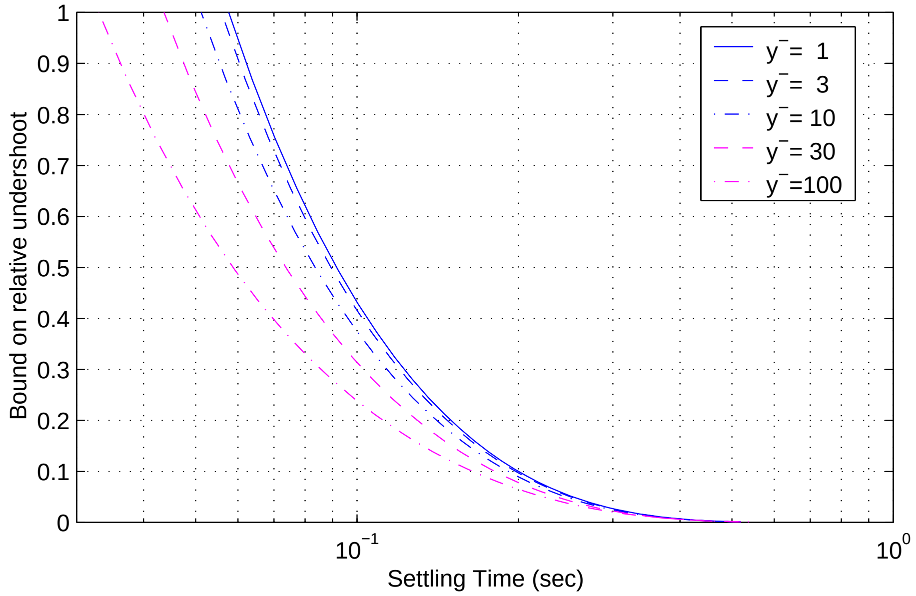

Figure 5: Magnetic suspension system - bound on the relative undershoot versus settling time for several values of $\bar{y}$.

5 Conclusions

'Non-minimum phase' behaviour can be understood in the linear and nonlinear case using the zero-dynamics formulation. In this formulation, the 'constraints' imposed by plant NMP behaviour can be examined. In particular, the permissible output behaviour must drive the state of the zero dynamics onto the stable manifold. Furthermore, in cases where we wish to achieve this in a finite time, a lower bound on the required output deviation is imposed. Several examples illustrate the use of these results. These generalise the linear system conclusions that real NMP zeros, fast settling and small undershoot are incompatible requirements.

References

[1] J. S. Freudenberg and D. P. Looze, *Frequency Domain Properties of Scalar and Multivariable Feedback Systems*. Springer-Verlag, 1988.

[2] M. M. Seron, J. H. Braslavsky, and G. C. Goodwin, *Fundamental Limitations in Filtering and Control*. Springer-Verlag, 1997.

[3] M. M. Seron, J. H. Braslavsky, P. V. Kokotoviċ, and D. Q. Mayne, "Feedback limitations in nonlinear systems: From Bode integrals to cheap control," *IEEE Trans. on Automatic Control*, vol. 44, pp. 829–833, April 1999.

[4] L. Qiu and E. Davison, "Performance limitations of non-minimum phase systems in the servomechanism problem," *Automatica*, vol. 29, no. 2, pp. 337–349, 1993.

[5] R. H. Middleton, "Trade-offs in linear control system design," *Automatica*, vol. 27, no. 2, pp. 281–292, 1991.

[6] A. Isidori, *Nonlinear Control Systems: An Introduction*. Springer-Verlag, 2nd ed., 1989.

[7] J. K. Ward, *Control of the Inverted Pendulum - An Integrated Approach.* PhD thesis, School of EE and CS, The University of Newcastle, 2308, Australia, 2002.

[8] K. L. Moore and S. P. Bhattacharyya, "A technique for choosing zero locations for minimal overshoot," *IEEE Trans. on Automatic Control*, vol. 35, no. 5, pp. 577–80, 1990.

[9] M. E. Halpern, "Rational L_1 suboptimal feedback system design," in *Proc. 14th IFAC World Congress, Vol. F-2d-08-2, Beijing, July 1999.*

[10] H. Khalil, *Nonlinear Systems.* Prentice Hall, second ed., 1996.

[11] T. Kailath, *Linear Systems.* Prentice Hall, 1980.

[12] R. H. Middleton and J. H. Braslavsky, "Towards quantitative time domain design tradeoffs in nonlinear control," in *Proc. 2002 American Control Conference, Anchorage, Alaska, May 2002.*

[13] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, "Cheap control performance of a class of non-right-invertible nonlinear systems," *to appear, IEEE Trans. on Automatic Control.*

[14] D. A. Sanchez, *Ordinary Differential Equations and Stability Theory: An Introduction.* San Francisco and London: W.H. Freeman and Company, 1968.

[15] F. Brauer and J. A. Nohel, *Qualitative Theory of Ordinary Differential Equations.* New York: W.A. Benjamin Inc., 1969.