Chapter 1
Introduction

1.1 ▪ Brief History

The use of high-gain observers in feedback control appeared first in the context of linear feedback as a tool for robust observer design. In their celebrated work on loop transfer recovery [36], Doyle and Stein used high-gain observers to recover, with observers, frequency-domain loop properties achieved by state feedback. The investigation of high-gain observers in the context of robust linear control continued in the 1980s, as seen in the work of Petersen and Hollot [116] on H_∞ control. The use of high-gain observers in nonlinear feedback control started to appear in the late 1980s in the work of Saberi [86, 129], Tornambe [151], and Khalil [37]. Two key papers, published in 1992, represent the beginning of two schools of research on high-gain observers. The work by Gauthier, Hammouri, and Othman [50] started a line of work that is exemplified by [21, 25, 35, 49, 51, 57, 154]. This line of research covered a wide class of nonlinear systems and obtained global results under global growth conditions. The work by Esfandiari and Khalil [39] brought attention to the peaking phenomenon as an important feature of high-gain observers. While this phenomenon was observed earlier in the literature [109, 117], the paper [39] showed that the interaction of peaking with nonlinearities could induce finite escape time. In particular, it showed that, in the lack of global growth conditions, high-gain observers could destabilize the closed-loop system as the observer gain is driven sufficiently high. It proposed a seemingly simple solution for the problem. It suggested that the control should be designed as a globally bounded function of the state estimates so that it saturates during the peaking period. Because the observer is much faster than the closed-loop dynamics under state feedback, the peaking period is very short relative to the time scale of the plant variables, which remain very close to their initial values. Teel and Praly [149, 150] built on the ideas of [39] and earlier work by Tornambe [152] to prove the first nonlinear separation principle and develop a set of tools for semiglobal stabilization of nonlinear systems. Their work drew attention to [39], and soon afterward many leading nonlinear control researchers started using high-gain observers; cf. [5, 8, 26, 54, 62, 65, 67, 71, 72, 89, 94, 98, 99, 102, 103, 118, 119, 123, 128, 133, 138, 139, 141, 158]. These papers have studied a wide range of nonlinear control problems, including stabilization, regulation, tracking, and adaptive control. They also explored the use of time-varying high-gain observers. Khalil and his coworkers continued to investigate high-gain observers in nonlinear feedback control for about twenty-five years converging a wide range of

problems; cf. [1, 2, 4, 32, 46, 73, 74, 76, 83, 104, 113, 134, 135, 136, 140, 153]. Atassi and Khalil [18] proved a separation principle that adds a new dimension to the result of Teel and Praly [149]; namely, the combination of fast observer with control saturation enables the output feedback controller to recover the trajectories of the state feedback controller as the observer gain is made sufficiently high.

1.2 ▪ Motivating Examples

We use a few examples to motivate the design of high-gain observers and discuss their main features. Example 1.1 illustrates the robustness of the observer and its peaking phenomenon. Example 1.2 shows that the interaction of peaking and nonlinearities can lead to a finite escape time. It shows also how saturating the controller overcomes the destabilizing effect of peaking. An important feature of high-gain observers in feedback control is illustrated in this example, where it is shown that the output feedback controller recovers the performance of the state feedback controller as the observer's dynamics become sufficiently fast. This performance recovery property is due to the combined effect of a fast observer and a saturated control. Example 1.3 shows that this property does not hold by designing a fast observer without saturated control. Example 1.4 introduces the reduced-order high-gain observer and compares it with the full-order one.

Example 1.1. Consider the two-dimensional system

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = \phi(x,u), \qquad y = x_1,$$

where $x = \mathrm{col}(x_1, x_2)$, ϕ is locally Lipschitz, and $x(t)$ and $u(t)$ are bounded for all $t \geq 0$. To estimate x, we use the observer

$$\dot{\hat{x}}_1 = \hat{x}_2 + h_1(y - \hat{x}_1), \qquad \dot{\hat{x}}_2 = \phi_0(\hat{x},u) + h_2(y - \hat{x}_1),$$

where $\phi_0(x,u)$ is a nominal model $\phi(x,u)$. We can take $\phi_0 = 0$, which simplifies the observer to a linear one. Whatever the choice of ϕ_0 is, we assume that

$$|\phi_0(z,u) - \phi(x,u)| \leq L\|x - z\| + M$$

for some nonnegative constants L and M, for all (x,z,u) in the domain of interest.[1] In the special case when $\phi_0 = \phi$ and ϕ is Lipschitz in x uniformly in u, the foregoing inequality holds with $M = 0$. The estimation error $\tilde{x} = x - \hat{x}$ satisfies the equation

$$\dot{\tilde{x}} = A_o\tilde{x} + B\delta(x,\tilde{x},u), \qquad \text{where} \qquad A_o = \begin{bmatrix} -h_1 & 1 \\ -h_2 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

and $\delta(x,\tilde{x},u) = \phi(x,u) - \phi_0(\hat{x},u)$. We view this equation as a perturbation of the linear system $\dot{\tilde{x}} = A_o\tilde{x}$. In the absence of δ, asymptotic error convergence is achieved by designing $H = \mathrm{col}(h_1, h_2)$ such that A_o is Hurwitz. In the presence of δ, we need to design H with the additional goal of rejecting the effect of δ on $\tilde{x}$. This is ideally achieved, for any δ, if the transfer function from δ to $\tilde{x}$,

$$G_o(s) = \frac{1}{s^2 + h_1 s + h_2}\begin{bmatrix} 1 \\ s + h_1 \end{bmatrix},$$

[1] Throughout the book, $\|x\| = \sqrt{x^T x}$.

is identically zero. While this is not possible, we can make $\sup_{\omega\in R}||G_o(j\omega)||$ arbitrarily small by choosing $h_2 \gg h_1 \gg 1$. In particular, taking

$$h_1 = \frac{\alpha_1}{\varepsilon}, \qquad h_2 = \frac{\alpha_2}{\varepsilon^2}$$

for some positive constants α_1, α_2, and ε, with $\varepsilon \ll 1$, it can be shown that

$$G_o(s) = \frac{\varepsilon}{(\varepsilon s)^2 + \alpha_1 \varepsilon s + \alpha_2}\begin{bmatrix} \varepsilon \\ \varepsilon s + \alpha_1 \end{bmatrix}.$$

Hence, $\lim_{\varepsilon\to 0} G_o(s) = 0$. This disturbance rejection property of the high-gain observer can be seen in the time domain by scaling the estimation error. Let

$$\eta_1 = \frac{\tilde{x}_1}{\varepsilon}, \qquad \eta_2 = \tilde{x}_2. \tag{1.1}$$

Then

$$\varepsilon\dot{\eta} = F\eta + \varepsilon B\delta, \qquad \text{where} \quad F = \begin{bmatrix} -\alpha_1 & 1 \\ -\alpha_2 & 0 \end{bmatrix}. \tag{1.2}$$

The matrix F is Hurwitz because α_1 and α_2 are positive. The matrices A_o and F/ε are related by the similarity transformation (1.1). Therefore, the eigenvalues of A_o are $1/\varepsilon$ times the eigenvalues of F. From (1.2) and the change of variables (1.1), we can make some observations about the behavior of the estimation error. Using the bound $|\delta| \le L||\tilde{x}|| + M \le L||\eta|| + M$ and the Lyapunov function $V = \eta^T P\eta$, where P is the solution of $PF + F^TP = -I$, we obtain

$$\varepsilon\dot{V} = -\eta^T\eta + 2\varepsilon\eta^T PB\delta \le -||\eta||^2 + 2\varepsilon L||PB||\,||\eta||^2 + 2\varepsilon M||PB||\,||\eta||.$$

For $\varepsilon L||PB|| \le \frac{1}{4}$,

$$\varepsilon\dot{V} \le -\tfrac{1}{2}||\eta||^2 + 2\varepsilon M||PB||\,||\eta||.$$

Therefore (see [80, Theorem 4.5]), $||\eta||$, and consequently $||\tilde{x}||$, is ultimately bounded by εcM for some $c > 0$, and

$$||\eta(t)|| \le \max\left\{be^{-at/\varepsilon}||\eta(0)||,\ \varepsilon cM\right\} \qquad \forall\ t \ge 0$$

for some positive constants a and b. This inequality and the scaling (1.1) show that

$$|\tilde{x}_1| \le \max\left\{be^{-at/\varepsilon}||x(0)||,\ \varepsilon^2 cM\right\}, \qquad |\tilde{x}_2| \le \max\left\{\frac{b}{\varepsilon}e^{-at/\varepsilon}||x(0)||,\ \varepsilon cM\right\}.$$

Hence, $||\tilde{x}(t)||$ approaches the ultimate bound exponentially fast, and the smaller ε, the faster the rate of decay, which shows that for sufficiently small ε the estimation error $\tilde{x}$ will be much faster than x. The ultimate bound can be made arbitrarily small by choosing ε small enough. If $M = 0$, which is the case when $\phi_0 = \phi$, then $\tilde{x}(t)$ converges to zero as t tends to infinity. Notice, however, that whenever $x_1(0) \ne \hat{x}_1(0)$, $\eta_1(0) = O(1/\varepsilon)$.[2] Consequently, the solution of (1.2) will contain a term of the form $(1/\varepsilon)e^{-at/\varepsilon}$ for some $a > 0$. While this exponential mode decays rapidly for small ε, it exhibits an impulsive-like behavior where the transient response peaks to $O(1/\varepsilon)$ values before it decays rapidly toward zero. In fact, the function $(a/\varepsilon)e^{-at/\varepsilon}$

[2] A function $f(\varepsilon)$ is $O(\varepsilon^n)$ if $|f(\varepsilon)| \le k\varepsilon^n$ for some $k > 0$, where n can be positive or negative.

approaches an impulse function as ε tends to zero. This behavior is known as the *peaking phenomenon*. It has a serious impact when the observer is used in feedback control, as we shall see in the next example. We use numerical simulation to illustrate the foregoing observations. Consider the system

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = -x_1 - 2x_2 + a x_1^2 x_2 + b \sin 2t, \qquad y = x_1,$$

with $a = 0.25$ and $b = 0.2$. It can be shown that for all $x(0) \in \Omega = \{1.5x_1^2 + x_1 x_2 + 0.5x_2^2 \leq \sqrt{2}\}$, $x(t)$ is bounded [80, Example 11.1]. We use the high-gain observer

$$\dot{\hat{x}}_1 = \hat{x}_2 + \frac{2}{\varepsilon}(y - \hat{x}_1), \qquad \dot{\hat{x}}_2 = -\hat{x}_1 - 2\hat{x}_2 + \hat{a}\hat{x}_1^2\hat{x}_2 + \hat{b}\sin 2t + \frac{1}{\varepsilon^2}(y - \hat{x}_1),$$

with two different choices of the pair $(\hat{a}, \hat{b})$. When $a = 0.25$ and $b = 0.2$ are known, we take $\hat{a} = 0.25$ and $\hat{b} = 0.2$. This is a case with no model uncertainty and $\phi_0 = \phi$. The other case is when the coefficients a and b are unknown. In this case we take $\hat{a} = \hat{b} = 0$. Figure 1.1 shows simulation results for both cases. Figures 1.1(a) and (b) show the estimation errors $\tilde{x}_1$ and $\tilde{x}_2$ in the no-uncertainty case for different values of ε. The estimation error $\tilde{x}_2$ illustrates the peaking phenomenon. We note that the peaking phenomenon is not present in $\tilde{x}_1$. While peaking is induced by $\tilde{x}_1(0)$, it does not appear in $\tilde{x}_1$ because $\tilde{x}_1 = \varepsilon\eta_1$. Figures 1.1(c) and (d) show $\tilde{x}_2$ for the uncertain model case when $\hat{a} = \hat{b} = 0$. Comparison of Figures 1.1(b) and (c) shows that the presence of uncertainty has very little effect on the performance of the observer when ε is sufficiently small. Figure 1.1(d) demonstrates the fact that the ultimate bound on $\tilde{x}_2$ is $O(\varepsilon)$. △

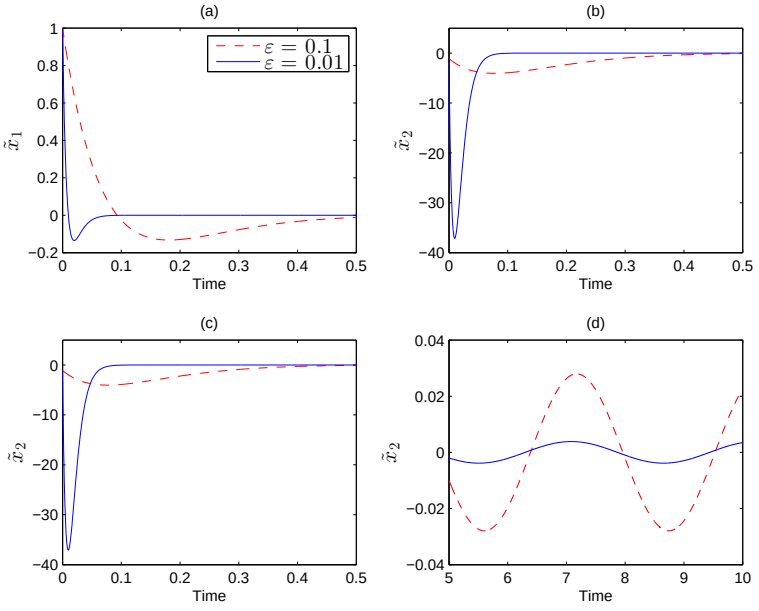

Figure 1.1. *Simulation of Example* 1.1. *Figures* (a) *and* (b) *show the estimation errors* $\tilde{x}_1$ *and* $\tilde{x}_2$ *in the case* $\hat{a} = a$ *and* $\hat{b} = b$. *Figures* (c) *and* (d) *show the transient and steady-state behavior of* $\tilde{x}_2$ *in the case* $\hat{a} = \hat{b} = 0$.

Example 1.2. Consider the two-dimensional system

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = \phi(x,u), \qquad y = x_1,$$

where $x = \mathrm{col}(x_1, x_2)$. Suppose $u = \gamma(x)$ is a locally Lipschitz state feedback controller that stabilizes the origin of the closed-loop system

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = \phi(x, \gamma(x)).$$

To implement this control with output feedback, we use the high-gain observer

$$\dot{\hat{x}}_1 = \hat{x}_2 + (\alpha_1/\varepsilon)(y - \hat{x}_1), \qquad \dot{\hat{x}}_2 = \phi_0(\hat{x}, u) + (\alpha_2/\varepsilon^2)(y - \hat{x}_1),$$

where ϕ_0 is a nominal model of ϕ, and α_1, α_2, and ε are positive constants with $\varepsilon \ll 1$. We saw in the previous example that if

$$|\phi_0(z,u) - \phi(x,u)| \le L\,||x - z|| + M$$

over the domain of interest, then for sufficiently small ε, the estimation errors $\tilde{x}_1 = x_1 - \hat{x}_1$ and $\tilde{x}_2 = x_2 - \hat{x}_2$ satisfy the inequalities

$$|\tilde{x}_1| \le \max\left\{ b e^{-at/\varepsilon} ||\tilde{x}(0)||,\ \varepsilon^2 c M \right\}, \qquad |\tilde{x}_2| \le \left\{ \frac{b}{\varepsilon} e^{-at/\varepsilon} ||\tilde{x}(0)||,\ \varepsilon c M \right\}$$

for some positive constants a, b, c. These inequalities show that reducing ε diminishes the effect of model uncertainty and makes $\tilde{x}$ much faster than x. The bound on $\tilde{x}_2$ demonstrates the peaking phenomenon; namely, $\tilde{x}_2$ might peak to $O(1/\varepsilon)$ values before it decays rapidly toward zero. The peaking phenomenon might destabilize the closed-loop system. This fact is illustrated by simulating the system

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = x_2^3 + u, \qquad y = x_1,$$

which can be globally stabilized by the state feedback controller

$$u = -x_2^3 - x_1 - x_2.$$

The output feedback controller is taken as

$$u = -\hat{x}_2^3 - \hat{x}_1 - \hat{x}_2, \qquad \dot{\hat{x}}_1 = \hat{x}_2 + (2/\varepsilon)(y - \hat{x}_1), \qquad \dot{\hat{x}}_2 = (1/\varepsilon^2)(y - \hat{x}_1),$$

where we take $\phi_0 = 0$. Figure 1.2 shows the performance of the closed-loop system under state and output feedback. Output feedback is simulated for three different values of ε. The initial conditions are $x_1(0) = 0.1$, $x_2(0) = \hat{x}_1(0) = \hat{x}_2(0) = 0$. Peaking is induced by $[x_1(0) - \hat{x}_1(0)]/\varepsilon = 0.1/\varepsilon$ when ε is sufficiently small. Figure 1.2 shows a counterintuitive behavior as ε decreases. Since decreasing ε causes the estimation error to decay faster toward zero, one would expect the response under output feedback to approach the response under state feedback as ε decreases. Figure 1.2 shows the opposite behavior, where the response under output feedback deviates from the response under state feedback as ε decreases. This is the impact of the peaking phenomenon. The same figure shows the control u on a much shorter time interval to exhibit peaking. Figure 1.3 shows that as we decrease ε to 0.004, the system has a finite escape time shortly after $t = 0.07$.

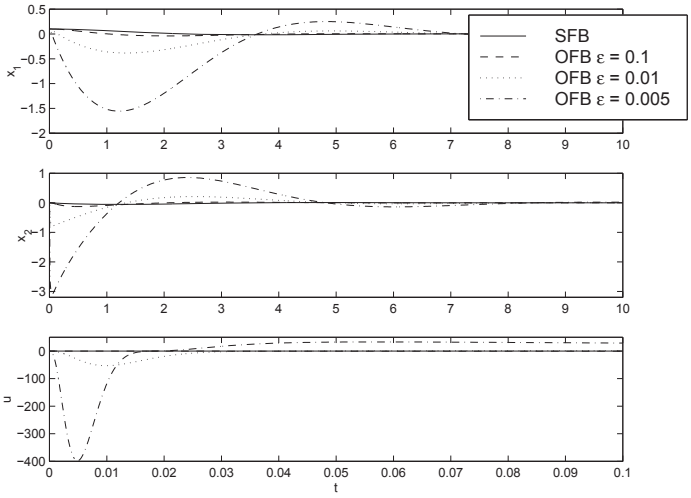

Figure 1.2. *Simulation of Example* 1.2. *Performance under state (SFB) and output (OFB) feedback. Reprinted with permission of Pearson Education, Inc., New York, New York* [80].

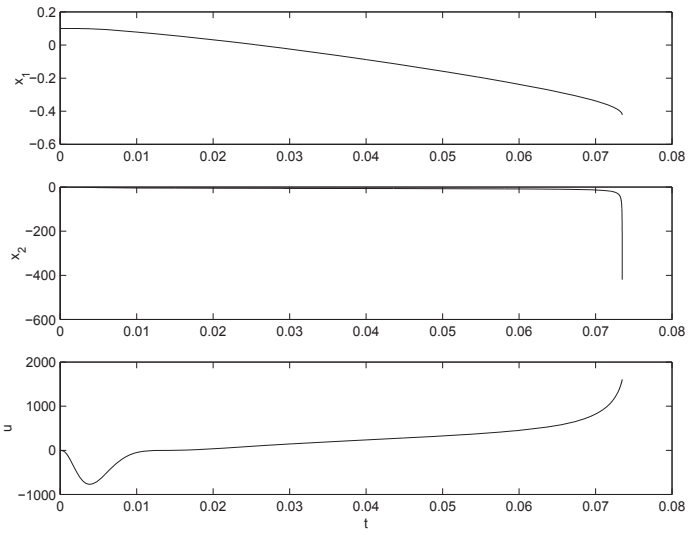

Figure 1.3. *Simulation of Example* 1.2. *Instability induced by peaking at* $\varepsilon = 0.004$. *Reprinted with permission of Pearson Education, Inc., New York, New York* [80].

Fortunately, we can isolate the observer peaking from the plant by saturating the control outside a compact set of interest. Writing the closed-loop system under state feedback as $\dot{x} = Ax$ and solving the Lyapunov equation $PA + A^T P = -I$, we have

$$A = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \qquad \text{and} \qquad P = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

Then $V(x) = x^T P x$ is a Lyapunov function for $\dot{x} = Ax$ and $\dot{V}(x) = -x^T x$. Suppose all initial conditions of interest belong to the set $\Omega = \{V(x) \leq 0.3\}$. Because Ω is positively invariant, $x(t) \in \Omega$ for all $t \geq 0$. By maximizing $|x_1 + x_2|$ and $|x_2|$ over

Ω, it can be shown that for all $x \in \Omega$, $|x_1 + x_2| \leq 0.6$ and $|x_2| \leq 0.6$. Hence, $|u| \leq |x_2|^3 + |x_1 + x_2| \leq 0.816$. Saturating u at ± 1 results in the globally bounded state feedback control

$$u = \text{sat}(-x_2^3 - x_1 - x_2),$$

where $\text{sat}(y) = \min\{|y|, 1\} \text{ sign}(y)$. For all $x(0) \in \Omega$, the saturated control produces the same trajectories as the unsaturated one because for $x \in \Omega$, $|u| < 1$ and the saturation is not active. In output feedback, the state x in the foregoing saturated control is replaced by its estimate $\hat{x}$, namely,

$$u = \text{sat}(-\hat{x}_2^3 - \hat{x}_1 - \hat{x}_2).$$

During the peaking period, the control saturates. Figure 1.4 shows the performance of the closed-loop system under saturated state and output feedback. The control u is shown on a shorter time interval that exhibits control saturation during peaking. The peaking period decreases with ε. The states x_1 and x_2 exhibit the intuitive behavior we expected earlier; namely, the response under output feedback approaches the response under state feedback as ε decreases. Note that we decrease ε to 0.001 beyond the value 0.004 where instability was detected in the unsaturated case. Not only does the system remain stable, but the response under output feedback is almost indistinguishable from the response under state feedback. What is also interesting is that the region of attraction under output feedback approaches the region of attraction under saturated state feedback as ε tends to zero. This is shown in Figures 1.5 and 1.6. The first figure shows the phase portrait of the closed-loop system under $u = \text{sat}(-x_2^3 - x_1 - x_2)$. It has a bounded region of attraction enclosed by a limit cycle. The second figure shows the intersection of the boundary of the region of attraction under $u = \text{sat}(-\hat{x}_2^3 - \hat{x}_1 - \hat{x}_2)$ with the x_1–x_2 plane. It approaches the limit cycle as ε tends to zero.

The behavior we see in Figures 1.4 and 1.6 will be realized with any globally bounded stabilizing function $\gamma(x)$. During the peaking period, the control $\gamma(\hat{x})$ saturates. Since the peaking period shrinks to zero as ε tends to zero, for sufficiently small ε the peaking period becomes so small that the state of the plant x remains close

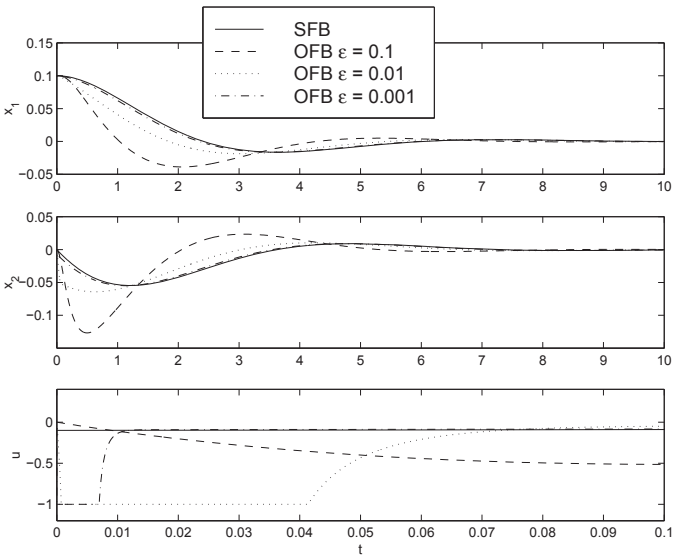

Figure 1.4. *Simulation of Example* 1.2. *Performance under state (SFB) and output (OFB) feedback with saturation. Reprinted with permission of Pearson Education, Inc., New York, New York* [80].

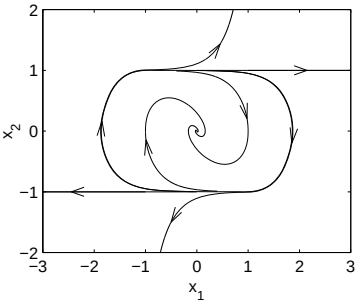

Figure 1.5. *Simulation of Example* 1.2. *Phase portrait of the closed-loop system under* $u = \text{sat}(-x_2^3 - x_1 - x_2)$. *Reprinted with permission of Pearson Education, Inc., New York, New York* [80].

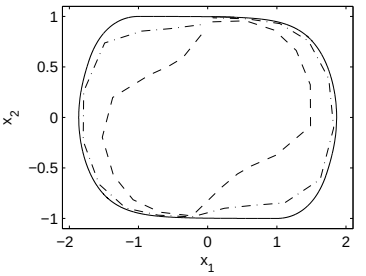

Figure 1.6. *Simulation of Example* 1.2. *Output feedback with* $\varepsilon = 0.08$ *(dashed) and* $\varepsilon = 0.01$ *(dash-dot). Reprinted with permission of Pearson Education, Inc., New York, New York* [80].

to its initial value. After the peaking period, the estimation error becomes $O(\varepsilon)$ and the feedback control $\gamma(\hat{x})$ becomes close to $\gamma(x)$. Consequently, the trajectories of the closed-loop system under output feedback asymptotically approach its trajectories under state feedback as ε tends to zero. This leads to recovery of the performance achieved under state feedback. The global boundedness of $\gamma(x)$ can be always achieved by saturating the state feedback control, or the state estimates, outside a compact set of interest. $\triangle$

The foregoing example shows that the design of the output feedback controller is based on a separation procedure, whereby the state feedback controller is designed as if the whole state was available for feedback, followed by an observer design that is independent of the state feedback design. By choosing ε small enough, the output feedback controller recovers the stability and performance properties of the state feedback controller. This is the essence of the separation principle that is presented in Section 3.1. The separation principle is known in the context of linear systems where the closed-loop eigenvalues under an observer-based controller are the union of the eigenvalues under state feedback and the observer eigenvalues; hence stabilization under output feedback can be achieved by solving separate eigenvalue placement problems for the state feedback and the observer. Over the last two decades there have been several results that present forms of the separation principle for classes of nonlinear systems. It is important to emphasize that the separation principle in the case of high-gain observers has a unique feature that does not exist in other separation-principle results, linear systems included, and that is the recovery of state trajectories by making the observer sufficiently fast. This feature has significant practical implications because it allows the designer to design the state feedback controller to meet transient response specification and/or constraints on the state or control variables. Then, by saturating the state estimate $\hat{x}$ and/or the control u outside compact sets of interest to make $\gamma(\hat{x})$ and $\phi_0(\hat{x}, u)$ globally bounded in $\hat{x}$, the designer can proceed to tune the parameter ε by decreasing it monotonically to bring the trajectories under output feedback close enough to the ones under state feedback. This feature is achieved not only by making the observer fast but also by combining this property with the global boundedness of γ and ϕ_0. We illustrate this point in the example to follow.

Example 1.3. Consider the linear system

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = u, \qquad y = x_1,$$

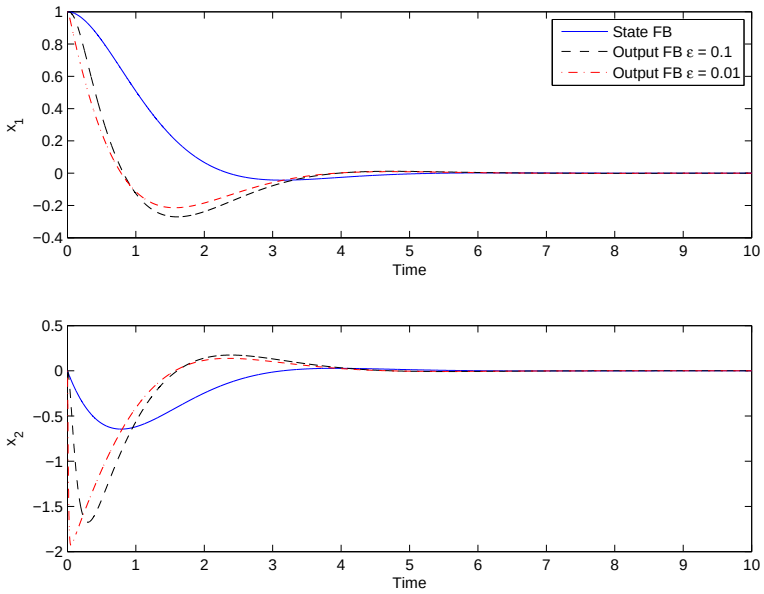

Figure 1.7. *Simulation of Example* 1.3. *The state trajectories under state and output feedback for linear example with unsaturated control. Reprinted with permission from John Wiley and Sons, Ltd* [84].

which is a special case of the system of Example 1.1 with $\phi = u$. A linear state feedback controller that assigns the eigenvalues at $-1 \pm j$ is given by $u = -2x_1 - 2x_2$. A high-gain observer is taken as

$$\dot{\hat{x}}_1 = \hat{x}_2 + (3/\varepsilon)(y - \hat{x}_1), \quad \dot{\hat{x}}_2 = u + (2/\varepsilon^2)(y - \hat{x}_1).$$

It assigns the observer eigenvalues at $-1/\varepsilon$ and $-2/\varepsilon$. The observer-based controller assigns the closed-loop eigenvalues at $-1 \pm j$, $-1/\varepsilon$, and $-2/\varepsilon$. The closed-loop system under output feedback is asymptotically stable for all $\varepsilon > 0$. As we decrease ε, we make the observer dynamics faster than the state dynamics. Will the trajectories of the system under output feedback approach those under state feedback as ε approaches zero? The answer is shown in Figure 1.7, where the state x is shown under state feedback and under output feedback for $\varepsilon = 0.1$ and 0.01. The initial conditions of the simulation are $x_1(0) = 1$ and $x_2(0) = \hat{x}_1(0) = \hat{x}_2(0) = 0$. Contrary to what intuition may suggest, we see that the trajectories under output feedback do not approach the ones under state feedback as ε decreases. In Figure 1.8 we repeat the same simulation when the control is saturated at ± 4, that is, $u = 4\,\mathrm{sat}((-2\hat{x}_1 - 2\hat{x}_2)/4)$. The saturation level 4 is chosen such that $4 > \max_\Omega |-2x_1 - 2x_2|$, where $\Omega = \{1.25x_1^2 + 0.5x_1x_2 + 0.375x_2^2 \leq 1.4\}$ is an estimate of the region of attraction under state feedback control that includes the initial state $(1, 0)$ in its interior. This choice of the saturation level saturates the control outside Ω. Figure 1.8 shows a reversal of the trend we saw in Figure 1.7. Now the trajectories under output feedback approach those under state feedback as ε decreases. This is a manifestation of the performance recovery property of high-gain observers when equipped with a globally bounded control. Figure 1.9 shows the control signal u with and without saturation during the peaking period for $\varepsilon = 0.01$. It demonstrates the role of saturation in suppressing the peaking phenomenon. △

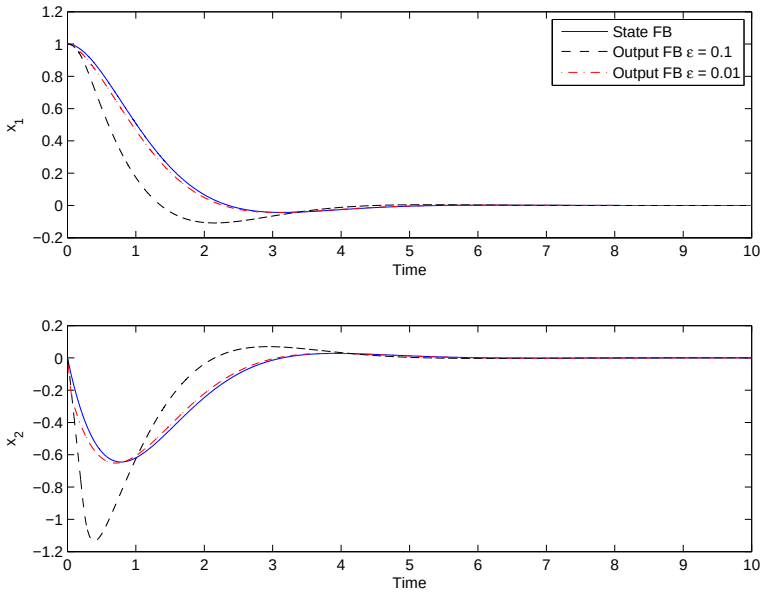

Figure 1.8. *Simulation of Example* 1.3. *The state trajectories under state and output feedback for linear example with saturated control. Reprinted with permission from John Wiley and Sons, Ltd* [84].

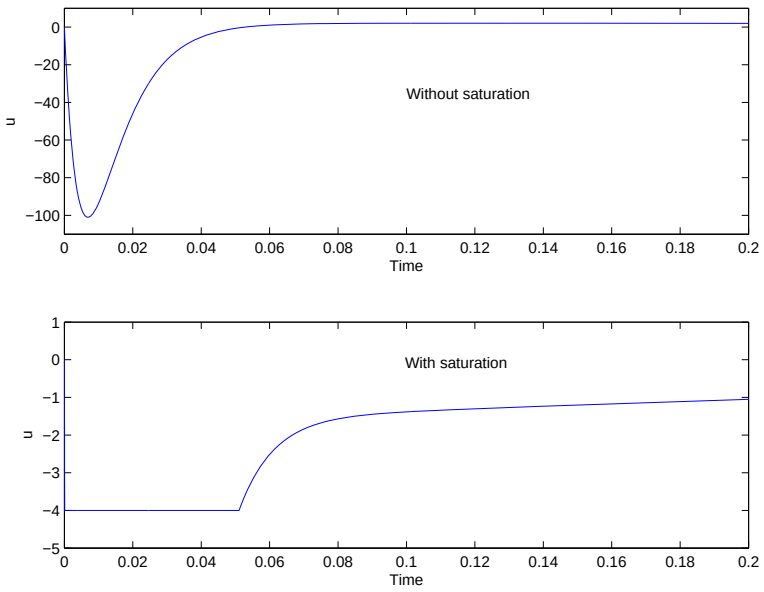

Figure 1.9. *Simulation of Example* 1.3. *The control signal for the linear example with and without control saturation when* $\varepsilon = 0.01$. *Reprinted with permission from John Wiley and Sons, Ltd* [84].

The full-order observer of Example 1.1 estimates both states x_1 and x_2. Since $y = x_1$ is measured, we can design a reduced-order observer of the form

$$\dot{z} = -h(z + hy) + \phi_o(\hat{x}, u), \qquad \hat{x}_2 = z + hy,$$

to estimate $\hat{x}_2$ only. When the observer gain h is chosen as $h = \alpha/\varepsilon$ for some positive constants α and ε with $\varepsilon \ll 1$, it becomes a high-gain observer that has features similar to the full-order high-gain observer. In particular, the estimation error decays rapidly while exhibiting the peaking phenomenon. The example to follow compares the full- and reduced-order observers.

Example 1.4. Consider the van der Pol oscillator

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = -x_1 + (1-x_1^2)x_2,$$

and suppose we want to estimate x_2 from a measurement of x_1. We design a full-order high-gain observer of dimension two and a reduced-order high-gain observer of dimension one. In both cases we do not include a model of the function $x_1+(1-x_1)^2x_2$ so that the observers would be linear. The full-order observer is taken as

$$\dot{\hat{x}}_1 = \hat{x}_2 + (2/\varepsilon)(y-\hat{x}_1), \qquad \dot{\hat{x}}_2 = (1/\varepsilon^2)(y-\hat{x}_1),$$

and its transfer function from y to $\hat{x}_2$ is

$$G_2(s) = \frac{s}{(\varepsilon s+1)^2}.$$

The reduced-order observer is taken as

$$\dot{z}_1 = -(1/\varepsilon)[z_1 + (1/\varepsilon)y], \qquad \hat{x}_2 = z_1 + (1/\varepsilon)y,$$

and its transfer function from y to $\hat{x}_2$ is

$$G_1(s) = \frac{s}{\varepsilon s+1}.$$

Comparison of the two transfer functions shows that the magnitude of the frequency response of $G_2(s)$ rolls off beyond the cutoff frequency of $1/\varepsilon$, while that of $G_1(s)$ remains constant beyond the cutoff frequency. If the measurement of x_1 is corrupted by noise, that is, $y = x_1 + v$, the full-order observer will have better attenuation of high-frequency noise. For example, if $v(t) = N\sin\omega t$ and $\omega = k/\varepsilon$ with $k \gg 1$, the magnitude of the steady-state component of $\hat{x}_2$ due to noise will be $(kN/\varepsilon)/(1+k^2) \approx N/(\varepsilon k)$ for the full-order observer and $(kN/\varepsilon)/\sqrt{(1+k^2)} \approx N/\varepsilon$ for the reduced-order observer. The high-frequency noise is attenuated much better by the full-order observer. The same observation can be seen from simulation. Figure 1.10 shows the estimation error $\tilde{x}_2 = x_2 - \hat{x}_2$ for the two observers. The simulation is run with $\varepsilon = 0.001$ and initial conditions $x_1(0) = x_2(0) = 1$, $\hat{x}_1(0) = \hat{x}_2(0) = z_1(0) = 0$. The measurement noise is a uniformly distributed random signal with values between ± 0.0001, generated by the Simulink icon "Uniform Random Number" with sample time 0.00001. The estimation error is plotted over the time period $[0, 0.01]$ to show the peaking behavior of $\hat{x}_2$ and over the period $[9.9, 10]$ to show the steady-state behavior where the effect of noise is significant. In both cases $\tilde{x}_2$ peaks to $O(1/\varepsilon)$ values, but peaking is larger in the reduced-order observer due to the fact that $\hat{x}_2$ has a component proportional to y/ε. What is significant in this simulation is the effect of measurement noise. Comparison of Figures 1.1(b) and (d) shows that the full-order observer attenuates the effect of measurement noise better by an order of magnitude. Moreover, it filters out the high-frequency content of the noise. $\triangle$

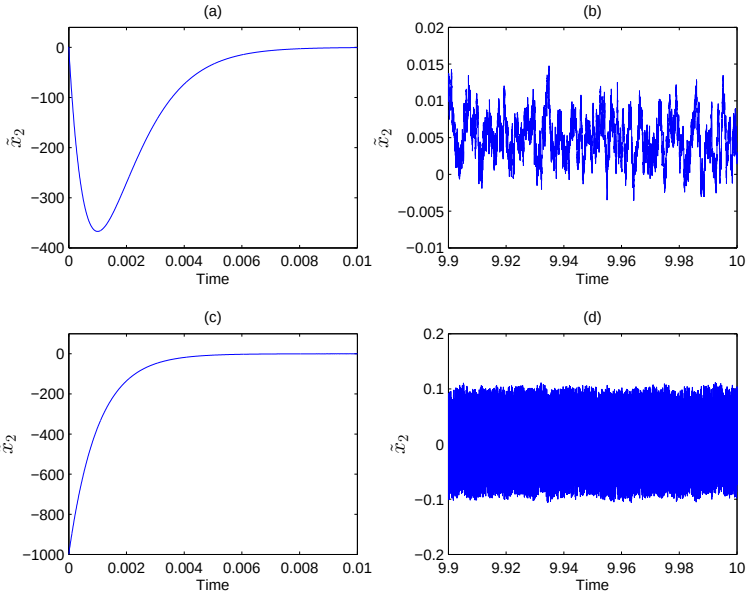

Figure 1.10. *Simulation of Example* 1.4. *Figures* (a) *and* (b) *are for the full-order observer. Figures* (c) *and* (d) *are for the reduced-order observer. All figures show the estimation error* $\tilde{x}_2 = x_2 - \hat{x}_2$.

1.3 ▪ Challenges

1.3.1 ▪ High-Dimensional Observers

A high-gain observer for the system

$$\begin{aligned} \dot{x}_i &= x_{i+1} \qquad \text{for } 1 \le i \le \rho - 1, \\ \dot{x}_\rho &= \phi(x, u), \\ y &= x_1 \end{aligned}$$

is given by

$$\begin{aligned} \dot{\hat{x}}_i &= \hat{x}_{i+1} + \frac{\alpha_i}{\varepsilon^i}(y - \hat{x}_1) \quad \text{for } 1 \le i \le \rho - 1, \\ \dot{\hat{x}}_\rho &= \phi_0(\hat{x}, u) + \frac{\alpha_\rho}{\varepsilon^\rho}(y - \hat{x}_1), \end{aligned}$$

where ϕ_0 is a nominal model of ϕ, ε is a sufficiently small positive constant, and α_1 to α_ρ are chosen such that the polynomial

$$s^\rho + \alpha_1 s^{\rho-1} + \cdots + \alpha_{\rho-1} s + \alpha_\rho$$

is Hurwitz; that is, its roots have negative real parts. This observer faces a numerical challenge if its dimension, ρ, is high. The observer gains are proportional to powers of $1/\varepsilon$, with $1/\varepsilon^\rho$ as the highest power. During the transient period, the internal states of the observer could peak to large values, which are proportional to powers of $1/\varepsilon$, with $1/\varepsilon^{\rho-1}$ as the highest power. These features pose a challenge in the numerical implementation of the observer when ρ is high because in digital implementation

both parameters and signals have to be represented within the finite word length of the digital system. It is worthwhile to note that saturating the state estimates or the control signal before applying the control to the plant does not prevent peaking in the observer's internal variables. To address this numerical challenge, modified high-gain observers are presented in Chapter 7.

1.3.2 ▪ Measurement Noise

The most serious challenge to the application of high-gain observers is measurement noise. This is not surprising because the observer estimates the derivatives of the output. When the output is corrupted by measurement noise, it is expected that noise will have a serious effect on the accuracy of the estimates. The following example explores the effect of measurement noise on a second-order observer.

Example 1.5. Reconsider the system of Example 1.1 and suppose the measurement y is corrupted by bounded noise v; that is,

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = \phi(x,u), \qquad y = x_1 + v,$$

where $|v(t)| \le N$. Equation (1.2) and the inequality satisfied by $\dot{V}$ change to

$$\varepsilon\dot{\eta} = F\eta + \varepsilon B\delta - \frac{1}{\varepsilon}Ev, \qquad \text{where} \qquad E = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix},$$

and

$$\varepsilon\dot{V} \le -\tfrac{1}{2}||\eta||^2 + 2\varepsilon M||PB||\,||\eta|| + \frac{2N}{\varepsilon}||PE||\,||\eta||.$$

Therefore, the ultimate bound on $||\tilde{x}||$ takes the form

$$||\tilde{x}|| \le c_1 M\varepsilon + \frac{c_2 N}{\varepsilon} \tag{1.3}$$

for some positive constants c_1 and c_2. This inequality shows a trade-off between model uncertainty and measurement noise. An illustration of the ultimate bound in Figure 1.11 shows that decreasing ε reduces the ultimate bound until we reach the value $\varepsilon_1 = \sqrt{c_2 N/(c_1 M)}$. Reducing ε beyond this point increases the ultimate bound.

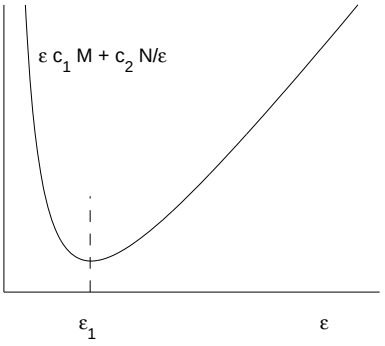

Figure 1.11. *Illustration of the ultimate bound* (1.3).

Thus, the presence of measurement noise puts a lower bound on ε. Another trade-off exists between the speed of convergence of the observer and the ultimate bound on the estimation error. For the separation properties we saw in Example 1.2, we need to choose ε sufficiently small to make the observer dynamics sufficiently faster than the dynamics of the closed-loop system under state feedback. A lower bound on ε limits the speed of convergence. For the high-gain observer to be effective, the ratio N/M should be relatively small so that ε can be chosen to attenuate the effect of uncertainty and make the observer sufficiently fast. Even if there was no model uncertainty, that is, $M = 0$, we still need N to be relatively small so that we can design the observer to be sufficiently fast without bringing the ultimate bound on the estimation error to unacceptable levels. It is worthwhile to note that the bound (1.3) does not take into consideration the low-pass filtering characteristics of the observer at frequencies higher than $1/\varepsilon$, as we saw in Example 1.4. $\triangle$

Ideas to reduce the effect of measurement noise are presented in Chapter 8.

1.4 ▪ Overview of the Book

Chapter 2 starts by describing the class of nonlinear systems for which high-gain observers are designed. The observer design is then presented in terms of a small positive parameter ε, which can be thought of as the observer's time constant or the reciprocal of the observer's bandwidth when the observer is linear. The observer gains are proportional to negative powers of ε. Theorem 2.1 gives an upper bound on the estimation error, which illustrates three key features of the observer: the peaking phenomenon, the fast decay of the error, and the small ultimate bound on the error. The main step in the observer design is a pole placement problem where the eigenvalues of a matrix are assigned in the left-half plane. This step can be carried out by solving Lyapunov or Riccati equations, as shown in Section 2.3. An interesting observation about the Lyapunov-equation design is that it assigns all the eigenvalues of the observer at $-1/\varepsilon$. For a system with a chain of ρ integrators, the dimension of the observer is ρ. Since the first state of the chain is the output, it is possible to design an observer of dimension $\rho-1$. This reduced-order observer is described in Section 2.4. The chapter concludes by presenting the observer design for a special class of multi-output systems, which covers important physical problems where the measured variables are positions and the estimated variables are velocities, accelerations, etc. References are given for more general classes of multi-output systems.

Chapter 3 deals with stabilization and tracking problems. Theorem 3.1 is the separation principle for the stabilization of time-invariant systems. Its proof is a must-read for anyone who wants to understand how high-gain observers work in feedback control. Elements of that proof are repeated in several proofs throughout the book. While the separation principle is presented only for time-invariant systems, extensions to time-varying systems are referenced. For minimum-phase systems, robust stabilization and tracking problems are presented in Sections 3.2 and 3.3, respectively. In both cases the design of output feedback control follows a separation approach where state feedback control is design first; then its performance is recovered by a high-gain observer. A key idea of Section 3.2 is the use of a high-gain observer to reduce a relative degree-higher-than-one system to a relative-degree-one system, for which the control design is straightforward.

Chapter 4 considers adaptive control of nonlinear systems modeled globally by an nth-order differential equation. The dynamics of the system are extended by adding

integrators at the input side, which results in a state model in which the state variables are either states of the added integrators or derivatives of the output. By using a high-gain observer to estimate the derivatives of the output, the output feedback controller can recover the performance of the state feedback controller. A Lyapunov-based state feedback adaptive control is designed in Section 4.2, and its output feedback implementation is given in Section 4.3. Convergence of the tracking error to zero is shown without additional conditions, but convergence of the parameter errors to zero is shown under a persistence of excitation condition. After illustrating the performance of the adaptive controller and the persistence of excitation condition by examples in Section 4.4, robustness to bounded disturbance is shown in Section 4.5. It is shown that the ultimate bound on the tracking error can be made arbitrarily small by adding a robust control component. The robustness properties of the adaptive controller allows its application to nonlinear systems where a nonlinear function is approximated by a function approximator, such as neural network, with a small bounded residual error; this is the subject of Section 4.6.

The nonlinear regulation problem is the subject of Chapter 5. The controller includes a servocompensator that implements an internal model of the exogenous (reference and disturbance) signals. The exogenous signals are generated by a linear exosystem that has simple eigenvalues on the imaginary axis. The problem is treated first for the special case of constant exogenous signals in Sections 5.1 and 5.2; then the more general cases is treated in the rest of the chapter. In addition to the usual tool of using a high-gain observer to recover the performance of state feedback control, the chapter has a number of results that are unique to the regulation problem. First, it shows that for nonlinear systems the internal model is not simply a model of the exosystem; it has to include harmonics of the sinusoidal signals, which are induced by nonlinearities. Second, it deals with impact the servocompensator has on the transient response of the system by introducing the conditional servocompensator, which ensures that the transient response of the system under output feedback can be made arbitrarily close to the transient response of a state feedback sliding mode controller that has no servocompensator. Third, the chapter deals with the case when the internal model is uncertain or unknown. The effect of uncertainty is studied in Section 5.5, and adaptation is used in Section 5.6 to estimate the unknown model.

Chapter 6 presents the extended high-gain observer. It shows two uses of the extended observer, one as a disturbance estimator and the other as a soft sensor of the internal dynamics. As a disturbance estimator, it is shown in Section 6.1 how to implement feedback linearization by estimating the uncertain terms and compensating for them. This is shown for single-input–single-output systems, then for multi-input-multi-output systems. It is also shown how to combine the extended high-gain observer with dynamic inversion to deal with nonaffine control or uncertain control coefficient. Most of the high-gain observer results in Chapters 3 to 5 apply only to minimum-phase systems. This is mainly because the high-gain observer does not estimate the states of the internal dynamics. In Section 6.3 it is shown that the extended high-gain observer can provide information about the internal dynamics by sensing a term in the external dynamics that can be viewed as an output for the internal dynamics. This is used for stabilization of a nonminimum phase system or for designing an observer that estimates the full state of the system. In the later case, two observers are used: the extended high-gain observer is used for the external dynamics, and an extended Kalman filter is used for the internal dynamics.

Chapters 7 and 8 address the challenges described in Section 1.3. The challenge with a ρ-dimensional observer when ρ is high is that the observer gain is proportional

to $1/\varepsilon^{\rho}$ and the peak of its internal states during the transient period is the order of $1/\varepsilon^{\rho-1}$. Chapter 7 addresses this challenge by introducing observers where the gain and the peaking signal are limited to the order of $1/\varepsilon$. Such observers are constructed as cascade connections of first- or second-order observers where the gains are of the order of $1/\varepsilon$. Peaking is limited by inserting saturation functions in between the cascaded observers. Two such observers are designed. The first observer is a simple cascade connection of one second-order observer and $\rho-2$ first-order observers. This observer is robust to model uncertainty, and its performance in feedback control is comparable to the standard high-gain observer, but the steady-state estimation errors are orders of magnitude higher than the standard observer's errors. Moreover, the estimation error does not converge to zero under conditions where the standard observer's error converges to zero. The second observer is a cascade connection of $\rho-1$ second-order observers with feedback injection. This observer has the same steady-state and convergence properties of the standard observer.

Chapter 8 starts by characterizing the effect of measurement noise on the estimation error. The general result for bounded noise shows that the ultimate bound on the estimation error is of the order of $N/\varepsilon^{\rho-1}$, where N is the bound on the noise and ρ is the dimension of the observer. While the rest of the chapter is concerned with this case, Section 8.1 shows less conservative bounds when the frequency of the noise is much lower or much higher than $1/\varepsilon$. The effect of noise on feedback control is studied in Section 8.2, where it is shown that a result similar to the separation principle of Section 3.1 can be shown if the amplitude of the noise is restricted. Even in this case, the presence of noise puts a lower bound on ε. In the tracking problem of Section 8.3 it is shown that the component of the tracking error due to measurement noise does not depend on a negative power of ε. Its ultimate bound is of the order of N. In Section 8.4 two techniques are discussed for reducing the effect of measurement noise. In the first technique, the high-frequency content of the noise is filtered out before feeding the measurement into the observer. The second technique uses a nonlinear gain to adjust ε so that a smaller ε is used during the transient period to achieve fast convergence and a larger one is used at steady state to reduce the effect of noise.

Digital implementation of high-gain observers is discussed in Chapter 9. The observer is discretized with a sampling period proportional to ε. The nonlinear observer is discretized using the Forward Difference method, while other discretization methods can be used with linear observers. Digital control with zero-order hold is shown to have properties similar to continuous-time controllers when ε and the sampling period are sufficiently small. Finally, a multirate digital control scheme is presented. This scheme is useful for computationally demanding controllers where the control sampling period cannot be reduced beyond a certain value, which might not be small enough to implement the fast high-gain observer. The multirate scheme allows the observer to run with a sampling period shorter than the controller's sampling period.