## A

## Linearization

## A. 1 Functions of one variable

Suppose that $f: \mathbf{R} \rightarrow \mathbf{R}$ is a function which has derivatives of all orders throughout an interval containing $c$, and suppose that

$$
\lim _{n \rightarrow \infty} \frac{f^{n+1}(z)}{(n+1)!}(x-c)^{n+1}=0
$$

for some number $z$ between $x$ and $c$. From Calculus, we know that $f(x)$ is represented by the Taylor series for $f(x)$ at $c$ :

$$
\begin{equation*}
f(x)=f(c)+f^{T}(c)(x-c)+\frac{f^{\prime \prime}(c)}{2!}(x-c)^{2}+\cdots+\frac{f^{(n)}(c)}{n!}(x-c)^{n}+\cdots \tag{A.1}
\end{equation*}
$$

Now, if the point $c$ is close to $x$, that is, if $x-c$ is small, then the second order and higher terms will contribute little to the sum; thus, $f(x)$ can be approximated by the first and second terms:

$$
f(x) \approx f(c)+f^{T}(c)(x-c)
$$

For our purposes, the importance of this approximation is that it is linear in $x$; for this reason, this approximation is known as linearization.

Example 19. Linearize the equation $f(x)=x^{3} /(x+1)$ at the point $c=2$ and find the largest relative error in the region $x \in[1.5,2.5]$.

Solution: We have that $f(c)=8 / 3$. Also,

$$
f^{T}(x)=\frac{2 x^{3}+3 x^{2}}{(x+1)^{2}} ; \quad f^{T}(2)=\frac{28}{9}
$$

Thus, our linear approximation of this function is:

$$
f(x) \approx \frac{28 x-32}{9}
$$

To find the largest relative error, note that the error is given by "error" $=f(x)-$ "approximation". The relative error, $\delta(x)$ is the error divided by $f(x)$ :

$$
\delta(x)=\frac{9 x^{3}-28 x^{2}+4 x+32}{9 x^{3}}
$$

To find the point with the largest relative error, we differentiate $\delta(x)$ and solve for the local minima and maxima. Our function has a local minimum at $x=2$ (this should be obvious!) and a local maximum at $x=-12 / 7$ which is outside our range. It follows that the maximum relative error in our region must occur at either of the boundary points: $\delta(1.5) \approx 0.18, \delta(2.5) \approx 0.05$.

## A. 2 Multivariable functions

In this previous section we have looked at a function of one variable $x$. What happens when $f$ depends on more than one variable? In this case we have a series analogous to that of Eq. A.1; we first state it for $f: \mathbf{R}^{2} \rightarrow \mathbf{R}$. In this case, $f$ is a function of two variables, say $x_{1}$ and $x_{2}: f=f\left(x_{1}, x_{2}\right)$. Provided certain regularity assumptions hold (infinite differentiability, etc.) we can write

$$
\begin{aligned}
& f\left(x_{1}, x_{2}\right)= \underbrace{f(c, d)}_{0 \text { order terms }}+\underbrace{\left.\frac{\partial f\left(x_{1}, x_{2}\right)}{\partial x_{1}}\right|_{c, d}\left(x_{1}-c\right)+\left.\frac{\partial f\left(x_{1}, x_{2}\right)}{\partial x_{2}}\right|_{c, d}\left(x_{2}-d\right)}_{1 \text { st order terms }} \\
&+\underbrace{\left.\frac{\partial^{2} f}{\partial x_{1}^{2}}\right|_{\left(x_{c, d}-c\right)^{2}+\left.2 \frac{\partial^{2} f}{\partial x_{1} \partial x_{2}}\right|_{c, d}\left(x_{1}-c\right)\left(x_{2}-d\right)+\left.\frac{\partial^{2} f}{\partial x_{2}^{2}}\right|_{c, d}\left(x_{2}-d\right)^{2}}}_{2 \text { nd order terms }} \begin{aligned}
& + \text { higher order terms. }
\end{aligned} \\
&
\end{aligned}
$$

Again, we can concentrate on the terms linear in $x_{1}$ and $x_{2}$ to approximate $f$ :

$$
f\left(x_{1}, x_{2}\right) \approx f(c, d)+\left.\frac{\partial f\left(x_{1}, x_{2}\right)}{\partial x_{1}}\right|_{\substack{x_{1}=c, x_{2}=d}} ^{\left(x_{1}-c\right)+\left.\frac{\partial f\left(x_{1}, x_{2}\right)}{\partial x_{2}}\right|_{x_{1}=c, x_{2}=d} ^{\left(x_{2}-d\right)} .} \begin{gather*}
 \tag{A.2}\\
\hline
\end{gather*}
$$

This equation can be expressed more elegantly as follows. Define the vectors:

$$
\mathbf{x}:=\left[\begin{array}{l}
x_{1} \\
x_{2}
\end{array}\right], \quad \mathbf{c}:=\left[\begin{array}{l}
c \\
d
\end{array}\right] \quad \text { and } \nabla:=\left[\begin{array}{c}
\frac{\partial}{\partial x_{1}} \\
\frac{\partial}{\partial x_{2}}
\end{array}\right] .
$$

This last vector is the well known Del operator which you would have come across in both Calculus II and E\&M. The approximation (A.2) can then be written as:

$$
f\left(x_{1}, x_{2}\right) \approx f(c, d)+\left.\nabla f\right|_{\mathbf{x}=\mathbf{c}} \cdot(\mathbf{x}-\mathbf{c})
$$

where " ." refers to the standard inner product.

The procedure for linearizing the function $f: \mathbf{R}^{n} \rightarrow \mathbf{R}$ is exactly the same. For notational convenience we use the vector $\mathbf{x} \in \mathbf{R}^{n}$ to denote the variables, and the operator:

$$
\nabla^{T}:=\left[\frac{\partial}{\partial x_{1}} \cdots \frac{\partial}{\partial x_{n}}\right] .
$$

The linear approximation to $f(\mathbf{x})$ is then:

$$
\begin{equation*}
f(\mathbf{x}) \approx f(\mathbf{c})+\left.\nabla f\right|_{\mathbf{x}=\mathbf{c}} \cdot(\mathbf{x}-\mathbf{c}) \tag{A.3}
\end{equation*}
$$

In essence, a linearization is just a fancy term for computing the hyperplane (another fancy word!) tangent to a point.

Example 20. Linearize the equation

$$
f(x, y)=4 x^{3}-6 x^{2} y^{3}+y^{5}
$$

about the point $(x, y)=(1,1)$.
Solution: We compute the following:

$$
f(1,1)=-1,\left.\quad \frac{\partial f}{\partial x}\right|_{(1,1)}=\left.12 x\left(x-y^{3}\right)\right|_{(1,1)}=0
$$

and $\left.\frac{\partial f}{\partial y}\right|_{(1,1)}=\left.y^{2}\left(5 y^{2}-18 x^{2}\right)\right|_{(1,1)}=13$. Then

$$
f(x, y) \approx-1+0(x-1)-13(y-1)=-13 y+12
$$

## A. 3 Linearizing non-linear differential equations.

We know apply our linearization procedure to non-linear differential equations. The key point that we need to keep in mind is that the partial derivatives must be taken with respect to each variable of the differential equation, including the order of the derivatives. For example, suppose that we have a differential equation depending on $y, \dot{y}, \ddot{y}, r$ and $\dot{r}$. We can write this differential equation as:

$$
\begin{equation*}
h(y, \dot{y}, \ddot{y}, r, \dot{r})=0 . \tag{A.4}
\end{equation*}
$$

We define the vector: $\mathbf{x}=[y \dot{y} \ddot{y} r \dot{r}]^{T}$ and write the differential equation as $h(\mathbf{x})=0$.

The next step is to find a point $\mathbf{x}_{0}$ at which we need to linearize $h(\mathbf{x})$. Since this is a differential equation, it only makes sense to linearize about constant solutions. Why? A linearization is an approximation that is only valid around a region close to $\mathbf{x}_{0}$. If the derivatives of the variables in $\mathbf{x}$ are changing, then the variables are not going to stay in that region for long, and so the approximation will not be valid for
very long. ${ }^{1}$ The point $\mathbf{x}_{0}$ is known as an operating point. Also, since $\mathbf{x}_{0}$ is a solution, it is important to remember that $h\left(\mathbf{x}_{0}\right)=0$.

In our example, the vector $\mathrm{x}_{0}$ would be points where

$$
\mathbf{x}_{0}=\left[\begin{array}{lllll}
y_{0} & 0 & 0 & r_{0} & 0
\end{array}\right]^{T} ; \quad \text { with } h\left(\mathbf{x}_{0}\right)=0 .
$$

Using (A.3), we can write

$$
0=h(\mathbf{x}) \approx h\left(\mathbf{x}_{0}\right)+\left.\nabla h\right|_{\mathbf{x}_{0}} \cdot\left(\mathbf{x}-\mathbf{x}_{0}\right)
$$

Note that the first term in the right hand side of the approximation is 0 . It follows that $\left.\nabla h\right|_{\mathbf{x}_{0}} \cdot\left(\mathbf{x}-\mathbf{x}_{0}\right)=0$. Writing

$$
\Delta \mathbf{x}:=\mathbf{x}-\mathbf{x}_{0}=\left[\begin{array}{lll}
\Delta y \Delta \dot{y} \Delta \ddot{y} \Delta r \Delta \dot{r}
\end{array}\right]^{T}
$$

we have that

$$
\begin{aligned}
0 & =\left.\nabla h\right|_{\mathbf{x}_{0}} \cdot\left(\mathbf{x}-\mathbf{x}_{0}\right) \\
& =\left.\frac{\partial h}{\partial y}\right|_{\mathbf{x}_{0}} \Delta y+\left.\frac{\partial h}{\partial \dot{y}}\right|_{\mathbf{x}_{0}} \Delta \dot{y}+\left.\frac{\partial h}{\partial \ddot{y}}\right|_{\mathbf{x}_{0}} \Delta \ddot{y}+\left.\frac{\partial h}{\partial r}\right|_{\mathbf{x}_{0}} \Delta r+\left.\frac{\partial h}{\partial \dot{r}}\right|_{\mathbf{x}_{0}} \Delta \dot{r} .
\end{aligned}
$$

Thus, we have a linear differential equation in terms of the $\Delta y, \Delta r$, etc.
Typically, we will have more than one equation. Suppose that there exist $n$ equations $h_{i}(\mathbf{x})=0$, for $i=1, \ldots, n$. This is handled in the exact same way as above. The only differences are that: a) the operating point has to satisfy all of the equations; b) we have to linearize all of the equations.

Fig 1: Circuit diagram for Example 21.
Example 21. Consider the circuit diagram of Fig. ??, where there is a non-linear term whose output voltage $y(t)$ is given by $y(t)=5 i(t)+20 i^{3}(t)$. Given that the circuit is supposed to operate at a current of 0.1 A , find a linear transfer function relating the output voltage to the input voltage $r(t)$.

[^0]Solution: In this example we have three signals: $i(t), r(t)$ and $y(t)$. There are two equations, the first one relates the output voltage to the current:

$$
\begin{equation*}
y(t)=5 i(t)+20 i^{3}(t) \tag{A.5}
\end{equation*}
$$

the second relates the current to the input voltage. This is given by

$$
\begin{equation*}
0.02 \frac{d i}{\mathrm{~d} t}+y+10=r \tag{A.6}
\end{equation*}
$$

In order to linearize these equations, we must write them in the form of Eq. A.4. The first of these can be written as

$$
\begin{equation*}
h_{1}(y, i)=5 i+20 i^{3}-y \tag{A.7}
\end{equation*}
$$

At the operating point of $i_{0}=.1 \mathrm{~A}$, we have that $y_{0}=0.52 \mathrm{~V}$ and $r_{0}=10.52 \mathrm{~V}$. Using the abbreviation "o.p." for operating point, A.7) can be approximated as

$$
\begin{aligned}
h_{1}(y, i) & \approx h_{1}(0.52,0.1)+\left.\frac{\partial h_{1}}{\partial y}\right|_{o . p .} \Delta y+\left.\frac{\partial h_{1}}{\partial i}\right|_{o . p .} \Delta i \\
0 & =0-\Delta y+5 \Delta i+60 i_{0}^{2} \Delta i=5.6 \Delta i-\Delta y .
\end{aligned}
$$

For the second equation we can write:

$$
\begin{equation*}
h_{2}\left(y, r, \frac{d i}{\mathrm{~d} t}\right)=y+10-r+0.02 \frac{d i}{\mathrm{~d} t} \tag{A.8}
\end{equation*}
$$

Again, using our linearization procedure, we can approximate

$$
\begin{aligned}
h_{2}\left(y, r, \frac{d i}{\mathrm{~d} t}\right) & \approx h_{2}(0.52,10.52,0)+\left.\frac{\partial h_{2}}{\partial y}\right|_{\text {o.p. }} \Delta y+\left.\frac{\partial h_{2}}{\partial r}\right|_{\text {o.p. }} \Delta r+\left.\frac{\partial h_{2}}{\partial \frac{d i}{\mathrm{~d} t}}\right|_{\text {o.p. }} \Delta \frac{d i}{\mathrm{~d} t} \\
0 & =0+\Delta y-\Delta r+0.02 \Delta \frac{d i}{\mathrm{~d} t}
\end{aligned}
$$

Taking Laplace Transforms of both of the linear equations, we get $5.6 \Delta I(s)=$ $\Delta Y(s)$ and $\Delta Y(s)=\Delta R(s)+0.02 s \Delta I(s)$. Eliminating the variable $I(s)$ in these two equations gives the desired transfer function

$$
G_{Y R}(s)=\frac{\Delta Y(s)}{\Delta R(s)}=\frac{280}{s+280}
$$

Fig 2: Inverted pendulum.

We look at one more example:
Example 22. Consider the inverted pendulum on a cart of Fig. ??. Find the transfer function from $\Delta f$ to $\Delta \theta$.

Solution: The first step is to use our knowledge of physics to obtain differential equations for the system. We do this by considering the following two free body diagrams. Note that we have added a force $f_{p}$ which is that which is acting along the pendulum arm.

Consider first the equations for the pendulum. There are two displacements that have to be considered: in the $x$ direction, and in the $y$ direction. The equations are always of the form $F=m a$. Using a little trigonometry, we have:

$$
\begin{aligned}
& x \text { direction: } \quad f_{p} \sin \theta=m_{2} \frac{\mathrm{~d}^{2}}{\mathrm{~d} t^{2}}(x+l \sin \theta) \\
& y \text { direction: } m_{2} g-f_{p} \cos \theta=m_{2} \frac{d^{2}}{\mathrm{~d} t^{2}}(l-l \cos \theta) .
\end{aligned}
$$

For the cart, we have to consider only one direction:

$$
x \text { direction: }-f-f_{p} \sin \theta=m_{1} \frac{d^{2}}{\mathrm{~d} t^{2}}(x)
$$

We note the following two identities:

$$
\frac{d^{2}}{\mathrm{~d} t^{2}}(\sin \theta)=\frac{d}{\mathrm{~d} t}(\dot{\theta} \cos \theta)=\ddot{\theta} \cos \theta-(\dot{\theta})^{2} \sin \theta
$$

and similarly

$$
\frac{d^{2}}{\mathrm{~d} t^{2}}(\cos \theta)=-(\dot{\theta})^{2} \cos \theta-\ddot{\theta} \sin \theta
$$

Substituting these identities into the three equations above gives us the following three non-linear differential equations:

$$
\begin{align*}
f_{p} \sin \theta & =m_{2}\left(\ddot{x}+l \ddot{\theta} \cos \theta-l(\dot{\theta})^{2} \sin \theta\right)  \tag{A.9}\\
m_{2} g-f_{p} \cos \theta & =m_{2} l\left((\dot{\theta})^{2} \cos \theta+\ddot{\theta} \sin \theta\right)  \tag{A.10}\\
-f-f_{p} \sin \theta & =m_{1} \ddot{x} \tag{A.11}
\end{align*}
$$

in 4 unknowns: $x, \theta, f$ and $f_{p}$.
Our next step is to find an operating point. We set all derivatives to 0 and get

$$
\begin{align*}
f_{p_{0}} \sin \theta_{0} & =0  \tag{A.12}\\
m_{2} g-f_{p_{0}} \cos \theta_{0} & =0  \tag{A.13}\\
-f_{0}-f_{p_{0}} \sin \theta_{0} & =0 \tag{A.14}
\end{align*}
$$

From Eq. A.12, we need either $f_{p_{0}}=0$ or $\sin \theta_{0}=0$.
Case 1: $f_{p_{0}}=0$.
This would imply in Eq. A. 13 that $m_{2} g=0$ which is physically impossible.
Case 2:
$\sin \theta_{0}=0$ with $\theta_{0}=0$. In this case, we have $f_{0}=0$ and $f_{p_{0}}=m_{2} g$. Note that $x_{0}$ is arbitrary, which should not be too surprising. Let's linearize about this point. The first Eq. A.9) can be written as

$$
h_{1}\left(\theta, \dot{\theta}, \ddot{\theta}, \ddot{x}, f_{p}\right)=-f_{p} \sin \theta+m_{2}\left(\ddot{x}+l \ddot{\theta} \cos \theta-l(\dot{\theta})^{2} \sin \theta\right)=0 .
$$

The partial derivatives are:

$$
\begin{aligned}
& \left.\frac{\partial h_{1}}{\partial \theta}\right|_{o . p .}=m_{2}\left(-l \ddot{\theta} \sin \theta-l(\dot{\theta})^{2} \cos \theta\right)-\left.f_{p} \cos \theta\right|_{o . p .}=-f_{p_{0}}=-m_{2} g \\
& \left.\frac{\partial h_{1}}{\partial \dot{\theta}}\right|_{o . p .}=-\left.2 m_{2} l \dot{\theta} \sin \theta\right|_{o . p .}=0 ;\left.\quad \frac{\partial h_{1}}{\partial \ddot{\theta}}\right|_{o . p .}=\left.m_{2} l \cos \theta\right|_{o . p .}=m_{2} l \\
& \left.\frac{\partial h_{1}}{\partial \ddot{x}}\right|_{o . p .}=m_{2} ; \text { and }\left.\frac{\partial h_{1}}{\partial f_{p}}\right|_{o . p .}=-\left.\sin \theta\right|_{o . p .}=0
\end{aligned}
$$

Remark:Perhaps the biggest source of confusion arises when partial derivatives have to be taken with respect to different derivatives of a variable; for example $\theta, \dot{\theta}$ and $\ddot{\theta}$ in this example. Note that because it is a partial derivative, $\theta$ and $\dot{\theta}$ are considered to be completely independent.

Our linearized equation is then

$$
m_{2}\left(l \frac{d^{2} \Delta \theta}{\mathrm{~d} t^{2}}-g \Delta \theta+\frac{d^{2} \Delta x}{\mathrm{~d} t^{2}}\right)=0
$$

Do the same thing at home to the two equations $h_{2}$ and $h_{3}$ obtained from Eq. A. 10 and Eq. A. 11 in order to get:

$$
\begin{aligned}
\Delta f_{p} & =0 \\
m_{2} g \Delta \theta+m_{1} \frac{d^{2}}{\mathrm{~d} t^{2}} \Delta x+\Delta f & =0
\end{aligned}
$$

We are left with only two equations. can eliminate $\frac{d^{2} \Delta x}{\mathrm{~d} t^{2}}$ to get one equation in two unknowns:

$$
m_{2} g \Delta \theta+m_{1} g \Delta \theta-m_{1} l \frac{d^{2} \Delta \theta}{\mathrm{~d} t^{2}}+\Delta f=0
$$

Using Laplace Transforms we get

$$
\frac{\Delta \Theta(s)}{\Delta F(s)}=\frac{1}{m_{1} s^{2} l-\left(m_{1}+m_{2}\right) g} .
$$

This transfer function is strictly proper; it is not stable, since the poles are at $\pm \sqrt{\frac{\left(m_{1}+m_{2}\right) g}{m_{1} l}}$.

Case 3: $\sin \theta_{0}=0$ with $\theta_{0}=\pi$. In this case, the pendulum is actually hanging down. (We'll assume that this is physically possible, which, with a well designed system, it is.) This obviously is not the most interesting of control systems, but it will help illustrate a point. Our operating point is now $f_{0}=0$ and $f_{p_{0}}=-m_{2} g$ with $x_{0}$ again arbitrary. Note that the only difference is in the change of sign in $f_{p_{0}}$. Our linearization procedure is almost identical (all the partical derivatives will be the same, the only change is on the operating point). Carrying out all the steps of Case 2, we get the following transfer function:

$$
\frac{\Delta \Theta(s)}{\Delta F(s)}=\frac{-1}{m_{1} s^{2} l+\left(m_{1}+m_{2}\right) g}
$$

This function is also not stable, since the poles are at $\pm j \sqrt{\frac{\left(m_{1}+m_{2}\right) g}{m_{1} l}}$; where $j=$ $\sqrt{-1}$.

Note, however that the poles are just at the boundary. Had we included friction into our model, we would have had poles in the left hand plane.


[^0]:    ${ }^{1}$ This concept may be somewhat confusing at first, particularly if your choice of variables is not the most suitable. For example, in a cruise control system in an automobile, an operating point is a set of conditions such that the velocity, motor speed etc. are constant. Obviously, you could include an equation to include displacement, which would not be constant (unless you are parked!) but this is clearly not what a cruise control system is about.

