Controlling the Angle of a Pendulum Situated on a Cart via Differential Dynamic Programming

Daniel A. Hagen

December 5, 2018

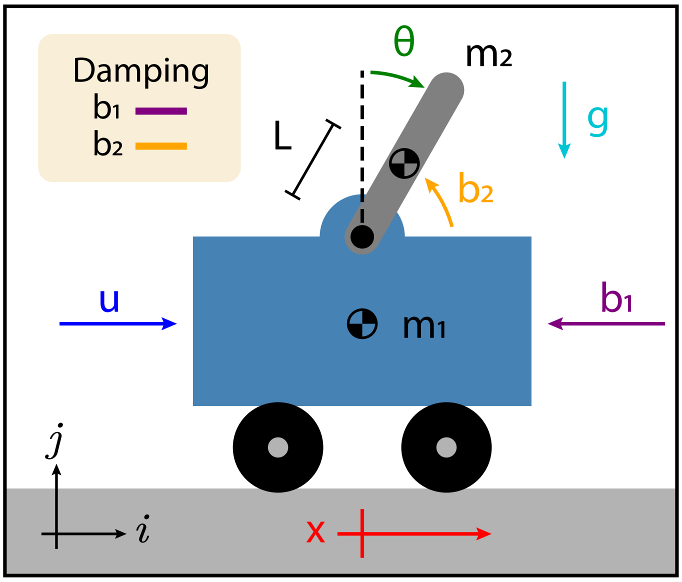

Figure 1: Cart-Pendulum Diagram. The axes in the bottom-left represent the origin frame of reference. The variables x represents the carts position along the i-axis, θ represents the rotation of the pendulum about the $-k$-axis (negative because of the definition of the k-axis, constructed by the right-hand rule), u represents the input of the system (force applied to the cart), b_i represent damping terms for the cart ($i = 1$) and the pendulum ($i = 2$), L represents the distance from the pendulum's joint center to its center of gravity, and g is gravity.

1 Pendulum-Cart Equations of Motion

Consider a cart with mass m_1 with a pendulum of mass m_2 attached above. If h_{wheels} is the height of the wheels, h_{cart} is the height of the cart, L is the distance from the pendulum's joint center to its center of gravity, and x and θ represent the horizontal position and the angle of pendulum at time t, respectively, then we can define position vectors,

$$\vec{r}_1 = \begin{pmatrix} x_o + x \\ h_{wheels} + \frac{1}{2}h_{cart} \end{pmatrix} \tag{1}$$

$$\vec{r}_2 = \vec{r}_1 + \begin{pmatrix} L\sin(\theta) \\ \frac{1}{2}h_{cart} + L\cos(\theta) \end{pmatrix} = \begin{pmatrix} x_o + x(t) + L\sin(\theta) \\ h_{wheels} + h_{cart} + L\cos(\theta) \end{pmatrix} \tag{2}$$

where x_o represents the initial position of the cart along the i-axis. The kinetic energy of the system, T, can therefore be defined as,

$$
\begin{aligned}
T &= \frac{1}{2} m_1 ||\dot{\vec{r}}_1||_2^2 + \frac{1}{2} m_2 ||\dot{\vec{r}}_2||_2^2 \\
&= \frac{1}{2} m_1 \left|\left| \begin{pmatrix} \dot{x} \\ 0 \end{pmatrix} \right|\right|_2^2 + \frac{1}{2} m_2 \left|\left| \begin{pmatrix} \dot{x} + L\cos(\theta)\dot{\theta} \\ -L\sin(\theta)\dot{\theta} \end{pmatrix} \right|\right|_2^2 \\
&= \frac{1}{2} m_1 \dot{x}^2 + \frac{1}{2} m_2 \left(\left(\dot{x} + L\cos(\theta)\dot{\theta} \right)^2 + \left(-L\sin(\theta)\dot{\theta} \right)^2 \right) \\
&= \frac{1}{2}(m_1 + m_2)\dot{x}^2 + m_2 \dot{x} \left(L\cos(\theta)\dot{\theta} \right) + \frac{1}{2} m_2 L^2 \dot{\theta}^2
\end{aligned}
\tag{3}
$$

and the potential energy can be defined as,

$$
U = U_{|\pi/2|} + m_2 g L \cos(\theta) \tag{4}
$$

where $U_{|\pi/2|}$ is the potential energy associated with the $\theta = \pm\pi/2$ (i.e., a constant). Therefore the *Lagrangian* of this system can be written as,

$$
\begin{aligned}
\mathcal{L} &= T - U \\
&= \frac{1}{2}(m_1 + m_2)\dot{x}^2 + m_2 \dot{x} \left(L\cos(\theta)\dot{\theta} \right) + \frac{1}{2} m_2 L^2 \dot{\theta}^2 - U_{|\pi/2|} - m_2 g L \cos(\theta)
\end{aligned}
\tag{5}
$$

The Lagrange's Equations can then be calculated as,

$$
\begin{aligned}
u - b_1 \dot{x} &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) - \frac{\partial \mathcal{L}}{\partial x} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \left((m_1 + m_2)\dot{x} + m_2 L \cos(\theta)\dot{\theta} \right) - 0 \\
&= (m_1 + m_2)\ddot{x} - m_2 L \sin(\theta)\dot{\theta}^2 + m_2 L \cos(\theta)\ddot{\theta}
\end{aligned}
\tag{6}
$$

$$
\begin{aligned}
-b_2 \dot{\theta} &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) - \frac{\partial \mathcal{L}}{\partial \theta} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \left(m_2 L \cos(\theta)\dot{x} + m_2 L^2 \dot{\theta} \right) - \left(-m_2 \dot{x} \left(L\sin(\theta)\dot{\theta} \right) + m_2 g L \sin(\theta) \right) \\
&= -m_2 \dot{x} \left(L\sin(\theta)\dot{\theta} \right) + m_2 L \cos(\theta)\ddot{x} + m_2 L^2 \ddot{\theta} + m_2 \dot{x} \left(L\sin(\theta)\dot{\theta} \right) - m_2 g L \sin(\theta) \\
&= m_2 L \cos(\theta)\ddot{x} + m_2 L^2 \ddot{\theta} - m_2 g L \sin(\theta)
\end{aligned}
\tag{7}
$$

This yields the system,

$$
\begin{cases}
(m_1 + m_2)\ddot{x} + m_2 L \cos(\theta)\ddot{\theta} = m_2 L \sin(\theta)\dot{\theta}^2 - b_1 \dot{x} + u & \text{(8a)} \\
m_2 L \cos(\theta)\ddot{x} + m_2 L^2 \ddot{\theta} = m_2 g L \sin(\theta) - b_2 \dot{\theta} & \text{(8b)}
\end{cases}
$$

which can be rewritten as,

$$
\begin{bmatrix} (m_1 + m_2) & m_2 L \cos(\theta) \\ m_2 L \cos(\theta) & m_2 L^2 \end{bmatrix} \begin{pmatrix} \ddot{x} \\ \ddot{\theta} \end{pmatrix} = \begin{pmatrix} m_2 L \sin(\theta)\dot{\theta}^2 - b_1 \dot{x} + u \\ m_2 g L \sin(\theta) - b_2 \dot{\theta} \end{pmatrix} \tag{9}
$$

Therefore the equations of motion can be written as,

$$
\begin{aligned}
\begin{pmatrix}\ddot{x}\\ \ddot{\theta}\end{pmatrix} &= \begin{bmatrix}(m_1+m_2) & m_2L\cos(\theta)\\ m_2L\cos(\theta) & m_2L^2\end{bmatrix}^{-1}\begin{pmatrix}m_2L\sin(\theta)\dot{\theta}^2 - b_1\dot{x} + u\\ m_2gL\sin(\theta) - b_2\dot{\theta}\end{pmatrix}\\
&= \tfrac{1}{(m_1+m_2)m_2L^2 - m_2^2L^2\cos^2(\theta)}\begin{bmatrix}m_2L^2 & -m_2L\cos(\theta)\\ -m_2L\cos(\theta) & (m_1+m_2)\end{bmatrix}\begin{pmatrix}m_2L\sin(\theta)\dot{\theta}^2 - b_1\dot{x} + u\\ m_2gL\sin(\theta) - b_2\dot{\theta}\end{pmatrix}\\
&= \tfrac{1}{m_2L^2(m_1+m_2\sin^2(\theta))}\begin{bmatrix}m_2L^2 & -m_2L\cos(\theta)\\ -m_2L\cos(\theta) & (m_1+m_2)\end{bmatrix}\begin{pmatrix}m_2L\sin(\theta)\dot{\theta}^2 - b_1\dot{x} + u\\ m_2gL\sin(\theta) - b_2\dot{\theta}\end{pmatrix}\\
&= \tfrac{1}{m_2L^2(m_1+m_2\sin^2(\theta))}\begin{pmatrix}m_2L^2\left(m_2L\sin(\theta)\dot{\theta}^2 - b_1\dot{x} + u\right) - m_2L\cos(\theta)\left(m_2gL\sin(\theta) - b_2\dot{\theta}\right)\\ -m_2L\cos(\theta)\left(m_2L\sin(\theta)\dot{\theta}^2 - b_1\dot{x} + u\right) + (m_1+m_2)\left(m_2gL\sin(\theta) - b_2\dot{\theta}\right)\end{pmatrix}\\
&= \tfrac{1}{m_1+m_2\sin^2(\theta)}\begin{pmatrix}m_2L\sin(\theta)\dot{\theta}^2 - b_1\dot{x} - \frac{m_2g}{2}\sin(2\theta) + \frac{b_2}{L}\cos(\theta)\dot{\theta} + u\\ -\frac{m_2}{2}\sin(2\theta)\dot{\theta}^2 + \frac{b_1}{L}\cos(\theta)\dot{x} + \frac{(m_1+m_2)g}{L}\sin(\theta) - \frac{(m_1+m_2)b_2}{m_2L^2}\dot{\theta} - \frac{1}{L}\cos(\theta)u\end{pmatrix}
\end{aligned} \tag{10}
$$

If we define $\vec{x} := (x_1, x_2, x_3, x_4)^T := (x, \theta, \dot{x}, \dot{\theta})^T$, we can rewrite the equations of motion as the state equations,

$$
\begin{cases}
\dot{x}_1 = x_3 & \text{(11a)}\\
\dot{x}_2 = x_4 & \text{(11b)}\\
\dot{x}_3 = \tfrac{1}{m_1+m_2\sin^2(x_2)}\left(m_2L\sin(x_2)x_4^2 - b_1x_3 - \dfrac{m_2g}{2}\sin(2x_2) + \dfrac{b_2}{L}\cos(x_2)x_4 + u\right) & \text{(11c)}\\
\dot{x}_4 = \tfrac{1}{m_1+m_2\sin^2(x_2)}\begin{pmatrix}-\dfrac{m_2}{2}\sin(2x_2)x_4^2 + \dfrac{b_1}{L}\cos(x_2)x_3\\ \quad + \dfrac{(m_1+m_2)g}{L}\sin(x_2) - \dfrac{(m_1+m_2)b_2}{m_2L^2}x_4 - \dfrac{1}{L}\cos(x_2)u\end{pmatrix} & \text{(11d)}
\end{cases}
$$

If $\dot{x}_i := f_i(\vec{x}, u)$ for all $i \in [1, 2, 3, 4]$, then defining $f(\vec{x}, u) := (f_1(\vec{x}, u), f_2(\vec{x}, u), f_3(\vec{x}, u), f_4(\vec{x}, u))^T$, the state dynamics can be rewritten as,

$$\dot{\vec{x}} = f(\vec{x}, u) \tag{12}$$

2 Defining the Cost Function

In order to perform any optimal control over this system, we must first define the cost function over which we wish to optimize. Assume that the controller to seeks to stabilize the pendulum at the unstable vertical position (i.e., $\vec{x}_T = (p_1, 0, 0, 0)$ where $p \in \mathbb{R}$). The cost function can then be defined as,

$$V(\vec{x}(t), t) = \min_u \left[\phi(\vec{x}(t_f), t_f) + \int_t^{t_f} \mathcal{L}(\vec{x}(\tau), u(\tau))d\tau\right] \tag{13}$$

where ϕ is the *terminal cost* and $\mathcal{L}$ is the *running cost.*

The terminal cost will penalizes large distances between the final state, $\vec{x}(t_f)$, and the desired state, $\vec{x}_T$. If we only are concerned with the pendulum reaching a desired angle ($x_{T,2}$) then the terminal cost can be given by *Eq. 14a.* Alternatively, if the controller desires that the pendulum not only reach the desired angle at t_f, but that it be there at zero angular velocity then we can use the terminal cost function given by *Eq. 14b.* Lastly, if the controller also wishes to have the cart position be constant by the end of the trial, then we can use the terminal cost given by *Eq. 14c* (Note: k_2, k_3, and k_4 are *positive* scaling factors).

$$\phi_1(\vec{x}(t_f), t_f) = \frac{k_2}{2}(x_2(t_f) - x_{T,2})^2 \tag{14a}$$

$$\phi_2(\vec{x}(t_f), t_f) = \frac{k_2}{2}(x_2(t_f) - x_{T,2})^2 + \frac{k_4}{2}x_4(t_f)^2 \tag{14b}$$

$$\phi_3(\vec{x}(t_f), t_f) = \frac{k_2}{2}x_2(t_f)^2 + \frac{k_3}{2}x_3(t_f)^2 + \frac{k_4}{2}x_4(t_f)^2 \tag{14c}$$

If we consider the desired state to be $\vec{x}_T = (x_{T,1}, x_{T,2}, 0, 0)^T$ where $x_{T,1}$ can be any cart position and $x_{T,2}$ represents the desired pendulum angle, then any linear quadratic terminal cost can be given by

$$\phi(\vec{x}(t_f), t_f) = \frac{1}{2}\left(\vec{x}(t_f) - \vec{x}_T\right)^T Q_T \left(\vec{x}(t_f) - \vec{x}_T\right) \tag{15a}$$

$$= \frac{1}{2}\left(\vec{x}(t_f) - \vec{x}_T\right)^T \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & k_2 & 0 & 0 \\ 0 & 0 & k_3 & 0 \\ 0 & 0 & 0 & k_4 \end{bmatrix} \left(\vec{x}(t_f) - \vec{x}_T\right) \tag{15b}$$

where $Q_T \in \mathbb{R}^{4\times4}$ is a symmetric matrix whose elements can be chosen to reflect the cost functions given by *Eq. 14* or to represent any other quadratic cost functions. Terminal cost functions do not need to be quadratic and can be chosen to inflict any desired cost/penalty on the system at any state, but for the sake of this derivation we will assume a quadratic terminal cost given by *Eq. 15b*.

The running cost represents the "cost of getting there" and can penalize things like total input energy (*Eq. 16a*), time spend away from the initial cart position (*Eq. 16b*), time not spent at the desired angle (*Eq. 16c*), or time not spent at the desired angular velocity (coincidentally, this also implies a penalty for total angular momentum, *Eq. 16d*). The total running cost can therefore be given any or all of the equations given in *Eq. 16* (Note: d_1, c_1, c_2 and c_4 are *positive* scaling factors).

$$\mathcal{L}_1(\vec{x}(t), u(t)) = \frac{d_1}{2}u(t)^2 \tag{16a}$$

$$\mathcal{L}_2(\vec{x}(t), u(t)) = \frac{c_1}{2}(x_1(t) - x_1(0))^2 \tag{16b}$$

$$\mathcal{L}_3(\vec{x}(t), u(t)) = \frac{c_2}{2}(x_2(t) - x_{T,2})^2 \tag{16c}$$

$$\mathcal{L}_4(\vec{x}(t), u(t)) = \frac{c_4}{2}x_4(t)^2 \tag{16d}$$

If we again only consider quadratic running cost functions (and some arbitrary vector $\vec{x}_R$), we generalize all quadratic cost functions to *Eq. 17a* or we can rewrite the running cost functions from *Eq. 16* in matrix form (*Eq. 17b*).

$$\mathcal{L}(\vec{x}(t), u(t)) = \frac{d_1}{2}u(t)^2 + \frac{1}{2}\left(\vec{x}(t) - \vec{x}_R\right)^T Q_R \left(\vec{x}(t) - \vec{x}_R\right) \tag{17a}$$

$$= \frac{d_1}{2}u(t)^2 + \frac{1}{2}\left(\vec{x}(t) - \begin{pmatrix} x_1(0) \\ x_{T,2} \\ 0 \\ 0 \end{pmatrix}\right)^T \begin{bmatrix} c_1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_4 \end{bmatrix} \left(\vec{x}(t) - \begin{pmatrix} x_1(0) \\ x_{T,2} \\ 0 \\ 0 \end{pmatrix}\right) \tag{17b}$$

For the sake of the derivation, we will assume that the terminal and running costs are given by *Eqs. 15a & 17a*.

3 Differential Dynamic Programming

In order to utilize differential dynamic programming, we assume a an initial action (input) trajectory. We then perform dynamic programming *around* the trajectory (*backward-pass*). We use this to find an update for our input, and then we create a new state trajectory from this new input (*forward-pass*). This process is repeated until convergence is achieved.

3.1 Linearized dynamics

For the dynamics given by *Eq. 12* and some prescribed reference state and action trajectories $(\bar{x}, \bar{u})$, the dynamics are then linearized as follows.

$$
\begin{aligned}
\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} &= f\left(\vec{x}, u\right) \\
&= f\left(\vec{x} + \bar{x} - \bar{x}, u + \bar{u} - \bar{u}\right) \\
&= f\left(\bar{x} + \delta\vec{x}, \bar{u} + \delta u\right) && \text{(where } \delta\vec{x} = \vec{x} - \bar{x} \text{ and } \delta u = u - \bar{u}\text{)} \\
&\approx f(\bar{x}, \bar{u}) + f_{\vec{x}}\,\delta\vec{x} + f_u\,\delta u && \begin{array}{c}\text{(where } f_{\vec{x}} \text{ and } f_u \text{ are the Jacobians} \\ \text{of } f \text{ w.r.t. } \vec{x} \text{ and } u\text{, respectively.)}\end{array} \\
&= \frac{\mathrm{d}\bar{x}}{\mathrm{d}t} + f_{\vec{x}}\,\delta\vec{x} + f_u\,\delta u
\end{aligned}
$$

Note that $f_{\vec{x}} \in \mathbb{R}^{4\times 4}$ and $f_u \in \mathbb{R}^{4\times 1}$. Therefore, moving $\mathrm{d}\bar{x}/\mathrm{d}t$ to the L.H.S. yields,

$$
\begin{aligned}
\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} - \frac{\mathrm{d}\bar{x}}{\mathrm{d}t} &= f_{\vec{x}}\,\delta\vec{x} + f_u\,\delta u \\
\frac{\mathrm{d}}{\mathrm{d}t}\left(\delta\vec{x}\right) &= f_{\vec{x}}\,\delta\vec{x} + f_u\,\delta u \\
&= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} & \frac{\partial f_3}{\partial x_4} \\ 0 & \frac{\partial f_4}{\partial x_2} & \frac{\partial f_4}{\partial x_3} & \frac{\partial f_4}{\partial x_4} \end{bmatrix} \delta\vec{x} + \tfrac{1}{m_1 + m_2\sin^2(x_2)} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -\frac{1}{L}\cos(x_2) \end{bmatrix} \delta u
\end{aligned} \tag{18}
$$

where,

$$
\begin{aligned}
\frac{\partial f_3}{\partial x_2} &= \tfrac{1}{(m_1 + m_2\sin^2(x_2))^2} \left(\begin{array}{l} \frac{\partial}{\partial x_2} \begin{pmatrix} m_2 L\sin(x_2)x_4^2 - \frac{m_2 g}{2}\sin(2x_2) \\ \qquad - b_1 x_3 + u + \frac{b_2}{L}\cos(x_2)x_4 \end{pmatrix} \left(m_1 + m_2\sin^2(x_2)\right) \\ - \begin{pmatrix} m_2 L\sin(x_2)x_4^2 - \frac{m_2 g}{2}\sin(2x_2) \\ \qquad - b_1 x_3 + u + \frac{b_2}{L}\cos(x_2)x_4 \end{pmatrix} \frac{\partial}{\partial x_2}\left(m_1 + m_2\sin^2(x_2)\right) \end{array} \right) \\
&= \tfrac{1}{(m_1 + m_2\sin^2(x_2))^2} \left(\begin{array}{l} \left(m_2 L\cos(x_2)x_4^2 - m_2 g\cos(2x_2) - \frac{b_2}{L}\sin(x_2)x_4\right)\left(m_1 + m_2\sin^2(x_2)\right) \\ - m_2\sin(2x_2) \begin{pmatrix} m_2 L\sin(x_2)x_4^2 - \frac{m_2 g}{2}\sin(2x_2) \\ \qquad - b_1 x_3 + u + \frac{b_2}{L}\cos(x_2)x_4 \end{pmatrix} \end{array} \right)
\end{aligned} \tag{19a}
$$

$$
\frac{\partial f_3}{\partial x_3} = -\frac{b_1}{m_1 + m_2\sin^2(x_2)} \tag{19b}
$$

$$
\frac{\partial f_3}{\partial x_4} = \frac{1}{m_1 + m_2\sin^2(x_2)}\left(2m_2 L\sin(x_2)x_4 + \frac{b_2}{L}\cos(x_2)\right) \tag{19c}
$$

$$\frac{\partial f_4}{\partial x_2} = \frac{1}{(m_1+m_2\sin^2(x_2))^2}\left(\begin{array}{l} \dfrac{\partial}{\partial x_2}\left(\begin{array}{l} -\dfrac{m_2}{2}\sin(2x_2)x_4^2 + \dfrac{(m_1+m_2)g}{L}\sin(x_2) \\ +\dfrac{b_1}{L}\cos(x_2)x_3 - \dfrac{(m_1+m_2)b_2}{m_2L^2}x_4 \\ -\dfrac{1}{L}\cos(x_2)u \end{array}\right)\Big(m_1+m_2\sin^2(x_2)\Big) \\ -\left(\begin{array}{l} -\dfrac{m_2}{2}\sin(2x_2)x_4^2 + \dfrac{(m_1+m_2)g}{L}\sin(x_2) \\ +\dfrac{b_1}{L}\cos(x_2)x_3 - \dfrac{(m_1+m_2)b_2}{m_2L^2}x_4 \\ -\dfrac{1}{L}\cos(x_2)u \end{array}\right)\dfrac{\partial}{\partial x_2}\Big(m_1+m_2\sin^2(x_2)\Big) \end{array}\right) \tag{20a}$$

$$= \frac{1}{(m_1+m_2\sin^2(x_2))^2}\left(\begin{array}{l} \left(\begin{array}{l} -m_2\cos(2x_2)x_4^2 + \dfrac{(m_1+m_2)g}{L}\cos(x_2) \\ -\dfrac{b_1}{L}\sin(x_2)x_3 + \dfrac{1}{L}\sin(x_2)u \end{array}\right)\Big(m_1+m_2\sin^2(x_2)\Big) \\ -m_2\sin(2x_2)\left(\begin{array}{l} -\dfrac{m_2}{2}\sin(2x_2)x_4^2 + \dfrac{(m_1+m_2)g}{L}\sin(x_2) \\ +\dfrac{b_1}{L}\cos(x_2)x_3 - \dfrac{(m_1+m_2)b_2}{m_2L^2}x_4 \\ -\dfrac{1}{L}\cos(x_2)u \end{array}\right) \end{array}\right)$$

$$\frac{\partial f_4}{\partial x_3} = \frac{b_1\cos(x_2)}{L(m_1+m_2\sin^2(x_2))} \tag{20b}$$

$$\frac{\partial f_4}{\partial x_4} = \frac{1}{m_1+m_2\sin^2(x_2)}\left(-m_2\sin(2x_2)x_4 - \frac{(m_1+m_2)b_2}{m_2L^2}\right) \tag{20c}$$

3.2 Discretized Linearized Dynamics

We can now discretize the linearized dynamics by utilizing a forward integration technique, like *forward Euler.*

$$\delta\vec{x}_{k+1} = \delta\vec{x}_k + \frac{\mathrm{d}}{\mathrm{d}t}\Big(\delta\vec{x}_k\Big)\cdot \mathrm{d}t$$

where $\mathrm{d}t = t_{k+1} - t_k$. Plugging in the dynamics yields,

$$\begin{aligned} \delta\vec{x}_{k+1} &= (I + f_{\vec{x}}\,\mathrm{d}t)\,\delta\vec{x}_k + f_u\,\delta u_k\,\mathrm{d}t \\ &= \Phi\,\delta\vec{x}_k + B\,\delta u_k \end{aligned} \tag{21}$$

Therefore we can define the discretized, linearized dynamics matrices as,

$$\begin{cases} \Phi(\bar{x}(t_k),\bar{u}(t_k)) := \begin{bmatrix} 1 & 0 & \mathrm{d}t & 0 \\ 0 & 1 & 0 & \mathrm{d}t \\ 0 & \frac{\partial f_3}{\partial x_2}\mathrm{d}t & 1+\frac{\partial f_3}{\partial x_3}\mathrm{d}t & \frac{\partial f_3}{\partial x_4}\mathrm{d}t \\ 0 & \frac{\partial f_4}{\partial x_2}\mathrm{d}t & \frac{\partial f_4}{\partial x_3}\mathrm{d}t & 1+\frac{\partial f_4}{\partial x_4}\mathrm{d}t \end{bmatrix} & (22a) \\ B(\bar{x}(t_k),\bar{u}(t_k)) := \dfrac{1}{m_1+m_2\sin^2(x_2)}\begin{bmatrix} 0 \\ 0 \\ \mathrm{d}t \\ -\frac{1}{L}\cos(x_2)\mathrm{d}t \end{bmatrix} & (22b) \end{cases}$$

3.3 Discretized Cost Function

Now that the dynamics have been discretized in time and linearized we can employ a similar approach to dynamic programming *along* the trajectory. But first we need to discretize the cost function in time as well. Consider the value function for a given state when t_f is some small timestep ($\mathrm{d}t$) away.

$$
\begin{aligned}
V\left(\vec{x}(t), t\right) &= \min_{u(t)} \left[\int_t^{t+\mathrm{d}t} \mathcal{L}\left(\vec{x}(s), u(s)\right) \mathrm{d}s + V\left(\vec{x}(t+\mathrm{d}t), t+\mathrm{d}t\right) \right] \\
&\approx \min_{u(t)} \Big[\mathcal{L}(\vec{x}(t), u(t))\, \mathrm{d}t + V\left(\vec{x}(t+\mathrm{d}t), t+\mathrm{d}t\right) \Big]
\end{aligned} \tag{23}
$$

where $\ell(\vec{x}(t), u(t)) := \mathcal{L}(\vec{x}(t), u(t))\, \mathrm{d}t$ is the left Riemann rectangular approximation of the integral. This can be rewritten in discrete time as,

$$
V\left(\vec{x}(t_k), t_k\right) \approx \min_{u(t_k)} \Big[\ell\left(\vec{x}(t_k), u(t_k)\right) + V\left(\vec{x}(t_{k+1}), t_{k+1}\right) \Big] \tag{24}
$$

3.4 Taylor Quadratic Expansion of Value Function

Let $V_{\vec{x}} \in \mathbb{R}^{4\times 1}$ be the gradient of V with respect to $\vec{x}$ and $V_u \in \mathbb{R}$ be the derivative of V w.r.t. u. Let $V_{\vec{x}\vec{x}} \in \mathbb{R}^{4\times 4}$ be the Hessian of V w.r.t. $\vec{x}$, $V_{uu} \in \mathbb{R}$ be the second derivative of V w.r.t. u, and $V_{\vec{x}u} = V_{u\vec{x}}^T \in \mathbb{R}^{4\times 1}$ be the gradient of V_u w.r.t $\vec{x}$.

Additionally, let $\ell_{\vec{x}} \in \mathbb{R}^{4\times 1}$ be the gradient of ℓ w.r.t. $\vec{x}$ and $\ell_u \in \mathbb{R}$ be derivative of ℓ with respect to u. Let $\ell_{\vec{x}\vec{x}} \in \mathbb{R}^{4\times 4}$ be the Hessian of ℓ w.r.t. $\vec{x}$, $\ell_{uu} \in \mathbb{R}$ be the second derivative of ℓ w.r.t. u, and $\ell_{\vec{x}u} = \ell_{u\vec{x}}^T \in \mathbb{R}^{4\times 1}$ be the gradient of ℓ_u w.r.t. $\vec{x}$. Therefore, we can expand the discrete value function as,

$$
\begin{aligned}
V\left(\vec{x}(t_k), t_k\right) &\approx \min_{u(t_k)} \Big[\ell\left(\vec{x}(t_k), u(t_k)\right) + V\left(\vec{x}(t_{k+1}), t_{k+1}\right) \Big] \\
&= \min_{u(t_k)} \begin{bmatrix} \ell\left(\vec{x}(t_k) - \bar{x}(t_k) + \bar{x}(t_k), u(t_k) - \bar{u}(t_k) + \bar{u}(t_k)\right) \\ + V\left(\vec{x}(t_{k+1}) - \bar{x}(t_{k+1}) + \bar{x}(t_{k+1}), t_{k+1}\right) \end{bmatrix} \\
&= \min_{u(t_k)} \begin{bmatrix} \ell\left(\bar{x}(t_k) + \delta\vec{x}(t_k), \bar{u}(t_k) + \delta u(t_k)\right) \\ + V\left(\bar{x}(t_{k+1}) + \delta\vec{x}(t_{k+1}), t_{k+1}\right) \end{bmatrix} \\
&= \min_{u(t_k)} \begin{bmatrix} \ell\left(\bar{x}(t_k), \bar{u}(t_k)\right) + \begin{bmatrix} \ell_{\vec{x}}^T & \ell_u^T \end{bmatrix} \begin{bmatrix} \delta\vec{x}(t_k) \\ \delta u(t_k) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \delta\vec{x}(t_k)^T & \delta u(t_k)^T \end{bmatrix} \begin{bmatrix} \ell\vec{x}\vec{x} & \ell\vec{x}u \\ \ell u\vec{x} & \ell uu \end{bmatrix} \begin{bmatrix} \delta\vec{x}(t_k) \\ \delta u(t_k) \end{bmatrix} \\
+ V\left(\bar{x}(t_{k+1}), t_{k+1}\right) + V_{\vec{x}}\left(\bar{x}(t_{k+1}), t_{k+1}\right)^T \delta\vec{x}(t_{k+1}) \\
+ \frac{1}{2}\delta\vec{x}(t_{k+1})^T\, V_{\vec{x}\vec{x}}\left(\bar{x}(t_{k+1}), t_{k+1}\right)\, \delta\vec{x}(t_{k+1}) \end{bmatrix} \\
&= \min_{\delta u(t_k)} \begin{bmatrix} \ell\left(\bar{x}(t_k), \bar{u}(t_k)\right) + \begin{bmatrix} \ell_{\vec{x}}^T & \ell_u^T \end{bmatrix} \begin{bmatrix} \delta\vec{x}(t_k) \\ \delta u(t_k) \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \delta\vec{x}(t_k)^T & \delta u(t_k)^T \end{bmatrix} \begin{bmatrix} \ell\vec{x}\vec{x} & \ell\vec{x}u \\ \ell u\vec{x} & \ell uu \end{bmatrix} \begin{bmatrix} \delta\vec{x}(t_k) \\ \delta u(t_k) \end{bmatrix} \\
+ V\left(\bar{x}(t_{k+1}), t_{k+1}\right) + V_{\vec{x}}\left(\bar{x}(t_{k+1}), t_{k+1}\right)^T \left(\Phi\, \delta\vec{x}(t_k) + B\, \delta u(t_k)\right) \\
+ \frac{1}{2}((\Phi\, \delta\vec{x}(t_k) + B\, \delta u(t_k)))^T\, V_{\vec{x}\vec{x}}\left(\bar{x}(t_{k+1}), t_{k+1}\right)\, \left(\Phi\, \delta\vec{x}(t_k) + B\, \delta u(t_k)\right) \end{bmatrix}
\end{aligned}
$$

We change the variable that we minimize over from $u(t_k)$ to $\delta u(t_k)$ because, from the definition, $\delta u(t_k) = u(t_k) - \bar{u}(t_k)$ where $\bar{u}(t_k)$ is constant for a given instant. Therefore minimizing over $u(t_k)$ is equivalent to minimizing over $\delta u(t_k)$. We could solve this minimization by then taking the gradient of the cost function with respect to $\delta u(t_k)$, but this is rather tedious. Instead let's define a new function Q as,

$$
Q(\vec{x}(t_k), u(t_k)) := \ell(\vec{x}(t_k), u(t_k)) + V(\vec{x}(t_{k+1}), u(t_{k+1})) \tag{25}
$$

Such that,

$$
\begin{aligned}
Q(\vec{x}(t_k), u(t_k)) &= Q\left(\vec{x}(t_k) - \bar{x}(t_k) + \bar{x}(t_k), u(t_k) - \bar{u}(t_k) + \bar{u}(t_k)\right) \\
&= Q\left(\bar{x}(t_k) + \delta\vec{x}(t_k), \bar{u}(t_k) + \delta u(t_k)\right) \\
&= Q(\bar{x}(t_k), \bar{u}(t_k)) + Q_{\vec{x}}^T \delta\vec{x}(t_k) + Q_u^T \delta u(t_k) \\
&\qquad + \frac{1}{2}\begin{bmatrix}\delta\vec{x}(t_k)^T & \delta u(t_k)^T\end{bmatrix}\begin{bmatrix}Q_{\vec{x}\vec{x}} & Q_{\vec{x}u} \\ Q_{u\vec{x}} & Q_{uu}\end{bmatrix}\begin{bmatrix}\delta\vec{x}(t_k) \\ \delta u(t_k)\end{bmatrix}
\end{aligned} \tag{26}
$$

where $Q_{\vec{x}} \in \mathbb{R}^{4\times 1}$ is the gradient of Q with respect to $\vec{x}$, $Q_u \in \mathbb{R}$ is the derivative of Q w.r.t. u, $Q_{\vec{x}\vec{x}} \in \mathbb{R}^{4\times 4}$ is the Hessian of Q w.r.t. $\vec{x}$, $Q_{uu} \in \mathbb{R}$ is the second derivative of Q w.r.t. u, and $Q_{\vec{x}u} = Q_{u\vec{x}}^T \in \mathbb{R}^{4\times 1}$ is the gradient of Q_u w.r.t $\vec{x}$. Therefore it is possible to define these values of Q in terms of our previous cost function in terms of ℓ and V.

$$
\begin{cases}
Q(\bar{x}(t_k), \bar{u}(t_k)) = \ell(\bar{x}(t_k), \bar{u}(t_k)) + V(\bar{x}(t_{k+1}), t_{k+1}) & \text{(27a)} \\
Q_{\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) = \ell_{\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) + \Phi(\bar{x}(t_k), \bar{u}(t_k))^T V_{\vec{x}}(\bar{x}(t_{k+1}), t_{k+1}) & \text{(27b)} \\
Q_u(\bar{x}(t_k), \bar{u}(t_k)) = \ell_u(\bar{x}(t_k), \bar{u}(t_k)) + B(\bar{x}(t_k), \bar{u}(t_k))^T V_{\vec{x}}(\bar{x}(t_{k+1}), t_{k+1}) & \text{(27c)} \\
Q_{\vec{x}u}(\bar{x}(t_k), \bar{u}(t_k)) = \ell_{\vec{x}u}(\bar{x}(t_k), \bar{u}(t_k)) + \Phi(\bar{x}(t_k), \bar{u}(t_k))^T V_{\vec{x}\vec{x}}(\bar{x}(t_{k+1}), t_{k+1}) B(\bar{x}(t_k), \bar{u}(t_k)) & \text{(27d)} \\
Q_{u\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) = \ell_{u\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) + B(\bar{x}(t_k), \bar{u}(t_k))^T V_{\vec{x}\vec{x}}(\bar{x}(t_{k+1}), t_{k+1}) \Phi(\bar{x}(t_k), \bar{u}(t_k)) & \text{(27e)} \\
Q_{\vec{x}\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) = \ell_{\vec{x}\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) + \Phi(\bar{x}(t_k), \bar{u}(t_k))^T V_{\vec{x}\vec{x}}(\bar{x}(t_{k+1}), t_{k+1}) \Phi(\bar{x}(t_k), \bar{u}(t_k)) & \text{(27f)} \\
Q_{uu}(\bar{x}(t_k), \bar{u}(t_k)) = \ell_{uu}(\bar{x}(t_k), \bar{u}(t_k)) + B(\bar{x}(t_k), \bar{u}(t_k))^T V_{\vec{x}\vec{x}}(\bar{x}(t_{k+1}), t_{k+1}) B(\bar{x}(t_k), \bar{u}(t_k)) & \text{(27g)}
\end{cases}
$$

We can now solve the equivalent problem,

$$
\min_{\delta u(t_k)} Q\left(\vec{x}(t_k), u(t_k)\right)
$$

3.5 Minimization of the Cost Function

To find the minimum of the cost function we take the gradient with respect to $\delta u(t_k)$ and find where it is equal to zero.

$$
\begin{aligned}
&\nabla_{\delta u(t_k)} \left[Q(\bar{x}(t_k), \bar{u}(t_k)) + Q_{\vec{x}}^T \delta\vec{x}(t_k) + Q_u^T \delta u(t_k) + \frac{1}{2}\begin{bmatrix}\delta\vec{x}(t_k)^T & \delta u(t_k)^T\end{bmatrix}\begin{bmatrix}Q_{\vec{x}\vec{x}} & Q_{\vec{x}u} \\ Q_{u\vec{x}} & Q_{uu}\end{bmatrix}\begin{bmatrix}\delta\vec{x}(t_k) \\ \delta u(t_k)\end{bmatrix} \right] \\
&= Q_u + Q_{u\vec{x}}\,\delta\vec{x}(t_k) + Q_{uu}\,\delta u(t_k) = 0
\end{aligned}
$$

Therefore the optimal control for time t_k is given by,

$$
\begin{aligned}
\delta u(t_k) &= -Q_{uu}^{-1}\Big(Q_u + Q_{u\vec{x}}\,\delta\vec{x}(t_k)\Big) \\
&= -Q_{uu}^{-1} Q_{u\vec{x}}\,\delta\vec{x}(t_k) - Q_{uu}^{-1} Q_u
\end{aligned} \tag{28}
$$

3.6 Solving for the values of V, $V_{\vec{x}}$, and $V_{\vec{x}\vec{x}}$

Substituting the values from the optimal input (δu^*) given by DDP back into the value function we find,

$$
\begin{aligned}
V(\vec{x}(t_k),t_k) &\approx \begin{bmatrix} Q(\bar{x}(t_k),\bar{u}(t_k)) + Q_{\vec{x}}^T\delta\vec{x}(t_k) + Q_u^T\delta u^*(t_k) \\ \qquad + \frac{1}{2}\begin{bmatrix}\delta\vec{x}(t_k)^T & \delta u^*(t_k)^T\end{bmatrix}\begin{bmatrix}Q_{\vec{x}\vec{x}} & Q_{\vec{x}u}\\ Q_{u\vec{x}} & Q_{uu}\end{bmatrix}\begin{bmatrix}\delta\vec{x}(t_k)\\ \delta u^*(t_k)\end{bmatrix}\end{bmatrix} \\
&= \begin{bmatrix} Q(\bar{x}(t_k),\bar{u}(t_k)) + Q_{\vec{x}}^T\delta\vec{x}(t_k) + Q_u^T\Big(-Q_{uu}^{-1}Q_u - Q_{uu}^{-1}Q_{u\vec{x}}\,\delta\vec{x}(t_k)\Big) \\ + \frac{1}{2}\begin{bmatrix}\delta\vec{x}(t_k)^T \\ \Big(-Q_{uu}^{-1}Q_u - Q_{uu}^{-1}Q_{u\vec{x}}\,\delta\vec{x}(t_k)\Big)^T\end{bmatrix}^T\begin{bmatrix}Q_{\vec{x}\vec{x}} & Q_{\vec{x}u}\\ Q_{u\vec{x}} & Q_{uu}\end{bmatrix}\begin{bmatrix}\delta\vec{x}(t_k)\\ \Big(-Q_{uu}^{-1}Q_u - Q_{uu}^{-1}Q_{u\vec{x}}\,\delta\vec{x}(t_k)\Big)\end{bmatrix}\end{bmatrix} \\
&= Q(\bar{x}(t_k),\bar{u}(t_k)) + Q_{\vec{x}}^T\delta\vec{x}(t_k) - Q_u^TQ_{uu}^{-1}Q_u - Q_u^TQ_{uu}^{-1}Q_{u\vec{x}}\,\delta\vec{x}(t_k) \\
&\quad + \frac{1}{2}\Bigg(\delta\vec{x}(t_k)^TQ_{\vec{x}\vec{x}}\,\delta\vec{x}(t_k) \\
&\qquad + \Big(-Q_u^TQ_{uu}^{-1} - \delta\vec{x}(t_k)^TQ_{\vec{x}u}Q_{uu}^{-1}\Big)Q_{u\vec{x}}\,\delta\vec{x}(t_k) \\
&\qquad + \delta\vec{x}(t_k)^TQ_{\vec{x}u}\Big(-Q_{uu}^{-1}Q_u - Q_{uu}^{-1}Q_{u\vec{x}}\,\delta\vec{x}(t_k)\Big) \\
&\qquad + \Big(-Q_u^TQ_{uu}^{-1} - \delta\vec{x}(t_k)^TQ_{\vec{x}u}Q_{uu}^{-1}\Big)Q_{uu}\Big(-Q_{uu}^{-1}Q_u - Q_{uu}^{-1}Q_{u\vec{x}}\,\delta\vec{x}(t_k)\Big)\Bigg) \\
&= \Big[Q(\bar{x}(t_k),\bar{u}(t_k)) - \frac{1}{2}Q_u^TQ_{uu}^{-1}Q_u\Big] \\
&\qquad + \Big[Q_{\vec{x}}^T - Q_u^TQ_{uu}^{-1}Q_{u\vec{x}}\Big]\,\delta\vec{x}(t_k) \\
&\qquad + \frac{1}{2}\delta\vec{x}(t_k)^T\Big[Q_{\vec{x}\vec{x}} - Q_{\vec{x}u}Q_{uu}^{-1}Q_{u\vec{x}}\Big]\,\delta\vec{x}(t_k) \\
&\approx V(\bar{x}(t_k),t_k) + V_{\vec{x}}(\bar{x}(t_k),t_k)\,\delta\vec{x}(t_k) + \frac{1}{2}\delta\vec{x}(t_k)^TV_{\vec{x}\vec{x}}(\bar{x}(t_k),t_k)\,\delta\vec{x}(t_k)
\end{aligned}
$$

Therefore, we can set the coefficients of $\delta\vec{x}$ equal to each other obtain a close, discrete approximation of the value function.

$$
\begin{cases}
V(\bar{x}(t_k),t_k) \approx Q(\bar{x}(t_k),\bar{u}(t_k)) - \frac{1}{2}Q_u^TQ_{uu}^{-1}Q_u & \text{(29a)} \\
\qquad = \ell(\bar{x}(t_k),\bar{u}(t_k)) + V(\bar{x}(t_{k+1}),t_{k+1}) - \frac{1}{2}Q_u^TQ_{uu}^{-1}Q_u & \text{(29b)} \\
V_{\vec{x}}(\bar{x}(t_k),t_k) \approx Q_{\vec{x}} - Q_{\vec{x}u}Q_{uu}^{-1}Q_u & \text{(29c)} \\
V_{\vec{x}\vec{x}}(\bar{x}(t_k),t_k) \approx Q_{\vec{x}\vec{x}} - Q_{\vec{x}u}Q_{uu}^{-1}Q_{u\vec{x}} & \text{(29d)}
\end{cases}
$$

4 Computational Work

In order to perform the *backward pass*, we first start with a discrete input array $\bar{u}$. We then calculate the discrete state array, $\bar{x}$, using a forward integration technique like *Forward Euler*. Now that the values of both $\bar{x}$ and $\bar{u}$ are known for a given trial, we can solve for all values of Φ and B (linearized dynamics) as well as all of the values and derivatives of ℓ (discretized integral approximation of the running cost). The equations for Φ and B are given by *Eq. 22*, but we define ℓ and its derivatives as,

$$
\begin{cases}
\ell(\bar{x}(t_k), \bar{u}(t_k)) := \left(\frac{d_1}{2} u(t)^2 + \frac{1}{2} \left(\bar{x}(t_k) - \vec{x}_R\right)^T Q_R \left(\bar{x}(t_k) - \vec{x}_R\right)\right) \mathrm{d}t \in \mathbb{R} & \text{(30a)} \\
\ell_{\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) := Q_R \left(\bar{x}(t_k) - \vec{x}_R\right) \mathrm{d}t \in \mathbb{R}^{4\times 1} & \text{(30b)} \\
\ell_u(\bar{x}(t_k), \bar{u}(t_k)) := d_1 \bar{u}(t_k) \mathrm{d}t \in \mathbb{R} & \text{(30c)} \\
\ell_{\vec{x}u}(\bar{x}(t_k), \bar{u}(t_k)) := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^{2\times 1} & \text{(30d)} \\
\ell_{u\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) := \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{1\times 2} & \text{(30e)} \\
\ell_{\vec{x}\vec{x}}(\bar{x}(t_k), \bar{u}(t_k)) := Q_R \mathrm{d}t \in \mathbb{R}^{4\times 4} & \text{(30f)} \\
\ell_{uu}(\bar{x}(t_k), \bar{u}(t_k)) := d_1 \mathrm{d}t \in \mathbb{R} & \text{(30g)}
\end{cases}
$$

where Q_R is given by *Eq. 17b*.

4.1 Backward Pass

To perform the backwards pass, we start at the final timestep (t_f) and work backwards in time. At t_f we only have the terminal cost as there are no more steps left. Therefore we can initialize $V(\bar{x}(t_f), t_f)$, $V_{\vec{x}}(\bar{x}(t_f), t_f)$, and $V_{\vec{x}\vec{x}}(\bar{x}(t_f), t_f)$ as,

$$
\begin{cases}
V(\bar{x}(t_f), t_f) = \frac{1}{2} \left(\bar{x}(t_f) - \vec{x}_T\right)^T Q_T \left(\bar{x}(t_f) - \vec{x}_T\right) \in \mathbb{R} & \text{(31a)} \\
V_{\vec{x}}(\bar{x}(t_f), t_f) = Q_T \left(\bar{x}(t_f) - \vec{x}_T\right) \in \mathbb{R}^{4\times 1} & \text{(31b)} \\
V_{\vec{x}\vec{x}}(\bar{x}(t_f), t_f) = Q_T \in \mathbb{R}^{4\times 4} & \text{(31c)}
\end{cases}
$$

where Q_T is given by *Eq. 15b*. Utilizing *Eqs. 27* and *30*, we can the solve for the previous timestep's values for V, $V_{\vec{x}}$ and $V_{\vec{x}\vec{x}}$ using *Eq. 29*. This backward iterative process is referred to as the *backward pass.*

4.2 Forward Pass

To then perform the forward pass, we must solve for the optimal values of $\delta u^*(t_k)$. In order to do this, we utilize the discrete linear dynamics given by *Eq. 22* as well as the equation for optimal $\delta u^*(t_k)$ given by *Eq. 28*. If $\delta x(t_o) = (0, 0, 0, 0)^T$, then $\delta u^*(t_o) = -Q_{uu}^{-1} Q_u$ and the remaining values for $\delta x(t_k)$ and $\delta u^*(t_k)$ can be found from,

$$
\begin{cases}
\delta u^*(t_k) = -Q_{uu}^{-1}(\bar{x}(t_k), \bar{u}(t_k)) \left(Q_{u\vec{x}}(\bar{x}(t_k), \bar{u}(t_k))\, \delta\vec{x}(t_k) + Q_u(\bar{x}(t_k), \bar{u}(t_k))\right) & \text{(32a)} \\
\delta\vec{x}(t_{k+1}) = \Phi(\bar{x}(t_k), \bar{u}(t_k))\, \delta\vec{x}(t_k) + B(\bar{x}(t_k), \bar{u}(t_k))\, \delta u^*(t_k) & \text{(32b)}
\end{cases}
$$

Performing this over all of the timesteps is referred to as the *forward pass.* Once the values of δu^* are known for every timestep, a new $\bar{u}$ can be defined as $\bar{u}_{\text{new}} = \bar{u} + \delta u^*$. The DDP process is then repeated for this new input until the cost function converges.

5 Results

Running DDP for this system and the cost function defined in *Eqs. 15b* & *17b* produced the following results.

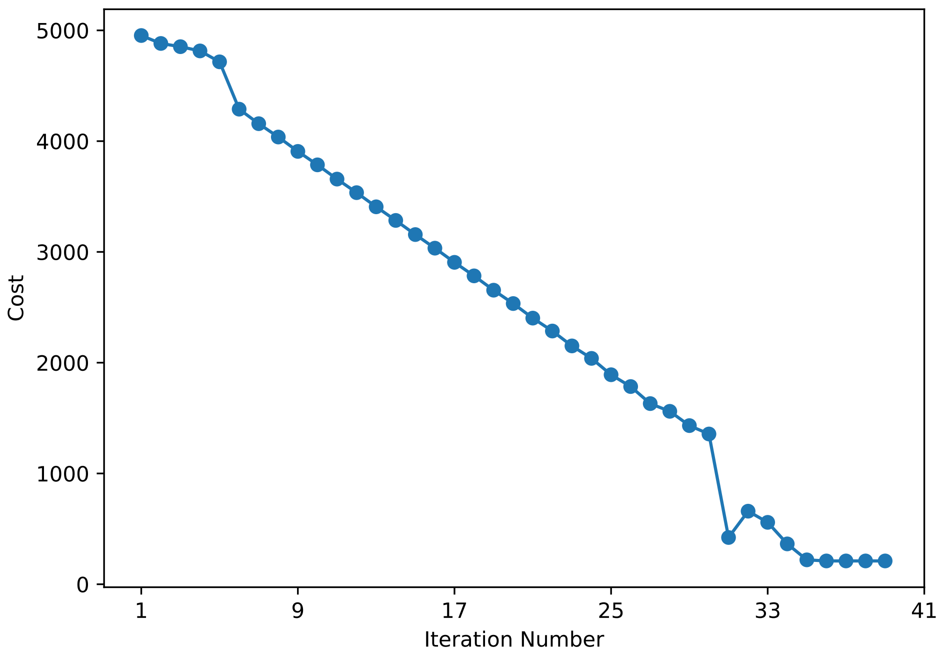

Figure 2: Cost versus iteration number.

To see an animation of the last iteration, Click Here!