Analytical Methods for Network Congestion Control

Steven H. Low
Engineering & Applied Science, Caltech

SYNTHESIS LECTURES ON COMMUNICATION NETWORKS

MORGAN & CLAYPOOL PUBLISHERS

ABSTRACT

The congestion control mechanism has been responsible for maintaining stability as the Internet scaled up in size, speed, traffic volume, coverage, and complexity by many orders of magnitude over the last three decades. In this book we develop a coherent theory of congestion control from the ground up to help understand and design these algorithms. We model network traffic as fluids that flow from sources to destinations and model congestion control algorithms as feedback dynamical systems. We show that the model is well defined, characterize its equilibrium points, and prove their stability. We will use several real protocols for illustration but the emphasis will be on various mathematical techniques for algorithm analysis.

Specifically we are interested in four questions:

1. How to model congestion control algorithms?
2. Are the models well defined?
3. How to characterize the equilibrium points of a congestion control model?
4. How to analyze the stability of these equilibrium points?

For each topic we first present analytical tools, from convex optimization, to control and dynamical systems, Lyapunov and Nyquist stability theorems, and to projection and contraction theorems. We then apply these basic tools to congestion control algorithms and prove rigorously their equilibrium and stability properties. A notable feature of this book is the careful treatment of projected dynamics that introduces discontinuity in our differential equations.

Even though our development is carried out in the context of congestion control, the set of system theoretic tools employed and the process of understanding a physical system, building mathematical models, and analyzing these models for insights have a much wider applicability than to congestion control.

KEYWORDS

Communication networks, congestion control, projected dynamics, convex optimization, network utility maximization, Lyapunov stability, passivity, gradient projection algorithm, contraction mapping, Nyquist stability

Contents

Questions for publisher:[1]

1. Dark background for Matlab generated figures.

2. Indentation immediately after begin{subequations} end{subequations}.

3. Footnote 4 is cited on p.84 but appears on p.86.

4. Trying to quickly browse across the following pages is exceedingly slow (perhaps due to loading of the figures on these pages?): p.42 Figure 2.1 (quick scrolling from p.41 to p.42 to p.43 in PDF sometimes freezes for a while), p.78 Figure 3.1 (most severe), p.88 Figure 3.5 (less severe).

[1] This is a preprint. The issues listed here have been resolved in the print version.

ERRATA

LIST OF CORRECTIONS

- May 29, 2017: Reference [36] in Chapter 1.6.
- June 23, 2017: Chapter 2.2.1.

July 10, 2017: Incorporated the above into PDF marked up from Morgan Claypool, together with many changes in Morgan and Claypool.

The following are not in the Morgan Claypool published version.

- Exercise 2.4: Add citation to Boyd and Vandenberghe 2004 [Exercise 3.6]
- August 16, 2017 (Zhi Low): Conditions C2.1 and C2.3: "zero row" should be "zero column".

REVISION IDEAS

August 6, 2017 (Zhi Low):

- Chapter 2, Exercise 2.9(a): What is the physical meaning of $RR^T p = Rq$? RHS says: Entry l is the sum of *round-trip* prices q_i of all flows i going through link l. LHS says: Entry ll' of RR^T is the number of slows going through both links l and l'. Hence, entry l of LHS is the sum of the price on every link l' that share a flow with link l weighted by the number of such flows. The equality basically equates two ways to count for prices.

  Given any q, assuming $N > L$, there may not be a p that satisfies $R^T p = q$. Algebraically, there is a solution for p if and only if q is in the column space of R^T, in which case p gives the linear combination of the columns of R^T that forms q.

  The simplest example is a single link (with price p) shared by two flows with q_1 and q_2. Give $q := (q_1, q_2)$ there is a solution for p if and only if $q_1 = q_2$ (in which case $p = q_1 = q_2$). If $q_1 \neq q_2$ then the expression $(RR^T)^{-1}Rq = (q_1 + q_2)/2$ is the projection of q onto the column space of R^T.

Preface

PURPOSES OF BOOK

The congestion control mechanism has been responsible for maintaining stability as the Internet scaled up in size, speed, traffic volume, coverage, and complexity by many orders of magnitude over the last three decades. Our primary goal is to develop a coherent theory of Internet congestion control from the ground up to help understand and design the equilibrium and stability properties of large-scale networks under end-to-end control.

In addition, we have two broader purposes in mind. First we wish to introduce a set of system theoretic tools and illustrate their application to concrete problems. Second we wish to demonstrate in depth the entire process of understanding a physical system, building mathematical models of the system, analyzing the models, exploring the practical implications of the analysis, and using the insights to improve a design. Even though our development is carried out in the context of congestion control, these basic analytical tools and the research process are much more broadly applicable.

The Internet, called ARPANet at the time, was born in 1969 with four nodes. The Transmission Control Protocol (TCP) was published by Vinton Cert and Robert Kahn in 1974 [14], split into TCP/IP (Transmission Control Protocol/Internet Protocol) in 1978, and deployed as a standard on the ARPANet by 1983. An Internet congestion collapse was detected in October 1986 on a 32 kbps link between the University of California Berkeley campus and the Lawrence Berkeley National Laboratory that is 400 yards away, during which the throughput dropped by a factor of almost 1,000 to 40 bps. Two years later Van Jacobson implemented and published the congestion control algorithm in the Tahoe version of TCP [26] based on an idea of Raj Jain, K.K. Ramakrishnan and Dah-Ming Chiu [27]. Before Tahoe, there were mechanisms in TCP to prevent senders from overwhelming receivers, but no effective mechanism to prevent the senders from overwhelming the network. This was not an issue because there were few hosts, until the mid-1980s. By November 1986 the number of hosts was estimated to have grown to 5,089 [1], but most of the backbone links have remained 50 – 56 bps (bits per second) since the beginning of the ARPANet. Jacobson's scheme adapts sending rates to the congestion level in the network, thus preventing the senders from overwhelming the network.

Jacobson anticipated even in his original paper [26] the network environments in which his algorithm will perform poorly: "... TCP spans a range from 800 Mbps Cray channels to 1200 bps packet radio links." The algorithm worked very well over a network with relatively low transmission capacity, small delay, and few random packet losses. This was mostly the case in the 1990s, but as the network speed underwent rapid upgrades (see Figure 1), as Internet exploded onto the global scene beyond research and education,

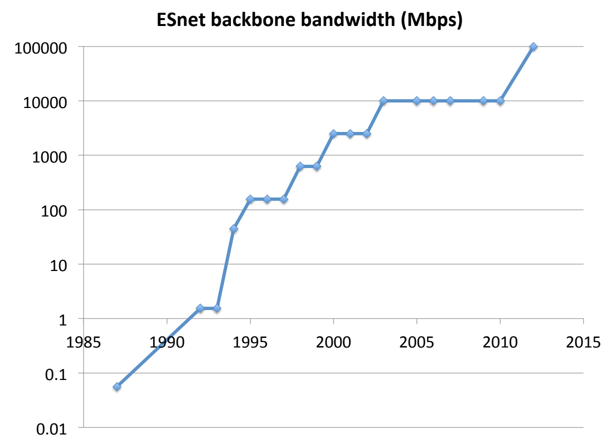


Figure 1: Highest link speed of US Department of Energy's Energy Sciences Network (ESnet) from 1987 (56 kbps) to 2012 (100Gbps) [2].

and as wireless infrastructure was integrated with and mobile services proliferated on the Internet, the strain on the original design started to show. This motivated a flurry of research activities on TCP congestion control in the 1990s. A mathematical understanding of Internet congestion control started in the late 1990s with Frank Kelly's work on network utility maximization [28]. An intensive effort ensued and lasted for a decade to develop a theory to reverse engineer existing algorithms and understand structural properties of large-scale networks under end-to-end congestion control, systematically design new algorithms based on analytical insights, and deploy some of these innovations in the field.

This book is a personal account of that effort, focusing on the theory development.

SUMMARY

We start in Chapter 1 with a summary of classical Internet congestion control protocols. We explain how to model them as dynamical systems using ordinary differential equations:

$$\begin{aligned} \dot{x} &= f(x(t), q(t)), & q(t) &= R^T p(t) \\ \dot{p} &= g(y(t), p(t)), & y(t) &= R\,x(t) \end{aligned}$$

and its variants, where $x(t), q(t) \in \mathbb{R}^N$, $p(t), y(t) \in \mathbb{R}^L$, and $R \in \{0,1\}^{L\times N}$ for a network with N nodes and L links. The graph structure is described by the routing matrix R. The decentralized nature of the system manifests itself in the structure of f and g:

$$\begin{aligned} \dot{x}_i &= f_i(x_i(t), q_i(t)), & q_i(t) &= \sum_l R_{li} p_l(t) \\ \dot{p}_l &= g_l(y_l(t), p_l(t)), & y_l(t) &= \sum_i R_{li} x_i(t) \end{aligned}$$

i.e., each node i (link l) updates its state $x_i(t)$ $(p_l(t))$ based only on local variables $(x_i(t), q_i(t))$ $(y_l(t), p_l(t))$. We prove the existence and uniqueness of solution trajectories to these equations. This ensures that the models are well defined. This class of network models is more general than congestion control and therefore the techniques developed here may be of wider applicability.

We prove in Chapter 2 that the equilibrium point of an arbitrary network under congestion control is the unique optimal solution of a simple convex optimization problem, called network utility maximization. Hence we can interpret congestion control as a distributed algorithm carried out by traffic sources and network resources to maximize utility over the Internet in real time. We explain several implications of this insight.

We present in Chapters 3 through 5 three different methods to study the global asymptotic stability of the equilibrium point, assuming there is no feedback delay. These methods are based on Lyapunov stability theorems, passivity theorems, gradient descent and contraction mapping theorems. The Lyapunov method is the basic tool for proving stability of general nonlinear systems. The passivity method allows one to analyze the stability of an interconnection of multiple dynamical systems in terms of the passivity of the component systems in open loop. The last method treats congestion control as a gradient algorithm for solving the dual of the network utility maximization.

Finally we describe in Chapter 6 the Nyquist stability method for analyzing local stability around the equilibrium point in the presence of feedback delay.

There is a large literature on congestion control and we have not attempted to provide a survey. Pointers are provided at the end of each chapter only to some papers that are directly related to or extend materials covered in that chapter. We present proofs for some, but not all, of these classical results to illustrate techniques or concepts that we find particularly useful.

Many applications, including congestion control, can be modeled by a system of nonlinear differential equations of the form:

$$\dot{x} \quad = \quad (f(x(t))^{+}_{x(t)}$$

where the projection operation $(\cdot)^{+}_{(\cdot)}$ on the right-hand side ensures that the state variable $x(t)$ remains nonnegative. For example $x(t)$ may represent the sending rates of traffic sources or the prices of an economy. The projection introduces discontinuity to the vector field, even when f itself is continuous, and complicates analysis. Analytical models often ignore projection even though nonnegative dynamics is prevalent in reality. A notable feature of this book is the careful treatment of the projected dynamics. In particular we include detailed proofs that extend standard results on the existence, uniqueness, equilibrium and stability properties of smooth unprojected systems to discontinuous projected systems. Some of the stability proofs for congestion control algorithms modeled by projected dynamics in Chapters 3 and 4 are new. As we will see projection mostly preserves these properties.

ACKNOWLEDGMENTS

The idea of this book started when Jean Walrand of Berkeley asked me in early 2004 to write a little book on TCP congestion control for Morgan & Claypool's Synthesis Lectures on Communication Networks of which he was the inaugural Editor. I agreed but did not start writing until the summer of 2010 when I visited Karl Åström and Anders Rantzer at Lund University in Sweden with my extended family. That was a memorable summer! It was also when my research switched from Internet to power systems, so writing again went onto the backburner after the first draft at Lund. Major revisions were done during the summer of 2015 when I visited Janusz Bialek at Skoltech in Russia to give a short course on analytical methods for Internet and power systems, and during spring 2016 when I visited Jiming Chen, Youxian Sun and Zaiyue Yang at Zhejiang University in China. I thank the tremendous encouragement and patience of Jean Walrand and the gentle prodding of the publisher Michael Morgan over more than a decade. It's a relief to have paid my debt. I also thank the warm hospitality of my hosts at Lund University, Skoltech, and Zhejiang University.

This book is a product of our FAST project at Caltech from 2000–2007 and I have learnt a lot from my collaborators, especially John Doyle, Harvey Newman, and Fernando Paganini, and from the first generation of Netlab members, including Lachlan Andrew, Lijun Chen, Cheng Jin, George Lee, Lun Li, Mortada Mehyar, Christine Ortega, A. Kevin Tang, Jiantao Wang, David Wei, Bartek Wydrowski. I thank the US National Science Foundation (especially Darleen Fisher), Army Research Office, Air Force Office of Scientific Research, Cisco, and Caltech's Lee Center for Advanced Networking for their generous financial support. Some of us took the effort to deploy our research in the real world through a startup FastSoft. Since 2014, FastTCP has been accelerating more than 1TB of Internet traffic every second. I experienced first hand the thrill and the challenge in crossing the gap from theory to practice and I thank my colleagues and supporters at FastSoft.

Linqi Guo has worked through the entire draft carefully and corrected numerous errors. I thank him for his meticulous reading and helpful suggestions. Teaching assistants of my networking course (cs/ee 143) at Caltech have contributed some of the exercises, especially Lingwen Gan, Ben Yuan, and Changhong Zhao.

Finally I thank my family, Jenny, Zhi, Zhiyou, my parents and my sister's family for their unwavering support and trust.

S. H. Low
Pasadena, CA, June 2017

NOTATIONS

We collect some of the notational conventions in this book.

Let $\mathbb{R}^n$, $n \geq 1$, be the set of n-dimensional real vectors, $\mathbb{R}^n_+$ the set of n-dimensional nonnegative real vectors, and $\mathbb{R}^{n\times m}$ the set of $n \times m$ real matrices. If x is a vector or matrix then x^T denotes its transpose. By default a vector x is taken to be a *column* vector and can be specified as either

$$x \;=\; \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{or} \quad x \;=\; (x_1, \ldots, x_n) \quad \text{or} \quad x \;=\; (x_i, i = 1, \ldots, n)$$

Inequalities are taken componentwise, i.e., $x \geq 0$ $(x > 0)$ means $x_i \geq 0$ $(x_i > 0)$ for $i = 1, \ldots, n$. If $x_i \in \mathbb{R}^{n_i}$, $i = 1, \ldots, k$, are defined then, unless otherwise specified, x denotes the vector $x := (x_i, i = 1, \ldots, k)$ with dimension $n := \sum_i n_i$. Conversely if a vector x is defined then x_i denotes its ith component in $\mathbb{R}^{n_i}$. Similarly for functions $f_i : \mathbb{R}^{k_i} \to \mathbb{R}^{m_i}$, $i = 1, \ldots, n$, and $f := (f_i, i = 1, \ldots, n) : \mathbb{R}^K \to \mathbb{R}^M$ where $K := \sum_i k_i$ and $M := \sum_i m_i$.

For a scalar function $f : \mathbb{R}^n \to \mathbb{R}$, $\frac{\partial f}{\partial x}$ is the row vector and $\nabla f(x)$ is the column vector, both with components $\frac{\partial f}{\partial x_i}$. For a vector function $f : \mathbb{R}^n \to \mathbb{R}^n$, $\frac{\partial f}{\partial x}$ is the $n \times n$ Jacobian matrix defined by

$$\left[\frac{\partial f}{\partial x}\right]_{ij} \;:=\; \frac{\partial f_i}{\partial x_j}$$

Given a set of utility functions $U_i(x_i) : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, N$, $U_i'(x_i)$ denote their derivatives. We sometimes use $U : \mathbb{R}^N \to \mathbb{R}$ to denote the sum $U(x) := \sum_i U_i(x_i)$. Since U is separable in x_i we use $U'(x)$ to denote the *vector* $U'(x) := (U_i'(x_i),\, i = 1, \ldots, N)$.

For a scalar $a \in \mathbb{R}$, $(a)^+ := \max\{a, 0\}$; for a vector a, $(a)^+$ is defined componentwise, i.e.,

$$(a)^+ \;:=\; \big([a_i]^+,\, \forall i\big)$$

For scalars $a, b \in \mathbb{R}$

$$(a)_b^+ \;:=\; \begin{cases} a & \text{if} \quad a > 0 \;\text{ or }\; b > 0 \\ 0 & \text{otherwise} \end{cases}$$

If $a, b \in \mathbb{R}^n$ are vectors of the same dimension then $(a)_b^+$ is defined componentwise, i.e.,

$$\big[(a)_b^+\big]_i \;:=\; (a_i)_{b_i}^+ \qquad \forall i$$

We use $\|\cdot\|$ to denote an arbitrary norm and $\|x\|_2 := \sqrt{\sum_i x_i^2}$ the Euclidean norm. $B_\delta(x^*) := \{x | \|x - x^*\| \leq \delta\}$ is a *closed* ball around x^* in $\mathbb{R}^n$ unless otherwise specified. $A \subseteq B$ means A is a subset of B and $A \subset B$ means A is a strict subset of B. Given $a_i, i = 1, \ldots, n$, $\mathrm{diag}(a_i, i = 1, \ldots, n)$ denotes the diagonal matrix with a_i as its ith diagonal entry.

CHAPTER 1

Congestion control models

We consider a network under end-to-end congestion control as a deterministic feedback dynamical system described by a set of ordinary differential equations (ODEs). In Chapter 1.1 we present a simple model that treats data packets in the network as fluids that flow from their sources to their destinations. In Chapter 1.2 we describe classical Internet congestion control protocols. In Chapter 1.3 we illustrate how to model these protocols as feedback dynamical systems. In Chapter 1.4 we present general ODE models of congestion control algorithms and discuss limitations and extensions of these models. In Chapter 1.5 we discuss conditions that guarantee the existence and uniqueness of the solution to an ODE model. Finally we show that these conditions are satisfied by the ODE models of congestion control, ensuring that these models are well defined.

1.1 NETWORK MODEL

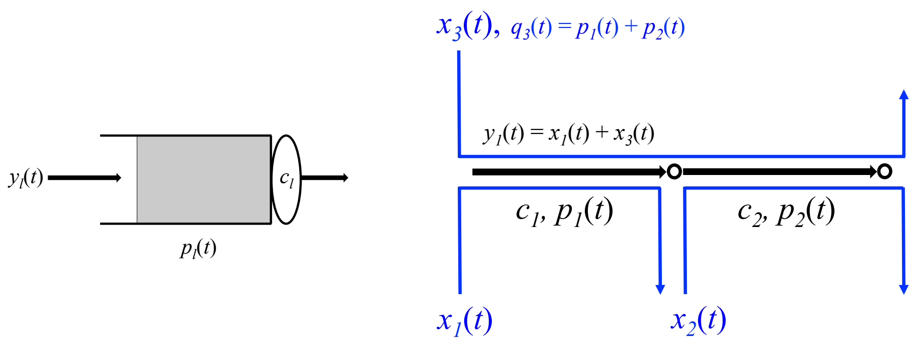
A network is an interconnected set of computing, storage, or communication resources shared by competing users. For our purposes, a user is typically not a human, but a traffic flow from a source to a destination through a subset of these resources. A computing or communication resource is characterized by how fast it can process or transmit information, in units of bits per second or packets per second. A storage resource queues up packets while they wait to be processed or transmitted. We will model each resource abstractly as a "link" that consists of a single server with a buffer (waiting space); see Figure 1.1(a). We often assume that the buffer capacity is infinite. We will call the users "sources" or "flows".

Formally, a network is a set of L links with finite capacities $c = (c_l, l \in L)$ in packets per second (pps). They are shared by a set of N sources. We abuse notation and use L and N to both denote sets and their cardinalities. Each source i uses a set $L_i \subseteq L$ of links. The sets L_i define an $L \times N$ routing matrix R with entries:

$$R_{li} \quad = \quad \begin{cases} 1 & \text{if } l \in L_i \\ 0 & \text{otherwise} \end{cases}$$

We refer to the set L_i of links as source i's path.

Each source i adapts its transmission rate $x_i(t)$ at time t, in pps, according to an algorithm based on some measure of congestion locally observed at source i. This local measure of congestion, denoted by $q_i(t)$ at time t, summarizes the congestion information on the path of source i. Each link l adapts, implicitly or explicitly, a congestion measure



(a) Link model

(b) Network model

Figure 1.1: (a) A link is modeled as a fluid queue with link capacity c_l, infinite buffer size, input rate $y_l(t)$ at time t, and a congestion price $p_l(t)$ at time t. (b) A network is modeled as a collection of links l shared by a set of sources i with sending rates $x_i(t)$ at time t. The routing matrix in this example is $R = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. Source i observes the sum $q_i(t) = \sum_l R_{li} p_l(t)$ of link prices in its path and link l observes the aggregate rate $y_l(t) = \sum_i R_{li} x_i(t)$ from sources sharing the link.

$p_l(t)$ at time t in response to the aggregate input traffic rate $y_l(t)$ locally at link l. Due to its economic interpretation, we call $p_l(t)$ the congestion price at link l or the link price.

We will model below the source algorithms that update the source rates $x_i(t)$ and link algorithms that update link prices $p_l(t)$ by a set of differential and algebraic equations. These local algorithms are interconnected by the routing matrix R that aggregates the link prices $p_l(t)$ on the path of source i into a scalar price $q_i(t)$ observed at source i:

$$q_i(t) \quad := \quad \sum_l R_{li} p_l(t), \qquad i \in N$$

The routing matrix R also aggregates the source rates $x_i(t)$ of flows i that traverse link l into an aggregate flow rate $y_l(t)$ at link l:

$$y_l(t) \quad := \quad \sum_i R_{li} x_i(t), \qquad l \in L$$

These definitions are illustrated in Figure 1.1(b).

1.2 CLASSICAL TCP/AQM PROTOCOLS

To model the algorithms adapting the source rates and link prices, we start by summarizing the basic mechanism of window-based congestion control. We then describe several classical congestion control protocols. In the next section we will present mathematical models of these algorithms.

1.2.1 WINDOW-BASED CONGESTION CONTROL

Transmission Control Protocol (TCP) is one of the transport layer protocols on the Internet (the other being User Datagram Protocol (UDP)). It provides a reliable bit stream end-to-end from a source to a destination over an unreliable datagram service provided by the Internet Protocol (IP) layer. It hides bit errors, packet losses and reordering in the underlying network and delivers to the destination a bit stream that is free of error, loss or duplicate, in the same order it has been sent by the source. A TCP connection operates in three phases. In phase one a virtual circuit is set up between two end points. In phase two they exchange data and acknowledgment packets. Both end points can send data packets, and receive acknowledgments, regardless of which of them initiated the connection. In phase three, the connection is terminated; either end point can initiate the termination. An application that exchanges a large amount of data, e.g., video streaming, usually spends most of its time in phase two. Congestion control is used in phase two to regulate the sending rate of a sender. This is achieved through a *window mechanism*, as we now explain.

Consider a sender that wishes to send a large data file, e.g. a video server transferring a movie to a subscriber. The movie is broken into small chunks and control information is added to each chunk to form a *packet*. Examples of control information include sender and receiver information (IP addresses and TCP port numbers), control bits for error detection, sequence number to detect packet loss and to enable the assembly of the original movie at the destination from received packets. These packets are numbered consecutively and then sent into the network one by one from the source towards the destination.[1] An idealized operation is illustrated in Figure 1.2. When a packet is correctly received at the destination

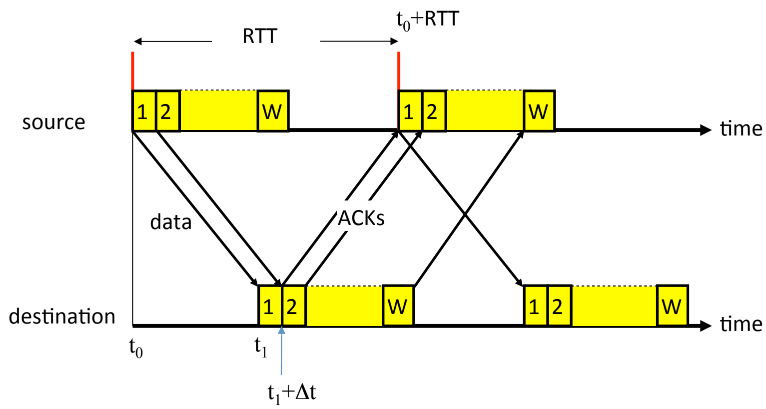


Figure 1.2: Window control mechanism (idealized). Packets are numbered 1, 2, ... W is the *window size*. In the first cycle, the transmission of packet 1 starts at time t_0 and finishes at $t_0 + \Delta t$. The first bit of packet 1 is received at t_1 and the last bit at $t_1 + \Delta t$. Immediately packet 2 starts to be transmitted at time $t_0 + \Delta$ and received at time $t_1 + \Delta$. The ACK for packet 1 is sent at $t_1 + \Delta$ and received at the sender at time t_0+RTT. The cycle then repeats.

an ACK is sent from the destination towards the source. Figure 1.2 assumes an ACK packet has a negligible size compared with a data packet and incurs zero transmission time.[2] The time between sending of a data packet and the arrival of its ACK at the source is called the round-trip time (RTT). If a packet is lost in the network or incurs an excessive delay, the source will not receive its ACK in time and will retransmit the packet after a timeout. This is the basic mechanism to compensate for bit errors and packet losses.

If the original packet was not lost but only delayed, the receiver will receive duplicate packets. The sequence number in the packet will indicate that the packet is a duplicate

[1] TCP is actually a "byte stream" where a packet is identified by the byte numbers the packet contains. This works in the face of packet fragmentation, but it is simpler for our purpose to think of packets being numbered by consecutive integers.

[2] If both end points send and receive data packets (two-way transfer) then each end point can piggybak its ACK in the data packet it sends to the other end. The header of a TCP packet contains both the field `sequence number` that indexes the data packet an end point originates and the field `ack number` that acknowledges the data packet it has received from the other end.

and will be discarded. The receiver will send an ACK for each duplicate packet in case the duplicate packets were sent because their ACKs were lost. Duplicate ACKs at the sender will be discarded.

As shown in Figure 1.2, a source keeps a variable called *window size* W that determines the maximum number of outstanding packets that are allowed to be transmitted but not yet acknowledged. When the window size is exhausted (i.e., the number of outstanding packets reaches W), the source must wait for an ACK before sending a new packet. The new packet can be sent only after the next ACK arrives correctly. In reality packets may incur bit errors, they can be dropped in the network, the RTTs can be random and packets may arrive at the destination out of order. Figure 1.2 ignores these complications and depicts an idealized scenario where all packets are of the same size and arrive at the destination correctly with constant RTT. In this idealized scenario the data transfer process is deterministic and periodic where exactly W packets are sent and acknowledged in each RTT. Moreover, in each RTT, the sender sends packets $1, 2, \ldots, W$ back-to-back, waits for their ACKs to return, and these ACKs trigger the next batch of W packets in the next RTT, and the cycle repeats.

Two features are important for our purpose. The first is the "self-clocking" feature that automatically slows down the source when a network becomes congested and ACKs are delayed. The second is that the window size controls the source sending rate: roughly W packets are sent every RTT (see Figure 1.2). Before Jacobson's proposal in 1988 the window size W was fixed by each TCP connection and the self-clocking feature was the only congestion control mechanism on the Internet. Jacobson's idea is to *dynamically* adapt W to network congestion.

By a *congestion control algorithm* we mean an algorithm that infers congestion and adjusts W during phase two of a TCP connection. Since data transfer can be in both directions such an algorithm can be executed in both directions. We will however fix the direction of a TCP connection in our mathematical model so we do not take into account the interaction of data transfers in both directions between the same pair of end points. Even though TCP congestion control is source-based, a congestion control algorithm actually involves two components: a source algorithm that dynamically adjusts sending rate (or window size) in response to congestion in its path, and a link algorithm that updates, implicitly or explicitly, a congestion measure and sends it back, implicitly or explicitly, to sources that use that link.

On the current Internet, the source algorithm is carried out by TCP, and the link algorithm is carried out by (active) queue management (AQM) schemes such as DropTail or RED in e.g. routers. Different protocols use different metrics to measure congestion, e.g. as we will see below, TCP Reno and its variants use loss probability as a congestion measure, and TCP Vegas and FAST use queueing delay as a congestion measure. Both are implicitly updated at the links and implicitly fed back through end-to-end loss or delay measurements at the sources. The equilibrium and dynamics of the network depend on the TCP/AQM protocol pair.

We now summarize some of the classical TCP protocols (Section 1.2.2) and AQM protocols (Section 1.2.3). We then derive mathematical models of these protocols in Section 1.3. These models will be used in the rest of the book to illustrate various methods to analyze congestion control algorithms.

1.2.2 TCP ALGORITHMS

TCP Tahoe and Reno

The first congestion control algorithm on the Internet was implemented by Jacobson in Tahoe (1988) and Reno (1990) versions of TCP. The window adjustment algorithm is based on an idea of Jain, Ramakrishnan and Chiu that a source should gently probe the network for spare capacity by linearly increasing its window and exponentially reduce its window when congestion is detected. This is called Additive Increase Multiplicative Decrease (AIMD). Congestion is detected when the source detects a packet loss.

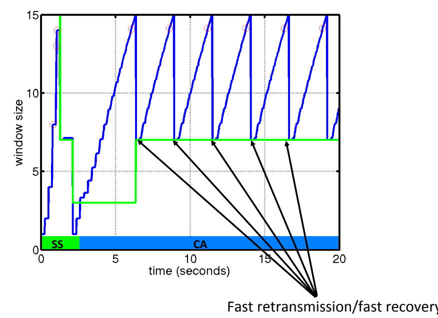
Specifically the Tahoe protocol works as follows. A connection starts cautiously with a small window size of one packet and the source increments its window size by one every time it receives an ACK. This roughly doubles the window size every round-trip time and is called the `slow-start` phase. When the window size reaches a threshold the source enters the `congestion avoidance` phase where it increases its window size by the reciprocal of the current window size every time it receives an acknowledgment. This increases the window size by one packet in each RTT and is referred to as additive increase. The threshold, called the slow-start threshold `ssthreshold`, that determines the transition from `slow-start` to `congestion avoidance` is meant to indicate the available capacity in the network and is adjusted each time a packet loss is detected. On detecting a loss the source sets `ssthreshold` to half of the current window size, retransmits the lost packet, and re-enters `slow-start` by resetting its window size to one packet.

A packet is deemed lost in one of two ways by TCP Tahoe. The first is if the sender does not receive its ACK within a pre-specified time called a timeout period. The second is if the sender receives three duplicate ACKs. Suppose the sender sends packets 0, 1, 2, 3, 4 back to back. Packet 0 is received correctly, packet 1 is lost, and packets 2, 3, 4 are received correctly. The receiver sends an ACK for packet 0. Each packet 2, 3, 4 that arrives at the receiver triggers a duplicate ACK for packet 0, indicating that the receiver is expecting packet 1. When these three duplicate ACKs arrive at the sender, packet 1 is deemed lost. Often, especially when window size is large, a packet loss is detected through duplicate ACKs much sooner than a timeout.

Two refinements were subsequently implemented in TCP Reno to recover from packet losses more efficiently. Call the time from detecting a loss (through three duplicate ACKs) to receiving the ACK for the retransmitted packet the `fast retransmit/fast recover` (fr/fr) phase. In TCP Tahoe the window size is frozen in the `fr/fr` phase. This means that a new packet can be transmitted only a round-trip time later. Moreover the "pipe" from the source to the destination is cleared when the retransmitted packet reaches the receiver and some of the routers in the path may become idle during this period, resulting in a loss of efficiency. The first refinement allows a Reno source to temporarily increment its window size by one on receiving each duplicate ACK while it is in the `fr/fr` phase. The rationale is that each duplicate ACK signals that a packet has left the network. When the window size becomes larger than the number of outstanding packets, a *new* packet can be transmitted in the `fr/fr` phase while the soure is waiting for a (nonduplicate) ACK for the retransmitted packet. The second refinement essentially sets the window size at the end of the `fr/fr` phase to half of the window size when `fr/fr` starts and then enters `congestion avoidance` directly. Hence `slow-start` is entered only rarely in TCP Reno when the connection first starts and when a loss is detected by a timeout rather than three duplicate ACKs.

```
for every ACK {
    W += 1/W       (AI)
}
for every loss {
 W = W/2      (MD)
}
```

(a) Pseudocode



(b) Window trajectory

Figure 1.3: Congestion avoidance behavior of TCP Reno.

The pseudocode for the `congestion avoidance` phase of Reno is shown in Figure 1.3(a) and the resulting window trajectory in Figure 1.3(b).

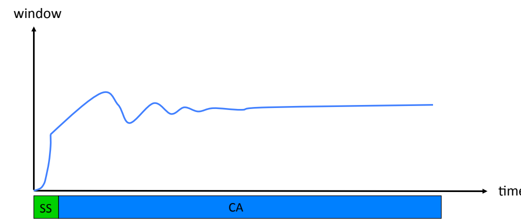
TCP Vegas

TCP Vegas improves upon TCP Reno through three main techniques. The first is a modified retransmission mechanism where timeout is checked on receiving the first duplicate ACK, rather than waiting for the third duplicate ACK (as Reno would), and results in a more timely detection of loss when the clock resolution was low back in the early 1990s (e.g. 500ms). The second technique is a more prudent way to grow the window size during the initial use of `slow-start` when a connection starts up and it results in fewer losses.

The third technique is a new congestion avoidance algorithm that corrects the oscillatory behavior of Reno. The pseudocode is shown in Figure 1.4(a) and the resulting window trajectory in Figure 1.4(b). Here `W` is the current window size, `RTT` is the current

```
for every ACK
{
   if W/RTT_min - W/RTT < α then W += 1/W
   if W/RTT_min - W/RTT > β then W -= 1/W
}
for every loss {
   W = W/2
}
```

(a) Pseudocode



(b) Window trajectory

Figure 1.4: Congestion avoidance behavior of TCP Vegas.

round-trip time, $\texttt{RTT}_{\texttt{min}}$ is the minimum round-trip time observed so far, and $\alpha < \beta$ are protocol parameters in packets/second.

To understand the pseudocode, consider a source i. The round-trip time $T_i(t)$ a packet of source i experiences is the sum of a fixed component d_i and a variable component $q_i(t)$, i.e., $T_i(t) = d_i + q_i(t)$. The fixed component d_i, we call round-trip propagation delay, accounts for signal propagation time and fixed processing time along its path. The variable component $q_i(t)$ accounts for queueing delay at routers in its path. The quantity $\texttt{RTT}_{\texttt{min}}$ ($\min_t T_i(t)$) is an estimate of a source's round-trip propagation delay d_i. Let $w_i(t)$ be source i's window size, $x_i(t)$ be its sending rate, and d_i be its round-trip propagation delay. Then the pseudocode in Figure 1.4(a) increments or decrements the window size by 1 packet according as (multiplying both sides of the conditionals by d_i)

$$w_i(t) \;-\; x_i(t)\,d_i \;<\; \alpha_i\,d_i \quad \text{or} \quad w_i(t) \;-\; x_i(t)\,d_i \;>\; \beta_i\,d_i$$

The quantity $x_i(t)d_i$ represents the number of source i's packets propagating in the communication channel (e.g. fiber optic cables) from source to destination. Hence the quantity $w_i(t) - x_i(t)d_i$ represents the number of i's packets queued at some routers in i's path when there are $w_i(t)$ outstanding packets. A Vegas source estimates the number of its own packets buffered in the path and tries to keep this number between $\alpha := \alpha_i d_i$ (originally 1 packet) and $\beta := \beta_i d_i$ (originally 3 packets) by adjusting its window size. The window size is incremented or decremented by approximately 1 packet in each round-trip time according as the current estimate is less than α or greater than β. Otherwise the window size is unchanged. The rationale is that each source should maintain a small number of its own packets in the pipe to take advantage of extra capacity when it becomes available.

Another interpretation of Vegas observes that

$$w_i(t) - x_i(t)d_i \;=\; x_i(t)(T_i(t) - d_i) \;=\; x_i(t)q_i(t)$$

where $q_i(t)$ is the round-trip *queueing* delay. Then the conditional in the pseudocode of Vegas becomes

$$x_i(t) \quad < \quad \alpha_i \frac{d_i}{q_i(t)} \qquad \text{or} \qquad x_i(t) \quad > \quad \beta_i \frac{d_i}{q_i(t)}$$

i.e., a Vegas source sets its rate to be proportional to the ratio of its round-trip propagation delay to queueing delay, the proportionality constant being between α_i and β_i. The more congested its path is, the higher the queueing delay and hence the lower the rate.

TCP FAST

TCP FAST can be thought of as a high-speed version of Vegas. The pseudocode for the congestion avoidance phase is shown in Figure 1.5. Both use queueing delay as the measure

periodically

{

$$W = \gamma\left(\frac{\text{baseRTT}}{\text{RTT}} W + \alpha\right) + (1-\gamma)W$$

}

Figure 1.5: The pseudocode of TCP FAST congestion avoidance agorithm.

of congestion (price) and both have the same equilibrium point. Vegas works well when the network is slow or small, but its response is too sluggish in high-speed long-distance networks because, regardless of how far the network state is from its equilibrium, Vegas adjusts the window by the same amount (one packet per RTT). In contrast the size of the window adjustment in FAST is proportional to the distance of the network state from its equilibrium. It converges rapidly towards the equilibrium when it is far away and smooths into the equilibrium when it is close. See Chapter 1.3 for more details.

1.2.3 AQM ALGORITHMS

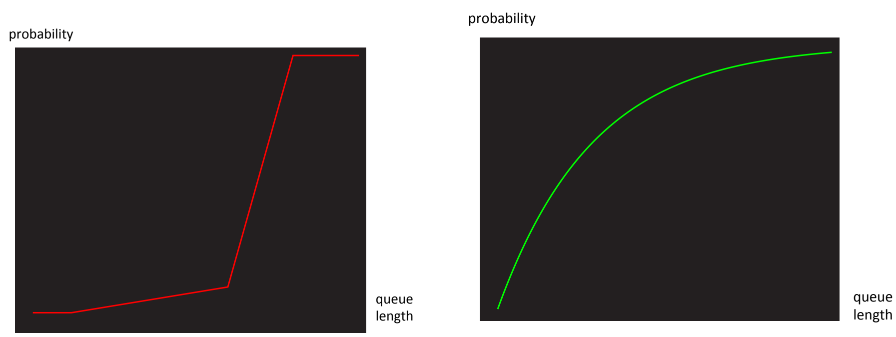
DropTail

Congestion control on the Internet is still predominantly source-based in that the link algorithm is implicit. A link (router) simply drops a packet that arrives at a full buffer. This is called DropTail (or Tail Drop) and the implicit link algorithm is carried out by the queue process. The congestion measure it updates depends on the TCP algorithm.

For TCP Reno and its variants, the congestion measure is packet loss probability. The end-to-end loss probability is observed at the source and is a measure of congestion on the end-to-end path. For TCP Vegas and FAST the congestion measure turns out to be link queueing delay when first-in-first-out service discipline is used. The congestion measure of a path is the sum of queueing delays at all constituent links.

RED

Random Early Detection (RED) is an alternative way to generate packet loss as a congestion measure. Instead of dropping only at a full buffer, RED maintains an exponentially weighted queue length and drops (or marks) packets with a probability that increases with the weighted queue length. When the weighted queue length is less than a minimum threshold no arrival packets are dropped. When it exceeds a maximum threshold all packets are dropped. When it is in between, a packet is dropped with a probability that is a piecewise linear and increasing function of the weighted queue length. See Figure 1.6(a).



(a) RED (b) REM

Figure 1.6: Loss probabilities of RED and REM as functions of (weighted) queue length.

REM

Random Exponential Marking (REM) consists of two simple ideas. First, in RED, the link price (loss probability) $p_l(t)$ depends on the queue length $b_l(t)$. As the number of flows sharing the link increases the steady-state price and hence queue length must also increase, leading to large delay. REM in contrast maintains a steady-state queue length around a target regardless of the number of flows sharing the link, thus decoupling the steady-state link delay from its congestion price. It achieves this by adjusting the link price $p_l(t)$ according to

$$\begin{aligned}\dot{p}_l &= \left(\alpha_l \left(y_l(t) - c_l\right) + \beta_l \left(b_l(t) - b_l^0\right)\right)^+_{p_l(t)} \\ \dot{b}_l &= \left(y_l(t) - c_l\right)^+_{b_l(t)}\end{aligned}$$

where

$$(a)^+_b \quad := \quad \begin{cases} a & \text{if } a > 0 \text{ or } b > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $y_l(t)$ and c_l are respectively the input rate and link capacity at link l, and $b_l(t)$ and $b_l^0 \geq 0$ are respectively the queue length and its target at link l. Hence $y_l(t) - c_l$ is a rate mismatch and $b_l(t) - b_l^0$ is a queue mismatch. REM increases $p_l(t)$ if the weighted sum of the rate mismatch and queue mismatch is positive, and decreases it otherwise. In equilibrium where $\dot{p} = 0$, REM matches rate and queue length to their target values:

$$y_l^* \;=\; c_l \quad \text{and} \quad b_l^* \;=\; b_l^0$$

unless l is not a bottleneck link in which case $p_l = 0$, $y_l^* \leq c_l$ and $b_l^* = 0$.

The second idea of REM is a novel way to convey the link prices $p_l(t)$ to the sources through loss probability. Specifically REM explicitly embeds the *sum* of link prices along a path into the end-to-end loss probability that can be observed at the source, as follows. Like RED, REM drops a packet that arrives at a link with a probability independently of all other packets. Unlike RED whose loss probability is a piecewise linear function of the weighted queue length, the loss probability $\tilde{p}_l(t)$ of REM however is exponential in the link price (see Figure 1.6(b)):

$$\tilde{p}_l(t) \;:=\; 1 - \phi^{-p_l(t)}$$

Then the end-to-end probability $\hat{q}_i(t)$ that a packet of source i is lost in its path is exponential in the end-to-end price $q_i(t) = \sum_l R_{li} p_l(t)$:

$$\hat{q}_i(t) \;:=\; 1 \;-\; \prod_l \left(1 - \tilde{p}_l(t)\right)^{R_{li}} \;=\; 1 \;-\; \phi^{-q_i(t)}$$

Hence if source i estimates its end-to-end loss probability $\hat{q}_i(t)$ it can compute the end-to-end price $q_i(t)$ as

$$q_i(t) \;=\; -\log_\phi \left(1 - \hat{q}_i(t)\right)$$

As we will see below, the end-to-end price $q_i(t)$ can be useful for general TCP algorithms.

1.3 MODELS OF CLASSICAL ALGORITHMS

Even though TCP is a source-based mechanism, we must consider congestion control as a feedback system where source rates $x(t)$ interact with link prices $p(t)$. Different protocols choose different algorithms to adapt $(x(t), p(t))$. In this section we present mathematical models of the protocol pairs Reno/RED, Vegas/DropTail, and FAST/DropTail. The model captures the `congestion avoidance` phase of these protocols.

1.3.1 RENO/RED

We only model the average behavior of the additive increase multiplicative decrease (AIMD) algorithm used to control the window size and does not differentiate between TCP Reno and its variants such as NewReno, SACK, etc. All these protocols (henceforth referred to as "Reno") increase the window by one packet every round-trip time if there is no loss in the round-trip time, and halve the window otherwise.

Let $w_i(t)$ be the window size of source i. Let T_i be the round-trip time (propagation plus queueing delay), which we assume to be constant. Let $x_i(t) := w_i(t)/T_i$ be the source rate at time t. The time unit is on the order of several round-trip times and source rate $x_i(t)$ should be interpreted as the average rate over this timescale. Dynamics smaller than the timescale of a round-trip time is not captured by the fluid model. Let $p_l(t)$ be the loss (or marking) probability at link l at time t. We make the key assumption that the end-to-end loss (or marking) probability $q_i(t)$ to which source algorithm reacts is the *sum* of link loss probabilities: $q_i(t) = \sum_l R_{li} p_l(t)$. This is reasonable when $p_l(t)$ are small, in which case

$$q_i(t) \quad = \quad 1 - \prod_l (1 - p_l(t))^{R_{li}} \quad \simeq \quad \sum_l R_{li} p_l(t)$$

Consider the pseudocode of AIMD in Figure 1.3(a). In period t it transmits at rate $x_i(t)$ packets per unit time and receives (positive and negative) ACKs at approximately the same rate, assuming every packet is acknowledged. Hence on average source i receives $x_i(t)(1 - q_i(t))$ number of positive ACKs per unit time and each positive ACK increases the window $w_i(t)$ by $1/w_i(t)$. It receives on average $x_i(t)q_i(t)$ negative ACKs (losses) per unit time and each halves the window. Hence the net change to the window in period t is roughly

$$\dot{w}_i \quad = \quad x_i(t)(1 - q_i(t)) \, \frac{1}{w_i(t)} \; - \; x_i(t) q_i(t) \, \frac{w_i(t)}{2}$$

Since T_i are assumed to be constant the sending rate $x_i(t) := w_i(t)/T_i$ has the same dynamics as the window $w_i(t)$ except for a constant scaling. Hence we will model the dynamics of the sending rate directly, as:

$$\dot{x}_i \quad = \quad \left(\frac{1 - q_i(t)}{T_i^2} \; - \; \frac{1}{2} \, q_i(t) \, x_i^2(t) \right)_{x_i(t)}^{+} \tag{1.1a}$$

where

$$(a)_b^+ \quad := \quad \begin{cases} a & \text{if } a > 0 \text{ or } b > 0 \\ 0 & \text{otherwise} \end{cases}$$

The quadratic term in (1.1a) captures the property that, if the rate doubles, the multiplicative decrease occurs at twice the frequency with twice the amplitude. The operation $(\cdot)_{x_i(t)}^+$ ensures that $x_i(t)$ stays nonnegative, i.e., $\dot{x}_i$ stays 0 when $x_i(t) = 0$ and the quantity in the bracket is negative.

There are variants of this model. One is that the window increases deterministically by 1 every round-trip time. This modifies the additive increase term in (1.1a) into:

$$\dot{x}_i \;=\; \left(\frac{1}{T_i^2} - \frac{1}{2}\,q_i(t)\,x_i^2(t)\right)_{x_i(t)}^{+} \tag{1.1b}$$

It is approximately the same as (1.1a) when the loss probabilities $p_l(t)$ are small.

Another variant is that, instead of halving the window on each negative acknowledgment, the window is halved once in each round-trip time that contains one or more negative acknowledgments. This modifies the multiplicative decrease term in (1.1a) into:

$$\dot{x}_i \;=\; \left(\frac{1-q_i(t)}{T_i^2} - \frac{1}{2T_i}\,q_i(t)\,x_i(t)\right)_{x_i(t)}^{+}$$

RED updates the queue length $b_l(t)$ according to:

$$\dot{b}_l \;=\; (y_l(t) - c_l)_{b_l(t)}^{+} \tag{1.1c}$$

The queue length $b_l(t)$ is an internal variable and (a "gentle" version of) RED drops (or marks) a packet with a probability $p_l(t)$ that is a piecewise linear increasing function of $b_l(t)$ (see Figure 1.6(a)):

$$p_l(t) \;=\; \begin{cases} 0, & b_l(t) \le b_1 \\ \rho_1(b_l(t) - b_1), & b_1 \le b_l(t) \le b_2 \\ \rho_2(b_l(t) - b_2) + m, & b_2 \le b_l(t) \le b_3 \\ 1, & b_l(t) \ge b_3 \end{cases} \tag{1.1d}$$

where $m \in (0,1)$ is the nonzero loss probability at the break point,

$$\rho_1 \;=\; \frac{m}{b_2 - b_1} \qquad \text{and} \qquad \rho_2 \;=\; \frac{1-m}{b_3 - b_2}$$

In summary the dynamical system that models Reno/RED consists of equations (1.1).[3]

The AQM model (1.1c)–(1.1d) contains internal variables $b_l(t)$ that are not observed at TCP sources. We can eliminate $b_l(t)$ and derive the dynamics of $p_l(t)$ from (1.1c) (see Figure 1.6(a)):

$$\dot{p}_l(t) \;=\; \begin{cases} \rho_1(y_l(t) - c_l)_{p_l(t)}^{+}, & 0 \le p_l(t) \le m \\ \rho_2(y_l(t) - c_l)_{-}^{p_l(t)-1}, & m < p_l(t) \le 1 \end{cases} \tag{1.2}$$

[3]In practice, the loss probability in RED depends not directly on the instantaneous queue length $b_l(t)$, but on an exponentially weighted average $r_l(t)$, modeled by

$$\dot{r}_l \;=\; -\alpha_l(r_l(t) - b_l(t))$$

where $\alpha_l \in (0,1)$ is an exponential weight. Then the loss probability $p_l(t)$ is given by (1.1d) with $b_l(t)$ replaced by $r_l(t)$.

where

$$(a)_b^+ \;:=\; \begin{cases} a & \text{if } a > 0 \text{ or } b > 0 \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad (a)_-^b \;:=\; \begin{cases} a & \text{if } a < 0 \text{ or } b < 0 \\ 0 & \text{otherwise} \end{cases}$$

Then Reno/RED can be modeled by ordinary differential equations (1.1a)(1.2) without internal variables and algebraic equations. Most of our analysis will be on models of this type.

1.3.2 VEGAS/DROPTAIL

Consider the pseudocode in Figure 1.4(a) of the congestion avoidance mechanism of TCP Vegas. Let d_i denote the round-trip propagation delay of source i, $q_i(t)$ the round-trip queueing delay, $T_i(t) := d_i + q_i(t)$ the round-trip time, $w_i(t)$ the window size and $x_i(t) := w_i(t)/T_i(t)$ the sending rate of flow i. For simplicity assume $\alpha = \beta$. As discussed earlier the conditional in the pseudocode is

$$w_i(t) - x_i(t)\, d_i \;<\; \alpha_i\, d_i \quad \text{or} \quad w_i(t) - x_i(t)\, d_i \;>\; \alpha_i\, d_i$$

The quantity $x_i(t)d_i$ represents the number of source i's packets propagating in the communication channel (e.g. fiber optic cables) from source to destination, and the quantity $w_i(t) - x_i(t)d_i$ represents the number of i's packets buffered at the routers in its path. A Vegas source estimates the number of its own packets buffered in the path and tries to keep this number at $\alpha := \alpha_i d_i$ packets by incrementing or decrementing its window size $w_i(t)$ by 1 packet per round-trip time. Since

$$w_i(t) - x_i(t)\, d_i \;=\; x_i(t)(T_i(t) - d_i) \;=\; x_i(t)\, q_i(t)$$

i.e., $x_i(t)q_i(t)$ is the number of i's own packets buffered in the queues in its path, we model the window dynamics as:

$$\dot{w}_i \;=\; \frac{1}{d_i + q_i(t)}\, \text{sign}\left(\alpha_i d_i \;-\; x_i(t)q_i(t)\right)^+_{w_i(t)} \tag{1.3a}$$

$$x_i(t) \;=\; \frac{w_i(t)}{d_i + q_i(t)} \tag{1.3b}$$

where sign(z) is -1 if $z < 0$, 0 if $z = 0$, and 1 if $z > 0$. The operation $(\cdot)^+_{w_i(t)}$ ensures that $w_i(t) \geq 0$. Hence (1.3a) says that the window is adjusted by 1 packet per round-trip time by comparing the number $x_i(t)q_i(t)$ of packets buffered in its path with the target $\alpha_i d_i$. In equilibrium each source i maintains $\alpha_i d_i$ packets in its path.

Hence Vegas uses as its congestion price the queueing delay at a link l

$$p_l(t) \;:=\; \frac{b_l(t)}{c_l}$$

where $b_l(t)$ is the queue length and evolves according to (1.1c). The price dynamics is therefore (dividing both sides of (1.1c) by c_l):

$$\dot{p}_l \quad = \quad \frac{1}{c_l}\left(y_l(t) - c_l\right)^+_{p_l(t)} \tag{1.3c}$$

where $y_l(t) := \sum_i R_{li}\, x_i(t)$. In summary the dynamical system that models Vegas/DropTail consists of equations (1.3).

The TCP model (1.3a)–(1.3b) contains algebraic equations. To reduce the model to ordinary differential equations, we can eliminate $x(t)$ by substituting (1.3b) into (1.3a)(1.3c) to obtain:

$$\begin{aligned}
\dot{w}_i &= \frac{1}{d_i + q_i(t)} \operatorname{sign}\left(\alpha_i d_i - \frac{w_i(t) q_i(t)}{d_i + q_i(t)}\right)^+_{w_i(t)} \\
\dot{p}_l &= \frac{1}{c_l}\left(\sum_i \frac{R_{li} w_i(t)}{d_i + q_i(t)} - c_l\right)^+_{p_l(t)}
\end{aligned}$$

1.3.3 FAST/DROPTAIL

TCP Reno and Vegas were designed in the late 1980s and early 1990s when network capacities are much smaller and a large majority of network traffics are relatively local. The window adjustment by one packet per round-trip time was adequate then, but much too slow as networks scale up in capacity and geographical coverage. This has motivated a large effort in the 2000s to develop an understanding of the mathematical structure of congestion control and a theory-aided design approach where performance analysis was done before, as opposed to after, implementation and deployment. TCP FAST came out of that effort.

Even though the implementation of FAST is very different from that of Vegas (compare their pseudocodes in Figure 1.5 and Figure 1.4(a)), the FAST design was inspired by the underlying mathematical structure of Vegas. In particular FAST attains the same equilibrium of Vegas and therefore achieves proportional fairness (see Chapter 2). It speeds up the dynamics of Vegas by adjusting the window size by an amount proportional to the deviation of the queue length in the path from its target value, instead of by one packet per round-trip time as Vegas does in (1.3a).

Specifically consider the FAST pseudocode in Figure 1.5. It can be modeled by

$$\dot{w}_i \quad = \quad \gamma\left(\alpha_i - x_i(t) q_i(t)\right)^+_{w_i(t)} \tag{1.4a}$$

$$x_i(t) \quad = \quad \frac{w_i(t)}{d_i + q_i(t)} \tag{1.4b}$$

where $q_i(t) := \sum_l R_{li}\, p_l(t)$. Like Vegas, FAST also uses as its congestion price the queueing delay at a link l whose dynamics is described by:

$$\dot{p}_l \quad = \quad \frac{1}{c_l}\left(y_l(t) - c_l\right)^+_{p_l(t)} \tag{1.4c}$$

where $y_l(t) := \sum_i R_{li}\, x_i(t)$. In summary the dynamical system that models FAST/DropTail consists of equations (1.4).

FAST can be interpreted as a high-speed version of Vegas in the following sense. Compare (1.3a) for Vegas and (1.4a) for FAST (replacing $\alpha_i d_i$ in (1.3a) by α_i). While Vegas increments/decrements the window by 1 packet depending on the sign of $\alpha_i - x_i(t)q_i(t)$, FAST updates the window by an amount proportional to $\alpha_i - x_i(t)q_i(t)$. For high-speed long-distance networks the magnitude of $\alpha_i - x_i(t)q_i(t)$ can be big during transient, in which case FAST closes the gap faster than Vegas.

The FAST model (1.4a)–(1.4b) contains algebraic equations. As for Vegas, we can eliminate $x(t)$ by substituting (1.4b) into (1.4a)(1.4c) to obtain a model consisting of ordinary differential equations:

$$\dot{w}_i = \gamma \left(\alpha_i - \frac{w_i(t) q_i(t)}{d_i + q_i(t)} \right)^+_{w_i(t)} \tag{1.5a}$$

$$\dot{p}_l = \frac{1}{c_l} \left(y_l(t) - c_l \right)^+_{p_l(t)} \tag{1.5b}$$

An alternative model is to replace the dynamic link model (1.5b) by a static model where the link queueing delay vector $p(t)$ is determined implicitly by the window sizes $w(t)$:

$$\dot{w}_i = \gamma \left(\alpha_i - \frac{w_i(t) q_i(t)}{d_i + q_i(t)} \right)^+_{w_i(t)} \tag{1.6a}$$

$$\sum_i R_{li} \frac{w_i(t)}{d_i + q_i(t)} \begin{cases} = c_l & \text{if } p_l(t) > 0 \\ \le c_l & \text{if } p_l(t) = 0 \end{cases} \tag{1.6b}$$

It can be proved that the model (1.6) is well defined: given $w(t)$, there is a unique queueing delay vector $p(t)$ that satisfies (1.6b) provided R has full row rank.

Another alternative model is to replace the dynamic source model (1.5a) by a static model (obtained by setting $\dot{w} = 0$ in (1.4a) and using $x_i(t) = w_i(t)/(d_i + q_i(t))$):

$$\dot{p}_l = \frac{1}{c_l} \left(y_l(t) - c_l \right)^+_{p_l(t)} \tag{1.7a}$$

$$x_i(t) = \frac{\alpha_i}{q_i(t)} \tag{1.7b}$$

1.4 A GENERAL SETUP

1.4.1 THE BASIC MODELS

We have seen in Chapter 1.3 that one can model a physical system in multiple ways, each making different assumptions. All the models there take the form of differential algebraic equations:

$$
\begin{aligned}
\dot{\tilde{x}}_i &= \tilde{f}_i(x_i(t), \tilde{x}_i(t), q_i(t)), & i \in N \\
x_i(t) &= f(x_i(t), \tilde{x}_i(t), q_i(t)), & i \in N \\
\dot{\tilde{p}}_l &= \tilde{g}_l(y_l(t), p_l(t), \tilde{p}_l(t)), & l \in L \\
p_l(t) &= g_l(y_l(t), p_l(t), \tilde{p}_l(t)), & l \in L
\end{aligned}
$$

where $(\tilde{x}(t), \tilde{p}(t))$ are (generally vector) internal variables and $(x(t), p(t))$ are source rates and link prices. The most important feature of this model is its decentralized nature: user i adapts its rate $x_i(t)$ based only on local information $(x_i(t), \tilde{x}_i(t))$ and the locally observed congestion measure $q_i(t)$, and link l adapts its price $p_l(t)$ based only on local information $(y_l(t), p_l(t), \tilde{p}_l(t))$. We represent this set of differential algebraic equations in vector form as

$$
\begin{aligned}
\dot{\tilde{x}} &= \tilde{f}(x(t), \tilde{x}(t), q(t)) \\
x &= f(x(t), \tilde{x}(t), q(t)) \\
\dot{\tilde{p}} &= \tilde{g}(y(t), p(t), \tilde{p}(t)) \\
p &= g(y(t), p(t), \tilde{p}(t))
\end{aligned}
$$

where $q(t) = R^T p(t)$, $y(t) = R\,x(t)$. As we have seen in Chapter 1.3 above, while including internal variables $(\tilde{x}(t), \tilde{p}(t))$ is convenient for modeling, it is usually easy for practical protocols to eliminate these internal variables to obtain a model involving only $(x(t), p(t))$. Such a model is usually more convenient for analysis.

Our analysis will therefore focus on the following three special cases without the internal variables.

1. *Primal algorithms.* This is the class of algorithms that have dynamics (memory) in the source rates (or window sizes) but not in the link prices:

$$
\begin{aligned}
\dot{x} &= f(x(t), q(t)) && (1.8a) \\
p &= g(y(t), p(t)) && (1.8b)
\end{aligned}
$$

2. *Dual algorithms.* This is the class of algorithms that have dynamics in the link prices but not in the source rates (or window sizes):

$$
\begin{aligned}
\dot{p} &= g(y(t), p(t)) && (1.9a) \\
x &= f(x(t), q(t)) && (1.9b)
\end{aligned}
$$

3. *Primal-dual algorithms.* This is the class of algorithms with dynamics in both source rates and link prices:

$$
\begin{aligned}
\dot{x} &= f(x(t), q(t)) && (1.10a) \\
\dot{p} &= g(y(t), p(t)) && (1.10b)
\end{aligned}
$$

We will refer to any of (1.8), (1.9), (1.10) together with

$$q(t) \;=\; R^T p(t) \quad \text{and} \quad y(t) \;=\; R\,x(t) \tag{1.11}$$

as the *basic model*. The structure of these models is illustrated in Figure 1.7. For example,

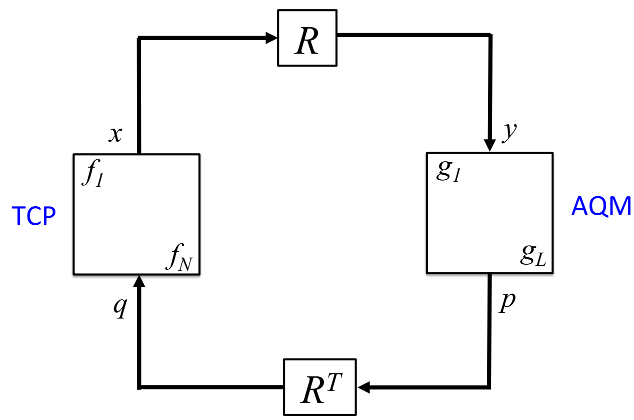


Figure 1.7: The basic model: a multi-source multi-link network.

FAST/DropTail can be modeled as either a primal, a dual, or a primal-dual algorithm; see Chapter 1.3.3.

For primal algorithms, we can often express $p(t)$ in terms of $x(t)$ using (1.8b) because the Jacobian $\frac{\partial g}{\partial p} - I$ is often nonsingular for TCP/AQM models. Substituting $p(t) = p(x(t))$ as a function of $x(t)$ into (1.8a)(1.11) then reduces the primal algorithm into a set of ODEs in $x(t)$. Similarly for dual algorithms, we can often express $x(t)$ in terms of $p(t)$ using (1.9b) and reduce the dual algorithm into a set of ODEs in $p(t)$ (substitute $x(p(t))$ into (1.9a)(1.11)). Therefore all three basic models can be described by ODEs. We hence often cast our discussion in terms of primal-dual algorithms (1.10) as the prototypical model, knowing that it applies to all three models.

Since the source rates $x(t)$ and prices $p(t)$ are nonnegative, hidden in the notation of the functions (f, g) is the projection to the nonnegative quadrant. We sometimes make this explicit by writing, instead of (1.10):

$$\begin{aligned} \dot{x} &= \left(f(x(t), q(t))\right)^+_{x(t)} \\ \dot{p} &= \left(g(y(t), p(t))\right)^+_{p(t)} \end{aligned}$$

where for any vectors $a, b \in \mathbb{R}^n$, the vector $(a)_b^+$ is defined componentwise, i.e., for all $i = 1, \ldots, n$,

$$\left[(a)_b^+\right]_i \quad := \quad (a_i)_{b_i}^+ \quad := \quad \begin{cases} a_i & \text{if } \ a_i > 0 \ \text{ or } \ b_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1.12}$$

As we will see later, the projection on the right-hand side of the differential equations mostly preserves the existence, uniqueness, equilibrium and stability properties of the unprojected system, but the analysis is considerably more complicated because of the discontinuity introduced by the projection. A notable feature of this book is the careful treatment of the projected dynamics.

We are interested in three questions:

1. Under what condition does the basic model (1.10) have a unique solution $(x(t), p(t), t \geq 0)$, given any initial point $(x(0), p(0))$?

2. Does equilibrium of (1.10) exist? Is it unique? How do we characterize it?

3. Is an equilibrium of (1.10) stable?

We answer question 1 in this Chapter, question 2 in Chapter 2, and question 3 in Chapters 3 through 6.

1.4.2 LIMITATIONS AND EXTENSIONS

We now discuss some limitations of our fluid models and various extensions.

1. **No feedback delay.** The basic models ignore feedback delay, i.e., they assume that a change in a link price $p_l(t)$ is instantly observed at all the sources that use that link, and a change in a source rate $x_i(t)$ affects the aggregate flow rates at all the links in the path of the source. In reality, of course, there is feedback delay due to signal propagation between links and sources, and due to packet queueing and processing in the network. Feedback delay can be modeled by replacing (1.11) with

$$\begin{aligned} q_i(t) &= \sum_l R_{li} p_l(t - \tau_{li}^b), \qquad i \in N \\ y_l(t) &= \sum_i R_{li} x_i(t - \tau_{li}^f), \qquad l \in L \end{aligned}$$

   Here τ_{li}^b models the backward delay from link l to source i and τ_{li}^f models the forward delays from source i to link l, both assumed to be constant. It is the time it takes for a change in the price at link l to affect the rate at source i, and the time it takes for a change in the rate at source i to affect the price at link l, respectively.

   We will use this delayed model in Chapter 6 to study linear stability in the presence of feedback delay.

2. **Window control.** A window-based control does not adapt its sending rate $x_i(t)$ directly, but rather its window size $w_i(t)$. This is modeled using the relationship between $x_i(t)$ and $w_i(t)$

$$x_i(t) \quad = \quad \frac{w_i(t)}{T_i(t)} \quad = \quad \frac{w_i(t)}{d_i + q_i(t)}$$

   as in (1.3) for Vegas and in (1.4) for FAST. For convenience however we often assume the round-trip time T_i to be constant and model the dynamics of $x_i(t)$ directly as in (1.1) for Reno.

3. **Round-trip timescale.** The basic models ignore flow arrivals and departures and fix set of flows and their routing matrix. They focus on the network dynamics at the round-trip timescale. This amounts to ignoring short-duration TCP flows.

4. **Heterogeneous protocols.** The basic models assume that, while different sources i can adapt their rates using different algorithms f_i, they all react to the same type of congestion prices $p_l(t)$. For instance all sources react to packet loss probabilities in their paths as in Reno, or all sources react to packet delay in their paths as in Vegas or FAST. If sources that use different types of congestion prices (e.g. Reno and Vegas) share the same network, the network behavior can be much more complicated. For instance while a network of homogeneous sources typically has a unique equilibrium point, a network of heterogeneous sources may not.

5. **Deterministic fluid.** We model traffic as a deterministic fluid and ignore randomness in, e.g. packet processing and queueing times, packet arrival process, or flow arrivals and departures.

6. **Queue output process.** The basic models assume a flow maintains its rate along its path so that every link in its path sees the *source* rate. In reality however as a flow goes through a queue, its output rate depends on the buffering process and is generally different from its input rate, except when the network is in a steady state where all queues are stabilized.

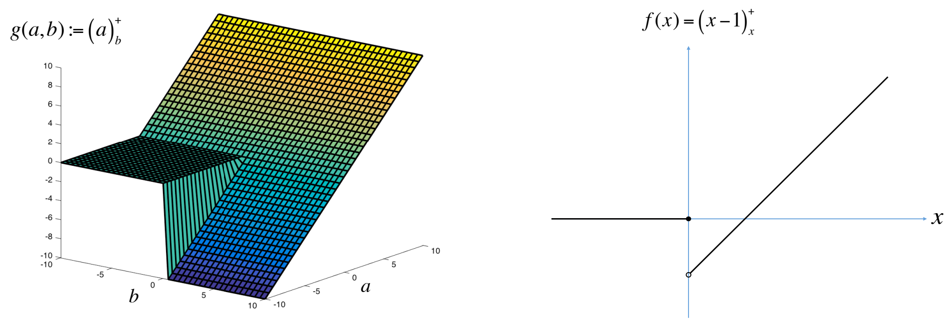
1.5 SOLUTION OF THE BASIC MODELS

The ODE model (1.10) describes the protocol action in the congestion avoidance phase. In this section we discuss conditions on the TCP/AQM models (f, g) that guarantee the existence of a unique solution trajectory $(x(t), p(t), t \geq 0)$ given any initial point $(x(0), p(0))$. We first summarize sufficient conditions for the existence and uniqueness of the solution to general ODE systems. We then illustrate how to use these results to prove that the TCP/AQM models of Chapter 1.3 are well defined.

Consider a system of projected ordinary differential equations:

$$\dot{x} = (f(x(t)))^+_{x(t)}, \qquad t \geq 0 \tag{1.13a}$$
$$x(0) = x_0 \geq 0 \quad \text{given} \tag{1.13b}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$, and the projection $(f(x))^+_x$ is defined in (1.12).[4] Even though the projection function $g(a,b) := (a)^+_b$, a, b in $\mathbb{R}$, is continuous in a given any b, it is discontinuous in b for $a < 0$. To visualize the effect of projection see Figure 1.8. Therefore the right-hand side of (1.13a) is discontinuous in $x(t)$ even if f is continuous. It turns out that this is a



(a) $g(a,b)$

(b) $f(x)$

Figure 1.8: (a) The function $g(a,b) = (a)^+_b$ is discontinuous along the line $b = 0$ for $a < 0$. (b) The function $f(x) := (x-1)^+_x$ is discontinuous at $x = 0$.

simple type of discontinuity in that the projection preserves the existence and uniqueness properties, i.e., the conditions we describe below that guarantee the existence of a unique solution to unprojected dynamics $\dot{x} = f(x(t))$ will also guarantee the existence of a unique solution to projected dynamics $\dot{x} = (f(x(t)))^+_{x(t)}$.

We hence first consider a system of ODEs without projection:

$$\dot{x} = f(x(t)), \qquad t \geq 0 \tag{1.14a}$$
$$x(0) = x_0 \quad \text{given} \tag{1.14b}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$. We explain different notions of Lipschitz continuity. We then state Lipschitz conditions on f that guarantee the existence and uniqueness of solutions to the unprojected dynamics (1.14). Finally we state precisely the result that projection preserves these properties.

[4]We abuse notation to use f to denote either a generic function or a TCP algorithm and x to denote either a generic variable or source rates, depending on the context.

1.5.1 EXISTENCE AND UNIQUENESS THEOREMS

Definition 1.1 Lipschitz continuity. The function $f(x)$ is said to be:

1. *locally Lipschitz at* x_0 if there exist $r = r(x_0) > 0$ and $L = L(x_0)$ such that $\|f(x) - f(y)\| \leq L\|x - y\|$ for all $x, y \in B_r(x_0)$.

2. *locally Lipschitz on a domain* [5] $D \subseteq \mathbb{R}^n$ if it is locally Lipschitz at each point in D.

3. *Lipschitz on a domain* D if the Lipschitz constant L is uniform at each point in D, i.e., there exists L such that $\|f(x) - f(y)\| \leq L\|x - y\|$ for all $x, y \in D$.

4. *globally Lipschitz* if it is Lipschitz on $D = \mathbb{R}^n$, i.e., there exists L such that $\|f(x) - f(y)\| \leq L\|x - y\|$ for all $x, y \in \mathbb{R}^n$.

Being Lipschitz is stronger than continuity. Here $\|\cdot\|$ can be any norm in $\mathbb{R}^n$, such as p-norms defined by:

$$\begin{aligned} \|x\|_p &:= \left(|x_1|^p + \cdots + |x_n|^p\right)^{1/p}, \qquad 1 \leq p < \infty \\ \|x\|_\infty &:= \max_i \; |x_i| \end{aligned}$$

The three most commonly used norms are $\|x\|_1$, $\|x\|_2$ (Euclidean norm), and $\|x\|_\infty$. Since all norms are equivalent in $\mathbb{R}^n$, Lipschitz continuity can be defined using any norm. Different choices of norms differ only in their Lipschitz constants.

Example 1.2

1. $f(x) = x^2$ is Lipschitz on any *bounded* domain of $\mathbb{R}$, but only locally Lipschitz on $\mathbb{R}$. To see this note that, by the mean value theorem, for any finite $x, y \in \mathbb{R}$ there exists $x \leq z \leq y$ such that

$$|f(x) - f(y)| \;=\; |f'(z)|\,|x - y| \;\leq\; \sup_{z \in C}(2z)|x - y| \;\leq\; L\,|x - y|$$

where $\sup_{z \in C}(2z)$ is finite for any bounded set $C \subset \mathbb{R}$ and L is an upper bound.

2. $f(x) = x^{-1}$ is locally Lipschitz everywhere in $\mathbb{R}$ except at $x = 0$ and Lipschitz on any compact set C that does not contain $x = 0$. Again

$$|f(x) - f(y)| \;=\; |f'(z)|\,|x - y| \;\leq\; \max_{z \in C} \frac{1}{z^2}\,|x - y| \;\leq\; L\,|x - y|$$

where $\max_{z \in C} z^{-2}$ is finite and L is an upper bound.

[5] A *domain* is an open and connected set.

3. $f(x) = \sin x$ is globally Lipschitz on $\mathbb{R}$ since $|f(x) - f(y)| = |\cos(z)||x - y| \le |x - y|$.

□

Lemma 1.3

1. *Let f be continuously differentiable on a domain (connected open set) $D \subseteq \mathbb{R}^n$. Then f is locally Lipschitz on D.*

2. *Suppose f is continuously differentiable on a convex domain $D \subseteq \mathbb{R}^n$ and let $L := \sup_{y \in D} \|\frac{\partial f}{\partial x}(y)\|$. Then f is Lipschitz on D with Lipschitz constant L if and only if L is finite.*

3. *Suppose f is continuously differentiable on a convex and compact set $D \subseteq \mathbb{R}^n$. Then $L := \sup_{y \in D} \|\frac{\partial f}{\partial x}(y)\|$ is finite and f is Lipschitz on D with Lipschitz constant L.*

4. *Let f be continuously differentiable on $\mathbb{R}^n$. Then f is globally Lipschitz on $\mathbb{R}^n$ if and only if $\frac{\partial f}{\partial x}$ is uniformly bounded on $\mathbb{R}^n$.*

The proof of the lemma, given in Appendix 1.5.3, illustrates several useful techniques.

The next result says that global Lipschitz continuity guarantees the existence and uniqueness of the solution to the unprojected dynamic (1.14).

Theorem 1.4 *If f is globally Lipschitz on $\mathbb{R}^n$, i.e., there exists an L such that*

$$\|f(x) - f(y)\| \quad \le \quad L\|x - y\| \qquad \forall x, y \in \mathbb{R}^n$$

then (1.14) *has a unique solution* $(x(t), t \ge 0)$.

Example 1.5 Consider $\dot{x} = Ax + b =: f(x)$ for $t \ge 0$. Then

$$\|f(x) - f(y)\| \quad = \quad \|A(x - y)\| \ \le \|A\| \cdot \|x - y\| \qquad \forall x, y \in \mathbb{R}^n$$

where $\|x\|$ is any norm on $\mathbb{R}^n$ and $\|A\|$ is the induced matrix norm. Therefore f is globally Lipschitz on $\mathbb{R}^n$ and the system has a unique solution $x(t), t \ge 0$ by Theorem 1.4. □

Global Lipschitz continuity is often too stringent. Local Lipschitz continuity is often enough. For instance we model FAST/DropTail as a dual algorithm in (1.7). The unprojected dynamics is

$$\begin{aligned} \dot{p}_l &= \frac{1}{c_l}(y_l(t) - c_l) \\ x_i(t) &= \frac{\alpha_i}{q_i(t)} \end{aligned}$$

Substituting $x = x(p)$ and $y = y(p)$ as functions of p, this becomes:

$$\dot{p}_l \quad = \quad \frac{1}{c_l}\left(\sum_i \frac{R_{li}\alpha_i}{\sum_k R_{ki}p_k(t)} - c_l\right)$$

The function on the right-hand side is not globally Lipschitz over $p > 0$, but is locally Lipschitz. This system has a unique solution $(p(t), t \geq 0)$ given any $p(0)$ according to the next result (see Chapter 1.5.2 for details).

Theorem 1.6 *Suppose f is locally Lipschitz on a domain $D \subseteq \mathbb{R}^n$. Let $W := W(x_0)$ be a compact (closed and bounded) subset of D that contains the given initial state x_0. Suppose every solution of* (1.14)*, if exists, lies entirely in W, i.e., $x(t) \in W$ for all $t \geq 0$. Then* (1.14) *has a solution $(x(t), t \geq 0)$ and it is unique.*

Example 1.7 Consider $\dot{x} = -x^3 =: f(x)$. Note that $f(x)$ is not globally Lipschitz on $\mathbb{R}$, but it is locally Lipschitz at all $x \in \mathbb{R}$. Moreover given any $x(0) = x_0$, a solution $x(t)$, if it exists, stays in the compact set $W := \{x \in \mathbb{R} \,|\, |x| \leq |x_0|\}$. Hence there exists a unique solution by Theorem 1.6. Indeed the solution is $x(t) = \left(2t + x_0^{-2}\right)^{-1/2}$ if $x_0 \neq 0$ and $x(t) \equiv 0$ if $x_0 = 0$. □

In general the conditions in Theorems 1.4 and 1.6 are sufficient but not necessary. If f is only locally Lipschitz on $\mathbb{R}^n$ (but there is no compact set W that contains any solution trajectory as required in Theorem 1.6), then a unique solution $x(t)$ always exists over $[0, T]$ for some $T > 0$, but T may be finite.[6] Indeed not even continuity is necessary for a solution to exist over some finite interval, as the next example shows.

Example 1.8 **Discontinuous right-hand side with classical solution.** Consider $\dot{x} = f(x(t))$, $t \geq 0$, over $\mathbb{R}$ where

$$f(x) \quad := \quad \begin{cases} -1 & \text{if} \quad x > 0 \\ 0 & \text{if} \quad x = 0 \\ 1 & \text{if} \quad x < 0 \end{cases}$$

[6]More generally, if f is continuous then, given $x(0)$, a (classical) solution $x(t)$ exists over some interval $[0, T]$ for some possibly finite $T > 0$, but it may not be unique. If, in addition, f is also locally one-sided Lipschitz (i.e. for all $x \in \mathbb{R}^n$ there exists an $L < \infty$ and a neighborhood $B_\epsilon(x)$ such that $(f(y) - f(y'))^T (y - y') \leq L \|y - y'\|^2$ for all $y, y' \in B_\epsilon(x)$) then the solution $x(t)$ is also unique. Continuity and one-sided Lipschitz continuity do not imply each other, but Lipschitz continuity implies both and hence implies the existence and uniqueness of the solution. See also Condition C1.1 in Appendix 1.5.4.

Since f is discontinuous at $x = 0$ it satisfies neither Theorem 1.4 nor Theorem 1.6 (not locally Lipschitz at $x = 0$). Yet, given any initial point $x(0) = x_0$, it has a unique continuously differentiable solution over a time interval $t \in [0, T)$ for some possibly finite $T > 0$. If $x_0 > 0$ then the solution is $x(t) = x_0 - t$ for $t \in [0, x_0)$. If $x_0 < 0$ then $x(t) = x_0 + t$ for $t \in [0, -x_0)$. If $x_0 = 0$ then $x(t) = 0$ for $t \in [0, \infty)$. □

When f is Lipschitz, the unique solution $x(t)$ to (1.14), if exists, is continuously differentiable (in t). For projected dynamics (1.13a) whose right-hand side is however discontinuous, a solution even if it exists, is generally not differentiable everywhere. To allow projected dynamics (1.13a) we relax the classical solution notion. Formally we define a function $x : [0, \infty) \to \mathbb{R}^n_+$ to be *a (Carathéodory) solution to* (1.13) if it is absolutely continuous and satisfies (1.13) almost everywhere (i.e., except possibly for a set of time instances $t \in [0, \infty)$ that has Lebesgue measure zero).[7]

Example 1.9 Carathéodory solution. Consider $\dot{x} = f(x(t))$, $t \geq 0$, over $\mathbb{R}$ where

$$f(x) \quad := \quad \begin{cases} 1 & \text{if} \quad x \neq 0 \\ a & \text{if} \quad x = 0 \end{cases}$$

where $a > 0$ and $a \neq 1$ so that f is discontinuous at $x = 0$ (though not due to projection). A unique (Carathéodory) solution is $x(t) = t$ for $t \in [0, \infty)$. If $x(0) > 0$ then this solution satisfies $\dot{x} = f(x(t))$ at all t. If $x(0) \leq 0$ then it satisfies $\dot{x} = f(x(t))$ except at one time t where $x(t) = 0$.

If $a = 0$ then f is still discontinuous at $x = 0$, but for $x(0) = 0$, there are two Carathéodory solutions: $x(t) = 0$ for $t \in [0, \infty)$ and $x(t) = t$ for $t \in [0, \infty)$. It turns out one can define another solution notion, called a Filippov solution, for which $x(t) = t$ is the unique solution. In general neither Carathéodory solution nor Filippov solution includes the other. □

The next result generalizes Theorems 1.4 and 1.6 to the case of projected dynamical systems. Its proof is given in Appendix 1.5.4.

[7]This is called a Carathéodory solution for an ordinary differential equation system with a discontinuous right-hand side (vector field). A function $f : I \to \mathbb{R}^n$ where I is an interval in $\mathbb{R}$ is *absolutely continuous* if given any $\epsilon > 0$ there exists $\delta > 0$ such that for any finite sequence of pairwise disjoint subintervals $[s_k, t_k]$ of I we have

$$\sum_k |t_k - s_k| < \delta \quad \Rightarrow \quad \sum_k \|f(t_k) - f(s_k)\| < \epsilon$$

Continuous differentiability implies Lipschitz continuity implies absolute continuity implies uniform continuity implies continuity. Absolute continuity also implies bounded variation implies differentiability almost everywhere.

Theorem 1.10 *Consider the projected dynamical system* (1.13) *where* $f : \mathbb{R}^n \to \mathbb{R}^n$.

1. *If f is Lipschitz on the nonnegative quadrant $\mathbb{R}^n_+$, i.e., there exists L such that*

$$\|f(x) - f(y)\| \quad \leq \quad L\|x - y\| \qquad \forall x, y \in \mathbb{R}^n_+$$

   then (1.13) *has a unique solution* $(x(t), t \geq 0)$.

2. *Suppose f is locally Lipschitz on a domain $D \subseteq \mathbb{R}^n_+$. Let $W := W(x_0)$ be a compact (closed and bounded) subset of D that contains the given initial state x_0. Suppose every solution of* (1.13), *if exists, lies entirely in W, i.e., $x(t) \in W$ for all $t \geq 0$. Then* (1.13) *has a unique solution for all $t \geq 0$.*

The norm in the theorem can be any norm in $\mathbb{R}^n$.

1.5.2 APPLICATION TO TCP/AQM MODELS

We now apply Theorem 1.10 to show that a basic model of TCP/AQM is well defined. Consider the following dual algorithm in Chapter 1.3.3 that models FAST/DropTail:

$$\dot{p}_l = \frac{1}{c_l} (y_l(t) - c_l)^+_{p_l(t)} \qquad (1.16a)$$

$$x_i(t) = \frac{\alpha_i}{q_i(t)} \qquad (1.16b)$$

where for any $a, b \in \mathbb{R}$, $(a)^+_b := a$ if $a > 0$ or $b > 0$ and 0 otherwise. Here

$$q_i(t) = \sum_l R_{li} p_l(t), \qquad y_l(t) = \sum_i R_{li} x_i(t)$$

We now prove that (1.16) has a unique solution trajectory.

Write $x_i(q_i) := x_i(p) := \alpha_i/q_i$ as a function of q_i or p. Similarly write $y_l(p) := \sum_i R_{li} x_i(p)$ as a function of p. Consider then the following system with $p(t)$ as its state:

$$\dot{p}_l = \frac{1}{c_l} (y_l(p(t)) - c_l)^+_{p_l(t)} =: \frac{1}{c_l} (g_l(p(t)))^+_{p_l(t)} \qquad (1.17a)$$

$$p(0) = p^0 \text{ given} \qquad (1.17b)$$

where $g(p) := y(p) - c$ and

$$y_l(p) := \sum_i R_{li}\, x_i(p) := \sum_i R_{li} \frac{\alpha_i}{\sum_k R_{ki} p_k} \qquad (1.17c)$$

Clearly $g(p)$ is continuously differentiable over $\mathbb{R}^L_+$ at $p > 0$. Hence it is locally Lipschitz on any domain $D \subseteq \mathbb{R}^L_+$ with $p > 0$.[8] Theorem 1.10.2 then guarantees the existence and

[8]It is not globally Lipschitz because $\frac{\partial g}{\partial p}$ is not uniformly bounded on $\mathbb{R}^L_+$; see Lemma 1.3.4.

uniqueness of solution to (1.17) if we can construct a compact set $W := W(p^0) \subseteq D$ that contains the given initial point p^0 such that every solution $(p(t),\, t \geq 0)$ of (1.17), if exists, lies in W. We now construct such an W in two steps.

First consider the function

$$V(p) \;\; := \;\; \sum_i \alpha_i \, \log x_i(p) \; + \; \sum_l p_l \left(c_l - y_l(p)\right)$$

We claim that if there is a solution $(p(t), t \geq 0)$ to (1.17) then V is nonincreasing along the trajectory $p(t)$. To see this we have

$$\begin{aligned}
\frac{\partial V}{\partial p_l}(p) \;\; &= \;\; \sum_i \frac{\alpha_i}{x_i(p)} \frac{\partial x_i}{\partial p_l}(p) \; + \; (c_l - y_l(p)) \; - \; \sum_k p_k \sum_i R_{ki} \frac{\partial x_i}{\partial p_l}(p) \\
&= \;\; (c_l - y_l(p)) \; + \; \sum_i \left(\frac{\alpha_i}{x_i(p)} - q_i(p) \right) \frac{\partial x_i}{\partial p_l}(p) \\
&= \;\; c_l - y_l(p)
\end{aligned}$$

where the last equality follows from (1.16b). Hence

$$\frac{d}{dt} V(p(t)) \;\; = \;\; \frac{\partial V}{\partial p} \, \dot{p} \;\; = \;\; \sum_l \frac{1}{c_l} \left(c_l - y_l(p(t))\right) \left(y_l(p(t)) - c_l\right)^+_{p_l(t)}$$

For each l we have

$$\left(c_l - y_l(p(t))\right) \left(y_l(p(t)) - c_l\right)^+_{p_l(t)} \;\; = \;\; \begin{cases} -\left(c_l - y_l(p(t))\right)^2 & \text{if } y_l(p(t)) > c_l \text{ or } p_l(t) > 0 \\ 0 & \text{if } y_l(p(t)) \leq c_l \text{ and } p_l(t) = 0 \end{cases}$$

Hence for all $p(t) \geq 0$ on the solution trajectory we have

$$\frac{d}{dt} V(p(t)) \;\; \leq \;\; 0$$

Second define $W := W(p^0)$ by

$$W \;\; := \;\; W(p^0) \;\; := \;\; \left\{ p \in \mathbb{R}^L_+ \mid p \geq 0,\, V(p) \leq V(p^0) \right\}$$

Clearly W contains p^0 and any solution $(p(t),\, t \geq 0)$, if exists, must lie in W since $\dot{V}(p(t))$ is nonincreasing. Moreover W is closed. We claim that it is also bounded and hence is a compact set. To see this substitute (1.17c) into $V(p)$ to get

$$\begin{aligned}
V(p) \;\; &= \;\; \sum_i \alpha_i \left(\log \alpha_i - \log \left(\sum_l R_{li} p_l \right) \right) \; + \; \sum_l p_l \left(c_l - \sum_i R_{li} x_i(p) \right) \\
&= \;\; \sum_i \alpha_i \, \log \alpha_i \; + \; \sum_l c_l p_l \; - \; \sum_i q_i(p) x_i(p) \; - \; \sum_i \alpha_i \log \left(\sum_l R_{li} p_l \right) \\
&= \;\; \sum_i \alpha_i \, (\log \alpha_i - 1) \; + \; \sum_l c_l p_l \; - \; \sum_i \alpha_i \log \left(\sum_l R_{li} p_l \right)
\end{aligned}$$

where the second and last equalities follow from

$$\sum_l \sum_i R_{li}\, p_l\, x_i(p) \quad = \quad \sum_i q_i(p)\, x_i(p) \quad = \quad \sum_i \alpha_i$$

If W is unbounded there must exists a link k and a sequence of prices $p^1, p^2, \ldots,$ such that $p^j \in W$ and $p_k^j > j$ for all integers j, i.e., the sequence $p^1, p^2, \ldots,$ in W marches off to infinity along its link k component. The linear term $c_k p_k$ in the $V(p)$ expression above dominates over the log term and hence $V(p^j)$ will become unbounded as $j \to \infty$. This contradicts $V(p^j) \leq V(p^0)$ and hence the sequence p^j cannot be in W, a contradiction.[9]

In summary g in (1.17a) is locally Lipschitz on any domain D and we have constructed a compact set W dependent on the given initial point p^0 that contains any solution of (1.17) if it exists. Theorem 1.10.2 therefore guarantees that indeed a unique solution $(p(t), t \geq 0)$ exists.

1.5.3 APPENDIX: PROOF OF LEMMA 1.3

Since all norms in $\mathbb{R}^n$ are equivalent we can without loss of generality assume the norm $\|\cdot\|$ is Euclidean (see Footnote 11). We however present a more involved proof with general norm because it illustrates several techniques that are useful in applications, especially a generalization of the mean value theorem for vector-valued functions under any norm (Lemma 1.12).

Dual norm. Consider any norm $\|\cdot\|$ on $\mathbb{R}^n$ and define its dual norm $\|\cdot\|_*$ by: for any $x \in \mathbb{R}^n$

$$\|x\|_* \quad := \quad \max_{y:\|y\|=1} x^T y$$

The maximization is attained since inner product is continuous and the feasible set is compact. If we think of x^T as an $1 \times n$ matrix then $\|x\|_*$ is the matrix norm induced by the general vector norm $\|\cdot\|$ on $\mathbb{R}^n$. A very useful inequality is:

$$x^T y \quad \leq \quad \|x\| \cdot \|y\|_* \qquad \forall x, y \in \mathbb{R}^n \tag{1.18a}$$

which follows directly from the definition of the dual norm. Since this holds for all x, y, it is equivalent to

$$\left|x^T y\right| \quad \leq \quad \|x\| \cdot \|y\|_* \qquad \forall x, y \in \mathbb{R}^n \tag{1.18b}$$

because one can always choose $-x$ in place of x if necessary. Moreover, given any $x \in \mathbb{R}^n$, there is a normalized $y(x) \in \mathbb{R}^n$ that attains equality, i.e.,

$$x^T y(x) \ = \ \|x\|_* \quad \text{and} \quad \|y(x)\| \ = \ 1 \tag{1.18c}$$

[9]This proves that $V(p)$ is radially unbounded; see also Remark 3.17 in Chapter 3.2.

A crucial fact for $\mathbb{R}^n$ is that the dual of a dual norm is the original norm, i.e., $\|\cdot\|_{**} = \|\cdot\|$. This and (1.18c) imply: given any $x \in \mathbb{R}^n$ there exists an $y(x) \in \mathbb{R}^n$ such that

$$x^T y(x) \;=\; \|x\| \quad \text{and} \quad \|y(x)\|_* \;=\; 1 \tag{1.18d}$$

because

$$\|x\| \;=\; \|x\|_{**} \;=\; \max_{y:\|y\|_*=1} x^T y \;=\; x^T y(x)$$

where $y(x)$ is a maximizer (which clearly exists).[10] The inequalities (1.18a)(1.18b) say that the inner product of any two vectors are upper bounded by the product of the norm of one of the vectors and its dual norm of the other vector. For Euclidean norm this is the Cauchy-Schwarz inequality, but (1.18a) holds for *any* norm. More remarkably, (1.18c)(1.18d) say both the norm and its dual norm of any vector can be attained by the inner product of the vector with some other vector, again for any norm that, unlike the Euclidean norm, may not be defined by inner product.

These properties (1.18) are useful in applications and we use them to prove Lemma 1.3. For instance, together with the Hölder inequality, (1.18a) implies that p and q norms are dual of each other where $p, q \in [1, \infty]$ and $1/p + 1/q = 1$ (if $p = \infty$ then $q := 1$).

Lemma 1.11 *Let $p, q \in [1, \infty]$ and $1/p + 1/q = 1$. The p-norm and the q-norm are dual of each other.*

Proof of Lemma 1.11. We prove the case of $1 < p < \infty$; the case of $p = 1$ or $p = \infty$ follows a similar idea. For $\mathbb{R}^n$, the Hölder inequality is

$$\sum_i |x_i y_i| \;\leq\; \|x\|_q \, \|y\|_p \qquad \forall x, y \in \mathbb{R}^n$$

Hence

$$\|x\|_q \;\geq\; \max_{y:\|y\|_p=1} \sum_i |x_i y_i| \;\geq\; \max_{y:\|y\|_p=1} \sum_i x_i y_i \;=\; \|x\|_*$$

Therefore $\|x\|_q \geq \|x\|_*$. To prove the reverse inequality we have from (1.18a)

$$\|x\|_* \;\geq\; (\|y\|_p)^{-1} \sum_i x_i y_i \;=\; \left(\sum_i |y_i|^p \right)^{-1/p} \sum_i x_i y_i \qquad \forall y \in \mathbb{R}^n$$

[10]For the p-norm the dual is the q-norm with $p^{-1} + q^{-1} = 1$ (see Lemma 1.11) and

$$(y(x))_i \;:=\; \frac{x_i^{p-1}}{\|x\|_p^{p-1}} \operatorname{sign}\left((x_i)^p\right)$$

so that $x^T y(x) = \|x\|_p$ and $\|y(x)\|_q = 1$.

Since $q = 1 + \frac{q}{p}$, choose

$$y_i \quad := \quad |x_i|^{q/p}\,\mathrm{sign}(x_i)$$

so that the above inequality becomes:

$$\|x\|_* \quad \geq \quad \left(\sum_i |x_i|^q\right)^{-1/p} \sum_i |x_i|^{1+q/p} \quad = \quad \left(\sum_i |x_i|^q\right)^{\frac{1}{q}} \quad = \quad \|x\|_q$$

Hence $\|x\|_* = \|x\|_q$ when $\|\cdot\| = \|\cdot\|_p$. $\square$

Lemma 1.11 and (1.18a) imply the following useful special cases:

$$\begin{aligned}
x^T y \quad &\leq \quad \|x\|_p\,\|y\|_q \qquad \left(p^{-1} + q^{-1} = 1\right) \\
x^T y \quad &\leq \quad \|x\|_2\,\|y\|_2 \qquad (p = q = 2, \text{ Cauchy-Schwarz inequality}) \\
\|x\|_2^2 \quad &\leq \quad \|x\|_1\,\|x\|_\infty \qquad (y := x, p = 1, q = \infty)
\end{aligned}$$

We now use (1.18a) and (1.18d) to prove the mean value theorem for vector-valued functions, which implies Lemma 1.3.

Lemma 1.12 *Consider any differentiable function $f : \mathbb{R}^n \to \mathbb{R}^m$. Given any x, y, w in $\mathbb{R}^n$ we have*

$$w^T(f(y) - f(x)) \quad = \quad w^T \frac{\partial f}{\partial x}(z)\,(y - x) \tag{1.19a}$$

$$\|f(y) - f(x)\| \quad \leq \quad \left\|\frac{\partial f}{\partial x}(z)\right\|\,\|y - x\| \tag{1.19b}$$

where $z := \alpha x + (1-\alpha)y$ for some $\alpha \in [0,1]$, $\|\cdot\|$ is any norm, and for matrix, it denotes the induced norm.

Proof of Lemma 1.12. Fix any x, y, w in $\mathbb{R}^n$. Let $z(\alpha) := (1-\alpha)x + \alpha y$ for $\alpha \in [0,1]$ so that $z(0) = x$ and $z(1) = y$, and $z(\alpha)$ traces the straight path from x to y. Define the function

$$g(\alpha) \quad := \quad g_w(\alpha) \quad := \quad w^T f(z(\alpha))$$

as a function of $\alpha \in [0,1]$. Since g is from $\mathbb{R}$ to $\mathbb{R}$ the standard mean value theorem implies that

$$g(1) - g(0) \quad = \quad g'(\beta)$$

for some $\beta \in [0,1]$ that depends on w. Since $g(0) = w^T f(x)$ and $g(1) = w^T f(y)$ this becomes (using chain rule)

$$w^T(f(y) - f(x)) \quad = \quad w^T \frac{\partial f}{\partial x}(z(\beta))\,(y - x)$$

proving (1.19a).

To prove (1.19b), use (1.18d) to choose $w \in \mathbb{R}^n$ such that[11]

$$w^T(f(y) - f(x)) \;=\; \|f(y) - f(x)\| \quad \text{and} \quad \|w\|_* = 1$$

Substituting this w into (1.19a) yields

$$\begin{aligned} \|f(y) - f(x)\| &= w^T(f(y) - f(x)) \;=\; w^T \frac{\partial f}{\partial x}(z(\beta))\,(y - x) \\ &\leq \|w\|_* \cdot \left\| \frac{\partial f}{\partial x}(z(\beta))(x - y) \right\| \\ &\leq \left\| \frac{\partial f}{\partial x}(z(\beta)) \right\| \cdot \|x - y\| \end{aligned}$$

proving (1.19b). In the above, the first inequality follows from (1.18a) and the second inequality follows from the definition of the induced norm of $\frac{\partial f}{\partial x}$. This completes the proof of Lemma 1.12. □

Proof of Lemma 1.3. To prove Lemma 1.3.1 fix any $x_0 \in D$. Take any $r > 0$ such that the convex and compact set $B_r(x_0) \subseteq D$. For any $x, y \in B_r(x_0)$ we have from (1.19b)

$$\|f(y) - f(x)\| \;\leq\; \left\| \frac{\partial f}{\partial x}(z) \right\| \|y - x\|$$

for some $z \in B_r(x_0)$ between x and y and for any norm $\|\cdot\|$. By assumptions in Lemma 1.3, $\frac{\partial f}{\partial x}$ not only exists but is also continuous. Since $B_r(x_0)$ is compact the maximization in

$$L \;:=\; L(x_0) \;:=\; \max_{z \in B_r(x_0)} \left\| \frac{\partial f}{\partial x}(z) \right\|$$

is attained. Hence f is locally Lipschitz at x_0 with Lipschitz constant L, as desired.

For Lemma 1.3.2, let

$$L \;:=\; \sup_{z \in D} \left\| \frac{\partial f}{\partial x}(z) \right\|$$

Suppose L is finite. Since D is convex, the above argument goes through with $B_r(x_0)$ replaced by D to prove that f is Lipschitz with L. Conversely suppose f is Lipschitz with a

[11] If the norm $\|\cdot\|$ is Euclidean then the argument below simplifies to: setting $w := f(y) - f(x)$ in (1.19a) yields

$$\begin{aligned} \|f(y) - f(x)\|_2^2 &= (f(y) - f(x))^T \frac{\partial f}{\partial x}(z(\beta))\,(y - x) \\ &\leq \|f(y) - f(x)\|_2 \cdot \left\| \frac{\partial f}{\partial x}(z(\beta)) \right\|_2 \|y - x\|_2 \end{aligned}$$

proving (1.19b). This is done in [23].

Lipschitz constant K, i.e., $\|f(x) - f(y)\| \leq K\|x - y\|$ for all $x, y \in D$. We now argue that L must be finite. Assume for the sake of contradiction that for each integer $k > 0$ there exists an $x_k \in D$ such that the induced norm

$$\left\| \frac{\partial f}{\partial x}(x_k) \right\| \quad := \quad \max_{\|y\| \neq 0} \frac{\left\| \frac{\partial f}{\partial x}(x_k)\, y \right\|}{\|y\|} \quad > \quad k$$

Let $y_k \neq 0$ attain the maximization above, i.e., for all integers $k > 0$, there exist $x_k \in D$ and $y_k \in \mathbb{R}^n$ such that

$$\left\| \frac{\partial f}{\partial x}(x_k)\, y_k \right\| \quad > \quad k\, \|y_k\| \tag{1.20a}$$

Apply (1.18d) to (1.20a) to choose an $w_k \in \mathbb{R}^n$ such that

$$w_k^T \frac{\partial f}{\partial x}(x_k)\, y_k \;=\; \left\| \frac{\partial f}{\partial x}(x_k)\, y_k \right\| \quad \text{and} \quad \|w_k\|_* \;=\; 1$$

i.e., for all integers $k > 0$ there exist $x_k \in D$ and y_k, w_k in $\mathbb{R}^n$ such that

$$w_k^T \frac{\partial f}{\partial x}(x_k)\, y_k \;>\; k\, \|y_k\| \quad \text{and} \quad \|w_k\|_* \;=\; 1 \tag{1.20b}$$

Fix any $\epsilon > 0$. Since $\frac{\partial f}{\partial x}$ is continuous there is a $\delta_k > 0$ such that for all integers $k > 0$

$$\left\| \frac{\partial f}{\partial x}(z) - \frac{\partial f}{\partial x}(x_k) \right\| \quad < \quad \epsilon \qquad \text{for any } z \in B_{\delta_k}(x_k) \;\subseteq\; D \tag{1.20c}$$

Note that y_k in (1.20a) can be chosen with arbitrarily small norm $\|y_k\| > 0$. Hence we can choose $\hat{x}_k$ in the ball $B_{\delta_k}(x_k)$ such that $\hat{x}_k - x_k = y_k$ for an $0 < \|y_k\| < \delta_k$. Then we have by (1.19a)

$$\begin{aligned} w_k^T (f(\hat{x}_k) - f(x_k)) \quad &= \quad w_k^T \frac{\partial f}{\partial x}(\hat{z}_k)\, (\hat{x}_k - x_k) \\ &= \quad w_k^T \frac{\partial f}{\partial x}(x_k)\, y_k \;+\; w_k^T \left(\frac{\partial f}{\partial x}(\hat{z}_k) - \frac{\partial f}{\partial x}(x_k) \right) y_k \end{aligned} \tag{1.21}$$

for some $\hat{z}_k \in B_{\delta_k}(x_k)$ between $\hat{x}_k$ and x_k.

We now examine each term in (1.21). On the left-hand side, (1.18a) implies

$$w_k^T (f(\hat{x}_k) - f(x_k)) \quad \leq \quad \|w_k\|_* \cdot \|f(\hat{x}_k) - f(x_k)\| \quad = \quad \|f(\hat{x}_k) - f(x_k)\| \tag{1.22}$$

by the choice of w_k. On the right-hand side, the first term is lower bounded in (1.20b). The magnitude of the second term can be upper bounded by:

$$\begin{aligned} \left| w_k^T \left(\frac{\partial f}{\partial x}(\hat{z}_k) - \frac{\partial f}{\partial x}(x_k) \right) y_k \right| \quad &\leq \quad \|w\|_* \cdot \left\| \left(\frac{\partial f}{\partial x}(\hat{z}_k) - \frac{\partial f}{\partial x}(x_k) \right) y_k \right\| \\ &\leq \quad \left\| \frac{\partial f}{\partial x}(\hat{z}_k) - \frac{\partial f}{\partial x}(x_k) \right\| \cdot \|y_k\| \\ &< \quad \epsilon\, \|y_k\| \end{aligned}$$

where the first inequality follows from (1.18b), the second inequality from the definition of matrix norm, and the last inequality from (1.20c) since $\hat{z}_k \in B_{\delta_k}(x_k)$. Hence the right-hand side of (1.21) satisfies

$$w^T \frac{\partial f}{\partial x}(x_k)\, y_k \;+\; w^T \left(\frac{\partial f}{\partial x}(\hat{z}_k) - \frac{\partial f}{\partial x}(x_k) \right) y_k \quad > \quad (k - \epsilon)\, \|y_k\|$$

Substituting this and (1.22) into (1.21), we have, for all integers $k > 0$ there exist $x_k, \hat{x}_k \in B_{\delta_k}(x_k) \;\subseteq\; D$ such that

$$\|f(\hat{x}_k) - f(x_k)\| \quad > \quad (k - \epsilon)\|\hat{x}_k - x_k\|$$

For large enough k, this contradicts that f is Lipschitz with Lipschitz constant K. This completes the proof of Lemma 1.3.2.

Lemma 1.3.3 follows from Lemma 1.3.2 since L is finite if D is compact. Lemma 1.3.4 follows from Lemma 1.3.2 with $D = \mathbb{R}^n$. This completes the proof of the lemma. □

1.5.4 APPENDIX: PROOF OF THEOREM 1.10

Consider the projected dynamical system

$$\dot{x} \quad = \quad \pi_K(x(t), f(x(t))), \qquad x(0) = x_0 \;\text{ given} \tag{1.23}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ and π denotes the projection such that $x(t)$ stays in the closed convex polyhedron K. Informally $\pi(x, v) = v$ if x is in the interior of K or if v points into K, and, if x is on the boundary of K and v points out of K, $\pi(x, v) = v + \gamma$ where γ is an inward normal to K at x such that a solution of (1.23) stays in K. [12] The system (1.13) is a special case of (1.23) where K is the nonnegative quadrant $\mathbb{R}^n_+$, i.e.,

$$\dot{x} \quad = \quad \pi_{\mathbb{R}^n_+}(x(t), f(x(t))) \quad = \quad (f(x(t)))^+_{x(t)}$$

For the proof in this subsection we assume $K \subseteq \mathbb{R}^n_+$. Consider the following condition:

C1.1: There exists an $B < \infty$ such that f satisfies

1. $\|f(x)\| \;\leq\; B(1 + \|x\|)$ for all $x \in K$, and
2. $(f(x) - f(y))^T (x - y) \;\leq\; B\|x - y\|^2$ for all $x, y \in K$.

Here $\|\cdot\|$ can be any norm since all norms are equivalent in $\mathbb{R}^n$ (the constant B may depend on the chosen norm). These conditions are called the linear growth condition and the one-sided Lipschitz condition. The standard Lipschitz condition $\|f(x) - f(y)\| \leq L\|x - y\|$

[12] See [20, equation (5)] for a precise definition of π.

implies both the one-sided Lipschitz condition and continuity of f, but the one-sided Lipschitz condition and continuity do not imply each other. In particular f can be discontinuous and satisfies C1.1. Hence the right-hand side of (1.23) can be discontinuous both because of the projection π and because $f(x)$ is discontinuous (even though Theorem 1.10 assumes f to be either globally or locally Lipschitz).

Recall that a solution to (1.23) is defined to be a function $(x(t), t \geq 0)$ that is absolutely continuous and satisfies (1.23) almost everywhere. Theorem 1.10 follows from the next result proved in [20, Theorems 2 and 3].

Lemma 1.13 *Suppose C1.1 holds. Then there is a unique solution* $(x(t), t \geq 0)$ *to* (1.23).

To prove Theorem 1.10 we first show that if f is Lipschitz over the nonnegative quadrant then f satisfies C1.1 with $K = \mathbb{R}^n_+$ and hence a unique solution to (1.13) exists (Theorem 1.10.1). Then we show that if f is only locally Lipschitz but any solution, when exists, stays in a compact set then a unique solution exists as well (Theorem 1.10.2).

Proof of Theorem 1.10.1. First suppose f in (1.13) is Lipschitz over the nonnegative quadrant $K := \mathbb{R}^n_+$, i.e., there exists $L < \infty$ such that

$$\|f(x) - f(y)\| \quad \leq \quad L\|x - y\|, \qquad x, y \in K$$

under any norm $\|\cdot\|$. Since f is defined on $\mathbb{R}^n$, $f(0)$ is finite. Hence we have for any $x \in K$

$$\|f(x)\| \quad \leq \quad \|f(x) - f(0)\| + \|f(0)\| \quad \leq \quad L\|x\| + \|f(0)\| \quad \leq \quad B(1 + \|x\|)$$

for any finite $B \geq \max\{L, \|f(0)\|\}$, where the second inequality follows from the Lipschitz continuity of f. Moreover, for any $x, y \in K$, the Cauchy-Schwarz inequality implies

$$(f(x) - f(y))^T(x - y) \quad \leq \quad \|f(x) - f(y)\|_2 \cdot \|x - y\|_2$$

Since all norms are equivalent in $\mathbb{R}^n$ we can convert the above inequality from the Euclidean norm $\|\cdot\|_2$ to the norm that defines the Lipschitz continuity of f to obtain

$$(f(x) - f(y))^T(x - y) \quad \leq \quad a\,\|f(x) - f(y)\| \cdot \|x - y\| \quad \leq \quad aL\|x - y\|^2$$

for some finite constant a, where the second inequality follows from the Lipschitz continuity of f. Hence f satisfies condition C1.1, and Theorem 1.10.1 follows from Lemma 1.13.

Proof of Theorem 1.10.2. Suppose now that f is only *locally* Lipschitz over a domain $D \subseteq \mathbb{R}^n_+$ under any norm, but any solution to (1.13), if exists, lies entirely in a compact set $W \subseteq D$ that contains the initial point x_0. We will prove the theorem through two lemmas.

Lemma 1.14 *The function f is Lipschitz on the compact set $W \subseteq D$, i.e., there exists $L < \infty$ such that*

$$\|f(x) - f(y)\| \quad \leq \quad L\|x - y\|, \qquad x, y \in W$$

Lemma 1.15 *Suppose f is Lipschitz on W under any norm. There is a function $\hat{f} : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ that is Lipschitz on $\mathbb{R}^n_+$ (with possibly a different Lipschitz constant if $\|\cdot\|$ is not the Euclidean norm) and coincides with f over W, i.e.,*

$$\hat{f}(x) \quad = \quad f(x), \qquad x \in W$$

Then since any solution $(x(t), t \geq 0)$ to (1.13), if it exists, lies entirely in W, the solutions to the system (1.13) coincide with the solutions to the following system

$$\dot{x} \quad = \quad \left(\hat{f}(x(t))\right)^+_{x(t)}, \qquad x(0) = x_0 \;\text{ given} \tag{1.24}$$

Since $\hat{f}$ is Lipschitz on $\mathbb{R}^n_+$ (with possibly a different Lipschitz constant from that of f), by Theorem 1.10.1, there is indeed a unique solution $(x(t), t \geq 0)$ to (1.24) and hence to (1.13). Therefore the proof of the theorem will be complete after Lemmas 1.14 and 1.15 are proved.

Proof of Lemma 1.14. Since W is compact, D is open, $W \subseteq D$, and f is locally Lipschitz in D we can construct an open cover in D of W such that f is Lipschitz on each open set in the cover, as follows. For each $x \in W$ there is an $r(x) > 0$ and $L(x) < \infty$ such that the *open* ball $B_{r(x)}(x) \subseteq D$ and

$$\|f(y) - f(z)\| \quad \leq \quad L(x)\|y - z\|, \qquad y, z \in B_{r(x)}(x)$$

Clearly the collection $\{B_{r(x)}(x) : x \in W\}$ is such an open cover with $W \subset \cup_{x \in W} B_{r(x)}(x) \subseteq D$. Since W is compact there is a finite subcover, i.e., there exists a finite collection $\{B_i := B_{r(x_i)}(x_i),\, x_i \in W\}$ such that

$$W \quad \subseteq \quad \cup_i B_i \quad \subseteq \quad D$$

We now prove by contradiction that f is Lipschitz on the compact set W.

Suppose not and for any integer $k > 0$ there are $y_k \neq z_k$ in W such that $\|f(y_k) - f(z_k)\| > k\,\|y_k - z_k\|$. Since W is compact there exists convergent subsequences y_{k_j} and z_{k_j} such that $\lim_j y_{k_j} = y$ and $\lim_j z_{k_j} = z$ with $y, z \in W$.[13] This implies three mutually exclusive cases illustrated in Figure 1.9, each of which we now argue violates

$$\|f(y_{k_j}) - f(z_{k_j})\| \quad > \quad k_j\,\|y_{k_j} - z_{k_j}\| \qquad \text{for all } j \tag{1.25}$$

leading to a contradiction.

[13] Even though $y_{k_j} \neq z_{k_j}$ for all j, it is possible that all convergent subsequences have $y = z$. Hence we cannot take limit in (1.25) and use $\|f(y) - f(z)\| \;\geq\; \lim_j\, k_j\,\|y_{k_j} - z_{k_j}\|$.

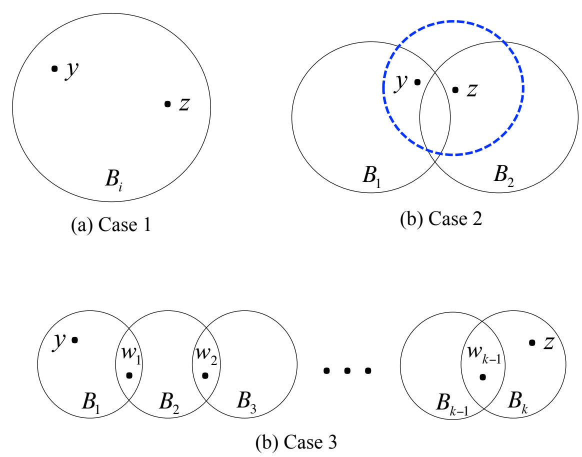


Figure 1.9: Proof of Lemma 1.14.

(a) Suppose $y, z \in B_i$ for some B_i over which f is Lipschitz. Since $\lim_j y_{k_j} = y$ and $\lim_j z_{k_j} = z$, B_i contains all but finitely many y_{k_j}, z_{k_j}. The Lipschitz continuity of f over B_i then violates (1.25) since $y_{k_j} \neq z_{k_j}$ for all j.

(b) Suppose y and z are in different covering sets but y, z are "close". Specifically since f is locally Lipschitz at y, there is an L_y and a $\delta_y > 0$ such that $\|f(a) - f(b)\| \leq L_y \|a - b\|$ for all a, b in $B_{\delta_y}(y)$. Similarly there is an L_z and a $\delta_z > 0$ such that $\|f(a) - f(b)\| \leq L_z \|a - b\|$ for all a, b in $B_{\delta_z}(z)$. If either $z \in B_{\delta_y}(y)$ or $y \in B_{\delta_z}(z)$ then (1.25) is violated (since $B_{\delta_y}(y)$ or $B_{\delta_z}(z)$ contains all but finitely many y_{k_j}, z_{k_j}).

(c) Suppose finally that y and z are in different covering sets B_1 and B_k respectively and $\|y - z\| > \min\{\delta_y, \delta_z\} =: r > 0$. Then taking limit in j on both sides of (1.25) yields

$$\|f(y) - f(z)\| \quad \geq \quad \left(\lim_j k_j\right)\left(\lim_j \|y_{k_j} - z_{k_j}\|\right) \quad \geq \quad r\left(\lim_j k_j\right) \qquad (1.26)$$

Since W is connected (because D is connected by definition) there is a sequence of covering sets, say, $B_1, \ldots, B_k$ and $w_1, \ldots, w_{k-1}$ such that

$$y \in B_1, w_1 \in B_1 \cap B_2, w_2 \in B_2 \cap B_3, \ldots, w_{k-1} \in B_{k-1} \cap B_k, z \in B_k$$

By construction f is Lipschitz with Lipschitz constant $L(x_i)$ over each open ball B_i of radius $r(x_i)$. We have

$$\begin{aligned}
\|f(y) - f(z)\| &= \|(f(y) - f(w_1)) + (f(w_1) - f(w_2)) + \cdots + (f(w_{k-1}) - f(z))\| \\
&\leq \|f(y) - f(w_1)\| + \|f(w_1) - f(w_2)\| + \cdots + \|f(w_{k-1}) - f(z)\| \\
&\leq L(x_1)\|y - w_1\| + L(x_2)\|w_1 - w_2\| + \cdots + L(x_k)\|w_{k-1} - z\| \\
&\leq 2\left(L(x_1)r(x_1) + \cdots + L(x_k)r(x_k)\right) \quad < \quad \infty
\end{aligned}$$

This contradicts (1.26).

This proves that f is Lipschitz on the compact set W, as desired. □

Proof of Lemma 1.15. Given a function $g : A \to H_2$ where A is a subset of a Hilbert space H_1 and H_2 is another Hilbert space.[14] If g is Lipschitz on A under the norm induced by the inner product (Euclidean norm for $\mathbb{R}^n$), then there is a function $\hat{g} : H_1 \to H_2$ that extends g (i.e., $\hat{g}(x) = g(x)$ for $x \in A$) and is Lipschitz on H_1 with the same Lipschitz constant as g. This is called the Kirszbraun theorem. We now apply this to our problem.

[14] A Hilbert space is a inner product space that is complete under the norm induced by the inner product. Examples are $\mathbb{R}^n, \mathbb{R}^n_+$ or any linear subspace of $\mathbb{R}^n$ endowed with the Euclidean norm (induced by the inner product on $\mathbb{R}^n$).

Lemma 1.14 shows that f is Lipschitz on $W \subseteq \mathbb{R}^n_+$ under some norm. Since all norms are equivalent in $\mathbb{R}^n$, f is Lipschitz on W *under the Euclidean norm* with a possibly different Lipschitz constant. The Kirszbraun theorem then implies that there is a Lipschitz-continuous extension $\hat{f} : \mathbb{R}^n_+ \to \mathbb{R}^n_+$ with $\hat{f} \equiv f$ on W.

□

Applying Theorem 1.10.1 to $\hat{f}$ completes the proof of Theorem 1.10.2. □

1.6 BIBLIOGRAPHICAL NOTES

The Transmission Control Protocol (TCP) was published by Vinton Cert and Robert Kahn in 1974 [14] and deployed as the TCP/IP standard on the ARPANet (precursor of the Internet) by 1983. There are many textbooks that describe in detail transport protocol services and different congestion control protocols in TCP e.g. [16, 45, 31]. The first congestion control algorithm on the Internet was implemented by Jacobson in Tahoe (1988) and Reno (1990) versions of TCP [26] based on the Additive Increase Multiplicative Decrease (AIMD) idea of [27]. The TCP Vegas is proposed in [12]. A mathematical model is introduced, validated and analyzed in [37]. TCP FAST is proposed and analyzed in [50] (including that the link model (1.6b) is well defined). The RED algorithm is proposed in [22] and REM in [4]. There are many other AQM proposals, including the PI controller [25] and AVQ [30]. Heterogeneous protocols where sources react to different types of congestion prices, e.g. some react to packet loss and others to packet delay are analyzed in [46, 47]. See [36] for an overview of Internet congestion control models.

There are many good texts on dynamical systems described by ordinary differential equations. The materials in Chapter 1.5.1 on the existence and uniqueness of solutions to Lipschitz continuous ODEs are mainly taken from [29, Chapter 3]. Its application to the TCP/AQM algorithm in Chapter 1.5.2 is new. Projected dynamics (1.13) is a simple kind of discontinuous ODE systems. For projection onto any closed convex polyhedron, the existence and uniqueness of the solution are established in [20, Theorems 2 and 3]. This is used to prove Theorem 1.10 in Appendix 1.5.4 (see [24] for the Kirszbraun theorem). It is also proved in [20, Lemma 2] that the solution is continuous in its initial state, which is used to prove LaSalle's invariance principle in Chapter 3.1 for projected dynamics. See [17] for a tutorial on general discontinuous ODE systems, from which Examples 1.8 and 1.9 are taken.

1.7 PROBLEMS

Exercise 1.1. Show that if f is locally Lipschitz at x_0 then it is continuous at x_0, i.e., given any $\epsilon > 0$ there exists a $\delta = \delta(\epsilon) > 0$ such that $\|f(x) - f(x_0)\| < \epsilon$ for all $x \in B_\delta(x_0)$.

Exercise 1.2. Consider $\dot{x} = f(x(t))$ for $t \geq 0$ where $f(x) = \sqrt{x}$, $x \geq 0$.

1. Show that $f(x)$ is not locally Lipschitz around 0.

2. Suppose $x(0) = 0$. Since f is continuous, solutions exists. Show that it is not unique by exhibiting two distinct solutions $(x(t), t \geq 0)$.

Exercise 1.3. Suppose functions f_1 and f_2 are globally Lipschitz with constants L_1 and L_2 respectively, i.e.,

$$\|f_1(y) - f_1(x)\| \ \leq \ L_1\|y - x\| \quad \text{and} \quad \|f_2(y) - f_2(x)\| \ \leq \ L_2\|y - x\|$$

Prove

1. $f_1 + f_2$ is Lipschitz;
2. $f_2 \circ f_1$ is Lipschitz;
3. $f_1 f_2$ is not necessarily Lipschitz, unless both are bounded.

Exercise 1.4. Consider the following simplified version of the Reno/RED model:

$$\begin{aligned}
\dot{x}_i &= \left(\frac{1}{T_i^2} - \frac{1}{2}\, q_i(t)\, x_i^2(t)\right)^+_{x_i(t)} \\
\dot{b}_l &= (y_l(t) - c_l)^+_{b_l(t)} \\
p_l(t) &= \min\{\rho_l b_l(t), 1\}
\end{aligned}$$

where for any $a, b \in \mathbb{R}$, $(a)^+_b := a$ if $a > 0$ or $b > 0$ and 0 otherwise, and $q_i(t) = \sum_l R_{li} p_l(t)$ and $y_l(t) := \sum_i R_{li} x_i(t)$. We can eliminate the prices $p_l(t)$ to obtain an ODE model involving only $(x(t), b(t))$:

$$\dot{x}_i = \left(\frac{1}{T_i^2} - \frac{1}{2}\, x_i^2(t) \sum_l R_{li} \, \min\{\rho_l b_l(t), 1\}\right)^+_{x_i(t)} =: [f_i\,(x_i(t), b(t))]^+_{x_i(t)} \tag{1.27a}$$

$$\dot{b}_l = \left(\sum_i R_{li} x_i(t) - c_l\right)^+_{b_l(t)} =: [g_l\,(x(t), b_l(t))]^+_{b_l(t)} \tag{1.27b}$$

Prove that the function (f, g) defined by the right-hand side of (1.27) is Lipschitz on any compact $D \in \mathbb{R}^{N+L}$.[15]

[15] To prove that given any initial state $(x(0), b(0)) = (x_0, b_0)$ the system (1.27) has a unique solution $(x(t), b(t),\ t \geq 0)$, we need to show, in addition, that there exists a compact subset $W := W(x_0, b_0)$ of D containing (x_0, b_0) such that every solution $(x(t), b(t), t \geq 0)$ of (1.27), if exists, lies in W. Theorem 1.10.2 then guarantees the existence and uniqueness of solution to (1.27) given any initial state (x_0, b_0).

CHAPTER 2

Equilibrium structure

Consider a set of ordinary differential equations (ODEs):

$$\dot{x} \;=\; f(x(t)), \quad t \geq 0, \qquad x(0) \;=\; x_0 \tag{2.1}$$

where $f : D \to \mathbb{R}^n$ and $D \subseteq \mathbb{R}^n$ is a domain (open and connected set). Here f may not be continuous, e.g., (2.1) can be a projected dynamical system.

Definition 2.1 A point $x^* \in D$ is an *equilibrium* of (2.1) if $f(x^*) = 0$.

In this chapter we study the equilibria of the family of primal-dual algorithms modeled by a set of ODEs (the basic model (1.10)). It turns out that, under mild assumptions, the equilibrium (x^*, p^*) of a primal-dual algorithm has a very simple characterization: x^* is the unique maximizer of the following convex problem called *network utility maximization*:

$$\max_{x \geq 0} \quad \sum_i U_i(x_i) \qquad \text{subject to} \quad Rx \leq c$$

and p^* is a minimizer of the associated Lagrangian dual problem. In this chapter we explain this equilibrium structure and explore some of its implications.

2.1 CONVEX OPTIMIZATION

We start by introducing some basic concepts in convex optimization.

2.1.1 CONVEX PROGRAM

A convex program is defined by a convex set and a convex function, as we now define.

Convex set. A set is called convex if, given any two points in the set, every point in between lies in the set.

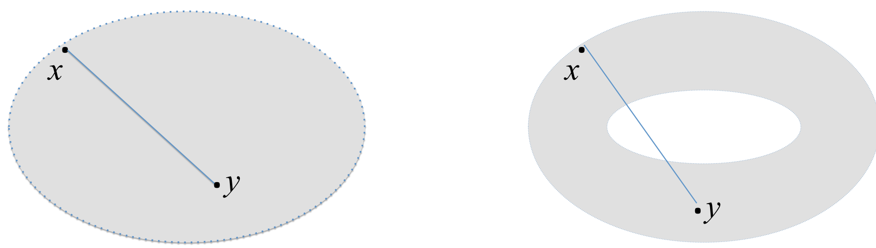
Definition 2.2 A set $D \subseteq \mathbb{R}^n$ is *convex* if, given any $x, y \in D$,

$$\alpha x + (1-\alpha) y \;\in\; D, \qquad \forall \alpha \in [0, 1]$$

For instance for any $x_0 \in D$ there exists $r > 0$ such that the r-ball around x_0,

$$B_r(x_0) \quad := \quad \{\, x \in D \mid \|x - x_0\|_2 \le r \,\}$$

is contained in D, where $\|x\|_2 := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ is the Euclidean norm. Moreover $B_r(x_0)$ is convex for any $r > 0$, $x_0 \in D$. The definition is illustrated in Figure 2.1.



(a) Convex set. (b) Nonconvex set.

Figure 2.1: Definition of a convex set: every point in between two points in the set lies in the set.

Three types of convex sets are the most useful in engineering applications. First is a set specified by linear inequalities:

$$\text{Affine set:} \qquad C \quad := \quad \{x \in \mathbb{R}^n \mid Ax \le b\}, \quad n \ge 1$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, $m \ge 1$. The second is a second-order cone (SOC) defined as:

$$\text{SOC:} \qquad C \quad := \quad \{(x,t) \in \mathbb{R}^{n+1} \mid \|x\|_2 \le t\}, \quad n \ge 1$$

A ball $B_t(0)$ is a cross section of the second-order cone defined by $\|x\|_2 \le t$ for a fixed t. The third is a set of semidefinite matrices defined as follows. A real matrix $X \in \mathbb{R}^{n \times n}$ is *symmetric* if $X = X^T$, i.e., $X_{ij} = X_{ji}$ for all $i, j = 1, \ldots, n$. A real matrix X is *positive semidefinite* (psd) if X is symmetric and $x^T X x = \sum_{i,j} X_{ij} x_i x_j \ge 0$ for all $x \in \mathbb{R}^n$. We write $X \succeq 0$ to denote that X is positive semidefinite. Given a symmetric matrix $X \in \mathbb{R}^{n \times n}$ the following are equivalent:

1. X is positive semidefinite.
2. All eigenvalues of X are nonnegative.
3. $X = BB^T$ for some matrix $B \in \mathbb{R}^{n \times m}$ and some natural number m.

The set of all positive semidefinite matrices is:

$$\text{psd matrices:} \qquad \mathbb{S}_+^n \quad := \quad \{X \in \mathbb{R}^{n \times n} \mid X \succeq 0\}, \quad n \ge 1$$

The proof that these three types of sets are convex is left as an exercise. Efficient algorithms exist to solve constrained optimization problems that minimize a certain cost function over an affine set, second-order cones, or semidefinite matrices.

Given these three basic convex sets we can create other convex sets through simple *convexity-preserving* operations. Let $\mathbb{X}$ and $\mathbb{Y}$ be linear subspaces. For example $\mathbb{X} := \mathbb{R}^n$ and $\mathbb{Y} := \mathbb{R}^m$.

1. *Linear transformation:* Let $f : \mathbb{X} \to \mathbb{Y}$ be linear.

   (a) If $A \subseteq \mathbb{X}$ is convex then $f(A) := \{f(x) \mid x \in A\} \subseteq \mathbb{Y}$ is convex.

   (b) If $B \subseteq \mathbb{Y}$ is convex then $f^{-1}(B) = \{x \mid f(x) \in B\} \subseteq \mathbb{X}$ is convex.

2. *Direct product:* Let $A \subseteq \mathbb{X}$, $B \subseteq \mathbb{Y}$ be convex. Then $A \times B := \{(x, y) \mid x \in A, y \in B\}$ is convex. In fact the direct product of an arbitrary (e.g., uncountably many) number of convex sets is convex.

3. *Finite sum:* Let $A, B \subseteq \mathbb{X}$ be convex. Then $A + B := \{a + b \mid a \in A, b \in B\}$ is convex. Therefore the sum of any finite number of convex sets is convex.

4. *Arbitrary intersection:* Let $A, B \subseteq \mathbb{X}$ be convex. Then the intersection $A \cap B$ is convex. In fact the intersection of an arbitrary collection of (e.g., uncountably many) convex sets is convex.

The proof that these set operations preserve convexity is left as an exercise. In contrast to intersection the union of two convex sets can be nonconvex.

Example 2.3 Consider the ellipsoid

$$E \quad := \quad \{x \in \mathbb{R}^n \mid x^T A x \le c\}$$

where $A \in \mathbb{R}^{n\times n}$ is a psd matrix and $c > 0$. E is convex because it can be derived from an application of convexity-preserving operation on a convex set as follows. Since A is psd it can be expressed as $A := BB^T$ for some $B \in \mathbb{R}^{n\times m}$. Hence $x^T A x = x^T BB^T x = \|B^T x\|_2^2$.

Let $y = B^T x$. Then the set $C := \{(y, t) \in \mathbb{R}^{m+1} \mid \|y\|_2 \le t\}$ is a (convex) SOC. Hence the set $D := \{y \in \mathbb{R}^m \mid \|y\|_2 \le c\}$ is convex since it is the intersection of two convex sets:

$$D \quad = \quad C \cap (\mathbb{R}^m \times \{t = c\})$$

Then $E = f^{-1}(D)$ where $f(x) := B^T x$ is a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$. Hence E is convex as desired.

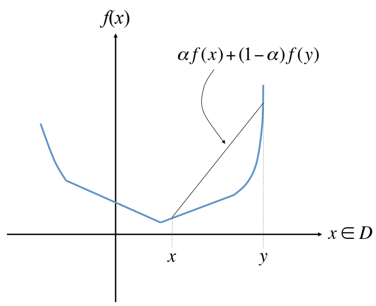
Convex function.

Definition 2.4 A function $f : D \to \mathbb{R}$ defined over a convex domain $D \subseteq \mathbb{R}^n$ is *convex* if, for all $x, y \in D$ and all $\alpha \in [0, 1]$,

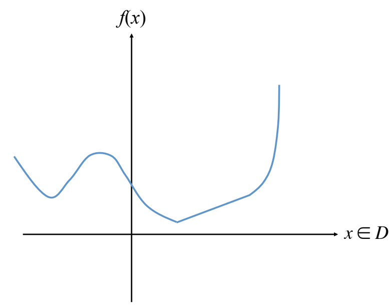
$$f(\alpha x + (1-\alpha)y) \quad \leq \quad \alpha f(x) + (1-\alpha)f(y)$$

It is *strictly convex* if the inequality is strict for $x \neq y$ and $\alpha \in (0, 1)$. A function f is *concave* (*strictly concave*) if $-f$ is convex (strictly convex).

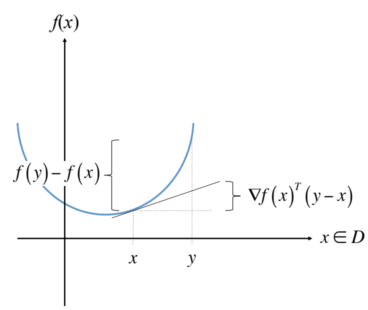
The definition says that the straight line connecting $f(x)$ and $f(y)$ lies above the function f between x and y, as illustrated in Figure 2.2(a).



(a) Convex function.



(b) Nonconvex function.



(c) Differentiable convex function.

Figure 2.2: Definition of a convex function: The straight line connection $f(x)$ and $f(y)$ lies above f between x and y. The linear approximation of a differentiable convex function f lies below f.

Example 2.5 If $f(x) = x^2$ then for any x, y and $\alpha \in [0, 1]$

$$\alpha f(x) + (1-\alpha)f(y) - f(\alpha x + (1-\alpha)y) \quad = \quad \alpha(1-\alpha)\,(x-y)^2 \;>\; 0$$

for $x \neq y$ and $\alpha \in (0,1)$. Hence f is strictly convex. □

Checking if a function is convex by verifying the convexity definition is often difficult. The following theorem provides three different ways to check the convexity of a function. Consider $f : D \to \mathbb{R}$ over a convex domain $D \subseteq \mathbb{R}^n$. Let $\nabla f(x)$ denote the *column* vector of partial derivatives of f (whereas $\frac{\partial f}{\partial x}$ denotes the row vector of partial derivatives). Let

$$\nabla^2 f(x) \;:=\; \frac{\partial^2 f}{\partial x^2} \;:=\; \left[\frac{\partial^2 f}{\partial x_i \, \partial x_j}\right]$$

denote the $n \times n$ Hessian matrix.

Theorem 2.6 *The function f is convex on D if and only if any one of the following holds:*

1. *For a differentiable function f,*

$$f(y) - f(x) \;\geq\; \nabla f(x)^T (y - x), \qquad \forall x, y \in D$$

2. *For a twice differentiable function f,*

$$\nabla^2 f(x) \;\succeq\; 0, \qquad \forall x \in D$$

   i.e., the Hessian matrix is positive semidefinite (all eigenvalues are nonnegative).

3. *For $x \in D$ and all $v \in \mathbb{R}^n$ the function*

$$g(t) \;:=\; f(x + tv)$$

   is convex on $\{t \in \mathbb{R} \mid x + tv \in D\}$.

The first-order condition in Theorem 2.6.1 says that the function f always lies above its linear approximation, i.e., $f(y)$ is always greater than or equal to the tangent plane to f at any point x. This is illustrated in Figure 2.2(c). See Appendix 5.4 for a proof of the first-order condition. The second-order condition in Theorem 2.6.2 roughly says that the gradient at any point x is increasing around x. The condition in Theorem 2.6.3 does not require differentiability of f and says that, if we take any cross section of the surface f defined by (x, v), i.e., from x in the direction of v or $-v$, the corresponding *scalar* function $g(t)$ is convex.

Theorem 2.6 provides an exact characterization for convexity, but not for *strict* convexity. For instance if $\nabla^2 f(x) \succ 0$ for all $x \in D$ then f is strictly convex in D, but the converse may not hold; e.g., $f(x) = x^4$ is strictly convex but $f''(x) = 0$ at $x = 0$.

A common mistake is to confuse the second-order condition in Theorem 2.6.2 that $\nabla^2 f(x)$ is positive semidefinite with the condition that

$$x^T \nabla^2 f(x) x \quad \geq \quad 0 \qquad \text{for all } x \in D$$

For any $x \in D$, $\nabla^2 f(x) \succeq 0$ if and only if

$$y^T \nabla^2 f(x) y \quad \geq \quad 0 \qquad \text{for all } y \in \mathbb{R}^n$$

i.e., regardless of what D is, the test on $\nabla^2 f(x)$ is for *all* $y \in \mathbb{R}^n$. This is illustrated in the next example.

Example 2.7 Consider the function

$$f(x_1, x_2) \quad = \quad x_1 x_2$$

over the domain

$$D \quad := \quad \big\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0\big\}$$

with

$$\nabla^2 f(x) \quad = \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

We have

$$x^T \nabla^2 f(x) x \quad = \quad 2\, x_1 x_2 \quad > \quad 0 \qquad \text{for all } x \in D$$

This however does *not* imply that f is strictly convex over D. The eigenvalues of $\nabla^2 f(x)$ are 1 and -1, and hence f is neither convex nor concave. Indeed the function value along the direction $x_1 = x_2$ corresponding to the eigenvalue-eigenvector pair $\big(1, [1\ 1]^T\big)$ is given by

$$g(t) \quad := \quad f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) \quad = \quad (x_1 + t)(x_2 + t), \qquad t > -\min\{x_1, x_2\}$$

Hence $g(t)$ is convex in t, i.e. f is convex along $x_1 = x_2$. Along the direction $x_1 = -x_2$ corresponding to the eigenvalue-eigenvector pair $\big(-1, [1\ \ -1]^T\big)$ the function value is

$$g(t) \quad := \quad f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + t \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) \quad = \quad (x_1 + t)(x_2 - t), \qquad -x_1 \leq t \leq x_2$$

Therefore $g(t)$ is concave in t, i.e., f is concave along $x_1 = -x_2$. This is illustrated in Figure 2.3. $\square$

Example 2.8 We illustrate Theorem 2.6 using $f(x) = \log x$ for $x > 0$.

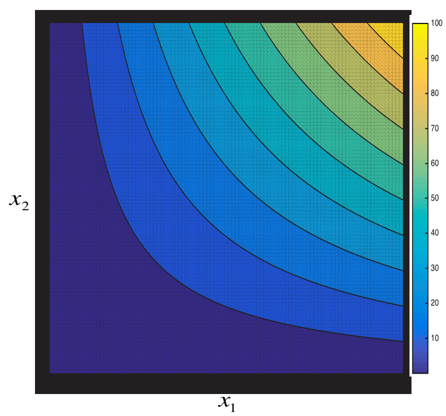


Figure 2.3: Contour plot of $f(x) = x_1 x_2$ which is neither convex nor concave over $D := \{(x_1, x_2) \mid x_1 > 0,\, x_2 > 0\}$.

1. We have $f'(x) = x^{-1}$ and for $x \neq y > 0$ (such that $\frac{y}{x} \neq 1$)

$$f(y) - f(x) \;=\; \log \frac{y}{x} \;<\; \frac{y}{x} - 1 \;=\; \frac{1}{x}(y - x) \;=\; f'(x)(y - x)$$

where the inequality follows from $\log z < z - 1$ for $z > 0$ and $z \neq 1$. Hence f is strictly concave by Theorem 2.6.1.

2. To use Theorem 2.6.2 we have

$$f''(x) \;=\; -\frac{1}{x^2} \;<\; 0$$

implying strict concavity of f.

$\square$

Example 2.9 We illustrate the three sufficient conditions of Theorem 2.6 using the convex $f : \mathbb{R}^2 \to \mathbb{R}$ defined by:

$$f(x) \;:=\; f(x_1, x_2) \;:=\; x_1^2 - 4x_1x_2 + 4x_2^2 \;=\; (x_1 - 2x_2)^2$$

For the first-order condition we have

$$\nabla f(x) \;:=\; \nabla f(x_1, x_2) \;=\; 2(x_1 - 2x_2) \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

and hence

$$\begin{aligned} & f(y) - f(x) - \nabla f(x)^T (y - x) \\ = \;& (y_1 - 2y_2)^2 - (x_1 - 2x_2)^2 \;-\; 2(x_1 - 2x_2)((y_1 - x_1) - 2(y_2 - x_2)) \\ = \;& (y_1 - 2y_2)^2 - 2(x_1 - 2x_2)(y_1 - 2y_2) + (x_1 - 2x_2)^2 \\ = \;& \left((y_1 - 2y_2) - (x_1 - 2x_2)\right)^2 \;\geq\; 0 \end{aligned}$$

satisfying the condition of Theorem 2.6.1.

For Theorem 2.6.2 we have

$$\nabla^2 f(x) \;=\; 2 \begin{bmatrix} 1 \\ -2 \end{bmatrix} \begin{bmatrix} 1 & -2 \end{bmatrix}$$

Therefore $\nabla^2 f(x)$ is positive semidefinite as

$$y^T \nabla^2 f(x)\, y \;=\; 2 \left(\begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right)^2 \;\geq\; 0$$

for any $y \in \mathbb{R}^2$.

For Theorem 2.6.3 we have

$$g(t) \quad := \quad f(x+tv) \;=\; ((x_1+tv_1)-2(x_2+tv_2))^2 \;=\; ((v_1-2v_2)\,t+(x_1-2x_2))^2$$

which is clearly a convex function in t for any fixed x and v. □

The addition, multiplication by a positive constant, and supremum operations preserve convexity. Specifically suppose f_1 and f_2 are two convex functions on the same domain. Then

1. $f := \alpha f_1 + \beta f_2$, $\alpha, \beta \geq 0$, is convex.
2. $f := \max\{f_1, f_2\}$ is convex. In fact $f(x) := \sup_{y\in Y} f(x;y)$ is convex in x for arbitrary set Y, provided that, for every $y \in Y$ fixed, $f(x;y)$ is convex in x.
3. $f(x,y) := |x| + |y|$ defined on $\mathbb{R}^2$ is convex as it can be expressed in terms of the supremum and addition operations ($f(x,y) = \max\{x,-x\} + \max\{y,-y\}$).

Convex functions define another important class of convex sets. Let $f : D \to \mathbb{R}$ where $D \subseteq \mathbb{R}^n$. If D is a convex set and f a convex function then for each $\alpha \in \mathbb{R}$ the *level set* $\{x \in D \mid f(x) \leq \alpha\}$ is convex. Let $f : D \to \mathbb{R}^m$ where $D \subseteq \mathbb{R}^n$ be a vector-valued function where $f := (f_1, \ldots, f_m)$ with $f_i : D \to \mathbb{R}$. Then the set specified by:

$$X \quad := \quad \{x \in D \mid f(x) \leq b\} \qquad \text{for some } b \in \mathbb{R}^m$$

is convex if each f_i is convex. This is because the level sets

$$X_i \quad := \quad \{x \in D \mid f_i(x) \leq b_i\}, \quad i = 1, \ldots, m$$

are all convex and $X = \cap_{i=1}^m X_i$ and hence is convex since intersection preserves convexity.

Convex program. Consider an optimization problem of the form:

$$\min_x \quad f(x) \quad \text{subject to} \quad x \in X \tag{2.2}$$

$X \subseteq \mathbb{R}^n$ is called the *feasible set* and $f : \mathbb{R}^n \to \mathbb{R}$ the *objective function*. The problem (2.2) is called a *convex program/problem* if f is a convex function and X is a convex set. For instance

$$X \quad := \quad \{x \in \mathbb{R}^n \mid g(x) \leq b\} \qquad \text{for some } b \in \mathbb{R}^m$$

for a vector-valued convex function $g : \mathbb{R}^n \to \mathbb{R}^m$. An $x \in X$ is called a *feasible solution* of (2.2). A feasible solution x^* that attains the minimum of f over X (i.e., $f(x^*) \leq f(x)$ for all $x \in X$) is called a (global) *optimal solution/optimum* or a (global) *minimizer*. A feasible solution x^* that attains the minimum of f over a neighborhood of x^* (i.e., $f(x^*) \leq f(x)$ for all $x \in B_r(x^*) \cap X$ for some $r > 0$) is called a *local optimal solution/optimum* or a *local minimizer*. By setting $U(x) = -f(x)$, the following maximization problem is called a convex program if $U(x)$ is a concave function and X is a convex set:

$$\max_x \quad U(x) \quad \text{subject to} \quad x \in X$$

An optimal solution may not exist and when it does, it may not be unique.

Theorem 2.10 *Consider the problem* (2.2).

1. *An optimal solution x^* exists if X is nonempty and compact (closed and bounded) and f is continuous.*

2. *The optimal solution x^* is unique if f is strictly convex.*

Note that the existence of an optimal solution x^* requires only that f be continuous, not necessarily convex.

Convexity is important for the efficient computation of an optimal solution. This is because for a convex objective function, local optimality implies global optimality. Moreover only the first-order condition is required to guarantee local optimality. Specifically, for an unconstrained minimization problem

$$\min_{x \in \mathbb{R}^n} \quad f(x)$$

a necessary condition for a point x^* to be a *local* minimizer is (assuming f is differentiable)

$$\nabla f(x^*) \quad = \quad 0$$

If f is convex then this is also sufficient for x^* to be *globally* optimal, as illustrated in Figure 2.2. For constrained minimization problem (2.2) where X is nonempty, closed and convex, the first-order necessary condition for $x^* \in X$ to be a local minimizer becomes: there is a neighborhood $B_r(x^*)$ for some $r > 0$ such that

$$\left(\nabla f(x^*)\right)^T (x - x^*) \quad \geq \quad 0 \qquad \forall x \in B_r(x^*) \cap X \tag{2.3}$$

i.e., moving away from x^* to any other feasible point x in $B_r(x^*)$ can only locally increase the function value f. If f is convex then this is both necessary and sufficient for x^* to be *globally* optimal. To see this, suppose (2.3) holds but there is another $\hat{x} \in X$ such that $f(\hat{x}) < f(x^*)$. Consider $z(\alpha) := \alpha\hat{x} + (1-\alpha)x^*$. Since X is convex $z(\alpha)$ is feasible for $\alpha \in [0, 1]$. Since f is convex we have, for *any* $\alpha \in (0, 1]$,

$$f(z(\alpha)) \quad \leq \quad \alpha f(\hat{x}) + (1-\alpha) f(x^*) \quad < \quad f(x^*)$$

But, for small enough $\alpha > 0$ so that $z(\alpha) \in B_r(x^*)$, this contradicts

$$f(z(\alpha)) \quad \geq \quad f(x^*) + \nabla^T f(x^*)(z(\alpha) - x^*) \quad \geq \quad f(x^*)$$

where the first inequality follows from Theorem 2.6.1 and the second inequality from (2.3). Hence x^* is globally optimal in X.

Example 2.11 Consider

$$\min_{x \in \mathbb{R}} \; f(x) := x^2 \qquad \text{subject to} \quad x \geq a$$

See Figure 2.4. It is clear from the figure that the unique minimizer is 0 where $f'(0) = 0$

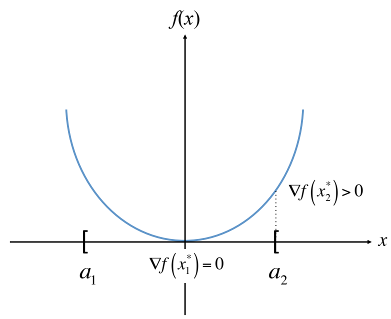


Figure 2.4: Example 2.11: $\min_{x \geq a} x^2$. If $a \leq 0$ then the unique minimizer is $x_1^* = 0$ where $f'(x^*) = 0$. If $a > 0$ then the unique minimizer is $x_2^* = a$ where $f'(x^*) > 0$.

if $a \leq 0$ and a where $f'(a) > 0$ if $a > 0$. We will derive this conclusion from the optimality condition (2.3) which is

$$f'(x^*)(x - x^*) \quad \geq \quad 0, \qquad \forall x \geq a \tag{2.4}$$

First suppose $a \leq 0$. If $a \leq x^* < 0$ then $f'(x^*) < 0$ and there exists a feasible $x > x^*$ where (2.4) cannot be satisfied. Similarly if $x^* > 0 \geq a$ then $f'(x^*) > 0$ and there exists a feasible $a \leq x < x^*$ where (2.4) cannot be satisfied. Hence the unique optimal is $x^* = 0$ where $f'(x^*) = 0$. Suppose next $a > 0$. Then $f'(x) > 0$ for any feasible $x \geq a$. Then the only way (2.4) can be satisfied is if $x^* = a$.

Therefore the optimality condition reduces for this example (for any $a \in \mathbb{R}$) to: x^* is optimal if and only if there exists a p^* such that

$$x^* \geq a, \;\; p^* \geq 0, \;\; f'(x^*) = p^*, \;\; p^*(x^* - a) = 0$$

This is called the Karush-Kuhn-Tucker (KKT) condition for optimality. $\square$

2.1.2 KKT THEOREM AND DUALITY

For our purposes a special case of (2.2) is sufficient where the objective function f is separable in x_i and the feasible set X is specified by a set of linear inequalities (i.e., an affine set). It is also more convenient for the application to TCP/AQM to consider, without loss of generality, maximization instead of minimization.

Specifically consider

$$\max_{x \geq 0} \sum_{i=1}^{N} U_i(x_i) \quad \text{subject to} \quad Rx \leq c \tag{2.5}$$

where $R \in \mathbb{R}^{L \times N}$ is any $L \times N$ matrix, $c \in \mathbb{R}^L$ is any vector, and U_i are concave. The main result Theorem 2.12 in this subsection holds for general linear constraints. After stating the theorem we will then specialize to TCP/AQM applications where entries of R are nonnegative and $c > 0$ is strictly positive. Hence there are L linear (or affine) constraints:

$$\sum_i R_{li} x_i \leq c_l, \qquad l = 1, \ldots, L$$

Even though the objective function is separable in the decision variables x_i, these variables are coupled through the constraint $Rx \leq c$. A direct solution will therefore require the coordination among the x_i's. When (2.5) models TCP congestion control (see below), this means that all TCP sources must coordinate in deciding their sending rates. This is clearly impractical on a large network like the Internet. The simple structure of (2.5) – separable objective linear constraints – however means that a distributed solution can be derived by considering the *Lagrangian dual* (or the *dual*) problem of (2.5).

To derive the dual problem define the Lagrangian of (2.5) to be the following function:

$$L(x,p) \quad := \quad \sum_i U_i(x_i) - p^T(Rx - c) \qquad \text{for } p \geq 0$$

To interpret, note that the Lagrangian $L(x,p)$ contains L terms each of the form:

$$p_l \left(\sum_i R_{li} x_i - c_l \right), \qquad l = 1, \ldots, L$$

one for each constraint $\sum_i R_{li} x_i \leq c_l$. Since $p_l \geq 0$, this term is nonnegative if x violates the constraint. Hence one can interpret $L(x,p)$ as the sum of the primal objective function $\sum_i U_i(x_i)$ and a penalty when any of the constraints in (2.5) is violated. Indeed if we minimize over each $p_l \geq 0$ the penalty term $p_l(\sum_i R_{li} x_i - c_l)$ we will ensure that no constraints can be violated. Hence the problem (2.5), also called the *primal* problem, is equivalent to

$$\max_{x\,:\,x \geq 0, Rx \leq c} \sum_i U_i(x_i) \quad = \quad \max_{x \geq 0} \min_{p \geq 0} \; L(x,p)$$

The dual objective function is defined to be:

$$D(p) \quad := \quad \max_{x \geq 0} \; L(x,p) \quad = \quad \sum_i \max_{x_i \geq 0} \left(U_i(x_i) - x_i \sum_l R_{li} p_l \right) \; + \; \sum_l c_l p_l, \quad p \geq 0$$

where the last equality follows because the separability of the objective function allows us to interchange the order of $\max_{x \geq 0}$ and $\sum_i$. The maximization over x in the definition of the dual objective function $D(p)$ is *unconstrained* (i.e., x may not satisfy $Rx \leq c$ but still satisfied the constraint $x \geq 0$ that has not been dualized) and decoupled, and typically much easier to solve than the original constrained problem (2.5). Note that a maximizer x_i in the definition of $D(p)$ can be unbounded, e.g. when $q_i := \sum_l R_{li} p_l \geq 0$ is smaller than $\lim_{x_i \to \infty} U_i'(x_i)$. In this case we define the maximizer $x_i := \infty$ and $D(p) := \infty$. The (Lagrangian) dual problem is defined to be:

$$\min_{p \geq 0} \qquad D(p) \;\; := \; \sum_i \max_{x_i \geq 0} \left(U_i(x_i) - x_i \sum_l R_{li} p_l \right) \; + \; \sum_l c_l p_l \tag{2.6}$$

Hence the dual problem is equivalent to:

$$\min_{p \geq 0} \; D(p) \quad = \quad \min_{p \geq 0} \max_{x \geq 0} \; L(x,p)$$

It is a convex problem whether or not U_i are concave.

It is easy to show that any primal feasible $x \in X$ and dual feasible $p \geq 0$ satisfy

$$U(x) \quad \leq \quad D(p)$$

i.e., $\max_{x \geq 0} \min_{p \geq 0} L(x,p) \leq \min_{p \geq 0} \max_{x \geq 0} L(x,p)$. Hence the optimal primal value is upper bounded by the optimal dual value:

$$\max_{x \,:\, x \geq 0, Rx \leq c} \; \sum_i U_i(x_i) \quad \leq \quad \min_{p \geq 0} \; D(p)$$

This inequality is called the *weak duality* theorem. It holds whether or not the primal is a convex problem. It also holds when the optimal primal and dual objective values are unbounded: if the primal optimal value is ∞ then the dual problem is feasible; if the dual optimal value is $-\infty$ then the primal problem is infeasible. For general nonlinear optimization the inequality can be strict in which case the gap

$$\min_{p \geq 0} D(p) \quad - \quad \max_{x : x \geq 0, Rx \leq c} \sum_i U_i(x_i)$$

is called the *duality gap*. For our convex program (2.5)–(2.6), *strong duality* holds and the duality gap is zero.[1] The dual problem (2.6) is sometimes called the Lagrange relaxation

[1]If, in addition to the linear constraints, there are nonlinear convex *inequality* constraints in (2.5), then strong duality holds provided Slater's condition, or other constraint qualifications, hold. Slater's condition is not needed for linear inequality constraints [11].

because, in the maximization of $L(x,p)$ over x, the hard constraint $Rx \leq c$ has been relaxed into a penalty term in the Lagrangian function. It is therefore not surprising that the relaxed problem (2.6) yields an upper bound on the original problem (2.5).

The key results on convex optimization that we will use are summarized in the following theorem.

Theorem 2.12 *Consider the convex problem* (2.5) *and its Lagrangian dual problem* (2.6) *and suppose that U_i are continuously differentiable and concave. A vector $x^* \in \mathbb{R}^N_+$ is primal optimal if and only if there exists a $p^* \in \mathbb{R}^L$ such that*

1. primal feasibility: $Rx^* \leq c$, $x^* \geq 0$.

2. dual feasibility: $p^* \geq 0$.

3. stationarity: $U_i'(x_i^*) \leq \sum_l R_{li} p_l^*$, *with equality if* $x_i^* > 0$, *for* $i = 1, \ldots, N$.

4. complementary slackness: $p_l^* \left(\sum_i R_{li} x_i^* - c_l\right) = 0$ *for* $l = 1, \ldots, L$.

In that case p^* is dual optimal, i.e., p^* minimizes (2.6). *The primal optimal solution x^* is unique if U_i are strictly concave. If, in addition, R has full row rank then the dual optimal solution p^* is unique provided $x^* > 0$.*

The complementary slackness condition means that if $p_l^* > 0$ then the constraint must be active $\sum_i R_{li} x^* = c_l$; equivalently if the constraint is inactive $\sum_i R_{li} x^* < c_l$ then $p_l^* = 0$.

2.2 NETWORK UTILITY MAXIMIZATION

We now apply the theory of convex optimization to understand the equilibrium structure of TCP/AQM algorithms. The idea is to identify equilibrium conditions of these dynamical systems with the KKT condition and derive the underlying utility functions that an equilibrium point implicitly optimizes. We first apply Theorem 2.12 to example systems Reno/RED, Vegas/DropTail and FAST/DropTail. We then prove the result for a class of dual algorithms and for general primal-dual algorithms. We discuss implications of network utility maximization on these algorithms.

We first specialize Theorem 2.12 to congestion control applications by imposing additional assumptions on the routing matrix R, the link capacity vector c, and the utility functions U_i that typically hold in practice. We explain these assumptions and their implications.

For congestion control models, the entries of R are assumed nonnegative, $c > 0$ is assumed to be strictly positive. This implies that the feasible set of (2.5) is compact and has nonempty interior, i.e., (2.5) is strictly feasible. Since U_i are continuous and the feasible set is compact, a primal optimal x^* always exists (assuming U_i are defined on the feasible set; see below). Strict feasibility is called the Slator's condition and it implies strong duality. Moreover, since the optimal primal value is finite, the dual problem is feasible and the Slator's condition implies that a dual optimal p^* exists. Hence the existence of primal-dual optimal (x^*, p^*) is always guaranteed for congestion control algorithms.

As we will see below the utility functions U_i for TCP are usually strictly concave increasing and continuously differentiable. For $p \in \mathbb{R}^L_+$ such that

$$q_i(p) \quad \in \quad \left[\lim_{x_i \to \infty} U_i'(x_i),\ \lim_{x_i \to 0} U_i'(x_i)\right]$$

the unique maximizer x_i in (2.6) is given explicitly by $x_i(p) = U_i'^{-1}(q_i(p))$, possibly ∞. If $q_i(p) > \lim_{x_i \to 0} U_i'(x_i) \geq 0$, we define $x_i(p) := 0$ and if $q_i(p) < \lim_{x_i \to \infty} U_i'(x_i)$, we define $x_i(p) := \infty$. In the rest of this book we denote this maximizer (when U_i are strictly concave increasing and continuously differentiable) by

$$x_i \quad := \quad x_i(p) \quad := \quad \left[U_i'^{-1}\left(\sum_l R_{li} p_l\right)\right]^+, \qquad p \in \mathbb{R}^L_+ \tag{2.7}$$

where $[a]^+ := \max\{a, 0\}$ for any $a \in \mathbb{R}$, U_i' are the derivatives of U_i and $U_i'^{-1}$ are their inverses.[2] If U_i is not defined on the entire feasible set, e.g., $U_i(x_i) = \log x_i$ for TCP Vegas and FAST, we will assume that $\lim_{x_i \to 0} U_i'(x_i) = \infty$ in which case $x_i(p) = U_i'^{-1}(q_i(p)) > 0$ without projection. Moreover this implies that, given any $q_i \geq 0$, a unique maximizer $x_i(p) = U_i'^{-1}(q_i(p)) > 0$ exists (without projection) and the dual function $D(p)$ is continuously differentiable.

Example 2.13 If $U_i(x_i) = \log(x_i + a_i)$ for $x_i \geq 0$ with $a_i > 0$ then $U_i'^{-1}(q_i) = \frac{1}{q_i} - a_i$ for $q_i > 0$, which can be negative for $q_i > 1/a_i$. The projection in (2.7) then sets $x_i(p) = 0$. In the context of congestion control, this means that if the end-to-end congestion price q_i is sufficiently large then set the sending rate x_i to zero. □

[2]If U_i is concave (not necessarily strictly so) and continuously differentiable then the inverse of U_i may not exist in which case $U_i'^{-1}$ is defined to be the generalized inverse:

$$U_i'^{-1}(p) \quad := \quad \inf\left\{\ x_i \in \mathbb{R} \mid U_i'(x_i) = \sum_l R_{li} p_l\ \right\}$$

If U_i is concave and differentiable then U_i' is not necessarily continuous in which case the generalized inverse $U_i'^{-1}$ is:

$$U_i'^{-1}(p) \quad := \quad \inf\left\{\ x_i \in \mathbb{R} \mid U_i'(x_i) \leq \sum_l R_{li} p_l\ \right\}$$

In the above it is understood that the infimum of an empty set is defined to be ∞.

2.2.1 EXAMPLE: RENO/RED

From (1.1b) in Chapter 1.3.1 Reno is modeled by

$$\dot{x}_i \quad = \quad \left(\frac{1}{T_i^2} \; - \; \frac{1}{2}\, q_i(t)\, x_i^2(t) \right)^+_{x_i(t)}$$

where $q(t) := R^T p(t)$ and for any $a, b \in \mathbb{R}$, $(a)_b^+ = a$ if $a > 0$ or $b > 0$ and 0 otherwise. Here, as in the rest of this chapter, R denotes a routing matrix. We take a simpler model of RED where the loss probability is linear in the backlog until it saturates at 1:

$$\begin{aligned} \dot{b}_l &= (y_l(t) - c_l)^+_{b_l(t)} \\ p_l(t) &= \min\{\rho b_l(t),\, 1\} \end{aligned}$$

where $y(t) := Rx(t)$ and $c := (c_l, l \in L) > 0$ is a link capacity vector.

The equilibrium of Reno/RED is defined by $\dot{x}_i = 0$ and $\dot{b}_l = 0$ for all i, l. By the definition of $(a)_b^+$, $\dot{x}_i$ can only be zero if the equilibrium rate $x_i^* > 0$ and

$$\frac{1}{T_i^2} \; - \; \frac{1}{2}\, q_i^*\, (x_i^*)^2 \quad = \quad 0$$

since $T_i > 0$. Similarly $\dot{b}_l = 0$ implies $y_l^* \le c_l$ with equality if $b_l^* > 0$. Hence the equilibrium of Reno/RED is characterized by:

$$x_i^* \;=\; \frac{1}{T_i} \frac{\sqrt{2}}{\sqrt{q_i^*}} \quad \text{and} \quad y_l^* \begin{cases} \le c_l & \\ = c_l & \text{if } b_l^* > 0 \end{cases} \qquad \text{for all } i, l \tag{2.8}$$

where $p_l^* = \min\{\rho_l b_l^*, 1\}$. Hence we have

$$\frac{1}{T_i^2} \frac{2}{(x_i^*)^2} \quad = \quad q_i^* \quad = \quad \sum_l R_{li}\, p_l^*$$

Define a function $U_i(x_i)$ such that its derivative equals the left-hand side expression:

$$U_i'(x_i) \quad := \quad \frac{1}{T_i^2} \frac{2}{x_i^2}$$

so that

$$U_i(x_i) \quad := \quad -\frac{1}{T_i^2} \frac{2}{x_i}$$

and $U(x) := \sum_i U_i(x_i)$.

Theorem 2.12 then implies that a point $(x^*, p^*) \in \mathbb{R}^{N+L}$ is an equilibrium of Reno/RED if and only if it is an optimal solution of

$$\max_{x \ge 0} \;\; U(x) \;:=\; -\sum_i \frac{2}{T_i^2} \frac{1}{x_i} \qquad \text{subject to} \qquad Rx \le c \tag{2.9}$$

and its Lagrangian dual. To see this, note that, from Theorem 2.12, (x^*, p^*) is an optimal solution of (2.9) and its dual if and only if

$$x^* \geq 0, \quad Rx^* \leq c, \quad p^* \geq 0, \quad \frac{2}{T_i^2 (x_i^*)^2} = \sum_l R_{li} p_l^* = q_i^*, \quad y_l^* = c_l \text{ if } b_l^* > 0$$

These conditions coincide with the definition of equilibrium for Reno/RED.

Finally we claim that (x^*, p^*) is unique if the routing matrix R has full row rank. Since U_i are strictly concave the optimal x^* is unique. Hence the optimal q_i^* is unique since

$$q_i^* \;=\; \frac{2}{T_i^2 (x_i^*)^2}$$

Since R has full row rank, RR^T is invertible and hence

$$q^* \;=\; R^T p^* \quad\Longrightarrow\quad Rq^* \;=\; RR^T p^* \quad\Longrightarrow\quad p^* \;=\; \left(RR^T\right)^{-1} Rq^*$$

which is unique.

Implications. From (2.8) the steady-state throughput of Reno is inversely proportional to the round-trip time T_i and inversely proportional to the square root of the end-to-end loss probability q_i. This means that a long-distance connection (where the propagation delay is long) has a smaller steady-state throughput than a local connection. It also means that a TCP flow whose path has a high loss probability suffers a low throughput, even if the losses are not due to buffer overflow (congestion) but due to random bit error or interference over a wireless link.

2.2.2 EXAMPLES: VEGAS/DROPTAIL; FAST/DROPTAIL

Vegas is modeled by (writing α_i instead of $\alpha_i d_i$):

$$\dot{w}_i \;=\; \frac{1}{d_i + q_i(t)} \operatorname{sign}\left(\alpha_i \;-\; x_i(t) q_i(t)\right)^+_{w_i(t)}, \quad x_i(t) \;=\; \frac{w_i(t)}{d_i + q_i(t)} \tag{2.10a}$$

$$\dot{p}_l \;=\; \frac{1}{c_l} \left(y_l(t) - c_l\right)^+_{p_l(t)} \tag{2.10b}$$

where $q(t) := R^T p(t)$, $y(t) := Rx(t)$, and $\operatorname{sign}(a) := 1$ if $a > 0$, -1 if $a < 0$ and 0 if $a = 0$. Let $T_i(t) := d_i + q_i(t)$ be the round-trip time of TCP flow i at time t.

Equilibrium (x^*, p^*) (or equivalently (w^*, p^*)) is defined by:

$$\dot{w}_i \;=\; 0 \quad \text{and} \quad \dot{p}_l \;=\; 0 \qquad \text{for all } i, l$$

yielding

$$x_i^* q_i^* \;=\; \alpha_i \quad \text{and} \quad y_l^* \begin{cases} \leq c_l & \\ = c_l & \text{if } p_l^* > 0 \end{cases} \qquad \text{for all } i, l$$

FAST is modeled by:

$$\dot{w}_i \quad = \quad \gamma\left(\alpha_i - x_i(t)q_i(t)\right)^+_{w_i(t)}, \quad x_i(t) \quad = \quad \frac{w_i(t)}{d_i + q_i(t)} \tag{2.11a}$$

$$\dot{p}_l \quad = \quad \frac{1}{c_l}\left(y_l(t) - c_l\right)^+_{p_l(t)} \tag{2.11b}$$

It has the same equilibrium as Vegas.

At equilibrium we have

$$\sum_l R_{li} p_l^* \quad = \quad q_i^* \quad = \quad \frac{\alpha_i}{x_i^*}$$

Hence equating

$$U_i'(x_i) \quad = \quad \sum_l R_{li} p_l \quad = \quad \frac{\alpha_i}{x_i}$$

yields the utility function $U_i(x_i) = \alpha_i \log x_i$. A point $(x^*, p^*) \in \mathbb{R}^{N+L}$ is an equilibrium of Vegas or FAST if and only if it is an optimal solution of

$$\max_{x \geq 0} \sum_i \alpha_i \log x_i \quad \text{subject to} \quad Rx \leq c \tag{2.12}$$

and its Lagrangian dual. This is because Theorem 2.12 states that (x^*, p^*) is an optimal solution of (2.12) if and only if

$$x^* \geq 0, \quad Rx^* \leq c, \quad p^* \geq 0, \quad \frac{\alpha_i}{x_i^*} = \sum_l R_{li} p_l^* =: q_i^*, \quad y_l^* = c_l \text{ if } b_l^* > 0$$

These conditions coincide with the definition of equilibrium for Vegas or FAST. Finally (x^*, p^*) is unique if the routing matrix R has full row rank.

Implications. As mentioned before, FAST can be considered a high-speed version of Vegas in the following sense. Vegas adjusts its window by 1 packet per round-trip time based on the sign of $\alpha_i - x_i(t)q_i(t)$ but regardless of how far $x_i(t)q_i(t)$ is from its target value α_i (compare (2.10a) and (2.11a)). FAST on the other hand adjusts its window based on both the sign and the magnitude of $\alpha_i - x_i(t)q_i(t)$. Hence under FAST $x_i(t)q_i(t)$ converges rapidly to a neighborhood of the target α_i when it is far away and slows down when it is close.

Unlike Reno, Vegas/FAST do not discriminate long-distance flows that have a large propagation delay. As explained in Chapter 1.3.2 the parameter α_i represents the target (steady-state) number of flow i's own packets in buffers along its path, not including packets that are propagating in the links. It determines both the throughput x^* (primal variable) and queueing delay p^* (dual variable): a larger α_i generally leads to a higher throughput x_i^* and larger queueing delay p^*.

2.2.3 EQUILIBRIUM OF DUAL ALGORITHMS

Consider the class of general dual algorithms:

$$\dot{p}_l = \gamma_l \left(y_l(t) - c_l\right)^+_{p_l(t)} := g_l(y_l(t), p_l(t)) \tag{2.13a}$$

$$x_i(t) = \left(U_i'^{-1}(q_i(t))\right)^+ \tag{2.13b}$$

where for any $a \in \mathbb{R}$, $(a)^+ := \max\{a, 0\}$ and $U_i'^{-1}$ are the inverses of the derivatives U_i' of the utility functions U_i. For any $a \in \mathbb{R}$ and $b \geq 0$, $(a)^+_b := a$ if $a > 0$ or $b > 0$ and $(a)^+_b := 0$ if $a \leq 0$ and $b = 0$. As before,

$$q_i(t) = \sum_l R_{li} p_l(t) \quad \text{and} \quad y_l(t) = \sum_i R_{li} x_i(t)$$

The previous examples suggest that the link algorithm (2.13a) can model congestion prices that are proportional to queue length or queueing delay under first-in-first-out queueing discipline.

Dual algorithms have dynamics only in the congestion prices, not in the source rates. Since $x(t)$, and hence $y(t)$, are statically determined by $p(t)$ through (2.13b), we also write $x(p) = x(q)$ and $y(p)$ to mean, componentwise (see (2.7)),

$$x_i(p) := x_i(q_i) := \left(U_i'^{-1}(q_i)\right)^+ \quad \text{and} \quad y_l(p) := \sum_l R_{li}\, x_i(p)$$

where $q_i := \sum_i R_{li} p_l$. We call p^*, as opposed to (x^*, p^*), an *equilibrium* of the dual algorithm (2.13) if

$$g(y(p^*), p^*) = 0$$

where g is defined in (2.13a). Consider the network utilization maximization:

$$\max_{x \geq 0} \quad \sum_i U_i(x_i) \qquad \text{subject to } Rx \leq c \tag{2.14a}$$

and its Lagrangian dual:

$$\min_{p \geq 0} \quad D(p) := \sum_i \max_{x_i \geq 0} (U_i(x_i) - x_i q_i) + \sum_l p_l c_l \tag{2.14b}$$

We make the following assumptions:

C2.1: The link capacities are positive and finite, i.e. $c > 0$. R has no zero column (every flow uses at least one link).

C2.2: U_i are twice continuously differentiable and strictly concave increasing on $\mathbb{R}_+$, with $U_i''(x_i) < 0$ for $x_i \geq 0$.[3] If an U_i is not defined on the entire feasible set, we will assume that $\lim_{x_i \to 0} U_i'(x_i) = \infty$.

As discussed near the beginning of Chapter 2.2, the fact that all entries of R are nonnegative and the condition $c > 0$ in C2.1 imply that (2.14a) is strictly feasible. Hence the primal optimal value is attained by an x^* (U_i are continuous), Slater's condition is satisfied, the dual problem is feasible, and a dual optimal solution p^* is attained. The condition C2.2 on the second derivative of U_i serves two purposes. First it ensures that the primal optimal x^* is unique and when R has full row rank, the dual optimal p^* will also be unique when $x^* > 0$. Second it ensures that the dual dynamic $g(y(p), p)$, though discontinuous in p, is a projection of a locally Lipschitz function to a closed convex set. This allows the application of Theorem 1.10.2 to guarantee the existence and uniqueness of solution to (2.13); see Remark 3.17 in Chapter 3.2. If U_i is not defined over the entire feasible set, e.g., $U_i(x_i) = \log x_i$ for TCP Vegas or FAST, the condition in C2.2 on $U_i'(x_i)$ ensures that, given any $q_i \geq 0$, a unique $x_i(p) = U_i'^{-1}(q_i(p)) > 0$ exists (without projection) and the dual function $D(p)$ is continuously differentiable.

Theorem 2.14 *Suppose conditions C2.1 and C2.2 hold. Then the dual algorithm* (2.13) *has an equilibrium p^* that is an optimal solution of the dual problem* (2.14b). *Moreover the equilibrium source rate $x(p^*)$ is the unique optimal solution of the primal problem* (2.14a). *If the routing matrix R has full row rank and $x^* > 0$ then p^* is unique.*

Proof. The discussion preceding the theorem shows that a dual optimal solution p^* exists. Theorem 2.12 implies that $x^* := x(p^*)$ is the unique primal optimal solution. Moreover they satisfy:

$$\begin{aligned} U'(x^*) \leq q^*, \quad x^* \geq 0, \quad (x^*)^T \left(U'(x^*) - q^*\right) = 0 \\ y^* = Rx^* \leq c, \quad p^* \geq 0, \quad (p^*)^T \left(y^* - c\right) = 0 \end{aligned}$$

This is exactly the condition for $g(y(p^*), p^*) = 0$, i.e., for p^* to be an equilibrium of (2.13).

To prove the uniqueness of p^*, note that $x^* > 0$ implies $q_i^* = U_i'(x_i^*)$ holds for all i. Hence q^* is unique since x^* is unique. Then

$$q^* \;=\; R^T p^* \quad \Longrightarrow \quad Rq^* \;=\; RR^T p^* \quad \Longrightarrow \quad p^* \;=\; \left(RR^T\right)^{-1} Rq^*$$

where $\left(RR^T\right)^{-1}$ exists since R has full row rank. Hence p^* is unique when R has full row rank and $x^* > 0$. This proves the existence and uniqueness of the equilibrium p^* and the primal-dual optimality of p^* and the associated source rates $x^* := x(p^*)$. $\square$

[3]We explicitly require $U_i''(x_i) < 0$ because we will use $U_i'^{-1}(q_i) = \left(U_i''(x_i(q_i))\right)^{-1}$ later. Strict concavity (which is implied by $U_i''(x_i) < 0$) is insufficient because a strictly concave function $f(x)$ can have $f''(x) = 0$, e.g., $f(x) = -x^4$ at $x = 0$.

In the proof above for the uniqueness of p^* we have assumed that q^* that is determined by $U_i'(x^*)$ is in the row space of R ($q^* = R^T p^*$ for some p^*); see Exercise 2.9 when this assumption does not hold. The condition $x^* > 0$ is needed for the uniqueness of p^* because otherwise it is possible to have multiple equilibrium p^* even when R has full row rank; see Exercise 2.11.

2.2.4 EQUILIBRIUM OF PRIMAL-DUAL ALGORITHMS

The examples above suggest that different TCP congestion control algorithms all solve the same prototypical network utility maximization (NUM) problem but may have different utility functions. This is indeed the case as long as the end-to-end congestion measure q_i to which the congestion control algorithm adapts is the *sum* of the link congestion measures p_l.

Consider the class of primal-dual algorithms:

$$\dot{x} = (f(x(t), q(t)))^+_{x(t)} \tag{2.15a}$$

$$\dot{p} = (g(y(t), p(t)))^+_{p(t)} \tag{2.15b}$$

where

$$q(t) = R^T p(t) \quad \text{and} \quad y(t) = R\,x(t) \tag{2.15c}$$

A point $(x, p) \geq 0$ is an equilibrium of (2.15) if and only if $\dot{x} = 0$ and $\dot{p} = 0$. Since

$$(f_i(x_i, q_i))^+_{x_i} := \begin{cases} f_i(x_i, q_i) & \text{if } f_i(x_i, q_i) > 0 \text{ or } x_i > 0 \\ 0 & \text{if } f_i(x_i, q_i) \leq 0 \text{ and } x_i = 0 \end{cases}$$

$\dot{x}_i = 0$ if and only if

$$f_i(x_i, q_i) \leq 0 \qquad \text{with equality if } x_i > 0, \qquad \forall i \tag{2.16a}$$

Similarly $\dot{p}_l = 0$ if and only if

$$g_l(y_l, p_l) \leq 0 \qquad \text{with equality if } p_l > 0, \qquad \forall l \tag{2.16b}$$

Formally we say that a point (x, p) is an *equilibrium* of the basic model (2.15) if and only if $x \geq 0$, $p \geq 0$ and (x, p) satisfies (2.16).

Consider in particular one of the equilibrium conditions $f_i(x_i, q_i) = 0$ when $x_i > 0$ (we will consider the case of $x_i = 0$ later). We will argue that, under appropriate assumptions, the set of (x_i, q_i) that satisfies this condition defines implicitly $q_i =: u_i(x_i)$ as a function of x_i. This will imply that any reasonable TCP design f_i inevitably induces a utility function that source i implicitly optimizes. The function f_i for Reno is (see Chapter 2.2.1):

$$f_i(x_i, q_i) := \frac{1}{T_i^2} - \frac{1}{2}\, q_i\, x_i^2 \tag{2.17a}$$

and that for FAST is from (2.11a):

$$f_i(w_i, q_i) \quad := \quad \gamma \left(\alpha_i - \frac{w_i \, q_i}{d_i + q_i} \right) \tag{2.17b}$$

Hence any equilibrium will have $x_i > 0$ and $f_i(x_i, q_i) = 0$ for Reno and $w_i > 0$ and $f_i(w_i, q_i) = 0$ for FAST. Each of these conditions defines implicitly q_i as a function of x_i or w_i.

In Appendix 2.4 we generalize these examples and prove conditions on the TCP algorithm $f_i(x_i, q_i)$ under which the equilibrium condition ($\dot{x}_i = 0$)

$$f_i(x_i, q_i) \ \leq \ 0 \qquad \text{with equality if } x_i > 0$$

uniquely defines q_i as an implicit function $u_i(x_i)$ of x_i, i.e. given any $x_i \geq 0$, there exists a unique $q_i = u_i(x_i)$ that satisfies the above equilibrium condition. Define the utility function of each source i as

$$U_i(x_i) \quad = \quad \int u_i(x_i) dx_i, \qquad x_i \geq 0 \tag{2.18}$$

that is unique up to a constant. Then U_i is a continuous function assuming u_i does not contain delta functions. Moreover $u_i(x_i) = q_i \geq 0$ for all $x_i \geq 0$ implies that U_i is nondecreasing. Under the conditions in Lemma 2.16, u_i is *strictly* decreasing and hence U_i is strictly concave. An increasing utility function represents a greedy source – a larger rate yields a higher utility – and concavity represents diminishing return.

Now consider the problem of network utility maximization:

$$\max_{x \geq 0} \qquad \sum_i U_i(x_i) \qquad \text{subject to } Rx \leq c \tag{2.19a}$$

and its Lagrangian dual:

$$\min_{p \geq 0} \qquad \sum_i \max_{x_i \geq 0} (U_i(x_i) - x_i q_i) + \sum_l p_l c_l \tag{2.19b}$$

An optimal rate vector x^* exists since the objective function in (2.19a) is continuous and the feasible solution set is compact (c is assumed to be positive and finite). It is unique if U_i are strictly concave. The dual problem has an optimal solution since the primal problem is feasible. The dual variable p is a precise measure of congestion in the network. As the sources are coupled through the shared links (the capacity constraint), solving for x^* directly however may require coordination among possibly all sources and hence is infeasible in a large network. *The key to understanding the equilibrium of* (2.15) *is to regard* $x(t)$ *as primal variables,* $p(t)$ *as dual variables,* $(f, g) = (f_i, g_l, i \in N, l \in L)$ *as a distributed primal-dual algorithm to solve the primal problem (2.19a) and its Lagrangian dual (2.19b), and an equilibrium point as an optimal primal and dual solution.*

We make the following assumptions on (f, g):

C2.3: For all $i \in N$ and $l \in L$, f_i and g_l are functions such that the solutions of (2.16) are nonnegative. The link capacities are positive and finite, i.e. $c > 0$. R has no zero column (every flow uses at least one link).

C2.4: For all $i \in N$, $f_i(x_i, q_i)$ are continuously differentiable on $\mathbb{R}_+ \times \mathbb{R}_+$. Moreover $\frac{\partial f_i}{\partial x_i}(x_i, q_i) < 0$ and $\frac{\partial f_i}{\partial q_i}(x_i, q_i) < 0$ if $x_i > 0$ and $q_i > 0$. For each $x_i > 0$ there exists $\hat{q}_i := \hat{q}_i(x_i) > 0$ such that $f_i(x_i, q_i) \leq 0$ for all $q_i \geq \hat{q}_i$. Finally $f_i(x_i, 0) > 0$.

C2.5: The condition (2.16b) is equivalent to $y_l \leq c_l$ with equality if $p_l > 0$.

Condition C2.3 guarantees that a primal-dual optimal solution (x^*, p^*) of (2.19) exists (see the discussion immediately after condition C2.2). Moreover the optimality (KKT) condition in Theorem 2.12 coincides with the equilibrium condition (2.16), implying that (x^*, p^*) is primal-dual optimal for (2.19) if and only if it is an equilibrium of (2.15). In particular, an equilibrium of (2.15) exists. C2.4 implies the assumptions in Lemma 2.16 in the Appendix 2.4 and hence a unique $u_i(x_i) \geq 0$ exists on $\mathbb{R}_+$ so that $f_i(x_i, u(x_i))$ satisfies the equilibrium condition (2.16a) (which is equivalent to (2.22) in Appendix 2.4). The discussion above then guarantees the existence and strict concavity of utility functions U_i and hence the uniqueness of optimal x^*. C2.5 guarantees primal feasibility and complementary slackness of (x^*, p^*).

We summarize our discussion.

Theorem 2.15 *Suppose assumptions C2.3 – C2.5 hold.*

1. *An equilibrium (x^*, p^*) of (2.15) exists.*
2. *Moreover (x^*, p^*) is an equilibrium of (2.15) if and only if it solves the primal problem (2.19a) and its dual (2.19b) with the utility function given by (2.18).*
3. *The utility functions U_i are strictly concave and hence the optimal rate vector x^* is unique.*
4. *The optimal price vector p^* is unique provided R has full row rank and $x^* > 0$.*

For the last claim, see the proof of Theorem 2.14 (as well as the remark about the uniqueness of p^* after the proof there).

2.3 IMPLICATIONS OF NETWORK UTILITY MAXIMIZATION

2.3.1 TCP/AQM PROTOCOLS

Various TCP/AQM protocols can be interpreted as different distributed algorithms (f, g) to solve the network utility maximization problem (2.19a) and its dual (2.19b) with different utility functions U_i. This computation is carried out by sources and links over the Internet in real time in the form of congestion control. Theorem 2.15 characterizes a large class of protocols (f, g) that admits such an interpretation. This interpretation is the consequence of end-to-end control: it holds as long as the end-to-end congestion measure to which the TCP algorithm reacts is the *sum* of the constituent link congestion measures, under some mild assumptions on the TCP and AQM algorithms that are typically satisfied (assumptions C2.3–C2.5). [4]

The definition of utility functions U_i depends only on TCP algorithms f_i. The role of AQM g_l is to ensure that the complementary slackness condition of problem (2.19) is satisfied (Theorem 2.12.4). The complementary slackness has a simple interpretation: AQM should match input rate to capacity to maximize utilization at every bottleneck link. Any AQM that stabilizes queues possesses this property and generates a Lagrange multiplier p^* that solves the dual problem.

The theorem also provides important insights that help understand, and design, practical protocols that have been deployed on the Internet (e.g., TCP FAST).

First the theorem implies that an arbitrary network under end-to-end control has a unique equilibrium point. The equilibrium point is determined by the utility functions U_i and network parameters such as link capacities c and routing matrix R. It is independent of, e.g., the order in which flows have arrived (for a given set of flow arrivals). The convexity of the underlying optimization problem that (f, g) attempts to solve also leads to a relatively simple dynamic behavior. Even though it may not be possible, nor critical, that optimality is exactly attained in a real network, the utility maximization framework offers a means to steer the network towards an operating point that is easily understandable.

Second, being the Lagrange multipliers, the prices p^* are indeed the right measure of congestion. For implicit congestion control the prices often represent (functions of) packet loss probabilities or queueing delays at the network links. Since the Lagrange multipliers are determined by (U, c, R) this means that the steady-state packet loss probabilities or queueing delays are independent of the buffer size[5]. In particular if p represents loss probabilities at the network links then doubling the buffer sizes will not reduce the steady-state

[4] If the sources react to the *maximum* congestion price in their paths, then the networks achieves maxmin fairness in equilibrium.

[5] If p_l represent queueing delay then their equilibrium values p_l^* are independent of the buffer sizes as long as the buffers are large enough to accommodate the equilibrium queues.

packet losses but will only increase the average queueing delay – buffers will fill up to attain the same loss probabilities before their sizes were doubled.

Third it is useful to treat a practical congestion control scheme simply as an implementation of a certain optimization algorithm. The optimization model then makes possible a systematic method for design and refinement, where modifications to a congestion control mechanism are guided by modifications to the optimization algorithm. For instance, the dual algorithm (2.13) in Chapter 2.2.3 can be interpreted as the first-order gradient projection algorithm to solve the dual problem (2.14b) (see Theorem 5.8 in Chapter 5.2 for more details). It is well known that Newton algorithms usually converge much faster than the first-order gradient projection algorithm, but the computation of Hessian requires global information and is hard to implement in practice. Guided by this insight, however, one can design a practical Newton-like scheme that can attain optimality with a much higher convergence rate without increasing the communication requirement.

2.3.2 UTILITY FUNCTION, THROUGHPUT AND FAIRNESS

A central issue in networking is how to allocate resources to competing users *efficiently* and *fairly* in a decentralized manner. These notions are straightforward when there is a single resource (link) but much subtler for a network of resources. For a single resource of capacity c an allocation policy is efficient if the aggregate equilibrium throughput $\sum_i x_i^* = c$ and it is fair if users are allocated an equal share $x_i^* = c/N$ for all i. On a network, the equilibrium throughputs x_i depend on their routes R and the link capacities c and are generally unequal among users i. The aggregate equilibrium throughput $T(x^*) := \sum_i x_i^*$ that the network can support, however, remains a reasonable measure of efficiency, though there are alternative notions.[6] A natural extension of equal sharing at a single resource to a network setting is maxmin fairness where the smallest rate is maximized. An insight from network utility maximization is that the network bandwidth is allocated among competing users according to their utility functions, given a network specified by (R, c). Hence a general notion of fairness can be defined in terms of utility functions. In this section we explain two counterintuitive behaviors of network utility maximization.

Consider the following class of utility functions parameterized by a scalar $\alpha \geq 0$:

$$U(x_i, \alpha) = \begin{cases} (1-\alpha)^{-1} \, x_i^{1-\alpha} & \text{if } \alpha \neq 1 \\ \log x_i & \text{if } \alpha = 1 \end{cases} \tag{2.20}$$

It includes many resource allocation policies considered in the literature as special cases: e.g. maximum throughput ($\alpha = 0$), proportional fairness ($\alpha = 1$), and maxmin fairness ($\alpha = \infty$). It provides a convenient way to compare fairness across allocation policies. Moreover it also includes several TCP congestion control algorithms as special cases: e.g. Reno ($\alpha = 2$),

[6]For example any Pareto-optimal allocation x^* that lies on the boundary of the feasible set $\{x : Rx \leq c\}$ can be defined to be efficient.

and Vegas, FAST ($\alpha = 1$). A bandwidth allocation policy can hence be defined in terms of this class of utility functions parameterized by α.

Suppose all flows adopt the same utility (2.20) with the same α. We sometimes refer to a network by (R, c, α) where R is the (fixed) routing matrix. We now explain how changing c or α impacts the aggregate throughput $T(x^*)$.

Suppose we add capacities to some links without reducing the capacity of the other links. The new equilibrium rates x^* will be determined by network utility maximization with the new link capacity vector and generally be different. Will the aggregate throughput $T(x^*)$ be always higher? Surprisingly, the answer is "not necessarily" even though the utility function is strictly increasing in x_i. Moreover, given any parameter $\alpha_0 > 0$ there exists a network specified by (R, c) such that for all $\alpha > \alpha_0$, adding any equal amount of capacity to all links will result in a strictly lower throughput $T(x^*)$!

Given two allocation policies identified with α_1 and α_2, we say the first policy is *fairer* than the second policy if $\alpha_1 > \alpha_2$ and *more efficient* if the aggregate throughput $T(\alpha_1) := \sum_i x_i(\alpha_1)$ in equilibrium under policy α_1 is higher than the aggregate throughput $T(\alpha_2) := \sum_i x_i(\alpha_2)$ under policy α_2. It is a folklore that a fairer policy (with a larger α) is always less efficient (with a smaller $T(\alpha)$). Using the network utility maximization model, we can show however that this is not the case for all networks. Indeed it is possible to characterize exactly the set of networks specified by (R, c) for which a fairer allocation is indeed always less efficient, i.e., $T(\alpha)$ is a decreasing function of α. This characterization has led to the discovery of the first counterexamples, i.e., networks (R, c) where a fairer allocation is more efficient.

Hence counterintuitive behaviors can arise in a network where sources interact through shared links in intricate and surprising ways. The popular practice of modeling a network by a single node will fail to capture such subtlety.

2.4 APPENDIX: EXISTENCE OF UTILITY FUNCTIONS

In this appendix we provide conditions on a TCP algorithm f_i that ensure that an underlying utility function U_i exists. The basic idea is that $\dot{x}_i = 0$ if and only if the following equilibrium condition holds:

$$f_i(x_i, q_i) \ \leq \ 0 \qquad \text{with equality if } x_i > 0$$

We will show that this implicitly defines $q_i = u_i(x_i)$ as a function of $x_i \geq 0$. The utility function U_i is then defined to be the integral of u_i. We first derive u_i in Lemma 2.16 for the case where $x_i > 0$ and $f(x_i, q_i) = 0$ and then extends the function u_i to the case where $x_i = 0$ and $f(0, q_i) \leq 0$.

Lemma 2.16 *Consider a continuously differentiable $f_i : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$. Suppose for each $x_i > 0$ in an arbitrary set $X_i \subseteq \mathbb{R}_+$ the following conditions (dependent on x_i) hold:*

1. $\frac{\partial f_i}{\partial q_i}(x_i, q_i) < 0$ *over a compact interval* $Q_i := [0, \overline{q}_i]$ *for some positive finite* $\overline{q}_i$;

2. *there exists* $\hat{q}_i$ *with* $0 < \hat{q}_i \leq \overline{q}_i$ *such that* $f_i(x_i, q_i) \leq 0$ *for all* $q_i \geq \hat{q}_i$.

3. $f_i(x_i, 0) > 0$.

Then given any $x_i \in X_i$ *with* $x_i > 0$, *there is a unique* $q_i =: u_i(x_i)$ *that satisfies* $0 \leq u_i(x_i) \leq \overline{q}_i$ *and* $f_i(x_i, u_i(x_i)) = 0$. *Moreover if for each* $(x_i, q_i) \in X_i \times Q_i$

4. $\frac{\partial f_i}{\partial x_i}(x_i, q_i) < 0$ *as long as* $x_i > 0$ *and* $q_i > 0$.

then $u_i(x_i)$ *is a strictly decreasing function on* $X_i \setminus \{0\}$.

Before proving the lemma we explain how the conditions in the lemma are motivated by practical TCP design f_i. The f_i of Reno and FAST in (2.17) both satisfy the assumptions in the lemma.[7] For example, given $x_i > 0$ for Reno, $\hat{q}_i \geq 2/\left(x_i T_i\right)^2$ from (2.17a) and $\overline{q}_i$ can be any number greater than or equal to $\hat{q}_i$. The first assumption says that, as long as $x_i > 0$, as congestion price q_i along source i's path increases, the rate adjustment f_i will strictly decrease. Moreover, when it is high enough (when $q_i \geq \hat{q}_i$), the sending rate x_i itself will be reduced ($f_i(x_i, q_i) \leq 0$). The first two assumptions thus imply that, as congestion price q_i increases, source i will eventually reduce its rate x_i; moreover the pace of reduction strictly increases as $q_i \geq \hat{q}_i$. The third assumption says that when there is no congestion $q_i = 0$, the rate will be raised ($f_i > 0$). The last assumption says that the higher the sending rate x_i the smaller the adjustment f_i. These assumptions are satisfied by Reno/Vegas/FAST. The conclusion that u_i is strictly decreasing means that the more severe the congestion, the smaller the sending rate. This implies the strict concavity of the utility functions.

Proof of Lemma 2.16. We first prove the existence of the implicit function u_i on $X_i \setminus \{0\}$. Fix any $x_i \in X_i$ with $x_i > 0$. Define for each $q_i \in Q_i$

$$h_i(q_i) \quad := \quad \left(q_i \; + \; \gamma f_i(x_i, q_i)\right)^+$$

where $(a)^+ := \max\{a, 0\}$ for any $a \in \mathbb{R}$ and $\gamma > 0$ will be determined below. We will prove that h_i is a contraction mapping from the closed set Q_i into Q_i. Then the contraction theorem (see Theorem 5.6 in Chapter 5) implies that there is a unique fixed point $q_i =: u_i(x_i)$ of h_i in Q_i, i.e.,

$$u_i(x_i) \quad = \quad \left(u_i(x_i) \; + \; \gamma f_i(x_i, u_i(x_i))\right)^+$$

[7]Note that if, e.g., FAST is designed to be

$$f_i(w_i, q_i) \quad := \quad \gamma\left(\frac{\alpha_i}{w_i} - \frac{q_i}{d_i + q_i}\right)$$

then f_i is not defined at $w_i = 0$ and hence not continuously differentiable on $\mathbb{R}_+ \times \mathbb{R}_+$ as required by the lemma.

Moreover $u_i(x_i) = q_i \geq 0$. If the fixed point $q_i = 0$ then the above equation becomes $0 = (f_i(x_i, 0))^+$, which is impossible under assumption 3 of the lemma. Hence the fixed point $q_i > 0$ and $f_i(x_i, u_i(x_i)) = 0$ as desired.

Hence, for the existence of u_i, we are left to prove that h_i is a mapping from Q_i to Q_i and then h_i is a contraction mapping.

First, if $\hat{q}_i \leq q_i \leq \overline{q}_i$ then $f_i(x_i, q_i) \leq 0$ by assumption 2 of the lemma and hence $h_i(q_i) \leq q_i \leq \overline{q}_i$, i.e., $h_i(q_i) \in Q_i$. Consider now $0 \leq q_i < \hat{q}_i$. Let $a := \max_{q_i' \in [0, \hat{q}_i]} f_i(x_i, q_i')$ which exists and is finite since f_i is continuous. Moreover $h(q_i) \leq (q_i \ + \ \gamma\, a)^+$. If $a \leq 0$ then $0 \leq h_i(q_i) \leq q_i \leq \overline{q}_i$ and hence $h_i(q_i) \in Q_i$. Otherwise $a > 0$ and $0 \leq h_i(q_i) \leq q_i + \gamma\, a \leq \overline{q}_i$ as long as $0 < \gamma \leq \gamma_1$ with

$$\gamma_1 \quad := \quad \frac{\overline{q}_i - q_i}{a} \quad > \quad 0$$

where the strict inequality follows from $q_i < \overline{q}_i$ and a is finite. This proves that h_i is a mapping from Q_i into Q_i.

To show that h_i is a contraction on Q_i, use the mean value theorem to obtain: for any $q_i, \tilde{q}_i$ in Q_i, we have (since $(\cdot)^+$ is non-expansive)

$$|h_i(q_i) - h_i(\tilde{q}_i)| \quad \leq \quad |(q_i - \tilde{q}_i) \, + \, \gamma\, (f_i(x_i, q_i) - f_i(x_i, \tilde{q}_i))| \quad = \quad \left| 1 + \gamma \, \frac{\partial f_i}{\partial q_i}(x_i, z) \right| |q_i - \tilde{q}_i|$$

for some z between q_i and $\tilde{q}_i$. Since f_i is continuously differentiable and Q_i is compact, assumption 1 of the lemma implies that

$$m_i(x_i) \ := \ - \max_{q_i' \in Q_i} \frac{\partial f_i}{\partial q_i}(x_i, q_i') \quad \text{and} \quad M_i(x_i) \ := \ - \min_{q_i' \in Q_i} \frac{\partial f_i}{\partial q_i}(x_i, q_i')$$

satsify $0 < m_i(x_i) < M_i(x_i) < \infty$. Then

$$1 - \gamma\, M_i(x_i) \quad \leq \quad 1 + \gamma\, \tfrac{\partial f_i}{\partial q_i}(x_i, z) \quad \leq \quad 1 - \gamma\, m_i(x_i)$$

Hence

$$\left| 1 + \gamma \, \frac{\partial f_i}{\partial q_i}(x_i, z) \right| \quad \leq \quad \max\{\, |1 - \gamma\, m_i(x_i)| \, , \ |\gamma\, M_i(x_i) - 1| \,\} \quad =: \quad \alpha(x_i) \quad =: \quad \alpha$$

Then $\alpha < 1$ as long as $0 < \gamma < \gamma_2$ where

$$\gamma_2 \quad := \quad \frac{2}{M_i(x_i)}$$

Hence choose $0 < \gamma < \min\{\gamma_1, \gamma_2\}$ and we have

$$|h_i(q_i) - h_i(\tilde{q}_i)| \quad \leq \quad \alpha \, |q_i - \tilde{q}_i|$$

for some $\alpha \in (0, 1)$. We have thus proved that, given any $x_i \in X_i$ with $x_i > 0$, the function h_i is a contraction mapping from Q_i into Q_i. This proves the existence of the implicit function u_i on $X_i \setminus \{0\}$.

Finally we prove that u_i is strictly decreasing on X_i under assumption 4 of the lemma. Consider without loss of generality $0 < x_1 < x_2$ in X_i and the corresponding $q_1 := u_i(x_1)$ and $q_2 := u_i(x_2)$ such that

$$f_i(x_1, q_1) \;=\; 0 \quad \text{and} \quad f_i(x_2, q_2) \;=\; 0$$

Assumption 3 of the lemma then implies that $q_1 > 0, q_2 > 0$. To show that $q_1 > q_2$ we have by the mean value theorem

$$\begin{aligned} 0 &= f_i(x_2, q_2) \;-\; f_i(x_1, q_1) \\ &= (f_i(x_2, q_2) - f_i(x_2, q_1)) \;+\; (f_i(x_2, q_1) - f_i(x_1, q_1)) \\ &= \frac{\partial f_i}{\partial q_i}(x_2, \tilde{q})\,(q_2 - q_1) \;+\; \frac{\partial f_i}{\partial x_i}(\tilde{x}, q_1)\,(x_2 - x_1) \end{aligned}$$

for some $\tilde{x} \in [x_1, x_2]$ and some $\tilde{q}$ between q_2 and q_1. Since both $\tilde{x} > 0$ and $q_1 > 0$, assumption 4 of the lemma implies that the second term is negative. Hence the first term must be positive and this is possible only if $q_1 > q_2$ under assumption 1 of the lemma. Therefore u_i is strictly decreasing on $X_i \setminus \{0\}$. This completes the proof of the lemma. □

Lemma 2.16 says that there is a unique function $u_i(x_i) \geq 0$ on $X_i \setminus \{0\}$ that satisfies one of the equilibrium conditions $f_i(x_i, u_i(x_i)) = 0$ in (2.16a). We now address the case where $x_i = 0$. Extend $u_i(x_i)$ to X_i by defining

$$u_i(0) \;:=\; \lim_{\substack{x_i \to 0 \\ x_i > 0}} u_i(x_i) \;\geq\; 0 \tag{2.21a}$$

which exists (possibly $+\infty$) since u_i is strictly decreasing for $x_i > 0$. Since f_i is continuous we have

$$f_i(0, u_i(0)) \;=\; \lim_{\substack{x_i \to 0 \\ x_i > 0}} f_i(x_i, u_i(x_i)) \;=\; 0 \tag{2.21b}$$

with the interpretation $f_i(0, u_i(0)) = \lim_{\substack{x_i \to 0 \\ x_i > 0}} f_i(0, u_i(x_i))$ if $u_i(0) = \infty$. Under conditions C2.3 – C2.5, u_i is well defined on $\mathbb{R}_+$, and hence $\{(x, u(x)) \,|\, x_i \geq 0\}$ is exactly the set of points that satisfy $f_i(x_i, u_i(x_i)) = 0$. Moreover, since f_i is continuously differentiable, assumption 4 in Lemma 2.16 implies

$$\frac{\partial f_i}{\partial x_i}(0, q_i) \;\leq\; 0$$

This condition and (2.21b) together imply

$$u_i(0) \;\leq\; q_i \quad \Longleftrightarrow \quad f_i(0, q_i) \;\leq\; 0$$

In summary, under the conditions of Lemma 2.16, the function $u_i : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ defined in the lemma and extended by (2.21a) has the desirable property: (x_i, q_i) satisfies (2.16a) if and only if

$$u_i(x_i) \quad \leq \quad q_i \qquad \text{with equality if } x_i > 0 \tag{2.22}$$

For Reno/Vegas/FAST (treating x_i rather than w_i as the variable), it is not possible that $f_i(0, q_i) < 0$ and $u_i(0) < q_i$ in equilibrium because $u_i(x_i) > 0$ for all $x_i \geq 0$. See Exercise 2.10 for an example f_i where this is possible.

2.5 BIBLIOGRAPHICAL NOTES

Excellent texts on convex optimization include [7, 11]. Network utility maximization is first formulated in [28]. Theorem 2.15 interprets Internet congestion control as maximizing utility and is from [34] but the proofs presented in Chapter 2.2.4 are new. In particular Lemma 2.16 justifies the existence of a utility function U_i corresponding to a TCP algorithm f_i by appealing to properties motivated by TCP algorithms such as Reno/Vegas/FAST, without invoking global implicit function theorems that usually require assumptions that are difficult to satisfy. This interpretation holds as long as the sources react to the *sum* of congestion prices in their paths. If they react to the *maximum* price in their paths then the network achieves maxmin fairness, i.e., it maximizes the minimum source rates; see [54].

The dual algorithm (2.13) in Chapter 2.2.3 is from [35] and is the gradient projection algorithm to solve the dual problem (2.14b). Practical Newton-like schemes that can attain optimality with a much higher convergence rate but the same communication requirement are proposed in [5, 51, 52, 55].

The class of utility functions (2.20) is first proposed in [39] for network congestion control. This class has also been used earlier in economics as social welfare functions. The counterintuitive behavior in fairness-efficiency tradeoff and in capacity-throughput relation described in Chapter 2.3.2 is studied in [48]. See [32] for an axiomatic theory of fairness. See also [9] for bounds on these tradeoffs for general systems whose feasible set is not necessarily described by the linear network capacity constraint $Rx \leq c$ in congestion control, and [10] for the effect of fairness on the number of random flow arrivals that can be supported on the network.

2.6 PROBLEMS

Exercise 2.1 (Convex sets)**.** Prove that the following sets are convex:

1. *Affine set:* $C = \{x \in \mathbb{R}^n \mid Ax = b\}$ where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, $m, n \geq 1$.

2. *Second-order cone:* $C = \{(x, t) \in \mathbb{R}^{n+1} \mid \|x\|_2 \leq t\}$, $n \geq 1$. Here $\|x\|_2 := \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ is the Euclidean norm.

3. *Semidefinite matrices:* $C = \{A \in \mathbb{S}^{n\times n} \mid A \succeq 0\}$, $n \geq 1$. where $\mathbb{S}^{n\times n}$ is the set of symmetric $n \times n$ real matrices and $A \succeq 0$ means $x^T Ax \geq 0$ for any $x \in \mathbb{R}^n$. Such a matrix is called *positive semidefinite*.

Exercise 2.2 (Operations preserving set convexity)**.** Operations that preserve convexity are of fundamental importance to the convex optimization theory. Let $\mathbb{X}$ and $\mathbb{Y}$ be linear subspaces. For example $\mathbb{X} := \mathbb{R}^n$ and $\mathbb{Y} := \mathbb{R}^m$.

1. *Linear transformation:* Let $f : \mathbb{X} \to \mathbb{Y}$ be linear. Prove:
   (a) If $A \subseteq \mathbb{X}$ is convex then $f(A) := \{f(x) \mid x \in A\}$ is convex.
   (b) If $B \subseteq \mathbb{Y}$ is convex then $f^{-1}(B) = \{x \in \mathbb{R}^n \mid f(x) \in B\}$ is convex.
2. *Arbitrary direct product:* Let $A \subseteq \mathbb{X}$, $B \subseteq \mathbb{Y}$ be convex.
   (a) Prove that the product space
   $$\mathbb{X} \times \mathbb{Y} := \{(x, y) \mid x \in \mathbb{X},\ y \in \mathbb{Y}\}$$
   with $+$ and $\cdot$ defined by
   $$\begin{aligned} (x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2, y_1 + y_2) && \forall (x_1, y_1),\ (x_2, y_2) \in \mathbb{X} \times \mathbb{Y}; \\ \lambda(x, y) &:= (\lambda x, \lambda y) && \forall \lambda \in \mathbb{R},\ \forall (x, y) \in \mathbb{X} \times \mathbb{Y} \end{aligned}$$
   is also a linear space. For example, if $\mathbb{X} = \mathbb{R}^m$ and $\mathbb{Y} = \mathbb{R}^n$ for some $m, n \geq 1$, then $\mathbb{X} \times \mathbb{Y} = \mathbb{R}^{m+n}$.
   (b) Prove that the direct product
   $$A \times B := \{(x, y) \mid x \in A, y \in B\}$$
   is convex. In fact the direct product of an arbitrary number of convex sets is convex.
3. *Finite sum:* Let $A, B \subseteq \mathbb{X}$ be convex. Prove that the set
   $$A + B := \{a + b \mid a \in A, b \in B\}$$
   is convex. Therefore the sum of any finite number of convex sets is convex.
4. *Arbitrary intersection:* Let $A, B \subseteq \mathbb{X}$ be convex. Prove that the intersection $A \cap B$ is convex. In fact the intersection of an arbitrary collection of convex sets is convex.
5. *Union can be nonconvex.* Let $A, B \subseteq \mathbb{X}$ be convex. Give an example where the union $A \cup B$ is nonconvex. [Hint: Consider $\mathbb{X} = \mathbb{R}$].

Exercise 2.3 (Convex functions)**.** Prove that the following functions are convex:

1. *Exponential:* $f(x) := e^{ax}$ where $a, x \in \mathbb{R}$.
2. *Entropy:* $f(x) := x \ln x$ defined on $\mathbb{R}_{++} := (0, \infty)$.
3. *Log-exponential:* $f(x_1, x_2) := \ln(e^{x_1} + e^{x_2})$, $x_i \in \mathbb{R}$.

Exercise 2.4 (Convex functions)**.** [11, Exercise 3.6]
For each of the following functions determine if it is convex, concave, or neither.

- $f(x) = e^x - 1$ on $\mathbb{R}$.
- $f(x) = x_1 x_2$ on $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0\}$.
- $f(x) = \frac{1}{x_1 x_2}$ on $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0\}$.
- $f(x) = x_1 / x_2$ on $\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0\}$.

Exercise 2.5 (Operations preserving function convexity)**.** Prove that addition, multiplication by nonnegative constants, and supremum operations preserve convexity. Specifically suppose f_1 and f_2 are two convex functions on the same domain. Prove that:

1. $f := \alpha f_1 + \beta f_2$, $\alpha, \beta \geq 0$, is convex.
2. $f := \max\{f_1, f_2\}$ is convex.
3. $f(x, y) := |x| + |y|$ defined on $\mathbb{R}^2$ is convex. [Hint: use result in 2.]

Exercise 2.6 (Level sets are convex)**.** Let $f : C \to \mathbb{R}$ where $C \subseteq \mathbb{R}^n$. Prove that the level set $\{x \in C \mid f(x) \leq \alpha\}$ is convex for any $\alpha \in \mathbb{R}$ provided that C is a convex set and f is a convex function.

Exercise 2.7 (Convex optimization)**.** Consider

$$(\mathbf{P}) : \min_x \quad f(x) \quad \text{s.t.} \qquad Ax = b; \quad g_i(x) \leq 0, \quad i = 1, \ldots, k,$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $k \geq 1$, and $f, g_1, \ldots, g_k$ are scalar functions defined on $\mathbb{R}^n$. Prove that if $f, g_1, g_2, \ldots, g_k$ are convex then the feasible set

$$C = \{x \mid Ax = b, g_i(x) \leq 0 \text{ for } i = 1, \ldots, k\}$$

is convex. In this case (P) is called a convex program. [Hint: The set C is the intersection of $k + 1$ convex sets. Use the results of previous problems.]

Exercise 2.8 (Duality theory)**.** Consider Problem (P) in Exercise 2.7. Let $\mu \in \mathbb{R}^m, \lambda \in \mathbb{R}^k_+ = [0, \infty)^k$, and define

$$L(x, \mu, \lambda) := f(x) + \mu^T(Ax - b) + \lambda^T g(x)$$

where $g(x) = (g_1(x), g_2(x), \dots, g_k(x))^T$.

1. *Unconstrained optimization.* Let $L(\mu, \lambda) := \min_{x \in \mathbb{R}^n} L(x, \mu, \lambda)$ denote the unconstrained optimization over x for fixed μ, λ. Assume that Problem (P) has an optimal solution and denote it by x^*. Show that $L(\mu, \lambda) \le f(x^*)$ for any $\mu \in \mathbb{R}^m$ and $\lambda \in \mathbb{R}^k_+$.

2. *Dual problem.* Consider the dual problem

$$(\mathbf{D}): \quad \max \quad L(\mu, \lambda) \qquad \text{s.t.} \qquad \lambda \ge 0$$

Assume (D) has an optimal solution and denote it by (μ^*, λ^*).

   (a) Show that $L(\mu^*, \lambda^*) - f(x^*) \le \sum_{i=1}^k \lambda_i^* g_i(x^*) \le 0$. It implies that Problem (D) provides a lower bound for Problem (P). Note that this holds whether or not $f, g_1, g_2, \dots, g_k$ are convex.

   (b) Assume now $f, g_1, g_2, \dots, g_k$ are convex. Show that the equality is attained, i.e., $L(\mu^*, \lambda^*) = f(x^*) + \sum_{i=1}^k \lambda_i^* g_i(x^*)$, if and only if

$$\partial_x L(x^*, \mu^*, \lambda^*) = 0$$

   assuming $f, g_1, g_2, \dots, g_k$ are differentiable.

   (c) Show that if there exists (x, μ, λ) such that x is feasible for (P), (μ, λ) is feasible for (D), $\partial_x L(x, \mu, \lambda) = 0$, and $\lambda_i g_i(x) = 0$ for $i = 1, \dots, k$, then x solves (P) and (μ, λ) solves (D). These are the KKT conditions.

Exercise 2.9 (Uniqueness of p^*)**.** We claim that when the routing matrix R has full row rank, then the prices p are unique. Moreover, given q, p is given by

$$R^T p = q \quad \Rightarrow \quad p = (RR^T)^{-1} Rq$$

This implicitly assumes that, given q, there is a solution p that satisfies $R^T p = q$, i.e., q lies in the row space of R.

1. Given any vector q, describe how you would check if a solution p exists.

2. Give an example where the given q does not satisfy your condition for existence of a solution p. What is the expression $(RR^T)^{-1} Rq$? (Hint: Try $R = [1\ 1]$.)

Exercise 2.10 (TCP with finite $u(x)$)**.** Consider a TCP design $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ given by

$$f(x,q) \quad := \quad a + e^{-x} - q$$

Show that f satisfies condition C2.4 and determine the unique function u in Lemma 2.16. Also show that it is possible to have $f(0,q) < 0$ and $u(0) < q$ in equilibrium.

Exercise 2.11 (Nonuniqueness of p^*)**.** Consider a network with two links with capacities c_1, c_2 and three flows with a routing matrix

$$R \quad := \quad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

i.e., flows 1 and 2 traverse links 1 and 2 respectively and flow 3 traverses both links. Let the TCP design of flows 1 and 2 be the function given in Exercise 2.10:

$$f_1(x_1,q_1) \;:=\; a_1 + e^{-x_1} - q_1, \qquad f_2(x_2,q_2) \;:=\; a_2 + e^{-x_2} - q_2$$

and that of flow 3 be:

$$f_3(x_3,q_3) \quad := \quad \gamma\,(\alpha - x_3\,q_3)$$

Show that, even though R has full row rank, the optimal link prices p^* may be nonunique, by an appropriate choice of parameters c_l, a_i, α.

Exercise 2.12 (Throughput vs. fairness)**.** Consider a linear network with L links indexed by $1,\ldots,L$, each of capacity $c = 1$. There are $L+1$ flows indexed by $0,\ldots,L$. Flows $l = 1,\ldots,L$ traverse only link l and the flow indexed by 0 traverses all the L links. Suppose all flows have the following utility function with the same $\alpha \geq 0$:

$$U_i(x_i) = \begin{cases} \frac{x_i^{1-\alpha}}{1-\alpha}, & \alpha \neq 1 \\ \log x_i, & \alpha = 1. \end{cases}$$

The rate at which each flow transmits is determined by the solution of the following utility maximization, subject to capacity constraints:

$$\max_{x \geq 0} \sum_{i=0}^{L} U_i(x_i) \quad \text{s.t.} \quad Rx \leq c$$

where $x = (x_0,\ldots,x_L)$, matrix R is a routing matrix. The expression $x \geq 0$ means $x_i \geq 0$ for $i = 0,\ldots,L$.

Calculate the aggregate throughput $T(\alpha) = \sum_{i=0}^{L} x_i(\alpha)$. Explain the dependence of $T(\alpha)$ on α. Also comment on the dependence of fairness on α.

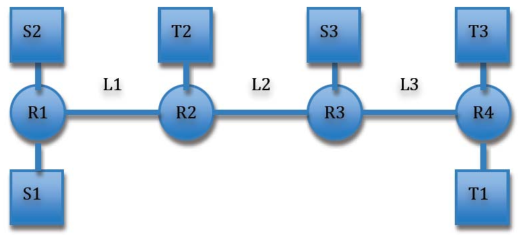


Figure 2.5: Network topology for TCP steady state analysis.

Exercise 2.13 (TCP steady state analysis)**.** Consider the network in Fig. 2.5, where R1–R4 are routers, L1–L3 are links, S1–S3 are source hosts, and T1–T3 are the corresponding destination hosts. The link capacities of L1, L2 and L3 are 2500 packets/s. The one way propagation delay of each link L1 – L3 is 10ms and assume there is no propagation delay between a host and a router. There are three flows: flow 1 from S1 to T1, flow 2 from S2 to T2, and flow 3 from S3 to T3. Flow 1 starts at t=0, flow 2 starts at t=10sec and flow 3 starts at t=20sec. All flows use TCP FAST, i.e., the window update is

$$w(t+\delta t) \quad = \quad \gamma\left(\frac{\mathrm{RTT}_{\min}}{\mathrm{RTT}}w(t)+\alpha\right)+(1-\gamma)w(t)$$

with $\alpha = 50$.

1. Calculate the steady-state throughput of each flow and queue length of each link, during 0s–10s, 10s–20s and after 20s, assuming each flow knows its $\mathrm{RTT}_{\min}$ (round-trip propagation delay) accurately. Assume before flow 2 starts, all packets are buffered at L1.

2. Repeat 1 but with each flow measuring its steady-state $\mathrm{RTT}_{\min}$ that includes queueing delay due to other flows that started before it does. Assume before flow 2 starts, all packets are buffered at L1.

CHAPTER 3

Global stability: Lyapunov method

In Chapter 2 we characterize the equilibrium of a network under end-to-end congestion control and discuss some implications on network performance such as throughput, delay, loss and fairness. This is useful because the first step in designing a congestion control algorithm should be to ensure it has desirable equilibrium properties. In this chapter we study whether the algorithm will indeed drive the network towards an equilibrium starting from an arbitrary initial state. Even though in reality a network is rarely in equilibrium, a stable control ensures that it is always pursuing a desirable state. It also makes the global behavior of the overall network easier to understand.

Numerous congestion control algorithms have been proposed in the literature. We will restrict our analysis to a subset that has proved useful in modeling classical congestion control algorithms. Our purpose is not the study of particular protocols, but use these examples to illustrate the main techniques in proving global stability and explain structural features of congestion control algorithms. We present in this and the following two chapters three different methods to prove global stability. They are methods based on the Lyapunov stability theory (Chapter 3), the passivity theory (Chapter 4), and convergence theorems for gradient algorithms (Chapter 5). In each chapter we first introduce the general method and the associated stability results and then apply them to congestion control algorithms.

For simplicity, the models studied in these chapters ignore feedback delay even though delay is critical in determining stability. In Chapter 6 we will study local stability around an equilibrium in the presence of feedback delay. Global stability in the presence of feedback delay is much more difficult and beyond the scope of this book.

3.1 LYAPUNOV STABILITY THEOREMS

Consider a time-invariant dynamical system:

$$\dot{x} \quad = \quad f(x(t)), \quad t \geq 0, \qquad x(0) = x_0 \tag{3.1}$$

where $f : D \to \mathbb{R}^n$ and $D \subseteq \mathbb{R}^n$ is a domain (open connected set).[1] Here we allow f to be either locally Lipschitz itself or a projection of a locally Lipschitz function $\tilde{f}$ to a closed convex set, specifically,

$$f(x) \quad := \quad \left(\tilde{f}(x)\right)_x^+$$

with a locally Lipschitz $\tilde{f}$. A point $x^* \in D$ is an equilibrium of (3.1) if $f(x^*) = 0$. Assume D contains an equilibrium point x^*.

Definition 3.1 Stability. An equilibrium $x^* \in D$ of (3.1) is:

1. *stable* if $\forall \epsilon > 0$, $\exists \delta = \delta(\epsilon) > 0$ such that

   $$\|x_0 - x^*\| < \delta \;\Rightarrow\; \|x(t) - x^*\| < \epsilon \quad \forall t \geq 0$$

   It is *unstable* if it is not stable.

2. *asymptotically stable* if it is stable and δ can be chosen such that

   $$\|x_0 - x^*\| < \delta \;\Rightarrow\; \lim_{t\to\infty} x(t) = x^*$$

3. *globally asymptotically stable* if $D = \mathbb{R}^n$ or $D = \mathbb{R}^n_+$, x^* is stable and given any initial point $x_0 \in D$, $\lim_{t\to\infty} x(t) = x^*$, i.e., the solution converges to x^*.

These definitions are illustrated in Figure 3.1.

Theorem 3.2 *If x^* is a globally asymptotically stable equilibrium then it is the unique equilibrium.*

In that case we can assume without loss of generality that $x^* = 0 \in D$ by studying the perturbed system $x(t) - x^*$. It is however more convenient for us not to make this assumption because our (unshifted) variables in TCP congestion control often stay nonnegative $x(t) \geq 0$.

A general method to prove the stability of an equilibrium point x^* of (3.1) is to find what is called a *Lyapunov function* $V(x)$ with the property that $V(x)$ is finite, lower bounded, and it decreases along the solution trajectory of (3.1). Specifically let $V : D \to \mathbb{R}$

[1]We require D to be an open set so as to avoid differentiability issues at the boundary. This is only for simplicity of exposition and does not lose generality if, starting from any x_0 on the boundary of D, f always drives $x(t)$ towards the interior.

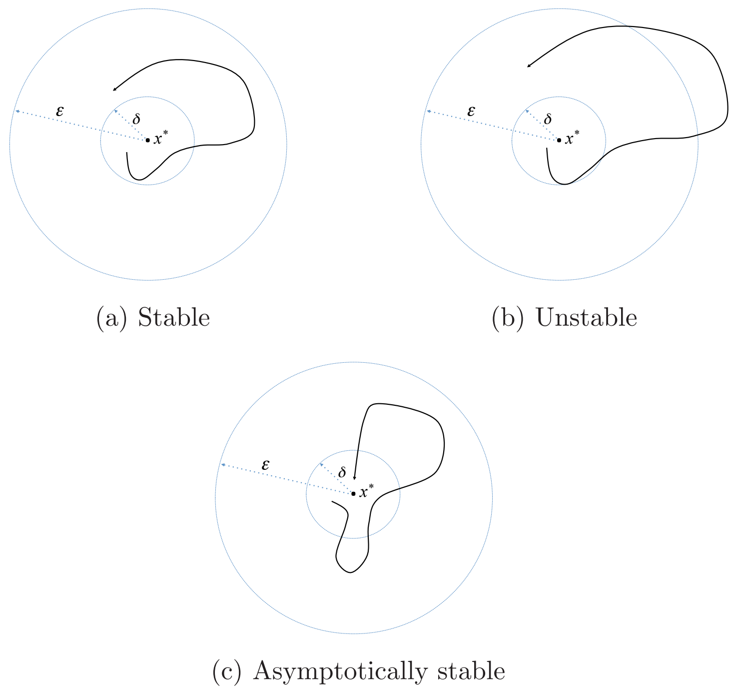


Figure 3.1: Stability of an equilibrium x^*.

be a continuously differentiable function, i.e., the *row* vector $\frac{\partial V}{\partial x}$ of partial derivatives of V exists and is continuous. Define

$$\dot{V}(x) \quad := \quad \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} \dot{x}_i \quad = \quad \frac{\partial V}{\partial x} f(x)$$

We use V both as a function $V(x)$ of $x \in \mathbb{R}^n$ and as a function $V(x(t))$ of $t \in \mathbb{R}_+$, depending on the context. As a function of t, $\dot{V}(x(t))$ is the rate of change in V along the solution trajectory of (3.1). If $f(x) = (\tilde{f}(x))_x^+$ then $f(x)$ is generally discontinuous in x even when $\tilde{f}$ is Lipschitz continuous. Hence even though $\dot{V}(x)$ always exists it is generally discontinuous in x when we allow projected dynamics.

Consider the following conditions:

C3.1: V is positive definite (lower bounded): $V(x^*)$ is finite and $V(x) > V(x^*)$ for all $x \neq x^*$ in D.

C3.2: V has a negative semidefinite rate: $\dot{V}(x) \leq 0$ for all x in D.

C3.2': V has a negative definite rate: $\dot{V}(x) < 0$ for all $x \neq x^*$ in D.

C3.2": There exists $0 < \delta_0 < \epsilon_0$ with $B_{\epsilon_0}(x^*) \subseteq D$ such that for all $0 < \delta \leq \delta_0$ and $\delta < \epsilon \leq \epsilon_0$, there exists $\alpha > 0$ such that

$$\dot{V}(x) \quad \leq \quad -\alpha \quad < \quad 0 \qquad \forall x \text{ with } \delta \leq \|x - x^*\| \leq \epsilon$$

C3.3: V is radially unbounded: $||x|| \to \infty$ implies $V(x) \to \infty$.

We comment on these conditions. Condition C3.1 is equivalent to the following condition often found in the literature:

C3.1': V is positive definite: $V(x^*) = 0$ and $V(x) > 0$ for all $x \neq x^*$ in D.

because, otherwise, we can always define $\tilde{V}(x) := V(x) - V(x^*)$ and V satisfies C3.1 and any of the other conditions if and only if $\tilde{V}$ satisfies C3.1' and the same set of other conditions. As we will see below, condition C3.2' is commonly used to guarantee asymptotic stability when f is locally Lipschitz. C3.2' is inadequate when $f(x) = (\tilde{f}(x))_x^+$ is a projected dynamics and generally discontinuous, and we require C3.2". Indeed C3.2' implies C3.2" when f is continuous since, in that case, $\dot{V}(x) = \frac{\partial V}{\partial x} f(x)$ is continuous in x. If V satisfies C3.3 then the level set $\Omega_c := \{x | V(x) - V(x^*) \leq c\}$ is bounded for all values of $c > 0$. See Exercise 3.2. Otherwise, for example, if $x^* = 0$ and $V(x) = (x_1 - x_2)^2$, then Ω_c is unbounded for any c (This V does not satisfy C3.1 either).

Functions V that satisfy these conditions are called *Lyapunov functions* because of the following key result on the stability of nonlinear dynamical systems.

Theorem 3.3 Lyapunov Stability (locally Lipschitz f). *Suppose f in (3.1) is locally Lipschitz on D. Let $x^* \in D$ be an equilibrium. Let $V : D \to \mathbb{R}$ be a continuously differentiable function.*

1. *If V satisfies C3.1, C3.2 then x^* is stable.*
2. *If V satisfies C3.1, C3.2' then x^* is asymptotically stable.*
3. *If $D = \mathbb{R}^n$ or $D = \mathbb{R}^n_+$ and V satisfies C3.1, C3.2', C3.3 then x^* is globally asymptotically stable.*

Proof.
Part 1. To prove stability, fix any $\epsilon > 0$. We will derive an $\delta > 0$ such that $x(0) \in B_\delta(x^*)$ implies $x(t) \in B_\epsilon(x^*)$ for all $t \geq 0$. We can assume without loss of generality that the set

$$B_\epsilon(x^*) \quad := \quad \{x \in \mathbb{R}^n \mid \|x - x^*\| \leq \epsilon\}$$

lies in D. Otherwise we can pick any $r \in (0, \epsilon]$ such that the set

$$B_r(x^*) \quad := \quad \{x \in \mathbb{R}^n \mid \|x - x^*\| \leq r\}$$

lies in D and replace B_ϵ with B_r in the following. The argument proceeds in three steps and is easy to visualize in the simple case (e.g. $V(x) = x_1^2 + x_2^2$ in $\mathbb{R}^2$) illustrated in Figure 3.2 and its caption.

Let v be the minimum $V(x)$ on the boundary of the closed ball $B_\epsilon(x^*)$ relative to $V(x^*)$:

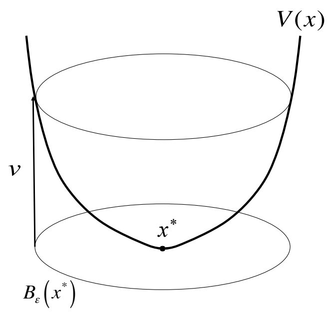
$$v \quad := \quad \min_{x:\|x-x^*\|=\epsilon} V(x) - V(x^*)$$

Condition C3.1 means that $v > 0$ and hence the set[2]

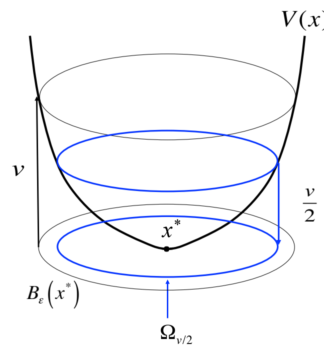
$$\Omega_{v/2} \quad := \quad \left\{ x \in B_\epsilon(x^*) \;\middle|\; V(x) - V(x^*) \leq \frac{v}{2} \right\}$$

is nonempty (since V is continuous). It is in the interior of $B_\epsilon(x^*)$ since $V(x) - V(x^*)$ equals $v/2$ on the boundary of $\Omega_{v/2}$ but equals v on the boundary of $B_\epsilon(x^*)$. Moreover $\Omega_{v/2}$ must

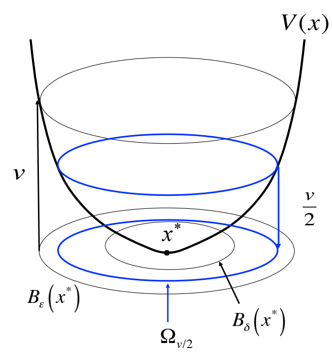
[2] The level set $\Omega_{v/2}$ is defined to be a subset of $B_\epsilon(x^*)$ and hence bounded. There can be x outside $B_\epsilon(x^*)$ where $V(x)$ drops below $v/2$ but these x are not in the set $\Omega_{v/2}$. Note that the choice of v to be the minimum value of V on the bounded set $\|x - x^*\| = \epsilon$ is important. For example if $V(x) := x_1^2/(1 + x_1^2) + x_2^2$ and $x^* = (0, 0)$ then $\{\, x \in \mathbb{R}^2 \mid V(x) - V(x^*) \leq c \,\}$ is unbounded if $c > 1$ because $(x_1, 0)$ is in the set for arbitrarily large $|x_1|$. However, the definition of v ensures that $v < 1$ and hence $\Omega_{v/2}$ is strictly contained in $B_\epsilon(x^*)$. See Remark 3.7 and Figure 3.5.



(a) $B_\epsilon(x^*)$ induces $v > 0$



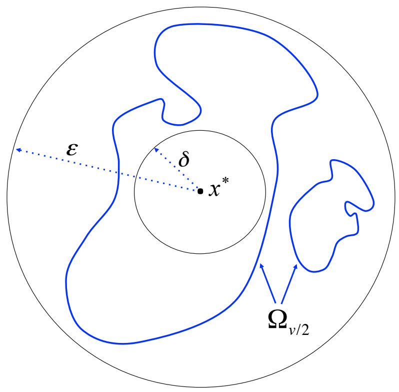
(b) $v > 0$ induces $\Omega_{v/2}$



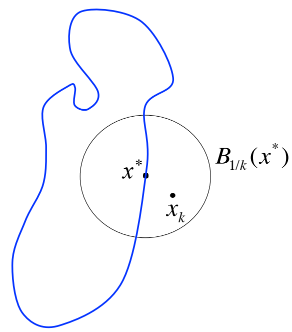
(c) $\Omega_{v/2}$ induces $B_\delta(x^*)$

Figure 3.2: Proof of part 1. Assume without loss of generality $V(x^*) = 0$. (a) Since V is continuous, $v := \min_{x:\|x-x^*\|=\epsilon} V(x)$ is finite and positive under condition C3.1. (b) Hence the $\Omega_{v/2} := \{x \in B_\epsilon(x^*) | V(x) \le v/2\}$ lies in the interior of $B_\epsilon(x^*)$. (c) $\Omega_{v/2}$ must contain a ball $B_\delta(x^*)$ such that $x(0) \in B_\delta(x^*)$ implies $x(t) \in \Omega_{v/2} \subset B_\epsilon(x^*)$.

contain a ball $B_\delta(x^*)$ for some $\delta > 0$ (see Figure 3.3(b)) since, otherwise, it means that for $k = 1, 2, \ldots,$ there exists an $x_k \in B_{1/k}(x^*) \cap B_\epsilon(x^*)$ with $V(x_k) - V(x^*) > v/2$. Since $x_k \to x^*$ and V is continuous, $V(x_k) \to V(x^*)$, contradicting $V(x_k) - V(x^*) > v/2$ for all k. These sets are illustrated in Figure 3.3(a).



(a) Level surface, ϵ- and δ-balls



(b) $B_\delta(x^*)$ must lie inside $\Omega_{v/2}$

Figure 3.3: $B_\delta(x^*) \subseteq \Omega_{v/2} \subset B_\epsilon(x^*)$.

We now show that

$$\|x(0) - x^*\| < \delta \quad \Rightarrow \quad \|x(t) - x^*\| < \epsilon \quad \text{for all } t \geq 0$$

Suppose $x(0) \in B_\delta(x^*)$. Then $x(0) \in \Omega_{v/2}$ and, since $\dot{V} \leq 0$ by condition C3.2,

$$V(x(t)) - V(x^*) \quad \leq \quad V(x(0)) - V(x^*) \quad \leq \quad \frac{v}{2}, \qquad t \geq 0$$

Hence $x(t) \in \Omega_{v/2} \subset B_\epsilon(x^*)$ for all $t \geq 0$ as desired, since $\Omega_{v/2}$ is in the interior of $B_\epsilon(x^*)$.[3]

Part 2. To prove asymptotic stability of x^* we have to prove that, under C3.1 and C3.2', $\delta > 0$ can be chosen such that $x(0) \in B_\delta(x^*)$ implies, in addition, $\lim_{t\to\infty} x(t) = x^*$. Fix any $\epsilon_0 > 0$. Construct an $\delta_0 > 0$ as in part 1 such that $x(0) \in B_{\delta_0}(x^*)$ implies $x(t) \in B_{\epsilon_0}(x^*)$ for all $t \geq 0$. This is convenient as it allows us to restrict attention to the compact set $B_{\epsilon_0}(x^*)$. We will prove part 2 in three steps.
Step 1: We show that $V(x(t)) \to V(x^*)$ as $t \to \infty$. Condition C3.2' implies that $V(x(t))$ monotonically decreases and hence $V(x(t)) \to v^*$ for some v^*. Suppose for the sake of

[3]Since $B_\epsilon(x^*)$ is compact, Theorem 1.6 or Theorem 1.10 of Chapter 1.5.1 implies that, starting from an initial state $x(0) \in B_\delta(x^*)$, a unique solution $(x(t), t \geq 0)$ of (3.1) indeed exists.

contradiction that $v^* > V(x^*)$ and hence $V(x(t)) \geq v^* > V(x^*)$ for all $t \geq 0$. Then there is a $\eta > 0$ such that $\|x(t) - x^*\| \geq \eta$ for all $t \geq 0$, for otherwise, for every natural number k there is a $t_k \geq 0$ such that $x(t_k) \in B_{1/k}(x^*)$. This sequence $x(t_k) \to x^*$ as $k \to \infty$ and since V is continuous $V(x(t_k)) \to V(x^*)$, contradicting $V(x(t)) \geq v^* > V(x^*)$ for all $t \geq 0$.

Then consider the slowest rate of decrease in $V(x)$ outside $B_\eta(x^*)$:

$$\alpha \quad := \quad \inf_x \; -\dot{V}(x) := -\frac{\partial V}{\partial x} f(x) \qquad \text{s. t.} \quad \eta \leq \|x - x^*\| \leq \epsilon_0 \tag{3.2}$$

Since V is continuously differentiable and f is locally Lipschitz, $\dot{V}(x)$ is continuous in x and hence α is attained in the compact feasible set $\{x \,|\, \eta \leq \|x - x^*\| \leq \epsilon_0\}$; moreover $\alpha > 0$ by condition C3.2'.[4]

Hence, for $t \geq 0$, since $x(t) \in B_{\epsilon_0}(x^*)$ we have

$$V(x(t)) \quad = \quad V(x(0)) + \int_0^t \dot{V}(x(\tau))d\tau \quad \leq \quad V(x(0)) - \alpha\, t \quad \longrightarrow \quad -\infty \qquad \text{as} \quad t \to \infty$$

contradicting $v^* > V(x^*) > -\infty$ (condition C3.1). Hence $V(x(t)) \to V(x^*)$ as $t \to \infty$.
Step 2: We show that $x(t)$ crosses *Lyapunov surfaces*, or *level sets*, with decreasing levels. Specifically, following the construction in part 1, construct the neighborhoods $B_{\epsilon_i}(x^*)$ and level sets $\Omega_{v_i/2}$ as follows. Start with any $\epsilon_0 > 0$ such that $B_{\epsilon_0}(x^*)$ lies in D. Recall the minimum value v_0 of $V(x) - V(x^*)$ on the boundary of $B_{\epsilon_0}(x^*)$:

$$v_0 \quad := \quad \min_{x:\|x-x^*\|=\epsilon_0} \; V(x) - V(x^*)$$

and the level set

$$\Omega_{v_0/2} \quad := \quad \left\{ \; x \in B_{\epsilon_0}(x^*) \;\middle|\; V(x) - V(x^*) \leq \frac{v_0}{2} \; \right\}$$

It is important that $\Omega_{v_0/2}$ is nonempty and in the interior of $B_{\epsilon_0}(x^*)$. As we argued in part 1 there is a nonempty neighborhood $B_{\epsilon_1}(x^*)$ in $\Omega_{v_0/2}$. Define

$$v_1 \quad := \quad \min_{x:\|x-x^*\|=\epsilon_1} \; V(x) - V(x^*)$$

on the boundary of $B_{\epsilon_1}(x^*)$ and the level set

$$\Omega_{v_1/2} \quad := \quad \left\{ \; x \in B_{\epsilon_1}(x^*) \;\middle|\; V(x) - V(x^*) \leq \frac{v_1}{2} \; \right\}$$

which is nonempty and in the interior of $B_{\epsilon_1}(x^*)$. As before $\Omega_{v_1/2}$ contains a nonempty neighborhood $B_{\epsilon_2}(x^*)$, and so on. This collection of sets are nested as shown in Figure 3.4:

$$B_{\epsilon_0}(v^*) \;\supset\; \Omega_{v_0/2} \;\supseteq\; B_{\epsilon_1}(x^*) \;\supset\; \Omega_{v_1/2} \;\supseteq\; \cdots \;\ni\; x^*$$

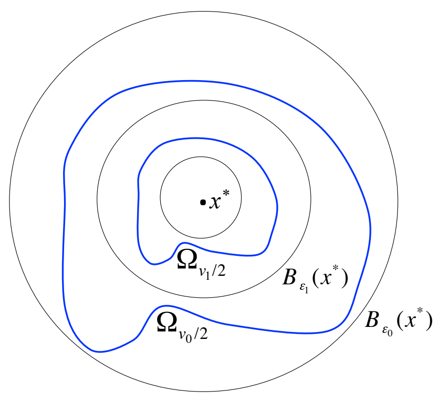


Figure 3.4: Construction of the nested sets $(B_{\epsilon_i}(x^*), \Omega_{v_i/2})$.

We now argue that since $V(x(t)) \to V(x^*)$ (step 1), the trajectory $x(t)$ enters levels sets $\Omega_{v_0/2}, \Omega_{v_1/2}, \cdots$ and therefore enters the neighborhoods $B_{\epsilon_0}(v^*), B_{\epsilon_1}(v^*), \cdots$, i.e., $x(t) \to x^*$. Specifically the nested set of pairs $(B_{\epsilon_i}(x^*), \Omega_{v_i/2})$ has the properties

- If $x(0) \in B_\delta(x^*)$ as in part 1 then $x(t) \in B_{\epsilon_0}(x^*)$ for all $t \geq 0$.
- $\Omega_{v_i/2}$ is nonempty and in the interior of $B_{\epsilon_i}(x^*)$, i.e., $\phi \neq \Omega_{v_i/2} \subset B_{\epsilon_i}(x^*)$.
- The levels $v_0 > v_1 > v_2 > \cdots$ are strictly decreasing with
$$v_{i+1} \quad \leq \quad \frac{v_i}{2}$$
- The sizes $\epsilon_0 > \epsilon_1 > \epsilon_2 > \cdots$ converge to zero, $\epsilon_i \to 0$ as $i \to \infty$ (because $v_i \to 0$ and $V(x) > V(x^*)$ for all $x \neq x^*$).

Step 1 shows that $x(t)$ will cross the level sets $\Omega_{v_i/2}$, $i = 0, 1, \ldots$.
Step 3: We show that $x(t) \to x^*$ as $t \to \infty$, i.e., given any $\epsilon > 0$ there exists T such that $x(t) \in B_\epsilon(x^*)$ for all $t > T$. Fix an $\epsilon > 0$. Let $\epsilon_i \leq \epsilon$ and consider $\Omega_{v_i/2} \subset B_{\epsilon_i}(x^*)$. Since $x(t)$ will eventually enter $\Omega_{v_i/2}$, there exists T such that, for all $t \geq T$, $x(t) \in \Omega_{v_i/2} \subset B_{\epsilon_i}(x^*) \subseteq B_\epsilon(x^*)$, as desired.

Part 3. To prove global asymptotic stability we must show that, given *any* $x_0 \in D = \mathbb{R}^n$ or $\mathbb{R}^n_+$, $\lim_{t\to\infty} x(t) = x^*$. Fix an $x_0 \in D$ and consider

$$\Omega_{v_0} \quad := \quad \{x \in D \mid V(x) - V(x^*) \leq V(x_0)\}$$

The same argument in Step 2 above applies if we can show that Ω_{v_0} is compact, or $\Omega_{v_0} \subseteq B_R(x^*)$ for some finite $R > 0$. If no such R exists then, for each $k = 1, 2, \ldots$, there exists an $x_k \in \Omega_{v_0}$ such that $V(x_k) \leq V(x^*) + V(x_0) < \infty$ and $\|x_k - x^*\| > k$. This contradicts condition C3.3. (Also see Exercise 3.2.) □

Example 3.4 Bounded level set. Suppose $V : D \to \mathbb{R}$ is continuously differentiable and satisfies C3.1. Prove that there is a level set Ω_c that contains x^*, is bounded, and $V(x) - V(x^*) \leq c$ for all $x \in \Omega_c$, provided c is small enough.

This essentially extracts the first part of the proof of Theorem 3.3. Start with an $\epsilon > 0$ small enough such that $B_\epsilon(x^*)$ lies in D. Define

$$v \quad := \quad \min_{x:\|x-x^*\|=\epsilon} V(x) - V(x^*)$$

[4]This is the only place where the continuity of $\dot{V}(x)$ in x is used. When f_i are projected dynamics, $\dot{V}(x)$ may be discontinuous and we require condition C3.2" to ensure $\alpha < 0$ is attained.

Condition C3.1 means that $v > 0$ and hence the set with $c := v/2$:

$$\Omega_{v/2} \quad := \quad \left\{ x \in B_\epsilon(x^*) \;\middle|\; V(x) - V(x^*) \le \frac{v}{2} \right\}$$

is nonempty (since V is continuous). It is in the interior of $B_\epsilon(x^*)$ since $V(x) - V(x^*)$ equals $v/2$ on the boundary of $\Omega_{v/2}$ but equals v on the boundary of $B_\epsilon(x^*)$.

Note that the choice of v to be the minimum value of V on the bounded set $\|x - x^*\| = \epsilon$ is important. Consider

$$V(x) \quad := \quad \frac{x_1^2}{1 + x_1^2} \; + \; x_2^2$$

$x^* = (0, 0)$. Then $\{ x \in \mathbb{R}^2 \,|\, V(x) - V(x^*) \le c \}$ is unbounded if $c > 1$ because $(x_1, 0)$ is in the set for arbitrarily large $|x_1|$. However, the definition of v ensures that $v < 1$ and hence $\Omega_{v/2}$ is strictly contained in $B_\epsilon(x^*)$. See Remark 3.7 and Figure 3.5. See also Exercise 3.2. □

When (3.1) is a projected dynamics so that $f(x)$ is discontinuous in x, condition C3.2' needs to be strengthened to C3.2".

Corollary 3.5 Lyapunov Stability (projected dynamics f). *Suppose $f(x) = \left(\tilde{f}(x)\right)_x^+$ in (3.1) where $\tilde{f}$ is locally Lipschitz on D. Let $x^* \in D$ be an equilibrium. Let $V : D \to \mathbb{R}$ be a continuously differentiable function.*

1. *If V satisfies C3.1, C3.2 then x^* is stable.*
2. *If V satisfies C3.1, C3.2" then x^* is asymptotically stable.*
3. *If $D = \mathbb{R}^n$ or $D = \mathbb{R}^n_+$ and V satisfies C3.1, C3.2", C3.3 then x^* is globally asymptotically stable.*

Proof. As mentioned before, $\dot{V}(x) := \frac{\partial V}{\partial x} f(x)$ always exists but is in general not continuous when f is a projected dynamics. The continuity of $\dot{V}(x)$ is used only in the proof of part 2 to show that α in (3.2) is strictly positive. With condition C3.2" there always exist a $\delta < \min\{\eta, \delta_0\}$, an ϵ with $\delta < \epsilon < \epsilon_0$, and an $\alpha > 0$ such that $\dot{V}(x) \le -\alpha < 0$ for all x with $\delta \le \|x - x^*\| \le \epsilon$. Hence $V(x(t)) \le V(x(0)) - \alpha t$ as before. The rest of the argument is identical to the proof for Theorem 3.3. □

Remark 3.6 Existence and uniqueness of solution. Since f is assumed to be only locally Lipschitz (or a projection of a locally Lipschitz function), a solution of (3.1) is not a priori guaranteed. The proof of Theorem 3.3.3 (or Corollary 3.5.3) for global stability shows that, if there is a Lyapunov function V that satisfies C3.1, C3.2 (it is not necessary to satisfy C3.2' or C3.2") and C3.3 then, starting from any initial point x_0, the solution trajectory $x(t)$, if exists, will stay entirely in a ball B_R that contains the level set $\{x \mid V(x) \leq V(x_0)\}$. Theorem 1.6 or Theorem 1.10 of Chapter 1.5.1 then implies that a unique solution $(x(t), t \geq 0)$ of (3.1) indeed exists.

Remark 3.7 Radially unbounded. We comment on which part of the proof for Theorem 3.3.2 breaks down if V is not radially unbounded and $\Omega_{v_0/2}$ is not bounded (that for Corollary 3.5.2 is identical). Recall that the proof of part 2 first shows that $V(x(t)) \to V(x^*)$ as $t \to \infty$ and then shows that, since $V(x(t))$ crosses Lyapunov surfaces to enter $\Omega_{v_i/2}$ and $\Omega_{v_i/2} \subset B_{\epsilon_i}$, we must have $x(t) \to x^*$. Both steps require that $\Omega_{v_0/2}$ be contained in B_R for some $R < \infty$. To see this consider a concrete Lyapunov candidate $V : \mathbb{R}^2 \to \mathbb{R}$ given by

$$V(x) \quad := \quad x_2^2$$

and assume the origin is the equilibrium point under study. V is not radially unbounded because $(x_1, 0) \to \infty$ but $V(x_1, 0) = 0$ as $x_1 \to \infty$. But V does not satisfy condition C3.1 since $V(x_1, 0) = V(0, 0)$ (the origin is the equilibrium point by assumption). Hence V is not a valid Lyapunov candidate for us. Consider instead

$$V(x) \quad := \quad \frac{x_1^2}{1 + x_1^2} \; + \; x_2^2$$

V is not radially unbounded because $(x_1, 0) \to \infty$ but $V(x_1, 0) \leq 1$ as $x_1 \to \infty$. The set

$$\{x \in \mathbb{R}^2 \mid V(x) = c\}$$

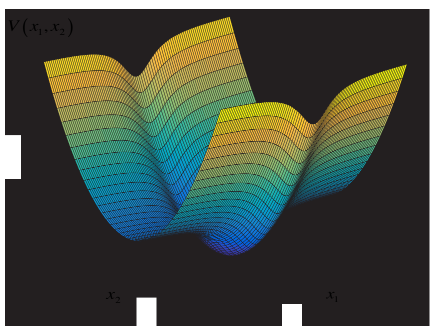
is unbounded for $c \geq 1$. Suppose the system starts at some $x(0) = (x_1(0), x_2(0))$ with $x_2(0) \geq \sqrt{2}$ so that $V(x(0))/2 \geq 1$ and the set

$$\left\{ x \in \mathbb{R}^2 \;\mid\; V(x) \leq \frac{V(x(0))}{2} \right\}$$

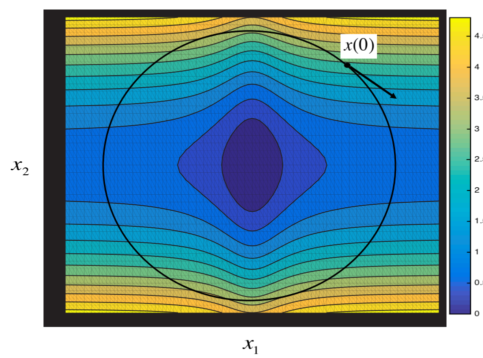
is unbounded. In an attempt to get a bounded set, suppose we take the intersection of the above set with the ball defined by $x(0)$:

$$\Omega_{v_0/2} \quad := \quad \left\{ x \in \mathbb{R}^2 \;\mid\; V(x) \leq \frac{V(x(0))}{2} \right\} \; \cap \; B_{\|x(0)\|}(0)$$

As shown in Figure 3.5, $x(0)$ is at the boundary of $\Omega_{v_0/2}$. Since we only know that $\dot{V} < 0$, starting from $x(0)$ on the boundary, $x(t)$ may leave $\Omega_{v_0/2}$ instead of entering it and therefore the trajectory of $x(t)$ may not stay in a bounded set even as $V(t)$ is strictly decreasing in time; see Figure 3.5(b). This means that the minimum rate α of decrease in (3.2) may be



(a) 3D plot



(b) Contour plot

Figure 3.5: Plots of $V(x) = x_1^2/(1 + x_1^2) + x_2^2$. Starting from an $x(0)$ such that $V(x(0)) > 1$, the trajectory $x(t)$ may not stay in any bounded set even if $\dot{V} < 0$.

0 over the region that the trajectory $x(t)$ may visit. Hence the proof for $V(x(t)) \to V(x^*)$ becomes invalid. Moreover since $\Omega_{v_i/2}$ may not be contained in any ball B_{ϵ_i}, even if $x(t)$ crosses Lyapunov surfaces, $x(t)$ is not guaranteed to stay in any small ball B_{ϵ_i} for sufficiently large t. Hence the argument that $x(t) \to x^*$ becomes invalid.

Example 3.8 Consider

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_2^2 \\ \dot{x}_2 &= x_1 x_2 - 2x_2^3 \end{aligned}$$

with the origin as an equilibrium. Show that the origin is globally asymptotically stable.

Consider the Lyapunov function candidate

$$V(x) \; := \; \frac{1}{2}\left(x_1^2 + x_2^2\right)$$

Clearly V is continuously differentiable, positive definite (condition C3.1), i.e., $V(x) > 0$ for all $x \neq 0$, and radially unbounded (C3.3). For condition C3.2' we have

$$\begin{aligned} \dot{V} &= x_1\,\dot{x}_1 + x_2\,\dot{x}_2 \\ &= -x_1^2 - 2\,x_2^4 + 2\,x_1 x_2^2 \;=\; -\left(x_1 - x_2^2\right)^2 - x_2^4 \;\leq\; 0 \end{aligned}$$

If $\dot{V} = 0$ then $x_2 = 0$ and $x_1 = x_2^2 = 0$. Hence $\dot{V} < 0$ for all $x \neq 0$ and satisfies C3.2'. Theorem 3.3 then implies that $x^* = 0$ is globally asymptotically stable. □

Finding a Lyapunov function is generally difficult. Often a given Lyapunov function has a rate that is negative semidefinite, but not negative definite, i.e., condition C3.2 is satisfied but not C3.2' or C3.2". It still guarantees global asymptotic stability if it satisfies an additional property. By "$x(t) \equiv a$" or "$x \equiv a$" we mean "$x(t) = a$ for all $t \geq 0$."

Theorem 3.9 LaSalle's Invariance Principle. *Let $x^* \in D$ be an equilibrium of (3.1) where f is either locally Lipschitz or a projection of a locally Lipschitz function. Let $V : D \to \mathbb{R}$ be a continuously differentiable function. Let $E := \{x \in D \,|\, \dot{V}(x) = 0\}$ and suppose no solution of (3.1) can stay identically in E other than the trivial solution $x(t) \equiv x^*$.*

1. *If V satisfies conditions C3.1, C3.2 then x^* is asymptotically stable.*
2. *If $D = \mathbb{R}^n$ or $D = \mathbb{R}^n_+$ and V satisfies C3.1, C3.2, C3.3 then x^* is globally asymptotically stable.*

Note that we allow f in (3.1) to be discontinuous when it is a projected dynamics. When condition C3.2 is replaced by C3.2' or C3.2", then $E = \{x^*\}$ and the theorem reduces to Theorem 3.3 and its corollary. Otherwise E generally contains other points $a \neq x^*$ where $\dot{V}(a) = 0$. The condition in Theorem 3.9 requires that no solution trajectory $(x(t), t \geq 0)$ of (3.1) can stay entirely in E except $x(t) \equiv x^*$. For instance E cannot contain a limit cycle, or if $\dot{V}(a) = 0$ with $a \neq x^*$ then $x(t) \equiv a$ cannot be a solution; see Example 3.13. The intuition is that an asymptotically stable system will approach points in E that are on a solution trajectory. If the only solution trajectory in E is $x(t) \equiv x^*$ then the system will approach x^*.

Before proving Theorem 3.9 we need to introduce a few concepts. Consider a general dynamical system

$$\dot{x} \quad = \quad f(x(t)) \qquad \text{where } f : D \to \mathbb{R}^n,\ t \geq 0 \tag{3.3}$$

and $D \subseteq \mathbb{R}^n$ is a domain. Given any initial state $x(0) = x_0$ let $\phi(t; x_0)$ denote a solution of (3.3), and $x(t) := \phi(t; x_0)$. A point $x \in D$ is called a *positive limit point* of $x(t)$ if (given x_0) there is a sequence $\{t_i\}$ with $t_i \to \infty$ as $i \to \infty$ such that $x(t_i) \to x$ as $i \to \infty$. The set of all positive limit points of $x(t)$ (given x_0) is called the *positive limit set* of $x(t)$. A set $A \subseteq D$ is called a *positive invariant set* or just an *invariant set* with respect to (3.3) if

$$x(0) \in A \quad \Longrightarrow \quad x(t) \in A \qquad \forall t \geq 0$$

We say that $x(t)$ *approaches a set* A as $t \to \infty$ if for all $\epsilon > 0$ there is an T such that

$$\inf_{y \in A} \|x(t) - y\| \quad < \quad \epsilon \qquad \forall t > T$$

for some norm $\|\cdot\|$.

We now illustrate these concepts with simple examples. If $\dot{x} = -x(t)$, $t \geq 0$, then the origin $\{0\}$ is the only positive limit set of the solution $x(t)$ for any initial state $x(0)$. More generally an asymptotically stable equilibrium point is a positive limit set of every solution starting sufficiently close to the equilibrium point. If $\dot{x} = \sin t$, $t \geq 0$, then $[x(0) - 1, x(0) + 1]$ is the positive limit set of the solution $x(t) = -\cos t$ for any initial state $x(0)$. More generally a stable limit cycle is a positive limit set of every solution starting sufficiently close to the limit cycle. In this case, even though a solution $x(t)$ approaches the limit cycle A as $t \to \infty$, $x(t)$ does not converge to any specific point in A. Hence $x(t)$ approaching a set does not imply the existence of $\lim_{t \to \infty} x(t)$. If $x(t)$ is bounded however then there exists a sequence $(t_i, i = 0, 1, \dots)$ such that $x(t_i)$ converges to a point as $i \to \infty$ by the Bolzano-Weierstrass theorem. An equilibrium point and a limit cycle are invariant sets. Theorem 3.3 implies that the level set $\Omega_c := \{x \in \mathbb{R}^n \mid V(x) \leq c\}$ with $\dot{V}(x) \leq 0$ for all $x \in \Omega_c$ is an invariant set. This set is bounded for any $c > 0$ if $V(x)$ is radially unbounded; see Exercise 3.2.

Finally we need the notion of continuity of solution in initial state. Suppose, given any initial state $x_0 \in D$, there is a unique solution to (3.3) in D, denoted by $\phi(t; x_0)$, $t \geq 0$. Hence $\phi(0; x_0) = x_0$ and $x(t) := \phi(t; x_0)$. We say that the solution $\phi(t; \cdot)$ of (3.3) is *continuous in its initial state* if, given any $x_0 \in D$ and any $\epsilon > 0$, there is an $\delta > 0$ such that for any $y_0 \in D$

$$\|x_0 - y_0\| \ < \ \delta \quad \Longrightarrow \quad \sup_{t \in [0,\infty)} \|\phi(t; x_0) - \phi(t; y_0)\| \ < \ \epsilon$$

Hence if $\phi(t; \cdot)$ is continuous then two trajectories that start close to each other will stay close to each other for all $t \geq 0$.

A positive limit set is important because any solution trajectory $x(t)$ approaches its positive limit set. The next fundamental result does not require f to be continuous and is therefore applicable to projected dynamics. See its proof in Appendix 3.4.[5]

Lemma 3.10 *Consider a compact set $\Omega \subset D$ and suppose, starting from any state in Ω, there is a unique solution to* (3.3) *and the solution is continuous in the initial state. Given an initial state $x(0) = x_0 \in \Omega$, if the corresponding solution $(x(t), t \geq 0)$ lies entirely in the compact set Ω, then its positive limit set $A^+ := A^+(x_0)$ is in Ω, nonempty, compact and invariant. Moreover $x(t)$ approaches A^+ as $t \to \infty$.*

The lemma only requires that (3.3) have a unique solution over the compact set Ω, but does not require Ω to be invariant because the conclusion is for an individual solution trajectory that lies entirely in Ω.

Proof of Theorem 3.9.

We prove part 1 of the theorem. Since V satisfies conditions C3.1 and C3.2, the stability of x^* follows from Theorem 3.3.1 and Corollary 3.5.1 (if a solution exists). In addition we need to prove that $x(t) \to x^*$ as $t \to \infty$, provided the initial state $x(0)$ is close enough to x^*. We now prove this in three steps. Step 1 shows that we can apply Lemma 3.10 to our system (3.1) and hence any solution $x(t)$ will approach its positive limit set. Step 2 shows that the positive limit set is in E. The final step concludes asymptotic stability since the only point in E a solution $x(t)$ can approach is x^*.

Step 1: We show that the (pre-)conditions in Lemma 3.10 are satisfied.

Lemma 3.11 *Suppose f in (3.1) is either locally Lipschitz or a projection of a locally Lipschitz function. Then there is a unique solution $\phi(t, x_0)$, $t \geq 0$, given any initial state $x(0) = x_0$ and the solution $\phi(t, x_0)$ is continuous in x_0.*

[5]If the definition of a positive limit point x of a solution $x(t)$ does not require $x \in D$ (so x can be in $\mathbb{R}^n \setminus D$) then the condition in Lemma 3.10 can be relaxed to "$x(t)$ is bounded and lies in D for all $t \geq 0$."

Proof of Lemma 3.11. As at the beginning of the proof for Theorem 3.3, since V satisfies C3.1 and is continuous, we can find an $\epsilon > 0$ such that the closed-ball $B_\epsilon(V^*) \subseteq D$. Let $v := \min_{x:\|x-x^*\|=\epsilon} V(x) - V(x^*)$ be the minimum value of $V(x) - V(x^*)$ on the boundary of $B_\epsilon(V^*)$ and define the level set

$$\Omega_{v/2} \quad := \quad \{\, x \in B_\epsilon(x^*) \,|\, V(x) - V(x^*) \leq v/2 \,\} \tag{3.4}$$

Then $\Omega_{v/2}$ is bounded and hence compact. Moreover it is invariant, i.e., $x(0) \in \Omega_{v/2}$ implies $x(t) \in \Omega_{v/2}$ for all $t \geq 0$, since $\dot{V}(x) \leq 0$ over D. This means that, given any initial state $x(0) \in \Omega_{v/2}$, a solution, if exists, lies entirely in the compact set $\Omega_{v/2}$. Then, since f is locally Lipschitz or a projection of a locally Lipschitz function, the existence and uniqueness of the solution $x(t)$ follow from Theorems 1.6 and 1.10.

When f is locally Lipschitz, the continuity of the solution $\phi(t; x_0)$ in its initial state x_0 is a standard result; see e.g. [29, Chapter 3.2]. When $f(x) = (\tilde{f}(x))_x^+$ where $\tilde{f}$ is locally Lipschitz, the continuity of $\phi(t; \cdot)$ is proved in [20, Lemma 2]. □

Step 2: We prove LaSalle's invariance principle. Since $\Omega_{v/2}$ defined in (3.4) is compact and invariant, Lemma 3.10 applies to any solution trajectory $x(t)$ that starts in $\Omega_{v/2}$. In particular $x(t)$ approaches its nonempty, compact, invariant positive limit set $A^+ = A^+(x_0)$. Let

$$E_{v/2} \quad := \quad \{x \in \Omega_{v/2} \,|\, \dot{V}(x) = 0\} \quad \subseteq \quad E$$

Lemma 3.12 *Under the conditions of Theorem 3.9, every solution that starts in $\Omega_{v/2}$ approaches the largest invariant set in $E_{v/2}$ as $t \to \infty$.*

Proof of Lemma 3.12. Fix any $x(0) = x_0$ in $\Omega_{v/2}$ and its unique solution $x(t)$ in $\Omega_{v/2}$. Since $\dot{V}(x) \leq 0$ on $\Omega_{v/2}$, $V(x(t))$ is nonincreasing in t and hence $V(x(t))$ approaches a limit v^* as $t \to \infty$. Since $V(x)$ is continuous and $\Omega_{v/2}$ is invariant and compact, v^* is finite. Moreover for any positive limit point $y \in A^+ := A^+(x_0)$ there is a sequence $(t_i, i = 0, 1, \dots)$ with $t_i \to \infty$ as $i \to \infty$ such that $\lim_i x(t_i) = y$. Continuity of V then implies

$$V(y) \quad = \quad V(\lim_i x(t_i)) \quad = \quad \lim_i V(x(t_i)) \quad = \quad v^*$$

i.e., $V(x) = v^*$ on A^+. Since A^+ is invariant (Lemma 3.10) we have $\dot{V}(x) = 0$ on A^+. This means $A^+ \subseteq E_{v/2}$. If M is the largest invariant set in $E_{v/2}$ then

$$A^+ \ \subseteq \ M \ \subseteq \ E_{v/2} \ \subseteq \ \Omega_{v/2} \ \subset \ D$$

Lemma 3.10 implies that $x(t)$ approaches A^+ and hence M as $t \to \infty$. □

Step 3: Since the only solution trajectory that can stay entirely in $E_{v/2}$ is $x(t) \equiv x^*$, $M = \{x^*\}$. Lemma 3.12 then implies that $x(t) \to x^*$ as $t \to \infty$.

This completes the proof of part 1 of Theorem 3.9.

To prove part 2 of Theorem 3.9, note that condition C3.3 implies that $\Omega_c := \{x \in \mathbb{R}^n \mid V(x) - V(x^*) \leq c\}$ is bounded for any $c > 0$ (see Exercise 3.2). Hence given any initial state $x_0 \in D = \mathbb{R}^n$ we can use $c := V(x(0))$. The set Ω_c will be compact and the argument for part 1 proves that $x(t) \to x^*$, proving global asymptotic stability. $\square$

Example 3.13 LaSalle's invariance principle Theorem 3.9. Consider

$$\begin{aligned} \dot{x}_1 &= -2\left(x_1(t) - x_2^2(t)\right) \\ \dot{x}_2 &= \left(-2x_1(t) - x_2(t) - x_2^3(t)\right)^+_{x_2(t)} \end{aligned}$$

where the domain D is $\mathbb{R} \times \mathbb{R}_+$, i.e., $x_2(t) \geq 0$ for all $t \geq 0$. The equilibrium points are given by

$$x_1 = x_2^2, \qquad \left(-2x_1 - x_2 - x_2^3\right)^+_{x_2} = 0$$

The second equality is equivalent to

$$-2x_1 - x_2 - x_2^3 = 0 \quad \text{or} \quad x_2 = 0,\ -2x_1 - x_2 - x_2^3 < 0$$

Substituting $x_1 = x_2^2$ we have $-2x_1 - x_2 - x_2^3 = -x_2(x_2+1)^2$ and hence $x^* = (0,0)$ is the unique equilibrium (recall that $x_2 \geq 0$). We now show that the origin is globally asymptotically stable.

Consider the Lyapunov function candidate

$$V(x) := \frac{1}{2}\left(x_1^2 + x_2^2\right)$$

Clearly V is continuously differentiable, positive definite (condition C3.1) and radially unbounded (C3.3). For condition C3.2 we have

$$\begin{aligned} \dot{V} &= x_1(t)\,\dot{x}_1(t) + x_2(t)\,\dot{x}_2(t) \\ &= -2\,x_1^2(t) + 2\,x_1(t)x_2^2(t) + x_2(t)\left(-2x_1(t) - x_2(t) - x_2^3(t)\right)^+_{x_2(t)} \end{aligned}$$

But the last term $x_2(t)\left(-2x_1(t) - x_2(t) - x_2^3(t)\right)^+_{x_2(t)} = x_2(t)\left(-2x_1(t) - x_2(t) - x_2^3(t)\right)$. Hence

$$\dot{V} = -\left(x_1 - x_2^2\right)^2 - \left(x_1 + x_2\right)^2 \leq 0$$

Moreover if $\dot{V}(x) = 0$ then $x_1 = x_2^2$ and $x_1 = -x_2$, i.e., $x_2(x_2+1) = 0$. Hence

$$\dot{V}(x) = 0 \quad \text{if and only if} \quad x_2 = 0 \ \text{ or } \ x_2 = -1$$

Hence any trajectory $x(t)$ that lies entirely in the set $E := \{x \mid \dot{V}(x) = 0\}$ must have either $x_2(t) \equiv 0$ or $x_2(t) \equiv -1$. Even though the latter trajectory lies in E, it is not a system solution because any solution $x(t)$ must have $x_2(t) \geq 0$ for all $t \geq 0$. Hence any solution trajectory $x(t)$ that stays identically in E must have $x_2 \equiv 0$, and hence, since $x_1 \equiv x_2^2$, $x_1 \equiv 0$, i.e., it must be the trivial solution $x(t) \equiv 0$. The origin is hence globally asymptotically stable by LaSalle's invariance principle Theorem 3.9. $\square$

3.2 STABILITY OF DUAL ALGORITHMS

Consider the class of dual algorithms considered in Chapter 2.2.3, reproduced here:

$$\dot{p}_l = \gamma_l \left(y_l(t) - c_l\right)^+_{p_l(t)} =: g_l(y_l(t), p_l(t)), \qquad l = 1, \ldots, L \tag{3.5a}$$

$$x_i(t) = \left(U_i'^{-1}(q_i(t))\right)^+, \qquad i = 1, \ldots, N \tag{3.5b}$$

where for any $a \in \mathbb{R}$, $[a]^+ := \max\{a, 0\}$ and $U_i'^{-1}$ are the inverses of the derivatives U_i' of the utility functions U_i. For any $a \in \mathbb{R}$ and $b \geq 0$, $(a)_b^+ = a$ if $a > 0$ or $b > 0$ and $(a)_b^+ = 0$ if $a \leq 0$ and $b = 0$. We assume $p(0) \geq 0$ which guarantees that $(x(t), p(t)) \geq 0$, $t \geq 0$. As before,

$$q_i(t) = \sum_l R_{li} p_l(t) \quad \text{and} \quad y_l(t) = \sum_i R_{li} x_i(t)$$

Using (3.5b) we write $x(p) = x(q)$ and $y(p)$ to mean, componentwise,

$$x_i(p) := x_i(q_i) := \left(U_i'^{-1}(q_i)\right)^+ \quad \text{and} \quad y_l(p) := \sum_i R_{li}\, x_i(p)$$

where $q_i := \sum_l R_{li} p_l$. We call p^*, as opposed to (x^*, p^*), an *equilibrium* of the dual algorithm (3.5) if

$$g(y(p^*), p^*) = 0$$

Recall also the network utility maximization:

$$\max_{x \geq 0} \quad \sum_i U_i(x_i) \qquad \text{subject to } Rx \leq c \tag{3.6a}$$

and its Lagrangian dual:

$$\min_{p \geq 0} \quad D(p) := \sum_i \max_{x_i \geq 0} (U_i(x_i) - x_i q_i) + \sum_l p_l c_l \tag{3.6b}$$

As discussed immediately after condition C2.1 and before Theorem 2.14, C2.1 guarantees that a primal-dual optimal solution (x^*, p^*) exists for (3.6). The condition C2.2 there (U_i

are twice continuously differentiable and strictly concave increasing with $U_i''(x_i) < 0$ for all $x_i \geq 0$) in addition ensures that the optimal primal solution x^* is unique. If R has full row rank and the optimal solution $x^* = x(p^*) > 0$ then the optimal dual solution p^* is also unique (see the proof of Theorem 2.14).

Before stating formally the stability property of the dual algorithm (3.5) we first show that LaSalle's invariance principle Theorem 3.9.2 is applicable to (3.5). The dual algorithm can be expressed in terms of only the price vector $p(t)$ as:

$$\dot{p} \quad = \quad \Gamma \left(\tilde{g}(p(t)) \right)^+_{p(t)}, \qquad p(0) \geq 0,\ t \geq 0$$

where $\tilde{g}(p) := y(p) - c$ and $\Gamma = \mathrm{diag}(\gamma_l)$ is a diagonal gain matrix. We claim that this is a projected dynamic with a locally Lipschitz $\tilde{g}(p)$.

Lemma 3.14 *If U_i are twice continuously differentiable with $U_i''(x_i) < 0$ for all $x_i \geq 0$, then $\tilde{g}(p)$ is locally Lipschitz.*

Proof. Fix any $p, \hat{p} \geq 0$ that are close enough. We will show that $|x_i(p) - x_i(\hat{p})| \leq L_i \|p - \hat{p}\|$ for some finite L_i for all i. This implies that

$$\|\tilde{g}(p) - \tilde{g}(\hat{p})\| \quad \leq \quad \sum_i R_{li} \, |x_i(p) - x_i(\hat{p})| \quad \leq \quad \left(\sum_i R_{li} L_i \right) \|p - \hat{p}\|$$

proving that $\tilde{g}(p)$ is Lipschitz.

Note that $(\cdot)^+$ is nonexpansive: for any $a, b \in \mathbb{R}$

$$|a^+ - b^+| \quad = \quad \begin{cases} |a - b| & & \text{if} \quad a > 0, b > 0 \\ a & \leq \ |a - b| & \text{if} \quad a > 0, b \leq 0 \\ b & \leq \ |b - a| & \text{if} \quad a \leq 0, b > 0 \\ 0 & \leq \ |a - b| & \text{if} \quad a \leq 0, b \leq 0 \end{cases}$$

i.e., $|a^+ - b^+| \leq |a - b|$. Hence

$$|x_i(p) - x_i(\hat{p})| \quad = \quad \left| \left(U_i'^{-1}(p) \right)^+ - \left(U_i'^{-1}(\hat{p}) \right)^+ \right| \quad \leq \quad \left| U_i'^{-1}(p) - U_i'^{-1}(\hat{p}) \right|$$

Since U_i are twice continuously differentiable, $U_i'^{-1}(p)$ are continuously differentiable with

$$\frac{\partial}{\partial p_l} U_i'^{-1}(p) \quad = \quad \frac{R_{li}}{U_i''(x_i(p))}$$

Lemma 1.3 then implies that $U_i'^{-1}(p)$ are locally Lipschitz and hence $|x_i(p) - x_i(\hat{p})| \leq L_i \|p - \hat{p}\|$ locally for some Lipschitz constants L_i. This completes the proof of the lemma. $\square$

Hence the dual algorithm (3.5) is a projected dynamical system with an underlying Lipschitz derivative $\tilde{g}(p)$, to which Theorem 3.9 is applicable for proving stability. This also guarantees the existence and uniqueness of a solution; see Remark 3.17 after the next theorem.

To simplify proofs we strengthen condition C2.2 to the following:

> C2.2': U_i are twice continuously differentiable and strictly concave increasing on $\mathbb{R}_+$, with $U_i''(x_i) < 0$ for $x_i \geq 0$. Moreover we assume that $\lim_{x_i \to 0} U_i'(x_i) = \infty$ and $\lim_{x_i \to \infty} U_i'(x_i) = 0$.

Condition C2.2' guarantees that for any $p \in \mathbb{R}_+^L$ such that $q_i(p) > 0$ there is a unique maximizer $x_i(p)$ for $\max_{x_i \geq 0}\ U_i(x_i) - x_i q_i(p)$ given by

$$x_i(p) \quad := \quad U_i'^{-1}(q_i(p)) \quad \geq \quad 0$$

By Danskin's theorem (e.g. [8, p. 649]) and the fact that a differentiable convex function is continuously differentiable, the dual objective function

$$D(p) \quad = \quad \sum_i \max_{x_i \geq 0} \left(U_i(x_i) - x_i\, q_i(p)\right) \; + \; \sum_l c_l\, p_l$$

is continuously differentiable in p and its derivative is

$$\frac{\partial D}{\partial p}(p) \quad = \quad c - Rx(p) \quad = \quad c - y(p)$$

Theorem 3.15 *Suppose the conditions C2.1 and C2.2' hold and R has full row rank. Then the unique equilibrium p^* of (3.5) is globally asymptotically stable, provided $x(p^*) > 0$.*

Proof. Theorem 2.14 proves the existence and uniqueness of the equilibrium p^* and the primal-dual optimality of p^* and the associated source rates $x^* := x(p^*) > 0$. To prove that p^* is globally asymptotically stable we will show that $D(p)$ is continuously differentiable and satisfies conditions C3.1, C3.2, C3.3 and hence is a Lyapunov function for the dual algorithm (3.5). Moreover the only solution of (3.5) that can stay identically in the set $\{p \,|\, \dot{D}(p) = 0\}$ is the trivial solution $p(t) \equiv p^*$. LaSalle's invariance principle (Theorem 3.9.2) then implies the global asymptotic stability of p^*.

We now prove that $D(p)$ satisfies conditions C3.1, C3.2, C3.3.

C3.1: The strict concavity of U_i and full row rank of R mean that the optimal p^* is unique and hence

$$D(p) \quad > \quad D(p^*) \qquad \text{for all } p \neq p^*$$

Moreover $D(p^*)$ is finite by Theorem 2.14.

C3.2: We have for all $p(t) \neq p^*$

$$\begin{aligned}\dot{D} &= \frac{\partial D}{\partial p} g(y(p(t)), p(t)) = (c - y(p(t)))^T (y(p(t)) - c)^+_{p(t)} \\ &= -\sum_l (y_l(p(t)) - c_l)\,(y_l(p(t)) - c_l)^+_{p_l(t)} \\ &= -\sum_l (y_l(p(t)) - c_l)^2 \;\mathbf{1}\,(y_l(p(t)) > c_l \text{ or } p_l(t) > 0)\end{aligned}$$

where the indicator function $\mathbf{1}(P) := 1$ if P is true and 0 otherwise. Hence $\dot{D} \leq 0$.

C3.3: We have to show that $D(p)$ is radially unbounded, i.e., $D(p) \to \infty$ as $\|p\| \to \infty$, under condition C2.2. For ease of reference in the future, this is proved as Lemma 3.16 below.

This completes the proof that $D(p)$ is a Lyapunov function for the dual algorithm (3.5).

LaSalle's invariance principle (Lemma 3.12) implies that any solution $p(t)$ of (3.5) will converge to the largest invariant set in

$$E := \{\, p \in \mathbb{R}^L_+ \mid \dot{D}(p) = 0 \,\}$$

From the proof above of condition C3.2, $\dot{D}(p) = 0$ if and only if, for each $l = 1, \ldots, L$,

$$p_l \geq 0 \quad \text{and} \quad y_l(p) \leq c_l \;\text{ with equality if } p_l > 0$$

This means that p is an equilibrium point of (3.5). Since the equilibrium is unique we must have $p = p^*$. Hence the only solution $p(t)$ of (3.5) that can stay identically in E is the trivial solution $p(t) \equiv p^*$. Theorem 3.9.2 then implies the global asymptotic stability of p^*. □

Lemma 3.16 *If U_i are strictly concave and continuously differentiable then the dual objective function $D(p)$ is radially unbounded.*

Proof. Since U_i is strictly concave and continuously differentiable, $U_i'^{-1}(q_i)$ is a strictly decreasing continuous function of q_i. Hence $y_l(p) = \sum_i R_{li} x_i(p) = \sum_i R_{li} \left(U_i'^{-1}(q_i)\right)^+$ is strictly decreasing in p_l since $q_i(p)$ is strictly decreasing in p_l if $R_{li} = 1$, unless $y_l(p) = 0$ for large enough p. This means that there is an $\epsilon > 0$ and, for *every* link $l = 1, \ldots, L$, there is a finite threshold $\tilde{p}_l$ such that for *any* $p \geq 0$,

$$p_l \geq \tilde{p}_l \implies c_l - y_l(p) > \epsilon, \qquad l = 1, \ldots, L \tag{3.7}$$

i.e., the flow rate $y_l(p)$ is smaller than c_l by more than ϵ if $p_l \geq \tilde{p}_l$ even if all other links $k \neq l$ have zero prices. Let $\tilde{p} := (\tilde{p}_l,\, l = 1, \ldots, L)$. Then (3.7) implies in particular that

$$c_l - y_l(\tilde{p}) > \epsilon, \qquad l = 1, \ldots, L \tag{3.8}$$

We will take $\|p\| \to \infty$ and compare the gap $D(p) - D(\tilde{p})$.

$$\begin{aligned}
& D(p) - D(\tilde{p}) \\
= & \sum_i (U_i(x_i(p)) - U_i(x_i(\tilde{p}))) \;-\; \sum_l (p_l y_l(p) - \tilde{p}_l y_l(\tilde{p})) \;+\; \sum_l c_l\,(p_l - \tilde{p}_l) \\
\geq & \sum_i U_i'(x_i(p))\,(x_i(p) - x_i(\tilde{p})) \;-\; \sum_l (p_l y_l(p) - \tilde{p}_l y_l(\tilde{p})) \;+\; \sum_l c_l\,(p_l - \tilde{p}_l) \\
\geq & \sum_i q_i(p)\,(x_i(p) - x_i(\tilde{p})) \;-\; \sum_l (p_l y_l(p) - \tilde{p}_l y_l(\tilde{p})) \;+\; \sum_l c_l\,(p_l - \tilde{p}_l)
\end{aligned} \tag{3.9}$$

where the first inequality follows from the concavity of U_i. To see the second inequality, note that since $x_i(p) = \arg\max_{x_i \geq 0} U_i(x_i) - x_i q_i(p)$, we have

$$U_i'(x_i(p)) = q_i(p) \quad \text{or} \quad U_i'(x_i(p)) < q_i(p) \text{ with } x_i(p) = 0$$

In either case we have, since $x_i(\tilde{p}) \geq 0$,

$$U'(x_i(p))\,(x_i(p) - x_i(\tilde{p})) \;\;\geq\;\; q_i(p)\,(x_i(p) - x_i(\tilde{p}))$$

Substitute $q_i = \sum_l R_{li} p_l$ and $y_l = \sum_i R_{li} x_i$ into (3.9) to obtain

$$\begin{aligned}
& D(p) - D(\tilde{p}) \\
\geq & \sum_l \sum_i R_{li}\, p_l\,(x_i(p) - x_i(\tilde{p})) \;-\; \sum_l \sum_i R_{li}\,(p_l x_i(p) - \tilde{p}_l x_i(\tilde{p})) \;+\; \sum_l c_l\,(p_l - \tilde{p}_l) \\
= & \sum_l (p_l - \tilde{p}_l) \left(c_l - \sum_i R_{li} x_i(\tilde{p}) \right)
\end{aligned} \tag{3.10}$$

Substituting (3.8) into (3.10) we have

$$\begin{aligned}
D(p) - D(\tilde{p}) \;\; & \geq \;\; \sum_l (p_l - \tilde{p}_l)\,(c_l - y_l(\tilde{p})) \\
& \geq \;\; \epsilon \sum_{l: p_l > \tilde{p}_l} (p_l - \tilde{p}_l) \;\;-\; \sum_{l: p_l \leq \tilde{p}_l} \tilde{p}_l (c_l + y_l(\tilde{p}))
\end{aligned} \tag{3.11}$$

where the last inequality follows from

$$p_l \leq \tilde{p}_l \quad \Longrightarrow \quad (p_l - \tilde{p}_l)\,(c_l - y_l(\tilde{p})) \;\geq\; -\tilde{p}_l (c_l + y_l(\tilde{p}))$$

The second term on the right-hand side of (3.11) is constant. Hence $D(p) \to \infty$ as $\|p\| \to \infty$ as desired. □

Remark 3.17 Assumptions C2.1 and C2.2 on dual algorithm. The assumptions C2.1 and C2.2 in Chapter 2.2.3 have implications on the existence and uniqueness of solution trajectory $(p(t), x(p(t)), t \geq 0)$ of the dual algorithm, the existence and uniqueness of the equilibrium point p^*, and its global asymptotic stability. All of these are brought together after completing the proof of Theorem 3.15.

1. *Solution.* The dual algorithm $g(p) := (\tilde{g}(p))_p^+$ is the projection of $\tilde{g}$. Lemma 3.14 shows that $\tilde{g}$ is locally Lipschitz under C2.2. The proof Theorem 3.15 constructs a Lyapunov function V that satisfies C3.1, C3.2, C3.3. Then, starting from any initial point $p_0 \geq 0$, a solution trajectory $p(t)$, if exists, lies entirely in the closed set $\Omega := \{p \in \mathbb{R}_+^L \mid V(p) \leq V(p_0)\}$. The radial unboundedness condition C3.3 implies Ω is bounded, and hence compact. Therefore a solution indeed exists and is unique (Remark 3.6 and Theorem 1.10.2).

2. *Equilibrium.* As discussed before Theorem 2.14, condition C2.1 guarantees that a primal-dual optimal solution (x^*, p^*) exists for (3.6). C2.2 (or C2.2') on the second derivative of U_i ensures that the primal optimal $x(p^*)$ is unique and when R has full row rank, the dual optimal p^* is also unique when $x(p^*) > 0$.

3. *Stability.* The Lyapunov function V and LaSalle's invariance principle prove the global asymptotic stability of the unique equilibrium (Theorem 3.9.2).

3.3 STABILITY OF PRIMAL-DUAL ALGORITHMS

We consider a class of primal-dual algorithms of the form

$$\begin{aligned}\dot{x} &= K\left(\frac{\partial L}{\partial x}(x,p)\right)_{x(t)}^+ \\ \dot{p} &= \Gamma\left(-\frac{\partial L}{\partial p}(x,p)\right)_{p(t)}^+\end{aligned}$$

where K and Γ are invertible gain matrices. Note that for any $a, b \in \mathbb{R}$,

$$(-a)_b^+ \quad \neq \quad -(a)_b^+$$

Here $L(x,p)$ is concave in x and convex in p. Primal-dual algorithms have dynamics in both the source rates x and congestion prices p. Often the function L is the Lagrangian of a constrained optimization problem and the algorithm is a first-order saddle point algorithm that iterates on steepest ascent in the x direction and steepest descent in the p direction.

For our purposes consider the network utility maximization and its dual problem (3.6). Its Lagrangian is

$$L(x,p) \quad := \quad \sum_i U_i(x_i) \; - \; p^T(Rx - c) \tag{3.12}$$

Then the primal-dual algorithm takes a decentralized form

$$\dot{x}_i = \kappa_i \left(U_i'(x_i(t)) - q_i(t)\right)_{x_i(t)}^+ =: f_i(x_i(t), q_i(t)), \qquad i = 1, \ldots, N \tag{3.13a}$$

$$\dot{p}_l = \gamma_l \left(y_l(t) - c_l\right)_{p_l(t)}^+ =: g_l(y_l(t), p_l(t)), \qquad l = 1, \ldots, L \tag{3.13b}$$

When U_i are twice continuously differentiable the dynamics of (3.13) are defined by the projection of locally Lipschitz functions. We assume $(x(0), p(0)) \geq 0$. A point (x^*, p^*) is an *equilibrium* of the primal-dual algorithm (3.13) if $f(x^*, p^*) = 0, g(x^*, p^*) = 0$, i.e.,

$$\begin{array}{rll} U'(x^*) \leq q^*, & x^* \geq 0, & (x^*)^T (U'(x^*) - q^*) = 0 \\ y^* = Rx^* \leq c, & p^* \geq 0, & (p^*)^T (y^* - c) = 0 \end{array}$$

As shown in Theorem 2.14 this means (x^*, p^*) is an equilibrium of (3.13) if and only if x^* is an optimal solution of the network utility maximization (3.6a) and p^* is an optimal solution of its dual (3.6b). Moreover when U_i are twice continuously differentiable and strictly concave increasing (condition C2.2), x^* is unique. When R has full row rank and $x^* > 0$, p^* is also unique.

The same conditions that guarantee the global asymptotic stability of dual algorithms also guarantee that of the primal-dual algorithm. We start with a simple lemma that will be repeatedly used to remove the projection operation.

Lemma 3.18 *For any scalars $a \geq 0, b \geq 0$ and $c \in \mathbb{R}$, we have*

$$\begin{array}{rcl} (a-b)(c)_a^+ & \leq & (a-b)c \\ (a-b)(c)_b^+ & \geq & (a-b)c \end{array}$$

where $(c)_a^+ := c$ if $a > 0$ or $c > 0$ and 0 otherwise. If $a := (a_i, i = 1, \ldots, n)$, $b := (b_i, i = 1, \ldots, n)$, $c := (c_i, i = 1, \ldots, n)$ are vectors in $\mathbb{R}^n$ then the following inequalities hold:

$$\begin{array}{rcl} (a-b)^T(c)_a^+ & \leq & (a-b)^T c \\ (a-b)^T(c)_b^+ & \geq & (a-b)^T c \end{array}$$

with the interpretation $(a-b)^T(c)_a^+ := \sum_i (a_i - b_i)(c_i)_{a_i}^+$ and likewise for $(a-b)^T(c)_b^+$.

Proof. We will prove the case when a, b, c are scalars; it implies the vector case. The first inequality is an equality if $a > 0$ or $c \geq 0$. When $a = 0$ and $c < 0$, the left-hand side is zero but the right-hand side is nonnegative. The second inequality is proved similarly. $\square$

While the dual objective function $D(p)$ serves as a Lyapunov function for proving the global asymptotic stability of the dual algorithm, the Lagrangian $L(x, p)$ cannot serve as a Lyapunov function for the primal-dual algorithm because the algorithm attempts to maximize L over x but minimize L over p.

Theorem 3.19 *Suppose the conditions C2.1 and C2.2 hold and R has full row rank. Then the unique equilibrium (x^*, p^*) of (3.13) is globally asymptotically stable, provided $x^* > 0$.*

Proof. The discussion preceding Lemma 3.18 shows that there exists a unique equilibrium $z^* := (x^*, p^*)$ that is primal-dual optimal for (3.6) since U_i are strictly concave and R has full row rank. Let $z := (x, p)$ and consider the quadratic Lyapunov function candidate

$$V(z) \;\; := \;\; \frac{1}{2}\,(z - z^*)^T \begin{bmatrix} K^{-1} & 0 \\ 0 & \Gamma^{-1} \end{bmatrix} (z - z^*)$$

where $K := \mathrm{diag}(\kappa_i, i = 1, \ldots, N)$ and $\Gamma := \mathrm{diag}(\gamma_l, l = 1, \ldots, L)$. Hence $V(z) = 0$ if and only if $z = z^*$. Since $K > 0, \Gamma > 0$, $V(z)$ is positive definite (i.e., $V(z) > 0$ for all $z \neq z^*$) and radially unbounded. Hence $V(z)$ is continuously differentiable and satisfies conditions C3.1 and C3.3. To prove the global asymptotic stability of z^* we now show that $\dot{V}(z) \leq 0$ for all $z \neq z^*$ (condition C3.2) and that the only solution to (3.13) that can maintain $\dot{V}(z) \equiv 0$ is the equilibrium trajectory $z(t) \equiv z^*$. LaSalle's Theorem 3.9.2 then implies the global asymptotic stability of z^*.

Write the primal-dual algorithm (3.13) in terms of the Lagrangian:

$$\dot{z} \;\; = \;\; \begin{bmatrix} K & 0 \\ 0 & \Gamma \end{bmatrix} \begin{bmatrix} \left(\frac{\partial L}{\partial x}(x(t), p(t))\right)^+_{x(t)} & 0 \\ 0 & \left(-\frac{\partial L}{\partial p}(x(t), p(t))\right)^+_{p(t)} \end{bmatrix}$$

Then

$$\begin{aligned}
\dot{V} \;\; &= \;\; (z(t) - z^*)^T \begin{bmatrix} K^{-1} & 0 \\ 0 & \Gamma^{-1} \end{bmatrix} \dot{z} \\
&= \;\; (x(t) - x^*)^T \left(\frac{\partial L}{\partial x}(x(t), p(t))\right)^+_{x(t)} \; + \; (p(t) - p^*)^T \left(-\frac{\partial L}{\partial p}(x(t), p(t))\right)^+_{p(t)} \\
&\leq \;\; (x(t) - x^*)^T \, \frac{\partial L}{\partial x}(x(t), p(t)) \; - \; (p(t) - p^*)^T \, \frac{\partial L}{\partial p}(x(t), p(t))
\end{aligned}$$

where the inequality follows from Lemma 3.18. Since L is strictly concave in x and convex (linear) in p we have, for $z \neq z^*$,[6]

$$\begin{aligned}
(x - x^*)^T \, \frac{\partial L}{\partial x}(x, p) \;\; &\leq \;\; L(x, p) - L(x^*, p) \\
(p - p^*)^T \, \frac{\partial L}{\partial p}(x, p) \;\; &\geq \;\; L(x, p) - L(x, p^*)
\end{aligned}$$

Hence

$$\begin{aligned}
\dot{V} \;\; &\leq \;\; L(x(t), p^*) - L(x^*, p(t)) \\
&= \;\; (L(x(t), p^*) - L(x^*, p^*)) \; + \; (L(x^*, p^*) - L(x^*, p(t))) \;\; \leq \;\; 0
\end{aligned} \tag{3.14}$$

[6]Even though L is strictly concave in x, the inequalities are both non-strict for $z(t) \neq z^*$ with $z(t) = (x^*, p(t))$.

where the second inequality follows from the saddle-point property of the optimal point (x^*, p^*):

$$L(x, p^*) \quad \le \quad L(x^*, p^*) \quad \le \quad L(x^*, p) \qquad \forall (x, p)$$

This completes the proof that $V(z)$ satisfies condition C3.2 and is hence a Lyapunov function.

To use LaSalle's Theorem 3.9 it suffices to show that only the equilibrium trajectory $z(t) \equiv z^*$ of (3.13) can stay identically in the set

$$E \quad := \quad \{ \ z \ge 0 \mid \dot{V}(z) = 0 \ \}$$

From (3.14) and the saddle-point property of z^* we have $\dot{V}(z(t)) \equiv 0$ (i.e., $\dot{V}(z(t)) = 0$ for all $t \ge 0$) only if

$$L(x(t), p^*) \ \equiv \ L(x^*, p^*) \qquad \text{and} \qquad L(x^*, p(t)) \ \equiv \ L(x^*, p^*) \tag{3.15}$$

The *strict* concavity of U and hence L means that the maximizer x^* of $L(\cdot, p^*)$ is unique and therefore we must have $x(t) \equiv x^*$ from the first equivalence in (3.15). In particular $\dot{x} \equiv 0$. Hence we have from (3.13a)

$$U'(x^*) \le q(p(t)), \quad x^* \ge 0, \quad (x^*)^T (U'(x^*) - q(p(t))) = 0, \qquad t \ge 0 \tag{3.16a}$$

We now argue that $(x(t) \equiv x^*, p(t))$ also satisfies

$$y^* = Rx^* \le c, \quad p(t) \ge 0, \quad (p(t))^T (y^* - c) = 0, \qquad t \ge 0 \tag{3.16b}$$

But (3.16) implies that $(x(t), p(t))$ are equilibrium points of (3.13) for all $t \ge 0$. Since the equilibrium is unique we must have $(x(t), p(t)) \equiv (x^*, p^*)$. Hence we have shown that if $z(t)$ stays identically in E then $z(t) \equiv z^*$ is the equilibrium trajectory. The theorem then follows from Theorem 3.9.2.

Therefore the proof is complete if we can establish (3.16b). The first condition $Rx^* \le c$ holds since x^* is a primal feasible point for (3.6a). The second condition $p(t) \ge 0$ of (3.16b) holds since $(x(t), p(t))$ is a solution of (3.13). Finally substitute (3.12) into the second equivalence in (3.15) to obtain

$$(p(t))^T (y^* - c) \quad \equiv \quad (p^*)^T (y^* - c) \quad = \quad 0$$

Hence (x^*, p) satisfies (3.16b) and the proof is complete. □

3.4 APPENDIX: PROOF OF LEMMA 3.10

Fix an initial state $x(0) = x_0$ in the compact set Ω and consider the unique solution $x(t)$, $t \ge 0$. We first prove that its positive limit set $A^+ := A^+(x_0)$ is nonempty, compact, and invariant.

The Bolzano-Weierstrass theorem states that every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence. Since $x(t)$ is bounded there is a sequence $(t_i, i = 0, 1, \dots)$ with $t_i \to \infty$ as $i \to \infty$ such that $\lim_{i\to\infty} x(t_i) = x$ for some x in $\mathbb{R}^n$. Since $x(t)$ lies entirely in the compact set Ω, $x \in \Omega \subset D$ and is hence a positive limit point of $x(t)$, proving that A^+ is nonempty.

Since all positive limit points lie in Ω, $A^+ \subseteq \Omega$ and is bounded. To show that A^+ is compact, we will show that A^+ is also closed, i.e., if $y^n \in A^+$, $n = 0, 1, \dots,$ and $\lim_n y^n = y$, then $y \in A^+$. For each n there is a sequence $(t_i^n, i = 0, 1, \dots)$ such that the solution $x(t)$ satisfies $\lim_i x(t_i^n) = y^n$. We now construct a sequence $\left(t_{i(n)}^n, n = 0, 1, \dots\right)$ and prove that the solution $x(t)$ on this sequence converges to y as $n \to \infty$, certifying that $y \in A^+$. For each n, since $\lim_i x(t_i^n) = y^n$, there is an $i(n)$ such that $i(n) \geq n$ and

$$\left\| x\left(t_{i(n)}^n\right) - y^n \right\| \quad < \quad \frac{1}{n}$$

By construction, $i(n) \to \infty$ as $n \to \infty$. Moreover given any $\delta > 0$ we have

$$\left\| x\left(t_{i(n)}^n\right) - y^n \right\| \quad < \quad \delta \qquad \forall n > \frac{1}{\delta}$$

To show that $x\left(t_{i(n)}^n\right)$ converges to y, fix an $\epsilon > 0$. Since $\lim_n y^n = y$ there is an N_y such that $\|y^n - y\| < \epsilon/2$ for all $n > N_y$. Let $N := \max\{N_y, \lceil 2/\epsilon \rceil\}$. We have for any $n > N$

$$\|x(t_{i(n)}^n) - y\| \quad \leq \quad \left\| x\left(t_{i(n)}^n\right) - y^n \right\| + \|y^n - y\| \quad < \quad \frac{\epsilon}{2} + \frac{\epsilon}{2} \quad = \quad \epsilon$$

as desired. This proves that A^+ is closed and bounded, and hence compact.

We now prove that A^+ is invariant. Note that $A^+ := A^+(x_0)$ is the positive limit set of the unique solution $x(t)$ with the given initial state $x(0) = x_0$. Suppose the system starts in state $y \in A^+ \subseteq \Omega$. Denote the unique solution by $\phi(t; y)$. Hence $\phi(0; y) = y$ and $x(t) = \phi(t; x_0)$. We will show that $\phi(t; y) \in A^+$ for all $t \geq 0$, by exhibiting a sequence of times $(\tau_i, i = 0, 1, \dots)$ for each $t \geq 0$ such that $x(\tau_i) = \phi(\tau_i; x_0) \to \phi(t; y)$ as $i \to \infty$. Since $y \in A^+$ there is a sequence $(t_i, i = 0, 1, \dots)$ such that $x(t_i) = \phi(t_i; x_0) \to y$ as $i \to \infty$. Hence for sufficiently large i, $x(t_i)$ is close to y. Compare two solution trajectories at different times: the state $\phi(t; y)$ reached at time t starting from the initial state y, and the state $\phi(t_i + t; x_0)$ reached at time $t_i + t$ starting from the initial state x_0. We have, for any $t \geq 0$,

$$\lim_i \|\phi(t + t_i; x_0) - \phi(t; y)\| \quad = \quad \lim_i \|\phi(t; x(t_i)) - \phi(t; y)\| \quad = \quad \|\phi(t; \lim_i x(t_i)) - \phi(t; y)\|$$

where the first equality follows from the uniqueness of solution trajectory, and the second equality follows from the continuity of $\phi(t; \cdot)$ in its initial state. Since $\lim_i x(t_i) = \lim_i \phi(t_i; x_0) = y$, we have

$$\lim_i \|\phi(t + t_i; x_0) - \phi(t; y)\| \quad = \quad 0 \qquad \forall t \geq 0$$

Since $x(t+t_i) = \phi(t+t_i; x_0)$ stays in Ω for all time, we have shown that, for each $t \geq 0$, $\phi(t; y)$ is a positive limit point of the solution $x(t)$. Hence $\phi(t; y) \in A^+$ for all $t \geq 0$ if $y \in A^+$, i.e., A^+ is invariant.

Finally we prove that $x(t)$, corresponding to the given initial state $x_0 \in \Omega$, approaches A^+ as $t \to \infty$. Suppose not and there is an $\epsilon > 0$ such that for all integer $k \geq 0$ there is a time $t_k \geq k$ such that

$$\inf_{y \in A^+} \|x(t_k) - y\| \quad \geq \quad \epsilon \tag{3.17}$$

Since $x(t)$ stays entirely in Ω, the sequence $(x(t_k), k = 0, 1, \ldots)$ has a subsequence $\left(x\left(t_{k_j}\right), j = 0, 1, \ldots\right)$ that converges to a positive limit point $x^* \in A^+$. Hence

$$\lim_j \left\|x\left(t_{k_j}\right) - x^*\right\| \quad = \quad 0$$

This contradicts (3.17) since $x\left(t_{k_j}\right)$ is a subsequence of $x(t_k)$. Hence $x(t)$ approaches A^+ as $t \to \infty$.

This completes the proof of Lemma 3.10. □

3.5 BIBLIOGRAPHICAL NOTES

There are many excellent texts on nonlinear systems and Lyapunov stability theory. We have included detailed proofs of these results for smooth systems, largely following [29], because we need to extend these results to discontinuous projected dynamics. For our purposes it suffices to still use a Lyapunov functions $V(x)$ that is continuously differentiable, but its rate $\dot{V}(x) := \frac{\partial V}{\partial x}(x)\, f(x)$ is generally discontinuous when f is projected dynamics. It turns out that continuity of $\dot{V}(x)$ is crucial only for the proof of the asymptotic stability in Theorem 3.3.2. It seems that the existence, uniqueness and continuity with respect to initial condition of solutions to projected dynamics are first established in [20], where projection is to a closed convex set in $\mathbb{R}^n$. Lyapunov stability theorems and LaSalle's invariance principle are extended in [6] to a more general setting (than what we use here) of differential equations with discontinuous right-hand side and nonsmooth Lyapunov functions. Our proofs in Chapter 3.1 are simple extensions of the classical result to projected dynamics. Strictly speaking, LaSalle's invariance principle is Lemma 3.12. Theorem 3.9 are due to Barbashin and Krasovskii who proved it before LaSalle's invariance principle. Some of the applications of these stability results to congestion control algorithms modeled by projected dynamics are new.

The dual algorithms in Chapter 3.2 is introduced in [35] to model TCP congestion control. Global stability proof of the primal-dual algorithm in Chapter 3.3 is adapted from [21] which also uses Krasovskii's method [29] for stability analysis. Unlike [21] that treats a projected dynamical system as a hybrid dynamical system where discontinuity is captured by transitions of discrete states, we prove stability directly using stability theorems for projected dynamics. See also [15] for proofs that build on results in [6]. The proof in [21] does not provide a complete argument why the Lagrange multiplier $p(t)$ converges to a point (instead of approaching the set of optimal p^*) when p^* is nonunique. We bypass this issue by assuming that the routing matrix R is of full row rank and $x^* > 0$, guaranteeing the uniqueness of p^*. See [56, 33] for a convergence argument using a quadratic Lyapunov function in the presence of multiple equilibrium points. Quadratic Lyapunov functions are first used to study primal-dual algorithms in [3]. See also [42, 44] for the analysis of various congestion control algorithms.

3.6 PROBLEMS

Exercise 3.1 (Global asymptotic stability)**.** Show that if x^* is globally asymptotically stable then it is the unique equilibrium.

Exercise 3.2 (Level sets)**.** Consider a continuous function $V : D \to \mathbb{R}$ where D is a domain (open connected set in $\mathbb{R}^n$).

1. Suppose V satisfies condition C3.1: $V(x^*)$ is finite and $V(x) > V(x^*)$ for all $x \neq x^*$ in D. Is the level set $\Omega_c := \{x \in \mathbb{R}^n | V(x) - V(x^*) \leq c\}$ always bounded for sufficiently small $c > 0$? Prove or give a counterexample.

2. Suppose V satisfies C3.1. Does there exists a level set Ω_c that contains x^*, is bounded, and $V(x) - V(x^*) \leq c$ for all $x \in \Omega_c$, for sufficiently small $c > 0$? Prove or give a counterexample.

3. Suppose V is radially unbounded: $V(x) \to \infty$ as $\|x\| \to \infty$. Show that the level set $\Omega_c := \{x \in \mathbb{R}^n | V(x) - V(x^*) \leq c\}$ is bounded for all $c > 0$.

Exercise 3.3 (Asymptotic stability)**.** Consider the dynamical system

$$\begin{aligned} \dot{x}_1 &= \quad -x_2 - x_1 \sin(x_1^2 + x_2^2) \\ \dot{x}_2 &= \quad x_1 - x_2 \sin(x_1^2 + x_2^2). \end{aligned}$$

Prove that the origin is asymptotically stable for $\|x\|_2^2 < \pi$.

Exercise 3.4. Consider the linear system $\dot{x} = Ax$ where $x \in \mathbb{R}^n$. Prove that A is asymptotically stable if and only if there is there is a positive definite matrix $P \succ 0$ that solves

$$A^T P + PA \quad = \quad -Q$$

for any positive definite matrix $Q \succ 0$.

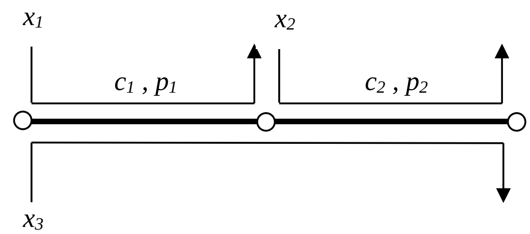


Figure 3.6: The network for Exercise 3.5.

Exercise 3.5 (Stability of TCP)**.** Consider the network in Fig. 3.6. Suppose the TCP algorithms are given by

$$\begin{aligned} x_1(t) &= \frac{1}{p_1(t)}, \\ x_2(t) &= \frac{1}{\sqrt{p_2(t)}}, \\ x_3(t) &= \frac{2}{(p_1(t)+p_2(t))^{1/3}} \end{aligned}$$

and the queue management algorithms are given by

$$\begin{aligned} \frac{d}{dt}p_1(t) &= \gamma(x_1(t)+x_3(t)-c_1), \\ \frac{d}{dt}p_2(t) &= \gamma(x_2(t)+x_3(t)-c_2). \end{aligned}$$

1. Find the utility functions of the 3 flows and write down the network utility maximization problem implicitly solved by this algorithm. [Hint: Write down the equilibrium condition and interpret that as the optimality condition of a network utility maximization problem.]

2. Is the equilibrium point (x^*, p^*) unique? Explain.

3. Prove that the equilibrium point (x^*, p^*) is asymptotically stable. [Hint: Try the dual objective function as a candidate Lyapunov function.]

Exercise 3.6 (Lyapunov stability of primal algorithms [28])**.** Consider the class of algorithms:

$$\dot{x}_i = \kappa_i\left(a_i - x_i(t)q_i(t)\right) =: f_i(x_i(t), q_i(t)) \tag{3.18a}$$
$$p_l(t) = g_l(y_l(t)) \tag{3.18b}$$

Primal algorithms have dynamics only in the source rates, not in the congestion prices. As before,

$$q_i(t) \;=\; \sum_l R_{li} p_l(t) \quad \text{and} \quad y_l(t) \;=\; \sum_i R_{li} x_i(t)$$

Since $p(t)$, and hence $q(t)$, are statically determined by $x(t)$ through (3.18b), we will call x^*, as opposed to (x^*, p^*), an equilibrium of the primal algorithm if $f(x^*) = f(x^*, q(x^*)) = 0$ where $q(x^*) := (q_i(x^*), i \in N)$ is given by

$$q_i(x^*) \;:=\; \sum_l R_{li}\, p_l(x^*) \;:=\; \sum_l R_{li}\, g_l(y_l^*), \qquad i = 1, ..., N$$

and $y_l^* := \sum_i R_{li} x_i^*$. Suppose g_l are nonnegative, continuous, increasing functions that are not identically zero.

Prove:

1. The primal algorithm (3.18) has a unique equilibrium x^* that is the unique solution of

$$\max_{x \in \mathbb{R}^N} \qquad V(x) \;:=\; \sum_i a_i \log x_i \;-\; \sum_l \int_0^{y_l} g_l(z)\, dz$$

2. Moreover x^* is globally asymptotically stable.

The exercise implies that the primal algorithm (3.18) solves a relaxation of the network utility maximization with log utility functions, where the capacity constraint $Rx \leq c$ is replaced by the penalty function $\int_0^{y_l} g_l(z)dz$ in the objective. The reason that the primal algorithm does not solve the exact network utility maximization problem with log utility function is that the functions $g_l(y_l)$ are independent of p_l and hence do not satisfy condition C2.5 in Chapter 2.2.4. This condition is required to ensure the complementary slackness condition of network utility maximization. As the function $g_l(z)$ is chosen to impose heavier penalty on violating the constraint, the equilibrium x^* will approach the solution of the network utility maximization.

CHAPTER 4

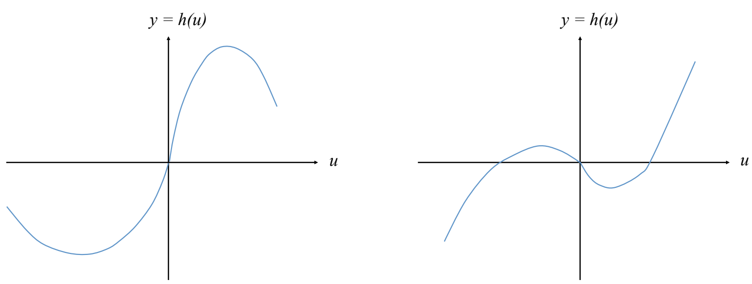
Global stability: passivity method

In this chapter we explain the concept of passive systems and how to use passivity theorems to prove the stability of congestion control algorithms. One of the most important features of passivity theory is that it allows us to study the stability of an interconnected system in terms of the passivity properties of its component systems.

4.1 PASSIVE SYSTEMS

Memoryless systems. We start with a memoryless system $y = h(u)$ where $h : \mathbb{R}^p \to \mathbb{R}^p$. The number p of inputs is equal to the number of outputs. We say that a (time-invariant)[1] memoryless system $y = h(u)$ is *passive* if $u^T y \geq 0$.

For a scalar system with $p = 1$, the graph of h only appears in the first and third quadrants, as shown in Figure 4.1. Hence a very useful property of a scalar passive system



(a) Passive system (b) Nonpassive system

Figure 4.1: Examples of nonlinear memoryless passive and nonpassive *scalar* systems with $p = 1$.

$y = h(u)$ where $u\,h(u) \geq 0$ is:

$$\int_0^u h(\sigma)\, d\sigma \ \geq \ 0, \qquad\qquad u \in \mathbb{R}$$

[1] All results extend to a time-varying system $y = h(t, u)$ where $h : \mathbb{R} \times \mathbb{R}^p \to \mathbb{R}^p$.

We will use this property below in constructing what are called storage functions.

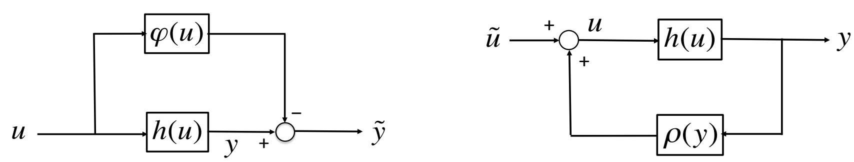
Suppose a memoryless system $y = h(u)$ satisfies

$$u^T y \quad \geq \quad u^T \varphi(u) \qquad \forall u \in \mathbb{R}^p$$

for some function φ, i.e., the input-output product is greater than a function that depends only on the input u. Then

$$u^T \tilde{y} \quad := \quad u^T(y - \varphi(u)) \quad \geq \quad 0 \qquad \forall u \in \mathbb{R}^p$$

i.e., by subtracting from the output y the feedforward term $\varphi(u)$, the new system $\tilde{y} := h(u) - \varphi(u)$ is passive; see Figure 4.2(a). Note that neither $h(u)$ nor $\varphi(u)$ need to be passive



(a) Input-feedforward passive (b) Output-feedback passive

Figure 4.2: (a) Input-feedforward passive system $u^T \tilde{y} \geq 0$. (b) Output-feedback passive system $\tilde{u}^T y \geq 0$.

themselves. If $u^T \varphi(u) > 0$ then there is "excess" passivity; otherwise there is "shortage" of passivity. Hence a system h that satisfies $u^T h(u) \geq u^T \varphi(u)$ can be transformed into a passive system through input feedforward. Such a system h is called *input-feedforward passive*. If $u^T \varphi(u) > 0$ for all $u \neq 0$ then h is called *input strictly passive* because passivity is strict in the sense that $u^T y \geq u^T \varphi(u) = 0$ only if $u = 0$.

Similarly suppose a system h satisfies

$$u^T y \quad \geq \quad y^T \rho(y), \qquad \forall u \in \mathbb{R}^p$$

for some function ρ, i.e., the input-output product is greater than a function that depends only on the output y. Then

$$\tilde{u}^T y \quad := \quad (u - \rho(y))^T y \quad \geq \quad 0, \qquad \forall u \in \mathbb{R}^p$$

i.e., by subtracting from the input u the feedback term $\rho(y)$ to obtain the new input $\tilde{u} := u - \rho(y)$, the system is made passive; see Figure 4.2(b). Such a system h is called *output-feedback passive*. If $y^T \rho(y) > 0$ for all $y \neq 0$ then h is called *output strictly passive*.

Dynamical systems. Consider:

$$\dot{x} \quad = \quad f(x, u) \tag{4.1a}$$
$$y \quad = \quad h(x, u) \tag{4.1b}$$

where $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ is locally Lipschitz or $f(x,u) := \left(\tilde{f}(x,u)\right)_x^+$ with a locally Lipschitz $\tilde{f} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$, $h : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$ is continuous, $f(0,0) = 0$, and $h(0,0) = 0$. Hence the origin is an equilibrium point without input ($u \equiv 0$). The system has the same number p of inputs and outputs.

Definition 4.1 The system (4.1) is *passive* if there exists a continuously differentiable positive semidefinite function $V(x)$ such that

$$u^T y \quad \geq \quad \dot{V} := \frac{\partial V}{\partial x}(x)\, f(x,u) \qquad \forall (x,u) \in \mathbb{R}^n \times \mathbb{R}^p \tag{4.2}$$

Moreover it is said to be

1. *lossless* if $u^T y = \dot{V}$.
2. *input-feedforward passive* if $u^T y \geq \dot{V} + u^T \varphi(u)$ for some function φ.
3. *input strictly passive* if $u^T y \geq \dot{V} + u^T \varphi(u)$ and $u^T \varphi(u) > 0$ for all $u \neq 0$.
4. *output-feedback passive* if $u^T y \geq \dot{V} + y^T \rho(y)$ for some function ρ.
5. *output strictly passive* if $u^T y \geq \dot{V} + y^T \rho(y)$ and $y^T \rho(y) > 0$ for all $y \neq 0$.
6. *strictly passive* if $u^T y \geq \dot{V} + \psi(x)$ for some positive definite function ψ.

In all cases the inequalities should hold for all $(x,u) \in \mathbb{R}^{n+p}$. The function V is called a *storage function*.

These definitions reduce to those for memoryless systems $y = h(u)$ with $V(x) \equiv 0$. We emphasize that the inequalities in Definition 4.1 need to hold not just on the solution-control trajectories of (4.1), but on all $(x,u) \in \mathbb{R}^{n+p}$. Even though $V(x)$ is a function of the state x, its rate $\dot{V}(x,u)$ is a function of both state x and input u.

To simplify notation we often write $y, V, \dot{V}$ for $y(x,u), V(x), \dot{V}(x,u)$ or for $y(x(t),u(t)), V(x(t)), \dot{V}(x(t),u(t))$ when there is no risk of confusion. We also write $x \equiv a$ or $x(t) \equiv a$ for $x(t) = a$, $\forall t \geq 0$; similarly for other functions such as u, h, y, etc.

Example 4.2 In this example we illustrate different passivity concepts.

1. *Integrator.* Consider

$$\dot{x} \;=\; u, \qquad y \;=\; x$$

To show that an integrator is passive consider the candidate storage function

$$V(x) \quad := \quad \frac{1}{2}x^2$$

Clearly V is continuously differentiable and positive semidefinite $V(x) \geq 0$. Furthermore

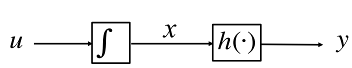
$$\dot{V} \quad = \quad \frac{\partial V}{\partial x}\,\dot{x} \quad = \quad x\,\dot{x} \quad = \quad uy$$

Hence an integrator is lossless.

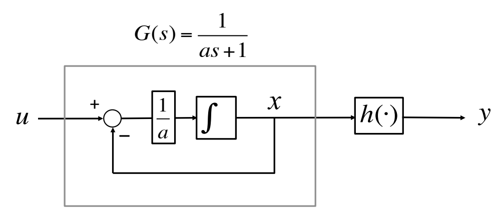
2. *Integrator in cascade with a memoryless passive system.* Consider

$$\dot{x} \;=\; u, \qquad y \;=\; h(x)$$

where h is continuous and passive. See Figure 4.3(a). Consider the candidate storage



(a) Integrator + memoryless



(b) First-order + memoryless

Figure 4.3: (a) Integrator in cascade with a memoryless passive system is passive with the storage function $V(x) := \int_0^x h(\sigma)d\sigma$. (b) First-order system in cascade with a memoryless passive system is passive with the storage function $V(x) := \int_0^x ah(\sigma)d\sigma$.

function

$$V(x) \quad := \quad \int_0^x h(\sigma)\,d\sigma$$

Since h is continuous, V is continuously differentiable. Since h is passive $xh(x) \geq 0$, $\int_0^x h(\sigma)d\sigma \geq 0$, i.e. V is positive semidefinite. Furthermore

$$\dot{V} \quad = \quad \frac{\partial V}{\partial x}\,\dot{x} \quad = \quad h(x)\,\dot{x} \quad = \quad uy$$

i.e., the system is lossless.

3. *First-order system in cascade with a memoryless passive system.* Suppose the integrator is replaced by the transfer function (see Chapter 6.2.1 for more discussion of transfer functions of linear time-invariant systems)

$$G(s) \;=\; \frac{1}{as+1}, \qquad a > 0$$

This transfer function $G(s)$ can be implemented by a first-order system as shown in Figure 4.3(b):

$$a\dot{x} \;=\; -x + u, \qquad y \;=\; h(x)$$

Consider the candidate storage function

$$V(x) \;:=\; \int_0^x a\,h(\sigma)\,d\sigma$$

which is continuously differentiable and positive semidefinite since h is continuous and passive. Furthermore

$$\dot{V} \;=\; \frac{\partial V}{\partial x}\,\dot{x} \;=\; h(x)\,a\dot{x} \;=\; h(x)\,(-x+u) \;=\; uy \;-\; xh(x)$$

i.e., $uy = \dot{V} + xh(x)$. Since h is passive, $xh(x) \geq 0$ and hence the system is passive. If $xh(x) > 0$ for $x \neq 0$ then the system is strictly passive.

□

Example 4.3 Input-feedforward and output-feedback passivity. Like the memoryless systems in Figure 4.3, the dynamical systems with memoryless input feedforward or output feedback as shown in Figure 4.4 are passive if the memoryless function h is passive. Indeed the system in Figure 4.4(a) is

$$\dot{x} \;=\; u, \qquad y \;=\; x + h(u)$$

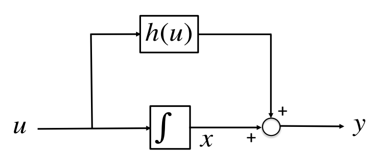
Consider the continuously differentiable positive definite function

$$V(x) \;:=\; \frac{1}{2}x^2$$

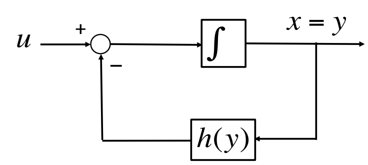
We have

$$\dot{V} \;=\; x\,\dot{x} \;=\; u(y - h(u))$$

Hence $uy = \dot{V} + uh(u)$. Therefore the system is input-feedforward passive if h is passive (i.e., $uh(u) \geq 0$) and the system is input strictly passive if $uh(u) > 0$ for $u \neq 0$.



(a) Integrator + memoryless input feedforward



(b) Integrator + memoryless output feedback

Figure 4.4: (a) Integrator with a memoryless passive input-feedforward h is passive with the storage function $V(x) := \frac{1}{2}x^2$. (b) Integrator with a memoryless passive output-feedback h is passive with the storage function $V(x) := \frac{1}{2}x^2$.

The system in Figure 4.4(b) is

$$\dot{x} \;=\; u - h(y), \qquad y \;=\; x$$

Use the same storage function V as above we have

$$\dot{V} \;=\; x\,\dot{x} \;=\; y(u - h(y))$$

Hence $uy = \dot{V} + yh(y)$. Therefore the system is output-feedback passive if h is passive (i.e., $yh(y) \geq 0$) and the system is output strictly passive if $yh(y) > 0$ for $y \neq 0$. □

The following properties of storage functions are key for proving Lyapunov stability below as they pertain to the properties C3.1, C3.2, C3.2', C3.2" of a Lyapunov function. The system (4.1) is called *zero-state observable* if no solution of (4.1) can stay identically in $\{x \,|\, y = h(x; 0) = 0\}$ except $x(t) \equiv 0$, i.e., if the state x must stay identically zero when both input u and output y are identically zero.

Lemma 4.4

1. *If the system* (4.1) *is strictly passive, i.e.,* $u^T y \geq \dot{V} + \psi(x)$*, with a storage function* V *and a positive definite function* ψ*, then we have without loss of generality*

$$\begin{aligned} V(x) &> 0 \quad \textit{for all } x \neq 0 \qquad \textit{and} \qquad V(0) = 0 \\ \dot{V}(x,0) &< 0 \quad \textit{for all } x \neq 0 \end{aligned}$$

   If ψ *is continuous then, for all* $0 < \delta < \epsilon$*, there exists* $\alpha > 0$ *such that*

$$\dot{V}(x,0) \;\leq\; -\alpha \;<\; 0 \qquad \forall x \textit{ with } \delta \leq \|x\| \leq \epsilon$$

2. *If the system* (4.1) *is output strictly passive and zero-state observable with a storage function* V *then we have without loss of generality*

$$\begin{aligned} V(x) &> 0 \quad \textit{for all } x \neq 0 \qquad \textit{and} \qquad V(0) = 0 \\ \dot{V}(x,0) &\leq 0 \quad \textit{for all } x \end{aligned}$$

   Moreover the only solution of $\dot{x} = f(x(t), 0)$ *with zero input* $u \equiv 0$ *that can stay identically in the set* $E := \{x \,|\, \dot{V}(x) = 0\}$ *is the trivial solution* $x \equiv 0$*.*

Proof. Suppose the system (4.1) is strictly passive, i.e.,

$$u^T h(x,u) \;\geq\; \dot{V}(x,u) + \psi(x) \qquad \text{for all } (x,u) \in \mathbb{R}^{n+p}$$

for some positive semidefinite storage function $V(x)$ and positive definite function $\psi(x)$. Take $u \equiv 0$ and we have

$$\dot{V}(x,0) \quad \leq \quad -\psi(x) \tag{4.3}$$

Since $\psi(x) > 0$ for all $x \neq 0$, (4.3) implies the second assertion:

$$\dot{V}(x,0) \quad < \quad 0 \qquad \text{for all } x \neq 0$$

For the third assertion note that

$$\sup_{x:\delta \leq \|x\| \leq \epsilon} \dot{V}(x,0) \quad \leq \quad - \inf_{x:\delta \leq \|x\| \leq \epsilon} \psi(x) \quad =: \quad -\alpha$$

Since $\psi(x)$ is continuous, α is attained and is strictly positive, proving the third assertion.

We now use (4.3) to prove $V(x) > 0$ for all $x \neq 0$. Since f is locally Lipschitz or a projection of a locally Lipschitz function $\tilde{f}$, for any $x_0 \in \mathbb{R}^n$ the autonomous system $\dot{x} = f(x,u)$ with zero input $u \equiv 0$ has a solution, denoted by $\phi(t; x_0)$ at time t, starting from x_0 at time 0, over some interval $[0, \delta]$. Then (4.3) implies[2]

$$V(\phi(\tau; x_0)) - V(x_0) \quad \leq \quad -\int_0^\tau \psi(\phi(t; x_0))\, dt \qquad \forall \tau \in [0, \delta]$$

Since $V(\phi(\tau; x_0)) \geq 0$ and ψ is positive definite we have for any $x_0 \in \mathbb{R}^n$

$$V(x_0) \quad \geq \quad \int_0^\tau \psi(\phi(t; x_0))\, dt \quad \geq \quad 0 \qquad \forall \tau \in [0, \delta] \tag{4.4}$$

Consider any $x_0 \neq 0$. Suppose $V(x_0) = 0$. Then (4.4) implies

$$\begin{array}{rrcll} & \psi(\phi(t; x_0)) & \equiv & 0 & \text{(for all } t \in [0, \delta]) \\ \Longrightarrow & \phi(t; x_0) & \equiv & 0 & (\psi \text{ is positive definite}) \\ \Longrightarrow & x_0 & = & 0 & (\phi(0; x_0) = x_0) \end{array}$$

contradicting $x_0 \neq 0$. Hence $V(x_0) > 0$ for every $x_0 \neq 0$. Finally we must have $V(0) \geq 0$ by the continuity of V. Indeed we can assume without loss of generality that $V(0) = 0$. Otherwise replace $V(x)$ by $\tilde{V}(x) := V(x) - V(0)$. The above argument applies to the storage function $\tilde{V}$ with $\tilde{V}(0) = 0$. This completes the proof of part 1.

For part 2, suppose (4.1) is output strictly passive and zero-state observable. We have $u^T y \geq \dot{V} + y^T \rho(y)$ for some positive semidefinite storage function V and some function ρ such that $y^T \rho(y) > 0$ for all $y \neq 0$. With $u \equiv 0$ we have the second assertion:

$$\dot{V}(x,0) \quad \leq \quad -h^T(x,0)\rho(h(x,0)) \quad \leq \quad 0 \qquad \forall x \in \mathbb{R}^n$$

[2] In the following we ignore some minor measurability issues if ψ or ρ is not continuous.

The proof of $V(x) > 0$ for $x \neq 0$ follows a similar argument (substituting $y(t) \equiv h(\phi(t;x_0),0)$ with $u \equiv 0$):

$$V(\phi(\tau;x_0)) - V(x_0) \quad \leq \quad -\int_0^\tau h^T(\phi(t;x_0),0)\,\rho(h(\phi(t;x_0),0))\,dt \qquad \forall \tau \in [0,\delta]$$

and hence for all $x_0 \in \mathbb{R}^n$

$$V(x_0) \quad \geq \quad \int_0^\tau h^T(\phi(t;x_0),0)\,\rho(h(\phi(t;x_0),0))\,dt \quad \geq \quad 0 \qquad \forall \tau \in [0,\delta]$$

Consider any $x_0 \neq 0$. Suppose $V(x_0) = 0$. Then for $t \in [0,\delta]$ we have

$$h^T(\phi(t;x_0),0)\,\rho(h(\phi(t;x_0),0)) \;\equiv\; 0 \quad \Rightarrow \quad h(\phi(t;x_0),0) \;\equiv\; 0 \quad \Rightarrow \quad \phi(t;x_0) \;\equiv\; 0$$

where the last implication follows from zero-state observability. Hence $x_0 = \phi(0;x_0) = 0$. This is a contradiction and therefore $V(x_0) > 0$ for any $x_0 \neq 0$. As in part 1 we can assume without loss of generality that $V(0) = 0$.

Finally consider any solution $x(t)$ of $\dot{x} = f(x,0)$ with zero input $u \equiv 0$ that stays identically in the set $E := \{x \,|\, \dot{V}(x) = 0\}$. Since

$$\dot{V}(x,0) \quad \leq \quad -y^T\rho(y) \quad \leq \quad 0 \qquad \forall x \text{ (and } y = h(x,0))$$

$\dot{V}(x(t),0) \equiv 0$ implies $y^T(t)\rho(y(t)) \equiv 0$. But $y^T\rho(y) > 0$ for $y \neq 0$, and hence $y(t) \equiv 0$. Zero-state observability then implies that $x(t) \equiv 0$.

This completes the proof of the lemma. $\square$

Passivity ensures stability and strict passivity ensures asymptotic stability. Consider the dynamical system (4.1) reproduced here for convenience:

$$\dot{x} \;=\; f(x,u) \tag{4.5a}$$
$$y \;=\; h(x,u) \tag{4.5b}$$

where $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$ is locally Lipschitz or $f(x,u) := \left(\tilde{f}(x,u)\right)^+_{x+c}$, $c \in \mathbb{R}$, with a locally Lipschitz $\tilde{f} : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^n$, $h : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$ is continuous. We assume $f(0,0) = 0$, and $h(0,0) = 0$, i.e., the origin $x = 0$ is an equilibrium at zero input $u \equiv 0$.

Theorem 4.5 *Consider the system (4.5). The origin of $\dot{x} = f(x,0)$ is*

1. *stable if the system is passive with a positive* definite *storage function V;*
2. *asymptotically stable if the system is*
   - *strictly passive $u^T y \geq \dot{V} + \psi(x)$ with a continuous positive definite function ψ, or*

- *output strictly passive and zero-state observable, i.e.,* $y(t) \equiv h(x(t), 0) \equiv 0 \Rightarrow x(t) \equiv 0$*;*

3. *globally asymptotically stable if the system is*

   - *strictly passive* $u^T y \geq \dot{V} + \psi(x)$ *with a continuous positive definite function* ψ*, or*
   - *output strictly passive and zero-state observable, i.e.,* $y(t) \equiv h(x(t), 0) \equiv 0 \Rightarrow x(t) \equiv 0$*,*

   and the storage function is radially unbounded.

Proof.

1. Let V be the (continuously differentiable) storage function of the system (4.5) with $u^T y(x, u) \geq \dot{V}(x, u)$ for all $(x, u) \in \mathbb{R}^{n+p}$. By assumption V is positive *definite*. Hence V is a Lyapunov function for the zero-input system ($u \equiv 0$) as it satisfies conditions C3.1 ($V(x) > 0$, for all $x \neq 0$) and C3.2 ($\dot{V}(x, 0) \leq 0$).

2. Lemma 4.4.1 shows that if the system (4.5) is strictly passive with a continuous ψ then its storage function V satisfies conditions C3.1, C3.2' and C3.2". V is therefore a Lyapunov function. The claim then follows from Theorem 3.3.2 for locally Lipschitz f or from Corollary 3.5.2 for projected dynamics.

   Lemma 4.4.2 shows that if the system (4.5) is output strictly passive and zero-state observable then its storage function V satisfies conditions C3.1 and C3.2 (but not necessarily C3.2' or C3.2") and is therefore a Lyapunov function. LaSalle's invariance principle then implies that the solution x of $\dot{x} = f(x, 0)$ will converge to the largest invariance set in $\{x \mid \dot{V}(x) = 0\}$. But the only solution of $\dot{x} = f(x, 0)$ that can stay identically in the set $\{x \mid \dot{V}(x) = 0\}$ is the trivial solution $x \equiv 0$ by Lemma 4.4.2. Hence (4.5) is asymptotically stable by Theorem 3.9.1.

3. When V is radially unbounded Theorem 3.9.2 implies that the system (4.5) is globally asymptotically stable.

□

4.2 FEEDBACK SYSTEMS

A main advantage of the passivity method is the important property that feedback connection preserves passivity. It allows us to analyze the stability of a feedback system through the passivity analysis of its open-loop components. We explain this property in this section.

Consider the feedback connection in Figure 4.5 where each of the components H_1 and H_2 is either a time-invariant dynamical system represented by the state space model

$$\dot{x}_i = f_i(x_i, e_i), \quad y_i = h_i(x_i, e_i)$$

or a (possibly time-varying) memoryless function represented by

$$y_i = h_i(t, e_i)$$

where

$$e_1 := u_1 - y_2, \qquad e_2 := u_2 + y_1$$

Suppose H_1 and H_2 are dynamical systems. If the set of equations

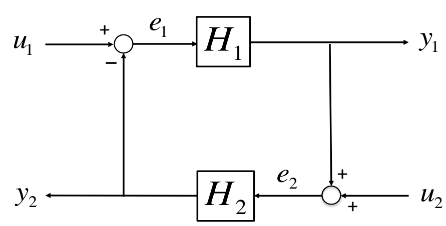


Figure 4.5: Feedback connection.

$$e_1 = u_1 - h_2(x_2, e_2) \tag{4.6a}$$
$$e_2 = u_2 + h_1(x_1, e_1) \tag{4.6b}$$

has a unique solution for (e_1, e_2) for every (x_1, x_2, u_1, u_2) then we can write (e_1, e_2) in terms of (x_1, x_2, u_1, u_2) and represent the feedback system by a state-space model of the form

$$\dot{x} = f(x, u), \quad y = h(x, u) \tag{4.7}$$

for some functions f, h where

$$x := \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad u := \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad y := \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

We assume for the closed-loop system (4.7) that the function f is locally Lipschitz or $f(x, u) = \left(\tilde{f}(x, u)\right)^+_{x+c}$, $c \in \mathbb{R}$, with a locally Lipschitz $\tilde{f}$, h is continuous, $f(0, 0) = 0$ and $h(0, 0) = 0$.

The functions f, h describing the closed-loop system (4.7) may not be easy to derive explicitly from the open-loop components f_i, h_i, $i = 1, 2$. If however h_1 is independent of e_1 or h_2 is independent of e_2 then it is easy to write down explicitly the unique solution for (e_1, e_2) in terms of (x_1, x_2, u_1, u_2). In this case the functions f and h of the closed-loop system are locally Lipschitz if f_i and h are.

If one component, say H_1, is a time-invariant dynamical system and the other H_2 is a time-varying memoryless function:

$$\dot{x}_1 \;=\; f_1(x_1, e_1), \quad y_1 \;=\; h_1(x_1, e_1), \quad y_2 \;=\; h_2(t, e_2)$$

then the closed-loop state-space model takes the form

$$\dot{x} \;=\; f(t, x, u), \quad y = h(t, x, u) \tag{4.8}$$

for some functions f, h where

$$x \;:=\; x_1, \quad u \;:=\; \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad y \;:=\; \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

provided the system of equations

$$\begin{aligned} e_1 \;&=\; u_1 - h_2(t, e_2) \\ e_2 \;&=\; u_2 + h_1(x_1, e_1) \end{aligned}$$

has a unique solution for (e_1, e_2) for every (x_1, t, u_1, u_2). Even though the open-loop component f_1 is time-invariant the closed-loop system (4.8) becomes time-varying if the memoryless system h_2 is time-varying. We assume for the closed-loop system (4.8) that the function f is piecewise continuous in t and locally Lipschitz in (x, u), or $f(t, x, u) = \left(\tilde{f}(t, x, u)\right)^{+}_{x+c}$, $c \in \mathbb{R}$, where $\tilde{f}$ is piecewise continuous in t and locally Lipschitz in (x, u). We also assume h is piecewise continuous in t and continuous in (x, u), $f(t, 0, 0) = 0$ and $h(t, 0, 0) = 0$.

Again the functions f, h describing the closed-loop system (4.8) may not be easy to derive explicitly from the open-loop components f_1, h_1, h_2, unless h_1 is independent of e_1 or h_2 is independent of e_2.

The next result says that if both open-loop systems H_i are passive with storage functions V_i then the closed-loop system is passive with the storage function $V_1 + V_2$. Furthermore if H_i are stable with V_i as their Lyapunov functions, then the closed-loop system is stable with $V_1 + V_2$ as its Lyapunov function. Note that a positive definite storage function of a passivity system implies stability of its equilibrium by Theorem 4.5.1.

Theorem 4.6 *Consider the feedback connection shown in Figure 4.5 and suppose the origin is an equilibrium of the closed-loop system* (4.7) *(or* (4.8)*) when input* $u \equiv 0$*, i.e.,* $f(0, 0) = 0$ *(or* $f(t, 0, 0) = 0$*).*

1. *If both open-loop components H_1 and H_2 are passive then the feedback connection from (u_1, u_2) to (y_1, y_2) is passive.*

2. *Furthermore, if the storage functions of H_1 and H_2 are positive definite, then the origin is stable.*

Proof. Since H_i are passive we have

$$e_i^T y_i \quad \geq \quad \dot{V}_i, \qquad i = 1, 2$$

for some storage functions V_i that are continuously differentiable and positive semidefinite (V_i is taken to be zero for a memoryless system). Then $V := V_1 + V_2$ is continuously differentiable and positive semidefinite. Moreover

$$\dot{V}_1 + \dot{V}_2 \quad \leq \quad \sum_i e_i^T y_i \quad = \quad (u_1 - y_2)^T y_1 \; + \; (u_2 + y_1)^T y_2 \quad = \quad u_1^T y_1 + u_2^T y_2$$

Hence V is a storage function for the feedback system, proving the passivity of the feedback connection. If V_i are positive definite then V is positive definite and Theorem 4.5.1 implies the stability of the origin. □

Passivity of components H_1 and H_2 with positive definite storage functions ensures the passivity and stability of the feedback connection. For asymptotic stability we consider first the case where both components are time-invariant dynamical systems and then the case where one of the components is a dynamical system and the other a memoryless function. The first case will be useful for proving the asymptotic stability of primal-dual algorithms where both the sources and the links have dynamics. The second case will be useful for primal or dual algorithms where there is dynamics in the sources or the links, but not both.

Consider the feedback connection of two time-invariant dynamical components H_i:

$$\dot{x}_i \quad = \quad f_i(x_i, e_i), \qquad i = 1, 2 \tag{4.9a}$$
$$y_i \quad = \quad h_i(x_i, e_i), \qquad i = 1, 2 \tag{4.9b}$$

whose closed-loop system is described by (4.7). The next result extends Theorem 4.5 from open-loop systems to closed-loop systems.

Theorem 4.7 *Consider the feedback system consisting of open-loop components specified by* (4.9) *with a closed-loop description* (4.7)*. The origin of the closed-loop system (when $u \equiv 0$) is asymptotically stable if any one of the following holds:*

1. *both open-loop components H_1 and H_2 are strictly passive so that $e_i^T y_i \geq \dot{V}_i + \psi_i(x_i)$, $i = 1, 2$, with continuous positive definite functions ψ_i.*

2. *both open-loop components H_1 and H_2 are output strictly passive and zero-state observable ($y_i(t) \equiv h(x_i(t), 0) \equiv 0 \Rightarrow x_i(t) \equiv 0$).*

3. *one of the feedback components is strictly passive and the other is output strictly passive and zero-state observable.*

Furthermore if the storage function of each component is radially unbounded then the origin is globally asymptotically stable.

Proof. The proof follows a similar argument as in the proof of Theorem 4.5, using the function $V(x) := V_1(x_1) + V_2(x_2)$ as the candidate Lyapunov function for the closed-loop system (4.7) where $V_i(x_i)$ are storage functions of components H_i.

When both H_i are strictly passive with storage functions V_i, Lemma 4.4.1 implies that they satisfy the following conditions for Lyapunov function

$$\begin{array}{rrcll} \text{C3.1:} & V_i(x_i) & > & 0 \quad \text{for all } x_i \neq 0 & \quad \text{and} \quad V_i(0) = 0 \\ \text{C3.2':} & \dot{V}_i(x_i) & < & 0 \quad \text{for all } x_i \neq 0 & \end{array}$$

This implies that $V(x) := V_1(x_1) + V_2(x_2)$ for the closed-loop system (4.7) also satisfies conditions C3.1 and C3.2'. Theorem 3.3.2 then implies that the origin of the closed-loop system is asymptotically stable if f in (4.7) is locally Lipschitz. If f is a projection of a locally Lipschitz $\tilde{f}$ then, since ψ_i are continuous, Lemma 4.4.1 implies that V satisfies C3.1 and C3.2". Corollary 3.5.2 then implies the asymptotic stability of the origin for the closed-loop system.

When both H_i are output strictly passive and zero-state observable with storage functions V_i, Lemma 4.4.2 implies that they satisfy the following conditions for Lyapunov function

$$\begin{array}{rrcll} \text{C3.1:} & V_i(x_i) & > & 0 \quad \text{for all } x_i \neq 0 & \quad \text{and} \quad V_i(0) = 0 \\ \text{C3.2:} & \dot{V}_i(x_i) & \leq & 0 \quad \text{for all } x_i & \end{array}$$

This implies that $V(x) := V_1(x_1) + V_2(x_2)$ also satisfies conditions C3.1 and C3.2. LaSalle's invariance principle then implies that the solution x of $\dot{x} = f(x, 0)$ converges to the largest invariance set in $E := \{x \mid \dot{V}(x) = 0\}$. We now argue that the only solution that can stay identically in E is the equilibrium solution $x \equiv 0$. Since H_i are output strictly passive we have $e_i^T y_i \geq \dot{V}_i(x_i, e_i) + y_i^T \rho_i(y_i)$ for some function ρ_i such that $y_i^T \rho_i(y_i) > 0$ for $y_i \neq 0$. Taking $u_i \equiv 0$ we have $e_1 = -y_2$, $e_2 = y_1$, and hence

$$\dot{V}(x, e) \;:=\; \dot{V}_1(x_1, e_1) + \dot{V}_2(x_2, e_2) \;\leq\; -y_1^T \rho_1(y_1) - y_2^T \rho_2(y_2) \;\leq\; 0$$

where $y_i = h_i(x_i, e_i)$. Hence the only solutions to $\dot{x} = f(x, 0)$ that can stay identically in E are those such that the resulting $y_i \equiv h_i(x_i, e_i) \equiv 0$. But $y_1 \equiv 0$ implies $e_2 \equiv 0$ and $y_2 \equiv$

0 implies $e_1 \equiv 0$ (since $u_i \equiv 0$). Thus $y_i \equiv h_i(x_i, 0) \equiv 0$ implies $x_i \equiv 0$ by the zero-state observability of H_i. Theorem 3.9.1 then implies that the origin of the closed-loop system (4.7) is asymptotically stable.

In the third case where H_1 is strictly passive and H_2 is output strictly passive we have (taking $u_i \equiv 0$)

$$\dot{V}(x) \quad \leq \quad -\psi_1(x_1) - y_2\rho_2(y_2) \quad \leq \quad 0 \quad \text{for all } x$$

for some positive definite function ψ_1 and some function ρ_2 such that $y_2^T\rho_2(y_2) > 0$ for all $y_2 \neq 0$. A similar argument as in the proof of Lemma 4.4 shows that $V(0) = 0$ and $V(x) > 0$ for all $x \neq 0$. Hence LaSalle's invariance principle implies that the solution x of $\dot{x} = f(x, 0)$ will converge to the largest invariance set in the set $E := \{x \,|\, \dot{V}(x) = 0\}$. We now argue that the only solution that can stay identically in E is the equilibrium solution $x \equiv 0$. Since ψ_1 is positive definite and $y_2^T\rho_2(y_2) > 0$ for all $y_2 \neq 0$, $\dot{V}(x) = 0$ implies $x_1 \equiv 0$ and $y_2 \equiv 0$. Note that $y_2 \equiv 0$ implies $e_1 \equiv 0$ as $u_1 \equiv 0$. Since $h_1(0, 0) = 0$ by assumption, this means that $y_1 \equiv h_1(x_1 \equiv 0, e_1 \equiv 0) \equiv 0$. Hence $e_2 \equiv 0$. Then $y_2 \equiv 0$ and zero-state observability of H_2 implies $x_2 \equiv 0$. Hence $x \equiv 0$ and the origin of the closed-loop system (4.7) is asymptotically stable.

Finally if $V_i(x_i)$ are radially unbounded so is $V(x)$. The origin is globally asymptotically stable by Theorem 3.9.2. □

Consider next the feedback connection of a dynamical system H_1:

$$\dot{x}_1 = f_1(x_1, e_1) \tag{4.10a}$$
$$y_1 = h_1(x_1, e_1) \tag{4.10b}$$

and a *time-invariant* memoryless system H_2:

$$y_2 = h_2(e_2) \tag{4.10c}$$

whose closed-loop system is described by (the time-invariant version of) (4.8).

Theorem 4.8 *Consider the feedback system consisting of open-loop components specified by* (4.10) *with a closed-loop description* (4.8). *Suppose*

- H_1 *is output-feedback passive, i.e.,*

$$e_1^T y_1 \quad \geq \quad \dot{V}_1 + y_1^T\rho_1(y_1) \tag{4.11}$$

for some ρ_1 *and a positive definite storage function* V_1, H_1 *is zero-state observable (i.e.,* $y_1 \equiv h_1(x_1, 0) \equiv 0 \Rightarrow x_1 \equiv 0$*); and*

- *H_2 is input-feedforward passive, i.e.,*

$$e_2^T y_2 \quad \geq \quad e_2^T \varphi_2(e_2) \tag{4.12}$$

for some φ_2.

Then the origin of the closed-loop system (when $u \equiv 0$) is asymptotically stable if

$$v^T[\rho_1(v) + \varphi_2(v)] \quad > \quad 0 \qquad \text{for all } v \neq 0 \tag{4.13}$$

Furthermore, if V_1 is radially unbounded, then the origin is globally asymptotically stable.

Even though, being memoryless, $y_2 = h_2(e_2)$ is trivially input-feedforward passive (by setting $\varphi_2 = h_2$), h_2 may not satisfy (4.13) but there may be some other function φ_2 that satisfies both (4.12) and (4.13).

Proof. Use the storage function of H_1 as a Lyapunov function candidate. Since $u \equiv 0$ we have $e_1 \equiv -y_2$ and $y_1 \equiv e_2$. Substituting into (4.11) yields

$$\dot{V}_1 \quad \leq \quad e_1^T y_1 - y_1^T \rho_1(y_1) \quad = \quad -e_2^T y_2 - y_1^T \rho_1(y_1)$$

Combining with (4.12) we have

$$\dot{V}_1 \quad \leq \quad -y_1^T(\varphi_2(y_1) + \rho_1(y_1)) \quad < \quad 0 \qquad \forall y_1 \neq 0$$

where the last inequality follows from (4.13). Hence LaSalle's invariance principle implies that the trajectory will converge to the largest invariant set in $E := \{x_1 \,|\, \dot{V}(x_1) = 0\}$. When $\dot{V}(x_1) \equiv 0$ we have $y_1 \equiv 0$, implying $e_2 \equiv 0$ and hence $y_2 \equiv h_2(0) \equiv 0$. This means $e_1 \equiv 0$ and hence zero-state observability of H_1 implies $x_1 \equiv 0$ since $y_1 \equiv h_1(x_1, 0) \equiv 0$. Asymptotic stability of the origin then follows from Theorem 3.9.1. Global asymptotic stability when V_1 is radially unbounded follows from Theorem 3.9.2. □

4.3 STABILITY OF PRIMAL ALGORITHMS

Consider the set of primal algorithms

$$\begin{aligned} \dot{x}_i &= \kappa_i \left(U_i'(x_i(t)) - q_i(t)\right)^+_{x_i(t)}, \qquad i = 1, \ldots, N \\ p_l(t) &= p_l(y_l(t)), \qquad l = 1, \ldots, L \end{aligned}$$

where $p_l : \mathbb{R}_+ \to \mathbb{R}_+$ is a function that models AQM at link l. Hence there is dynamics in the sources but not in the links. Here

$$q_i(t) \;:=\; \sum_l R_{li}\, p_l(t) \quad \text{and} \quad y_l(t) \;:=\; \sum_i R_{li}\, x_i(t)$$

In vector form they are

$$\begin{aligned} \dot{x} &= K\,(U'(x(t)) - q(t))^{+}_{x(t)} && (4.14a)\\ p(t) &= p(y(t)) && (4.14b)\\ q(t) &= R^T p(t), \quad y(t) \;=\; R\,x(t) && (4.14c) \end{aligned}$$

where $K := \mathrm{diag}(\kappa_i, i = 1, \ldots, N)$. Since $p(t)$, and hence $q(t) = R^T p(t)$, are statically determined by $x(t)$ through (4.14b)(4.14c) we may write $p(x(t)) = p\,(Rx(t))$ and $q(x(t))$ in place of $p(t)$ and $q(t)$ respectively. We call x^*, as opposed to (x^*, p^*), an equilibrium of the primal algorithm (4.14) if $(U'(x^*) - q(x^*))^{+}_{x^*} = 0$ which is equivalent to:

$$U_i'(x_i^*) \;\le\; q_i(x^*) \quad \text{with equality if } \; x_i^* > 0, \qquad i = 1, \ldots, N$$

We will assume that conditions C2.1 and C2.2 hold. Then (4.14) has a unique solution $(x(t), t \ge 0)$ for every initial point $x(0) \ge 0$; see Remark 3.17 in Chapter 3.2. Moreover a unique equilibrium point x^* exists since U_i are strictly concave. The equilibrium link prices $p(Rx^*)$ are also unique since they are determined by x^* (regardless of the row rank of the routing matrix R). Finally $(x^*, p(x^*))$ is primal-dual optimal for a network utility maximization problem and its dual (see Chapter 2.2). We now use Theorem 4.8 to study the asymptotic stability of x^*.

For simplicity we assume $U_i'(x_i) \to \infty$ as $x_i \to 0$ (e.g., $U_i(x_i) = \log x_i$) so that the projection is never active, i.e., $(U'(x) - q(x)))^{+}_{x} = U'(x) - q(x)$. The result can be easily extended to the general case with projection; see Chapter 4.4.

It is convenient to shift the equilibrium point of the primal algorithm to the origin. Define the perturbed variables (dropping t to simplify notation)

$$\tilde{x} \;:=\; x - x^* \quad\Rightarrow\quad x \;=\; x^* + \tilde{x}$$

Then

$$\tilde{y} \;:=\; R\,\tilde{x} \;=\; R(x - x^*) \;=\; y - y^*$$

and

$$\begin{aligned} \tilde{p} &:= p(y) - p(y^*) \;=\; p(y^* + \tilde{y}) - p(y^*) \quad =: \; \tilde{p}(\tilde{y})\\ \tilde{q} &:= R^T \tilde{p} \;=\; R^T (p(y) - p(y^*)) \;=\; q - q^* \end{aligned}$$

Then the perturbed source rates satisfy (since $U'(x^*) = q^*$)

$$\begin{aligned} \dot{\tilde{x}} \;=\; \dot{x} &= K\,(U'(x^* + \tilde{x}) \;-\; q)\\ &= K(\,\underbrace{U'(x^* + \tilde{x}) - U'(x^*)}_{\tilde{U}'(\tilde{x})} \;-\; \underbrace{(q - q^*)}_{\tilde{q}}\,) \quad =: \; K\left(\tilde{U}'(\tilde{x}) - \tilde{q}\right) \end{aligned}$$

We will prove that the origin of the following perturbed system is globally asymptotically stable:

$$\dot{\tilde{x}} = K\left(\tilde{U}'(\tilde{x}) - \tilde{q}\right) =: f(\tilde{x}, -\tilde{q}) \quad (4.15a)$$

$$\tilde{p} = \tilde{p}(\tilde{y}) \quad (4.15b)$$

where $\tilde{q} = R^T\tilde{p}$ and $\tilde{y} = R\tilde{x}$. It is treated as a feedback connection of a time-invariant dynamical system and a time-invariant memoryless system, as shown in Figure 4.6. Note that the second argument of f in (4.15a) is negative because the input to the upper component in Figure 4.6 is the error term $u_1 - q = -q$ with zero input $u_1 \equiv 0$.

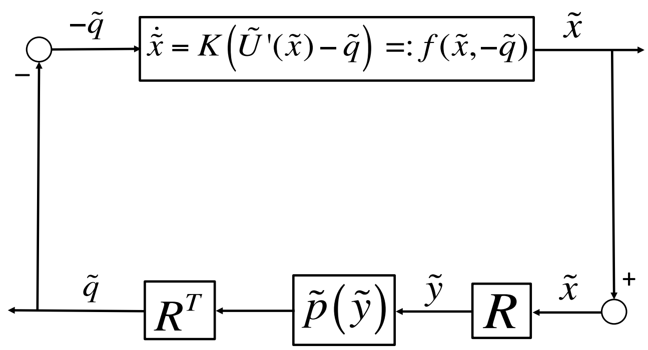


Figure 4.6: The primal algorithm (4.15) is modeled as a feedback connection of a dynamical system and a memoryless system.

Recall condition C2.2' used in Theorem 3.15 that strengthens condition C2.2 by requiring $U_i'(x_i) \to \infty$ as $x_i \to 0$ as well so that $D(p)$ is continuously differentiable.

Theorem 4.9 *Suppose conditions C2.1 and C2.2' hold, and the price functions $p_l(\cdot)$ are locally Lipschitz and nondecreasing. Then the origin of the primal algorithm* (4.15) *is globally asymptotically stable.*

Proof. As discussed above the conditions in the theorem guarantee that (4.15) has a unique solution $(x(t), t \geq 0)$ and the origin is its unique equilibrium. The closed-loop system of (4.15a) is the following system of differential equations in $\tilde{x}$:

$$\dot{\tilde{x}} = K\left(\tilde{U}'(\tilde{x}) - \tilde{q}(\tilde{x})\right) \quad (4.16)$$

We first show that the right-hand side is locally Lipschitz so that Theorem 4.8 is applicable. We claim that the function $\tilde{q}(\tilde{x})$ is locally Lipschitz in $\tilde{x}$. To see this, fix any $\tilde{x}$. Consider any $\hat{x}$ close enough to $\tilde{x}$ so that the corresponding $R\tilde{x}$ and $R\hat{x}$ are close enough. Then since

$p(\cdot)$ is locally Lipschitz, we have $\|p(R(x^* + \hat{x})) - p(R(x^* + \tilde{x}))\| \leq L_{\tilde{x}} \|R\hat{x} - R\tilde{x}\|$ for some finite constant $L_{\tilde{x}}$. Hence

$$\begin{aligned} \|\tilde{q}(\hat{x}) - \tilde{q}(\tilde{x})\| &= \left\|R^T(\tilde{p}(R\hat{x}) - \tilde{p}(R\tilde{x}))\right\| \\ &\leq \left\|R^T\right\| \|p\,(R(x^* + \hat{x})) - p\,(R(x^* + \tilde{x}))\| \\ &\leq L_{\tilde{x}} \left\|R^T\right\| \|R\| \|\hat{x} - \tilde{x}\| \end{aligned}$$

i.e., the function $\tilde{q}(\tilde{x})$ is locally Lipschitz in $\tilde{x}$. Hence the right-hand side of the closed-loop system (4.16) is locally Lipschitz in $\tilde{x}$ ($U'(\tilde{x})$ is also locally Lipschitz since U_i are twice continuously differentiable), and we can apply Theorem 4.8.

Consider the following positive definite candidate storage function for the open-loop dynamical system f in (4.15a):

$$V_1(\tilde{x}) \quad := \quad \frac{1}{2} \sum_i \frac{\tilde{x}_i^2}{\kappa_i}$$

with

$$\dot{V}_1 \quad = \quad \sum_i \tilde{x}_i \frac{\dot{\tilde{x}}_i}{\kappa_i} \quad = \quad \sum_i \tilde{x}_i \left(\tilde{U}_i'(\tilde{x}_i) - \tilde{q}_i\right) \tag{4.17}$$

Since the input to the open-loop system f is $-\tilde{q}$ and the output is the state $\tilde{x}$, (4.17) implies that f is output-feedback passive:

$$(-\tilde{q})^T \tilde{x} \quad = \quad \dot{V}_1 \; + \; \tilde{x}^T \left(-\tilde{U}'(\tilde{x})\right)$$

and zero-state observable. The function ρ_1 in Theorem 4.8 is $-\tilde{U}'$. The lower component of Figure 4.6 is memoryless with $\tilde{x}$ at the input and $\tilde{q}$ at the output:

$$\tilde{q} \quad = \quad R^T \tilde{p}\,(R\tilde{x})$$

and hence automatically input-feedforward passive since

$$\tilde{x}^T \tilde{q} \quad = \quad \tilde{x}^T \left(R^T \tilde{p}\,(R\tilde{x})\right)$$

The function φ_2 in Theorem 4.8 can be taken as $R^T \tilde{p}\,(R\tilde{x})$.

We are left to prove condition (4.13) in Theorem 4.8 which translates to:

$$\tilde{x}^T \left(-\tilde{U}'(\tilde{x}) + R^T \tilde{p}\,(R\tilde{x})\right) \quad = \quad -\tilde{x}^T \tilde{U}'(\tilde{x}) \; + \; \tilde{y}^T \tilde{p}\,(\tilde{y}) \quad > \quad 0 \qquad \text{for all } \tilde{x} \neq 0$$

We first claim that $-\tilde{x}^T \tilde{U}'(\tilde{x}) \;=\; -\sum_i \tilde{x}_i \tilde{U}_i'(\tilde{x}_i) \;>\; 0$ for all $\tilde{x} \neq 0$. By definition

$$\tilde{U}_i'(\tilde{x}_i) \quad = \quad U_i'(x_i^* + \tilde{x}_i) - U_i'(x_i^*)$$

Since U_i is strictly concave, $-\tilde{U}_i'(\tilde{x}_i)$ is strictly increasing in $\tilde{x}_i$. Moreover $-\tilde{U}_i'(0) = 0$ and hence the graph of $-\tilde{U}_i'(\tilde{x}_i)$ lies in the first and third quadrants ($\tilde{x}$ can be negative). Therefore $-\tilde{x}^T \tilde{U}'(\tilde{x}) > 0$ as long as $\tilde{x} \neq 0$. Similarly

$$\tilde{y}^T \tilde{p}(\tilde{y}) \;=\; \sum_l \tilde{y}_l \left(p_l(y_l^* + \tilde{y}_l) - p_l(y_l^*)\right) \;\geq\; 0$$

where the last inequality follows because p_l is nondecreasing and hence $\tilde{y}_l \left(p_l(y_l^* + \tilde{y}_l) - p_l(y_l^*)\right) \geq 0$.

Hence all conditions of Theorem 4.8 are satisfied, including the radial unboundedness of V_1. The origin is therefore globally asymptotically stable. □

4.4 STABILITY OF PRIMAL-DUAL ALGORITHMS

We now discuss the stability of the following class of primal-dual algorithms:

$$\begin{aligned} \dot{x}_i &= \kappa_i \left(U_i'(x_i(t)) - q_i(t)\right)^+_{x_i(t)} \\ \dot{p}_l &= \gamma_l (y_l(t) - c_l)^+_{p_l(t)} \end{aligned}$$

where $(a)^+_b = a$ if $a > 0$ or $b > 0$ and 0 otherwise. This is in vector form:

$$\dot{x} = K \left(U'(x(t)) - q(t)\right)^+_{x(t)} \tag{4.18a}$$

$$\dot{p} = \Gamma (y(t) - c)^+_{p(t)} \tag{4.18b}$$

where $K := \mathrm{diag}(\kappa_i, i = 1, \ldots, N)$ and $\Gamma := \mathrm{diag}(\gamma_l, l = 1, \ldots, L)$. A point (x^*, p^*) is an *equilibrium* of the primal-dual algorithm (4.18) if

$$(U_i'(x_i^*) - q_i^*)^+_{x_i^*} \;=\; 0 \qquad \text{and} \qquad (y_l^* - c_l)^+_{p_l^*} \;=\; 0$$

As for the primal algorithms we assume that conditions C2.1 and C2.2 hold and the routing matrix R has full row rank. Then (4.18) has a unique solution $(x(t), p(t), t \geq 0)$ for every initial point $(x(0), p(0)) \geq 0$ (see Chapter 1.5). Moreover a unique equilibrium point (x^*, p^*) exists, if $x^* > 0$, and is primal-dual optimal for a network utility maximization problem and its dual (see Chapter 2.2).

As in Chapter 4.3 we shift the equilibrium point to the origin by working with the perturbed variables defined as (dropping t to simplify notation):

$$\tilde{x} \;:=\; x - x^* \quad \text{and} \quad \tilde{p} \;:=\; p - p^*$$

Then

$$\tilde{y} \;:=\; R\,\tilde{x} \;=\; R(x - x^*) \;=\; y - y^* \quad \text{and} \quad \tilde{q} \;:=\; R^T \tilde{p} \;=\; R^T (p - p^*) \;=\; q - q^*$$

In terms of the perturbed variables the primal-dual algorithm is:[3]

$$\dot{\tilde{x}} \;=\; \dot{x} \;=\; K\left(U'(x^*+\tilde{x}) \;-\; R^T p^* - R^T \tilde{p}\right)^+_{x^*+\tilde{x}} \;=:\; f(\tilde{x},\, -\tilde{p}) \tag{4.19a}$$

$$\dot{\tilde{p}} \;=\; \dot{p} \;=\; \Gamma\,(y^* - c + \tilde{y})^+_{p^*+\tilde{p}} \;=:\; g(\tilde{p}, \tilde{y}) \tag{4.19b}$$

It can be treated as a zero-input feedback connection shown in Figure 4.7.

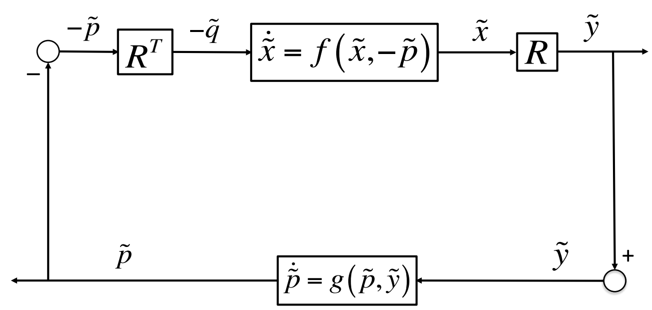


Figure 4.7: The primal-dual algorithm (4.19) is modeled as a feedback connection of two time-invariant dynamical systems.

Compare the primal iterations (4.19a) with (4.15a): since we ignore the projection to nonnegative $x_i(t)$ in (4.15a) the equilibrium for (4.15a) has a simple characterization: $\tilde{U}'(\tilde{x}) \;=\; \tilde{q}$ which, since $\tilde{x} = 0$, $\tilde{q} = 0$, is equivalent to

$$U'(x^*) \;=\; q^*$$

In contrast the equilibrium of (4.19a) is characterized by $U_i'(x_i^* + \tilde{x}_i) \le q_i^* + \tilde{q}_i$ with equality if $x_i^* + \tilde{x}_i > 0$, which is equivalent to

$$U_i'(x_i^*) \;\le\; q_i^* \quad \text{with equality if } \; x_i^* > 0, \qquad i = 1, \ldots, N$$

Lemma 4.10 *Suppose the utility functions U_i are strictly concave.*

1. *The open-loop forward system from $-\tilde{p}$ to $\tilde{y}$ is strictly passive.*

[3]Note that the second argument of f in (4.19a) is negative because the input to the integrator in the forward path in Figure 4.7 is $-q$.

2. *The open-loop backward system from $\tilde{y}$ to $\tilde{p}$ is passive.*

Proof of Lemma 4.10.

1. Consider the candidate storage function

$$V_1(\tilde{x}) \quad := \quad \frac{1}{2}\tilde{x}^T K^{-1}\tilde{x}$$

Clearly V_1 is continuously differentiable, positive definite and radially unbounded. Moreover we have from (4.19a)

$$\begin{aligned} \dot{V}_1(\tilde{x}) &= \tilde{x}^T K^{-1}\dot{\tilde{x}} \;=\; \tilde{x}^T \left(U'(x^* + \tilde{x}) - R^T p^* - R^T \tilde{p}\right)_x^+ \\ &\leq \tilde{x}^T \left(U'(x^* + \tilde{x}) - R^T p^* - R^T \tilde{p}\right) \end{aligned}$$

where the inequality follows from Lemma 3.18 and $\tilde{x} = x - x^*$. Hence, using $x^* + \tilde{x} = x$ and $R^T p^* + R^T \tilde{p} = R^T p = q$ to simplify notation, we have

$$\dot{V}_1(\tilde{x}) \quad \leq \quad \tilde{x}^T\left(U'(x) - q\right) \quad = \quad \tilde{x}^T\left(U'(x) - q\right) \; - \; \tilde{x}^T\left(U'(x^*) - q^*\right)_{x^*}^+$$

where the equality follows because $\left(U'(x^*) - q^*\right)_{x^*}^+ = 0$. Lemma 3.18 implies

$$0 \quad = \quad \tilde{x}^T\left(U'(x^*) - q^*\right)_{x^*}^+ \quad \geq \quad \tilde{x}^T\left(U'(x^*) - q^*\right)$$

and hence

$$\begin{aligned} \dot{V}_1(\tilde{x}) &\leq \tilde{x}^T\left(U'(x) - U'(x^*) - (q - q^*)\right) \\ &= \tilde{x}^T\left(U'(x^* + \tilde{x}) - U'(x^*)\right) \; - \; \tilde{x}^T R^T \tilde{p} \\ &= \tilde{x}^T \underbrace{\left(U'(x^* + \tilde{x}) - U'(x^*)\right)}_{\tilde{U}'(\tilde{x})} \; - \; \tilde{p}^T \tilde{y} \end{aligned}$$

yielding

$$-\tilde{p}^T\tilde{y} \quad \geq \quad \dot{V}_1(\tilde{x}) + \psi(\tilde{x})$$

where $\psi(\tilde{x}) := -\tilde{x}^T\tilde{U}'(\tilde{x})$. We now claim that $\psi(\tilde{x}) > 0$ for all $\tilde{x} \neq 0$ (as also proved in Theorem 4.9), implying that the (open-loop) forward path is strictly passive. To see this note that $-\tilde{x}^T\tilde{U}'(\tilde{x}) \;=\; -\sum_i \tilde{x}_i\tilde{U}_i'(\tilde{x}_i)$ where by definition

$$\tilde{U}_i'(\tilde{x}_i) \quad = \quad U_i'(x_i^* + \tilde{x}_i) - U_i'(x_i^*)$$

Since U_i is strictly concave, $-\tilde{U}_i'(\tilde{x}_i)$ is strictly increasing in $\tilde{x}_i$. Moreover $-\tilde{U}_i'(0) = 0$ and hence the graph of $-\tilde{U}_i'(\tilde{x}_i)$ lies in the first and third quadrants ($\tilde{x}$ can be negative). Therefore $-\tilde{x}^T\tilde{U}'(\tilde{x}) > 0$ as long as $\tilde{x} \neq 0$. This completes the proof that the open-loop forward system from $-\tilde{p}$ to $\tilde{y}$ is strictly passive.

2. Consider the candidate storage function

$$V_2(\tilde{p}) \quad := \quad \frac{1}{2}\,\tilde{p}^T\,\Gamma^{-1}\,\tilde{p}$$

Clearly V_2 is continuously differentiable, positive definite and radially unbounded. Moreover

$$\dot{V}_2(\tilde{p}) \quad = \quad \tilde{p}^T\,\Gamma^{-1}\,\dot{p} \quad = \quad \tilde{p}^T\,(y-c)_p^+ \quad \leq \quad \tilde{p}^T\,(y-c)$$

where the inequality follows from Lemma 3.18 and $\tilde{p} := p - p^*$. Hence

$$\dot{V}_2(\tilde{p}) \quad \leq \quad \tilde{p}^T\,(y - y^* + (y^* - c)) \quad = \quad \tilde{p}^T\tilde{y} + \tilde{p}^T(y^* - c)$$

Using Lemma 3.18 again we have $\tilde{p}^T(y^* - c) \leq \tilde{p}^T(y^* - c)_{p^*}^+ = 0$ and hence

$$\tilde{y}^T\tilde{p} \quad \geq \quad \dot{V}_2(\tilde{p})$$

i.e., the backward system from $\tilde{y}$ to $\tilde{p}$ is passive.

□

We cannot apply Theorem 4.7 because the backward system is neither *strictly* passive nor output strictly passive (though it is zero-state observable). However, since V_1 and V_2 are both positive definite, we will prove that $V(\tilde{x},\tilde{p}) := V_1(\tilde{x}) + V_2(\tilde{p})$ is indeed a Lyapunov function for the primal-dual algorithm (4.19) and apply LaSalle's invariance principle Theorem 3.9.2.

Theorem 4.11 *Suppose conditions C2.1 and C2.2 hold, the routing matrix R has full row rank, and the equilibrium $x^* > 0$. Then the origin of the primal-dual algorithm* (4.19) *is globally asymptotically stable.*

Proof. As discussed above (after (4.18)), the conditions in the theorem guarantee that (4.19) has a unique solution $(x(t), p(t), t \geq 0$ and the origin is its unique equilibrium. Moreover condition C2.2 implies that the dynamics of (4.19) are the projections of locally Lipschitz functions as Theorem 3.9.2 requires (this can be proved in a similar manner as Lemma 3.14).

Consider the candidate Lyapunov function consisting of the sum of the storage functions V_1, V_2 defined in the proof of Lemma 4.10:

$$V(\tilde{x},\tilde{p}) \quad := \quad \frac{1}{2}\left(\tilde{x}^T K^{-1}\tilde{x} \;+\; \tilde{p}^T\Gamma^{-1}\tilde{p}\right) \quad = \quad V_1(\tilde{x}) + V_2(\tilde{p})$$

Clearly V is continuously differentiable, positive definite, and radially unbounded. The proof of Lemma 4.10 shows that

$$\dot{V}_1(\tilde{x}) \;\leq\; \tilde{x}^T\tilde{U}'(\tilde{x}) - \tilde{p}^T\tilde{y} \qquad \text{and} \qquad \dot{V}_2(\tilde{p}) \;\leq\; \tilde{y}^T\tilde{p}$$

and hence

$$\dot{V}(\tilde{x},\tilde{p}) \quad \le \quad \tilde{x}^T \tilde{U}'(\tilde{x})$$

It is also proved there that

$$\tilde{x}^T \tilde{U}'(\tilde{x}) \ := \ \tilde{x}^T \left(U'(x^* + \tilde{x}) - U'(x^*)\right) \quad < \quad 0, \qquad \tilde{x} \neq 0$$

Hence we have $\dot{V}(\tilde{x},\tilde{p}) \le 0$. Hence V satisfies conditions C3.1, C3.2, and C3.3. (V does not satisfy $\dot{V}(\tilde{x},\tilde{p}) < 0$ for $(\tilde{x},\tilde{p}) \neq 0$ because it is possible that $\dot{V}(0,\tilde{p}) = 0$ for some $\tilde{p}$). LaSalle's invariance principle implies that the trajectory $(x(t), p(t))$ will converge to the largest invariance set in

$$E \quad := \quad \{(\tilde{x},\tilde{p}) \,|\, \dot{V}(\tilde{x},\tilde{p}) = 0\}$$

We now prove that the only solution to (4.19) that can stay identically in E is the equilibrium solution $(\tilde{x},\tilde{p}) \equiv 0$. Theorem 3.9.2 then implies the global asymptotic stability of the origin.

Since $\tilde{x}^T \tilde{U}'(\tilde{x}) < 0$ for all $\tilde{x} \neq 0$, $\dot{V} \equiv 0$ implies $\tilde{x}^T \tilde{U}'(\tilde{x}) \equiv 0$ and hence $\tilde{x} \equiv 0$. Hence $\dot{\tilde{x}} \equiv 0$. From this and (4.19a) we have (since $p(t) = p^* + \tilde{p}(t)$)

$$\left(U'(x^*) - R^T p(t)\right)^+_{x^*} \quad \equiv \quad 0 \tag{4.20}$$

Since

$$\dot{V}(\tilde{x},\tilde{p}) \quad = \quad \tilde{x}^T K^{-1} \dot{\tilde{x}} + \tilde{p}^T \Gamma^{-1} \dot{\tilde{p}} \quad = \quad \tilde{x}^T \left(U'(x(t)) - R^T p(t)\right)^+_{x(t)} \ + \ \tilde{p}^T (y(t) - c)^+_{p(t)}$$

(4.20) and $\tilde{x} \equiv 0$ imply $y(t) \equiv y^*$ and hence

$$\dot{V}(\tilde{x} \equiv 0, \tilde{p}) \quad = \quad \tilde{p}^T (y^* - c)^+_{p(t)}$$

Suppose now $(0, \tilde{p}(t)) \in E$ for all $t \ge 0$. Then $\dot{V} \equiv 0$ implies

$$\sum_l \tilde{p}_l(t)\,(y_l^* - c_l)^+_{p_l(t)} \quad \equiv \quad 0 \tag{4.21}$$

We claim that this implies, for each t and each l,

$$\tilde{p}_l(t)\,(y_l^* - c_l)^+_{p_l(t)} \quad = \quad 0 \tag{4.22}$$

To see this, fix any time t and any link l. Clearly (4.22) holds if $\tilde{p}_l(t) = 0$. If $\tilde{p}_l(t) < 0$ then $p_l^* > p_l(t) \ge 0$ and hence $y_l^* = c_l$ since (x^*, p^*) is the (unique) equilibrium point, implying (4.22). On the other hand if $\tilde{p}_l(t) > 0$ then $p_l(t) > p_l^* \ge 0$ and hence $\tilde{p}_l(t)\,(y_l^* - c_l)^+_{p_l(t)} = \tilde{p}_l(t)\,(y_l^* - c_l) \ \le \ 0$. Therefore every term on the left-hand side of (4.21) is nonpositive, and therefore every term must itself be zero.

If $\tilde{p}_l(t) \neq 0$ then (4.22) implies $(y_l^* - c_l)^+_{p_l(t)} = 0$, whereas if $\tilde{p}_l(t) = 0$ then

$$(y_l^* - c_l)^+_{p_l(t)} \quad = \quad (y_l^* - c_l)^+_{p^*} \quad = \quad 0$$

where the last equality follows because (x^*, p^*) is the equilibrium point of (4.18). Hence, for all $\tilde{p}_l(t)$, we have

$$(y_l^* - c_l)^+_{p_l(t)} \quad = \quad 0 \qquad \text{for all } l, t \tag{4.23}$$

But (4.20)(4.23) means that $(x^*, p(t))$ is an equilibrium point of the primal-dual algorithm (4.18) for every t. Since the equilibrium of (4.18) is unique under assumptions of the theorem we must have $p(t) \equiv p^*$ and $\tilde{p}(t) \equiv 0$. This shows that the only solution trajectory that can stay identically in E is the equilibrium solution $(\tilde{x}, \tilde{p}) = (0, 0)$.

The global asymptotic stability of the origin then follows from Theorem 3.9.2. $\square$

4.5 BIBLIOGRAPHICAL NOTES

The passivity theorems in Chapter 4.1 follow [29]. Passivity theory is first applied to congestion control in [53]. The stability proofs here for congestion control modeled by projected dynamics are new.

CHAPTER 5

Global stability: gradient projection method

In this chapter we introduce another approach to proving stability of congestion control algorithms. Using a discrete-time model this approach regards the dynamics of the closed-loop system as a gradient projection algorithm for the Lagrangian dual problem of network utility maximization. We derive sufficient conditions for this algorithm to converge to dual-optimal link prices, from which primal-optimal rates can be recovered.

5.1 CONVERGENCE THEOREMS

Consider the problem

$$\min_x \quad f(x) \quad \text{subject to} \quad x \in X \tag{5.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and $X \subseteq \mathbb{R}^n$ is nonempty, closed and convex. Let the column vector $\nabla f(x)$ denote the gradient of f evaluated at x, i.e., $[\nabla f(x)]_i := \partial f / \partial x_i,\ i = 1, \ldots, n$. Recall that a point x^* is a *local minimizer* if $f(x^*)$ is minimum on a neighborhood of x^*, i.e., there exists $r > 0$ such that $f(x^*) \leq f(x)$ for all $x \in B_r(x^*) \cap X$. A necessary optimality condition for general f is: if $x^* \in X$ is a local minimizer for (5.1) then there is a neighborhood $B_r(x^*)$ for some $r > 0$ such that

$$(\nabla f(x^*))^T (x - x^*) \quad \geq \quad 0 \qquad \forall x \in B_r(x^*) \cap X \tag{5.2}$$

i.e., moving away from x^* to any other feasible point x in $B_r(x^*)$ can only locally increase the function value f. If f is a convex function (X is assumed convex) then this is both necessary and sufficient for x^* to be a *global* minimum[1] of (5.1). This is illustrated in Figure 5.1.

The steepest descent (first-order gradient) algorithm for solving (5.1) is given by the following iteration: starting from an initial point $x(0) = x_0$,

$$x(t+1) \quad = \quad [\ x(t) - \gamma \nabla f(x(t))\ ]_X \tag{5.3}$$

where $\gamma > 0$ is a stepsize. Here $[x]_X$ denotes the projection of x onto X, i.e., for any $x \in \mathbb{R}^n$,

$$[x]_X \quad := \quad \arg\min_{y \in X} \quad \|x - y\|_2 \tag{5.4}$$

[1]For a proof see the discussion after Theorem 2.10.

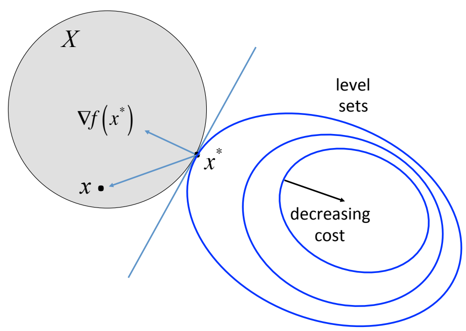


Figure 5.1: Moving away from an optimal point x^* to any other feasible point x can only locally increase the cost.

where $\|\cdot\|_2$ is the *Euclidean* norm. Hence $[x]_X$ is the point in X that is closest to $x \in \mathbb{R}^n$ in the Euclidean norm. To prove the convergence of the gradient projection algorithm (5.3), we need the following properties of the projection operation. They are illustrated in Figure 5.2.

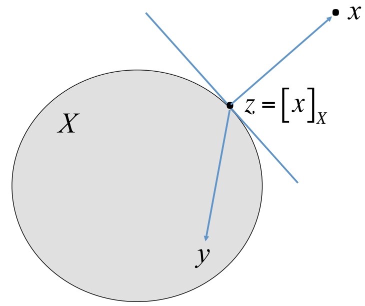


Figure 5.2: The point $z := [x]_X$ is the unique closest point to x in the convex set X under the *Euclidean* norm. For all other points $y \in X$, the inner product of $y - z$ and $x - z$ is nonpositive.

Lemma 5.1 Projection Theorem. *Suppose $X \subseteq \mathbb{R}^n$ is a nonempty, closed and convex set.*

1. *For every $x \in \mathbb{R}^n$ there exists a unique $[x]_X$ defined by* (5.4)*.*
2. *For every $x \in \mathbb{R}^n$, $z = [x]_X$ if and only if $z \in X$ and $(y - z)^T (x - z) \leq 0$ for all $y \in X$.*
3. *The projection mapping $T : \mathbb{R}^n \to \mathbb{X}$ defined by $T(x) := [x]_X$ is continuous and* non-expansive *under the Euclidean norm, i.e.,*

$$\|[y]_X - [x]_X\|_2 \quad \leq \quad \|y - x\|_2 \qquad \forall x, y \in \mathbb{R}^n$$

Note that Lemma 5.1 does not require X to be bounded (compact), only closed. This is because since X is nonempty there is an $w \in X$. Hence the minimization in the projection (5.4) can be equivalently restricted to the compact set $\{y \in X \mid \|x - y\|_2 \leq \|x - w\|_2\}$.

We assume:

C5.1: The objective function f is lower bounded on X, continuously differentiable and convex. The feasible set X is nonempty, closed and convex.

C5.1 guarantees that (5.1) is feasible and the gradient projection algorithm (5.3) is well defined. Since X is not necessarily compact (bounded), the optimal may not be attained (e.g., $X = \mathbb{R}$ and $f(x) = e^{-x}$). Moreover the sequence $(x(t), t = 0, 1, \dots)$ generated by the gradient projection algorithm (5.3) may not stay bounded and hence may not have any convergent subsequence (the Bolzano-Weierstrass theorem states that a sequence $(x(t), t = 0, 1, \dots)$ has a convergent subsequence if it is bounded).

To guarantee that the gradient projection algorithm makes progress towards minimzing f, we need:

C5.2: The gradient of f is Lipschitz continuous with a Lipschitz constant K, i.e.,

$$\|\nabla f(y) - \nabla f(x)\|_2 \quad \leq \quad K \, \|y - x\|_2 \qquad \forall x, y \in \mathbb{R}^n$$

Note that the norm is Euclidean.[2] C5.2 implies the following useful result, proved in Appendix 5.3.

Lemma 5.2 **Descent Lemma.** *If $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and satisfies C5.2 then*

$$f(x+y) \quad \leq \quad f(x) + y^T \nabla f(x) + \frac{K}{2}\|y\|_2^2 \qquad \forall x, y \in \mathbb{R}^n$$

The main result on the gradient projection algorithm (5.3) is the following characterization. Conditions C5.1 and C5.2 do not guarantee that the sequence $(x(t), t = 0, 1, \dots)$ generated by the gradient projection algorithm has any convergent subsequence, but if it does (e.g., if X is also bounded) then it converges to an optimal point x^* of (5.1) provided the stepsize γ is sufficiently small.

This implies that, when X is bounded and hence compact, and f is strictly convex so that a unique optimal point x^* exists, then $(x(t), t = 0, 1, \dots)$ itself converges to x^* under the gradient projection algorithm.

Theorem 5.3 *Suppose conditions C5.1 and C5.2 hold, and suppose $0 < \gamma < 2/K$. If the sequence $(x(t), t = 0, 1, \dots)$ produced by the gradient projection algorithm (5.3) has a convergent subsequence $(x(t_k), k = 1, 2, \dots)$ then its limit x^* is an optimal solution of (5.1).*

Proof. We prove the theorem in three steps. First we show the sequence $(f(x(t)), t = 0, 1, \dots)$ of objective values converges monotonically. Moreover the difference sequence

[2] In contrast, the norm that defines a *contraction mapping* can be arbitrary (see Definition 5.5 below).

$(x(t+1) - x(t), t = 0, 1, \dots)$ converges to zero. Specifically, by the Descent Lemma 5.2, we have

$$f(x(t+1)) \;\leq\; f(x(t)) \;+\; (\, x(t+1) - x(t)\,)^T \, \nabla f(x(t)) \;+\; \frac{K}{2}\, \|x(t+1) - x(t)\|_2^2 \tag{5.5}$$

Lemma 5.1.2 implies that for all t

$$(\, y - x(t+1)\,)^T \; (\, x(t) - \gamma \nabla f(x(t)) \;-\; x(t+1)\,) \;\leq\; 0 \qquad \forall y \in X \tag{5.6}$$

In particular let $y = x(t)$ and we have, after rearranging,

$$(\, x(t+1) - x(t)\,)^T \, \nabla f(x(t)) \quad \leq \quad -\frac{1}{\gamma}\, \|x(t+1) - x(t)\|_2^2$$

Substituting into (5.5) we have

$$f(x(t+1)) \quad \leq \quad f(x(t)) \;-\; \left(\frac{1}{\gamma} - \frac{K}{2} \right) \|x(t+1) - x(t)\|_2^2 \tag{5.7}$$

Hence the sequence $(f(x(t)), t = 0, 1, \dots)$ is strictly decreasing as long as $x(t+1) \neq x(t)$ provided $\gamma < 2/K$. Since f is lower bounded on X (condition C5.1), the sequence $(f(x(t)), t = 0, 1, \dots)$ is bounded and monotone and thus converges. Rearranging (5.7), we also have

$$\|x(t+1) - x(t)\|_2^2 \quad \leq \quad \left(\frac{1}{\gamma} - \frac{K}{2} \right)^{-1} (f(x(t)) - f(x(t+1)))$$

Since $f(x(t))$ converges this means that the *differences* $x(t+1) - x(t)$ converge to zero (though this does not guarantee that $x(t)$ itself converges).

Second suppose there is a subsequence $(x(t_k), k = 1, 2, \dots)$ that converges to x^*. Consider the sequence $(x(t_k + 1), k = 1, 2, \dots)$. By Lemma 5.1.3, the iteration $x(t+1) = [x(t) - \gamma \nabla f(x(t))]_X$ defined by (5.3) is a projection and hence a continuous function of $x(t)$. Hence the sequence $(x(t_k + 1), k = 1, 2, \dots)$, being the image of a continuous function on $x(t_k)$, also converges. We now show that it converges to x^* as $k \to \infty$. Fix any $\epsilon > 0$. We have to show that there exists an K such that

$$\|x(t_k + 1) - x^*\|_2 \quad < \quad \epsilon \qquad \forall k > K$$

Since $x(t_k) \to x^*$ there exists an K' such that

$$\|x(t_k) - x^*\|_2 \quad < \quad \frac{\epsilon}{2} \qquad \forall k > K' \tag{5.8a}$$

Step 1 above shows that $x(t_k + 1) - x(t_k)$ converges to zero and hence there exists K'' such that

$$\|x(t_k + 1) - x(t_k)\|_2 \quad < \quad \frac{\epsilon}{2} \qquad \forall k > K'' \tag{5.8b}$$

Combining (5.8) we have for $k > K := \max\{K', K''\}$

$$\|x(t_k+1) - x^*\|_2 \quad \leq \quad \|x(t_k+1) - x(t_k)\|_2 \;+\; \|x(t_k) - x^*\|_2 \quad < \quad \epsilon$$

as desired.

Finally note that (5.6) holds for all t. In particular consider $t = t_k, k = 1, 2, \ldots$. Taking $k \to \infty$, (5.6) yields

$$\left(y - \lim_k x(t_k+1) \right)^T \left(\lim_k x(t_k) \;-\; \gamma \lim_k \nabla f(x(t_k)) \;-\; \lim_k x(t_k+1) \right) \;\leq\; 0, \quad \forall y \in X$$

Since f is continuously differentiable and $\lim_k x(t_k) = \lim_k x(t_k+1) = x^*$, we have

$$\gamma \;\left(y - x^* \right)^T \nabla f(x^*) \quad \geq \quad 0 \qquad \forall y \in X$$

Hence x^* satisfies the optimality condition (5.2) and is globally optimal since f is a convex function over a convex set X. $\square$

When f satisfies a stronger form of convexity then the gradient projection algorithm indeed converges and does so geometrically. Suppose f is twice continuously differentiable (not just continuously differentiable as guaranteed by condition C5.1). Then f is strictly convex if and only if $\nabla^2 f(x) \succ 0$ (positive definite) for all x. If, for some $\alpha > 0$, $\nabla^2 f(x) \succeq \alpha I$ for all $x \in \mathbb{R}^n$ then the gradient projection algorithm (5.3) is a contraction mapping and $x(t)$ converges geometrically to the unique optimal solution x^* of (5.1) with rate α. We now make this precise.

Consider:

C5.3: For some $\alpha > 0$, f satisfies

$$\left(\nabla f(y) - \nabla f(x) \right)^T (y - x) \quad \geq \quad \alpha \, \|y - x\|_2^2 \qquad \forall x, y \in \mathbb{R}^n \tag{5.9}$$

We say f is *strongly convex* if it satisfies condition C5.3. The next result shows that it is stronger than strict convexity. It is proved in Appendix 5.4.

Lemma 5.4 Strong convexity. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable. If f satisfies C5.3 then f is strictly convex. Indeed (5.9) is equivalent to $\nabla^2 f(x) \succeq \alpha I$ for all $x \in \mathbb{R}^n$ when f is twice continuously differentiable.*

A key implication of Lemma 5.4 is that the gradient projection algorithm (5.3) is a contraction mapping, as we now explain.

Definition 5.5 Contraction. Consider a function $T : X \to X$ from a subset X of $\mathbb{R}^n$ into itself. T is called a *contraction mapping* or simply a *contraction* if there exists an $\alpha \in [0, 1)$ such that

$$\|T(y) - T(x)\| \quad \leq \quad \alpha\|y - x\| \qquad \forall x, y \in X$$

for an arbitrary norm $\| \cdot \|$.

A function T can be a contraction under a certain norm, but not under a different norm, so the proper choice of norm is critical.

Theorem 5.6 Contraction theorem. *Suppose $T : X \to X$ is a contraction mapping on a closed subset X of $\mathbb{R}^n$. Then*

1. *There exists a unique fixed point x^* such that $x^* = T(x^*)$.*
2. *Starting from any initial point $x(0) \in X$, the contraction iteration $x(t + 1) := T(x(t))$ converges geometrically to x^*; in particular*

$$\|x(t) - x^*\| \quad \leq \quad \alpha^t \,\|x(0) - x^*\| \qquad \forall t \geq 0$$

Proof. Consider the contraction iteration $x(t + 1) := T(x(t))$. Definition 5.5 implies

$$\|x(t + 1) - x(t)\| \;\leq\; \alpha \,\|x(t) - x(t - 1)\| \quad \leq \quad \cdots \quad \leq \quad \alpha^t \,\|x(1) - x(0)\|$$

Hence, for all $t \geq 0$ and $s \geq 1$, we have

$$\begin{aligned}
\|x(t + s) - x(t)\| \quad &= \quad \left\| \sum_{m=0}^{s-1} (x(t + m + 1) - x(t + m)) \right\| \\
&\leq \quad \sum_{m=0}^{s-1} \|x(t + m + 1) - x(t + m)\| \quad \leq \quad \|x(1) - x(0)\| \,\alpha^t \sum_{m=0}^{s-1} \alpha^m \\
&\leq \quad \frac{\alpha^t}{1 - \alpha} \,\|x(1) - x(0)\|
\end{aligned}$$

Since $\alpha \in [0, 1)$, $x(t)$ is a Cauchy sequence and hence must converge to a point x^* in $\mathbb{R}^n$. Since X is closed, $x^* \in X$. Since T is continuous,

$$x^* \quad = \quad \lim_t x(t + 1) \quad = \quad \lim_t T(x(t)) \quad = \quad T(\lim_t x(t)) \quad = \quad T(x^*)$$

and hence x^* is a fixed point of T. Moreover, the fixed point is unique for, otherwise, if x^* and y^* are both fixed points then

$$\|y^* - x^*\| \quad = \quad \|T(y^*) - T(x^*)\| \quad \leq \quad \alpha \,\|y^* - x^*\|$$

implying $y^* = x^*$ since $\alpha \in [0,1)$. This completes the proof of part 1.

For part 2, we have for all $t \geq 1$,

$$\|x(t) - x^*\| \quad = \quad \|T(x(t-1)) - T(x^*)\| \quad \leq \quad \alpha \, \|x(t-1) - x^*\|$$

Hence $\|x(t) - x^*\| \;\leq\; \alpha^t \, \|x(0) - x^*\|$. $\square$

Consider the mapping defined by the gradient projection algorithm (5.3):

$$T(x) \quad := \quad [\, x - \gamma \, \nabla f(x) \,]_X$$

The key observation is that, if f is strongly convex then T is a contraction. Theorem 5.6 then implies that the gradient projection algorithm converges geometrically to the unique optimal solution of (5.1). In particular condition C5.2 (Lipschitz continuity of ∇f) guarantees strict descent for sufficiently small stepsize $\gamma > 0$ and condition C5.3 (strong convexity of f) guarantees geometric convergence.

Theorem 5.7 *Suppose conditions C5.1–C5.3 hold. Then there is a unique optimal solution x^* for (5.1) and the gradient projection algorithm (5.3) converges geometrically to x^*, provided the stepsize γ satisfies:*

$$\begin{array}{lll} \text{if } \alpha < K: & 0 < \gamma < \dfrac{2\alpha}{K^2} & \\ \text{if } \alpha \geq K: & 0 < \gamma < \dfrac{\alpha}{K^2} - d \quad \text{or} & \dfrac{\alpha}{K^2} + d \;<\; \gamma \;<\; \dfrac{2\alpha}{K^2} \end{array}$$

where $d := \sqrt{\alpha^2 - K^2}/K^2$. Then

$$\|x(t) - x^*\| \quad \leq \quad \beta^t \, \|x(0) - x^*\| \qquad \forall t \geq 0$$

where $\beta := \sqrt{K^2\gamma^2 - 2\alpha\gamma + 1} \,\in\, (0,1)$.

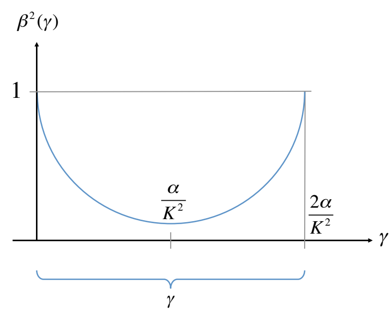
Proof. The gradient project algorithm (5.3) is the following iteration $x(t+1) = T(x(t))$ where $T : X \to X$ is defined by $T(x) := [\, x - \gamma \, \nabla f(x) \,]_X$. We will show that T is a contraction under conditions C5.2 and C5.3. Then the assertions follow from Theorem 5.6.

We have under the Euclidean norm

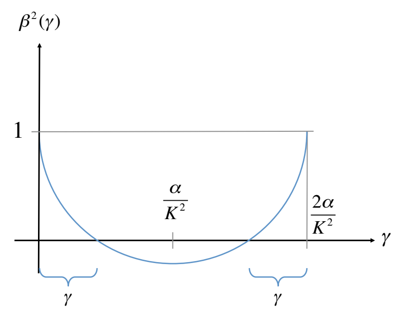
$$\begin{aligned} \|T(y) - T(x)\|_2^2 \quad &= \quad \|\, [y - \gamma \, \nabla f(y)]_X \;-\; [x - \gamma \, \nabla f(x)]_X \,\|_2^2 \\ &\leq \quad \|\, (y - x) - \gamma \, (\nabla f(y) - \nabla f(x))\|_2^2 \\ &= \quad \|y - x\|_2^2 \;-\; 2\gamma \, (\nabla f(y) - \nabla f(x))^T (y - x) \;+\; \gamma^2 \|\nabla f(y) - \nabla f(x))\|_2^2 \end{aligned}$$

where the inequality above follows from the fact that the projection operation is nonexpansive (Lemma 5.1.3). Conditions C5.3 and C5.2 guarantee that $(\nabla f(y) - \nabla f(x))^T (y - x) \geq \alpha \|y - x\|_2^2$ and $\|\nabla f(y) - \nabla f(x))\|_2^2 \leq K^2 \|y - x\|_2^2$ respectively. Hence

$$\|T(y) - T(x)\|_2^2 \quad \leq \quad \left(1 - 2\alpha\gamma + \gamma^2 K^2\right) \|y - x\|_2^2$$



(a) $\alpha < K$



(b) $\alpha \geq K$

Figure 5.3: The function $\beta^2(\gamma)$. (a) If $\alpha < K$ then T is a contraction for any stepsize $\gamma \in (0, 2\alpha/K^2)$. (b) If $\alpha \geq K$ then T is a contraction if $\gamma \in (0, \alpha/K^2 - d)$ or if $\gamma \in (\alpha/K^2 + d, 2\alpha/K^2)$ where $d := \sqrt{\alpha^2 - K^2}/K^2$.

Hence T is a contraction if and only if $\beta^2(\gamma) := 1 - 2\alpha\gamma + \gamma^2 K^2 \in [0,1)$. The function $\beta^2(\gamma)$ is shown in Figure 5.3. Therefore the condition on the stepsize γ in the theorem guarantees T is a contraction with parameter $\beta(\gamma) \in (0,1)$. The assertions then follow from Theorem 5.6. $\square$

If any subsequence of $x(t)$ converges then Theorem 5.3 guarantees that it converges to an optimal solution x^*. Theorem 5.7 guarantees that, indeed, $x(t)$ converges to the unique optimal x^* and does so geometrically. The bound $2/K$ on the stepsize γ in Theorem 5.3 depends only on the first-order information (the Lipschitz constant K of the gradient ∇f of the objective function). The bound $2\alpha/K^2$ on the stepsize γ in Theorem 5.7 depends also on the second-order information α, the strength of the convexity of f.

5.2 STABILITY OF DUAL ALGORITHMS

We now apply Theorems 5.3 and 5.7 to prove the stability of dual algorithms. Dual algorithms have dynamics only in the congestion prices, but not in the source rates. They take the form, in discrete time,

$$x_i(t) \;=\; U_i'^{-1}(q_i(t)) \tag{5.10a}$$
$$p_l(t+1) \;=\; (\, p_l(t) + \gamma(y_l(t) - c_l)\,)^+ \;=:\; g_l(y_l(t), p_l(t)) \tag{5.10b}$$

where $(z)^+ := \max\{z, 0\}$. Here $q_i(t) := \sum_l R_{li} p_l(t)$ and $y_l(t) := \sum_i R_{li} x_i(t)$. Since x is statically determined by p through (5.10a) we abuse notation and often write $x(p)$ and $y(p) = Rx(p)$. We say p^* is an equilibrium (fixed point) if $g(p^*) = g(y(p^*), p^*) = p^*$.

We make the following assumptions on the utility functions U_i:

C5.4:

(a) Conditions C2.1 and C2.2' hold. Moreover the routing matrix R has full row rank.

(b) The sequence $x(t)$ generated by the dual algorithm (5.10) stays in a compact set where the curvatures of U_i are bounded away from zero uniformly: $-U_i''(x_i) \geq 1/\overline{\alpha} > 0$ for all i, $x_i \geq 0$. Let P denote the corresponding set that contains $p(t)$ generated by (5.10).

(c) The curvatures of U_i are bounded above uniformly: $0 < -U_i''(x_i) \leq 1/\underline{\alpha}$ for all i, $x_i \geq 0$.

Condition C5.4(a) implies that the utility of flow i strictly increases as the sending rate increases, but there is a strictly diminishing return. It also assumes that the sending rate decreases to zero as the congestion price q_i increases without bound. and the sending rate increases without bound as q_i approaches 0. This assumption guarantees that for any $p \in \mathbb{R}^L_+$ there is a unique maximizer to $\max_{x_i \geq 0}\; U_i(x_i) - x_i q_i(p)$ given by $x_i(p)$ in (5.10a) without projection. Moreover the dual objective function $D(p)$ is continuously differentiable in p and its derivative is $\nabla D(p) = c - Rx(p)$. These are reasonable assumptions satisfied by practical algorithms (e.g., Reno, Vegas, FAST). The other conditions are technical and required for the convergence proof. As we will see below, conditions C5.4(a)(b) guarantee that the dual algorithm (5.10) converges to the unique optimal solution of the dual problem (5.11b). Condition C5.4(c) further implies that the dual algorithm is a contraction mapping and hence converges geometrically to the optimal solution. Condition C5.4(b) cannot hold for arbitrarily large x_i since it will contradict C5.4(a) that $U_i(x_i)$ is increasing. However since the dual objective function $D(p)$ below is radially unbounded (Lemma 3.16), the iterates $x(t)$ generated by the dual algorithm (5.10) will stay in a compact set.

Consider the network utility maximization:

$$\max_{x \geq 0} \sum_i U_i(x_i) \quad \text{subject to} \quad Rx \leq c \tag{5.11a}$$

and its Lagrangian dual:

$$\min_{p \geq 0} \qquad D(p) := \sum_i \left(U_i(x_i(p)) - x_i(p) \sum_l R_{li} p_l \right) + \sum_l p_l c_l \tag{5.11b}$$

The dual objective function $D(p)$ is convex and, since U_i are strictly concave (condition C5.4(a)), $D(p)$ is also continuously differentiable. By the KKT Theorem 2.12 of Chapter 2.1.2, the necessary and sufficient condition for a point $(x^* := x(p^*), p^*)$ to be optimal is:

$$p^* \geq 0, \quad x^* = \left(U_i'^{-1}(R^T p^*) \right)^+, \quad Rx^* \leq c \;\; \text{with} \; \sum_i R_{li} x_i^* = c_l \;\; \text{if} \; p_l^* > 0$$

This coincides with the equilibrium condition for the dual algorithm (5.10) since condition C5.4(a) means that the projection on x^* is never active. Hence $(x(p^*), p^*)$ is an equilibrium point of the dual algorithms (5.10) if and only if it solves the network utility maximization (5.11a) and its dual (5.11b). The equilibrium prices p^* are generally nonunique, but the equilibrium rates $x(p^*)$ are unique (for any p^*) since the utility functions U_i are strictly concave under condition C5.4(a). When R has full row rank then p^* is unique as well (since $x^* > 0$).

Define $\overline{L} := \max_{i \in N} \sum_l R_{li}$ and $\overline{N} := \max_{l \in L} \sum_i R_{li}$. In words $\overline{L}$ is the length of a longest path used by the sources, $\overline{N}$ is the number of sources sharing a most crowded link. Conditions C5.4(a)-(c) ensures that the algorithm (5.10) defines a contraction mapping, and hence converges geometrically to the unique equilibrium.

Theorem 5.8 *Suppose conditions C5.4(a)(b) hold.*

1. *The dual algorithm* (5.10) *has a unique equilibrium* p^* *that is the unique minimizer of the dual problem* (5.11b)*; moreover* $x(p^*)$ *is the unique maximizer of the primal* (5.11a)*.*

2. *Starting from any initial point* $p(0) \geq 0$*,* $p(t)$ *converges to* p^**, provided*

$$\gamma \quad < \quad \frac{2}{\overline{\alpha}\overline{L}\overline{N}}$$

3. *If, in addition, C5.4(c) holds then* $p(t)$ *converges geometrically to* p^* *provided* $\gamma > 0$ *is sufficiently small.*

Proof.
Part 1. The discussion preceding the theorem proves part 1.

Part 2. We will prove in Lemmas 5.9 and 5.10 that the dual objective function $D(p)$ satisfies conditions C5.1 and C5.2. Then Theorem 5.3 implies that any limit point $\hat{p}$ of the sequence $(p(t), t = 0, 1, \dots)$ generated by the dual algorithm (5.10), if exists, is optimal, and hence it must be equal to the unique optimal point p^*. Moreover it means that, provided the stepsize satisfies

$$0 \;<\; \gamma \;<\; \frac{2}{\overline{\alpha}\overline{L}\overline{N}}$$

$(D(p(t)), t = 0, 1, \dots)$ is a decreasing sequence (see the proof of Theorem 5.3). Since, from Lemma 3.16, $D(p)$ is radially unbounded under condition C5.4(a), this implies that the sequence $p(t)$ stays in the compact set $\{p | D(p) \leq D(p(0))\}$ (see Exercise 3.2). Hence the sequence $p(t)$ has a convergent subsequence by the Bolzano-Weierstrass theorem. Since all such subsequences converge to p^*, the sequence $(p(t), t = 0, 1, \dots)$ itself converges to the optimal point p^*.

Hence we are left with verifying C5.1 and C5.2. For C5.1 the key observation is that (5.10) is a gradient projection algorithm for solving the dual problem (5.11b) with the nonempty, closed and convex feasible set $\{p \in \mathbb{R}^L \,|\, p \geq 0\}$.

Lemma 5.9 *Under condition C5.4(a) the dual objective function* $D(p)$ *is lower bounded, continuously differentiable, and strictly convex.*

Proof of Lemma 5.9. By duality $D(p)$ is lower bounded by the optimal primal objective value which is clearly attained. Since U_i are strictly concave $x_i(p)$ is unique for each $p \geq 0$. The assumption $\lim_{x_i \to 0} U_i'(x_i) = \infty$ implies that the maximizer $x_i(p)$ satisfies

$$U_i'(x_i(p)) \quad = \quad q_i(p)$$

for all $p \geq 0$. Hence $D(p)$ in (5.11b) is differentiable with $\nabla D(p) = c - y(p)$. Moreover the derivative $\nabla D(p)$ is continuous since U_i are continuously differentiable. Hence the dual algorithm (5.10) is equivalent to the gradient projection (steepest descent) algorithm

$$p(t+1) \quad = \quad [p(t) - \gamma \nabla D(p(t))]^+$$

for solving the dual problem (5.11b).

Taking the derivative of $\nabla D(p) = c - y(p)$ yields the Hession of D as:

$$\nabla^2 D(p) \quad = \quad -R\left[\frac{\partial x}{\partial p}(p)\right] \quad = \quad RB(p)R^T$$

where $B(p) = \text{diag}(\beta_i(p), i \in N)$ is an $N \times N$ diagonal matrix with diagonal elements

$$\beta_i(p) \quad = \quad \frac{1}{-U_i''(x_i(p))} \quad > \quad 0 \tag{5.12}$$

Since R is of full row rank, $\nabla^2 D(p) = RB(p)R^T \succ 0$ for all $p \geq 0$.[3] Hence $D(p)$ is strictly convex. $\square$

The following result shows that $D(p)$ satisfies condition C5.2.

Lemma 5.10 *Under conditions C5.4(a)(b), ∇D is Lipschitz with*

$$||\nabla D(p) - \nabla D(\hat{p})||_2 \quad \leq \quad \overline{\alpha}\overline{L}\overline{N}\ ||p - \hat{p}||_2$$

for all $p, \hat{p}$ in the set P in condition C5.4(b).

Proof of Lemma 5.10. We will show that $||\nabla^2 D(p)||_2 = ||RB(p)R^T||_2 \leq \overline{\alpha}\overline{L}\overline{N}$ for all $p \in P$. The lemma then follows from [41, Theorem 9.19]

$$||\nabla D(p) - \nabla D(\hat{p})||_2 \quad \leq \quad ||\nabla^2 D(p)||_2\ ||p - \hat{p}||_2 \quad \leq \quad \overline{\alpha}\overline{L}\overline{N}\ ||p - \hat{p}||_2$$

Now

$$||RB(p)R^T||_2^2 \quad \leq \quad ||RB(p)R^T||_\infty \cdot ||RB(p)R^T||_1$$

i.e., $||RB(p)R^T||_2^2$ is upper bounded by the product of the maximum row sum and the maximum column sum of the $L \times L$ matrix $RB(p)R^T$. Since $RB(p)R^T$ is symmetric,

[3] Indeed, for all $p \geq 0, w \geq 0$, we have, for $w \neq 0$,

$$w^T \nabla^2 D(p) w \quad = \quad w^T RB(p)R^T w \quad \geq \quad \left(\min_i \beta_i(p)\right)\ ||R^T w||_2^2 \quad > \quad 0$$

where the last inequality follows because U_i are strictly concave and $R^T w$ is nonzero for a nonzero w since R has full row rank.

$||RB(p)R^T||_1 = ||RB(p)R^T||_\infty$, and hence

$$\begin{aligned} ||RB(p)R^T||_2 &\le ||RB(p)R^T||_\infty \\ &= \max_l \sum_{l'} [RB(p)R^T]_{ll'} \\ &= \max_l \sum_i \beta_i(p) R_{li} \left(\sum_{l'} R_{l'i} \right) \end{aligned}$$

By definition, $(\sum_{l'} R_{l'i}) \le \overline{L}$ and $\beta_i(p) \le \overline{\alpha}$. Hence

$$||RB(p)R^T||_2 \le \overline{\alpha}\,\overline{L} \max_l \left(\sum_i R_{li} \right) \le \overline{\alpha}\,\overline{L}\,\overline{N}$$

as desired. □

This completes the proof of part 2 of the theorem.

Part 3. We first show that $D(p)$ is not only strictly convex, but also strongly convex and hence satisfies condition C5.3. Since R has full row rank the minimum eigenvalue $\underline{\lambda}$ of RR^T is positive.

Lemma 5.11 *Conditions C5.4(a) and (c) imply that*

$$(\nabla D(p) - \nabla D(\hat{p}))^T (p - \hat{p}) \ge \underline{\alpha}\underline{\lambda} ||p - \hat{p}||_2^2, \qquad \forall p, \hat{p} \ge 0$$

where $\underline{\lambda} > 0$ *is the minimum eigenvalue of* RR^T.

Proof of Lemma 5.11. We claim that $\nabla^2 D(p) - \underline{\alpha}\underline{\lambda}\, I \succeq 0$ for all $p \ge 0$. Fix any $p \ge 0$. For all $w \in \mathbb{R}^L$ we have

$$\begin{aligned} w^T \nabla^2 D(p) w &= w^T RB(p)R^T w \ge \min_i \beta_i(p)\, w^T RR^T w \ge \underline{\alpha}\, w^T RR^T w \\ &\ge \underline{\alpha}\, \underline{\lambda}\, \|w\|_2^2 \end{aligned} \tag{5.13}$$

where the second last inequality follows since $\min_i \beta_i(p) \ge \underline{\alpha}$ for all $p \ge 0$ by condition C5.4(c), and the last inequality follows since $\underline{\lambda} > 0$ is the smallest eigenvalue of the positive definite matrix RR^T. □

Lemmas 5.11 and 5.4 imply that $D(p)$ is strongly convex. Since $D(p)$ satisfies conditions C5.1-C5.3, Theorem 5.7 implies part 3 of Theorem 5.8. This completes the proof. □

5.3 APPENDIX: PROOF OF LEMMA 5.2

We estimate the difference $f(x + y) - f(x)$ by considering the *scalar* function $g(s)$ defined by the intersection of the $f(x)$ surface with the vertical plane at x in the direction y. Fix any $x, y \in \mathbb{R}^n$ and define

$$g(s) := f(x + sy) \qquad \text{for } s \in [0, 1]$$

Then

$$f(x+y) - f(x) \quad = \quad g(1) - g(0) \quad = \quad \int_0^1 g'(s)ds$$

Using

$$g'(s) \quad = \quad y^T \nabla f(x+sy)$$

we have

$$\begin{aligned} f(x+y) - f(x) &= \int_0^1 y^T \nabla f(x+sy)ds \\ &= \int_0^1 \left(y^T \nabla f(x) \ + \ y^T \left(\nabla f(x+sy) - \nabla f(x)\right)\right) ds \\ &\leq y^T \nabla f(x) \ + \ \int_0^1 \|y\|_2 \ \|\nabla f(x+sy) - \nabla f(x)\|_2 \, ds \\ &\leq y^T \nabla f(x) \ + \ \|y\|_2 \int_0^1 K \, \|sy\|_2 \, ds \\ &= y^T \nabla f(x) \ + \ \frac{K}{2} \, \|y\|_2^2 \end{aligned}$$

where the first inequality follows from the Cauchy-Schwarz inequality and the second inequality follows from condition C5.2. This proves the descent lemma. □

5.4 APPENDIX: PROOF OF LEMMA 5.4

We first prove the following characterization of convex functions.

Lemma 5.12 *Consider a differentiable function* $f : \mathbb{R}^n \to \mathbb{R}$.

1. f *is convex if and only if for all* $x, y \in \mathbb{R}^n$

$$f(y) \quad \geq \quad f(x) + \frac{\partial f}{\partial x}(x)\,(y-x) \tag{5.14}$$

2. f *is strictly convex if and only if strictly inequality holds in* (5.14) *for* $x \neq y$.

Proof of Lemma 5.12. We first prove the result for a scalar differentiable convex function $g : \mathbb{R} \to \mathbb{R}$. Then we extend the result to scalar differentiable strictly convex function g. Finally we use these results to prove the lemma for a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$.

Step 1: convex $g : \mathbb{R} \to \mathbb{R}$. Suppose g is differentiable. We now prove that the following are equivalent:

(a) g is convex.

(b) $g(t) - g(s) \geq g'(s)(t-s)$ for any $s \neq t \in \mathbb{R}$.

(c) $g'(t) \geq g'(s)$ for any $t \geq s$ in $\mathbb{R}$, i.e. g has nondecreasing slope.

Suppose (a): g is convex. Fix any $s, t \in \mathbb{R}$. For any $\alpha \in [0,1]$ we have $g(s + \alpha(t-s)) \leq (1-\alpha)g(s) + \alpha g(t)$ and hence

$$g(t) - g(s) \quad \geq \quad \frac{g(s + \alpha(t-s)) - g(s)}{\alpha}$$

Taking limit

$$\lim_{\alpha \downarrow 0} \frac{g(s + \alpha(t-s)) - g(s)}{\alpha(t-s)} \, (t-s) \quad = \quad g'(s) \, (t-s)$$

we have (b). Conversely suppose (b) and we want to prove (a), i.e.

$$\alpha g(t) + (1-\alpha) g(s) \; - \; g(z) \quad \geq \quad 0 \tag{5.15}$$

for any $z := s + \alpha(t-s)$, $\alpha \in [0,1]$. Compare the difference $g(t) - g(z)$ and $g(s) - g(z)$ in terms of gradient at the common point z:

$$g(t) - g(z) \; \geq \; g'(z)(t-z) \quad \text{ and } \quad g(s) - g(z) \; \geq \; g'(z)(s-z)$$

To obtain (5.15), multiply the first inequality by α and the second inequality by $1-\alpha$ and sum, noting that $t - z = (1-\alpha)(t-s)$ and $s - z = -\alpha(t-s)$ so that the right-hand sides of these two inequalities sum to zero. This proves (a) $\Leftrightarrow$ (b).

Now suppose (b). Fix any $t \geq s$ and compare $g(t) - g(s)$ in terms of slope at s and at t:

$$g'(s)(t-s) \quad \leq \; g(t) - g(s) \; \leq \quad g'(t)(t-s)$$

yielding (c). Conversely suppose (c) and fix any $t \geq s$. By the mean value theorem we have, for some $z \in [s,t]$, $g(t) - g(s) \; = \; g'(z)(t-s) \; \geq \; g'(s)(t-s)$, which is (b). This proves (b) $\Leftrightarrow$ (c).

Step 2: strictly convex g. We claim that the following are equivalent:

(a') g is strictly convex.

(b') $g(t) - g(s) > g'(s)(t-s)$ for any $s, t \in \mathbb{R}$.

(c') $g'(t) > g'(s)$ for any $t > s$ in $\mathbb{R}$, i.e. g has strictly increasing slope.

Three of the four proof steps above go through with the only change that inequalities become strict, except the step proving (a) $\Rightarrow$ (b) where strict inequality is generally not preserved when we take the limit $\alpha \downarrow 0$. Hence we have (a') $\Leftarrow$ (b') $\Leftrightarrow$ (c').

We now prove (a') $\Rightarrow$ (c'). Suppose g is strictly convex but there exist $s < t$ with $g'(s) = g'(t)$. But g being convex means $g'(s) \leq g'(z) \leq g'(t)$ for all $z \in [s,t]$ from Step 1. Hence g is a straight line on $[s,t]$ with $g'(z) = g'(s)$ for all $z \in [s,t]$, contradicting strict convexity. This proves (c').

Step 3: $f : \mathbb{R}^n \to \mathbb{R}$. We now use the results above on scalar functions to prove the lemma. Suppose f is convex and fix any $x, y \in \mathbb{R}^n$. Define the function $g : \mathbb{R} \to \mathbb{R}$ by

$$g(s) \quad := \quad f(x + sy)$$

as a function of the scalar $s \in \mathbb{R}$. Then it is easy to show that $g(s)$ is convex. Indeed, this is Theorem 2.6.3 and it says that f is (strictly) convex on $\mathbb{R}^n$ if and only if any of its one-dimensional cross section is (strictly) convex. By the mean value theorem there exists an $s \in [0,1]$ such that

$$f(x+y) - f(x) \quad = \quad g(1) - g(0) \quad = \quad g'(s)$$

By (c) above we have $g'(s) \geq g'(0) = (\nabla f(x))^T y$ and hence

$$f(x+y) - f(x) \quad \geq \quad (\nabla f(x))^T y$$

establishing (5.14). Moreover if f is strictly convex then the inequalities above are strict.

Conversely suppose (5.14) holds. To prove the convexity of f, use the same proof above for (b) $\Rightarrow$ (a). Take $z := x + \alpha(y - x)$ for any $\alpha \in [0,1]$. We have

$$f(y) - f(z) \;\geq\; (\nabla f(z))^T (y - z) \qquad \text{and} \qquad f(x) - f(z) \;\geq\; (\nabla f(z))^T (x - z)$$

Multiply the first inequality by α and the second inequality by $1 - \alpha$ and sum to obtain:

$$\alpha f(y) + (1-\alpha) f(x) \;-\; f(z) \quad \geq \quad (\nabla f(z))^T (\alpha(y - z) - (1-\alpha)(z - x)) \quad = \quad 0$$

proving the convexity of f. Moreover if the inequalities above are strict then f is strictly convex. □

Proof of Lemma 5.4. We first use Lemma 5.12.2 to prove that if f satisfies condition C5.3 then f is strictly convex. As in the proof of Lemma 5.3, fix any $x, y \in \mathbb{R}^n$ and consider the (scalar) function along the path from x to y:

$$g(s) \quad := \quad f(x + sy) \qquad \text{for } s \in [0,1]$$

Then

$$\begin{aligned} f(x+y)-f(x) &= \int_0^1 g'(s)ds \;=\; \int_0^1 y^T \nabla f(x+sy)ds \\ &= \int_0^1 \left(y^T \nabla f(x) \;+\; y^T \left(\nabla f(x+sy) - \nabla f(x)\right)\right) ds \\ &\geq y^T \nabla f(x) \;+\; \int_0^1 \frac{1}{s} \alpha \, \|sy\|_2^2 \; ds \\ &= y^T \nabla f(x) \;+\; \frac{\alpha}{2} \, \|y\|_2^2 \end{aligned}$$

where the inequality follows from C5.3. Since $\alpha > 0$, Lemma 5.12.2 implies the strict convexity of f.[4]

We now show that if $\nabla^2 f(x) \succeq \alpha I$ for all $x \in \mathbb{R}^n$ then f is strongly convex, i.e., f satisfies C5.3. Fix any x, y and let

$$h(s) \;:=\; \nabla f(x + s(y-x))^T (y-x)$$

Then

$$h'(s) \;=\; (y-x)^T \, \nabla^2 f(x + s(y-x)) \, (y-x)$$

and

$$\begin{aligned} \left(\nabla f(y) - \nabla f(x)\right)^T (y-x) &= h(1) - h(0) \;=\; \int_0^1 h'(s)ds \\ &= \int_0^1 (y-x)^T \, \nabla^2 f(x + s(y-x)) \, (y-x) \; ds \\ &\geq \alpha \, \|y-x\|_2^2 \end{aligned}$$

where the inequality follows from $\nabla^2 f(x) \succeq \alpha I$. Hence $f(x)$ is strongly convex.

Conversely suppose f is strongly convex. To estimate $\nabla^2 f(x)$ we have for any $x, y \in \mathbb{R}^n$

$$\begin{aligned} y^T \, \nabla^2 f(x) \, y &= \lim_{\lambda \to 0} \frac{1}{\lambda} \left(\frac{\partial f}{\partial x}(x + \lambda y) - \frac{\partial f}{\partial x}(x) \right) y \\ &\geq \lim_{\lambda \to 0} \frac{1}{\lambda^2} \left(\alpha \, \|\lambda y\|_2^2 \right) \;=\; \alpha \, \|y\|_2^2 \end{aligned}$$

where the inequality follows from the strong convexity of f. Hence $\nabla^2 f(x) \succeq \alpha I$ as desired. This completes the proof of Lemma 5.4. □

[4] If f satisfies both C5.2 (Lispschitz ∇f with parameter K) and C5.3 (strong convexity with parameter α) then the proof here and that of Lemma 5.2 show that

$$y^T \, \nabla f(x) + \frac{\alpha}{2} \|y\|_2^2 \;\leq\; f(x+y) - f(x) \;\leq\; y^T \, \nabla f(x) + \frac{K}{2} \|y\|_2^2$$

5.5 BIBLIOGRAPHICAL NOTES

Materials in Chapter 5.1 on convergence theorems are mainly taken from [8]. The application of these results to proving the convergence of the dual algorithm (Theorems 5.8.1 and 5.8.2) is from [35]; the extension to gradient projection algorithm as a contraction mapping (Theorem 5.8.3) is new. We only consider the synchronous case here, but the results extend to the asynchronous case where sources and links may update at different times with different frequencies and where packets may be lost, experience different delays, or arrive out of order; see [35].

CHAPTER 6

Local stability with delay

In Chapters 3 to 5 we study global stability of congestion control algorithms assuming there is no feedback delay, i.e., when a source changes its sending rate the effect immediately reaches all the links in its path, and when a link updates its congestion price the new price is immediately sensed at all sources using that link. This is of course unrealistic. In this chapter we incorporate feedback delay into our model and study the stability of the closed-loop system. Global stability of nonlinear systems in the presence of feedback delay is generally very difficult. Our goal is more modest: we will study local stability around the equilibrium point in the presence of feedback delay using a linearized model.

6.1 LINEAR MODEL WITH FEEDBACK DELAY

We are interested in the stability of the primal algorithm or the dual algorithm specified as follows:

$$\begin{aligned} \text{Primal algorithm:} \quad \dot{x} &= f(x(t), q(t)), \qquad p = g(y(t)) && \text{(6.1a)} \\ \text{Dual algorithm:} \quad \dot{p} &= g(y(t), p(t)), \qquad x = f(q(t)) && \text{(6.1b)} \end{aligned}$$

where

$$q(t) \;:=\; R^T p(t) \quad \text{and} \quad y(t) \;:=\; Rx(t) \tag{6.1c}$$

Hence there is dynamics either in the sources or in the links, but not both. Suppose (x^*, p^*) is an equilibrium point for the primal or dual algorithm, i.e.

$$\begin{aligned} \text{Primal algorithm:} \quad 0 &= f(x^*, q^*), \qquad p^* = g(y^*) \\ \text{Dual algorithm:} \quad 0 &= g(y^*, p^*), \qquad x^* = f(q^*) \end{aligned}$$

where $q^* = R^T p^*$ and $y^* = Rx^*$. In terms of the perturbation variables

$$\tilde{x}(t) \;:=\; x(t) - x^* \quad \text{and} \quad \tilde{p}(t) \;:=\; p(t) - p^*$$

the linearized model of the primal algorithm (6.1a) around the equilibrium (x^*, p^*) is:

$$\dot{\tilde{x}} \;=\; \frac{\partial f}{\partial x}(x^*, q^*)\,\tilde{x}(t) + \frac{\partial f}{\partial q}(x^*, q^*)\,\tilde{q}(t), \qquad \tilde{p} \;=\; \frac{\partial g}{\partial y}(y^*)\,\tilde{y}(t) \tag{6.2a}$$

The linearized model of the dual algorithm (6.1b) around the equilibrium (x^*, p^*) is:

$$\dot{\tilde{p}} \;=\; \frac{\partial g}{\partial y}(y^*, p^*)\,\tilde{y}(t) + \frac{\partial g}{\partial p}(y^*, p^*)\,\tilde{p}(t), \qquad \tilde{x} \;=\; \frac{\partial f}{\partial q}(q^*)\,\tilde{q}(t) \tag{6.2b}$$

Note that the updates in (6.2) only involve local variables and hence the equations contain no feedback delay. The feedback delay is modeled by replacing (6.1c) with:

$$\tilde{q}_i(t) \;:=\; \sum_l R_{li}\,\tilde{p}_l(t - \tau_{li}^b) \quad \text{and} \quad \tilde{y}_l(t) \;:=\; \sum_i R_{li}\,\tilde{x}_i(t - \tau_{li}^f)$$

or in vector form:

$$\tilde{q}(t) \;:=\; R^T\,\tilde{p}(t - \tau^b) \quad \text{and} \quad \tilde{y}(t) \;:=\; R\,\tilde{x}(t - \tau^f) \tag{6.2c}$$

Here τ_{li}^b models the backward delay from link l to source i and means that a change in link price p_l at time $t - \tau_{li}^b$ is sensed at source i only at time t. Similarly τ_{li}^f models the forward delay from source i to link l and means that a change in sending rate x_i at time $t - \tau_{li}^f$ affects the input traffic at link l only at time t. These feedback delays are assumed constant. We will make use of a key assumption that for all sources i

$$\tau_{li}^f + \tau_{li}^b \;=\; \tau_i \qquad \text{for all links } l \text{ in } i\text{'s path}$$

i.e., for each source i, the sum of forward delay to a link in its path and the backward delay from that link is the same for *every* link in its path. This sum is the round-trip time τ_i and depends only on the source i. This assumption holds, e.g., if all the data packets and their ack packets of a TCP connection follow the same path and the congestion level on the path remains unchanged over the timescale of interest.

We are interested in the stability of the linear delayed system (6.2). To this end we will first explain the Nyquist stability theory of linear time-invariant system with delay and then apply the theory to study the local stability of (6.2).

6.2 NYQUIST STABILITY THEORY

We will model congestion control algorithms as a feedback system with time delay consisting of interconnection of component systems each characterized by a transfer function. An example is shown in Figure 6.1 which is analyzed in Section 6.3 below. We will derive sufficient conditions in terms of protocol parameters and network delays that guarantee the asymptotic stability of the closed-loop system. The main tool is the Nyquist stability theory that expresses the stability condition of a closed-loop system in terms of properties of its open-loop component systems. For us these properties correspond to the design of the TCP and AQM algorithms and their parameters. In this section we first summarize various concepts and linear system results that build up to the Nyquist stability criterion for linear time-invariant (LTI) multi-input-multi-output systems without time delay. Then we extend the stability criterion to LTI multi-input-multi-output system with time delay.

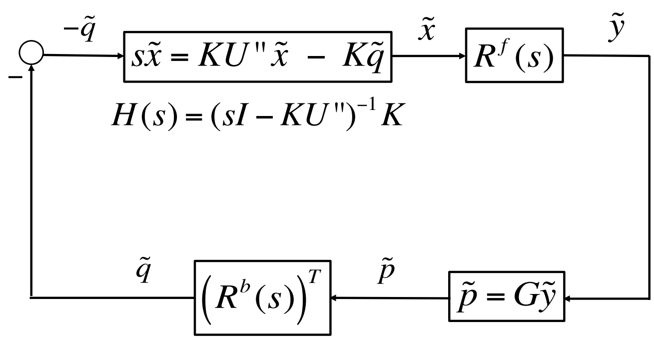


Figure 6.1: A delayed linear system model of a congestion control algorithm in Section 6.3.

6.2.1 LTI SYSTEMS, TRANSFER FUNCTIONS AND REALIZATIONS

An multi-input multi-output (MIMO) LTI system can be specified by its input-output behavior through a function (operator) that maps an input signal to an output signal. Such a function can be realized by a state-space implementation. In this subsection we summarize these concepts.

We start with the input-output description. An m-input p-output LTI system can be specified by an operator h that maps an input signal u in an appropriate space to an output signal y in an appropriate space according to:[1]

$$y(t) \quad = \quad \int_0^t h(\tau)u(t-\tau)d\tau, \qquad t \geq 0$$

Here, for each $t \geq 0$, the matrix $h(t) \in \mathbb{R}^{p\times m}$, input $u(t) \in \mathbb{R}^m$, and output $y(t) \in \mathbb{R}^p$. Note that the output $y(t)$ at time t depends not just on the input value $u(t)$ at time t, but the *function* $u(\tau)$ over $[0, t]$. Indeed the output $y(t)$ is the weighted sum (integral) of the past input $u(t-\tau)$, with input that is τ time in the past weighted by the matrix $h(\tau)$. The function h is called an *impulse response* of the LTI system because of the following reason. Consider an input signal with a single impulse in the jth component, i.e., $u(t) := \delta(t)e_j$ where $\delta(t)$ is the delta function and e_j is the unit vector with an 1 at the jth place and 0 everywhere else. The ith component $y_i(t)$ of the output at time t is:

$$y_i(t) \quad = \quad \int_0^t h_{ij}(\tau)\delta(t-\tau)d\tau \quad = \quad \int_0^t h_{ij}(\tau)\delta(\tau-t)d\tau \quad = \quad h_{ij}(t)$$

i.e., $h_{ij}(t)$ is the system response (output) $y_i(t)$ to an impulse at input u_j.

An equivalent specification of the LTI system is the Laplace transform matrix $H(s) \in \mathbb{C}^{p\times m}$ of its impulse response $h(t)$. We call $H(s)$ its *transfer function* (or transfer function matrix). It maps the Laplace transform $u(s) \in \mathbb{C}^m$ of the input $u(t)$ to the Laplace transform $y(s) \in \mathbb{C}^p$ of its output $y(t)$ over its region of convergence:[2]

$$y(s) \quad = \quad H(s)u(s)$$

We will model below congestion control by interconnection of transfer functions. Henceforth we will focus on transfer functions $H(s)$ rather than its impulse response $h(t)$.

[1] In general

$$y(t) \quad = \quad \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau, \qquad t \geq 0$$

If the LTI system is *causal*, i.e., $h(t) = 0$ for $t < 0$, then $y(t) = \int_0^\infty h(\tau)u(t-\tau)d\tau$. If the input signal $u(t)$ is also *causal*, i.e., $u(t) = 0$ for $t < 0$, then $y(t) = \int_0^t h(\tau)u(t-\tau)d\tau$.

[2] We abuse notation to use u, y to denote both the signals in the time domain and their Laplace transforms.

A transfer function $H(s)$ can be realized by a state-space implementation of the form:

$$\dot{x} = Ax(t) + Bu(t), \qquad t \geq 0 \tag{6.3a}$$
$$y = Cx(t) + Du(t), \qquad t \geq 0 \tag{6.3b}$$

where $x(t) \in \mathbb{R}^n$ is the *state* of the system, and $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, $D \in \mathbb{R}^{p \times m}$ are real matrices. We refer to such a realization by its matrices (A, B, C, D). A transfer function $H(s)$ can be realized by different state-space implementations with different matrices (A, B, C, D). These realizations have the same input-output behavior but can have different internal behavior (see examples below). For a realization (A, B, C, D) the impulse response $h(t)$ and its transfer function $H(s)$ are respectively

$$h(t) = Ce^{At}B + D\delta(t) \quad \text{and} \quad H(s) = C(sI - A)^{-1}B + D$$

where $\delta(t)$ is the delta function. For congestion control modeling we have $D = 0$ and therefore we usually refer to realizations (A, B, C) of impulse response $h(t)$ and transfer function $H(s)$ given by

$$h(t) = Ce^{At}B \quad \text{and} \quad H(s) = C(sI - A)^{-1}B$$

Results are mostly given in the following for the case where $D = 0$, but they do not change qualitatively when D is nonzero.

Proper and strictly proper transfer functions. For an MIMO system without time delay, with $m \geq 1$ inputs and $p \geq 1$ outputs, the transfer function is an $p \times m$ complex matrix function $H(s)$ of s (for s in its region of convergence in $\mathbb{C}$). Each entry $H_{ij}(s)$ is a rational function of the form:

$$H_{ij}(s) = \frac{n_{ij}(s)}{d_{ij}(s)}$$

where $n_{ij}(s)$ and $d_{ij}(s)$ are polynomials of s without common factors (co-prime). As explained above $H_{ij}(s)$ can be interpreted as a single-input-single-output (SISO) transfer function mapping input u_j to output y_i. The rational function $H_{ij}(s)$ is called *proper* if the degree of the numerator polynomial $n_{ij}(s)$ is no more than the degree of the denominator polynomial $d_{ij}(s)$; it is called *strictly proper* if the degree of $n_{ij}(s)$ is smaller than that of $d_{ij}(s)$.[3] For our purposes we will call the transfer function matrix $H(s)$ proper if all its entries are proper rational functions and strictly proper if they are strictly proper rational functions.

[3]More generally a rational transfer function $H_{ij}(s)$ is *proper* if $H_{ij}(s)$ converges to a finite complex number as $|s| \to \infty$. It is *strictly proper* if $H_{ij}(s) \to 0$ as $|s| \to \infty$.

If a transfer function matrix $H(s)$ has a state-space realization (A, B, C, D) then $H(s) = C(sI - A)^{-1}B + D$ takes the form

$$H(s) \quad = \quad \frac{1}{\det(sI - A)} C \operatorname{adj}(sI - A) B \; + \; D$$

where $\operatorname{adj}(sI - A)$ is the adjoint matrix. The determinant $\det(sI - A)$ in the denominator is a polynomial of degree n (the coefficient of s^n is 1). All entries of the adjoint matrix $\operatorname{adj}(sI - A)$ are polynomials of degrees less than n. Hence each entry of the matrix $C\operatorname{adj}(sI - A)B$ is a polynomial of degree less than n. Therefore $H(s)$ is proper if D is nonzero and strictly proper if D is zero. In particular $H(s) \to D$ as $|s| \to \infty$. Conversely if a rational transfer function $H(s)$ is proper (i.e. each entry of the matrix $H(s)$ is proper) then there exists a state-space realization (A, B, C, D). Example 6.3 shows one way to construct a realization (A, B, C, D) for a SISO transfer function that is proper.

For a time-delayed LTI system the transfer function is infinite dimensional and its transfer function $H(s)$ is not a rational function. $H(s)$ is called *proper* if there exists $\alpha > 0$ such that each entry $H_{ij}(s)$ satsifies

$$\sup \{ |H_{ij}(s)| \mid s \in \mathbb{C}, \ \operatorname{Re} s > \alpha \} \quad < \quad \infty$$

$H(s)$ is called *strictly proper* if

$$\lim_{\alpha \to \infty} \sup \{ H_{ij}(s) \mid s \in \mathbb{C}, \ \operatorname{Re} s > \alpha \} \quad = \quad 0$$

An LTI system is causal if and only if it is proper. For instance $H(s) = e^{-\tau s}$ is a transfer function of the time-delay system $y(t) = u(t - \tau)$ and it is strictly proper.

Example 6.1 Single integrator. A single integrator $\dot{y} = u(t)$ has a transfer function $H(s) = 1/s$. Its impulse response is the step function $h(t) = 1$ for $t \geq 0$ and 0 for $t < 0$. It can be realized by the following state-space implementation:

$$\dot{x} \; = \; u(t), \qquad y(t) \; = \; x(t)$$

In the Laplace domain this is $sx(s) = u(s)$ and $y(s) = x(s)$, yielding the transfer function $H(s) = 1/s$. The same transfer function can also be realized by a different implementation:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} \; = \; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} x(t) \; + \; \begin{bmatrix} 1 \\ -1 \end{bmatrix} u(t), \qquad y \; = \; \begin{bmatrix} 1 & 0 \end{bmatrix} x(t)$$

In Laplace domain this is

$$sx(s) \; = \; Ax(s) + Bu(s), \qquad y(s) \; = \; Cx(s)$$

where

$$x(s) \;:=\; \begin{bmatrix} x_1(s) \\ x_2(s) \end{bmatrix}, \qquad A \;:=\; \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \;\; B \;:=\; \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \;\; C \;:=\; \begin{bmatrix} 1 & 0 \end{bmatrix}$$

yielding the transfer function

$$H(s) \;=\; y(s)/u(s) \;=\; C(sI - A)^{-1}B \;=\; \frac{1}{s}$$

Even though both realizations have the same input-output behavior they have different internal properties. We will comment on this in Chapter 6.2.2 on minimal realization. □

Example 6.2 Single-input single-output (SISO) system. Consider the transfer function of a single-input single-output system

$$H(s) \;:=\; \frac{y(s)}{u(s)} \;=\; \frac{s-1}{(s+1)(s+2)}$$

Since $(s^2 + 3s^2 + 2)y(s) = (s-1)u(s)$, this represents the following LTI system in the time domain:

$$\ddot{y} + 3\dot{y} + 2y(t) \;=\; \dot{u} - u(t)$$

Choose a realization (A, B, C) with $n = 2$ states:

$$A \;:=\; \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \;\; B \;:=\; \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \;\; C \;:=\; \begin{bmatrix} c_1 & c_2 \end{bmatrix}$$

Then

$$\frac{s-1}{(s+1)(s+2)} \;=\; C(sI - A)^{-1}B \;=\; \frac{1}{\det(sI - A)} \begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} s - a_{22} & -a_{12} \\ -a_{21} & s - a_{11} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

where

$$\det(sI - A) \;=\; s^2 - (a_{11} + a_{22})s + (a_{11}a_{22} - a_{12}a_{21}) \;=\; (s+1)(s+2)$$

Hence $a_{11} + a_{22} = -3$, $a_{11}a_{22} - a_{12}a_{21} = 2$ and, for s not equal to the eigenvalues of A,

$$\begin{bmatrix} c_1 & c_2 \end{bmatrix} \begin{bmatrix} s - a_{22} & -a_{12} \\ -a_{21} & s - a_{11} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \;=\; s - 1$$

There are many choices of (A, B, C) that satisfy these conditions. Here is one example

$$A \;:=\; \begin{bmatrix} -4 & -2 \\ 3 & 1 \end{bmatrix}, \;\; B \;:=\; \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \;\; C \;:=\; \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{6.4}$$

with the state-space implementation:

$$
\begin{aligned}
\dot{x}_1 &= -4x_1(t) - 2x_2(t) + u(t) \\
\dot{x}_2 &= 3x_1(t) + x_2(t) \\
y(t) &= x_1(t)
\end{aligned}
$$

□

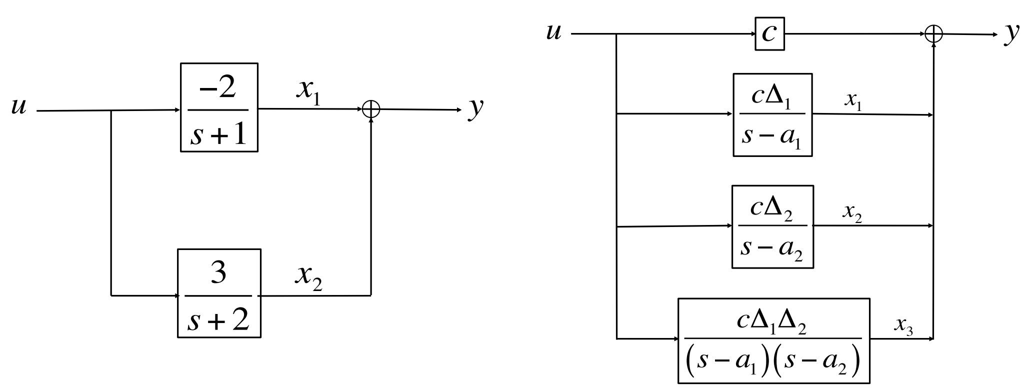
Example 6.3 Proper transfer functions. Consider again the transfer function in Example 6.2:

$$
H(s) \quad := \quad \frac{y(s)}{u(s)} \quad = \quad \frac{s-1}{(s+1)(s+2)}
$$

Another method to design a state-space realization is observe that the transfer function $H(s)$ is strictly proper, i.e., the degree of the numerator is less than that of the denominator. Then we can always write $H(s)$ as two integrators:

$$
H(s) \quad = \quad \frac{s-1}{(s+1)(s+2)} \quad = \quad \frac{-2}{s+1} + \frac{3}{s+2}
$$

as shown in the block diagram in Figure 6.2(a) This can be implemented in state space as:



(a) Strictly proper transfer function

(b) Proper transfer function

Figure 6.2: Realization of the transfer functions in Example 6.3. (a) Strictly proper transfer function $H(s) = \frac{s-1}{(s+1)(s+2)}$. (b) Proper transfer function $H(s) = \frac{c(s-b_1)(s-b_2)}{(s-a_1)(s-a_2)}$.

$$
\begin{aligned}
\dot{x}_1 &= -x_1(t) - 2u(t) \\
\dot{x}_2 &= -2x_2(t) + 3u(t) \\
y(t) &= x_1(t) + x_2(t)
\end{aligned}
$$

Notice that $y(t)$ depends only on the state $(x_1(t), x_2(t))$, not directly on the input $u(t)$. This is the consequence of the transfer function $H(s)$ being *strictly* proper. When $H(s)$ is proper but not strictly proper, i.e., when the numerator and denominator of $H(s)$ have the same degree, then $y(t)$ depends on $u(t)$ directly, as the next example shows.

Consider the transfer function

$$H(s) \quad = \quad \frac{c(s-b_1)(s-b_2)}{(s-a_1)(s-a_2)}$$

where $a_i \neq b_j$, i.e., the numerator and the denominator are co-prime. Decompose the rational function:

$$H(s) \quad = \quad \frac{c((s-a_1)+\Delta_1)((s-a_2)+\Delta_2)}{(s-a_1)(s-a_2)} \quad = \quad c + \frac{c\Delta_1}{s-a_1} + \frac{c\Delta_2}{s-a_2} + \frac{c\Delta_1\Delta_2}{(s-a_1)(s-a_2)}$$

where $\Delta_1 := a_1 - b_1$ and $\Delta_2 := a_2 - b_2$. Hence $H(s)$ can be represented by the block diagram in Figure 6.2(b). The last block is second-order, corresponding to

$$\ddot{x}_3 - (a_1+a_2)\dot{x}_3 + a_1 a_2 x_3(t) \quad = \quad c\Delta_1\Delta_2 u(t)$$

and requires two states (x_3, x_4) to implement:

$$\begin{aligned} \dot{x}_3 &= x_4(t) \\ \dot{x}_4 &= (a_1+a_2)x_4(t) - a_1 a_2 x_3(t) + c\Delta_1\Delta_2 u(t) \end{aligned}$$

Hence a state-space realization of $H(s)$ is:

$$\dot{x} \;=\; Ax(t) + Bu(t), \qquad y(t) \;=\; Cx(t) + Du(t)$$

where $x(t) := (x_1(t), x_2(t), x_3(t), x_4(t))$ and

$$A \;:=\; \begin{bmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -a_1 a_2 & a_1+a_2 \end{bmatrix}, \quad B \;:=\; \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ 0 \\ \Delta_1\Delta_2 \end{bmatrix}, \quad C^T \;:=\; \begin{bmatrix} c \\ c \\ c \\ 0 \end{bmatrix}, \quad D \;:=\; c$$

□

Example 6.4 MIMO system. Consider the transfer function of a 2-input 2-output system:

$$H(s) \quad := \quad \begin{bmatrix} \frac{s-1}{(s+1)(s+2)} & \frac{s+1}{(s-1)(s+2)} \\ 0 & s^{-1} \end{bmatrix}$$

The input-output relation in the Laplace domain is

$$\begin{aligned} y_1(s) &= \frac{s-1}{(s+1)(s+2)}\, u_1(s) \;+\; \frac{s+1}{(s-1)(s+2)}\, u_2(s) \\ y_2(s) &= s^{-1}\, u_2(s) \end{aligned}$$

or equivalently

$$\begin{aligned} \left(s^3+2s^2-s-2\right) y_1(s) &= \left(s^2-2s+1\right) u_1(s) \;+\; \left(s^2+2s+1\right) u_2(s) \\ s\, y_2(s) &= u_2(s) \end{aligned}$$

This is in the time domain

$$\begin{aligned} \dddot{y}_1 + 2\ddot{y}_1 - \dot{y}_1 - 2y_1(t) &= \ddot{u}_1 - 2\dot{u}_1 + u_1(t) \;+\; \ddot{u}_2 + 2\dot{u}_2 + u_2(t) \\ \dot{y}_2 &= u_2(t) \end{aligned}$$

It can be implemented in state space using the same method in Example 6.3 (see Exercise 6.5). □

6.2.2 STABILITY OF LTI SYSTEMS

In this subsection we discuss two notions of stability and explain their relationship. The first one is associated with a state-space realization (A, B, C) and the second with an input-output description through a transfer function $H(s)$. Both will be used later in analyzing the stability of congestion control algorithms.

Given a realization (A, B, C) of an LTI system we are interested in the asymptotic stability of the origin (equilibrium) when the input is zero, i.e., the asymptotic stability of the equilibrium point of the autonomous (undriven) system:

$$\dot{x} \;=\; Ax(t), \qquad t \geq 0 \tag{6.5}$$

We say that the origin of (6.5) is *exponentially stable* if, there exist $\alpha > 0$ and $\beta \geq 0$ such that, for any initial point $x(0) := x_0$, the solution $(x(t), t \geq 0)$ of (6.5) satisfies

$$\|x(t)\| \;\leq\; \beta\, e^{-\alpha t}\, \|x_0\|, \qquad t \geq 0$$

i.e., the trajectory $x(t)$ converges exponentially (geometrically) to the origin. For LTI systems the origin is exponentially stable if and only if it is asymptotically stable as defined in Definition 3.1 of Chapter 3.[4] We will hence use either stability notion interchangeably for LTI systems.

[4]If an LTI system is asymptotically stable then it is globally asymptotically stable. We will use asymptotic stability instead of global asymptotic stability because our linear system originates from the linearization of a nonlinear system and therefore describes local behavior around an equilibrium.

Theorem 6.5 Asymptotic stability. *The origin of* (6.5) *is asymptotically stable if and only if all eigenvalues* $\lambda(A)$ *of the system matrix* A *are in the open left-half plane (LHP).*

In this case we say that the system matrix A is *stable* or *Hurwitz*.

Example 6.6 For the first realization $\dot{x} = u(t)$ of the integrator in Example 6.1 the undriven system is $\dot{x} = 0 = Ax(t)$ with $A = 0$. Its eigenvalue is 0 and hence the origin is not asymptotically stable ($x(t) = x(0)$ for all $t \geq 0$). For the second realization the undriven system $\dot{x} = Ax(t)$ has a system matrix

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

with eigenvalues 0 and 1, and hence not asymptotically stable either. For the realization (A, B, C) in Example 6.2, the eigenvalues of the system matrix A are -1 and -2. Hence the origin is asymptotically stable.

□

Asymptotic stability is an internal property about the state $x(t)$ of the system in terms of the system matrix A. In contrast the second stability notion pertains to the input-output behavior in terms of the transfer function $H(s)$, as we now explain.

For any signal $u : [0, \infty) \to \mathbb{R}^m$ we define its p-norm for $p = 1, 2, \infty$ as:

$$\begin{aligned}
\|u\|_1 &:= \int_0^\infty \|u(t)\|_1 dt &&= \int_0^\infty \sum_i |u_i(t)| dt \\
\|u\|_2 &:= \left[\int_0^\infty \|u(t)\|_2^2 dt\right]^{1/2} &&= \left[\int_0^\infty \sum_i u_i^2(t) dt\right]^{1/2} \\
\|u\|_\infty &:= \sup_{t \geq 0} \|u(t)\|_\infty &&= \sup_{t \geq 0} \max_i |u_i(t)|
\end{aligned}$$

Note that $\|u(t)\|_p$ is a norm of a vector in $\mathbb{R}^m$ while $\|u\|_p$ is norm of a signal in an appropriate function space.

Consider a causal LTI system specified by its impulse response

$$y(t) = \int_0^t h(\tau) u(t - \tau) d\tau, \qquad t \geq 0 \tag{6.6a}$$

or its transfer function

$$y(s) = H(s) u(s) \tag{6.6b}$$

Definition 6.7 The LTI system (6.6) is *L_p-stable* for $p = 1, 2, \infty$, if there exists a finite $\gamma \in \mathbb{R}$ such that

$$\|y\|_p \quad \leq \quad \gamma \, \|u\|_p$$

for *every* input u.

The smallest constant γ that satisfies the condition in Definition 6.7 is often called the system gain. Hence an LTI system is L_p-stable if an input of finite p-norm $\|u\|_p$ produces an output of finite p-norm $\|y\|_p$. In particular the case with $p = \infty$ is referred to as *bounded input bounded output (BIBO) stability*. For LTI systems of finite dimension, BIBO stability is equivalent to L_p-stability.

Consider again the following LTI system in state-space form:

$$\dot{x} \;=\; Ax(t) + Bu(t), \qquad t \geq 0 \tag{6.7a}$$
$$y \;=\; Cx(t), \qquad t \geq 0 \tag{6.7b}$$

with impulse response $h(t)$ and transfer function $H(s)$ given by

$$h(t) \;=\; Ce^{At}B \quad \text{and} \quad H(s) \;=\; C(sI - A)^{-1}B \tag{6.7c}$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$. Before we characterize BIBO stability in terms of the transfer function $H(s)$, we need to define poles of a transfer function matrix.

The transfer function of a SISO system (6.7) with $p = m = 1$ is a rational function

$$H(s) \;=\; \frac{C\,\mathrm{adj}(sI - A)\,B}{\det(sI - A)} \;=\; \frac{\prod_{k=1}^{k_1}(s - z_k)}{\prod_{k=1}^{k_2}(s - \lambda_k)} \tag{6.8}$$

where the set of λ_k is a subset of the eigenvalues of A, counting multiplicity. As discussed in Chapter 6.2.1, the transfer function is strictly proper ($D = 0$), i.e. $k_1 < k_2 \leq n$. We assume that the numerator polynomial and the denominator polynomial are co-prime, i.e., all common factors $s - z_k = s - \lambda_{k'}$ have been canceled (pole-zero cancellation). We call a root λ_k of its denominator polynomial a *pole* of the transfer function $H(s)$, and a root z_k of its numerator polynomial a *zero* of $H(s)$.

For an MIMO system (6.7) without time delay, with $m \geq 1$ inputs and $p \geq 1$ outputs, the transfer function is an $p \times m$ complex matrix function $H(s) = C(sI - A)^{-1}B$ where each entry is a rational function of the form in (6.8). Let

$$H_{ij}(s) \;=\; \frac{n_{ij}(s)}{d_{ij}(s)}$$

where $n_{ij}(s)$ and $d_{ij}(s)$ are co-prime, i.e., all common factors have been canceled.

Definition 6.8 The transfer function matrix $H(s) = C(sI - A)^{-1}B$ has a pole at $s = \lambda$ if *some* of its entries of $H_{ij}(s)$ has a pole at $s = \lambda$, i.e., $d_{ij}(\lambda) = 0$ for some ij.

From (6.8) we have

$$\{\,\text{poles of } H(s)\,\} \quad \subseteq \quad \{\,\text{eigenvalues of } A\,\} \tag{6.9}$$

There may be eigenvalues of A that are not poles of $H(s)$ because of pole-zero cancellation. The definition of zeros of a transfer function matrix $H(s)$ is subtle. Roughly $H(s)$ has a zero at $s = z$ if the rank of $H(s)$ drops at $s = z$. This is complicated by the fact that $s = z$ can be both a zero and a pole of $H(s)$.

The next result characterizes L_p-stability and BIBO stability in terms of the poles of the transfer function $H(s)$ and eigenvalues of the system matrix A. It implies in particular that, for finite dimensional LTI system, BIBO stability is equivalent to L_p-stability.

Theorem 6.9 **L_p-stability.** *Consider the finite dimensional LTI system* (6.7) *with transfer function* $H(s)$ *and system matrix* A*.*

1. *For any* $p = 1, 2, \infty$*, the system* (6.7) *is* L_p*-stable if and only if all the poles of* $H(s)$ *are in the open LHP.*

2. *If the origin of the system* (6.7) *is asymptotically stable then the system is BIBO stable.*

By Theorem 6.5 the origin is asymptotically stable if and only if all the eigenvalues of A are in the open LHP. This implies Theorem 6.9.2. As explained above (see (6.8)), the set of poles of $H(s)$ can be a strict subset of the eigenvalues of A when there is pole-zero cancellation. This is why the converse of Theorem 6.9.2 generally does not hold, i.e., there are systems that are BIBO stable but their realizations are not asymptotically stable. We now describe an important class of realizations (A, B, C) where the poles of $H(s)$ coincide with the eigenvalues of A. Such a system is BIBO stability if and only if its origin is asymptotically stable.

We will call a realization (A, B, C) *controllable* if the rank of the following controllability matrix

$$[B \;\; AB \;\; A^2B \;\; \cdots \;\; A^{n-1}B]$$

is n, and *observable* if the rank of the following observability matrix

$$\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is n. If (A, B) is controllable then, given any initial state x_0, any target state x_1, and any time interval $[t_0, t_1]$, there exists an input $u(t)$ over $[t_0, t_1]$ that will drive the state from x_0 at time t_0 to x_1 at time t_1. If (A, C) is observable then any initial state x_0 can be uniquely determined from input/output measurements (i.e., measurements of $u(t), y(t)$ and their derivatives) over any finite time interval. If (A, B, C) is both controllable and observable then it is called a *minimal realization.* A minimal realization has no internal modes that are uncontrollable by applying an appropriate input nor internal modes that are unobservable from input/output measurements. In this case $H(s)$ has no pole-zero cancellation and instead of (6.9) we have

$$\{\text{poles of } H(s)\} \quad = \quad \{\text{eigenvalues of } A\}$$

We hence have the following corollary of Theorem 6.9.

Corollary 6.10 Stability of minimal realization. *Consider the finite dimensional LTI system* (6.7) *with transfer function* $H(s)$ *and a realization* (A, B, C). *If the realization is minimal, then the following are equivalent:*

1. *For any* $p = 1, 2, \infty$, *the system* (6.7) *is* L_p*-stable (including BIBO stability).*
2. *The origin of the system* (6.7) *is asymptotically stable.*
3. *All the poles of* $H(s)$ *are in the open LHP.*
4. *All eigenvalues of* A *are in the open LHP.*

Eigenvector dyadic expansion of $H(s)$ with diagonalizable A. We close this subsection by exhibiting the particularly simple structure of $H(s)$ and its poles in the special case where A is diagonalizable. A complex matrix $M \in \mathbb{C}^{n\times n}$ is called *diagonalizable* if it is similar to a diagonal matrix $\Lambda = \text{diag}(\lambda_i)$, i.e., there exists a nonsingular matrix V such that

$$V^{-1}MV \quad = \quad \Lambda$$

From $MV = V\Lambda$ we see that λ_i are eigenvalues of M with the columns v_i of V as the corresponding eigenvectors. Hence M is diagonalizable if and only if it has an eigenbasis, i.e. a basis consisting of eigenvectors.

Lemma 6.11 *Suppose* $M \in \mathbb{C}^{n\times n}$ *is diagonalizable. Then*

1. M *has a basis of right eigenvectors* $(v_1, \ldots, v_n)$ *such that* $Mv_i = \lambda_i v_i$. *Let* $V \in \mathbb{C}^{n\times n}$ *with* v_i *as its columns.*

2. *M has a basis of left eigenvectors $(u_1, \ldots, u_n)$ such that $u_i^* M = \lambda_i u_i^*$. Let $U \in \mathbb{C}^{n\times n}$ with u_i as its columns.*

3. *The left and right eigenvectors are orthogonal, i.e. $u_i^* v_j = 0$ if $i \neq j$ and 1 if $i = j$ or more compactly $U^* V = I$.*

4. *The identity matrix I has the eigenvector dyadic expansion $I = \sum_i v_i u_i^* = VU^*$.*

5. *The matrix M has the eigenvector dyadic expansion*

$$M = \sum_i \lambda_i v_i u_i^* = V\Lambda U^*$$

6. *If M is nonsingular then $\lambda_i \neq 0$ and*

$$M^{-1} = \sum_i (\lambda_i)^{-1} v_i u_i^* = V\Lambda^{-1} U^*$$

7. *Every vector $x \in \mathbb{C}^n$ can be expressed in terms of the basis V as:*

$$x = \sum_i v_i u_i^* x = \sum_i a_i v_i$$

where $a_i = u_i^* x$ is the projection of x onto the left eigenvector space. Every $y = Mx$ can be expressed in terms of V as:

$$y = M\left(\sum_i a_i v_i\right) = \sum_i \lambda_i a_i v_i$$

In particular Lemmas 6.11.3–6.11.4 imply that V and U^* are inverses of each other. When its eigenvectors do not form a basis, a matrix M is not diagonalizable. It can still be transformed into a block diagonal matrix called the Jordan form. Indeed any matrix $M \in \mathbb{C}^{n\times n}$ can be uniquely written as

$$M = M_1 + M_2$$

such that M_1 is diagonalizable and M_2 is nilpotent.[5] Moreover M and M_1 have identical eigenvalues and M_1, M_2 commute.

We now apply these properties to transfer function $H(s) := C(sI - A)^{-1}B$, for the case where $A \in \mathbb{R}^{n\times n}$ is diagonalizable with eigenvalues $\lambda_i \in \mathbb{C}$ and the corresponding left and right eigenvectors (u_i, v_i), $i = 1, \ldots, n$. Then $(sI - A)$ is diagonalizable and has

[5] A matrix A is nilpotent if A^k is the zero matrix for some finite positive integer k.

eigenvalues $(s - \lambda_i)$. Moreover, from Lemma 6.11.4–6, $(sI - A)^{-1}$ has a unique eigenvector dyadic expansion (V and U^* are inverses of each other):

$$(sI - A)^{-1} \quad = \quad (sVU^* - V\Lambda U^*)^{-1} \quad = \quad V(sI - \Lambda)^{-1}U^* \quad = \quad \sum_i (s - \lambda_i)^{-1}\, v_i u_i^*$$

Hence

$$H(s) \quad = \quad \sum_i (s - \lambda_i)^{-1}\, (Cv_i) \cdot (u_i^* B)$$

Since $(Cv_i) \cdot (u_i^* B)$ is a dyad with column Cv_i and row $u_i^* B$, $(s - \lambda_i)^{-1}$ disappears from the sum if and only if either Cv_i is the zero column vector or $u_i^* B$ is the zero row vector.

Theorem 6.12 *Consider the transfer function $H(s) := C(sI - A)^{-1}B$ and suppose $A \in \mathbb{R}^{n\times n}$ is diagonalizable with eigenvalues $\lambda_i \in \mathbb{C}$ and the corresponding left and right eigenvectors (u_i, v_i), $i = 1, \ldots, n$. Then $\lambda \in \mathbb{C}^n$ is a pole of $H(s)$ if and only if all the following three conditions hold:*

1. *λ is an eigenvalue of A, i.e., $\lambda = \lambda_i$ for some i.*
2. *Cv_i is not the zero column vector.*
3. *$u_i^* B$ is not the zero row vector.*

6.2.3 FEEDBACK SYSTEMS AND LOOP FUNCTIONS

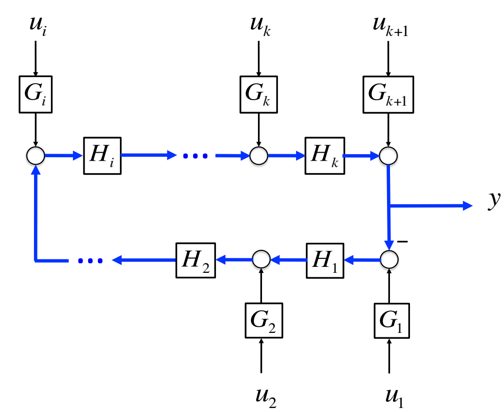
We are interested in the stability of feedback connection of component dynamical systems. Nyquist stability theory provides us with a method to assess the stability of the closed-loop system in terms of the stability of the component systems in open loop. A central concept is the loop function, which we now explain.

Consider a general feedback loop, as shown in Figure 6.3(a), consisting of the interconnection of k LTI systems (A_i, B_i, C_i) with corresponding transfer functions

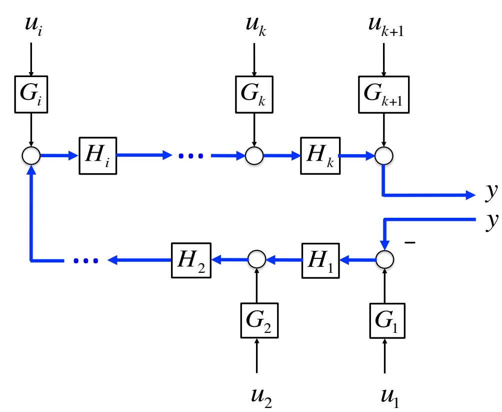
$$H_i(s) \quad := \quad C_i(sI - A_i)^{-1}B_i, \qquad i = 1, \ldots, k$$

Inputs u_i are inserted through a transfer function G_i at the input of H_i; these u_i can represent control inputs, reference signals, disturbances, or measurement noises. Since the system is LTI, the effect on the output y of all inputs $u_1, u_2, \ldots, u_k, u_{k+1}$ is the superposition of the effects of individual inputs when all other inputs are zero (going round the loop from the output y towards the inputs in the opposite direction of the signal flow; see Figure 6.3):

$$H_k \cdots H_1\, G_1 u_1,\;\; H_k \cdots H_2\, G_2 u_2,\;\; \ldots,\;\; H_k G_k u_k,\;\; G_{k+1} u_{k+1}$$



(a) Feedback system



(b) Loop function

Figure 6.3: Loop function of a feedback system.

To write down the effect of the output feedback, we break the loop at the output y to obtain an "open-loop" system shown in Figure 6.3(b). The effect on y of the output feedback is (negative sign for negative feedback):

$$-H_k \cdots H_1\, y$$

By superposition of linear systems we have

$$y \;=\; -\underbrace{H_k \cdots H_1}_{\text{loop function } L}\, y \;+\; \sum_{i=1}^{k} H_k \cdots H_i G_i\, u_i \;+\; G_{k+1}\, u_{k+1}$$

Hence the transfer functions from inputs u_i to output y can be read off from

$$y \;=\; (I+L)^{-1} \sum_{i=1}^{k} H_k \cdots H_i G_i\, u_i \;+\; (I+L)^{-1} G_{k+1}\, u_{k+1} \tag{6.10a}$$

where the *loop function* is:

$$L \;:=\; H_k \cdots H_1 \tag{6.10b}$$

It will play a critical role in Nyquist stability method as we will see below.[6] In particular the transfer function H_{yu_i} from input u_i to output y is

$$H_{yu_i} \;=\; (I+L)^{-1}\, H_k \cdots H_i\, G_i, \qquad i = 1, \ldots, k \tag{6.10c}$$

Example 6.13 **Block diagram manipulations.** In this example we find the transfer function H_{yu} from u to y and the loop function of the feedback system in Figure 6.4(a). Break the loop at the output y, as shown in Figure 6.4(b), and denote the transfer function of the local feedforward loop from u_1 to y_1 by H_1 and that of the local feedback loop from u_3 to y_3 by H_3. Applying superposition to the local loops we have

$$H_1 \;=\; G_1 - G_2 \quad \text{and} \quad H_3 \;=\; (I + G_3 G_4)^{-1} G_3$$

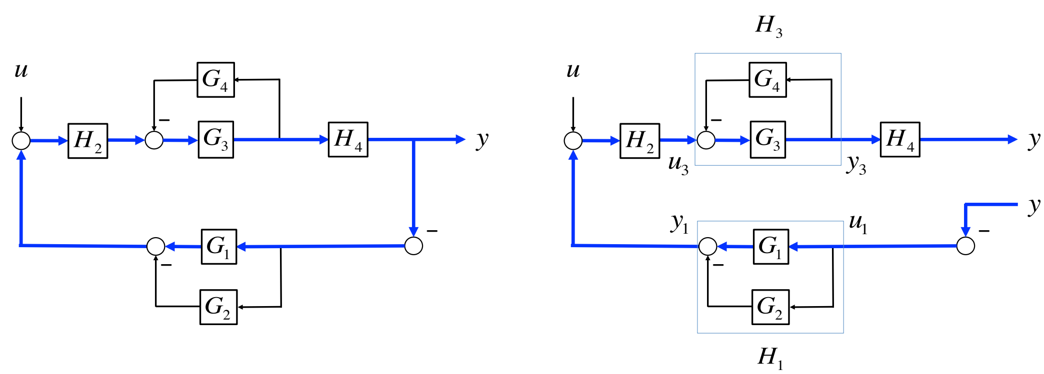
Applying again (6.10) we find the loop function to be:

$$L \;=\; H_4 H_3 H_2 H_1 \;=\; H_4\, (I + G_3 G_4)^{-1} G_3\, H_2\, (G_1 - G_2)$$

Hence the transfer function from u to y is

$$H_{yu} \;=\; (I+L)^{-1}\, (H_4 H_3 H_2) \;=\; (I+L)^{-1} \left(H_4\, (I + G_3 G_4)^{-1} G_3\, H_2 \right)$$

[6]The transfer function from $G_{k+1}u_{k+1}$ to y is $S := (I+L)^{-1}$ and is called the *sensitivity function*. The transfer function from $G_1 u_1$ to y is $T := (I+L)^{-1} L$ and is called the *complementary sensitivity function* since $S + T = I$. S and T determine the dynamic performance of the closed-loop system; see e.g. [43].



(a) Feedback system

(b) Loop function

Figure 6.4: Example 6.13.

6.2.4 STABILITY OF CLOSED-LOOP SYSTEMS

Consider the feedback connection in Figure 6.5 and let the component systems be

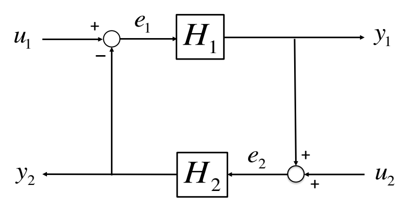


Figure 6.5: Feedback connection.

$$\dot{x}_i = A_i x_i(t) + B_i e_i(t), \quad y_i = C_i x_i(t), \qquad i = 1, 2 \tag{6.11a}$$

where $x_i(t) \in \mathbb{R}^{n_i}$, $e_i(t) \in \mathbb{R}^{m_i}$, $y_1(t) \in \mathbb{R}^{m_2}$, $y_2(t) \in \mathbb{R}^{m_1}$, $A_i \in \mathbb{R}^{n_i \times n_i}$, $B_1 \in \mathbb{R}^{n_1 \times m_1}$, $B_2 \in \mathbb{R}^{n_2 \times m_2}$, $C_1 \in \mathbb{R}^{m_2 \times n_1}$, and $C_2 \in \mathbb{R}^{m_1 \times n_2}$. The external inputs $u_i(t)$ are thus in $\mathbb{R}^{m_i}$. The transfer functions of the component systems are

$$H_i(s) = C_i(sI - A_i)^{-1} B_i, \qquad i = 1, 2 \tag{6.11b}$$

We sometimes refer to $H_i(s)$ as the *open-loop systems* and $\det(sI - A_i)$ as the *open-loop characteristic polynomials*. Motivated by Theorem 6.5, we say that the open-loop system $H_i(s)$ is *open-loop stable* if all roots of $\det(sI - A_i)$ are in the open LHP, i.e., the open-loop system matrix A_i is stable. This implies that $H_i(s)$ is BIBO stable (Theorem 6.9); the converse is true if (A_i, B_i, C_i) is a minimal realization of $H_i(s)$ (Corollary 6.10).

The state of the closed-loop system is $x := [x_1^T \quad x_2^T]^T \in \mathbb{R}^{n_1+n_2}$, input is $u := [u_1^T \quad u_2^T]^T \in \mathbb{R}^{m_1+m_2}$, and output is $y := [y_1^T \quad y_2^T]^T \in \mathbb{R}^{m_1+m_2}$. It can be shown that the closed-loop system is also LTI, described by

$$\dot{x} = Ax(t) + Bu(t), \quad y = Cx(t) \tag{6.12a}$$

where A, B, C can be expressed in terms of A_i, B_i, C_i as (Exercise 6.7):

$$A := \begin{bmatrix} A_1 & -B_1C_2 \\ B_2C_1 & A_2 \end{bmatrix} \qquad B := \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} \qquad C := \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix} \tag{6.12b}$$

We are interested in sufficient conditions under which the closed-loop system (6.12) is asymptotically stable when the input is zero $u \equiv 0$.

By Theorem 6.5 the origin of the closed-loop system is asymptotically stable (with $u \equiv 0$) if and only if all eigvenvalues of A are in the open LHP, or equivalently all roots of the characteristic polynomial $\det(sI - A) = 0$ are in the open LHP. As we see from

(6.12b), directly determining the eigenvalues of the closed-loop system matrix A can be complicated. The following fundamental result allows us to relate the eigenvalues of the closed-loop system matrix A to the eigvenvalues of the system matrices A_i of the open-loop (component) systems and the loop function $H_1 H_2$. It is proved in Appendix 6.5.

Theorem 6.14 *The characteristic polynomial* $\det(sI - A)$ *of the closed-loop system* (6.12) *is given by*

$$\det(sI - A) \quad = \quad \det(sI - A_1) \cdot \det(sI - A_2) \cdot \det(I + H_1(s)H_2(s)) \tag{6.13}$$

where $H_i(s) = C_i(sI - A_i)^{-1}B_i$.

This theorem has several important implications. First the left-hand side of (6.13) is a polynomial in s of degree $n_1 + n_2$. On the right-hand side, $\det(sI - A_i)$ are polynomials but $\det(I + H_1(s)H_2(s))$ are generally rational functions. Therefore any poles of $\det(I + H_1(s)H_2(s))$ must be canceled by zeros of $\det(sI - A_i)$.

Second the poles of the transfer function $H_i(s) = C_i(sI - A_i)^{-1}B_i$ are in general a subset of the eigenvalues of A_i (zeros of $\det(sI - A_i)$); see Chapter 6.2.2. If there is an eigenvalue λ of A_i that is not a pole of $H_i(s)$, the associated eigenvector is then associated with an uncontrollable and/or unobservable mode of the (open-loop) linear system (A_i, B_i, C_i); in this case, we say that the system (A_i, B_i, C_i) has a hidden mode at λ. The system (A_i, B_i, C_i) can be excited along the direction of the eigenvector associated with the hidden mode λ but the effect on the system state $x_i(t)$ will not manifest itself at the output $y_i(t)$. If the realization (A_i, B_i, C_i) is minimal (controllable and observable) then s is a pole of $H_i(s)$ if and only if it is a eigenvalue of A_i; see Corollary 6.10.

Third the closed-loop system (A, B, C) in (6.12) is asymptotically stable if and only if the zeros of $\det(sI - A)$, or the eigenvalues of the closed-loop system matrix A, are all in the open left-half plane (LHP). Let λ_i be the eigenvalues of the closed-loop system matrix A and $\lambda_{1j}, \lambda_{2j}$ be the eigenvalues of the open-loop system matrices A_1, A_2 respectively. Then (6.13) can be rewritten as

$$\prod_i (s - \lambda_i) \quad = \quad \prod_j (s - \lambda_{1j}) \cdot \prod_j (s - \lambda_{2j}) \cdot \det(I + H_1(s)H_2(s))$$

Hence the closed-loop system (A, B, C) will be asymptotically stable if and only if the rational function $\det(I + H_1(s)H_2(s))$ has no zeros (i.e. roots of its numerator polynomial) in the closed right-half plane (RHP) and any zeros of the polynomials $\det(sI - A_1) \cdot \det(sI - A_2)$ in the closed RHP are exactly canceled by the poles of $\det(I + H_1(s)H_2(s))$ (i.e. roots of its denominator polynomial). In particular if the open-loop systems are asymptotically stable, i.e., all $\lambda_{1j}, \lambda_{2j}$ of A_1, A_2 are in the open LHP for all j, then the closed-loop system is asymptotically stable if and only if $\det(I + H_1(s)H_2(s))$ has no zero (nor pole) in the closed RHP.

The Nyquist stability theorem provides a graphical test for this condition. It characterizes the asymptotic stability of the *closed-loop* system in terms of the eigenvalues of the open-loop system matrices A_1, A_2 and properties of the loop function $H_1(s)H_2(s)$. Let

$$p_{0+} \quad := \quad \text{number of closed RHP zeros of the polynomial} \\ \det(sI - A_1)\det(sI - A_2), \text{ counting multiplicity}$$

If A_1, A_2 have an eigenvalue at λ_1 in the closed RHP with combined (algebraic) multiplicity of 3 then λ_1 contributes 3 to p_{0+}. If both A_1 and A_2 are open-loop asymptotically stable then $p_{0+} = 0$. From the discussion above, the closed-loop system is asymptotically stable if and only if $\det(I + H_1(s)H_2(s))$ has no zero in the closed RHP and exactly p_{0+} number of poles in the closed RHP. Define the rational function:

$$f(s) \quad := \quad \det(I + H_1(s)H_2(s))$$

In the complex s-plane ($s = \sigma + j\omega$) consider the closed curve D that consists of the imaginary $j\omega$-axis and of the right half-circle centered at the origin with arbitrarily large radius. If the rational function $f(s)$ has a pole on the $j\omega$-axis the curve D is indented to the *left* by an arbitrarily small half-circle centered at the pole.[7] By convention D is a simple (not self-intersecting) closed curve oriented *clockwise*; see Figure 6.6(a). Let $f(D)$ denote the closed oriented curve $\{f(s), s \in D\}$ traced by $f(s)$ as s traverses the curve D clockwise. Even though D is a simple curve $f(D)$ is not necessarily a simple curve.

We call D a *Nyquist path* and $f(D)$ the corresponding *Nyquist plot*.

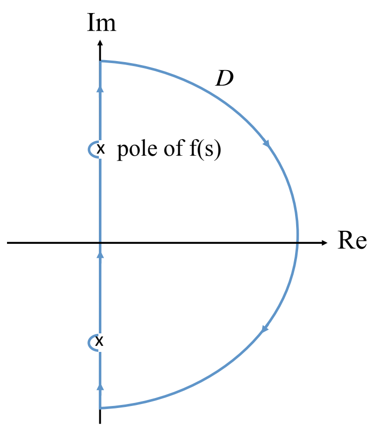
Theorem 6.15 **Nyquist stability criterion.** *Consider the closed-loop system* (A, B, C) *in* (6.12).

1. *The origin is asymptotically stable if and only if the Nyquist plot $f(D)$ encircles the origin p_{0+} times* counterclockwise *and does not go through the origin.*

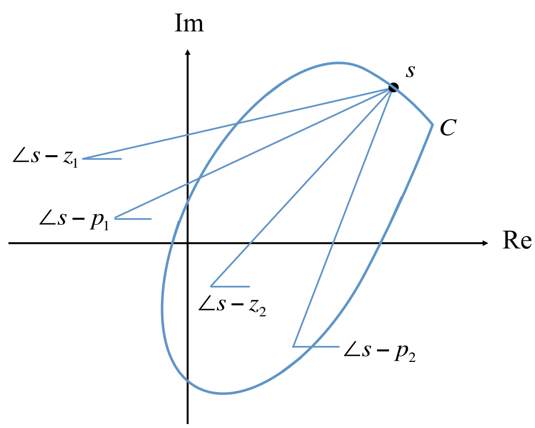
2. *Suppose the component systems are open-loop asymptotically stable (i.e., all eigenvalues of A_i, $i = 1, 2$, are in the open LHP). Then the origin of the closed-loop system is asymptotically stable if and only if the Nyquist plot $f(D)$ does not encircle the origin* counterclockwise *nor does it go through the origin.*

We now sketch a proof of Theorem 6.15. The proof relies on the argument principle, as follows. Consider the complex plane in Figure 6.6.

[7] By indenting to the left, the closed curve D includes purely imaginary poles of $f(s)$ and therefore the stability criterion in Theorem 6.15 disallows eigenvalues of the *closed-loop* system on the imaginary axis. This guarantees asymptotic stability. If we had indented to the right instead then the stability criterion allows eigenvalues of the *closed-loop* system on the imaginary axis and hence guarantees only stability, not asymptotic stability.



(a) Nyquist path D



(b) The argument principle

Figure 6.6: Nyquist path D and the argument principle.

The argument principle. We are given a *closed, oriented, and simple* (not self-intersecting) curve C and a complex-valued function

$$g(s) \quad := \quad \frac{\prod_i (s - z_i)}{\prod_k (s - p_k)}$$

where z_i and p_k are zeros and poles respectively of g with $z_i \neq p_k$ for all i, k. Let $g(s)$ be a proper rational function with $g(s)$ finite and nonzero for all $s \in C$ (i.e., $g(s)$ has neither poles nor zeros on C). As s traverses C clockwise, $g(s)$ traverses $g(C)$ and encircles the origin $p - n$ times *counterclockwise*, where

$$\begin{aligned} p &:= \text{\# of poles of } g \text{ inside } C \text{ counting multiplicity} \\ n &:= \text{\# of zeros of } g \text{ inside } C \text{ counting multiplicity} \end{aligned}$$

Recall that a proper rational function is one where the degree of the numerator polynomial does not exceed the degree of the denominator polynomial, and $g(C) := \{g(s),\ s \in C\}$. To see why the argument principle holds, consider the argument of g:

$$\angle g(s) \quad = \quad \sum_i \angle(s - z_i) \ - \ \sum_k \angle(s - p_k)$$

For any zero z_i or pole p_i outside of the contour C in Figure 6.6, as s completes a cycle on C, the net change of $\angle(s - z_i)$ or of $\angle(s - p_k)$ is zero. On the other hand, for a zero z_i or pole z_k inside C, as s completes a cycle on C clockwise, $\angle(s - z_i)$ or $\angle(s - p_k)$ makes exactly one cycle *clockwise*, i.e., z_i introduces -2π and p_k introduces 2π to the curve $g(s)$ as it traverses $g(C)$. Hence

$$\text{\# counterclockwise encirclements of the origin by } g(C) \quad = \quad p \ - \ n$$

since each 2π change in angle corresonds to a counterclockwise encirclement of the origin.

Proof of Theorem 6.15. Theorem 6.5 implies that the closed-loop system matrix A is asymptotically stable if and only if all its eigenvalues, i.e., all zeros of $\det(sI - A)$, are in the open LHP. From Theorem 6.14 and the discussion immediately thereafter, the zeros of $\det(sI - A)$ must come from the zeros of $f(s) := \det(I + H_1(s)H_2(s))$ and those zeros of the polynomial $\det(sI - A_1)\det(sI - A_2)$ that are *not* canceled by the poles of $\det(I + H_1(s)H_2(s))$. In other words A will be asymptotically stable if and only if

- every closed RHP zero of $\det(sI - A_1)\det(sI - A_2)$ is canceled by a closed RHP pole of $f(s)$.
- $f(s)$ has no closed RHP zero.

The rational function $f(s)$ is proper. Since the curve D is indented to the *left* by an arbitrarily small half-circle centered at any pole of $f(s)$ on the $j\omega$-axis, D contains all closed RHP poles of $f(s)$. Moreover $f(s)$ remains nonzero (and finite since f is proper) on D. Hence we can apply the argument principle to conclude that these two conditions are equivalent to that the number of counterclockwise encirclements of the origin by $f(D)$ is exactly $p = p_{0+}$ and $n = 0$. □

Before extending the Nyquist stability criterion to the time-delayed system we state the following simple result which is used in the stability analysis of congestion control algorithms below.

Lemma 6.16 *Let $\sigma(AB)$ denote the set of* nonzero *eigenvalues of the product matrix AB. Then $\sigma(AB) = \sigma(BA)$ and hence $det(I + AB) = det(I + BA)$.*

Proof. If (λ, v) are eigenvalue and eigenvector of AB with $\lambda \neq 0$, then $ABv = \lambda v \Rightarrow BA(Bv) = \lambda Bv$, and hence (λ, Bv) are eigenvalue and eigenvector of BA. Hence AB and BA have the same nonzero eigenvalues. The second assertion follows since both determinants equal, $\prod_i (1 + \lambda_i(AB)) = \prod_i (1 + \lambda_i(BA))$ where $\lambda_i(M)$ denotes the eigenvalues of matrix M. □

The proof of Lemma 6.16 also shows that if (λ, v) are eigenvalue and eigenvector of AB with $\lambda \neq 0$, then $v \neq 0$ by the definition of eigenvector. Therefore $ABv = \lambda v$ implies $Bv \neq 0$. Hence AB and BA no only share nonzero eigenvalues, their nonzero eigenvalues also have the same geometric multiplicities.[8] Lemma 6.16 implies that

$$f(s) \quad := \quad \det(I + H_1(s)H_2(s)) \quad = \quad \det(I + H_2(s)H_1(s))$$

in Theorem 6.15, i.e., it does not matter where we "break the loop."

6.2.5 GENERALIZED NYQUIST STABILITY CRITERION

The application of Theorem 6.15 requires the computation of the Nyquist plot $f(j\omega) := \det(I + H_1(j\omega)H_2(j\omega))$ as ω goes from $-\infty$ to $+\infty$ along the imaginary axis with possible left indentations. In the absence of delay the transfer functions $H_i(s)$ are rational functions. With time delay they involve $e^{-\tau s}$ and are transcendental functions. Instead of checking the Nyquist plot $f(j\omega)$, the stability analysis of the closed-loop system in the presence of feedback delay checks the properties of a certain family of circuits formed from the eigenloci of the loop function $L(j\omega) = H_1(j\omega)H_2(j\omega)$. A rigorous derivation of the generalized

[8]The algebraic multiplicity of any eigenvalue λ is the power of the factor $(s - \lambda)$ in the characteristic polynomial. The geometric multiplicity of λ is the dimension of the span of eigenvectors associated with λ. The geometric multiplicity of any eigenvalue is upper bounded by its algebraic multiplicity. The sum of algebraic multiplicities of all eigenvalues of an $n \times n$ (complex) matrix is n.

Nyquist stability criterion for the time-delayed case is beyond the scope of this book, but the basic idea is as follows.

Let $L(s)$ denote a loop function. Theorem 6.15 then checks the argument of the Nyquist plot $f(s)$ as s traverses the Nyquist path D. Since

$$\det(I + L(s)) \quad = \quad \prod_{i=1}^{m}(1 + \lambda_i(s))$$

where λ_i, $i = 1, \ldots, m$, are the eigenvalues of the loop function $L(s)$ evaluated at s, we have

$$\angle\det(I + L(s)) \quad = \quad \angle\prod_{i=1}^{m}(1 + \lambda_i(s)) \quad = \quad \sum_{i=1}^{m}\angle(1 + \lambda_i(s))$$

As s travels along the Nyquist path D, we can label the eigenvalues of $L(s)$ such that, for each i, $s \to \lambda_i(s)$ is a continuous function (this is because eigenvalues of a matrix are continuous functions of the elements of the matrix). This gives m continuous eigenloci $\lambda_i(D)$.

Thus it seems that, instead of counting the number of encirclements of the origin by $\det(I + L(s))$, we can equivalently sum the number of encirclements of the point $(-1, 0)$ by $\lambda_i(D)$, $i = 1, \ldots, m$. Unfortunately this does not work because some eigenloci $\lambda_i(D)$ may not form *closed* paths, as the following example shows.

Example 6.17 Consider the loop function

$$L(s) \quad := \quad \begin{bmatrix} 0 & 1 \\ \frac{s-1}{s+1} & 0 \end{bmatrix}$$

with two eigenloci

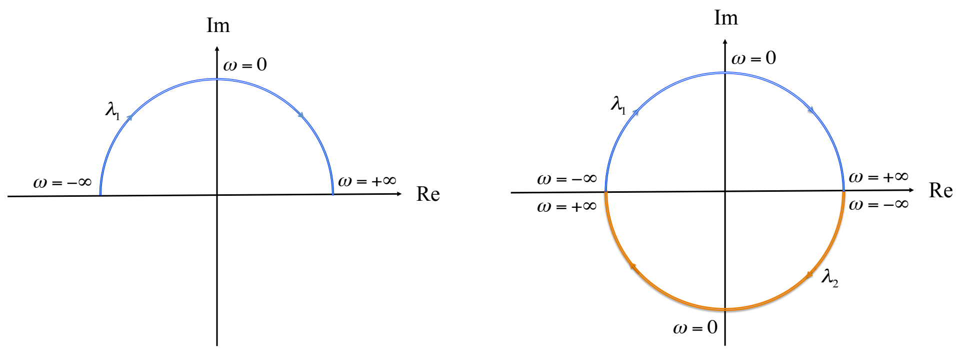
$$\lambda_1(j\omega) \quad = \quad \sqrt{\frac{j\omega - 1}{j\omega + 1}} \quad = \quad \frac{j(1 - j\omega)}{\sqrt{1 + \omega^2}} \quad = \quad e^{j\left(\frac{\pi}{2} - \tan^{-1}\omega\right)}$$

and

$$\lambda_2(j\omega) \quad = \quad -\lambda_1(j\omega) \quad = \quad -e^{j\left(\frac{\pi}{2} - \tan^{-1}\omega\right)}$$

These eigenloci $\lambda_1(D)$ and $\lambda_2(D)$ are shown in Figure 6.7 where the Nyquist path D is the $j\omega$-axis for $\omega \in [-\infty, \infty]$. Consider $\lambda_1(j\omega)$ and Figure 6.7(a). As ω goes from $-\infty$ to ∞, $\lambda_1(j\omega)$ traces out a semicircular arc in the upper half plane from $e^{j\pi}$ to e^{j0} in the clockwise direction. Similarly the eigenlocus of $\lambda_2(D)$ is the semicircle in the lower half plane as shown in Figure 6.7(b). □

Even though some eigenloci $\lambda_i(D)$ may not form closed paths, it can be shown however that we can always form an indexed family of *closed* paths $(\gamma_k(D), k = 1, \ldots, p)$ from



(a) $\lambda_1(j\omega)$ (b) $\lambda_i(j\omega)$, $i = 1, 2$, and γ_1

Figure 6.7: (a) Eigenloci $\lambda_i(D)$ may not be a closed curve. (b) We can always form an indexed family of *closed* curves ($\gamma_k, k = 1, \ldots, p$) from the set of eigenloci ($\lambda_i, i = 1, \ldots, m$).

the set of eigenloci ($\lambda_i, i = 1, \ldots, m$) that encircle the same points in the complex plane. For Example 6.17, from Figure 6.7(b), as ω goes from $-\infty$ to $+\infty$, $\lambda_1(j\omega)$ traces out the upper semicircle from $(-1, 0)$ to $(1, 0)$ in the clockwise direction while *simultaneously* $\lambda_2(j\omega)$ traces out the lower semicircle in the same direction. Moreover for *any* choice of $(\gamma_k(s), k = 1, \ldots, p)$, the sum of the number of encirclements of -1 is equal to the number of encirclements of the origin by $f(D) = \det(I + L(D))$. We summarize this discussion as a theorem (c.f. Theorem 6.15).

Theorem 6.18 *Suppose the loop function L is proper. Construct an indexed family of* closed *paths $(\gamma_k(s), k = 1, \ldots, p)$ from the eigenloci $(\lambda_i(s), i = 1, \ldots, m)$ of the loop function L as described above. The closed-loop system (A, B, C) in* (6.12) *is asymptotically stable if and only if $(\gamma_k(s), k = 1, \ldots, p)$ does not go through -1 on the real line and encircles -1 on the real line p_{0+} times counterclockwise, as s traverses the Nyquist path D.*

Hence we can test closed-loop stability by checking the number of encirclements of -1 by any such family of closed curves $(\gamma_k(s), k = 1, \ldots, p)$. Even though our discussion assumes a proper rational loop function L, Theorem 6.18 holds for more general proper functions, e.g., for systems involving delay or trigonometric functions.

For our purposes we only need a corollary of Theorem 6.18.

Corollary 6.19 *Suppose*

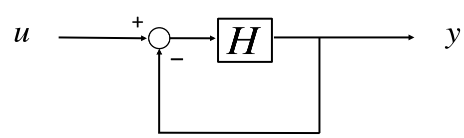
- *The loop function L is proper.*
- *The open-loop systems are all asymptotically stable, i.e. the eigenvalues of A_i are all in the open LHP.*

If all eigenloci of the loop function $L(s)$ stay entirely to the right of (and do not pass through) -1 on the real line, as s traverses the Nyquist path D, then the origin of the closed-loop system is asymptotically stable.

In applying the corollary it is sufficient to check the Nyquist plot of eigenloci of $L(s)$ as s traverses $[0, \infty)$ or $(-\infty, 0]$ instead of the complete Nyquist path D.

6.2.6 UNITY FEEDBACK SYSTEMS

When we study the primal or dual algorithm, there is dynamics in either sources or links but not both. As we will see in Sections 6.3 and 6.4 these algorithms can be modeled by the unity feedback system in Figure 6.8. The Nyquist stability criterion simplifies in this case because the loop function $L(s) = H_1(s)H_2(s)$ becomes $L(s) = H(s)$ for a unity feedback system, as we now explain.



$$y(s) = -H(s)y(s) + H(s)u(s)$$

$$\Longrightarrow \quad y(s) = \left(I + H(s)\right)^{-1} H(s)\, u(s)$$

Figure 6.8: Unity feedback system.

Let A^{cl} denote the system matrix of the closed-loop system in Figure 6.8. Let $(A^{\mathrm{ol}}, B^{\mathrm{ol}}, C^{\mathrm{ol}})$ be a realization of the open-loop transfer function $H(s)$. Theorem 6.14 reduces to: the characteristic polynomial $\det(sI - A^{\mathrm{cl}})$ of the closed-loop system is given by (see Exercise 6.6):

$$\det(sI - A^{\mathrm{cl}}) \quad = \quad \det(sI - A^{\mathrm{ol}}) \cdot \det(I + H(s)) \tag{6.14}$$

Hence, for unity feedback system, Theorems 6.14 and 6.15 on Nyquist stability for systems without feedback delay, as well as Theorem 6.18 and Corollary 6.19 for delayed systems hold with the loop function $L(s) = H(s)$.

Moreover the discussion preceding Corollary 6.10 establishes that: if $(A^{\mathrm{ol}}, B^{\mathrm{ol}}, C^{\mathrm{ol}})$ is minimal then

$$\{\text{eigenvalues of } A^{\mathrm{ol}}\} = \{\text{poles of } H(s)\}$$

This implies that, for a minimal realization, every eigenvalue of A^{ol} on the right-hand side of (6.14) is canceled by exactly one pole of $H(s)$. In particular the closed paths $\gamma_k(s)$ from the eigenloci of the loop function $L(s)$ will always encircle -1 on the real axis p_{0+} times counterclockwise as s traverses the Nyquist path D.

Example 6.20 **Unity feedback system.** Consider the delayed system

$$\dot{x} \;=\; u(t-\tau), \qquad y(t) \;=\; x(t)$$

with the transfer function $H(s) = e^{-\tau s}/s$. The unity feedback system in Figure 6.8 is the following

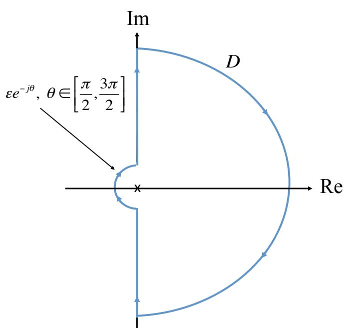
$$\dot{x} \quad = \quad -x(t-\tau) \tag{6.15}$$

We now use Theorem 6.18 to show the stability of the closed-loop system (6.15) provided the delay $\tau < \pi/2$.

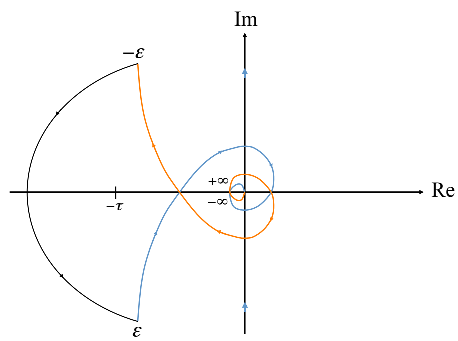
Since the loop function

$$L(s) \quad = \quad H(s) \quad = \quad \frac{e^{-\tau s}}{s}$$

has a pole at the origin, the number p_{0+} of poles in the closed RHP is 1. Consider the Nyquist path D shown in Figure 6.9(a) that has a left indentation $s = \left(\epsilon\, e^{-j\theta}, \theta \in [\pi/2, 3\pi/2]\right)$ around the origin for an arbitrarily small $\epsilon > 0$. By Theorem 6.18 the closed-loop system (6.15) is asymptotically stable if $L(s)$ does not go through -1 on the real line and encircles -1 on the real line once in the counterclockwise direction as s traverses the Nyquist path D. We now show that the Nyquist plot $(L(s), s \in D)$ is as shown in Figure 6.9(b). For



(a) Nyquist path D



(b) Nyquist plot $L(s), s \in D$

Figure 6.9: (a) Nyquist path D that indents left at the origin. The indentation is the path $(\epsilon e^{-j\theta}, \theta \in [\pi/2, 3\pi/2])$ for an arbitrarily small $\epsilon > 0$. (b) The corresponding Nyquist plot $(L(s) = \frac{1}{s}e^{-\tau s}, s \in D)$. The plot $L(s)$ on the indentation $s = (\epsilon e^{-j\theta}, \theta \in [\pi/2, 3\pi/2])$ is the black curve with a counterclockwise orientation.

$\omega \in [\epsilon, \infty]$ we have

$$L(j\omega) \;=\; \tau\,\frac{e^{-j\tau\omega}}{j\tau\omega} \;=\; \tau\,\frac{e^{-j\omega'}}{j\omega'} \;=\; -\tau\left(\frac{\sin\omega'}{\omega'} + j\,\frac{\cos\omega'}{\omega'}\right), \quad \omega' \in [\epsilon, \infty] \qquad (6.16)$$

where $\omega' := \tau\omega$. Hence the Nyquist plot $L(j\omega)$ for $\omega \in [\epsilon, \infty]$ is the blue curve in Figure 6.9(b). Similarly $L(j\omega)$ for $\omega \in [-\infty, -\epsilon]$ is the orange curve in Figure 6.9(b). For the left indentation $s = \left(\epsilon\, e^{-j\theta}, \theta \in [\pi/2, 3\pi/2]\right)$ around the origin, the Nyquist plot is

$$L(s) \;=\; \frac{e^{-\tau\epsilon e^{-j\theta}}}{\epsilon e^{-j\theta}} \;\approx\; \frac{e^{j\theta}}{\epsilon} \qquad \text{for small } \epsilon > 0$$

as θ goes from $\pi/2$ to $3\pi/2$. Hence $L(s)$ is the black curve in Figure 6.9(b) oriented counterclockwise. Therefore the stability condition in Theorem 6.18 is satisfied if and only if the blue segment of $L(j\omega)$ in (6.16) crosses the real axis (strictly) to the left of -1. At this crossing, $\cos(\tau\omega) = 0$, giving $\tau\omega = \pi/2$. Hence the blue curve crosses the real axis at

$$-\frac{\sin\tau\omega}{\omega} \;=\; -\frac{2\tau}{\pi}$$

Therefore the closed-loop system is asymptotically stable if and only if $\tau < \pi/2$. □

We now use Theorem 6.18 and Corollary 6.19 to study the linear stability of congestion control algorithms in the presence of feedback delay.

6.3 STABILITY OF PRIMAL ALGORITHMS

Consider the following primal algorithm:

$$\begin{aligned} \dot{x}_i(t) &= \kappa_i(U_i'(x_i(t)) - q_i(t))^+_{x_i(t)}, & p_l(t) &= g_l(y_l(t)) \\ q_i(t) &= \sum_l R_{li}\, p_l(t - \tau^b_{li}), & y_l(t) &= \sum_i R_{li}\, x_i(t - \tau^f_{li}) \end{aligned}$$

where $\kappa_i > 0$, the utility functions U_i are strictly concave with $U_i''(x_i) < 0$ for all $x_i \geq 0$, U_i' are the first derivatives of U_i, and g_l are nonnegative and strictly decreasing. Here τ^b_{li} represents the backward delay from link l to source i while τ^f_{li} represents the forward delay from source i to link l; both can depend on the pair (l, i). Let (x^*, p^*) be an equilibrium. Without loss of generality we assume $x_i^* > 0, p_l^* > 0$ for all i, l; otherwise we remove the zero entries from consideration in the linearized model below. Linearizing around (x^*, p^*) and using $\tilde{x} := x - x^*, \tilde{p} := p - p^*$, etc, we obtain the following linear delayed model in the

Laplace domain:[9]

$$\begin{aligned} s\tilde{x}_i &= \kappa_i U_i''(x_i^*)\tilde{x}_i - \kappa_i \tilde{q}_i, & \tilde{p}_l &= g_l'(y_l^*)\tilde{y}_l \\ \tilde{q}_i &= \sum_l R_{li} e^{-\tau_{li}^b s}\tilde{p}_l, & \tilde{y}_l &= \sum_i R_{li} e^{-\tau_{li}^f s}\tilde{x}_i \end{aligned}$$

where U_i'' are the second derivatives of U_i. This can be written compactly in matrix form by defining the forward (routing) matrix $R^f(s)$ and backward matrix $R^b(s)$ in the Laplace domain where

$$\left[R^f(s)\right]_{li} := R_{li} e^{-s\tau_{li}^f}, \qquad \left[R^b(s)\right]_{li} := R_{li} e^{-s\tau_{li}^b} \tag{6.17}$$

Then the linearized system is, in Laplace domain:

$$s\tilde{x} = KU''\tilde{x} - K\tilde{q}, \qquad \tilde{p} = G\tilde{y} \tag{6.18a}$$

$$\tilde{q} = \left(R^b(s)\right)^T \tilde{p}, \qquad \tilde{y} = R^f(s)\tilde{x} \tag{6.18b}$$

where $K :=\text{diag}(\kappa_i, i \in N) \succ 0$, $U'' := \text{diag}(U_i''(x_i^*), i \in N) \prec 0$ and $G := \text{diag}(g_l'(y_l^*), l \in L) \succ 0$. This is represented as a feedback system in Figure 6.10(a). Note that $H(s) \neq -(sI - KU'')^{-1}K$ because the input to H is $-\tilde{q}$, not $\tilde{q}$.

Assume:

C6.1: All utility functions U_i are strictly concave with $U_i''(x_i) < 0$ for all $x_i \geq 0$ and the routing matrix R has full row rank.

C6.2: For all sources i, $\tau_{li}^f + \tau_{li}^b = \tau_i$ for all links l in i's path.

Condition C6.1 implies that the equilibrium (x^*, p^*) exists and is unique. Condition C6.2 says that the forward delay from a source i to *any* links l in its path plus the backward delay from that link back to the source is equal to the round-trip delay τ_i of source i. It is a reasonable assumption if all packets of each source i follow the same round-trip path. We further assume τ_i are constants. The main implication of Condition C6.2 is the following relationship between the forward and backward matrix:

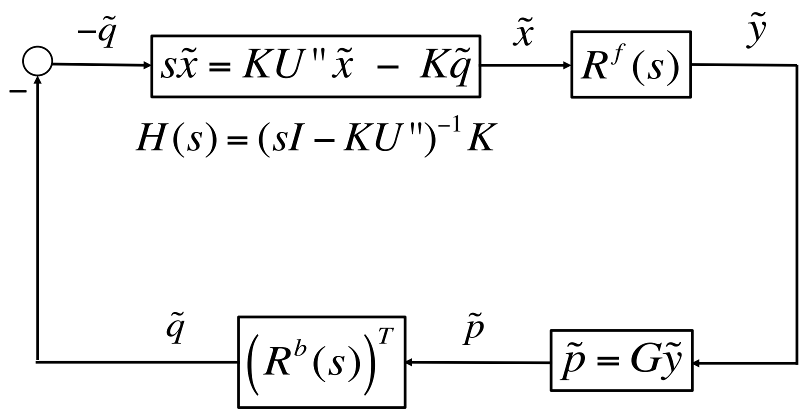
$$R^b(s) = R^f(-s)\text{diag}\left(e^{-\tau_i s}\right) \tag{6.19}$$

Let $\overline{N} := \max_l \sum_i R_{li}$ be an upper bound on the number of sources through any link and $\overline{L} := \max_i \sum_l R_{li}$ be an upper bound on the number of links used by any source.

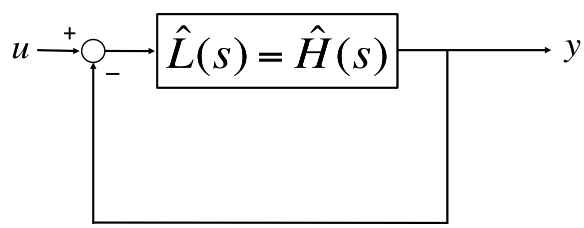
Theorem 6.21 *Assume conditions C6.1 and C6.2 hold. The origin of the linearized primal algorithm* (6.18) *is asymptotically stable if*

$$\max_i \tau_i \kappa_i \cdot \max_l g_l'(y_l^*) \leq \frac{\pi}{2\,\overline{N}\,\overline{L}}$$

[9]We abuse notation to use $\tilde{x}$ to denote both the time function $\tilde{x}(t)$ and its Laplace transform $\tilde{x}(s)$. The meaning should be clear from the context.



(a) Primal algorithm



(b) Equivalent system

Figure 6.10: (a) Delayed linear feedback system representing the primal algorithm (6.18). (b) Equivalent unity feedback system with loop function $\hat{L}(s) = \hat{H}(s) := \hat{R}(j\omega)\ \mathrm{diag}\left(\frac{\pi}{2}\frac{e^{-j\tau_i\omega}}{j\tau_i\omega+\alpha_i}\right)\hat{R}^H(j\omega)$.

An upper bound on $\tau_i \kappa_i$ means that to maintain stability the source control gain κ_i should be small if the round-trip time τ_i is large. An upper bound on $g_l'(y_l^*)$ means that the congestion price $p_l^* = g_l(y_l^*)$ at a link should not be too sensitive to its input rate y_l^* in equilibrium. The bounds $\overline{N}, \overline{L}$ in the stability condition are conservative. The proof of the theorem suggests how to modify the algorithm in (6.18) to remove these bounds (see the dual algorithm in the next subsection).

We will apply Corollary 6.19 to prove Theorem 6.21. The loop transfer function of (6.18) from $\tilde{p}$ to $\tilde{p}$ is (see Figure 6.10):

$$L(s) \quad = \quad GR^f(s)\,(sI - KU'')^{-1}\,K\,\left(R^b(s)\right)^T \tag{6.20}$$

Substituting (6.19) into (6.20), the loop transfer function becomes

$$\begin{aligned} L(s) \quad &= \quad GR^f(s)\,(sI - KU'')^{-1}\,K\ \mathrm{diag}\left(e^{-\tau_i s}\right)\left(R^f(-s)\right)^T \\ &= \quad \mathrm{diag}\left(g_l'\right) R^f(s)\ \mathrm{diag}\left(\frac{\kappa_i\, e^{-\tau_i s}}{s - \kappa_i\, U_i''}\right)\left(R^f(-s)\right)^T \end{aligned}$$

Since all matrices in $L(s)$ except $R^f(s)$ are diagonal we have

$$L(s) \quad = \quad \mathrm{diag}\left(g_l'\right) R^f(s)\ \mathrm{diag}\left(\sqrt{\frac{2}{\pi}\tau_i\kappa_i}\right)\mathrm{diag}\left(\frac{\pi}{2}\frac{e^{-\tau_i s}}{\tau_i s + \alpha_i}\right)\mathrm{diag}\left(\sqrt{\frac{2}{\pi}\tau_i\kappa_i}\right)\left(R^f(-s)\right)^T \tag{6.21}$$

where $\alpha_i := -\tau_i\kappa_i\, U_i''(x_i^*) \; > \; 0$. For any matrix M let $(\sigma(M(j\omega)), j\omega \in D)$ denote the eigenloci of M that are not identically zero as s traverses the Nyquist path D. Lemma 6.16 implies that we can change the order of some component matrices in L in (6.21) as far as nonzero eigenloci are concerned. Hence

$$\sigma(L(j\omega)) \quad = \quad \sigma\left(\hat{L}(j\omega)\right) \tag{6.22a}$$

where

$$\hat{L}(j\omega) \quad = \quad \hat{R}(j\omega)\ \mathrm{diag}\left(\frac{\pi}{2}\frac{e^{-j\tau_i\omega}}{j\tau_i\omega + \alpha_i}\right)\hat{R}^H(j\omega) \tag{6.22b}$$

$$\hat{R}(j\omega) \quad := \quad \mathrm{diag}\left(\sqrt{g_l'}\right) R^f(j\omega)\ \mathrm{diag}\left(\sqrt{\frac{2}{\pi}\tau_i\kappa_i}\right) \tag{6.22c}$$

and $\hat{R}^H(j\omega) = \hat{R}^T(-j\omega)$ is the Hermitian transpose. This means that the (nonzero) eigenloci of the loop function $L(j\omega)$ coincide with those of $\hat{L}(j\omega)$. Moreover we can treat $\hat{L}(j\omega)$ as the loop function of a unity feedback system (see Chapter 6.2.6). This is illustrated in Figure 6.10(b). For this equivalent unity feedback system, the corresponding open-loop transfer function is $\hat{H}(s) = \hat{L}(s)$. From (6.22b)(6.22c) the poles of $\hat{H}(s)$ are $-\alpha_i/\tau_i < 0$, i.e. all of them are in the open LHP and $p_{0+} = 0$. By Corollary 6.19 the closed-loop system is asymptotically stable if and only if the eigenloci of $L(j\omega)$, or equivalently $\hat{L}(j\omega)$, neither pass through nor encircle -1 in the complex plane as ω traverses the Nyquist path from 0 to $+\infty$.

To bound the eigenloci in (6.22) as ω goes from 0 to $+\infty$, we will use the following property.

Lemma 6.22 *Let $D := \text{diag}(d_i, i = 1, \ldots, n) \in \mathbb{C}^{n\times n}$ be a complex diagonal matrix and $A \in \mathbb{C}^{m\times n}$ be any nonzero complex matrix. Then*

$$\sigma\left(ADA^H\right) \quad\subseteq\quad \rho\left(AA^H\right) \cdot co\left(0, d_i,\, i = 1, \ldots, n\right)$$

where $\rho\left(AA^H\right)$ is the spectral radius of AA^H and $co(a_k)$ is the convex hull of points a_k.

Proof. Let (λ, v) be an eigenvalue and eigenvector pair of ADA^H such that $\|v\|_2 = 1$. Then

$$ADA^H v \;=\; \lambda\, v \quad\Rightarrow\quad \lambda \;=\; v^H ADA^H v \;=\; \sum_i d_i\, |\hat{v}_i|^2 \tag{6.23}$$

where $\hat{v} := A^H v$. If $v \neq 0$ is in the null space of A^H then $\hat{v}$ is the zero vector and $\lambda = 0$. In any case

$$\|\hat{v}\|_2^2 \;=\; v^H AA^H v \;\leq\; \rho\left(AA^H\right) \cdot \|v\|_2^2 \;=\; \rho\left(AA^H\right) \tag{6.24}$$

since AA^H is a Hermitian positive semidefinite matrix. Hence from (6.23) we have

$$\lambda \;=\; \rho\left(AA^H\right) \sum_i \frac{|\hat{v}_i|^2}{\rho\left(AA^H\right)}\, d_i \;=\; \rho\left(AA^H\right) \left(\sum_i \frac{|\hat{v}_i|^2}{\rho\left(AA^H\right)}\, d_i \;+\; \left(1 - \frac{\|\hat{v}\|_2^2}{\rho\left(AA^H\right)}\right) \cdot 0 \right)$$

The lemma then follows from (6.24). $\square$

Proof of Theorem 6.21. The closed loop system is asymptotically stable if and only if the eigenloci in (6.22) neither pass through nor encircle the critical point -1 in the complex plane. We now use Lemma 6.22 to bound the eigenloci (strictly) to the right of -1.

From Lemma 6.22

$$\sigma\left(\hat{R}(j\omega)\ \text{diag}\left(\frac{\pi}{2}\frac{e^{-\tau_i s}}{\tau_i s + \alpha_i}\right) \hat{R}^H(j\omega)\right) \quad\subseteq\quad \rho\left(\hat{R}(j\omega)\hat{R}^H(j\omega)\right) \cdot \text{co}\left(0,\ \frac{\pi}{2}\frac{e^{-\tau_i s}}{\tau_i s + \alpha_i}\right)$$

We examine each term on the right-hand side in turn.

Observe the sets of eigenloci in the last term satisfy:

$$\left\{\frac{\pi}{2}\frac{e^{-j\tau_i\omega}}{j\tau_i\omega + \alpha_i},\ \omega \in [0, +\infty]\right\} \quad=\quad \left\{\frac{\pi}{2}\frac{e^{-j\omega}}{j\omega + \alpha_i},\ \omega \in [0, +\infty]\right\}$$

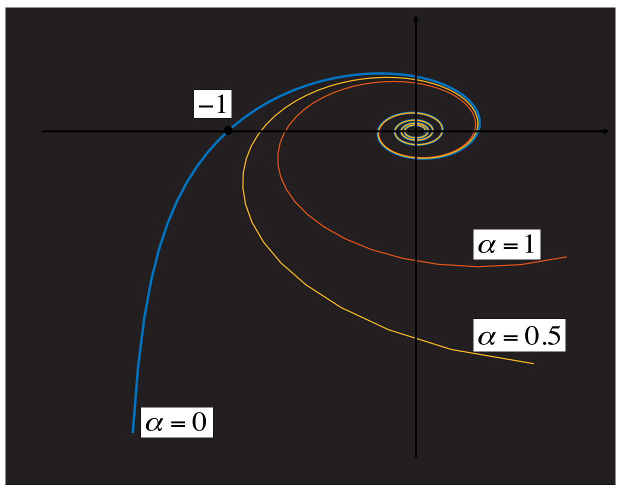


Figure 6.11: The Nyquist plots $\left\{\frac{\pi}{2}\frac{e^{-j\omega}}{j\omega+\alpha_i},\ \omega \in [0,+\infty]\right\}$ for $\alpha_i = 0, 0.5, 1$ in the complex plane. They cross the real axis to the right of -1 as long as $\alpha_i > 0$. When $\alpha = 0$, see Figure 6.9 in Example 6.20 for the complete Nyquist plot.

The set on the right-hand side is shown in Figure 6.11. These sets cross the real axis (strictly) to the right of -1 on the complex plane for any $\alpha_i > 0$.

Hence we only need to check that $\rho\left(\hat{R}(j\omega)\hat{R}^H(j\omega)\right) \le 1$ under the condition of Theorem 6.21. For any matrix A, $\rho(A) \le \|A\|$ for any induced matrix norm. Using the $\|\cdot\|_\infty$ norm (maximum absolute row sum), (6.22c), and Lemma 6.16, we have

$$\begin{aligned}
\rho\left(\hat{R}(j\omega)\hat{R}^H(j\omega)\right) &= \rho\left(\operatorname{diag}(g_l')\, R^f(j\omega)\, \operatorname{diag}\left(\frac{2}{\pi}\tau_i\kappa_i\right)\left(R^f(-j\omega)\right)^T\right) \\
&\le \left\|\operatorname{diag}(g_l')\, R^f(j\omega)\, \operatorname{diag}\left(\frac{2}{\pi}\tau_i\kappa_i\right)\left(R^f(-j\omega)\right)^T\right\|_\infty \\
&\le \left\|\operatorname{diag}(g_l')\, R^f(j\omega)\right\|_\infty \cdot \left\|\operatorname{diag}\left(\frac{2}{\pi}\tau_i\kappa_i\right)\left(R^f(-j\omega)\right)^T\right\|_\infty \\
&\le \max_l \sum_i \left| g_l'\, R_{li}\, e^{-j\tau_{li}^f\omega}\right| \cdot \max_i \sum_l \left|\frac{2}{\pi}\tau_i\kappa_i\, R_{li}\, e^{j\tau_{li}^f\omega}\right| \\
&= \frac{2}{\pi}\cdot \max_l g_l' \sum_i R_{li} \cdot \max_i \tau_i\kappa_i \sum_l R_{li} \\
&\le \frac{2\,\overline{N}\,\overline{L}}{\pi} \cdot \max_l g_l' \cdot \max_i \tau_i\kappa_i
\end{aligned}$$

Hence the condition in the theorem guarantees $\rho\left(\hat{R}(j\omega)\hat{R}^H(j\omega)\right) \le 1$. □

6.4 STABILITY OF DUAL ALGORITHMS

Consider the dual algorithm

$$\begin{aligned}
x_i(t) &= U_i^{\prime -1}(q_i(t)), & \dot{p}_l &= \gamma_l(y_l(t) - c_l)^+_{p_l(t)} \\
q_i(t) &= \sum_l R_{li}p_l(t-\tau_{li}^b), & y_l(t) &= \sum_i R_{li}x_i(t-\tau_{li}^f)
\end{aligned}$$

where $\gamma_l > 0$ and $U_i^{\prime -1}$ are nonnegative and strictly decreasing. Let (x^*, p^*) be an equilibrium and assume without loss of generality that $(x^*, p^*) > 0$. Using $\tilde{x} := x - x^*, \tilde{p} := p - p^*$, etc, the linearized system around (x^*, p^*) is, in the Laplace domain,

$$\begin{aligned}
\tilde{x}_i &= \frac{1}{U_i''(x_i^*)}\tilde{q}_i, & s\tilde{p}_l &= \gamma_l\tilde{y}_l \\
\tilde{q}_i &= \sum_l R_{li}e^{-\tau_{li}^b s}\tilde{p}_l, & \tilde{y}_l &= \sum_i R_{li}e^{-\tau_{li}^f s}\tilde{x}_i
\end{aligned}$$

As for the primal algorithm we use the delayed routing matrix $R^f(s)$ and $R^b(s)$ in (6.17) to write the dual algorithm compactly in matrix form:

$$\tilde{x} = \operatorname{diag}\left(\frac{1}{U_i''(x_i^*)}\right)\tilde{q}, \qquad \tilde{p} = \frac{1}{s}\Gamma\,\tilde{y} \tag{6.25a}$$

$$\tilde{q} = \left(R^b(s)\right)^T \tilde{p}, \qquad \tilde{y} = R^f(s)\,\tilde{x} \tag{6.25b}$$

where $\Gamma :=\operatorname{diag}(\gamma_l, l \in L) \succ 0$. This is represented as a feedback system in Figure 6.12 (note the minus sign in $H(s) := -\Gamma/s$ because of the negative feedback).

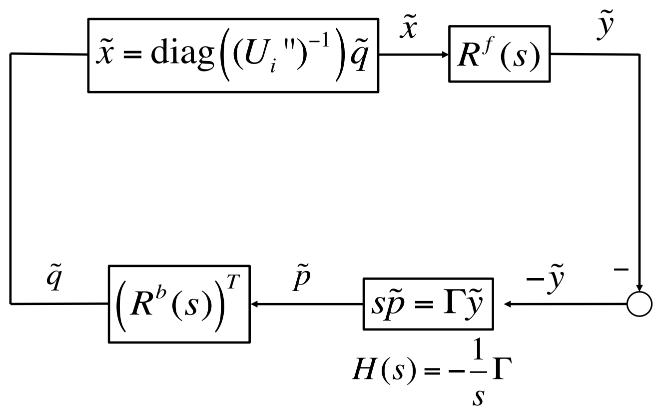


Figure 6.12: Delayed linear system representing the dual algorithm (6.25).

The loop transfer function from $\tilde{y}$ to $\tilde{y}$ is:

$$L(s) = R^f(s)\operatorname{diag}\left(\frac{1}{U_i''(x^*)}\right)(R^b(s))^T\frac{-\Gamma}{s}$$

Under condition C6.2 we can again relate the forward and backward routing matrix:

$$R^b(s) = R^f(-s)\operatorname{diag}\left(e^{-\tau_i s}\right)$$

Substituting into L, the loop transfer function becomes

$$L(s) = R^f(s)\operatorname{diag}\left(\frac{1}{-U_i''(x^*)}\right)\operatorname{diag}\left(\frac{e^{-\tau_i s}}{s}\right)(R^f(-s))^T\,\Gamma$$

Applying Lemma 6.22 we have

$$\sigma(L(s)) = \sigma(\hat{L}(s)) \tag{6.26a}$$

where

$$\hat{L}(s) \;=\; \hat{R}(s)\,\mathrm{diag}\left(\frac{\pi}{2}\frac{e^{-\tau_i s}}{\tau_i s}\right)\hat{R}^H(s) \tag{6.26b}$$

$$\hat{R}(s) \;:=\; \mathrm{diag}(\sqrt{\gamma_l})\,R^f(s)\,\mathrm{diag}\left(\sqrt{\frac{2}{\pi}\frac{\tau_i}{-U_i''(x^*)}}\right) \tag{6.26c}$$

The set of eigenloci here has the same structure as that in (6.22).

As for the primal algorithm we can treat $\hat{L}(s)$ as the loop function of a unity feedback system as shown in Figure 6.10(b) with $\hat{L}(s)$ given by (6.26b). For this equivalent unity feedback system the open-loop transfer function is $\hat{H}(s) = \hat{L}(s)$. From Lemma 6.22 we have

$$\sigma\left(\hat{R}(s)\,\mathrm{diag}\left(\frac{\pi}{2}\frac{e^{\tau_i s}}{\tau_i s}\right)\hat{R}^H(s)\right) \;\subseteq\; \rho\left(\hat{R}(s)\hat{R}^H(s)\right)\cdot \mathrm{co}\left(0,\,\frac{\pi}{2}\frac{e^{\tau_i s}}{\tau_i s}\right)$$

Since it is a unity feedback system, the closed paths obtained from the eigenloci of the loop function $L(s)$ will encircle -1 on the real axis p_{0+} times as s traverses the Nyquist path D with left indentations at the origin; see Chapter 6.2.6. Example 6.20 there also shows that the closed-loop system is asymptotically stable if

$$\rho\left(\hat{R}(j\omega)\hat{R}^H(j\omega)\right) \;<\; 1 \tag{6.27}$$

Then the eigenloci of the loop function $L(j\omega)$ stays entirely to the right of -1 on the real axis as ω traverse the imaginary axis from 0 to ∞ (see Corollary 6.10).

The condition (6.27) can be enforced by constraining algorithm parameters $\gamma_l, U_i''(x^*)$ based on feedback delays τ_i. For the primal algorithm considered in Theorem 6.21, the (sufficient) stability condition imposes upper bounds on source control gain κ_i and link price sensitivity $g_l'(y_l^*)$. These bounds become tighter as delay τ_i increase. Moreover the condition involves bounds $\overline{N}, \overline{L}$ that are likely to be conservative. To address these issues we now specialize to a dual algorithm for which the linearized closed-loop system remains asymptotically stable for *arbitrary* feedback delay and network topology. In other words the utility functions U_i are carefully chosen to obtain asymptotic stability that is scalable to arbitrary delay and link capacity.

Consider the utility functions

$$U_i(x_i) \;=\; \frac{M_i\tau_i}{\alpha_i}x_i\left(1-\log\left(\frac{x_i}{\overline{x}_i}\right)\right), \qquad x \le \overline{x}_i \tag{6.28a}$$

where τ_i is the round-trip time of source i, M_i is an upper bound on the maximum number of (bottleneck) links in the path of source i, $\alpha_i > 0$ is a design parameter, and $\overline{x}_i$ is the peak rate for source i. Scale the price function by the link capacity c_l:

$$\dot{p}_l \;=\; \frac{1}{c_l}(y_l(t)-c_l)^+_{p_l(t)} \tag{6.28b}$$

This price function can be implemented by having the source algorithm reacts to queueing delay as TCP Vegas and FAST do; see Chapter 1.3. Then the stability parameters in (6.26) are

$$U_i^{''}(x_i^*) \;=\; -\frac{M_i\tau_i}{x_i^*\alpha_i} \quad \text{and} \quad \gamma_l \;=\; \frac{1}{c_l} \tag{6.29}$$

Theorem 6.23 *Suppose conditions C6.1 and C6.2 hold. The origin of the linearized dual algorithm* (6.25) *with utility functions and link capacity scaling* (6.28) *is asymptotically stable as long as*

$$\max_i \alpha_i \;<\; \frac{\pi}{2}$$

Proof. As discussed above the theorem is proved if we can establish (6.27). Substitute (6.26c) into the left-hand side of (6.27) we have (again using Lemma 6.16)

$$\begin{aligned}
&\rho\left(\hat{R}(j\omega)\hat{R}^H(j\omega)\right)\\
=\;&\rho\left(\mathrm{diag}(\gamma_l)\,R^f(j\omega)\,\mathrm{diag}\left(\frac{2}{\pi}\frac{\tau_i}{U_i''(x^*)}\right)\left(R^f(-j\omega)\right)^T\right)\\
=\;&\rho\left(\mathrm{diag}\left(\frac{1}{c_l}\right)R^f(j\omega)\,\mathrm{diag}\left(-\frac{2}{\pi}\frac{x_i^*\alpha_i}{M_i}\right)\left(R^f(-j\omega)\right)^T\right)\\
=\;&\rho\left(\mathrm{diag}\left(\frac{1}{c_l}\right)R^f(j\omega)\mathrm{diag}\,(x_i^*)\cdot\mathrm{diag}\left(-\frac{2\alpha_i}{\pi}\right)\cdot\mathrm{diag}\left(\frac{1}{M_i}\right)\left(R^f(-j\omega)\right)^T\right)
\end{aligned}$$

where the last second equality follows from (6.29). Since the spectral radius of a matrix is no larger than any induced matrix norm, we use $\|\cdot\|_\infty$ (the maximum absolute row sum) to get

$$\begin{aligned}
&\rho\left(\hat{R}(j\omega)\hat{R}^H(j\omega)\right)\\
\leq\;&\left\|\mathrm{diag}\left(\frac{1}{c_l}\right)R^f(j\omega)\,\mathrm{diag}\,(x_i^*)\right\|_\infty\cdot\left\|\mathrm{diag}\left(-\frac{2\alpha_i}{\pi}\right)\right\|_\infty\cdot\left\|\mathrm{diag}\left(\frac{1}{M_i}\right)\left(R^f(-j\omega)\right)^T\right\|_\infty
\end{aligned}$$

Notice

$$\begin{aligned}
\left\|\mathrm{diag}\left(\frac{1}{c_l}\right)R^f(j\omega)\,\mathrm{diag}\,(x_i^*)\right\|_\infty &= \max_l \frac{1}{c_l}\sum_i\left|R_{li}\,e^{-j\tau_{li}^f\omega}\,x_i^*\right| \;\leq\; 1\\
\left\|\mathrm{diag}\left(\frac{1}{M_i}\right)\left(R^f(-j\omega)\right)^T\right\|_\infty &= \max_i \frac{1}{M_i}\sum_l\left|R_{li}\,e^{j\tau_{li}^f\omega}\right| \;\leq\; 1\\
\left\|\mathrm{diag}\left(-\frac{2\alpha_i}{\pi}\right)\right\|_\infty &= \max_i \frac{2\alpha_i}{\pi} < 1
\end{aligned}$$

Hence $\rho\left(\hat{R}(j\omega)\hat{R}^H(j\omega)\right) < 1$ and the proof is complete. □

6.5 APPENDIX: PROOF OF THEOREM 6.14

Factorize the system matrix A in (6.12b) of the closed-loop system into:

$$\begin{bmatrix} sI - A_1 & B_1C_2 \\ -B_2C_1 & sI - A_2 \end{bmatrix} = \begin{bmatrix} sI - A_1 & 0 \\ 0 & sI - A_2 \end{bmatrix} \begin{bmatrix} I & (sI - A_1)^{-1}B_1C_2 \\ -(sI - A_2)^{-1}B_2C_1 & I \end{bmatrix}$$

Apply the Schur's determinant identity (proved below)

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det D \cdot \det(A - BD^{-1}C) \tag{6.30}$$

to obtain

$$\begin{aligned} & \det(sI - A) \\ = \; & \det(sI - A_1) \cdot \det(sI - A_2) \cdot \det\left(I + (sI - A_1)^{-1}B_1C_2(sI - A_2)^{-1}B_2C_1\right) \end{aligned}$$

From Lemma 6.16 we therefore have

$$\begin{aligned} & \det(sI - A) \\ = \; & \det(sI - A_1) \cdot \det(sI - A_2) \cdot \det\left(I + C_1(sI - A_1)^{-1}B_1\, C_2(sI - A_2)^{-1}B_2\right) \\ = \; & \det(sI - A_1) \cdot \det(sI - A_2) \cdot \det\left(I + H_1(s)\, H_2(s)\right) \end{aligned}$$

We are left to derive the Schur's determinant identity (6.30). Perform two successive block Gaussian eliminations to eliminate the off-diagonal blocks:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix} = \begin{bmatrix} A - BD^{-1}C & B \\ 0 & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix}$$

Hence we have

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I & 0 \\ D^{-1}C & I \end{bmatrix}$$

from which (6.30) follows. This completes the proof of Theorem 6.14. □

6.6 BIBLIOGRAPHICAL NOTES

The control theory of multi-input-multi-output (MIMO) systems is much subtler than that of single-input-single-output (SISO) systems, and we have only presented a minimal set of results that are needed to analyze local stability of TCP congestion control algorithms in the presence of feedback delay. See, e.g. [13, 18, 38] for a detailed exposition. When A is diagonalizable, the eigenvector dyadic expansion of A leads to a transparent structure of the transfer function matrix and its poles, as explained in Lemma 6.11 and Theorem 6.12, both taken from [13, Chapter 3.3]. Theorems 6.14 and 6.15 are taken from [13, Chapter 11]. Theorem 6.18 is proved in [19]. Even though our discussion assumes a proper rational loop function L, Theorem 6.18 extends to more general proper functions, e.g., for systems involving delay or trigonometric functions; see [19, Section III]. The dual algorithm in Theorem 6.23 is proposed and analyzed in [40].

6.7 PROBLEMS

Exercise 6.1. Consider the single-input single-output LTI system in Example 6.2 specified by the transfer function:

$$H(s) \quad := \quad \frac{y(s)}{u(s)} \quad = \quad \frac{s-1}{(s+1)(s+2)}$$

Consider realizations (A, B, C) with $n = 2$ states:

$$A \; := \; \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B \; := \; \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}, \quad C \; := \; \begin{bmatrix} c_1 & c_2 \end{bmatrix}$$

Suppose we restrict $B = [1 \;\; 0]^T$. Show that all realizations (A, B, C) are controllable.

Exercise 6.2. Consider the single-input single-output LTI system in Exercise 6.1. Suppose we restrict $C = [1 \;\; 0]$. Show that all realizations (A, B, C) are observable.

Exercise 6.3. The transfer function in Exercise 6.1 is strictly proper and is realized in Example 6.3 by a different implementation:

$$\begin{aligned} \dot{x}_1 &= -x_1(t) - 2u(t) \\ \dot{x}_2 &= -2x_2(t) + 3u(t) \\ y(t) &= x_1(t) + x_2(t) \end{aligned}$$

Is this realization controllable? Observable?

Exercise 6.4. Consider the second transfer function in Example 6.3:

$$H(s) \quad = \quad \frac{c(s-b_1)(s-b_2)}{(s-a_1)(s-a_2)}$$

where $a_i \neq b_j$, i.e., the numerator and the denominator are co-prime. This transfer function is proper but not strictly proper. Is the following realization in Example 6.3

$$A \; := \; \begin{bmatrix} a_1 & 0 & 0 & 0 \\ 0 & a_2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -a_1 a_2 & a_1 + a_2 \end{bmatrix}, \quad B \; := \; \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ 0 \\ \Delta_1 \Delta_2 \end{bmatrix}, \quad C^T \; := \; \begin{bmatrix} c \\ c \\ c \\ 0 \end{bmatrix}, \quad D \; := \; c$$

controllable? Observable?

Exercise 6.5. Design a state-space realization of the MIMO transfer function in Example 6.4.

Exercise 6.6. Consider the unity feedback system in Figure 6.8 with the open-loop transfer function $H(s) = C^{\text{ol}}(sI - A^{\text{ol}})^{-1}B^{\text{ol}}$. Show that the characteristic polynomial of the system matrix A^{cl} of the closed-loop system satisfies:

$$\det(sI - A^{\text{cl}}) \quad = \quad \det(sI - A^{\text{ol}}) \cdot \det\left(I + H(s)\right)$$

This implies that the Nyquist stability results (Theorems 6.14, 6.15, 6.18 and Corollary 6.19) hold with the loop function $L(s) = H(s)$ for unity feedback systems.

Exercise 6.7. Consider the feedback connection in Figure 6.13 and let the component

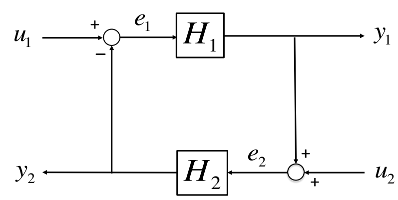


Figure 6.13: Feedback connection.

systems be

$$\dot{x}_i \quad = \quad A_i x_i + B_i e_i, \qquad y_i \quad = \quad C_i x_i, \qquad i = 1, 2$$

where $x_i \in \mathbb{R}^{n_i}$, $e_i \in \mathbb{R}^{m_i}$, $y_1 \in \mathbb{R}^{m_2}$, $y_2 \in \mathbb{R}^{m_1}$, $A_i \in \mathbb{R}^{n_i \times n_i}$, $B_1 \in \mathbb{R}^{n_1 \times m_1}$, $B_2 \in \mathbb{R}^{n_2 \times m_2}$, $C_1 \in \mathbb{R}^{m_2 \times n_1}$, and $C_2 \in \mathbb{R}^{m_1 \times n_2}$. The external inputs u_i are thus in $\mathbb{R}^{m_i}$. The transfer functions of the component systems are

$$H_i(s) \quad = \quad C_i(sI - A_i)^{-1}B_i, \qquad i = 1, 2$$

The state of the closed-loop system is $x := [x_1^T \;\; x_2^T]^T \in \mathbb{R}^{n_1+n_2}$, input is $u := [u_1^T \;\; u_2^T]^T \in \mathbb{R}^{m_1+m_2}$, and output is $y := [y_1^T \;\; y_2^T]^T \in \mathbb{R}^{m_1+m_2}$. Show that the closed-loop system is also LTI, described by

$$\dot{x} \quad = \quad Ax(t) + Bu(t), \qquad y \quad = \quad Cx(t)$$

where

$$A \; := \begin{bmatrix} A_1 & -B_1C_2 \\ B_2C_1 & A_2 \end{bmatrix} \qquad B \; := \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} \qquad C \; := \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix}$$

Exercise 6.8 ([49])**.** Let $P = P^* \succ 0$ and $\Lambda = \text{diag}(\lambda_i)$ be $n \times n$ matrices. Then the set of eigenvalues $\sigma(P\Lambda) \subseteq \rho(P) \cdot \text{co}\{0, \lambda_i\}$ where $\rho(P)$ is the spectral radius of P and $\text{co}\, A$ is the convex full of set A.

Exercise 6.9 ([49]). Consider the following primal algorithm

$$\begin{aligned} \dot{x}_i(t) &= \kappa_i(w_i - x_i(t-\tau_i)q_i(t)), & p_l(t) &= g_l(y_l(t)) \quad (6.31a) \\ q_i(t) &= \sum_l R_{li}p_l(t-\tau_{li}^b), & y_l(t) &= \sum_i R_{li}x_i(t-\tau_{li}^f) \quad (6.31b) \end{aligned}$$

Derive sufficient condition for linearized stability. (Hint: The difference between this primal algorithm and that in the chapter is the product $x_i(t-\tau_i)q_i(t)$. Consider the variable $z_i(t) := x_i(t-\tau_i)q_i(t)$.)

Exercise 6.10. Is the linear model of the dual algorithm (6.25) a minimal realization, assuming time delays are zero?

Exercise 6.11. Consider the following system

$$\begin{aligned} \dot{x}_1(t) &= y_1(t) - y_2(t-\tau_{12}) - u_1(t) \\ \dot{x}_2(t) &= y_2(t) - y_1(t-\tau_{21}) - u_2(t) \\ \dot{y}_i(t) &= \gamma_i x_i(t), \qquad i = 1, 2 \end{aligned}$$

where $\tau_{12} > 0$, $\tau_{21} > 0$, $\gamma_i > 0$.

1. Prove that we can choose γ_i so that the closed-loop system with zero input $u_i(t) \equiv 0$ is stable, but not asymptotically stable.

2. Design state feedback $u_i(t) = u(x_i(t))$ so that, with appropriate choice of γ_i, the closed-loop system is asymptotically stable. (Hint: Try $u_i(x_i) = a_i x_i$ for some $a_i > 0$.)

Bibliography

[1] Hobbes' Internet timeline 23. http://www.zakon.org/robert/internet/timeline/. Accessed April 30, 2016.

[2] US Department of Energy. https://www.es.net/engineering-services/the-network/network-maps/historical-network-maps/. Accessed May 11, 2016.

[3] K. J. Arrow, L. Hurwicz, H. Uzawa, and H. B. Chenery. *Studies in linear and non-linear programming*. Stanford University Press Stanford, 1958.

[4] Sanjeewa Athuraliya, Victor H. Li, Steven H. Low, and Qinghe Yin. REM: active queue management. *IEEE Network*, 15(3):48–53, May/June 2001. Extended version in *Proceedings of ITC17*, Salvador, Brazil, September 2001. http://netlab.caltech.edu.

[5] Sanjeewa Athuraliya and Steven H. Low. Optimization flow control with Newton–like algorithm. *Journal of Telecommunication Systems*, 15(3/4):345–358, 2000.

[6] Andrea Bacciotti and Francesca Ceragioli. Nonpathological lyapunov functions and discontinuous Carathéodory systems. *Automatica*, 42:453–458, 2006.

[7] Dimitri P. Bertsekas. *Nonlinear Programming*. Athena Scientific, 1995.

[8] Dimitri P. Bertsekas and John N. Tsitsiklis. *Parallel and distributed computation*. Prentice-Hall, 1989.

[9] Dimitris Bertsimas, Vivek F. Farias, and Nikolaos Trichakis. On the efficiency-fairness trade-off. *Journal of Management Science*, 58(2):2234–2250, 2012.

[10] T. Bonald and L. Massoulie. Impact of fairness on Internet performance. In *Proceedings of ACM Sigmetrics*, pages 82–91, June 2001.

[11] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[12] Lawrence S. Brakmo and Larry L. Peterson. TCP Vegas: end-to-end congestion avoidance on a global Internet. *IEEE Journal on Selected Areas in Communications*, 13(8):1465–80, October 1995. http://cs.princeton.edu/nsg/papers/jsac-vegas.ps.

[13] Frank M. Callier and Charles A. Desoer. *Linear System Theory*. Springer-Verlag, 1991.

[14] V. G. Cerf and R. E. Kahn. A protocol for network intercommunication. *IEEE Transactions on Communications*, 22(5), May 1974.

[15] Ashish Cherukuri, Enrique Mallada, and Jorge Cortés. Asymptotic convergence of constrained primal-dual dynamics. *Systems and Control Letters*, 2016.

[16] Douglas E. Comer. *Internetworking with TCP/IP, I: principles, protocol, and architecture*. Pearson, 6 edition, 2013.

[17] Jorge Cortés. Discontinuous dynamical systems: a tutorial on solutions, nonsmooth analysis, and stability. *IEEE Control Systems Magazine*, 28(3):36–73, 2008.

[18] Mohammed Dahleh, Munther A. Dahleh, and George Verghese. Lectures on dynamic systems and control. Lecture notes for MIT OpenCourseWare, 2011.

[19] C. A. Desoer and Y. T. Wang. On the generalized Nyquist stability criterion. *IEEE Trans. on Automatic Control*, 25(2):187–196, April 1980.

[20] P. Dupuis and A. Nagurney. Dynamical systems and variational inequalities. *Annals of Operations Research*, 44:9–42, 1993.

[21] D. Feijer and F. Paganini. Stability of primal-dual gradient dynamics and applications to network optimization. *Automatica*, 46(12):1974–1981, 2010.

[22] S. Floyd and V. Jacobson. Random early detection gateways for congestion avoidance. *IEEE/ACM Trans. on Networking*, 1(4):397–413, August 1993. `ftp://ftp.ee.lbl.gov/papers/early.ps.gz`.

[23] William S. Hall and Martin L. Newall. The mean value theorem for vector-valued functions: a simple proof. *Mathematics Magazine*, 52:157–158, 1979.

[24] Juha Heinonen. Lectures on Lipschitz analysis. Technical report, University of Jyväskylä, Finland, 2005.

[25] Chris Hollot, Vishal Misra, Don Towsley, and Wei-Bo Gong. On designing improved controllers for AQM routers supporting TCP flows. In *Proceedings of IEEE Infocom*, April 2001. `http://www-net.cs.umass.edu/papers/papers.html`.

[26] V. Jacobson. Congestion avoidance and control. *Proceedings of SIGCOMM'88, ACM*, August 1988. An updated version is available via `ftp://ftp.ee.lbl.gov/papers/congavoid.ps.Z`.

[27] R. Jain, K. K. Ramakrishnan, and D. M. Chiu. Congestion avoidance in computer networks with a connectionless network layer. Technical Report DEC-TR-506, Digital Equipment Corporation, August 1987.

[28] Frank P. Kelly, Aman Maulloo, and David Tan. Rate control for communication networks: Shadow prices, proportional fairness and stability. *Journal of Operations Research Society*, 49(3):237–252, March 1998.

[29] Hassan K. Khalil. *Nonlinear systems.* Prentice Hall, 3rd edition, 2002.

[30] S. Kunniyur and R. Srikant. End-to-end congestion control: utility functions, random losses and ECN marks. *IEEE/ACM Transactions on Networking*, 11(5):689 – 702, October 2003.

[31] James F. Kurose and Keith W. Ross. *Computer networking: a top-down approach.* Pearson, 6 edition, 2013.

[32] Tian Lan, David Kao, Mung Chiang, and Ashutosh Sabharwal. An axiomatic theory of fairness in network resource allocation. In *Proc. IEEE Infocom*, March 2010.

[33] Na Li, Changhong Zhao, and Lijun Chen. Connecting automatic generation control and economic dispatch from an optimization view. *IEEE Trans. on Control of Network Systems*, 3(3):254–264, September 2016.

[34] Steven H. Low. A duality model of TCP and queue management algorithms. *IEEE/ACM Trans. on Networking*, 11(4):525–536, August 2003. `http://netlab.caltech.edu`.

[35] Steven H. Low and David E. Lapsley. Optimization flow control, I: basic algorithm and convergence. *IEEE/ACM Transactions on Networking*, 7(6):861–874, December 1999. `http://netlab.caltech.edu`.

[36] Steven H. Low, Fernando Paganini, and John C. Doyle. Internet congestion control. *IEEE Control Systems Magazine*, 22(1):28–43, February 2002.

[37] Steven H. Low, Larry Peterson, and Limin Wang. Understanding Vegas: a duality model. *J. of ACM*, 49(2):207–235, March 2002. `http://netlab.caltech.edu`.

[38] J. M. Maciejowski. *Multivariable feedback design.* Addison-Wesley, 1989.

[39] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, 8(5):556–567, October 2000.

[40] Fernando Paganini, Zhikui Wang, John C. Doyle, and Steven H. Low. Congestion control for high performance, stability and fairness in general networks. *IEEE/ACM Transactions on Networking*, 13(1), February 2005.

[41] Walter Rudin. *Principles of mathematical analysis*. McGraw-Hill, Inc., 3 edition, 1976.

[42] Srinivas Shakkottai and R. Srikant. *Network Optimization and Control*, volume 2 of *Foundations and Trend in Networking*. NOW Publisher, 2008.

[43] Sigurd Skogestad and Ian Postlethwaite. *Multivariable feedback control: analysis and design*. John Wiley & Sons, 1996.

[44] R. Srikant. *The Mathematics of Internet Congestion Control*. Birkhauser, 2004.

[45] W. Stevens. *TCP/IP illustrated: the protocols*, volume 1. Addison–Wesley, 1999. 15th printing.

[46] A. Tang, J. Wang, S. H. Low, and M. Chiang. Equilibrium of heterogeneous congestion control: Existence and uniqueness. *IEEE/ACM Transactions on Networking*, 15(4):824–837, August 2007.

[47] A. Tang, X. Wei, S. H. Low, and M. Chiang. Equilibrium of heterogeneous congestion control: optimality and stability. *IEEE/ACM Transactions on Networking*, 15(4):824–837, August 2010.

[48] Ao Tang, Jiantao Wang, and Steven H. Low. Counter-intuitive throughput behavior of networks under end-to-end control. *IEEE/ACM Transactions on Networking*, April 2006.

[49] Glenn Vinnicombe. On the stability of end-to-end congestion control for the Internet. Technical report, Cambridge University, CUED/F-INFENG/TR.398, December 2000.

[50] D. X. Wei, C. Jin, S. H. Low, and S. Hegde. TCP FAST: motivation, architecture, algorithms, performance. *IEEE/ACM Transactions on Networking*, 14(6):1246–1259, December 2006.

[51] E. Wei, A. Ozdaglar, and A. Jadbabaie. A distributed newton method for network utility maximization, I: Algorithm. *IEEE Transactions on Automatic Control*, 58(9):2162–2175, 2013.

[52] E. Wei, A. Ozdaglar, and A. Jadbabaie. A distributed newton method for network utility maximization, II: convergence. *IEEE Transactions on Automatic Control*, 58(9):2176–2188, 2013.

[53] J.T. Wen and M. Arcak. A unifying passivity framework for network flow control. *IEEE Transaction on Automatic Control*, 49(2):162–174, February 2004.

[54] Bartek Wydrowski and Moshe Zukerman. MaxNet: A congestion control architecture for maxmin fairness. *IEEE Communications Letters*, 6(11):512–514, November 2002.

[55] M. Zargham, A. Ribeiro, A. Ozdaglar, and A. Jadbabaie. Accelerated dual descent for network optimization. *IEEE Transactions on Automatic Control*, 59(4):905–920, April 2014.

[56] C. Zhao, T. Topcu, N. Li, and S. H. Low. Design and stability of load-side primary frequency control in power systems. *IEEE Trans. on Automatic Control*, 59(5):1177–1189, May 2014.