

## CHAPTER 3

# The Acrobot and Cart-Pole

### 3.1 INTRODUCTION

A great deal of work in the control of underactuated systems has been done in the context of low-dimensional model systems. These model systems capture the essence of the problem without introducing all of the complexity that is often involved in more real-world examples. In this chapter we will focus on two of the most well-known and well-studied model systems - the Acrobot and the Cart-Pole. These systems are trivially underactuated - both systems have two degrees of freedom, but only a single actuator.

### 3.2 THE ACROBOT

The Acrobot is a planar two-link robotic arm in the vertical plane (working against gravity), with an actuator at the elbow, but no actuator at the shoulder (see Figure 3.1). It was first described in detail in [61]. The companion system, with an actuator at the shoulder but not at the elbow, is known as the Pendubot[76]. The Acrobot is so named because of its resemblance to a gymnast (or acrobat) on a parallel bar, who controls his motion predominantly by effort at the waist (and not effort at the wrist). The most common control task studied for the acrobot is the swing-up task, in which the system must use the elbow (or waist) torque to move the system into a vertical configuration then balance.

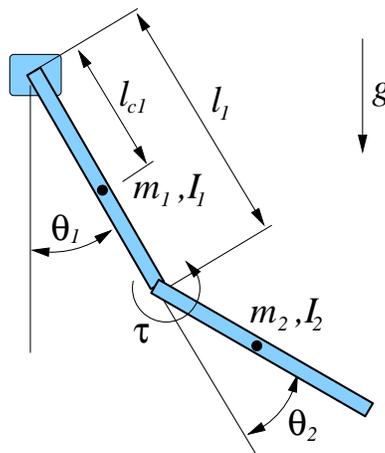


FIGURE 3.1 The Acrobot

The Acrobot is representative of the primary challenge in underactuated robots. In order to swing up and balance the entire system, the controller must reason about and exploit the state-dependent coupling between the actuated degree of freedom and the unactuated degree of freedom. It is also an important system because, as we will see, it

closely resembles one of the simplest models of a walking robot.

### 3.2.1 Equations of Motion

Figure 3.1 illustrates the model parameters used in our analysis.  $\theta_1$  is the shoulder joint angle,  $\theta_2$  is the elbow (relative) joint angle, and we will use  $\mathbf{q} = [\theta_1, \theta_2]^T$ ,  $\mathbf{x} = [\mathbf{q}, \dot{\mathbf{q}}]^T$ . The zero state is the with both links pointed directly down. The moments of inertia,  $I_1, I_2$  are taken about the pivots<sup>1</sup>. The task is to stabilize the unstable fixed point  $\mathbf{x} = [\pi, 0, 0, 0]^T$ .

We will derive the equations of motion for the Acrobot using the method of Lagrange. The kinematics are given by:

$$\mathbf{x}_1 = \begin{bmatrix} l_1 s_1 \\ -l_1 c_1 \end{bmatrix}, \quad \mathbf{x}_2 = \mathbf{x}_1 + \begin{bmatrix} l_2 s_{1+2} \\ -l_2 c_{1+2} \end{bmatrix}. \quad (3.1)$$

The energy<sup>2</sup> is given by:

$$T = T_1 + T_2, \quad T_1 = \frac{1}{2} I_1 \dot{q}_1^2 \quad (3.2)$$

$$T_2 = \frac{1}{2} (m_2 l_1^2 + I_2 + 2m_2 l_1 l_{c2} c_2) \dot{q}_1^2 + \frac{1}{2} I_2 \dot{q}_2^2 + (I_2 + m_2 l_1 l_{c2} c_2) \dot{q}_1 \dot{q}_2 \quad (3.3)$$

$$U = -m_1 g l_{c1} c_1 - m_2 g (l_1 c_1 + l_2 c_{1+2}) \quad (3.4)$$

Entering these quantities into the Lagrangian yields the equations of motion:

$$(I_1 + I_2 + m_2 l_1^2 + 2m_2 l_1 l_{c2} c_2) \ddot{q}_1 + (I_2 + m_2 l_1 l_{c2} c_2) \ddot{q}_2 - 2m_2 l_1 l_{c2} s_2 \dot{q}_1 \dot{q}_2 \quad (3.5)$$

$$-m_2 l_1 l_{c2} s_2 \dot{q}_2^2 + (m_1 l_{c1} + m_2 l_1) g s_1 + m_2 g l_2 s_{1+2} = 0 \quad (3.6)$$

$$(I_2 + m_2 l_1 l_{c2} c_2) \ddot{q}_1 + I_2 \ddot{q}_2 + m_2 l_1 l_{c2} s_2 \dot{q}_1^2 + m_2 g l_2 s_{1+2} = \tau \quad (3.7)$$

In standard, manipulator equation form, we have:

$$\mathbf{H}(\mathbf{q}) = \begin{bmatrix} I_1 + I_2 + m_2 l_1^2 + 2m_2 l_1 l_{c2} c_2 & I_2 + m_2 l_1 l_{c2} c_2 \\ I_2 + m_2 l_1 l_{c2} c_2 & I_2 \end{bmatrix}, \quad (3.8)$$

$$\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} -2m_2 l_1 l_{c2} s_2 \dot{q}_2 & -m_2 l_1 l_{c2} s_2 \dot{q}_2 \\ m_2 l_1 l_{c2} s_2 \dot{q}_1 & 0 \end{bmatrix}, \quad (3.9)$$

$$\mathbf{G}(\mathbf{q}) = \begin{bmatrix} (m_1 l_{c1} + m_2 l_1) g s_1 + m_2 g l_2 s_{1+2} \\ m_2 g l_2 s_{1+2} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (3.10)$$

## 3.3 CART-POLE

The other model system that we will investigate here is the cart-pole system, in which the task is to balance a simple pendulum around its unstable equilibrium, using only horizontal forces on the cart. Balancing the cart-pole system is used in many introductory courses in control, including 6.003 at MIT, because it can be accomplished with simple linear control (e.g. pole placement) techniques. In this chapter we will consider the full swing-up and balance control problem, which requires a full nonlinear control treatment.

<sup>1</sup>[77] uses the center of mass, which differs only by an extra term in each inertia from the parallel axis theorem.

<sup>2</sup>The complicated expression for  $T_2$  can be obtained by (temporarily) assuming the mass in link 2 comes from a discrete set of point masses, and using  $T_2 = \sum_i m_i \dot{\mathbf{r}}_i^T \dot{\mathbf{r}}_i$ , where  $l_i$  is the length along the second link of point  $\mathbf{r}_i$ . Then the expressions  $I_2 = \sum_i m_i l_i^2$  and  $l_{c2} = \frac{\sum_i m_i l_i}{\sum_i m_i}$ , and  $c_1 c_{1+2} + s_1 s_{1+2} = c_2$  can be used to simplify.

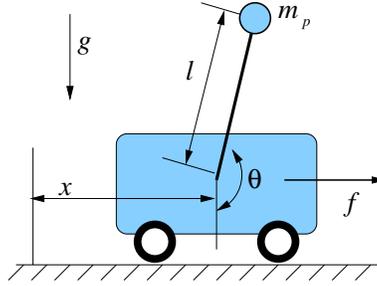


FIGURE 3.2 The Cart-Pole System

Figure 3.2 shows our parameterization of the system.  $x$  is the horizontal position of the cart,  $\theta$  is the counter-clockwise angle of the pendulum (zero is hanging straight down). We will use  $\mathbf{q} = [x, \theta]^T$ , and  $\dot{\mathbf{x}} = [\dot{\mathbf{q}}, \dot{\mathbf{q}}]^T$ . The task is to stabilize the unstable fixed point at  $\mathbf{x} = [0, \pi, 0, 0]^T$ .

### 3.3.1 Equations of Motion

The kinematics of the system are given by

$$\mathbf{x}_1 = \begin{bmatrix} x \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} x + l \sin \theta \\ -l \cos \theta \end{bmatrix}. \quad (3.11)$$

The energy is given by

$$T = \frac{1}{2}(m_c + m_p)\dot{x}^2 + m_p\dot{x}\dot{\theta}l \cos \theta + \frac{1}{2}m_p l^2 \dot{\theta}^2 \quad (3.12)$$

$$U = -m_p g l \cos \theta. \quad (3.13)$$

The Lagrangian yields the equations of motion:

$$(m_c + m_p)\ddot{x} + m_p l \ddot{\theta} \cos \theta - m_p l \dot{\theta}^2 \sin \theta = f \quad (3.14)$$

$$m_p l \ddot{x} \cos \theta + m_p l^2 \ddot{\theta} + m_p g l \sin \theta = 0 \quad (3.15)$$

In standard form, using  $\mathbf{q} = [x, \theta]^T$ ,  $\mathbf{u} = f$ :

$$\mathbf{H}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \mathbf{B}\mathbf{u},$$

where

$$\mathbf{H}(\mathbf{q}) = \begin{bmatrix} m_c + m_p & m_p l \cos \theta \\ m_p l \cos \theta & m_p l^2 \end{bmatrix}, \quad \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} 0 & -m_p l \dot{\theta} \sin \theta \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{G}(\mathbf{q}) = \begin{bmatrix} 0 \\ m_p g l \sin \theta \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

In this case, it is particularly easy to solve directly for the accelerations:

$$\ddot{x} = \frac{1}{m_c + m_p \sin^2 \theta} \left[ f + m_p \sin \theta (l \dot{\theta}^2 + g \cos \theta) \right] \quad (3.16)$$

$$\ddot{\theta} = \frac{1}{l(m_c + m_p \sin^2 \theta)} \left[ -f \cos \theta - m_p l \dot{\theta}^2 \cos \theta \sin \theta - (m_c + m_p)g \sin \theta \right] \quad (3.17)$$

In some of the follow analysis that follows, we will study the form of the equations of motion, ignoring the details, by arbitrarily setting all constants to 1:

$$2\ddot{x} + \ddot{\theta} \cos \theta - \dot{\theta}^2 \sin \theta = f \quad (3.18)$$

$$\ddot{x} \cos \theta + \ddot{\theta} + \sin \theta = 0. \quad (3.19)$$

### 3.4 BALANCING

For both the Acrobot and the Cart-Pole systems, we will begin by designing a linear controller which can balance the system when it begins in the vicinity of the unstable fixed point. To accomplish this, we will linearize the nonlinear equations about the fixed point, examine the controllability of this linear system, then using linear quadratic regulator (LQR) theory to design our feedback controller.

#### 3.4.1 Linearizing the Manipulator Equations

Although the equations of motion of both of these model systems are relatively tractable, the forward dynamics still involve quite a few nonlinear terms that must be considered in any linearization. Let's consider the general problem of linearizing a system described by the manipulator equations.

We can perform linearization around a fixed point,  $(\mathbf{x}^*, \mathbf{u}^*)$ , using a Taylor expansion:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \approx \mathbf{f}(\mathbf{x}^*, \mathbf{u}^*) + \left[ \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right]_{\mathbf{x}=\mathbf{x}^*, \mathbf{u}=\mathbf{u}^*} (\mathbf{x} - \mathbf{x}^*) + \left[ \frac{\partial \mathbf{f}}{\partial \mathbf{u}} \right]_{\mathbf{x}=\mathbf{x}^*, \mathbf{u}=\mathbf{u}^*} (\mathbf{u} - \mathbf{u}^*) \quad (3.20)$$

Let us consider the specific problem of linearizing the manipulator equations around a (stable or unstable) fixed point. In this case,  $\mathbf{f}(\mathbf{x}^*, \mathbf{u}^*)$  is zero, and we are left with the standard linear state-space form:

$$\dot{\mathbf{x}} = \left[ \mathbf{H}^{-1}(\mathbf{q}) [\mathbf{B}\mathbf{u} - \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} - \mathbf{G}(\mathbf{q})] \right], \quad (3.21)$$

$$\approx \mathbf{A}(\mathbf{x} - \mathbf{x}^*) + \mathbf{B}(\mathbf{u} - \mathbf{u}^*), \quad (3.22)$$

where  $\mathbf{A}$ , and  $\mathbf{B}$  are constant matrices. If you prefer, we can also define  $\bar{\mathbf{x}} = \mathbf{x} - \mathbf{x}^*$ ,  $\bar{\mathbf{u}} = \mathbf{u} - \mathbf{u}^*$ , and write

$$\dot{\bar{\mathbf{x}}} = \mathbf{A}\bar{\mathbf{x}} + \mathbf{B}\bar{\mathbf{u}}.$$

Evaluation of the Taylor expansion around a fixed point yields the following, very simple equations, given in block form by:

$$\mathbf{A} = \left[ \begin{array}{cc} \mathbf{0} & \mathbf{I} \\ -\mathbf{H}^{-1} \frac{\partial \mathbf{G}}{\partial \mathbf{q}} & -\mathbf{H}^{-1} \mathbf{C} \end{array} \right]_{\mathbf{x}=\mathbf{x}^*, \mathbf{u}=\mathbf{u}^*} \quad (3.23)$$

$$\mathbf{B} = \left[ \begin{array}{c} \mathbf{0} \\ \mathbf{H}^{-1} \mathbf{B} \end{array} \right]_{\mathbf{x}=\mathbf{x}^*, \mathbf{u}=\mathbf{u}^*} \quad (3.24)$$

Note that the term involving  $\frac{\partial \mathbf{H}^{-1}}{\partial q_i}$  disappears because  $\mathbf{B}\mathbf{u} - \mathbf{C}\dot{\mathbf{q}} - \mathbf{G}$  must be zero at the fixed point. Many of the  $\mathbf{C}\dot{\mathbf{q}}$  derivatives drop out, too, because  $\dot{\mathbf{q}}^* = 0$ .

**Linearization of the Acrobot.**

Linearizing around the (unstable) upright point, we have:

$$\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})_{\mathbf{x}=\mathbf{x}^*} = \mathbf{0}, \quad (3.25)$$

$$\left[ \begin{array}{c} \frac{\partial \mathbf{G}}{\partial \mathbf{q}} \\ \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{q}}} \end{array} \right]_{\mathbf{x}=\mathbf{x}^*} = \left[ \begin{array}{cc} -g(m_1 l_{c1} + m_2 l_1 + m_2 l_2) & -m_2 g l_2 \\ -m_2 g l_2 & -m_2 g l_2 \end{array} \right] \quad (3.26)$$

The linear dynamics follow directly from these equations and the manipulator form of the Acrobot equations.

**Linearization of the Cart-Pole System.**

Linearizing around the (unstable) fixed point in this system, we have:

$$\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})_{\mathbf{x}=\mathbf{x}^*} = \mathbf{0}, \quad \left[ \begin{array}{c} \frac{\partial \mathbf{G}}{\partial \mathbf{q}} \\ \frac{\partial \mathbf{G}}{\partial \dot{\mathbf{q}}} \end{array} \right]_{\mathbf{x}=\mathbf{x}^*} = \left[ \begin{array}{cc} 0 & 0 \\ 0 & -m_p g l \end{array} \right] \quad (3.27)$$

Again, the linear dynamics follow simply.

## 3.4.2 Controllability of Linear Systems

Consider the linear system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u,$$

where  $\mathbf{x}$  has dimension  $n$ . A system of this form is called *controllable* if it is possible to construct an unconstrained control signal which will transfer an initial state to any final state in a finite interval of time,  $0 < t < t_f$  [65]. If every state is controllable, then the system is said to be completely state controllable. Because we can integrate this linear system in closed form, it is possible to derive the exact conditions of complete state controllability.

**The special case of non-repeated eigenvalues.**

Let us first examine a special case, which falls short as a general tool but may be more useful for understanding the intuition of controllability. Let's perform an eigenvalue analysis of the system matrix  $\mathbf{A}$ , so that:

$$\mathbf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i,$$

where  $\lambda_i$  is the  $i$ th eigenvalue, and  $\mathbf{v}_i$  is the corresponding (right) eigenvector. There will be  $n$  eigenvalues for the  $n \times n$  matrix  $\mathbf{A}$ . Collecting the (column) eigenvectors into the matrix  $\mathbf{V}$  and the eigenvalues into a diagonal matrix  $\mathbf{\Lambda}$ , we have

$$\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}.$$

Here comes our primary assumption: let us assume that each of these  $n$  eigenvalues takes on a distinct value (no repeats). With this assumption, it can be shown that the eigenvectors  $\mathbf{v}_i$  form a linearly independent basis set, and therefore  $\mathbf{V}^{-1}$  is well-defined.

We can continue our eigenmodal analysis of the linear system by defining the modal coordinates,  $\mathbf{r}$  with:

$$\mathbf{x} = \mathbf{V}\mathbf{r}, \quad \text{or} \quad \mathbf{r} = \mathbf{V}^{-1}\mathbf{x}.$$

In modal coordinates, the dynamics of the linear system are given by

$$\dot{\mathbf{r}} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}\mathbf{r} + \mathbf{V}^{-1}\mathbf{B}\mathbf{u} = \mathbf{\Lambda}\mathbf{r} + \mathbf{V}^{-1}\mathbf{B}\mathbf{u}.$$

This illustrates the power of modal analysis; in modal coordinates, the dynamics diagonalize yielding independent linear equations:

$$\dot{r}_i = \lambda_i r_i + \sum_j \beta_{ij} u_j, \quad \beta = \mathbf{V}^{-1}\mathbf{B}.$$

Now the concept of controllability becomes clear. Input  $j$  can influence the dynamics in modal coordinate  $i$  if and only if  $\beta_{ij} \neq 0$ . In the special case of non-repeated eigenvalues, having control over each individual eigenmode is sufficient to (in finite-time) regulate all of the eigenmodes[65]. Therefore, we say that the system is controllable if and only if

$$\forall i, \exists j \text{ such that } \beta_{ij} \neq 0.$$

Note a linear feedback to change the eigenvalues of the eigenmodes is not sufficient to accomplish our goal of getting to the goal in finite time. In fact, the open-loop control to reach the goal is easily obtained with a final-value LQR problem5, and (for  $\mathbf{R} = \mathbf{I}$ ) is actually a simple function of the controllability Grammian[21].

#### A general solution.

A more general solution to the controllability issue, which removes our assumption about the eigenvalues, can be obtained by examining the time-domain solution of the linear equations. The solution of this system is

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau.$$

Without loss of generality, lets consider the that the final state of the system is zero. Then we have:

$$\mathbf{x}(0) = - \int_0^{t_f} e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)d\tau.$$

You might be wondering what we mean by  $e^{\mathbf{A}t}$ ; a scalar raised to the power of a matrix..? Recall that  $e^z$  is actually defined by a convergent infinite sum:

$$e^z = 1 + z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \dots$$

The notation  $e^{\mathbf{A}t}$  uses the same definition:

$$e^{\mathbf{A}t} = \mathbf{I} + \mathbf{A}t + \frac{1}{2}(\mathbf{A}t)^2 + \frac{1}{6}(\mathbf{A}t)^3 + \dots$$

Not surprisingly, this has many special forms. For instance,  $e^{\mathbf{A}t} = \mathbf{V}e^{\mathbf{\Lambda}t}\mathbf{V}^{-1}$ , where  $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$  is the eigenvalue decomposition of  $\mathbf{A}$  [82]. The particular form we will use here is

$$e^{-\mathbf{A}\tau} = \sum_{k=0}^{n-1} \alpha_k(\tau)\mathbf{A}^k.$$

This is a particularly surprising form, because the infinite sum above is represented by this finite sum; the derivation uses Sylvester's Theorem[65, 21]. Then we have,

$$\begin{aligned} \mathbf{x}(0) &= - \sum_{k=0}^{n-1} \mathbf{A}^k \mathbf{B} \int_0^{t_f} \alpha_k(\tau) u(\tau) d\tau \\ &= - \sum_{k=0}^{n-1} \mathbf{A}^k \mathbf{B} w_k, \text{ where } w_k = \int_0^{t_f} \alpha_k(\tau) u(\tau) d\tau \\ &= - [\mathbf{B} \quad \mathbf{A}\mathbf{B} \quad \mathbf{A}^2\mathbf{B} \quad \dots \quad \mathbf{A}^{n-1}\mathbf{B}]_{n \times n} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{n-1} \end{bmatrix} \end{aligned}$$

The matrix containing the vectors  $\mathbf{B}, \mathbf{A}\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}$  is called the controllability matrix. In order for the system to be complete-state controllable, for every initial condition  $\mathbf{x}(0)$ , we must be able to find the corresponding vector  $\mathbf{w}$ . This is only possible when the columns of the controllability matrix are linearly independent. Therefore, the condition of controllability is that this controllability matrix is full rank.

Although we only treated the case of a scalar  $u$ , it is possible to extend the analysis to a vector  $\mathbf{u}$  of size  $m$ , yielding the condition

$$\text{rank} [\mathbf{B} \quad \mathbf{A}\mathbf{B} \quad \mathbf{A}^2\mathbf{B} \quad \dots \quad \mathbf{A}^{n-1}\mathbf{B}]_{n \times (nm)} = n.$$

In Matlab<sup>3</sup>, you can obtain the controllability matrix using  $\mathbf{C}_m = \text{ctrb}(A, B)$ , and evaluate its rank with  $\text{rank}(\mathbf{C}_m)$ .

### Controllability vs. Underactuated.

Analysis of the controllability of both the Acrobot and Cart-Pole systems reveals that the linearized dynamics about the upright are, in fact, controllable. This implies that the linearized system, if started away from the zero state, can be returned to the zero state in finite time. This is potentially surprising - after all the systems are underactuated. For example, it is interesting and surprising that the Acrobot can balance itself in the upright position without having a shoulder motor.

The controllability of these model systems demonstrates an extremely important point: *An underactuated system is not necessarily an uncontrollable system.* Underactuated systems cannot follow arbitrary trajectories, but that does not imply that they cannot arrive at arbitrary points in state space. However, the trajectory required to place the system into a particular state may be arbitrarily complex.

The controllability analysis presented here is for LTI systems. A comparable analysis exists for linear time-varying (LTV) systems. One would like to find a comparable analysis for controllability that would apply to nonlinear systems, but I do not know of any general tools for solving this problem.

<sup>3</sup>using the control systems toolbox

### 3.4.3 LQR Feedback

Controllability tells us that a trajectory to the fixed point exists, but does not tell us which one we should take or what control inputs cause it to occur? Why not? There are potentially infinitely many solutions. We have to pick one.

The tools for controller design in linear systems are very advanced. In particular, as we describe in 6, one can easily design an optimal feedback controller for a regulation task like balancing, so long as we are willing to define optimality in terms of a quadratic cost function:

$$J(\mathbf{x}_0) = \int_0^{\infty} [\mathbf{x}(t)^T \mathbf{Q} \mathbf{x}(t) + \mathbf{u}(t)^T \mathbf{R} \mathbf{u}(t)] dt, \quad \mathbf{x}(0) = \mathbf{x}_0, \mathbf{Q} = \mathbf{Q}^T > 0, \mathbf{R} = \mathbf{R}^T > 0.$$

The linear feedback matrix  $\mathbf{K}$  used as

$$\mathbf{u}(t) = -\mathbf{K}\mathbf{x}(t),$$

is the so-called optimal linear quadratic regulator (LQR). Even without understanding the detailed derivation, we can quickly become practitioners of LQR. Conveniently, Matlab has a function,  $\mathbf{K} = \text{lqr}(A, B, Q, R)$ . Therefore, to use LQR, one simply needs to obtain the linearized system dynamics and to define the symmetric positive-definite cost matrices,  $\mathbf{Q}$  and  $\mathbf{R}$ . In their most common form,  $\mathbf{Q}$  and  $\mathbf{R}$  are positive diagonal matrices, where the entries  $Q_{ii}$  penalize the relative errors in state variable  $x_i$  compared to the other state variables, and the entries  $R_{ii}$  penalize actions in  $u_i$ .

Analysis of the close-loop response with LQR feedback shows that the task is indeed completed - and in an impressive manner. Often times the state of the system has to move violently away from the origin in order to ultimately reach the origin. Further inspection reveals the (linearized) closed-loop dynamics have right-half plane zeros - the system in non-minimum phase (acrobot had 3 right-half zeros, cart-pole had 1).

*[To do: Include trajectory example plots here]*

Note that LQR, although it is optimal for the linearized system, is not necessarily the best linear control solution for maximizing basin of attraction of the fixed-point. The theory of *robust control* (e.g., [96]), which explicitly takes into account the differences between the linearized model and the nonlinear model, will produce controllers which outperform our LQR solution in this regard.

## 3.5 PARTIAL FEEDBACK LINEARIZATION

In the introductory chapters, we made the point that the underactuated systems are not feedback linearizable. At least not completely. Although we cannot linearize the full dynamics of the system, it is still possible to linearize a portion of the system dynamics. The technique is called *partial feedback linearization*.

Consider the cart-pole example. The dynamics of the cart are effected by the motions of the pendulum. If we know the model, then it seems quite reasonable to think that we could create a feedback controller which would push the cart in exactly the way necessary to counter-act the dynamic contributions from the pendulum - thereby linearizing the cart dynamics. What we will see, which is potentially more surprising, is that we can also use a feedback law for the cart to feedback linearize the dynamics of the passive pendulum joint.

We'll use the term *collocated* partial feedback linearization to describe a controller which linearizes the dynamics of the actuated joints. What's more surprising is that it is often possible to achieve *noncollocated* partial feedback linearization - a controller which linearizes the dynamics of the unactuated joints. The treatment presented here follows from [78].

### 3.5.1 PFL for the Cart-Pole System

#### **Collocated.**

Starting from equations 3.18 and 3.19, we have

$$\begin{aligned}\ddot{\theta} &= -\dot{x}c - s \\ \ddot{x}(2 - c^2) - sc - \dot{\theta}^2s &= f\end{aligned}$$

Therefore, applying the feedback control law

$$f = (2 - c^2)\ddot{x}^d - sc - \dot{\theta}^2s \quad (3.28)$$

results in

$$\begin{aligned}\ddot{x} &= \ddot{x}^d \\ \ddot{\theta} &= -\ddot{x}^d c - s,\end{aligned}$$

which are valid globally.

#### **Non-collocated.**

Starting again from equations 3.18 and 3.19, we have

$$\begin{aligned}\ddot{x} &= -\frac{\ddot{\theta} + s}{c} \\ \ddot{\theta}(c - \frac{2}{c}) - 2 \tan \theta - \dot{\theta}^2s &= f\end{aligned}$$

Applying the feedback control law

$$f = (c - \frac{2}{c})\ddot{\theta}^d - 2 \tan \theta - \dot{\theta}^2s \quad (3.29)$$

results in

$$\begin{aligned}\ddot{\theta} &= \ddot{\theta}^d \\ \ddot{x} &= -\frac{1}{c}\ddot{\theta}^d - \tan \theta.\end{aligned}$$

Note that this expression is only valid when  $\cos \theta \neq 0$ . This is not surprising, as we know that the force cannot create a torque when the beam is perfectly horizontal.

## 3.5.2 General Form

For systems that are trivially underactuated (torques on some joints, no torques on other joints), we can, without loss of generality, reorganize the joint coordinates in any underactuated system described by the manipulator equations into the form:

$$\mathbf{H}_{11}\ddot{\mathbf{q}}_1 + \mathbf{H}_{12}\ddot{\mathbf{q}}_2 + \phi_1 = 0, \quad (3.30)$$

$$\mathbf{H}_{21}\ddot{\mathbf{q}}_1 + \mathbf{H}_{22}\ddot{\mathbf{q}}_2 + \phi_2 = \tau, \quad (3.31)$$

with  $\mathbf{q} \in \mathbb{R}^n$ ,  $\mathbf{q}_1 \in \mathbb{R}^m$ ,  $\mathbf{q}_2 \in \mathbb{R}^l$ ,  $l = n - m$ .  $\mathbf{q}_1$  represents all of the passive joints, and  $\mathbf{q}_2$  represents all of the actuated joints, and the  $\phi$  terms capture all of the Coriolis and gravitational terms, and

$$\mathbf{H}(\mathbf{q}) = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix}.$$

Fortunately, because  $\mathbf{H}$  is uniformly positive definite,  $\mathbf{H}_{11}$  and  $\mathbf{H}_{22}$  are also positive definite.

**Collocated linearization.**

Performing the same substitutions into the full manipulator equations, we get:

$$\ddot{\mathbf{q}}_1 = -\mathbf{H}_{11}^{-1} [\mathbf{H}_{12}\ddot{\mathbf{q}}_2 + \phi_1] \quad (3.32)$$

$$(\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\mathbf{H}_{12})\ddot{\mathbf{q}}_2 + \phi_2 - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\phi_1 = \tau \quad (3.33)$$

It can be easily shown that the matrix  $(\mathbf{H}_{22} - \mathbf{H}_{21}\mathbf{H}_{11}^{-1}\mathbf{H}_{12})$  is invertible[78]; we can see from inspection that it is symmetric. PFL follows naturally, and is valid globally.

**Non-collocated linearization.**

$$\ddot{\mathbf{q}}_2 = -\mathbf{H}_{12}^+ [\mathbf{H}_{11}\ddot{\mathbf{q}}_1 + \phi_1] \quad (3.34)$$

$$(\mathbf{H}_{21} - \mathbf{H}_{22}\mathbf{H}_{12}^+\mathbf{H}_{11})\ddot{\mathbf{q}}_1 + \phi_2 - \mathbf{H}_{22}\mathbf{H}_{12}^+\phi_1 = \tau \quad (3.35)$$

Where  $\mathbf{H}_{12}^+$  is a Moore-Penrose pseudo-inverse. This inverse provides a unique solution when the rank of  $\mathbf{H}_{12}$  equals  $l$ , the number of passive degrees of freedom in the system (it cannot be more, since the matrix only has  $l$  rows). This rank condition is sometimes called the property of ‘‘Strong Inertial Coupling’’. It is state dependent. Global Strong Inertial Coupling if every state is coupled.

**Task Space Linearization.**

In general, we can define some combination of active and passive joints that we would like to control. This combination is sometimes called a ‘‘task space’’. Consider an output function of the form,

$$\mathbf{y} = \mathbf{f}(\mathbf{q}),$$

with  $\mathbf{y} \in \mathbb{R}^p$ , which defines the task space. Define  $\mathbf{J}_1 = \frac{\partial \mathbf{f}}{\partial \mathbf{q}_1}$ ,  $\mathbf{J}_2 = \frac{\partial \mathbf{f}}{\partial \mathbf{q}_2}$ ,  $\mathbf{J} = [\mathbf{J}_1, \mathbf{J}_2]$ .

**THEOREM 4** (Task Space PFL). If the actuated joints are commanded so that

$$\ddot{\mathbf{q}}_2 = \bar{\mathbf{J}}^+ \left[ \mathbf{v} - \dot{\mathbf{J}}\dot{\mathbf{q}} + \mathbf{J}_1\mathbf{H}_{11}^{-1}\phi_1 \right], \quad (3.36)$$

where  $\bar{\mathbf{J}} = \mathbf{J}_2 - \mathbf{J}_1 \mathbf{H}_{11}^{-1} \mathbf{H}_{12}$ , and  $\bar{\mathbf{J}}^+$  is the right Moore-Penrose pseudo-inverse,

$$\bar{\mathbf{J}}^+ = \bar{\mathbf{J}}^T (\bar{\mathbf{J}} \bar{\mathbf{J}}^T)^{-1},$$

then we have

$$\ddot{\mathbf{y}} = \mathbf{v}. \quad (3.37)$$

subject to

$$\text{rank}(\bar{\mathbf{J}}) = p, \quad (3.38)$$

**Proof.** Differentiating the output function we have

$$\begin{aligned} \dot{\mathbf{y}} &= \mathbf{J} \dot{\mathbf{q}} \\ \ddot{\mathbf{y}} &= \dot{\mathbf{J}} \dot{\mathbf{q}} + \mathbf{J}_1 \ddot{\mathbf{q}}_1 + \mathbf{J}_2 \ddot{\mathbf{q}}_2. \end{aligned}$$

Solving 3.30 for the dynamics of the unactuated joints we have:

$$\ddot{\mathbf{q}}_1 = -\mathbf{H}_{11}^{-1} (\mathbf{H}_{12} \ddot{\mathbf{q}}_2 + \phi_1) \quad (3.39)$$

Substituting, we have

$$\ddot{\mathbf{y}} = \dot{\mathbf{J}} \dot{\mathbf{q}} - \mathbf{J}_1 \mathbf{H}_{11}^{-1} (\mathbf{H}_{12} \ddot{\mathbf{q}}_2 + \phi_1) + \mathbf{J}_2 \ddot{\mathbf{q}}_2 \quad (3.40)$$

$$= \dot{\mathbf{J}} \dot{\mathbf{q}} + \bar{\mathbf{J}} \ddot{\mathbf{q}}_2 - \mathbf{J}_1 \mathbf{H}_{11}^{-1} \phi_1 \quad (3.41)$$

$$= \mathbf{v} \quad (3.42)$$

Note that the last line required the rank condition (3.38) on  $\bar{\mathbf{J}}$  to ensure that the rows of  $\bar{\mathbf{J}}$  are linearly independent, allowing  $\bar{\mathbf{J}} \bar{\mathbf{J}}^+ = \mathbf{I}$ .  $\square$

In order to execute a task space trajectory one could command

$$\mathbf{v} = \ddot{\mathbf{y}}^d + \mathbf{K}_d (\dot{\mathbf{y}}^d - \dot{\mathbf{y}}) + \mathbf{K}_p (\mathbf{y}^d - \mathbf{y}).$$

Assuming the internal dynamics are stable, this yields converging error dynamics,  $(\mathbf{y}_d - \mathbf{y})$ , when  $\mathbf{K}_p, \mathbf{K}_d > 0$  [75]. For a position control robot, the acceleration command of (3.36) suffices. Alternatively, a torque command follows by substituting (3.36) and (3.39) into (3.31).

---

### EXAMPLE 3.1 End-point trajectory following with the Cart-Pole system

Consider the task of trying to track a desired kinematic trajectory with the endpoint of pendulum in the cart-pole system. With one actuator and kinematic constraints, we might be hard-pressed to track a trajectory in both the horizontal and vertical coordinates. But we can at least try to track a trajectory in the vertical position of the end-effector.

Using the task-space PFL derivation, we have:

$$\begin{aligned} y &= f(\mathbf{q}) = -l \cos \theta \\ \dot{y} &= l \dot{\theta} \sin \theta \end{aligned}$$

If we define a desired trajectory:

$$y^d(t) = \frac{l}{2} + \frac{l}{4} \sin(t),$$

then the task-space controller is easily implemented using the derivation above.

---

**Collocated and Non-Collocated PFL from Task Space derivation.**

The task space derivation above provides a convenient generalization of the partial feedback linearization (PFL) [78], which encompasses both the collocated and non-collocated results. If we choose  $\mathbf{y} = \mathbf{q}_2$  (collocated), then we have

$$\mathbf{J}_1 = 0, \mathbf{J}_2 = \mathbf{I}, \dot{\mathbf{J}} = 0, \bar{\mathbf{J}} = \mathbf{I}, \bar{\mathbf{J}}^+ = \mathbf{I}.$$

From this, the command in (3.36) reduces to  $\ddot{\mathbf{q}}_2 = v$ . The torque command is then

$$\boldsymbol{\tau} = -\mathbf{H}_{21}\mathbf{H}_{11}^{-1}(\mathbf{H}_{12}v + \phi_1) + \mathbf{H}_{22}v + \phi_2,$$

and the rank condition (3.38) is always met.

If we choose  $\mathbf{y} = \mathbf{q}_1$  (non-collocated), we have

$$\mathbf{J}_1 = \mathbf{I}, \mathbf{J}_2 = 0, \dot{\mathbf{J}} = 0, \bar{\mathbf{J}} = -\mathbf{H}_{11}^{-1}\mathbf{H}_{12}.$$

The rank condition (3.38) requires that  $\text{rank}(\mathbf{H}_{12}) = l$ , in which case we can write  $\bar{\mathbf{J}}^+ = -\mathbf{H}_{12}^+\mathbf{H}_{11}$ , reducing the rank condition to precisely the ‘‘Strong Inertial Coupling’’ condition described in [78]. Now the command in (3.36) reduces to

$$\ddot{\mathbf{q}}_2 = -\mathbf{H}_{12}^+[\mathbf{H}_{11}\mathbf{v} + \phi_1] \tag{3.43}$$

The torque command is found by substituting  $\ddot{\mathbf{q}}_1 = v$  and (3.43) into (3.31), yielding the same results as in [78].

## 3.6 SWING-UP CONTROL

### 3.6.1 Energy Shaping

Recall the phase portraits that we used to understand the dynamics of the undamped, unactuated, simple pendulum ( $u = b = 0$ ) in section 2.2.2. The orbits of this phase plot were defined by contours of constant energy. One very special orbit, known as a *homoclinic* orbit, is the orbit which passes through the unstable fixed point. In fact, visual inspection will reveal that any state that lies on this homoclinic orbit must pass into the unstable fixed point. Therefore, if we seek to design a nonlinear feedback control policy which drives the simple pendulum from any initial condition to the unstable fixed point, a very reasonable strategy would be to use actuation to regulate the energy of the pendulum to place it on this homoclinic orbit, then allow the system dynamics to carry us to the unstable fixed point.

This idea turns out to be a bit more general than just for the simple pendulum. As we will see, we can use similar concepts of ‘energy shaping’ to produce swing-up controllers for the acrobat and cart-pole systems. It’s important to note that it only takes one actuator to change the total energy of a system.

Although a large variety of swing-up controllers have been proposed for these model systems [30, 5, 94, 79, 54, 12, 61, 49], the energy shaping controllers tend to be the most natural to derive and perhaps the most well-known.

## 3.6.2 Simple Pendulum

Recall the equations of motion for the undamped simple pendulum were given by

$$ml^2\ddot{\theta} + mgl \sin \theta = u.$$

The total energy of the simple pendulum is given by

$$E = \frac{1}{2}ml^2\dot{\theta}^2 - mgl \cos \theta.$$

To understand how to control the energy, observe that

$$\begin{aligned}\dot{E} &= ml^2\dot{\theta}\ddot{\theta} + \dot{\theta}mgl \sin \theta \\ &= \dot{\theta} [u - mgl \sin \theta] + \dot{\theta}mgl \sin \theta \\ &= u\dot{\theta}.\end{aligned}$$

In words, adding energy to the system is simple - simply apply torque in the same direction as  $\dot{\theta}$ . To remove energy, simply apply torque in the opposite direction (e.g., damping).

To drive the system to the homoclinic orbit, we must regulate the energy of the system to a particular desired energy,

$$E^d = mgl.$$

If we define  $\tilde{E} = E - E^d$ , then we have

$$\dot{\tilde{E}} = \dot{E} = u\dot{\theta}.$$

If we apply a feedback controller of the form

$$u = -k\dot{\theta}\tilde{E}, \quad k > 0,$$

then the resulting error dynamics are

$$\dot{\tilde{E}} = -k\dot{\theta}^2\tilde{E}.$$

These error dynamics imply an exponential convergence:

$$\tilde{E} \rightarrow 0,$$

except for states where  $\dot{\theta} = 0$ . The essential property is that when  $E > E^d$ , we should remove energy from the system (damping) and when  $E < E^d$ , we should add energy (negative damping). Even if the control actions are bounded, the convergence is easily preserved.

This is a nonlinear controller that will push all system trajectories to the unstable equilibrium. But does it make the unstable equilibrium locally stable? *No*. Small perturbations may cause the system to drive all of the way around the circle in order to once again return to the unstable equilibrium. For this reason, one trajectories come into the vicinity of our swing-up controller, we prefer to switch to our LQR balancing controller to performance to complete the task.

## 3.6.3 Cart-Pole

Having thought about the swing-up problem for the simple pendulum, let's try to apply the same ideas to the cart-pole system. The basic idea, from [23], is to use collocated PFL to simplify the dynamics, use energy shaping to regulate the pendulum to its homoclinic orbit, then to add a few terms to make sure that the cart stays near the origin. The collocated PFL (when all parameters are set to 1) left us with:

$$\ddot{x} = u \quad (3.44)$$

$$\ddot{\theta} = -uc - s \quad (3.45)$$

The energy of the pendulum (a unit mass, unit length, simple pendulum in unit gravity) is given by:

$$E(\mathbf{x}) = \frac{1}{2}\dot{\theta}^2 - \cos \theta.$$

The desired energy, equivalent to the energy at the desired fixed-point, is

$$E^d = 1.$$

Again defining  $\tilde{E}(\mathbf{x}) = E(\mathbf{x}) - E^d$ , we now observe that

$$\begin{aligned} \dot{\tilde{E}}(\mathbf{x}) &= \dot{E}(\mathbf{x}) = \dot{\theta}\ddot{\theta} + \dot{\theta}s \\ &= \dot{\theta}[-uc - s] + \dot{\theta}s \\ &= -u\dot{\theta}\cos\theta. \end{aligned}$$

Therefore, if we design a controller of the form

$$u = k\dot{\theta}\cos\theta\tilde{E}, \quad k > 0$$

the result is

$$\dot{\tilde{E}} = -k\dot{\theta}^2\cos^2\theta\tilde{E}.$$

This guarantees that  $|\tilde{E}|$  is non-increasing, but isn't quite enough to guarantee that it will go to zero. For example, if  $\theta = \dot{\theta} = 0$ , the system will never move. However, if we have that

$$\int_0^t \dot{\theta}^2(t')\cos^2\theta(t')dt' \rightarrow \infty, \quad \text{as } t \rightarrow \infty,$$

then we have  $\tilde{E}(t) \rightarrow 0$ . This condition is satisfied for all but the most trivial trajectories.

Now we must return to the full system dynamics (which includes the cart). [23] and [80] use the simple pendulum energy controller with an addition PD controller designed to regulate the cart:

$$\ddot{x}^d = k_E\dot{\theta}\cos\theta\tilde{E} - k_px - k_d\dot{x}.$$

[23] provided a proof of convergence for this controller with some nominal parameters.

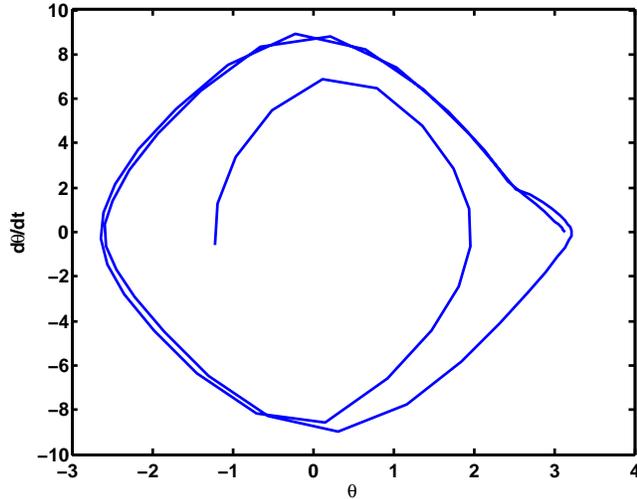


FIGURE 3.3 Cart-Pole Swingup: Example phase plot of the pendulum subsystem using energy shaping control. The controller drives the system to the homoclinic orbit, then switches to an LQR balancing controller near the top.

#### 3.6.4 Acrobot

Swing-up control for the acrobot can be accomplished in much the same way. [79] - pump energy. Clean and simple. No proof. Slightly modified version (uses arctan instead of sat) in [77]. Clearest presentation in [80].

Use collocated PFL. ( $\ddot{q}_2 = \ddot{x}^d$ ).

$$E(\mathbf{x}) = \frac{1}{2} \dot{\mathbf{q}}^T \mathbf{H} \dot{\mathbf{q}} + U(\mathbf{x}).$$

$$E_d = U(\mathbf{x}^*).$$

$$\bar{u} = \dot{q}_1 \tilde{E}.$$

$$\ddot{x}^d = -k_1 q_2 - k_2 \dot{q}_2 + k_3 \bar{u},$$

Extra PD terms prevent proof of asymptotic convergence to homoclinic orbit. Proof of another energy-based controller in [94].

#### 3.6.5 Discussion

The energy shaping controller for swing-up presented here are pretty faithful representatives from the field of nonlinear underactuated control. Typically these control derivations require some clever tricks for simplifying or canceling out terms in the nonlinear equations, then some clever Lyapunov function to prove stability. In these cases, PFL was used to simplify the equations, and therefore the controller design.

These controllers are important, representative, and relevant. But clever tricks with nonlinear equations seem to be fundamentally limited. Most of the rest of the material presented in this book will emphasize more general computational approaches to formulating and solving these and other control problems.

### 3.7 OTHER MODEL SYSTEMS

The acrobot and cart-pole systems are just two of the model systems used heavily in underactuated control research. Other examples include:

- Pendubot
- Inertia wheel pendulum
- Furuta pendulum (horizontal rotation and vertical pend)
- Hovercraft
- Planar VTOL

MIT OpenCourseWare  
<http://ocw.mit.edu>

6.832 Underactuated Robotics  
Spring 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.