We previously spent some time discussing the phase plane for second order systems. We’ll now re-
visit that discussion in order to elaborate on some of the finer points. Suppose we have the following
system:
\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} =
\begin{bmatrix}
\lambda_1 & 0 \\
0 & \lambda_2
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\] (1)

where the solution is:
\[
x_1(t) = e^{\lambda_1 t} x_1(0)
\]
\[
x_2(t) = e^{\lambda_2 t} x_2(0)
\]

This system may have arisen when diagonalizing the original system. We will now consider several
different cases.

**Negative Real Roots.** Suppose that $\lambda_1$ and $\lambda_2$ are both real, and both less than 0. For example, if
$\lambda_1 = -1$ and $\lambda_2 = -2$, then we have:
\[
x_1(t) = e^{-t} x_1(0)
\]
\[
x_2(t) = e^{-2t} x_2(0)
\]

Using the equation for $x_1(t)$, we have:
\[
e^{-t} = \frac{x_1(t)}{x_1(0)}
\]
Taking the natural logarithm of both sides, we have:
\[
t = -\ln \left( \frac{x_1(t)}{x_1(0)} \right)
\]

Using the equation for $x_2(t)$ we have:
\[
x_2(t) = e^{-2t} x_2(0)
\]
\[
= e^{-2 \left( -\ln \frac{x_1(t)}{x_1(0)} \right)} x_2(0)
\]
\[
= \left[ e^{\left( \ln \frac{x_1(t)}{x_1(0)} \right)} \right]^2 x_2(0)
\]
\[
= \left[ \frac{x_1(t)}{x_1(0)} \right]^2 x_2(0)
\]
\[
= \frac{x_2(0)}{(x_1(0))^2} (x_1(t))^2
\]
which is a parabola. So we can plot $x_1(t)$ versus $x_2(t)$, and the result is shown in Figure 1.

![Figure 1: Trajectories for Negative Real Roots](image)

Parabolic trajectories, heading towards the origin, for negative real roots, $\lambda_2 > \lambda_1$.

trajectories indicate the progression of change in $x_1$ and $x_2$ as a function of time. Note that all trajectories lead to the origin, regardless of the starting position. Further note that the eigenvalue associated with $x_2$ is larger than the eigenvalue associated with $x_1$, and therefore dominates the response. In other words, $x_2$ goes to zero faster than $x_1$, resulting in the trajectories that we see here.

**One Positive, One Negative Real Root.** Suppose we again have real roots, but now we have $\lambda_1 < 0$ and $\lambda_2 > 0$. For example, what if we have:

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

Performing the same analysis as before, we have:

$$
\begin{align*}
t &= -\ln \left( \frac{x_1(t)}{x_1(0)} \right)
\end{align*}
$$

and:

$$
\begin{align*}
x_2(t) &= e^{2t}x_2(0) \\
&= e^{2\left(-\ln \frac{x_2(t)}{x_2(0)}\right)}x_2(0) \\
&= \left[e^{\ln \frac{x_2(t)}{x_2(0)}}\right]^{-2} x_2(0) \\
&= \left[\frac{x_1(t)}{x_1(0)}\right]^{-2} x_2(0) \\
&= x_2(0)\left(x_1(0)\right)^{2} \frac{1}{\left(x_1(t)\right)^{2}}
\end{align*}
$$
The origin is now called a \textit{saddle point}, and the trajectories are shown in Figure 2. Note that if we start the system on the $x_1$ axis (i.e. $x_2(0) = 0$), it goes to the origin. If we start it anywhere else, it blows up along the $x_2$ axis, even though $x_1$ goes to zero. This system is unstable.

The same can be done for complex roots, and is left as an exercise for the reader (hint: use polar coordinates).

\textbf{Stability of Linear Systems}

Just to reinforce what we already have the heuristics for, we will summarize the stability of linear systems here. Suppose we have a second order system, with eigenvalues $\lambda_1, \lambda_2$.

- If $\lambda_1, \lambda_2$ real and $\lambda_1 < 0, \lambda_2 < 0$, then the system is stable.
- If $\lambda_1, \lambda_2$ real and either $\lambda_1 > 0, \lambda_2 > 0$ or both, then the system is unstable.
- If $\lambda_1, \lambda_2$ complex and real parts of $\lambda_1$ and $\lambda_2$ are less than 0, then the system is stable. If the one or both of the roots has zero real part, then the system is purely oscillatory and is called marginally stable.
- If $\lambda_1, \lambda_2$ complex and real parts of either root is greater than 0, then the system is unstable.

We will come back to explore this more rigorously later on in the course.

\textbf{Non-linear Systems}

We have been discussing primarily linear systems, but in general, as we saw in an extreme case with the neuron, most systems are in fact nonlinear in their behavior. Let us describe systems in the
following way:

\[ \dot{x}(t) = f(x) \]

or, for a second order system:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
f_1(x_1, x_2) \\
f_2(x_1, x_2)
\end{bmatrix}
\]

For example, consider the following system:

\[
\begin{align*}
\dot{x}_1(t) &= \cos x_1 + x_2^2 \\
\dot{x}_2(t) &= \ln x_1
\end{align*}
\]

We can also write this:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
\cos x_1 + x_2^2 \\
\ln x_1
\end{bmatrix}
\]

So, unlike with linear systems, we cannot separate this into the form \( \dot{x} = Ax \). For the linear systems, the state derivatives were linear combinations of the states, but for nonlinear systems, this is not the case. Note that linear systems should be thought of as a special case of linear systems in which we can write \( \dot{x} = f(x) = Ax \).

The set of equations \( \dot{x} = f(x) \) is still considered a state space representation. Can we still obtain this from the original, higher order differential equation? You bet. Consider the following example. Suppose we had our familiar spring mass damper problem, but the damping shows up in such a way as to have some whacky nonlinearity:

\[
m\ddot{y} + b \ln(y)\dot{y} + ky = 0
\]

It’s obviously not linear, because the equation does not just involve a linear combination of the derivatives. Let us define states \( x_1 \) and \( x_2 \), such that:

\[
\begin{align*}
x_1 &= y \\
x_2 &= \dot{y}
\end{align*}
\]

We then have:

\[
\begin{align*}
\dot{x}_1 &= \dot{y} = x_2 \\
\dot{x}_2 &= \dot{\dot{y}} = \frac{-b \ln(y)\dot{y} - ky}{m} \\
&= \frac{-b \ln(x_1)x_2 - kx_1}{m}
\end{align*}
\]

So we have:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
x_2 \\
\frac{1}{m}(-b \ln(x_1)x_2 - kx_1)
\end{bmatrix}
\]
Equilibrium Points

To this point we haven’t formally discussed the concept of equilibrium points, because it has been pretty intuitive. For example, in the spring-mass problem, we know that when the spring is in its unstretched position, this is the equilibrium. For the spring-mass-damper problem, the damper dissipated the energy, bringing the system to its resting place at the equilibrium, also where the spring and damper were in the undisplaced state.

Let’s formalize this discussion a bit. If we consider the system \( \dot{x} = f(x) \), then \( x_{eq} \) is called an equilibrium point of the system if \( f(x_{eq}) = 0 \). What is this saying? This is saying that the point in the phase plane, where the derivatives are zero is the equilibrium point, and at this point the system will not move. The time rate of change of the states is zero.

What about for linear systems? Consider our old friend:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} =
\begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\]  

What values of \( x_1 \) and \( x_2 \) make \( Ax = 0 \)? Let’s see:

\[
\begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} =
\begin{bmatrix}
0 \\
0
\end{bmatrix}
\]

So we need the following:

\[
a_{11}x_1 + a_{12}x_2 = 0 \\
a_{21}x_1 + a_{22}x_2 = 0
\]

As long as the two rows of \( A \) are not linearly related, then the only values of \( x_1 \) and \( x_2 \) that satisfy this set of equations is \( x_1 = 0 \) and \( x_2 = 0 \), which is the origin. So for linear systems, there is only one equilibrium point, and it is at the origin. In the case that the rows are linearly related, it means that you have expressed a first order system with a second order description, and that you need to reduce it.

Nonlinear systems can have multiple equilibria, and they are not necessary at the origin. As an example, consider the following system:

\[
\ddot{y} + 0.6\dot{y} + 3y + y^2 = 0
\]

Define states \( x_1, x_2 \) such that:

\[
x_1 = y \\
x_2 = \dot{y}
\]

So we have:

\[
\begin{align*}
\dot{x}_1 &= x_2 = f_1(x_1, x_2) \\
\dot{x}_2 &= 0.6x_2 - 3x_1 - x_1^2 = f_2(x_1, x_2)
\end{align*}
\]

In order for \( f_1(x_1, x_2) \) to be zero, \( x_2 = 0 \). For \( f_2(x_1, x_2) \) to be simultaneously zero, we have \(-3x_1 - x_1^2 = 0\). This leads to two equilibrium points, \((x_1, x_2) = (0, 0)\) and \((x_1, x_2) = (-3, 0)\).
Nonlinear Systems in the Phase Plane

We have used the phase plane to draw the trajectories for some simple linear systems, and it turns out to be a pretty valuable tool. But we had a pretty good grasp of the dynamics of those systems before we started. How do we do this in general? We’ll use the method of isoclines, which will be described below.

Consider again the following system:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
f_1(x_1, x_2) \\
f_2(x_1, x_2)
\end{bmatrix}
\]

Suppose we form the following ratio:

\[
\frac{\dot{x}_2}{\dot{x}_1} = \frac{dx_2}{dx_1} = \frac{f_2(x_1, x_2)}{f_1(x_1, x_2)}
\]

and we set this ratio equal to a constant \( \alpha \). We then have:

\[
f_2(x_1, x_2) = \alpha f_1(x_1, x_2)
\]

So all points on the line \( f_2 = \alpha f_1 \) for a given \( \alpha \) have same tangent curve, hence the name isocline.

Let’s see what this means for the spring-mass problem. Consider the following:

\[
m\ddot{y} + ky = 0
\]

Let \( x_1 \) and \( x_2 \) be defined such that:

\[
\begin{align*}
x_1 &= y \\
x_2 &= \dot{y}
\end{align*}
\]

So we have:

\[
\begin{align*}
\dot{x}_1 &= \dot{y} = x_2 \\
\dot{x}_2 &= \ddot{y} = -\frac{k}{m}y = -\frac{k}{m}x_1
\end{align*}
\]

Suppose that \( \frac{k}{m} = 1 \) for simplicity. If we form the ratio, we have:

\[
\frac{f_2(x_1, x_2)}{f_1(x_1, x_2)} = \frac{-x_1}{x_2} = \alpha
\]

So we have:

\[
\alpha x_2 + x_1 = 0
\]
If we draw the $x_1, x_2$ plane, we can then pick an $\alpha$, draw the corresponding line in the plane. The trajectories in this plane will all pass through this line with slope $\alpha$. We can see that this defines concentric circles in the plane, as we determined intuitively before.

How about nonlinear systems? Consider the Van der Pol equation:

$$\ddot{y} + 0.2(y^2 - 1)\dot{y} + y = 0$$

Define $x_1$ and $x_2$ such that:

$$x_1 = y$$
$$x_2 = \dot{y}$$

which gives us:

$$\dot{x}_1 = \dot{y} = x_2$$
$$\dot{x}_2 = \ddot{y} = -0.2(y^2 - 1)\dot{y} - y$$
$$= -0.2(x_1^2 - 1)x_2 - x_1$$

Again, for different values of $\alpha$, we plot the corresponding curve in the phase plane. Throughout each curve, the trajectory crosses the curve at the slope $\alpha$. A few curves for different values of $\alpha$ are shown in Figure 3. For nicer pictures, refer to Khoo, Chapter 9, Section 9.3.

**Figure 3: The Van der Pol Isoclines**

Isoclines for a variety of values for $\alpha$.

**Linearization**

Suppose we have a function $y = f(x)$. We can write this function as an Taylor series expansion about a point $x_0$:

$$f(x) = f(x_0) + f'(x)\bigg|_{x_0} (x - x_0) + \frac{f''(x)}{2}\bigg|_{x_0} (x - x_0)^2 + HOTS$$
where HOTS stands for *higher order terms*. We can approximate the function as:

\[
y = f(x) \approx f(x_0) + f'(x) \bigg|_{x_0} (x - x_0)
\]

This is essentially a straight line, or linear approximation of \( f(x) \) around \( x_0 \).

For systems, we have an analogous situation. Consider the nonlinear system \( \dot{x} = f(x) \). Suppose we linearize about the equilibrium point:

\[
\dot{x} \approx f(x_{eq}) + \frac{\partial f}{\partial x} \bigg|_{x_{eq}} (x - x_{eq})
\]

where \( \frac{\partial f}{\partial x} \) is called the Jacobian. Note that the first term \( f(x_{eq}) \) is zero, by the definition of the equilibrium point, so we’re left with:

\[
\dot{x} \approx \frac{\partial f}{\partial x} \bigg|_{x_{eq}} (x - x_{eq})
\]

For our second order system:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
f_1(x_1, x_2) \\
f_2(x_1, x_2)
\end{bmatrix}
\]

we have, after linearization:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\
\frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2}
\end{bmatrix} \begin{bmatrix}
x_{1,eq} \\
x_{2,eq}
\end{bmatrix}
\]

which gives us something like \( \dot{x} = Ax \).

**Example.** In order to drive this point home, let’s do an example. If you’re rusty with partial differentials, then review the appendix first. Consider the following:

\[
\begin{align*}
\dot{x}_1 &= x_2^2 - x_1 \cos x_2 = f_1(x_1, x_2) \\
&= -2x_2 + (x_1 + 1)x_1 + x_1 \sin x_2 = f_2(x_1, x_2)
\end{align*}
\]

We want to linearize this system about the equilibrium point. We find an equilibrium point at \( x_1 = x_2 = 0 \). So we first compute the partial derivatives:

\[
\begin{align*}
\frac{\partial f_1}{\partial x_1} &= -\cos x_2 \\
\frac{\partial f_1}{\partial x_2} &= -x_1 - \sin x_2 \\
\frac{\partial f_2}{\partial x_1} &= -2x_2 + x_1 \sin x_2 \\
\frac{\partial f_2}{\partial x_2} &= 2x_2 + x_1 \sin x_2
\end{align*}
\]
So we have:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
-x_2 \\
2x_1 + 1 + \sin x_2
\end{bmatrix} \begin{bmatrix}
-\cos x_2 \\
2x_1 + x_1 \sin x_2
\end{bmatrix} \begin{bmatrix}
x_1 - x_{1,eq} \\
x_2 - x_{2,eq}
\end{bmatrix}
\]

which results in:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
-1 & 0 \\
1 & -2
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix}
\]

**Stability of Nonlinear Systems**

The story for stability in nonlinear systems is pretty straightforward. The stability is discussed in terms of each equilibrium point, rather than for the whole system. If we put the system at an equilibrium point, it stays there forever. However, if we perturb it slightly, what happens? If the system returns to the equilibrium point, then it is locally stable about that equilibrium point. If it does not, it is not locally stable about that equilibrium point.

The linearization sometimes provides us some information concerning the local stability. If the linearized system is stable, then we can say that the equilibrium point is locally stable. If the linearized system is unstable, then we can say that the equilibrium point is locally unstable. However, if the real part of any of the eigenvalues is zero, we can conclude nothing about the local stability from the linearization. Period.