# Linear Systems Notes for CDS-140a

October 27, 2008

## 1 Linear Systems

Before beginning our study of linear dynamical systems, it only seems fair to ask the question "why study linear systems?" One might hope that most/all real-world systems are linear systems, so that our development of linear systems theory is directly applicable to physical problems. Unsurprisingly, however, most real systems are nonlinear. However, developing a robust theory of linear systems is advantageous for a number of reasons:

- It gives us practice in analyzing dynamical system.
- It builds up a set of general techniques that can be used to analyze both linear and nonlinear systems.
- In nonlinear systems, we often are interested in local behaviors, which done using a linear approximation of the system near the point of interest.

### 1.1 Definition

An autonomous linear dynamical system (simply a *linear system* from now on) is a collection of autonomous linear ordinary differential equations. Such systems can be written in *first-order form*:

$$\begin{cases} \dot{x}(t) = Ax(t) \\ x(0) = x_0 \end{cases} \tag{1.1.1}$$

where $x$ maps $\mathbb{R}$ to $\mathbb{R}^n$, $A$ is an $n \times n$ matrix, and $b \in \mathbb{R}^n$. We call $x(t)$ the *state* of the system at time $t$. Whenever $b = 0$, we call the system *homogeneous*, otherwise it is *inhomogeneous*. If $A$ is diagonal, the equations in the system decouple into $n$ independent equations, each one in a single variable. In this case we say the system is *uncoupled*.

We will consider only homogeneous systems until 1.12. Also, we will restrict our analysis to the reals, and will explicitely note whenever we need to utilize complex numbers.

## 1.2 Formal Solution

Consider the ordinary differential equation $\frac{d}{dt}x(t) = \alpha x(t)$ with $x(0) = x_0$. We know that this problem has $x(t) = e^{\alpha t}x_0$ as its solution. Motivated by this, we can consider formally writing the solution to our problem (Equation 1.1.1) in the same way:

$$x(t) = e^{At}x_0 \tag{1.2.1}$$

However, as yet we do not have a definition for the exponential of a matrix. For a square matrix $A$, we define the matrix exponential to be the Taylor-series expansion of the real-valued exponential:

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots \tag{1.2.2}$$

where $I$ is, as usual, the $n \times n$ identity matrix. As we will see in Section 1.5, this series converges for all matrices. Although the series is defined and converges for all square matrices A, it is computationally intractible to calculate the matrix exponential by using the series expansion. Instead, we will develop techniques to exponentiate matrices without resorting to the series definition. To explore these methods, let's first look at some special cases and examples.

## 1.3 Diagonal matrices

Let A be a diagonal matrix with diagonal elements $\lambda_1, \ldots, \lambda_n$:

$$A = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \tag{1.3.1}$$

Then referring back to Equation 1.1.1, we can explicitely write the linear system as a set of differential equations:

$$\begin{aligned} \dot{x}_1(t) &= \lambda_1 x_1(t) \\ &\vdots \\ \dot{x}_n(t) &= \lambda_n x_n(t) \end{aligned} \tag{1.3.2}$$

Since these equations are completely decoupled, we can immediately write down their solutions as:

$$\begin{aligned} x_1(t) &= e^{\lambda_1 t} x_1(0) \\ &\vdots \\ x_n(t) &= e^{\lambda_n t} x_n(0) \end{aligned} \tag{1.3.3}$$

Taking for granted that the solutions to linear differential equations are unique (we will prove this in Section 1.8), inspection shows that the matrix exponential $e^{At}$ is:

$$e^{At} = \begin{bmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{bmatrix} \tag{1.3.4}$$

Alternately, we can derive this same expression by applying Equation 1.2.2 and using induction on the diagonal entries.

Note that in the $n = 2$ case, solutions to the problem have a special property. Examining Equation 1.3.3, we can see that any solution of the linear system $(x_1(t), x_2(t)) = \left(e^{\lambda_1 t} x_1(0), e^{\lambda_2 t} x_2(0)\right)$ always satisfies

$$\frac{x_1(t)^{\lambda_2}}{x_2(t)^{\lambda_1}} = \text{constant}$$

We can classify linear systems according to the eigenvalues of the matrix $A$. Let's look first at two-dimensional systems with two distinct real eigenvalues (more generally, two linearly independent eigenvectors). There are three cases. eigenvalues.

### 1.3.1 Example: $\lambda_1 > 0$, $\lambda_2 > 0$

Consider the system

$$\begin{aligned} \dot{x}_1(t) &= x_1(t) \\ \dot{x}_2(t) &= 2x_2(t) \end{aligned}$$

By inspection we see that $\lambda_1 = 1$ and $\lambda_2 = 2$. Thus the solution to the system is:

$$\begin{aligned} x_1(t) &= e^{t} x_1(0) \\ x_2(t) &= e^{2t} x_2(0) \end{aligned}$$

See Figure 1.3.1 for a phase portrait. Because the trajectories appear to start at the origin and move away, we call the origin an "unstable node" or a "source" (this will be defined formally later).

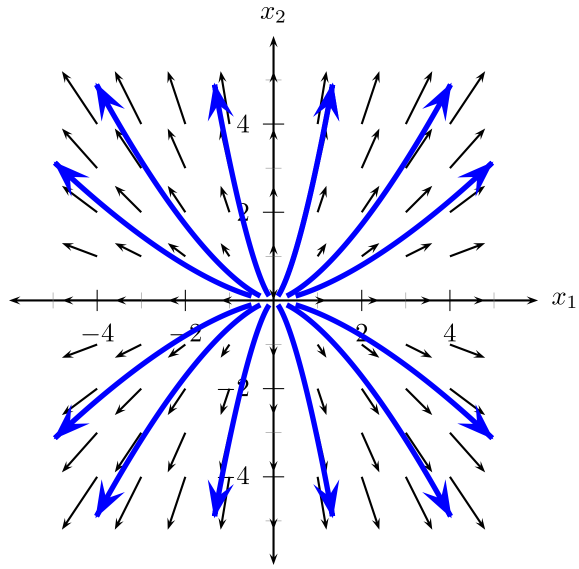


Figure 1.3.1: Phase Portrait and Selected Solution Curves for Example 1.3.1 with $\lambda_1 > 0$, $\lambda_2 > 0$

### 1.3.2 Example: $\lambda_1 < 0$, $\lambda_2 < 0$

Consider the system

$$\dot{x}_1(t) = -0.5x_1(t)$$
$$\dot{x}_2(t) = -x_2(t)$$

By inspection we see that $\lambda_1 = -0.5$ and $\lambda_2 = -1$. Thus the solution to the system is:

$$x_1(t) = e^{-0.5t}x_1(0)$$
$$x_2(t) = e^{-t}x_2(0)$$

See Figure 1.3.2 for a phase portrait. Note that the phase portrait is qualitatively similar to Figure 1.3.1, with the direction of the arrows pointing toward the origin instead of away from it. Because all the trajectories appear to converge to the origin, we call this type of equilibrium solution a "stable node" or a "sink" (this will be defined formally later).

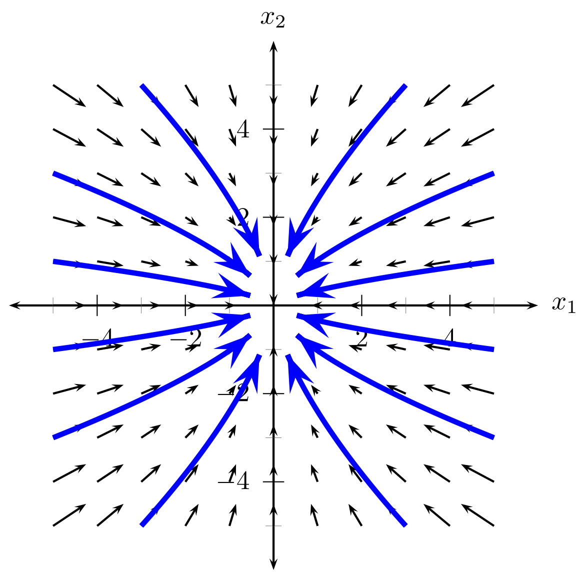


Figure 1.3.2: Phase Portrait and Selected Solution Curves for Example 1.3.2 with $\lambda_1 < 0$, $\lambda_2 < 0$

### 1.3.3 Example: $\lambda_1 > 0$, $\lambda_2 < 0$

Consider the system

$$\begin{aligned}\dot{x}_1(t) &= -1.5x_1(t)\\ \dot{x}_2(t) &= x_2(t)\end{aligned}$$

By inspection we see that $\lambda_1 = -1.5$ and $\lambda_2 = 1$. Thus the solution to the system is:

$$\begin{aligned}x_1(t) &= e^{-1.5t}x_1(0)\\ x_2(t) &= e^{t}x_2(0)\end{aligned}$$

See Figure 1.3.3 for a phase portrait. Notice that the solution curves appear to be hyperbolas. In this case we say that the origin is a “hyperbolic” point or a “saddle” point (this will be defined formally later).

## 1.4 Diagonalizable Matices

The technique for exponentiating diagonal matrices may seem to be of limited use, but it can also be used with a much larger class of matrices. These matrices are known as *diagonalizable* and are, in some sense, equivalent to diagonal

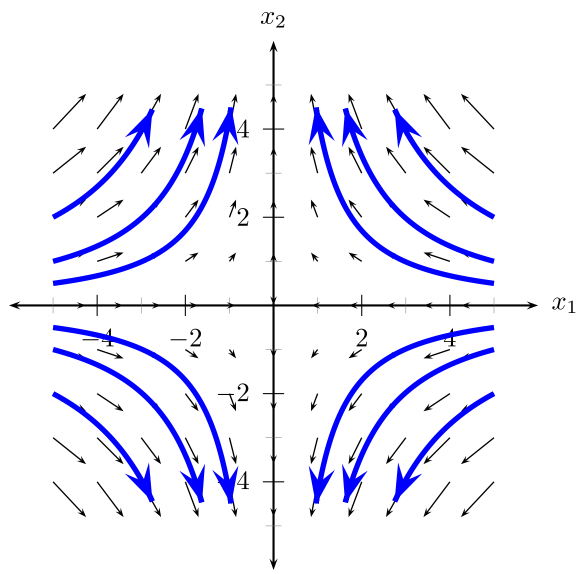


Figure 1.3.3: Phase Portrait and Selected Solution Curves for Example 1.3.3 with $\lambda_1 < 0$, $\lambda_2 > 0$

matrices. First recall that two matrices $A$ and $B$ are said to be similar if there exists an invertible matrix $S$ such that $A = SBS^{-1}$. Then $A$ and $B$ are said to be equivalent up to a similarity transformation.

Given a matrix $A \in \mathbb{R}^{n\times n}$ with $n$ linearly independent eigenvectors (equivalently, the eigenvectors for the matrix span $\mathbb{R}^n$), then A is similar to a diagonal matrix and said to be diagonalizable. More specifically, define $\Lambda$ and $P$:

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \tag{1.4.1}$$

$$P = \begin{bmatrix} | & & | \\ v_1, & \cdots, & v_n \\ | & & | \end{bmatrix} \tag{1.4.2}$$

where $\{\lambda_1, \dots, \lambda_n\}$ are the eigenvalues and $\{v_1, \dots, v_n\}$ are the corresponding eigenvectors of $A$. Then

$$A = P\Lambda P^{-1} \tag{1.4.3}$$

Consider the definition of the matrix exponential (Equation 1.2.2) and noticing that for $A$ similar to $B$,

$$A^k = SB^kS^{-1} \tag{1.4.4}$$

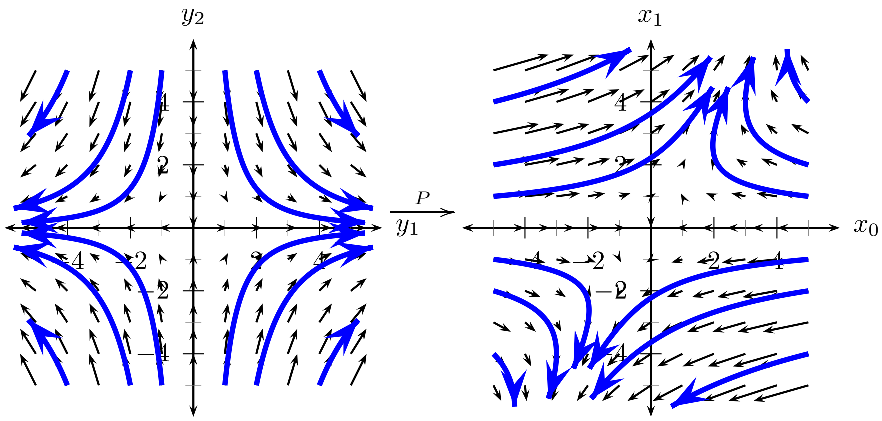


Figure 1.4.1: Coordinate Transformation of an ODE System

it is clear that for $A$ diagonalizable

$$\begin{aligned} e^A &= PIP^{-1} + P\Lambda P^{-1} + \frac{P\Lambda^2 P^{-1}}{2!} + \frac{P\Lambda^3 P^{-1}}{3!} + \cdots \\ &= Pe^{\Lambda} P^{-1} \end{aligned} \tag{1.4.5}$$

Thus by Equation 1.2.1, the solution for a system with a diagonalizable matrix $A$ is

$$x(t) = Pe^{\Lambda t} P^{-1} x(0) \tag{1.4.6}$$

If we define a new set of variables $y$ by

$$y = P^{-1} x \tag{1.4.7}$$

then it is clear that the system can be recast in terms of these new variables

$$\dot{y}(t) = \Lambda y(t) \tag{1.4.8}$$

with initial conditions $y(0) = y_0 = P^{-1} x_0$, which has the solution

$$y = e^{\Lambda t} y_0 \tag{1.4.9}$$

This means that we can solve a diagonalizable system by rewriting the problem in a different coordinate system, solving the diagonal system in the new coordinates, then applying a transformation to return back to the original coordinates (see Figure 1.4.1).

## 1.5 Matrix Exponential Convergence Analysis

Before examining nondiagonalizable matrices, we take a brief detour to prove the convergence of the matrix exponential. In order to show that the matrix exponential defined in Equation 1.2.2 exists (i.e. for all matrices $A$ the series converges), we need to endow our vector space with some topological structure. We will define a norm on the vector space.

For $B \in \mathbb{R}^{n\times n}$ we defined its norm as:

$$\|B\| = \sup_{\|x\|=1} \|Bx\| \tag{1.5.1}$$

where the norm of the vector is the standard Euclidean norm. This norm is often referred to as the *operator norm*. Together with this norm, the space of $n \times n$ matrices (for an arbitrary fixed $n \in \mathbb{Z}^+$) forms a normed linear space. That is, $\forall B, C \in \mathbb{R}^{n\times n}, \alpha \in \mathbb{R}$, the following properties hold:

$$\begin{aligned} \|B\| &\geq 0 \\ \|\alpha B\| &= |\alpha|\, \|B\| \\ \|B + C\| &\leq \|B\| + \|C\| \end{aligned}$$

Additionally, the operator norm also satisfies

$$\|BC\| \leq \|B\|\, \|C\| \tag{1.5.2}$$

Because of (1.5.2), we see that

$$\begin{aligned} \left\| \frac{A^k t^k}{k!} \right\| &\leq \frac{\left|t^k\right|}{k!} \left\|A^k\right\| \\ &\leq \frac{|t|^k}{k!} \|A\|^k \end{aligned}$$

Since each term in the series that defines the matrix exponential is dominated by an equivalent term in the series expansion for $e^{\|A\||t|}$, and because we know that $e^{\alpha t}$ is everywhere convergent for all values of $\alpha$, by the comparison test, we know that the series in Equation 1.2.2 is everywhere convergent and thus $e^{At}$ is defined for any $n \times n$ matrix $A$.

## 1.6 More on the Matrix Exponential

**Proposition 1.1.** *Given $S, N \in \mathbb{R}^{n\times n}$ that commute (i.e. $SN = NS$), $e^{S+N} = e^S e^N$*

This can be proved by examining the power series expansion for $e^{S+N}$, but convergence of the series needs to be carefully considered. Another method to prove convergence is to examine a particular system of differential equations and take advantage of the uniqueness of solutions to linear ODEs (yet to be proved).

*Proof.* Consider the system $\dot{x} = (S + N)\, x$. From our earlier discussion it's clear that $x(t) = e^{(S+N)t}x_0$ solves the system.

Now consider $x(t) = e^{St}e^{Nt}x_0$. Then by using the power rule for derivatives, $\dot{x}(t) = Se^{St}e^{Nt}x_0 + e^{St}Ne^{Nt}x_0$. Since $S$ and $N$ commute, by looking at the power series expansion for $e^{St}$, we see that $N$ commutes with $e^{St}$, and so we have $\dot{x}(t) = (S + N)\, e^{St}e^{Nt}x_0 = (S + N)\, x(t)$.

By the uniqueness of solutions to linear ODEs, we have $e^{(S+N)t} = e^{St}e^{Nt}$, and by setting $t = 1$, $e^{S+N} = e^S e^N$. □

**Proposition 1.2.** *Given* $T \in \mathbb{R}^{n\times n}$, $\left(e^T\right)^{-1} = e^{-T}$

*Proof.* $I = e^0 = e^{T+(-T)} = e^T e^{-T}$

$\therefore \left(e^T\right)^{-1} = e^{-T}$ □

These two propositions will be useful for computing the matrix exponential for nondiagonalizable matrices.

## 1.7 Nondiagonalizable Matrices

Some matrices are nondiagonalizable, but nonetheless, they can be written in a one of two forms that are very convenient for exponentiating: real canonical form and Jordan form.

In the two dimensional case, the semi-simple form reduces to a rotation matrix, which will be trated in this section. The general semi-simple form will be analyzed in Section 1.10.1

### 1.7.1 Jordan Form Matrices

One simple example of a nondiagonalizable matrix is:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \tag{1.7.1}$$

Note that the matrix $A$ has a single eigenvalue $\lambda = 1$ with multiplicity 2, but only a one-dimensional eigenspace for that eigenvalue, namely $v = (1, 0)^T$. It is this deficiency that exactly characterizes nondiagonalizable matrices.

Note that the corollary discussed above gives a simple way of calculating $e^{tA}$. First, note that $A = S + N$ where

$$S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$N = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{1.7.2}$$

Since $S$ commutes with $N$ ($S = I$ and the identity matrix commutes with every matrix), we know that $e^{tA} = e^{tS+tN} = e^{tS}e^{tN}$. Since $S$ is diagonal, we can compute that easily:

$$e^{tS} = \begin{bmatrix} e^t & 0 \\ 0 & e^t \end{bmatrix} \tag{1.7.3}$$

Note also that $N$ is nilpotent. That is, $N^k = 0$ for some $k \in \mathbb{Z}^+$ (in this case $k = 2$). Because of this property, the infinite series definition for $e^{tN}$ becomes a finite sum which can easily be computed. In this cases

$$e^{tN} = I + tN = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \tag{1.7.4}$$

Thus we have computed the matrix exponential for a nondiagonalizable matrix

$$e^{tA} = e^{tS}e^{tN} = \begin{bmatrix} e^t & te^t \\ 0 & e^t \end{bmatrix} \tag{1.7.5}$$

We can consider the general case of matrices of this form:

$$A = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \tag{1.7.6}$$

Then by applying the same reasoning as above, we get the matrix exponential

$$e^{tA} = \begin{bmatrix} e^{at} & bte^{at} \\ 0 & e^{at} \end{bmatrix} \tag{1.7.7}$$

As an example, consider a system with the matrix

$$A = \begin{bmatrix} -1 & 2 \\ 0 & -1 \end{bmatrix} \tag{1.7.8}$$

Then by the arguments above

$$e^{tA} = \begin{bmatrix} e^{-t} & 2te^{-t} \\ 0 & e^{-t} \end{bmatrix}$$

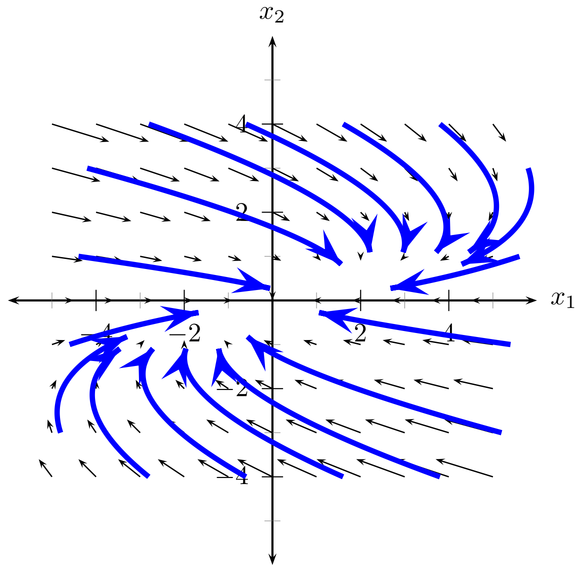


Figure 1.7.1: Phase Portrait and Selected Solution Curves for System Defined by Equation 1.7.8

See Figure 1.7.1 for a phase portrait. Notice that the solution curves appear to converge to the origin, so as before we have a sink. However, note that in this case the trajectories don't approach the origin "monotonically", they appear to shoot past the origin, then turn around and return. This type of node is occassionally known as a "degenerate" node. This behavior is known as "secular" or "transient" behavior.

### 1.7.2 Rotation Matrices

Another type of nondiagonalizable matrix are those that have the form

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \tag{1.7.9}$$

Note that $A = S + N$, where

$$\begin{aligned} S &= \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} = aI \\ N &= \begin{bmatrix} 0 & -b \\ b & 0 \end{bmatrix} \end{aligned} \tag{1.7.10}$$

and since $S = aI$ commutes with N, $e^{tA} = e^{tS}e^{tN}$. Since $y(t) = e^{tN}y(0)$ solves $\dot{y}(t) = Ny(t)$, we can compute $e^{tN}$ by solving the system of differential equations directly.

Explicitly writing out the differential equations, we have:

$$\begin{cases} \dot{x} = -by \\ \dot{y} = bx \end{cases} \tag{1.7.11}$$

Notice that this simplifies to $\ddot{x} + bx = 0$, the equation for simple harmonic motion, which has a solution $x(t) = \alpha\cos(bt) + \beta\sin(bt)$ for some constants $\alpha$, $\beta$. Then the system of equations has the solution:

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{bmatrix} \begin{bmatrix} x(0) \\ y(0) \end{bmatrix}$$

Therefore we have

$$e^{tN} = \begin{bmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{bmatrix}$$

and accordingly,

$$e^{tA} = e^{at} \begin{bmatrix} \cos(bt) & -\sin(bt) \\ \sin(bt) & \cos(bt) \end{bmatrix} \tag{1.7.12}$$

Notice that the trigonometric portion of the matrix is equivalent to a rotation of $bt$ radians about the origin in the phase plane.

As an example, consider a system with the matrix

$$A = \begin{bmatrix} -1 & 2 \\ -2 & -1 \end{bmatrix} \tag{1.7.13}$$

Then by the arguments above

$$e^{tA} = \begin{bmatrix} e^{-t}\cos(2t) & e^{-t}\sin(2t) \\ -e^{-t}\sin(2t) & e^{-t}\cos(2t) \end{bmatrix}$$

See Figure 1.7.2 for a phase portrait. Notice that the solution curves appear to rotate around the origin as they converge to the origin. In this case we say the origin is a "spiral" point or a "focus" (this will be defined formally later).

## 1.8 Existence and Uniqueness

Up to this point we have taken for granted that solutions to linear homogeneous ODEs exist and are unique. We will now prove both of these assertions.

**Theorem 1.3.** *If $A \in \mathbb{R}^{n \times n}$ and $x_0 \in \mathbb{R}^n$, then*

$$\begin{cases} \dot{x} = Ax \\ x(0) = x_0 \end{cases}$$

*has a unique solution*

$$x(t) = e^{At}x_0$$

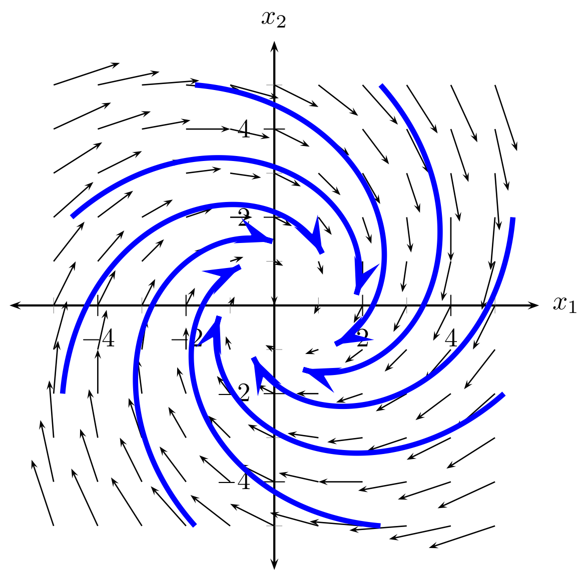


Figure 1.7.2: Phase Portrait and Selected Solution Curves for System Defined by Equation 1.7.13

*Existence.* Let $x(t) = e^{At}x_0$. Clearly $x(0) = x_0$. We now want to show $\frac{dx}{dt} = Ax$.

By the definition of the derivative

$$\begin{aligned}
\frac{dx}{dt} &= \lim_{h\to 0} \frac{x(t+h) - x(t)}{h} \\
&= \lim_{h\to 0} \frac{e^{(t+h)A}x_0 - e^{tA}x_0}{h} \\
&= \lim_{h\to 0} e^{tA}\frac{e^{hA} - I}{h}x_0 \\
&= \lim_{h\to 0} e^{tA}\frac{1}{h}\left(hA + \frac{h^2A^2}{2!} + \cdots\right)x_0 \\
&= \lim_{h\to 0} e^{tA}\left(A + \frac{hA^2}{2!} + \frac{h^2A^3}{3!} + \cdots\right)x_0 \\
&= \lim_{h\to 0} e^{tA}Ax_0 + he^{tA}\left(\frac{A^2}{2!} + \frac{hA^3}{3!} + \frac{h^2A^4}{4!} + \cdots\right)x_0
\end{aligned}$$

Because $A$ and $e^{tA}$ commute (since the matrix exponential is defined as a power-series expansion in powers of $A$) and the series in the last line converges to a matrix $B$, the second term in the limit above vanishes, leaving us with

$$\begin{aligned}
\frac{dx}{dt} &= Ae^{tA}x_0 \\
&= Ax
\end{aligned} \tag{1.8.1}$$

$\square$

*Uniqueness.* Let $x(t)$ and $\tilde{x}(t)$ be solutions to the ODE

$$\begin{cases} \frac{d}{dt}x(t) = Ax(t) \\ x(0) = x_0 \end{cases}$$

We will show that these two solutions must be the same.

Consider $e^{-tA}\tilde{x}(t)$, which we will show is a constant function of t.

$$\begin{aligned} \frac{d}{dt}e^{-tA}\tilde{x}(t) &= -Ae^{-tA}\tilde{x}(t) + e^{-tA}A\tilde{x}(t) \\ &= 0 \end{aligned}$$

since $e^{-tA}$ commutes with $A$.

$$\begin{aligned} \therefore \quad & e^{-tA}\tilde{x}(t)\,\text{isconstant} \\ \therefore \quad & e^{-tA}\tilde{x}(t) = e^{-0A}\tilde{x}(0) = x_0 \\ \therefore \quad & \tilde{x}(t) = e^{tA}x_0 \\ \therefore \quad & \tilde{x}(t) = x(t) \end{aligned}$$

$\square$

## 1.9 Classification of Planar Linear Systems

Now that we have looked at a number of planar phase portraits and attempted to qualitatively classify them, we seek to rigorously classify the phase plots of all planar linear systems in terms of the matrix $A$. It turns out that the correct way to do this is to examine the eigenvalues.

Consider the linear system

$$\dot{x} = Ax \tag{1.9.1}$$

where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{1.9.2}$$

Recall that the characteristic equation for the matrix $A$ is

$$\begin{aligned} 0 &= \det(A - \lambda I) \\ &= (a - \lambda)(d - \lambda) - bc \\ &= \lambda^2 - \tau\lambda + \delta \end{aligned} \tag{1.9.3}$$

where $\tau = \mathrm{tr}\,(A) = a+d$, $\delta = \det(A) = ad-bc$. Recall that these quantities are invariant under change of basis (a.k.a. similarity transformation). In the higher dimensional case, the characteristic polynomial has only invariants as the coefficients of lambda. In this particular case, this can be shown by remembering, from linear algebra, that $\det(CD) = \det(C)\det(D)$ and $\mathrm{tr}\,(CD) = \mathrm{tr}\,(DC)$. Thus, for $A = SBS^{-1}$,

$$\begin{aligned}\det(A) &= \det\left(SBS^{-1}\right)\\ &= \det(S)\det\left(S^{-1}\right)\det(B)\\ &= \det(B)\end{aligned}$$

and

$$\begin{aligned}\mathrm{tr}\,(A) &= \mathrm{tr}\left(SBS^{-1}\right)\\ &= \mathrm{tr}\left(BSS^{-1}\right)\\ &= \mathrm{tr}\,(B)\end{aligned}$$

In other words, these quantities are a property of the linear transformation represented by the matrix $A$, and so are the same for all matrix representations of that particular linear transform. Thus, because $\tau$ and $\delta$ are invariants, it is a good idea to try to classify the phase plots in terms of these quantities. Before we begin, take note that

$$\lambda = \frac{\tau \pm \sqrt{\tau^2 - 4\delta}}{2} \tag{1.9.4}$$

*Case 1: $\delta < 0$ (Hyperbolic Points)*

Recall that $\delta = \lambda_1\lambda_2$. Then we know that $\lambda_1, \lambda_2 \in \mathbb{R}$, and we can reorder the eigenvalues so that $\lambda_1 < 0$ and $\lambda_2 > 0$. In this case we say the phase plot has a "saddle point" or a "hyperbolic point." See Figure 1.3.3 for an example.

*Case 2: $\tau^2 - 4\delta \geq 0$, $\delta > 0$ (Nodes)*

Since the discriminant of the quadratic is positive, $\lambda_1, \lambda_2 \in \mathbb{R}$. We also know that $\mathrm{sign}\,(\lambda_1) = \mathrm{sign}\,(\lambda_2)$ because $\delta > 0$. Also, $\tau = \lambda_1 + \lambda_2$, so $\mathrm{sign}\,(\tau) = \mathrm{sign}\,(\lambda_1) = \mathrm{sign}\,(\lambda_2)$. In this case we say that the phase plot has a "node." If the eigenvalues are positive, we further classify it as an "unstable node" or "source." Conversely, if the eigenvalues are negative, we classify it as a "stable node" or "sink." Examples of this include diagonalizable matrices, see Figures 1.3.1 and 1.3.2, and one type of nondiagonalizable matrices, see Figure 1.7.1. This last case is an example of what is known as "transient" or "secular" growth.

*Case 3: $\tau^2 - 4\delta < 0$, $\delta \geq 0$ (Spirals)*

In this case, since the discriminant of the quadratic is negative we have $\lambda_1, \lambda_2 \in \mathbb{C}$. Furthermore, since the matrix is real-valued, $\lambda_1 = \overline{\lambda_2} = \lambda$. These matrices

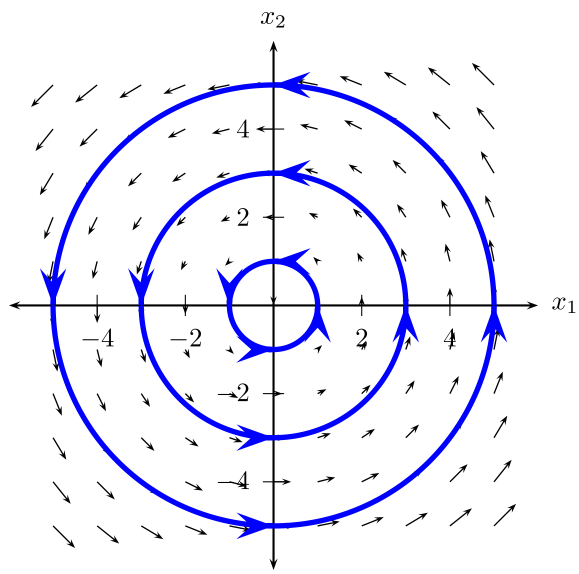


Figure 1.9.1: A Phase Plot Showing a Center

are similar to matrices of the form in Equation 1.7.9, with the correspondence that $\lambda = a \pm ib$. With these eigenvalues the phase plots are said to have a "focus" or a "spiral." We further clasify them as "unstable" if $\tau > 0$ (or equivalently $a > 0$) and "stable" if $\tau < 0$ (or $a < 0$). See Figure 1.7.2 for an example. If the eigenvalues are pure-imaginary, we say the phase portrait exhibits a "center." See Figure 1.9.1 for an example.

*Case 4: $\delta = 0$ (Degenerate Cases)*

Consider first the case when $\tau \neq 0$. Then the matrix has the form

$$A = \begin{bmatrix} 0 & 0 \\ 0 & \lambda \end{bmatrix} \tag{1.9.5}$$

Rewriting this as a differential equation and we have

$$\begin{cases} \dot{x} = 0 \\ \dot{y} = \lambda y \end{cases}$$

which has solutions of the form $(x, y) = \left(x_0, e^{\lambda t} y_0\right)$. In this case, then, trajectories in the phase plot are just vertical lines that approach the x-axis. Also, the x-axis consists completely of fixed-points. See Figure 1.9.2 for an example.

If we now consider the case when $\tau = 0$, we have two more possibilities to consider. One is the trivial case where $A = 0$, which is uninteresting from an

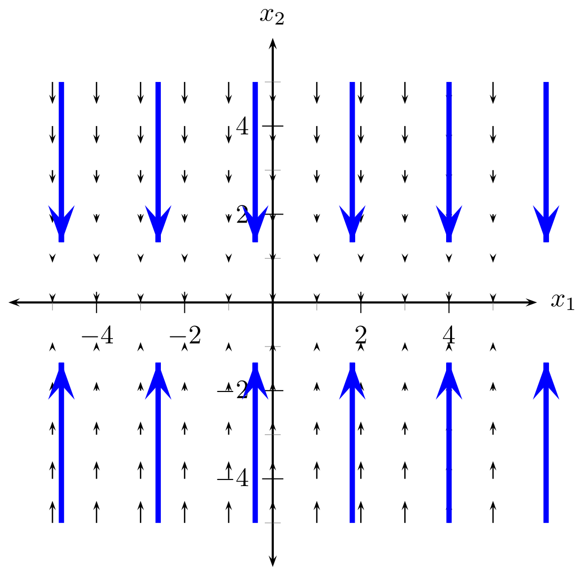


Figure 1.9.2: Phase Portrait and Selected Solution Curves for System Defined by Equation 1.9.5 with $\lambda = -1$

analysis point of view. Alternatively, we could have

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{1.9.6}$$

Rewriting this as a differential equation and we have

$$\begin{cases} \dot{x} = y \\ \dot{y} = 0 \end{cases}$$

which has solutions of the form $(x, y) = (y_0 t + x_0, y_0)$ In this case, then, trajectories in the phase plot are just horizontal liines. Also, the x-axis consists completely of fixed-points. See Figure 1.9.3 for an example.

With this we have classified all possible phase portraits for planar linear systems. The results of this section are summarized graphically in Figure 1.9.4. The $(\tau, \delta)$-plane is divided into qualitatively similar regions and characterized accordingly.

## 1.10 Canonical Forms

One might ask, and rightly so, how do we know we've covered all the cases for a planar system? That is, how do we know that by classifying systems according to the eigenstructure of the matrices, we have classified every system? To do this we need to appeal to linear algebra techniques and discus canonical forms.

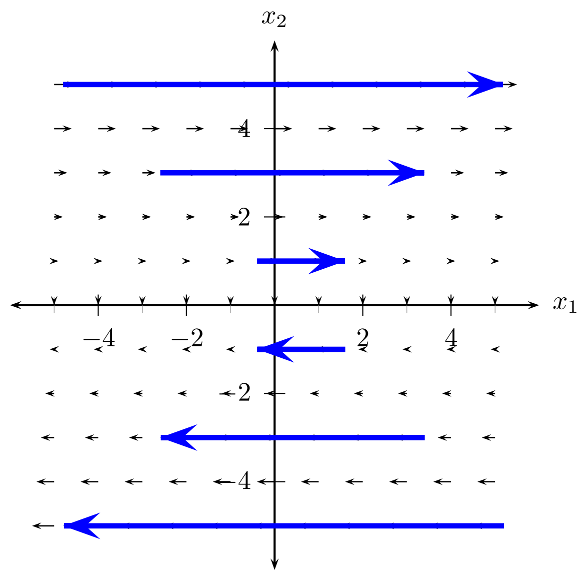


Figure 1.9.3: Phase Portrait and Selected Solution Curves for System Defined by Equation 1.9.6

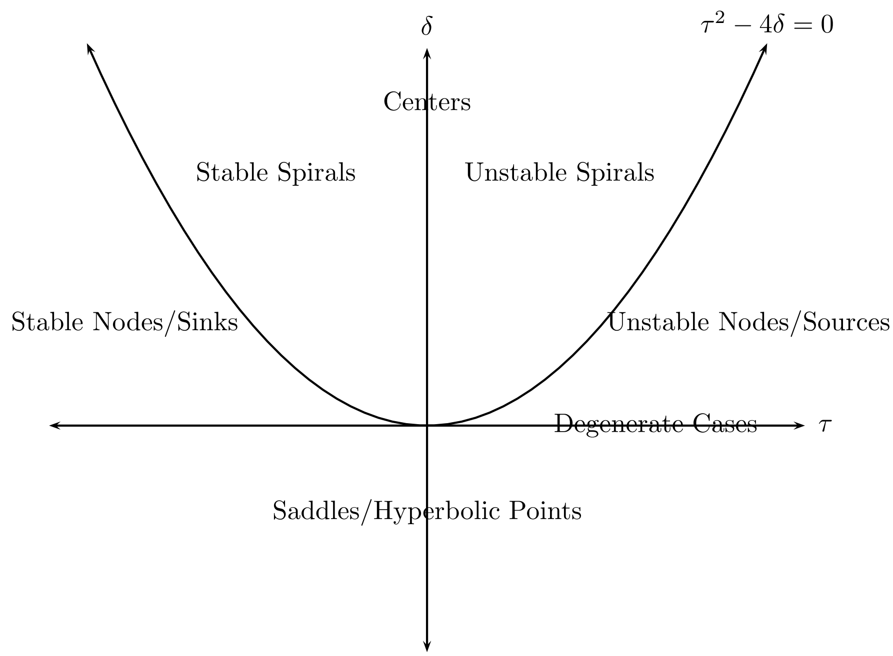


Figure 1.9.4: Phase Plot Classification for a System with Matrix $A$ Based on Values of $\tau = \operatorname{tr}(A), \delta = \det(A)$

### 1.10.1 Real Canonical Form

Consider a matrix $A \in \mathbb{R}^{2n\times 2n}$ with $2n$ distinct complex (non-real) eigenvalues. Since the matrix is real-valued, we know that the eigenvalues come in complex-conjugate pairs, so the collection of eigenvalues looks like:

$$\lambda_k = a_k + ib_k, \overline{\lambda_k} = a_k - ib_k$$

and has eigenvectors:

$$w_k = u_k + iv_k, \overline{w_k} = u_k - iv_k$$

where $u_k, v_k \in \mathbb{R}^{2n}$.

If we consider the field associated with the vector space of matrices to be the complex numbers $\mathbb{C}$, then $A$ is clearly diagonalizable. However, if we consider the field to be the real numbers $\mathbb{R}$, then $A$ is not diagonalizable (consider, for example, the rotation matrix discussed in Equation 1.7.9, which has eigenvalues $\lambda = a \pm ib$ and is not diagonalizable). However, if we construct the matrix $P$ by

$$P = \begin{bmatrix} | & | & & | & | \\ u_1, & v_1, & \cdots, & u_n, & v_n \\ | & | & & | & | \end{bmatrix} \tag{1.10.1}$$

(P turns out to be invertible), then the matrix $B$ (which is similar to $A$ under $P$):

$$B = PAP^{-1} = \begin{bmatrix} \begin{bmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{bmatrix} & & & \\ & \begin{bmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{bmatrix} & & \\ & & \ddots & \\ & & & \begin{bmatrix} a_n & -b_n \\ b_n & a_n \end{bmatrix} \end{bmatrix} \tag{1.10.2}$$

which is the *real canonical form* for the matrix $A$. This form is advantageous because it is easy to exponentiate, since for a matrix $K$ in block form i.e. for square matrices $C, D$ (not necessarily the same dimension),

$$K = \begin{bmatrix} C & 0 \\ 0 & D \end{bmatrix} \implies e^K = \begin{bmatrix} e^C & 0 \\ 0 & e^D \end{bmatrix}$$

More generally, if a matrix $A$ has distinct eigenvalues

$$(\lambda_1, \ldots, \lambda_p, a_1 + ib_1, a_1 - ib_1, \ldots, a_q + ib_q, a_q - ib_q)$$

with corresponding eigenvectors

$$(v_1, \ldots, v_p, u_1 + iw_1, u_1 - iw_1, \ldots, u_q + iw_q, u_q - iw_q)$$

where $\lambda_1, \ldots, \lambda_p, a_1, \ldots, a_q, b_1, \ldots, b_q \in \mathbb{R}$ and $v_1, \ldots, v_p, u_1, \ldots, u_q, w_1, \ldots, w_q \in \mathbb{R}^{2n}$ (i.e. we have $p$ real eigenvalues/eigenvectors and $2q$ complex eigenvalues/eigenvectors), then by forming

$$P = \begin{bmatrix} | & & | & | & | & & | & | \\ v_1, & \cdots, & v_p, & u_1, & w_1, & \cdots, & u_q, & w_q \\ | & & | & | & | & & | & | \end{bmatrix}$$

we produce a canonical form

$$PAP^{-1} = \begin{bmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & 0 & \\ & & \lambda_p & & & & \\ & & & \begin{bmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{bmatrix} & & & \\ & & & & \ddots & & \\ & 0 & & & & & \begin{bmatrix} a_q & -b_q \\ b_q & a_q \end{bmatrix} \end{bmatrix} \tag{1.10.3}$$

If, by change of basis (similarity transform), an arbitrary matrix A can be written in the form from Equation 1.10.3, then A is said to be *semi-simple*, and this form is the real-canonical form for the matrix. Note that this class of matrices is larger than the class used to define the semi-simple matrices in that it allows for repeated eigenvalues, as long as there is no eigenspace deficiency[1].

**Theorem 1.4** ($S + N$ Decomposition[2])**.** *Any matrix $A$ has a unique decomposition $A = S + N$ where $S$ is semi-simple, $N$ is nilpotent, and $S$ and $N$ commute.*

Note that the $S + N$ decomposition is unnecessary if we permit $\mathbb{C}$-valued matrices, in that case every matrix is diagonalizable if the field over which it is defined is algebraically-closed (i.e. $\mathbb{C}$). However, because we require that all matrices be $\mathbb{R}$-valued, we need the $S + N$ decomposition.

In order to compute this decomposition for a matrix $A \in \mathbb{R}^{n \times n}$, follow this procedure (note: we assume that there is some eigenvalue $\lambda$ that has algebraic multiplicity $r > 1$, if all the eigenvalues have multiplicity unity, then $S = A, N = 0$)

1. Find all the eigenvalues

[1] We have an eigenspace deficiency when, for an eigenvalue that is repeated $n$ times, there are fewer than $n$ linearly independent eigenvectors (or, when the algebraic and geometric multiplicities of an eigenvalue differ)

[2] a.k.a. the Jordan-Chevalley Decomposition

2. For each eigenvalue $\lambda$ with algebraic multiplicity[3] $r > 1$, compute $r$ generalized eigenvectors. That is, find $r$ linearly independent vectors $v_1, \ldots, v_r$ that each solves one of the equations:

$$\begin{cases} (A - \lambda I)\, v_1 = 0 \\ (A - \lambda I)\, v_2 = v_1 \\ \qquad\qquad \vdots \\ (A - \lambda I)\, v_r = v_{r-1} \end{cases}$$

   It is necessary to search for generalized eigenvectors because there may be a deficiency in the eigenspace, i.e. there are fewer than $r$ eigenvectors for the eigenvalue $\lambda$ in $A$.

3. Form D, the semi-simple "diagonal" matrix consisting of the eigenvalues as shown in Equation 1.10.3

4. Form the matrix $P$

$$P = \left[ v_1, \ldots, v_m, w_{m+1}, u_{m+1}, \ldots, w_{\frac{1}{2}(n-m)}, u_{\frac{1}{2}(n-m)} \right]$$

   where $v_k$s are the eigenvectors associated with the real eigenvalues and $w_k \pm i u_k$ are the eigenvectors associated with the complex eigenvalues.

5. Form $S = PDP^{-1}$

6. Form $N = A - S$

We assert that this provides the $S + N$ decomposition, although we have proven neither that $SN = NS$ nor that $N$ is nilpotent.

### 1.10.2 Jordan Canonical Form

We will neither state the the general Jordan Canonical Form for a matrix, nor prove its existence. Instead we will examine all the possible Jordan forms for a matrix $A \in \mathbb{R}^{3\times 3}$ with real eigenvalues. From this it is fairly straightforward to deduce the general form.

In essence, the Jordan Canonical Form states that all matrices are similar to either a diagonal or an "almost-diagonal" matrices, and that this similarity can be determined solely by examining the eigenvalues of a matrix and their algebraic and geometric multiplicity.

[3]Recall that algebraic multiplicity of an eigenvalue $\lambda$ is its multiplicity as a root of the characteristic equation, and the geometric multiplicity for the eigenvalue is the nullity of $A - \lambda I$, defined as $\dim(\ker(A - \lambda I))$

If the eigenvectors form a basis for $\mathbb{R}^3$, then the matrix $A$ is diagonalizable.

If we have eigenvalues $(\lambda, \mu, \mu)$ and $\mu \neq \lambda$, then, denoting similarity between two matrices as $A \sim B$, either $A$ is diagonalizable as described above, or

$$A \sim \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 1 \\ 0 & 0 & \mu \end{bmatrix}$$

In this case we have a one-dimensional eigenspace for $\lambda$ and a one-dimensional eigenspace for $\mu$, so we have a deficient eigenspace for $\mu$. To find the missing basis vector for $\mathbb{R}^3$, we search for a generalized eigenvector:

1. Let $v_1$ be an eigenvector for $\lambda$
2. Let $v_2$ be an eigenvector for $\mu$
3. Define $v_3$ as a solution to $(A - \mu I)\, v_3 = v_2$

Note that the definition of $v_3$ requires that the nullspace of $A - \lambda I$ be a subset of the range of $A - \lambda I$; otherwise we can't assume that there is a $v_3$ that solves the equation. This turns out to be a non-issue. Then defining $P$ by

$$P = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix}$$

we have

$$A = P \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 1 \\ 0 & 0 & \mu \end{bmatrix} P^{-1}$$

Consider

$$J = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 1 \\ 0 & 0 & \mu \end{bmatrix}$$

then $J$ has a $S + N$ decomposition (by inspection)

$$S = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{bmatrix} \qquad N = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

If for some matrix $A$, $P^{-1}AP = J$, then

$$A = PJP^{-1} = PSP^{-1} + PNP^{-1} = S_0 + N_0$$

Clearly $S_0$ is semi-simple and $N_0$ is nilpotent, so if we have the Jordan Form of a matrix $A$ and the $S + N$ decomposition for the matrix $J$, we can easily produce the $S + N$ decomposition for the original matrix $A$.

As a concrete example, consider

$$A = \begin{bmatrix} 2 & 10 \\ 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$$

We could compute its $S + N$ decomposition using the Jordan Form, or we can write it by inspection:

$$A = S + N = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 10 \\ 0 & 0 \end{bmatrix}$$

Finally, consider the case when $A$ has a single eigenvalue $\lambda$ with algebraic multiplicity three. Then either $A$ is diagonalizable, or

$$A \sim \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix} \text{ or } \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

Letting $N$ be the nilpotent matrix in the $S + N$ decomposition for $A$, note that if $A$ is diagonalizable, for $K = \ker(A - \lambda I)$, $\dim(K) = 3$ and $N^1 = 0$. Since there is no deficiency in the eigenspace, there are three linearly independent eigenvectors. In the first of the two cases above, $\dim(K) = 2$ and $N^2 = 0$, and we need to look for two linearly independent eigenvectors and one generalized eigenvector. In the second case, $\dim(K) = 1$ and $N^3 = 0$, and we must look for one eigenvector and linearly independent generalized eigenvectors.

Let's go through the process of finding the generalized eigenvector in the case where $\dim(K) = 2$. In this case, we know we can find the eigenvectors $v_1$ and $v_2$. Let $R = \mathcal{R}(A - \lambda I)$, then by the Rank-Nullity Theorem, $n = 3 = \dim(K) + \dim(R) = 2 + \dim(R) \implies \dim(R) = 1$. Thus, if we pick $0 \neq v \in R$, then $\exists u \ni (A - \lambda I)\, u = v$. Then $(A - \lambda I)\, v \in R$, but $\dim(R) = 1 \implies (A - \lambda I)\, v = \alpha v$. Assuming $\alpha \neq 0$, and since $v \neq 0$, then $Av - (\lambda - \alpha)\, v = 0$, so $\lambda - \alpha$ is an eigenvalue. Since we know it's not an eigenvalue (all of the eigenvalues are $\lambda$), we have a contradiction and so our assumption about $\alpha$ is wrong and thus $\alpha = 0$ and $v$ is an eigenvector. Therefore $(A - \lambda I)\, v = (A - \lambda I)^2 u = 0$, and so $u$ is a generalized eigenvector. Note that in this case we have $R \subset K$.

Consider as an example

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \sim \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Note that in this case we can again write the $S + N$ decomposition by inspection:

$$A = S + N = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Consider as another example the system of differential equations:

$$\begin{cases} \dot{x} = -x + 2y + 3z \\ \dot{y} = -4y + 5z \\ \dot{z} = -4z \end{cases}$$

and the question, do all trajectories approach the origin as $t \to \infty$? Then in matrix form, letting $x = (x, y, z)^T$, $\dot{x} = Ax$, where

$$A = \begin{bmatrix} -1 & 2 & 3 \\ 0 & -4 & 5 \\ 0 & 0 & -4 \end{bmatrix}$$

Since the eigenvalues are $\{-1, -4\}$, the Jordan Canonical Form says that

$$A \sim \begin{bmatrix} -1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{bmatrix} \text{ or } \begin{bmatrix} -1 & 0 & 0 \\ 0 & -4 & 1 \\ 0 & 0 & -4 \end{bmatrix}$$

We know that the trajectories of both of these matrices approach the origin as $t \to \infty$, so the trajectories of the system do as well (if you do the computations, it turns out $A$ is similar to the second Jordan matrix).

This examples serves as a warning about computing the $S + N$ decomposition. Naively, one might posit that the decomposition is

$$A = S + N = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & -4 \end{bmatrix} + \begin{bmatrix} 0 & 2 & 3 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix}$$

based on inspection. However, $S$ and $N$ do not commute, so they do not form the $S + N$ decomposition of $A$.

## 1.11 General Classification

Given $A \in \mathbb{R}^{n \times n}$, let

- $E^S$ = the span of all generalized eigenvectors of $A$ corresponding to eigenvalues with negative real part (Stable subspace)
- $E^C$ = the span of all generalized eigenvectors of $A$ corresponding to eigenvalues with zero real part (Center subspace)
- $E^U$ = the span of all generalized eigenvectors of $A$ corresponding to eigenvalues with positive real part (Unstable subspace)

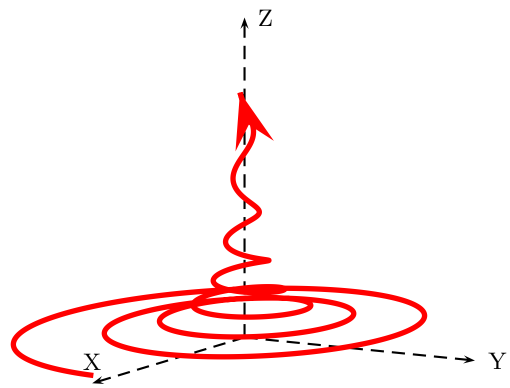


Figure 1.11.1: A Solution Curve for Example Given By Equation 1.11.1

Since $E^S \cap E^C = E^C \cap E^U = E^U \cap E^S = \{0\}$, and $\mathbb{R}^n = E^S \cup E^C \cup E^U$, we can write $\mathbb{R}^N$ as a direct sum of these three subspaces $\mathbb{R}^N = E^S \oplus E^C \oplus E^U$. That is, any point in the space $\mathbb{R}^N$ has a unique decomposition into components in these three subspaces: $\forall x \in \mathbb{R}^N \ \exists! \ \alpha \in E^c, \beta \in E^s, \gamma \in E^u \ni x = \alpha + \beta + \gamma$.

Consider the example

$$\begin{bmatrix} -1 & 2 & 0 \\ -2 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} -1+2i & 0 & 0 \\ 0 & -1-2i & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.11.1}$$

Then $E^S = \text{xy} - \text{plane}$, $E^U = \text{z} - \text{axis}$, and $E_C = \{0\}$. See Figure 1.11.1 for a plot with an example trajectory. Note that the trajectory approaches the z-axis.

It is important to remember that these three subspaces $E^S$, $E^C$, and $E^U$ are *invariant subspaces*. That is, if your initial point lies solely in one of these three subspaces, the trajetory will remain in that subspace indefinitely.

**Theorem 1.5** (Stability Theorem)**.** *The following are equivalent:*

1. $E^S = \mathbb{R}^n$
2. $\forall x_0 \in \mathbb{R}^n$, $\lim_{t\to\infty} e^{tA}x_0 = 0$ *(All trajectories approach the origin)*
3. $\exists \ a, c, m, M > 0 \ni me^{-at} \|x_0\| \leq \left\|e^{tA}x_0\right\| \leq Me^{-ct} \|x_0\|$ *(All trajectories approach the origin exponentially)*

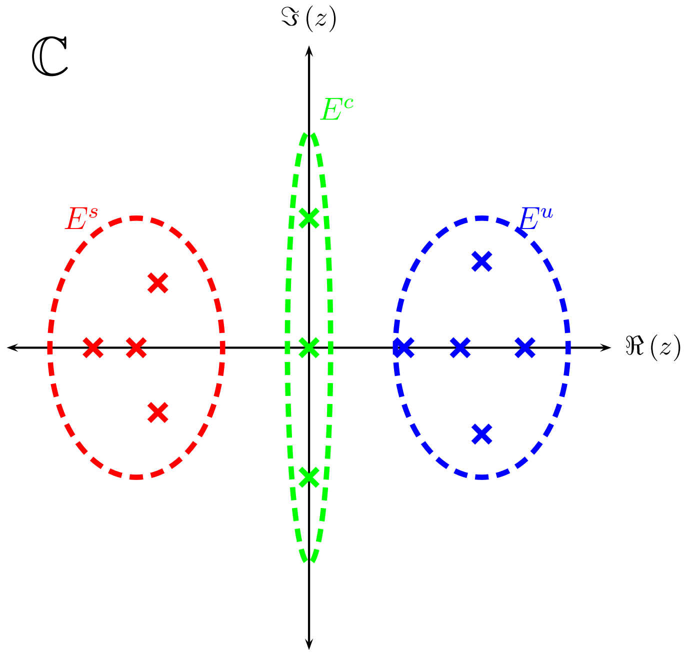


Figure 1.11.2: Eigenvalues (Crosses) for a Hypothetical System and their Subspace Classification

Considering the limit of trajectories as $t$ approaches both $+\infty$ and $-\infty$, this theorem states that the exponential growth rate of the system due to a particular eigenvalue is determined solely by the real part of the eigenvalue (See Figure 1.11.2). We must be careful, however, when considering the growth rate of points in the center subspace. While our language implies that trajectories located in the center subspace are centers, this is not completely true. There are also degenerate cases (i.e. see Figure 1.9.3, which has $\lambda_1 = \lambda_2 = 0$) which do not exhibit rotation.

## 1.12 Variation of Constants/Duhamel's Formula

Up to this point we have considered only homogeneous linear systems. In order to solve nonhomogeneous linear systems, we apply a technique known as Variation of Constants, or alternately Duhamel's Formula.

**Theorem 1.6** (Variation of Constants/Duhamel's Formula)**.** *For the nonhomogeneous linear system*

$$\begin{cases} \dot{x}(t) = Ax(t) + b(t) \\ x(0) = x_0 \end{cases} \tag{1.12.1}$$

*the solution is given by*

$$x(t) = e^{At}x_0 + e^{At}\int\limits_0^t e^{-A\tau}b(\tau)\,d\tau \tag{1.12.2}$$

*Proof.* Since we have been given an explicit formula for the solution, we only need to verify that it solves the system.

$$\begin{aligned} x(0) &= e^{A0}x_0 + e^{A0}\int\limits_0^0 e^{-A\tau}b(\tau)\,d\tau \\ &= x_0 \end{aligned}$$

$$\begin{aligned} \dot{x} &= Ae^{At}x_0 + e^{At}\left[e^{-At}b(t)\right] + Ae^{At}\int\limits_0^t e^{-A\tau}b(\tau)\,d\tau \\ &= A\left[e^{At}x_0 + e^{At}\int\limits_0^t e^{-A\tau}b(\tau)\,d\tau\right] + b(t) \\ &= Ax(t) + b(t) \end{aligned}$$

$\square$

In addition to solving nonhomogeneous linear systems, Duhamel's formula can assist in solving nonlinear systems. Consider for example the system

$$\dot{x} = Ax + f(x) \tag{1.12.3}$$

where $Ax$ represents the entire linear portion of the system and $f(x)$ is the nonlinear portion of the system and consists of all the nonlinear terms. Then by thinking about the system as

$$\dot{x}(t) = Ax(t) + f(x(t))$$

Duhamel says that we can write down an equation that the system must satisfy

$$x(t) = e^{At}x_0 + e^{At}\int\limits_0^t e^{-A\tau}f(x(\tau))\,d\tau \tag{1.12.4}$$

which is an integral equation. In some situations, there is enough information given in the problem to, or the problem satisfies certain criteria that, help characterize the solution when brought to bear on the integral equation.