University of California
Department of Mechanical Engineering

ECE230A/ME243A LINEAR SYSTEMS

Fall 1999 (B. Bamieh)

Lecture 2: Linearization; Local Theory of Second Order Systems [1]

2.1 Linearization

Much of our attention in this course will be focused on *linear* models. Linear models frequently arise as descriptions of small perturbations away from a nominal solution of the system. Consider, for example, the continuous-time (CT) state-space model

$$
\begin{aligned}
\dot{x}(t) &= f(x(t), u(t), t) \\
y(t) &= g(x(t), u(t), t)
\end{aligned}
\tag{2.1}
$$

where $x(t)$ is the n-dimensional state-vector at time t, $u(t)$ is the m-dimensional vector of inputs, and $y(t)$ is the p-dimensional vector of outputs. Suppose $x_o(t)$, $u_o(t)$ and $y_o(t)$ constitute a *nominal solution* of the system, i.e. a collection of continuous-time (CT) signals that jointly satisfy the equations in (2.1). Now let the control and initial condition be perturbed from their nominal values to $u(t) = u_o(t) + \delta u(t)$ and $x(0) = x_o(0) + \delta x(0)$ respectively, and let the state trajectory accordingly be perturbed to $x(t) = x_o(t) + \delta x(t)$. Substituting these new values into (2.1) and performing a (multivariable) Taylor series expansion to first-order terms, we find

$$
\begin{aligned}
\dot{\delta x}(t) &\approx \left[\frac{\partial f}{\partial x}\right]_o \delta x(t) + \left[\frac{\partial f}{\partial u}\right]_o \delta u(t) \\
\delta y(t) &\approx \left[\frac{\partial g}{\partial x}\right]_o \delta x(t) + \left[\frac{\partial g}{\partial u}\right]_o \delta u(t)
\end{aligned}
\tag{2.2}
$$

where the $n \times n$ matrix $[\partial f/\partial x]_o$ denotes the Jacobian of $f(.,.,.)$ with respect to x, i.e. a matrix whose ij-th entry is the partial derivative of the ith component of $f(.,.,.)$ with respect to the jth component of x, and where the other Jacobian matrices in (2.2) are similarly defined. The subscript $_o$ indicates that the Jacobians are evaluated along the nominal trajectory, i.e. at $x(t) = x_o(t)$ and $u(t) = u_o(t)$. The linearized model (2.2) is evidently linear, of the form

$$
\begin{aligned}
\dot{\delta x}(t) &= A(t)\,\delta x(t) + B(t)\,\delta u(t) \\
\delta y(t) &= C(t)\,\delta x(t) + D(t)\,\delta u(t).
\end{aligned}
\tag{2.3}
$$

When the original nonlinear model is time-invariant, the linearized model will also be time-invariant *if the nominal solution is constant* (i.e. if the nominal solution corresponds to a constant *equilibrium*); however, the linearized model *may be time varying* if the nominal solution is time varying (even if the original nonlinear model is time-invariant), and will be *periodic* — i.e., have periodically varying coefficients — if the nominal solution is periodic

---

[1]These notes are part of a set under development by Mohammed Dahleh (UC Santa Barbara), Munther Dahleh (MIT), and George Verghese (MIT). Comments and corrections would be greatly appreciated. Please do not copy for or distribute to people not enrolled in ECE230A/ME243A without permission from one of the coauthors.

(as happens when the nominal solution corresponds to operation in some cyclic or *periodic steady state*).

The same development can be carried out for discrete-time (DT) systems, but we focus in this lecture on the CT case.

Example 2.1.1 (Linearizing a Nonlinear Circuit Model)

Consider linearizing the state-space model we obtained for the nonlinear circuit in the previous Lecture. We ended up there with a nonlinear model of the form

$$
\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{C_1}\left(\frac{x_2-x_1}{R} - \mathcal{N}(x_1)\right) \\ \frac{1}{C_2}\left(x_3 - \frac{x_2-x_1}{R}\right) \\ -\frac{1}{L}x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L}v \end{bmatrix} \quad . \tag{2.4}
$$

For the linearization, all that happens is each x_j is replaced by δx_j, and $\mathcal{N}(x_1)$ is replaced by $[d\mathcal{N}(x_1)/dx_1]_o\, \delta x_1$, resulting in a linear state-space model of the form

$$
\delta\dot{x}(t) = A\,\delta x(t) + B\,\delta v(t) \tag{2.5}
$$

with

$$
A = \begin{pmatrix} -\frac{1}{RC_1} - \left[\frac{d\mathcal{N}}{dx_1}\right]_o & \frac{1}{RC_1} & 0 \\ \frac{1}{RC_2} & -\frac{1}{RC_2} & \frac{1}{C_3} \\ 0 & \frac{1}{L} & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{L} \end{pmatrix} \tag{2.6}
$$

Example 2.1.2 (Linearizing the Inverted Pendulum)

Recall from the example in the previous Lecture the equations that describe the dynamics of the inverted pendulum. Those equations are nonlinear due to the presence of the terms $\sin(\theta)$, $\cos(\theta)$, and $(\dot{\theta})^2$. We can linearize these equations around $\theta = 0$ and $\dot{\theta} = 0$, by assuming that $\theta(t)$ and $\dot{\theta}(t)$ remain small. Recall that for small θ

$$
\sin(\theta) \approx \theta - \frac{1}{6}\theta^3
$$

$$
\cos(\theta) \approx 1 - \frac{1}{2}\theta^2,
$$

and using these relations the linearized system of equations takes the form

$$
\left(1 - \frac{ml}{M_tL}\right)\ddot{s} + \frac{ml}{M_tL}\frac{g}{L}\theta = \frac{1}{M_t}u,
$$

$$
\left(1 - \frac{ml}{M_tL}\right)\ddot{\theta} - \frac{g}{L}\theta = -\frac{1}{M_tL}u \ .
$$

Using as state vector

$$
x = \begin{bmatrix} s \\ \dot{s} \\ \theta \\ \dot{\theta} \end{bmatrix},
$$

the following state-space model can be easily obtained:

$$
\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -\alpha\frac{ml}{M_tL}g & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \alpha\frac{g}{L} & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{\alpha}{M_t} \\ 0 \\ -\frac{\alpha}{LM_t} \end{pmatrix}u
$$

$$
y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}x,
$$

2

where the constant α is given by

$$\alpha = \frac{1}{\left(1 - \frac{ml}{M_t L}\right)}.$$

2.2   Phase Portraits of Second-Order Systems

To develop some intuition about the behavior of dynamic models, we will study the zero-input or undriven response of second-order time-invariant systems in some detail. This response is best described in terms of *phase portraits*. Given a second-order system described by the undriven state-space model

$$
\begin{aligned}
\dot{x}_1(t) &= f(x_1(t), x_2(t)) & (2.7) \\
\dot{x}_2(t) &= f(x_1(t), x_2(t)) \,, & (2.8)
\end{aligned}
$$

the phase portrait is a two-dimensional plot — for a variety of initial conditions — of system trajectories in the state space. The state-space for the above second-order example is a plane with the x_1 and x_2 components of the state vector represented on the two axes. Each trajectory in this example is a curve corresponding to the solution $x(t) = [x_1(t), x_2(t)]^T$ that satisfies the state equations above for a specified initial condition x_0. The temporal information for a trajectory may either be specified or omitted on the phase portrait.

Obtaining the phase portrait of a linear system is straightforward, as we shall demonstrate first. For a nonlinear system, one can simulate the system for various initial conditions and obtain such a plot. This numerical approach may result in misleading or incorrect conclusions, however. A powerful approach for understanding the nonlinear behavior is via linearization around distinguished nominal trajectories, such as equilibrium points. We shall demonstrate through examples that in many situations the linearized model gives full local information about the phase portrait of the nonlinear model, although it is inconclusive in other situations.

2.2.1   Linearization Revisited

The local theory of nonlinear systems is based on the linearization of the equations describing the system about an equilibrium point.

Definition 2.1 *(Equilibrium Point) An* **equilibrium point** *of a general undriven CT state-space system*

$$\dot{x}(t) = f(x(t), 0) \qquad (2.9)$$

is a point x_o in the state space such that $f(x_o, 0) = 0$, i.e., a point that the system will remain at, if started there. If in addition, there exists a $\delta > 0$ such that x_o is the only equilibrium point in the region $\|x - x_o\| < \delta$, then x_o is referred to as an isolated *equilibrium point.*

(In the DT case, an equilibrium point of the undriven system $x(k+1) = f(x(k), 0)$ is defined by the condition $x_o = f(x_o, 0)$, which again specifies a point that the system will remain at, if started there.)

In what follows we will only be considering isolated equilibrium points of unforced systems. Now suppose the second-order system given in equations (2.7, 2.8) has an equilibrium point at x_o. Let

$$
[\partial f/\partial x]_o = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\[2mm] \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} (x_o) = A
$$

be the Jacobian of f evaluated at x_o. Then the linearized system may be derived using Taylor series expansion and ignoring the high order terms. The linearized system can then be written as:

$$\dot{\delta x} = A\delta x. \tag{2.10}$$

where $\delta x = x - x_o$. If x is close to x_o, this approximation would appear to provide useful information about the nonlinear system. This, however, is not always the case, as shown in the following simple example.

Example 2.2.1   Consider the one-dimensional nonlinear system:

$$\dot{x} = -x^3.$$

It is clear that all trajectories of this system converge to the origin, irrespective of the initial condition. The linearized system around the origin is given by:

$$\dot{z} = 0.$$

The behavior of this system is drastically different from the previous one.

2.2.2   Classifying Equilibrium Points

Assume we linearize a second-order system about an equilibrium point x_o. Then x_o may be classified by the general behavior of the trajectories in the phase plane of the linearized system. This behavior is dictated by the eigenvalues, λ_1 and λ_2, of the A matrix in (2.10).

The possible classifications of x_o, along with the corresponding eigenvalue conditions, are given below:

(a)	stable node	$\lambda_1, \lambda_2 < 0$	$\lambda_1, \lambda_2 \in \mathbb{R}$
(b)	unstable node	$\lambda_1, \lambda_2 > 0$	$\lambda_1, \lambda_2 \in \mathbb{R}$
(c)	saddle	$\lambda_1 < 0 < \lambda_2$	$\lambda_1, \lambda_2 \in \mathbb{R}$
(d)	stable focus	$Re(\lambda) < 0$	$\lambda \in \mathbf{C}$
(e)	unstable focus	$Re(\lambda) > 0$	$\lambda \in \mathbf{C}$
(f)	center	$Re(\lambda) \equiv 0$	$\lambda \in \mathbf{C}$

Table 1

In order to visualize how the equilibrium point classifications arise, it will be helpful to use a transformation that puts A in a simplified form. Let us change variables in the state-space model, choosing $z = C(x - x_o)$ for some invertible constant matrix C; this corresponds to what is termed a *similarity transformation*. Then (2.10) becomes

$$\dot{z}(t) = CAC^{-1}z(t) \quad , \quad z(0) = C\delta x(0) . \tag{2.11}$$

By proper choice of the transformation matrix C, we can always ensure that CAC^{-1} takes precisely one of the following real forms: diagonal, nontrivial Jordan, and complex conjugate; which particular form it takes depends entirely on A.

4

Diagonal Form

$$CAC^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

In this case, the equilibrium point may be either a stable node, an unstable node, or a saddle point, as indicated in Table 1. This can be seen by considering the trajectories in the $z_1 - z_2$ phase plane. To obtain these trajectories, solve the transformed linear equation (2.11) to get

$$
\begin{aligned}
z_1(t) &= e^{\lambda_1 t} z_{o1} \\
z_2(t) &= e^{\lambda_2 t} z_{o2} \ .
\end{aligned}
\tag{2.12}
$$

Now eliminate t and solve for z_2 in terms of z_1, to obtain

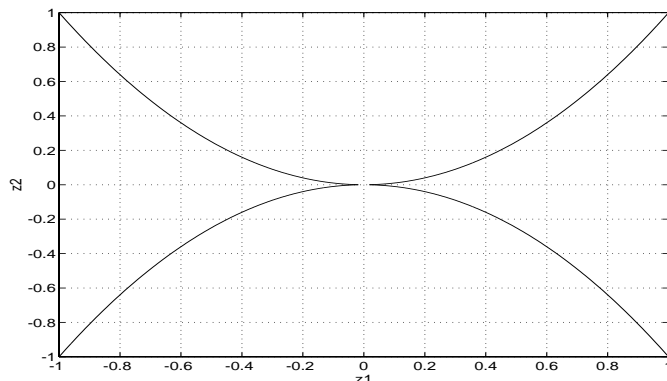$$z_2 = z_{2o} \left(\frac{z_1}{z_o 1} \right)^{\frac{\lambda_2}{\lambda_1}} \tag{2.13}$$

Figure 2.1: stable node, eigenvalues $\lambda_1 = -1$, $\lambda_2 = -2$

If both eigenvalues are negative, the phase portrait in the $z_1 - z_2$ phase plane for several trajectories near the origin will look as shown in Figure 2.1. Notice in this case how the trajectories converge to the origin. This is typical of the behavior near a stable node.

If on the other hand both eigenvalues are strictly positive, an unstable node results. The phase portrait for an unstable node is similar to Figure 2.1 in this case, except that the trajectories emanate from the equilibrium point rather than converging to it. Now consider the case where $\lambda_1 < 0 < \lambda_2$; in this situation, the trajectories resemble those of Figure 2.2. Trajectories such as these are indicative of a saddle point.

Jordan Form

$$CAC^{-1} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}$$

In this case there is one real eigenvalue of multiplicity two, but only a single associated eigenvector. There is then no C that can bring A to diagonal form; the above (nontrivial)
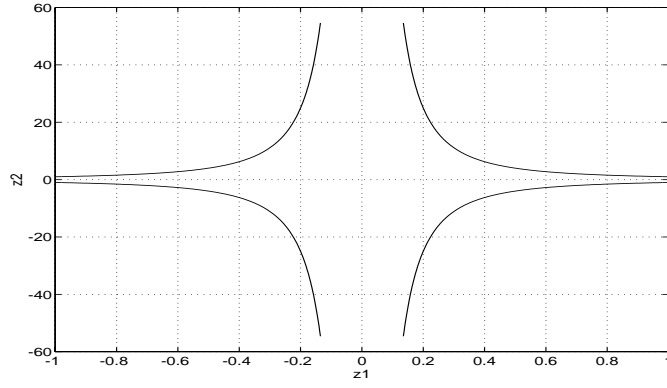
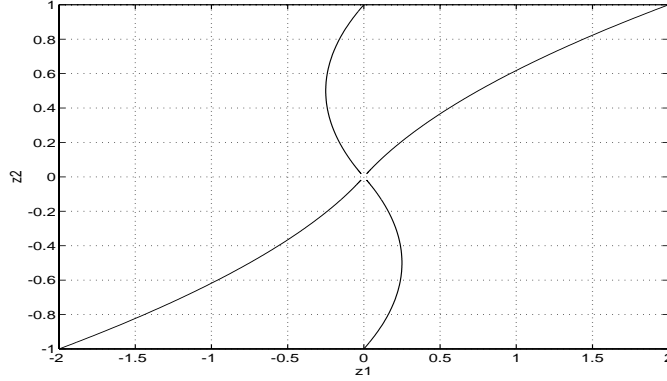Figure 2.2: saddle point, eigenvalues $\lambda_1 = -1, \lambda_2 = 2$

Figure 2.3: stable node, eigenvalues $\lambda = -1$

Jordan form is the closest we can come to diagonal. The corresponding linear differential equations (2.11) may be solved to yield:

$$
\begin{aligned}
z_1(t) &= z_{o1}e^{\lambda t} + z_{o2}te^{\lambda t} \\
z_2(t) &= z_{o2}e^{\lambda t} .
\end{aligned}
$$

Typical trajectories on the $z_1 - z_2$ phase portrait look similar to those in Figure 2.3. In this case $\lambda < 0$ and the equilibrium point is a stable node. In the case that $\lambda > 0$ the trajectories emanate from the equilibrium point, indicating an unstable node.

Complex Conjugate Form

$$
CAC^{-1} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}
$$

Here $\alpha + j\beta$ and $\alpha - j\beta$ are the complex conjugate eigenvalues of A, and $\beta > 0$. In the case where $Re(\lambda) < 0$ the equilibrium point is a stable focus. A trajectory of this case is shown in Figure 2.4 . If $Re(\lambda) > 0$ then the equilibrium point is an unstable focus. In this case a typical trajectory would look similar to Figure 2.4, except that it would be diverging from the equilibrium point. Lastly, if the eigenvalues of A are purely imaginary, then the equilibrium point is called a center. The trajectories about a center are circles.

It should be noted that all of the phase portraits which have been shown in this section are of the $z_1 - z_2$ plane. The phase portraits with respect to $x_1 - x_2$ may be obtained via a
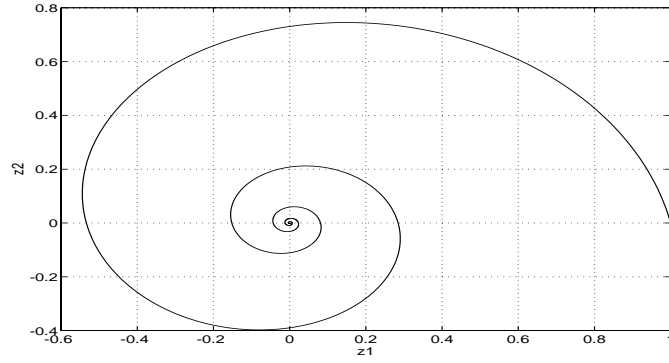
Figure 2.4: stable focus, $\alpha < 0$

linear transformation.

When Does Linearization Predict Local Behavior?

The linearized system predicts the local behavior of the nonlinear system around an equilibrium point if the equilibrium poin is hyperbolic:

Definition 2.2 *(Hyperbolic Equilibrium Point) An equilibrium point x_o is* **hyperbolic** *if $\frac{\partial f}{\partial x}(x_o) = A$ has no eigenvalues on the imaginary axis.*

If the equilibrium point is non-hyperbolic, then the linearized system exhibits oscillatory behavior. The behavior of the nonlinear system around this equilibrium then becomes very dependent on the higher-order terms in the Taylor's series expansion that were neglected in the linearization. On the other hand, if the equilibrium point is hyperbolic, the trajectories of the nonlinear system in the neighborhood of this equilibrium will exhibit one of the possible patterns depicted earlier. More precisely, the *Hartman-Grobman theorem* (which we shall not prove) states that, if the equilibrium point is hyperbolic, then there exists a continuous map with a continuous inverse that transforms every trajectory of the nonlinear system to a trajectory of the linearized system. In summary, the previous classification of an equilibrium point is relevant locally for the corresponding nonlinear system, except in the case of (f).

Example 2.2.2

$$\begin{aligned} \dot{x} &= x + e^{-y} & (2.14) \\ \dot{y} &= -y \,. & (2.15) \end{aligned}$$

What are the equilibrium points of this system, and what type of equilibrium points are they? What, if anything, can be said of the local behavior of the nonlinear system around each of the equilibrium points?

There is only one equilibrium point, at $(-1, 0)$, and the linearized system at this point is

7

$$\dot{z} = \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} z \ .$$

The eigenvalues of this system are at -1 and 1, so the equilibrium point is a saddle point. Since the equilibrium point is hyperbolic, the Hartman-Grobman theorem tells us that the nonlinear system also exhibits a saddle point locally about $(-1, 0)$.

Example 2.2.3

$$\dot{x} = -y + \alpha x(x^2 + y^2) \tag{2.16}$$
$$\dot{y} = x + \alpha y(x^2 + y^2) \ . \tag{2.17}$$

What are the equilibrium points of this system, and what type of equilibrium points are they? What, if anything, can be said of the local behavior of the nonlinear system around each of the equilibrium points?

There is only one equilibrium point, at $(0, 0)$, and the linearized system at this point is

$$\dot{z} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} z \ .$$

The eigenvalues of this system are $+j$ and $-j$, and hence the equilibrium point is a center. Since this equilibrium point is non-hyperbolic, we can draw no conclusion about the behavior of the nonlinear system near $(0, 0)$. The behavior of this nonlinear system can be evaluated analytically by making the substitutions $r^2 = x^2 + y^2$, and $\theta = \tan^{-1}\left(\frac{y}{x}\right)$, and noting that

$$\dot{\theta} = \frac{-y\dot{x} + x\dot{y}}{x^2 + y^2} \ . \tag{2.18}$$

Making these substitutions in the original model leads to the transformed model

$$\dot{r} = \alpha r^3$$
$$\dot{\theta} = 1 \ .$$

From these equations it is clear that for any nonzero initial value of the radius r, the system will diverge from the equilibrium point if $\alpha > 0$ and converge to it if $\alpha < 0$. This behavior is not at all predicted by the behavior of the linearization.

2.3 Exercises

Exercise 2.3.1 1. *Consider a pendulum comprising a mass m at the end of a light but rigid rod of length r. The angle of the pendulum from its equilibrium position is denoted by θ. Suppose a torque $u(t)$ can be applied about the axis of support of the pendulum (e.g. suppose the pendulum is attached to the axis of an electric motor, with the current through the motor being converted to torque). A simple model for this system takes the form*

$$mr^2\ddot{\theta}(t) + f\dot{\theta}(t) + mgr\sin\theta(t) = u(t)$$

where the term $f\dot{\theta}$ represents a frictional torque, with f being a positive coefficient, and g is the acceleration due to gravity.

(a)   Find a state-space representation for this model. Is your state-space model linear? time invariant?

(b)   What nominal input $u_o(t)$ corresponds to the nominal motion $\theta_o(t) = \Omega t$ for all t, where Ω is some fixed constant?

(c)   Linearize your state-space model in (a) around the nominal solution in (b). Is the resulting model time invariant or periodically varying?