University of Cincinnati

Date: 1/29/2015

I, Sadiqah Al Marzoog Ph.D., hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in Mathematical Sciences.

It is entitled:
Qualitative Analysis of Biofilms in Water Networks

Student's name: Sadiqah Al Marzoog Ph.D.

This work and its defense approved by:

Committee chair: Benjamin Vaughan, Ph.D.
Committee member: Donald French, Ph.D.
Committee member: Stephan Pelikan, Ph.D.
Qualitative Analysis of Biofilms in Water Networks

A dissertation submitted to the Graduate School of the

University of Cincinnati

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the Department of Mathematical Sciences
of the College of Arts and Sciences by

Sadiqah Al Marzooq

MS Financial Mathematics–University of Dayton
BS Mathematical Science–Science College of Dammam

January 29th, 2015

Committee Chair: Benjamin L. Vaughan, Ph.D.
Committee Members: Donald A. French, Ph.D. and Stephan Pelikan, Ph.D.
Abstract

In industrialized nations, the availability of potable water is provided through sophisticated water distribution systems that consist a network of pipes, pumping stations, reservoirs, and other components that deliver safe drinking water to consumers. Biological contamination of the water distribution system, due to factors such as a breach in a pipe or other causes, can have an adverse affect on water quality. In this thesis, we develop and analyze introductory mathematical models that provides qualitative predictions of the dynamics of a release of pathogenic bacterial species into the water distribution system.

Bacteria within a water distribution system frequently form biofilm on the interior surfaces of pipes. Biofilms are aggregates of various microorganisms that attach themselves to solid surfaces using self-secreted extracellular polymeric substances (EPS). The presence of a biofilm can affect the hydraulic roughness and the material life of a pipe, promote microbial-influenced corrosion, induce taste and/or odour changes in drinking water, reduce the effectiveness of disinfectants, and increase bacterial levels. To study the dynamics of planktonic (individual cells floating or swimming in the liquid) and biofilm-bound bacteria within the large network of pipes within a water distribution system, we developed compartmental models governing the concentration of introduced pathogens within the bulk fluid and the biofilms lining a network of pipes. These mathematical models are studied under time-constant and time-periodic flow regimes using Lyapunov stability and Floquet theory as well as numerical simulation. We developed efficient algorithms for predicting the long-time behavior of the pathogen population within the network and prove mathematically these predictions are equivalent to results from Lyapunov and Floquet theory. The analytical results are validated using numerical simulations of the full non-linear system on a range of water distribution network sizes.
Acknowledgements

My sincere gratitude and respect goes to my committee chair and advisor, Professor Benjamin L. Vaughan, for his hard efforts, guidance and time he spent with me to develop my background in biofilms and water distribution networks. I really appreciate his extreme patience and helping in MATLAB programming and writing this thesis. I admit that this work would not have been done without his continuous support and I look forward to cooperate with him other research projects in future.

It is a pleasure to thank my other advanced exam committee members, Professor Donald French and Professor Sookkyung Lim, for their insightful comments and discussions during the advanced exam. I am thankful to Professor French in particular for his valuable classes in numerical analysis and for reviewing some theorems in this thesis. I would like also to thank the Department of Mathematics at University of Cincinnati, especially Professor Timothy Hodges, Graduate Program Director, and Professor Magda Peligrad for their inseparable supports and encouragement.

I am grateful for the High Education Ministry at Saudi Arabia for their funding source that gave me the opportunity to study at graduate schools in the U.S. over seven years and to achieve my objectives.

My special acknowledgement to my parents, Samerah and Abdullah, who raised me with love, taught me how to work hard and pursue my goals until I reach it. Thanks a lot my parents, my brothers and sisters for your faith in me, your love and prayers. For most, I would like to thank my little brother and supportive, Abdullkareem, for his unyielding help and inspiration.

Lastly, but not the least, I cannot express my appreciation and respect to my husband, Zuhair, for his love, patience, and caring through my educational life. I must also thank my kids Abdullah, Zainab, and Ahmad. You are really my joy.
Contents

Abstract iii

Acknowledgments iv

List of Figures viii

List of Tables ix

1 Introduction 1

2 Constant Flow Network 5

2.1 Single Connection Model 5

2.1.1 Mathematical Model 5

2.1.2 Steady States and Stability 7

2.1.3 Numerical Results 9

2.2 Constant Flow Network Model 10

2.2.1 Mathematical Model 12

2.2.2 Determining the Flow Rate 18

2.2.3 Theorems 21

2.2.4 Efficient Algorithm for Determining the Pathogen Persistence within the Network 27

2.2.5 Theoretical and Numerical Results 29

2.2.6 Discussion 41

3 Periodic Flow Network 43

3.1 Single Connection Model 43
**List of Figures**

2.1 Compartmental Model of Pathogen Dynamics ........................................ 6
2.2 Transcritical Bifurcation Diagram for a Single Connection ......................... 8
2.3 Phase Portraits for the Nonlinear System Governing a Single Connection ........ 9
2.4 Numerical Solutions for a Single Connection ........................................... 11
2.5 Model Water Distribution Network .......................................................... 12
2.6 Compartmental Network Model of Pathogen Dynamics ............................... 16
2.7 Example Network .................................................................................. 20
2.8 Simple Flow Network with a Loop ............................................................ 21
2.9 Numerical Solution for Two Connection Network–Washout/Washout Case .... 32
2.10 Numerical Solution for Two Connection Network–Washout/Persist Case ...... 32
2.11 Numerical Solution for Two Connection Network–Persist/Persist Case .......... 33
2.12 Numerical Solution for Two Connection Network–Persist/Induced Case ...... 34
2.13 Flow Through the Medium-Size Network Example ................................... 35
2.14 Linear Stability for Large Network Example .............................................. 36
2.15 Steady State Numerical Solution for the Medium-Size Network Example .... 37
2.16 Flow Through the Large Network Example ............................................... 38
2.17 Linear Stability for Large Network Example .............................................. 39
2.18 Steady State Numerical Solution for the Large Network Example ................. 40
3.1 Bifurcation Diagram for a Single Connection with Periodic Flow ................. 49
3.2 Bifurcation Diagram for a Single Connection with Periodic Flow for Various $\omega$ 50
3.3 Bifurcation Diagram for a Single Connection with Periodic Flow for Various $\alpha$ and $\beta$ with Fixed $Q_c$ ............................................................................. 51
3.4 Numerical Solutions for a Single Connection with a Periodic Flow ........ 53
3.5 Numerical Solutions for a Two Pipe Network with Periodic Flows--Washout/Washout Case ............................................................................................................. 71
3.6 Numerical Solutions for a Two Pipe Network with Periodic Flows--Washout/Persist Case ............................................................................................................. 72
3.7 Numerical Solutions for a Two Pipe Network with Periodic Flows--Persist/Persist Case ............................................................................................................. 73
3.8 Numerical Solutions for a Two Pipe Network with Periodic Flows--Persist/Induced Case .................................................................................................................. 74
3.9 Flow Through the Medium-Sized Network Example with Periodic Flow Rates 75
3.10 Stability for Medium Network Example .................................................. 76
3.11 Steady State Numerical Solution for the Medium Network Example .......... 77
3.12 Flow Through the Large Network Example with an Unidirectional Periodic Flow ...................................................................................................................... 78
3.13 Linear Stability for Large Network Example with an Unidirectional Periodic Flow ...................................................................................................................... 79
3.14 Steady State Numerical Solution for the Large Network Example with an Unidirectional Flow ........................................................................................................ 80
A.1 Connectivity for the Medium-Sized Network Example .......................... 89
B.1 Connectivity of the Large Network Example ............................................ 91
C.1 Connectivity of the Medium Periodic Network Example ......................... 93
List of Tables

2.1 Stability of Two Pipe System ........................................... 31

3.1 Stability of Two Pipe Periodic System ................................. 69

A.1 Attachment and Detachment Rates for the Medium Network Example with Time-Constant Flow ........................................... 90

B.1 Demands in Large Network Example .................................... 92

B.2 Pressures in Large Network Example ................................... 92

D.1 Pressures in Large Network Example ................................... 95
Chapter 1

Introduction

The availability of potable (drinking) water is essential for human health and wellness. In industrialized nations, potable water is typically delivered through drinking water distribution systems. These systems consist of a network of pipes along with storage facilities (reservoirs, water towers, etc.) and pressurized components (pumping stations) to deliver the water to consumers. Water quality is a measure of condition of water within the network relative to some standard for the safety of the water for human contact and consumption. Contaminants such as microorganisms, salts, metals, pesticides, herbicides, pharmaceuticals, and radioactive materials can negatively affect the water quality. In particular, biological contamination from accidental exposure to fecal matter due to cracks, back-flow incidents, unsecured reservoirs, or leaching pipes can create a significant health risk [LeChevallier et al., 2003, Momba et al., 2000, Uber et al., 2004, Murray et al., 2006]. The presence of coliform bacteria such as *Escherichia Coli*, are used as a marker for fecal contamination. Waterborne diseases, such as botulism (*Clostridium botulinum*), cholera (*Vibrio cholerae*), dysentery (*Shigella dysenteriae*), and typhoid fever (*Salmonella typhi*), are thought to be responsible for 1.8 million deaths every year with an estimated 88% due to unsafe water supplies [World Health Organization et al., 2002].

There has been extensive research on design of water distribution networks to optimize flow rates and water quality while reducing costs as well as optimal maintenance of existing
systems. For an introduction to water engineering, see Mays [2010] and Davis [2011]. Of particular interest is the dynamics of non-native bacteria, possibly pathogenic, released into the water distribution system. Deterioration of water quality as it passes through drinking water distribution systems, especially as it relates to microbial contamination, has been well documented [LeChevallier et al., 1996, van der Kooij, 1992, Carabet et al., 2011, Camper et al., 1996, Power and Nagy, 1999, Carter et al., 2000, Besner et al., 2002, Skrabber et al., 2005, Zhang and DiGiano, 2002].

In water pipes, indigenous bacteria colonize the solid surfaces in structures called biofilms. Biofilms are aggregates of various microorganisms that attach themselves to solid surfaces using secreted extracellular polymeric substances (EPS). Biofilms are ubiquitous and are a crucial element in a wide range of harmful as well as beneficial phenomena including tooth decay, remediation of wastewater, maintenance of naval ships, bacterial growth on surgical implants, and complications in cystic fibrosis patients. While the existence of heterotrophic biofilm in a drinking water distribution system does not directly correlate with human health risk [Bartram, 2003], but the presence of biofilm may encourage the retention and possibly even growth of pathogens released within the system.

Bacteria within a biofilm have a distinct survival advantage compared with planktonic (free swimming) bacteria in a water distribution system. It has been observed that biofilms assist in the survival of microorganisms under inhospitable conditions found within distribution systems through improved nutrient absorption and improved resistance to disinfectants [LeChevallier et al., 1996, Camper, 1998, Szabo, 2006]. In particular, the improved resistance to disinfectants makes the removal of potentially pathogenic bacteria difficult. The attachment of planktonic pathogens to an established biofilm can lead to a reservoir of contamination [Camper, 1998] where pathogens survive disinfection and possibly grow, reappearing long after their initial introduction into the system [Camper and McFeters, 2000, Szabo, 2006] when detachment occurs and mobilizes them once again within the distribution system. Changing flows related to changes in instantaneous demand in water distribution
systems have been shown to contribute to detachment of biofilm and increases in water turbidity [Lehtola et al., 2006].

Research on the growth and structure of bacterial biofilms has appeared over the last four decades. A classic resource on the subject is the Costerton et al. review in Annual Review of Microbiology [Costerton et al., 1995], and, more recently, Flemming and Wingender [Flemming and Wingender, 2010] provide a comprehensive discussion of the biofilm matrix in Nature Reviews: Microbiology. There has also been extensive research in the mathematical modeling of biofilms. A comprehensive resource is the review by the IWA Task force on Biofilm Modeling [Eberl et al., 2006]. A key mechanism in the dynamics of non-native bacteria is the attachment/detachment processes which involve the exchange of bacteria from the biofilm to the fluid (detachment) and vice versa (attachment). In the case of detachment, when large scale pieces of the biofilm can separate from the mas in a process called sloughing [van Loosdrecht et al., 2002, Trulear and Characklis, 1982]. When the detachment involves individual or small numbers of bacteria, the process is called erosion [van Loosdrecht et al., 2002].

Simple mathematical erosion models have been implemented using individual-based biofilm models [Xavier et al., 2004, 2005b,a] where the detachment process is assumed to be as a simple first order transfer of material between the biofilm and the fluid and also in the original Wanner-Gujer model of biofilm growth [Wanner and Gujer, 1986], where erosion is modeled by an ad-hoc limiting term in the biofilm growth equation. More advanced models examine sloughing by incorporating the interaction between fluid shear forces on an elastic model of the biofilm in mechanical simulations [Picioreanu et al., 2001, Laspldou et al., 2005, Duddu et al., 2009]. In these models, large parts of the biofilm detach when the mechanical stresses exceed the adhesive strength of the EPS.

Of particular interest is the macroscale dynamics of non-native bacteria introduced into a water distribution network. The microscale and mesoscale dynamics of biofilms have been studied extensively theoretically [Eberl et al., 2006], there is a need to incorporate this re-
search with the larger scale network dynamics. This requires network-level models of fluid
dynamics within the network [Rossman and Boulos, 1996, Rossman, 2000, Eiger et al., 1994]coupled with models of various complexity describing the interaction between the network
topology, fluid mechanics of the flow, the planktonic bacteria, and the biofilms. Previous
work on modeling the dynamics of biological contaminants have focus on quantitative behav-
ior such as numerical measures of water quality [Uber et al., 2004, Murray et al., 2006]. We
are interested in qualitative predictions of the system’s behavior such as possible colonization
of indigenous biofilms within the network by pathogens, leading to persistent biofouling of
the drinking water, and the predictions of the critical pipes that dominate the contamination.

This thesis is organized as follows. We introduce a set of time-constant fluid flow mod-
els governing the dynamics of non-native bacteria in single pipes and within a network in
Chapter 2. These models are studied using Lyapunov stability and numerical simulations
in Section 2.1.2 for the single pipe model and Section 2.2 for the network model. Since
flows within realistic water distribution networks vary in time due to changes in demand,
we introduce a time-periodic flow model for a single pipe and for a network in Chapter 3.
These models are studied using Floquet theory and numerical simulations in Section 3.1.4
for a single pipe model and Section 3.2.1 the network model.
Chapter 2

Constant Flow Network

We will begin by studying a model water distribution network where the flow is constant in time. The main goal is to understand how the presence of a bacterial biofilm affects the retention time of introduced pathogens within the system and to predict the conditions where the pathogens might persist, affecting water quality. We will introduce non-linear system of ordinary differential equations in this chapter to understand the fundamental mechanisms and then analyze the stability of the system using linear stability analysis.

2.1 Single Connection Model

To begin to understand the dynamics a pathogen release within a large water distribution network, we begin by examining a single connection within the network. This connection can consist of a single pipe or a collection of pipes in serial which is treated as a single unit.

2.1.1 Mathematical Model

For a single connection, we use the well-mixed assumption to divide the concentration of pathogens is into two compartments: planktonic pathogens within the bulk fluid, $C_f$, and biofilm attached pathogens, $C_b$. The concentration within each compartment varies due to the following mechanisms (see Figure 2.1): attachment to the biofilm from the bulk fluid,
detachment from the biofilm into the bulk fluid, transport out the connection due to advection with the effluent, and logistic growth within the biofilm. Here, we neglect growth/decay within the bulk fluid and assume that there are no pathogens within in influent. For simplicity, we additionally assume that the attachment/detachment rates are independent of the flow. Under these assumptions, the concentration of the pathogens is governed by the following system of differential equations:

\[
\begin{align*}
C'_f &= -\frac{Q}{V}C_f + \beta \frac{S_A}{V}C_b - \alpha C_f \\
C'_b &= \alpha \frac{V}{S_A}C_f - \beta C_b + rC_b \left(1 - \frac{C_b}{K}\right)
\end{align*}
\]

where \(Q\) is the flow rate in hr\(^{-1}\), \(V\) is the fluid volume in the connection in mL, \(S_A\) is the surface area of the connection in cm\(^2\), \(\alpha\) is the attachment rate in hr\(^{-1}\), \(\beta\) is the detachment rate in hr\(^{-1}\), \(r\) is the logistic growth rate in hr\(^{-1}\), and \(K\) is the pathogen carrying capacity of the biofilm in CFU/cm\(^2\). The units of \(C_f\) are CFU/cm\(^3\) and \(C_b\) are CFU/cm\(^2\). Here, we assume that all parameters are positive constants. Initially, we assumed that there was an introduction of pathogens into the bulk fluid and there were no pathogens initially within the biofilm. Hence, the initial conditions are \(C_f(0) = C_0\) and \(C_b(0) = 0\).

Using the scales \(C_f = K (S_A/V) u\), \(C_b = Kv\), and \(t = r^{-1} \tau\), the dimensionless equations
are
\[
\begin{align*}
    u' &= \beta v - \left( \hat{\alpha} + \hat{Q} \right) u \\
    v' &= \hat{\alpha} u - \hat{\beta} v + v \left( 1 - v \right)
\end{align*}
\]
with the initial conditions \( u(0) = u_0 \) and \( v(0) = 0 \), where \( \hat{\alpha} = \alpha/r \), \( \hat{\beta} = \beta/r \), \( \hat{Q} = Q/ (Vr) \), and \( u_0 = C_0 V/(S_A K) \).

### 2.1.2 Steady States and Stability

The dimensionless system (2.2) has two possible steady states: the trivial steady state \((u^*, v^*) = (0, 0)\) and
\[
    u^* = \frac{\hat{\beta}}{\hat{\alpha} + \hat{Q}} v^* \quad \text{and} \quad v^* = 1 - \frac{\hat{\beta} \hat{Q}}{\hat{\alpha} + \hat{Q}}.
\]
We denote the trivial steady state as the washout steady state since the pathogens are completely removed from the system as \( t \to \infty \) and the non-zero steady state as the persistence steady state since the pathogens are not removed from the system as \( t \to \infty \). Note that the second steady state, (2.3), is positive if \( \hat{\beta} \leq 1 \) or \( \hat{\beta} > 1 \) and \( \hat{Q} < \hat{\alpha} / \left( \hat{\beta} - 1 \right) = Q_c \).

The linear stability of each steady state is determined by the eigenvalues of the Jacobian matrix
\[
    J = \begin{bmatrix}
        - \left( \hat{\alpha} + \hat{Q} \right) & \hat{\beta} \\
        \hat{\alpha} & 1 - \hat{\beta} - 2v^*
    \end{bmatrix}
\]
The trace and determinant of the Jacobian are
\[
    \tau = \text{tr} (J) = 1 - \left( \hat{\alpha} + \hat{\beta} + \hat{Q} + 2v^* \right) \quad \text{and} \quad \delta = \det (J) = \hat{Q} \left( \hat{\beta} - 1 \right) - \hat{\alpha} + 2 \left( \hat{\alpha} + \hat{Q} \right) v^*.
\]
Note that for any steady state, \( v^* \), we have
\[
    \tau^2 - 4\delta = \left( 1 - \hat{\alpha} - \hat{\beta} - \hat{Q} - 2v^* \right)^2 + 4\hat{\alpha} \hat{\beta} > 0
\]
so the linear system is not a spiral or center and does not oscillate.

When $\hat{\beta} \leq 1$, there are two positive steady states. In this case, the washout steady state is unstable since

$$\delta = -\left(\hat{\alpha} + \hat{Q}(1 - \hat{\beta})\right) < 0$$

and the linearized system is a saddle. For the persistence steady state, we have

$$\tau = -\frac{\hat{\alpha}(1 + \hat{\beta}) + \hat{Q}(1 - \hat{\beta}) + (\hat{\alpha} + \hat{Q})^2}{\hat{\alpha} + \hat{Q}} < 0$$

and

$$\delta = \hat{Q}(1 - \hat{\beta}) + \hat{\alpha} > 0$$

so the linearized system is a stable node.

When $\hat{\beta} > 1$, the system undergoes a transcritical bifurcation as $\hat{Q}$ increases (see Figure 2.2). In this case, we have

$$\tau = -\left(\hat{\alpha} + \hat{Q} \left(1 - \hat{\beta}\right)\right) - 2\nu^* < 0$$

for all $\hat{Q}$ if $\hat{\beta} > 1$ so stability of both steady states is given by $\delta$. For the persistence steady state,

$$\delta = \hat{\alpha} - \hat{Q}(\hat{\beta} - 1)$$

Figure 2.2: Bifurcation diagram showing the exchange of stability between the persistence and washout steady states as the dimensionless flow rate $\hat{Q}$ increases across the critical flow rate $Q_c$. 
which changes sign when \( \hat{Q} = Q_c = \hat{\alpha} / (\hat{\beta} - 1) \) with the steady state stable for \( \hat{Q} < Q_c \) and unstable when \( \hat{Q} > Q_c \). For the washout steady state,

\[
\delta = \hat{Q} (\hat{\beta} - 1) - \hat{\alpha}
\]

which changes from unstable to stable as \( \hat{Q} \) increases through \( \hat{Q} = Q_c \).

The global stability of each steady state is shown in the phase portraits (Figures 2.3a-2.3c). We denote the trivial steady state being stable as “washout” since \( u, v \to 0 \) as

\[
\delta = \hat{Q} (\hat{\beta} - 1) - \hat{\alpha}
\]

which changes from unstable to stable as \( \hat{Q} \) increases through \( \hat{Q} = Q_c \).

Figure 2.3: Phase portraits of the nonlinear system (2.2) for the three cases. The dotted blue line indicates the \( v \)-nullcline, the green dashed line the \( u \)-nullcline, the red arrows that direction field, and the purple lines are individual trajectories. An open circle denotes a linearly unstable steady state and a closed circle a linearly stable steady state. The case where \( \hat{\beta} < 1 \) is shown in 2.3a. In this case, there are two steady states and all trajectories approach the non-trivial steady state. For \( \hat{\beta} > 1 \), there are two steady states when \( \hat{Q} < Q_c \) with all trajectories approaching the non-trivial steady state as shown in 2.3b. When \( \hat{\beta} > 1 \) and \( \hat{Q} > Q_c \) there is only one the trivial steady state and all trajectories approach it.

\[ t \to \infty \]. The pathogens exit the system. When the non-trivial steady is stable, we denote the pathogens “persist” since \( u, v \to u^*, v^* \neq 0 \) as \( t \to \infty \). The pathogens will then exist for all time within the system.

### 2.1.3 Numerical Results

To validate the theoretical result given in Section 2.1.2, we numerical solve the system to determine the concentration of the pathogen for a single connection. For simplicity, we
assume that the quantities \( r, K, V, \) and \( S_A \) are set to be one.

We solve the dimensionless system, (2.1), numerically using MATLAB’s \texttt{ode45} function Shampine and Reichelt [1997]. Initially, we introduce small amount of pathogens into the bulk fluid and nothing within the biofilm. We let \( \hat{\alpha} = 1 \) in each simulation with the initial conditions \( u(0) = 10^{-2} \) and \( v(0) = 0 \). There are three different cases:

(i) For the case where \( 0 < \hat{\beta} < 1 \), we choose \( \hat{\beta} = \frac{3}{10} \) and \( \hat{Q} = \frac{1}{2} \). The numerical solution for the pathogen concentration within the biofilm is shown in Figure 2.4a. We see the pathogen concentration increase monotonically to the steady state value, indicating that the pathogen will persist within the biofilm. This behavior is expected since, when \( \hat{\beta} < 1 \), the growth rate is faster than the detachment rate, allowing the pathogen to colonize the biofilm. Note that this behavior occurs for any \( \hat{Q} \) when \( \hat{\beta} < 1 \).

(ii) For \( \hat{\beta} > 1 \) and \( \hat{Q} < Q_c \), we choose \( \hat{\beta} = 2 \) so that \( Q_c = 1 \). We let \( \hat{Q} = \frac{2}{3} < Q_c \). The numerical solution for the pathogen concentration within the biofilm is shown in Figure 2.4b. The pathogen concentration initially increases due to attachment of planktonic pathogens to the biofilm then it drops to the steady state value where the detachment rate, which is larger than the growth rate, is balanced by growth and attachment. This behavior is observed for any \( \hat{Q} < Q_c \).

(iii) For \( \hat{\beta} > 1 \) and \( \hat{Q} > Q_c \), we choose \( \hat{\beta} = 2 \) which gives \( Q_c = 1 \). We let \( \hat{Q} = \frac{4}{3} > Q_c \). The numerical solution for the pathogen concentration within the biofilm is shown in Figure 2.4c. The pathogen concentration initially increases due to attachment of planktonic pathogens to the biofilm then it drops to zero, indicating the pathogens washout of the system.

### 2.2 Constant Flow Network Model

Using the results from the analysis of a single connection, we now examine a model of a water distribution network where the flow is constant for all time.
Figure 2.4: Numerical results for the three different cases discussed in Section 2.1.2. The pathogen concentration, $\hat{v}$, is shown in (a) for the case $\hat{\beta} < 1$. Here, the pathogen increases until it reaches the non-zero steady state, indicating the pathogen will persist for all $t > 0$. For $\hat{\beta} > 1$ and $\hat{Q} < Q_c$, shown in (b), the pathogen concentration approaches a non-negative steady state, again indicating the pathogen persists within the pipe for all $t > 0$. When $\hat{\beta} > 1$ and $\hat{Q} > Q_c$, the trivial steady state is stable and the pathogen washes out of the system, as shown in (c).
2.2.1 Mathematical Model

In our model, a water distribution network is approximated by a series of connections, which can consist of one or more individual pipes, that are joined by a set of junctions, as shown in Figure 2.5.

![Figure 2.5: A model water distribution network. The circles are junctions and the lines are connections with the arrow indicating the direction of the flow.](image)

**Connectivity**

The network is defined mathematically by the junction-to-junction connectivity matrix, $C \in \mathbb{Z}^{M \times M}$, where $M$ is the number of junctions. The value of each element $C_{i,j}$ can be either zero or one with $C_{i,j} = 1$ indicating that there is a connection that links the $i^{th}$ junction to the $j^{th}$ junction. When $C_{i,j} = 0$, there is no connection between the $i^{th}$ and $j^{th}$ connections. Since a water distribution network is bi-directional, the connectivity matrix is symmetric, $C_{i,j} = C_{j,i}$, and has a zero diagonal since a junction cannot be connected directly with itself. The total number of connections can be obtained from $C$ using the relation

$$N = \sum_{j=i+1}^{M} \sum_{i=1}^{M-1} C_{i,j}$$

(2.6)

For our purposes, an alternative connectivity matrix that relates connections to junctions is useful. We define the connection-to-junction connectivity matrix, $H \in \mathbb{Z}^{N \times M}$, where $N$
is the number of connections, that takes the values 0, ±1. Here, \(H_{i,j} = \pm 1\) indicates that the \(i^{th}\) connection enters (positive)/exits (negative) the \(j^{th}\) junction and \(H_{i,j} = 0\) indicates that the \(i^{th}\) connection does not connect to the \(j^{th}\) junction. Since each connection joins two and only two junctions, there is exactly two non-zero elements in each row of \(H\) and the sum of each row is zero. Note that this connectivity matrix provides directional information about the flow through the connections. In the absence of actual data on the flow direction, all connections exit the lower indexed junction and enters the higher indexed junction. The junction-to-junction and connection-to-junction matrices for the network in Figure 2.5 are

\[
\begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\end{bmatrix}
\quad \text{and} \quad
\begin{bmatrix}
-1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 1 & 1 \\
\end{bmatrix}
\]

Once we know the flow rates within each connection, which we denote by the diagonal flow matrix \(Q \in \mathbb{R}^{N \times N}\) and is described in Section 2.2.2, we can compute a directed connection-to-junction connectivity matrix, \(H_Q = \text{sign}(Q)H\), where if \(B = \text{sign}(A)\) then \(B_{i,j} = A_{i,j}/|A_{i,j}|\) if \(A_{i,j} \neq 0\) and \(B_{i,j} = 0\) if \(A_{i,j} = 0\).

A modification of directed connection-to-junction connectivity matrix that is needed for the network model is the interior restricted version of \(H_Q\), which we denote by \(H_I \in \mathbb{Z}^{N \times M_I}\), where \(M_I\) is the number of interior junctions, respectively with an interior junction defined as a junction with only one connection. The matrix \(H_I\) is computed by removing the
corresponding column for each junction that has only one connection.

The directed connection-to-junction matrix, $H_Q$, and the interior, directed connection-to-junction matrix, $H_I$, for the network in Figure 2.5 are

\[
H_q = \begin{bmatrix}
-1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 1
\end{bmatrix}
\quad \text{and} \quad
H_i = \begin{bmatrix}
1 & 0 & 0 & 0 \\
-1 & 1 & 0 & 0 \\
-1 & 0 & 1 & 0 \\
-1 & 0 & 0 & 1 \\
0 & -1 & 0 & 0 \\
0 & -1 & 0 & 1 \\
0 & 0 & -1 & 1 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1
\end{bmatrix}
\tag{2.8}
\]

Another useful connectivity matrix is the parent connectivity matrix, $H_p \in \mathbb{Z}^{N \times N}$, where $H_{p_{i,j}} = 0,1$ with $H_{p_{i,j}} = 1$ indicating that the $i^{th}$ connection is the parent of the $j^{th}$ connection. The parent connectivity matrix, $H_p$, for the network in Figure 2.5 is

\[
H_p = \begin{bmatrix}
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\tag{2.9}
\]
Descendants and Ancestors

For a given network with parent connectivity matrix $H_p$, we define the set of children of the $i^{th}$ connection as

$$C(i) = \{ j \mid H_{pi,j} = 1 \}$$

and the set of parents of the $j^{th}$ connection as

$$P(j) = \{ i \mid j \in C(i) \}.$$  

The set of all descendants of the $i^{th}$ connection is

$$D(i) = \{ j \mid j \in C(i) \lor j \in D(k) \text{ for any } k \in C(i) \}.$$  

and the set of ancestors of the $j^{th}$ connection is defined as

$$A(j) = \{ i \mid j \in D(i) \}.$$  

These sets satisfy the follow relations:

$$C(i) \subseteq D(i), \quad P(i) \subseteq A(i), \quad D(i) \cap A(i) = \emptyset.$$  

The last relation is demonstrated in Section 2.2.2.

Network Model

We can now derive a network model governing the concentration of pathogens within the connections and the interior junctions of a water distribution network. Similar to the single connection model derived in Section 2.1, we assume that each connection is divided into two compartment, the bulk fluid and the biofilm. For each junction, we assume that the biofilm compartment is negligible and only assume that there is a bulk fluid compartment. For each
compartment, we assume the concentration of pathogens change due to flow in from the parent junction, flow out to the child junction, and attachment/detachment from the biofilm. We assume no growth in the bulk fluid (see Figure 2.6). For the biofilm compartment in the connection, we assume attachment/detachment from the bulk fluid and logistic growth. For the junctions, we assume the pathogen concentration changes due to only flow in from parent connections, flow out to child connections, and flows out of the system due to demands at the junctions. The demand at a junction is the removal of water from the system at the point of consumption such as a water tap. The governing equations are

$$\begin{align*}
\begin{cases}
u' &= Zc - Qu + f(u, v) \\
c' &= Wu - (T + D)c \\
v' &= g(u, v)
\end{cases}
\end{align*}$$

(2.15)

where $u$ is the vector of the pathogen concentration within the bulk fluid for each connection, $v$ is the vector of the pathogen concentration within the biofilm for each connection, and $c$ is the vector of the pathogen concentration within the fluid in the junctions. The vector functions $f(u, v)$ and $g(u, v)$ describe the attachment/detachment/growth dynamics of the pathogens. They are diagonal operators and are defined as

$$f(u, v) = \Gamma^{-1}Bv - Au$$

(2.16)

and

$$g(u, v) = \Gamma Au + (R - B)v + Y^{-1}R(v \odot v)$$

(2.17)
where the $\mathbf{v} \odot \mathbf{v}$ indicates component-wise multiplication of the vector. Each matrix is a diagonal matrix of connection parameters where, for the $i^{th}$ connection, $A_{i,i} = \alpha_i$ is the pathogen attachment rate from the bulk fluid to the biofilm, $B_{i,i} = \beta_i$ is the pathogen detachment rate from the biofilm to the bulk fluid, $\Gamma_{i,i} = \gamma_i = V_i/S_{A_i}$ is the bulk fluid volume to connection surface area ratio, $Y_{i,i} = K_i$ is the carrying capacity of the pathogens within the biofilm, and $R_{i,i} = r_i$ is the pathogen growth rate in the biofilm. Each are considered to be constant and positive so each diagonal matrix is invertible. The matrices $\mathbf{Z}$ and $\mathbf{W}$ indicate the flow into the connections and junctions, respectively. The matrix $\mathbf{Q}$ is a diagonal matrix that denotes the flow rate through the connections into the child junctions with $Q_{i,i} > 0$. The matrices $\mathbf{T}$ and $\mathbf{D}$ denote the flows out of the junction into the child connections and demands at each junction, respectively. The inflow matrices are defined by

$$
\mathbf{Z} = \mathbf{Q} (\mathbf{H}_{I} < 0) \in \mathbb{R}^{N \times M_I} \quad \text{and} \quad \mathbf{W} = (\mathbf{H}_{I}^T > 0) \mathbf{Q} \in \mathbb{R}^{M_I \times N} \quad (2.18)
$$

and the outflow matrices are $\mathbf{T} = (\mathbf{H}_{I}^T < 0) \mathbf{Q} (\mathbf{H}_{I} < 0) \in \mathbb{R}^{M_I \times M_I}$ and $\mathbf{D}$ is a diagonal matrix where $D_{i,j} = D_i$ is the demand at the $i^{th}$ junction.

We assume that under realistic flow conditions the pathogen concentration within the junctions respond quickly to changes in concentration so we use a quasi-steady state assumption that $\mathbf{c}' \approx 0$ to obtain

$$
\mathbf{c} \approx (\mathbf{T} + \mathbf{D})^{-1} \mathbf{W} \mathbf{u}. \quad (2.19)
$$

Substituting (2.19) into (2.15), we obtain the quasi-steady state network model for the pathogen concentrations within the connections:

$$
\begin{aligned}
\mathbf{u}' &= (\mathbf{K} - \mathbf{Q}) \mathbf{u} + \mathbf{f} (\mathbf{u}, \mathbf{v}) \\
\mathbf{v}' &= \mathbf{g} (\mathbf{u}, \mathbf{v})
\end{aligned} \quad (2.20)
$$

where the matrix $\mathbf{K} = \mathbf{Z} (\mathbf{T} + \mathbf{D})^{-1} \mathbf{W} \in \mathbb{R}^{N \times N}$ denotes the connection-to-connection inflows. It is a singular matrix with a zero diagonal due to the inability of connections to be
self-parents. Note that $K$ can be reordered, in general, to be become a lower triangular with zero diagonals by reindexing the connections using parent/child relationships between the connections since a connection cannot be its own ancestor (see Section 2.2.2).

### 2.2.2 Determining the Flow Rate

To compute the dynamics of pathogens within a water distribution network, we need a method for computing the fluid flow profiles within the network. Since our focus is on the dynamics of the pathogens and not on the actual flow, we will assume, for simplicity, the flow is constant in time throughout the network and each connection can be modeled as a linear resistance pipe with pressures defined at each junction. The flow rate, $Q$, through a linear resistance pipe is

$$Q = \frac{P_a - P_b}{R} \quad (2.21)$$

where $P_a$ and $P_b$ are the pressures at the ends of the pipe and $R$ is the hydraulic resistance of the pipe. For a network, the flow rates are computed using the connection-to-junction connectivity matrix, $H$, in the relation

$$r_f \odot q = Hp \quad (2.22)$$

where $q \in \mathbb{R}^N$ is the vector of flow rates for each connection, $r_f \in \mathbb{R}^N$ is the vector of hydraulic resistances for each connection, $p \in \mathbb{R}^M$ is the vector of pressures at the junctions. For each interior junction, we prescribe a demand, $D_i$, and for each non-interior junction, we prescribe a pressure, $P_i$. From (2.22) and the pressure conditions at the non-interior junctions, we have $N + (M - M_I)$ equations for $N + M$ unknowns. The remaining $M_I$ equations come from conservation of mass at each interior junction. At each interior junction,
we have the relationship

\[
\sum_{k=1}^{N_{\text{inflow}}} q_j - \sum_{k=1}^{N_{\text{outflow}}} q_k = D
\]

(2.23)

The equations for the entire network, using the connection-to-junction connectivity matrix \( H \) is

\[
\Delta H^T = d
\]

(2.24)

where \( \Delta \in \mathbb{R}^{M \times M} \) is a diagonal matrix with \( \Delta_{i,i} = 1 \) is the \( i \)th junction is an interior junction and \( \Delta_{i,i} = 0 \) otherwise. The vector \( d \in \mathbb{R}^M \) is the vector of the demand at each junction. For non-interior junctions, the values of \( d \) are undefined and not used. Combining (2.22), (2.24), and the conditions on the pressure at non-interior junctions, the linear system governing the flow rates through the connections and the pressures at the junctions is

\[
\begin{bmatrix}
I - \Delta & \Delta H^T \\
H & R_f
\end{bmatrix}
\begin{bmatrix}
p \\
q
\end{bmatrix}
= \begin{bmatrix}
b \\
0
\end{bmatrix},
\]

(2.25)

where \( p \) is the vector of pressures at each junction, \( I \) is the identity matrix, \( R_f = \text{diag} (r_f) \) is the diagonal matrix of hydraulic resistances for each connection, and \( b = \Delta d_0 + (I - \Delta) p_0 \). Here, \( d_0_i = D_i \) for each interior junction and \( p_0_i = P_i \) for each non-interior junction. Solving this system provides the flow rate \( q \), which we use to define the flow rate matrix \( Q = \text{diag} (q) \).

Flow Example

To illustrate the computation of the flow through a network, we given an example of a network consisting of two pipes, which is shown in Figure 2.7. We assume that we are given the pressure at the two end junctions, \( P_{\text{in}} \) and \( P_{\text{out}} \), as well as the demand given at the interior
junction, $D_1$. Each connection has a hydraulic resistance, $R_j$, and we seek to compute the flow rates through each connection, $Q_j$, and the pressure at the interior junction, $P_1$.

![Figure 2.7: Example network consisting of two connections and three junctions.](image)

Using the linear resistance model, (2.21), we have the relations

$$Q_1 = \frac{P_{\text{in}} - P_1}{R_1} \quad \text{and} \quad Q_2 = \frac{P_1 - P_{\text{out}}}{R_2}.$$  \hfill (2.26)

We additionally have a third relation due to conservation of mass at the interior junction:

$$Q_1 = Q_2 + D_1.$$  \hfill (2.27)

Combining (2.26) and (2.27), we obtain the flow rates

$$Q_1 = \frac{P_{\text{in}} - P_{\text{out}} - D_1R_1}{R_1 + R_2} \quad \text{and} \quad Q_2 = \frac{P_{\text{in}} - P_{\text{out}} - D_1R_2}{R_1 + R_2}$$  \hfill (2.28)

and the pressure at the interior junction

$$P_1 = \left( \frac{P_{\text{in}} - P_{\text{out}} + D_1R_2}{R_1 + R_2} \right) R_2 + P_{\text{out}}.$$  \hfill (2.29)

**Flow Loops**

Note that it is not possible in our model for a connection to be its own ancestor. To demonstrate this, we start with a network with a loop, which is shown in Figure 2.8. The subscripts for $Q$ and $P$ indicate the index of the connection/junction number each flow/pressure correspond to, respectively. Let us denote each connection by $C#$ and each junction by $J#$ and assume that $C2$ is the parent of $C4$, which is then a parent of $C3$, as denoted by the arrows

20
in Figure 2.8. This leads to C3 become the parent of C2, which completes the loop and C2 is its own ancestor. Note that for the flow to have the directions indicated by the arrow, \( Q_1, Q_2, Q_4, Q_5 > 0 \) and \( Q_3 < 0 \). For this network, the governing equations for the flow rate are

\[
Q_1 = \frac{P_1 - P_2}{R_1}, \quad Q_2 = \frac{P_2 - P_3}{R_2}, \quad Q_3 = \frac{P_2 - P_4}{R_3}, \quad Q_4 = \frac{P_3 - P_4}{R_4}, \quad \text{and} \quad Q_5 = \frac{P_4 - P_5}{R_5}.
\]

(2.30)

with the equations for conservation of mass

\[
Q_1 = Q_2 + Q_3 + D_2, \quad Q_2 = Q_3 + D_3, \quad \text{and} \quad Q_3 + Q_4 = Q_5 + D_4
\]

(2.31)

Since \( Q_1, Q_2, Q_4, Q_5 > 0 \), we require \( P_1 > P_2 > P_3 > P_4 > P_5 \). To have the loop, \( Q_3 < 0 \), which requires \( P_4 > P_2 \) but this creates a contradiction. Hence, we are not able to have a flow loop in this model.

### 2.2.3 Theorems

The general equations governing the concentration of pathogens within the network is given by (2.20). Since we are interested in determining the stability of the trivial steady state, which corresponds to the pathogen washing out of the entire network, we linearize the non-linear system (2.20) about \( u = v = 0 \) to obtain the linear system \( x' = Jx \) where \( x = [u, v]^T \)
and

\[
J = \begin{bmatrix}
K - Q - A & \Gamma^{-1}B \\
\Gamma A & R - B 
\end{bmatrix}
\]  

(2.32)

is the Jacobian matrix for the non-linear system linearized around the trivial steady state. The stability of the trivial steady state depends on the real part of the eigenvalues of the Jacobian matrix \( J \).

**Remark 2.2.1.** We assume here that \( K \) is a lower triangular matrix with zero diagonal without the loss of generality. If it is not, we can re-index the connections within the network using the flow information. Since \( K_{i,j} \neq 0 \) denote that the \( j \)th connection is a parent of the \( i \)th connection, we can number the connections without parents first, and then numbering their child connections next, and so forth through each generation, so that the matrix \( K \) will be lower triangular and, since no connection is its own parent, it will have zeros along the diagonal.

Our goal is to determine if the pathogens persist within the network, which corresponds to the trivial steady state being linearly unstable. To do this, we start by proving that the sign of the eigenvalues for the linearized network system is independent of the matrix \( K \), which contains the network connectivity information.

**Theorem 2.2.2.** The eigenvalues of \( J \) given in (2.32) are independent of \( K \).

**Proof.** The eigenvalues \( \lambda \) of \( J \) satisfy the matrix equation:

\[
(J - \lambda I) x = 0
\]  

(2.33)

where \( x = [u, v]^T \). This is equivalent to

\[
(K - Q - A - \lambda I) u + \Gamma^{-1}B v = 0
\]  

(2.34)

\[
\Gamma A u + (R - B - \lambda I) v = 0
\]  

(2.35)
so, solving for $u$, (2.35) yields

$$u = A^{-1} \Gamma^{-1} (R - B - \lambda I) v$$

(2.36)

since $A$ and $\Gamma$ are invertible.

We now substitute the value of $u$ from (2.36) into (2.34) to get

$$(K - Q - A - \lambda I) A^{-1} \Gamma^{-1} (R - B - \lambda I) v - \Gamma^{-1} Bv = 0$$

(2.37)

Since all the matrices except for $K$ are diagonal matrices, we can simplify (2.37) to obtain

$$[(K - Q - A - \lambda I) (R - B - \lambda I) - AB] v = 0.$$

(2.38)

The eigenvalue problem is now of the form $Lv = 0$ where

$$L = (K - Q - A - \lambda I) (R - B - \lambda I) - AB$$

(2.39)

is a lower triangular matrix since $K$ is lower triangular and the other matrices are diagonal. Thus, for the null space not to be empty, $L$ must be singular so $\det (L) = 0$. This implies that

$$\prod_{i=1}^{N} L_{i,i} = 0$$

(2.40)

so

$$\prod_{i=1}^{N} (q_i + \alpha_i + \lambda) (r_i - \beta_i - \lambda) + \alpha_i \beta_i = 0$$

(2.41)

since $K_{i,i} = 0$, $R_{i,i} = r_i$, $B_{i,i} = \beta_i$, $Q_{i,i} = q_i$, and $A_{i,i} = \alpha_i$. The yields the characteristic equations

$$(q_i + \alpha_i + \lambda) (r_i - \beta_i - \lambda) + \alpha_i \beta_i = 0$$

(2.42)

for $i = 1, 2, \ldots, N$ and if $\lambda$ satisfies (2.42) for any $i$, then $\lambda \in \sigma (A)$. Since the characteristic
Since the communication between connections in the network occur through the matrix $K$, Theorem 2.2.2 proves that the linear stability of the trivial steady state for the network is the same as the linear stability of the trivial steady state when each connection is viewed in isolation. This allows us to examine the connections independently and we are not required to examine the linear stability for the entire network. If one connection in isolation has a linearly unstable trivial steady state, then the network will have an linearly unstable trivial steady state and if all connections have a linearly stable trivial steady state in isolation, then the trivial steady state for the entire network will be stable.

We now prove the eigenvectors of the Jacobian matrix $J$ depend on the matrix $K$ and the non-zero elements of the eigenvectors indicate the descendants of individual connections.

**Theorem 2.2.3.** The eigenvectors of $J$ given by (2.32) depend on the matrix $K$ and the non-zero elements of the $i^{th}$ $v$-eigenvector correspond to all the descendants of the $i^{th}$ connection.

*Proof.* To find the eigenvectors of $J$, we solve $Lv = 0$. Notice that $L$ can be written as $L = K + D$ where $K$ is lower triangular with zero diagonal matrix and $D$ is a diagonal matrix with entries

$$D_{i,i} = -\left[ (\hat{q}_i + \hat{\alpha}_i + \lambda) \left( \hat{r}_i - \hat{\beta}_i - \lambda \right) + \hat{\beta}_i \hat{\alpha}_i \right] \quad (2.43)$$

so that

$$L = \begin{bmatrix}
D_{1,1} & 0 & \cdots & \cdots & \cdots & 0 \\
K_{2,1} & D_{2,2} & 0 & \cdots & \cdots & 0 \\
K_{3,1} & K_{3,2} & D_{3,3} & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
K_{N,1} & \ddots & \ddots & \ddots & D_{N-1,N-1} & 0 \\
K_{N,1} & \ddots & \ddots & \ddots & K_{N,N-1} & D_{N,N}
\end{bmatrix} \quad (2.44)$$
There will be three different cases based on the algebraic and geometric multiplicity of the eigenvalues \( \lambda_i \). We will assume that the eigenvalues all have multiplicity of one so \( \lambda_i \neq \lambda_j \) if \( i \neq j \) for all \( i, j = 1, 2, \ldots, N \). For higher multiplicities, see Appendix E.

We start by choosing \( \lambda_1 \) such that \( D_{1,1} = 0 \). This implies that \( v_1 = c \) is an arbitrary non-zero constant. Looking at the second row, we have

\[
K_{2,1}v_1 + D_{2,2}v_2 = 0 \tag{2.45}
\]

so \( v_2 = -(K_{2,1}/D_{2,2})v_1 = -(K_{2,1}/D_{2,2})c \) since \( D_{2,2} \neq 0 \) under the assumption \( \lambda_2 \neq \lambda_1 \).

Note that \( v_2 = 0 \) if connection 1 is not a parent of connection 2, \( 1 \notin \mathcal{P}(2) \).

For the third row, we have

\[
K_{3,1}v_1 + K_{2,1}v_2 + D_{3,3}v_3 = 0 \tag{2.46}
\]

so that

\[
v_3 = -\frac{1}{D_{3,3}}(K_{3,1}v_1 + K_{3,2}v_2) = -\frac{1}{D_{3,3}} \left( K_{3,1} - \frac{K_{3,2}K_{2,1}}{D_{2,2}} \right) c. \tag{2.47}
\]

Note that \( v_3 \neq 0 \) if \( 1 \in \mathcal{P}(3) \) \( (K_{3,1} \neq 0) \) or \( 2 \in \mathcal{P}(3) \) \( (K_{3,2} \neq 0) \) and \( 1 \in \mathcal{P}(2) \) \( (v_2 \neq 0) \). These conditions implies that connection 1 is an ancestor of connection 3 (we have \( 1 \in \mathcal{A}(3) \)). Note that \( v_3 = 0 \) can only occur if \( 1 \notin \mathcal{A}(3) \). To prove this, assume that \( 1 \in \mathcal{A}(3) \) so \( K_{3,1} \neq 0 \) and/or \( K_{3,2} \neq 0 \) and \( K_{2,1} \neq 0 \). If any of these conditions are not true, then \( v_3 \neq 0 \) so we assume that \( K_{3,1}, K_{3,2}, K_{2,1} \neq 0 \). These conditions imply that \( 1 \in \mathcal{P}(2), 1 \in \mathcal{P}(3) \), and \( 2 \in \mathcal{P}(3) \) but this is not possible since it would make connection 1 both the parent and grandparent of connection 3. Hence, \( v_3 \neq 0 \) iff \( 1 \in \mathcal{A}(3) \).

For any row \( i > 3 \), we have

\[
\sum_{j=1}^{i-1} K_{i,j}v_j + D_{i,i}v_i = 0. \tag{2.48}
\]
Solving (2.48) for \( v_i \) yields the recursive relation

\[
v_i = -\frac{1}{D_{i,i}} \sum_{j=1}^{i-1} K_{i,j} v_j = ac
\]

(2.49)

where \( a \) is a constant that depends on \( K \) and \( D \). If the \( j \)-th connection is a parent of the \( i \)-th connection, \( j \in P(i) \), then \( K_{i,j} \neq 0 \). Likewise, if \( j \not\in P(i) \) then \( K_{i,j} = 0 \). Hence, (2.49) reduces to

\[
v_i = -\frac{1}{D_{i,i}} \sum_{j=1 \atop j \in P(i)}^{i-1} K_{i,j} v_j
\]

(2.50)

and, by induction, \( v_i = 0 \) if \( 1 \not\in A(i) \). As a consequence, the non-zero entries of \( \mathbf{v} \) correspond to the descendants of the first connection.

For the \( m \)-th eigenvalue, \( \lambda_m \), with \( m > 1 \), we assume the ordering of the eigenvalues are such that \( D_{m,m} = 0 \) when \( \lambda = \lambda_m \) for \( m = 1, 2, \ldots, N \). Note that we still are assuming the algebraic multiplicity of the eigenvalues are one so \( D_{i,i} = 0 \) for \( i \neq m \). From the first row, we have \( D_{1,1} v_1 = 0 \) so \( v_1 = 0 \) since \( D_{1,1} \neq 0 \). Likewise, we have

\[
K_{2,1} v_1 + D_{2,2} v_2 = 0
\]

(2.51)

from the second row so \( v_2 = 0 \) since \( v_1 = 0 \) and \( D_{2,2} \neq 0 \). We now assume \( v_j = 0 \) for \( j = 1, \ldots, i - 1 \). We have

\[
\sum_{j=1}^{i-1} K_{i,j} v_j + D_{i,i} v_i = 0
\]

(2.52)

for the \( i \)-th row with \( i < m \). Since \( v_j = 0 \) for \( j = 1, \ldots, i - 1 \) and \( D_{i,i} \neq 0 \), we have, by induction, \( v_i = 0 \) for \( i < m \).

We now examine the \( m \)-th row of \( \mathbf{L} \):

\[
\sum_{j=1}^{m-1} K_{m,j} v_j + D_{m,m} v_m = 0.
\]

(2.53)
Since $v_j = 0$ for $j = 1, \ldots, m - 1$ and $D_{m,m} = 0$, we have $v_m = c$, where $c$ is an arbitrary non-zero constant. We proceed as above when $\lambda = \lambda_1$ to obtain that the non-zero entries of $m^{th}$ $v$ correspond to the ancestors of the $m^{th}$ connection.

\[\Box\]

2.2.4 Efficient Algorithm for Determining the Pathogen Persistence within the Network

To determine the stability of the trivial steady state for the entire network, one could linearized and compute the $2N$ eigenvalues of the Jacobian matrix. If any eigenvalue has a positive real part, the pathogens could persist within the network, based upon the initial pathogen distribution within the network. For large networks, the eigenvalues can be computed using QR factorization, which is $O(8N^3)$ for this problem. For efficiency, we can use the results of Theorems 2.2.2 and 2.2.3 to reduce the effort in determining the stability to $O(N)$.

We start by noting that the eigenvalues for the entire linearized network model and the eigenvalues of the linearized model of the individual connections in isolation are the same from Theorem 2.2.2. Hence, the trivial steady state for the entire network model is unstable if the trivial steady state for any individual connection is unstable—if the pathogens would persist for an individual connection, it will persist within the network.

We can additionally determine the distribution of the pathogens within the entire network using the eigenvalues computed for each individual connection and the parent connectivity matrix, $H_p$. From Theorem 2.2.3, if the pathogens within a connection persist, the eigenvector will contain a non-negative entry for each child of that connection. This implies that the persistence of the pathogen within the parent causes the pathogens to persist within each child. When the pathogens would washout of a child connection in isolation but the child has a ancestor where the pathogens persist, the pathogen concentration within the child approach a non-zero steady state. We say that the child is induced in this situation.
To summarize, each connection in isolation can be labelled as persistent or washout based on the bifurcation conditions for the individual connection (see Section 2.1.2). If any connection is classified as persistent, then the pathogen can persist within the network and it all connections are classified as washout, then the pathogens washout of the network. When a connection is persistent, all it descendants are either persistent (if they are persistent in isolation) or they are induced (if they are washout in isolation).

We can use the parent connectivity matrix, \( H_p \), iteratively to determine which connections within the network are induced. The algorithm is

1. Compute the stability of the trivial steady state for each connection in isolation. Classify each connection as “persistent” or “washout” using the bifurcation conditions.

2. Define the elements of the vector \( s^{(0)} \) as \( s^{(0)}_i = 1 \) if the \( i \)th connection is “persistent” and \( s^{(0)}_i = 0 \) if the \( i \)th connection is “washout”.

3. Define the iteration matrix \( Z = I + H_p^T \).

4. Set \( s = s^0 \)

5. Repeat

   (a) \( s = Zs \)

   (b) If \( s^0_i = 1 \), then set \( s_i = 1 \) for all \( i = 1, \ldots, N \).

   (c) If \( s^0_i = 0 \) and \( s_i > 0 \), then set \( s_i = 2 \) for all \( i = 1, \ldots, N \).

Until \( ||s - s^0|| = 0 \)

After this algorithm is completed, the vector \( s \) indicates the connection is “persistent” when \( s_i = 1 \), “induced” when \( s_i = 2 \), and “washout” when \( s_i = 0 \).

Note that this approach indicates which connections are the biggest concerns since a connection that is “persistent” can cause the pathogens to appear in all its descendants and if the pathogens were removed from the “persistent” connections by some external means,
then all the descendants would have the pathogen washout. This could aid in the efficient removal of pathogens form the network.

2.2.5 Theoretical and Numerical Results

To validate the theorems and the algorithm given in Section 2.2.4, we examine some simple networks.

Small Network Example

In this example, we study a small network consists of two pipes in serial, as shown in Figure 2.7. The dimensionless model for this system is given by:

\[
\begin{align*}
    u'_1 &= \hat{\beta}v_1 - (\hat{q}_1 + \hat{\alpha})u_1 \\
    u'_2 &= \hat{q}_2 (u_1 - u_2) + \hat{\beta}v_2 - \hat{\alpha}u_2 \\
    v'_i &= \hat{\alpha}u_i - \hat{\beta}v_i + v_i (1 - v_i) \text{ for } i = 1, 2.
\end{align*}
\]  

(2.54)

with \( \hat{q}_2 = \hat{q}_1 - D \). Note we have assumed that \( \alpha, \beta, r, \) and \( K \) are the same for each connection. There are three possible steady states for this system:

(i) \( v^*_1 = v^*_2 = 0 \). This is the trivial steady state and corresponds to the pathogens being washed out of the entire network.

(ii) \( v^*_1 = 0 \) and \( v^*_2 \neq 0 \). This corresponds to the pathogens persisting in the second connection but being washed out of the first connection.

(iii) If \( v^*_1 \neq 0 \) and \( v^*_2 \neq 0 \). This corresponds to the either the pathogens persisting in both connections or the pathogen persisting in the first connection and the second connection being induced by the first.

Note that there is no steady state where \( v^*_1 \neq 0 \) and \( v^*_2 = 0 \). Also, we have

\[
u_i = \frac{1}{\hat{\alpha}} \left( v_i + \hat{\beta} - 1 \right) v_i
\]  

(2.55)
from the last equation in 2.54 and if \( v_i = 0 \), then \( u_i = 0 \) and if \( v_i \neq 0 \), then \( u_i \neq 0 \) in general.

The Jacobian matrix for the entire system is

\[
J = \begin{bmatrix}
-(\hat{q}_1 + \hat{\alpha}) & \hat{\beta} & 0 & 0 \\
\hat{\alpha} & 1 - \hat{\beta} - 2v_1^* & 0 & 0 \\
\hat{q}_2 & 0 & -\hat{q}_2 + \hat{\alpha} & \hat{\beta} \\
0 & 0 & \hat{\alpha} & 1 - \hat{\beta} - 2v_2^*
\end{bmatrix}
\]  

(2.56)

The Jacobian (2.56) is a block lower triangular matrix so the characteristic equation for the eigenvalue problem is

\[
|J - \lambda I| = 0
\]

is

\[
[(\hat{q}_1 + \hat{\alpha} + \lambda) (1 - \hat{\beta} - 2v_1^* - \lambda) + \hat{\alpha}\hat{\beta}] [(\hat{q}_2 + \hat{\alpha} + \lambda) (1 - \hat{\beta} - 2v_2^* - \lambda) + \hat{\alpha}\hat{\beta}] = 0
\]  

(2.57)

since

\[
\det \left( \begin{bmatrix} A & 0 \\ C & B \end{bmatrix} \right) = \det (A) \det (B)
\]  

(2.58)

Note that each term in brackets in the characteristic equation, (2.56), is the characteristic equation for each pipe in isolation, which is expected from Theorem 2.2.2 for the trivial steady state. Note that \( \lambda \in \mathbb{R} \).

The first case where \( v_1^* = v_2^* = 0 \) is linearly stable when \( \hat{\beta} > 1 \) and \( q_1, q_2 < Q_c = \hat{\alpha}/(\hat{\beta} - 1) \). These are the conditions for washout in both connections in isolation.

The second case where \( v_1^* = 0 \) and \( v_2^* \neq 0 \) is linearly stable when \( \hat{\beta} > 1 \) with \( q_1 < Q_c \) and \( q_2 > Q_c \). This corresponds to the conditions for the first connection to washout and the second connection to persist.

The third steady state is stable when \( \hat{\beta} \leq 1 \) or \( \hat{\beta} > 1 \) with \( q_1 < Q_c \) regardless of \( q_2 \). This corresponds to persistence in both connections in isolation (\( \hat{\beta} \leq 1 \) or \( \hat{\beta} > 1 \) and \( q_2 < Q_c \)) or the first connection is persistent and the second connection is induced.
The stability of the system and the isolated connections is summarized in Table 2.1 and confirms the results from Theorems 2.2.2 and 2.2.3.

<table>
<thead>
<tr>
<th>In Isolation</th>
<th>In Network</th>
<th># of Eigenvalues</th>
<th>Form of the Eigenvectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pipe #1</td>
<td>Pipe #1</td>
<td>where Re (µ) &gt; 0</td>
<td>when Re (µ) &gt; 0</td>
</tr>
<tr>
<td>Washout</td>
<td>Washout</td>
<td>0</td>
<td>[0, 0, c, d]^T</td>
</tr>
<tr>
<td>Washout</td>
<td>Persist</td>
<td>1</td>
<td>[a, b, c, d]^T</td>
</tr>
<tr>
<td>Persist</td>
<td>Washout</td>
<td>1</td>
<td>2, [a_1, b_1, c_1, d_1]^T, [0, 0, c_2, d_2]^T</td>
</tr>
<tr>
<td>Persist</td>
<td>Persist</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Table of stable steady states in isolation and within the two pipe network. If the pathogen persists either pipe in isolation, then the pathogen will persist within the network.

To validate the analysis, we run numerical simulations for each of the four different cases. In each of the simulations, the initial conditions are \( u_1(0) = u_2(0) = 10^{-3} \) and \( v_1(0) = v_2(0) = 0 \) with the attachment rates \( \alpha_1 = \alpha_2 = 9 \), detachment rates \( \beta_1 = \beta_2 = 3 \), and growth rates \( r_1 = r_2 = 1 \). The resistance of each pipe to the flow is \( R_1 = R_2 = 1 \).

Figure 2.9 shows the numerical results when both connections are predicted to have the pathogens “washout” in isolation, which is the first case. We can when they are connected together in the network, both solutions go to zero confirming the the prediction of the analysis. Note that there is a larger spike in the second (downstream) connection due to the pathogens originally within the first (upstream) connection being transported by the flow into the second connection. Figure 2.10 shows the numerical solution for the pathogen concentration within the biofilm for the second case where the first connection is predicted to have the pathogens “washout” and, in the second connection, the pathogens “persist” in isolation. The results match the analysis when both connections are in the network, with the pathogen concentration in the biofilm going to zero in the first connection and going to a non-zero steady state in the second connection. Figure 2.11 shows the numerical solution for the pathogen concentration within the biofilm for the fourth case where the pathogen is predicted to “persist” within both connection when they are in isolation. When they are connected in
Figure 2.9: Numerical solution for pathogen concentration within the biofilm for the two connection network example. Both connections have the pathogens washout in isolation (a) and when part of the network (b), confirming the results of the linear stability analysis. Here, we set $Q_1 = 5 > Q_c$ and $D_1 = 0$, yielding $Q_2 = Q_1 - D_1 = 5 > Q_c$.

Figure 2.10: Numerical solution for pathogen concentration within the biofilm for the two connection network example. The parent connection has the pathogens washout and the child connection has the pathogens persist in isolation (a) and when part of the network (b), confirming the results of the linear stability analysis. Here, we set $Q_1 = 5 > Q_c$ and $D_1 = 3$ so that $Q_2 = Q_1 - D_1 = 2 < Q_c$. 
serial in the network, the analysis predicts the pathogens will “persist” and this is validated by the numerical solution. Note the slight increase in the steady state pathogen concentration in the second connection due to pathogens from the first connection flowing into the second continuously. Figure 2.12 shows the third case, which is the most important result for studying pathogen dynamics within the network. Here, pathogens in the first connection are predicted to “persist” while the pathogens in the second connection are predicted to “washout” when the connection are in isolation. The numerical simulation confirm this when they are not part of the network (Figure 2.12a). When they are connected in serial within the network, the first (upstream) connection “induces” a non-zero pathogen concentration in the second (downstream) connection as $t \to \infty$. This validates the prediction of the analysis.

Figure 2.11: Numerical solution for pathogen concentration within the biofilm for the two connection network example. In both connections, the pathogens persist in isolation (a) and when part of the network (b), confirming the results of the linear stability analysis. Here, we set $Q_1 = 0.65 < Q_c$ and $D_1 = -3$ so that $Q_2 = Q_1 - D_1 = 3.65 < Q_c$. Note that we set a negative demand, which implies fluid is added at the junction.
Figure 2.12: Numerical solution for pathogen concentration within the biofilm for the two connection network example. The parent connection has the pathogens persist and the child connection has the pathogens washout in isolation (a) but, when part of the network (b), the pathogen concentration does not go to zero in the child connection, indicating that the parent connection induces a non-zero pathogen concentration in the child connection, confirming the results of the linear stability analysis. Here, we set $Q_1 = 0.65 < Q_c$ and $D_1 = -5$ so that $Q_2 = Q_1 - D_1 = 5.65 > Q_c$. Note that we set a negative demand, which implies fluid is added at the junction.
Medium-Size Network Example

Here, we introduce a slightly larger water distribution network. This network consists of $M = 11$ junctions and $N = 11$ connections. The topology of the network and the flow through the network is shown in Figure 2.13.

![Figure 2.13: Topology and flow for the medium-size network example. The lines indicate connections (pipes) and the filled circles indicate the junctions. The triangles indicate the direction of the flow through the connections. The circled numbers indicate the index of each connection and the uncircled numbers indicate the index assigned to each junction.]

We determine the linear stability of each connection independently of the network and compare it to the steady state numerical solution. Since the linear stability assumes a general perturbation from the trivial steady state, we introduce pathogens into the bulk fluid in all connections and no pathogens within the biofilm in any connections.

Figure 2.14 shows the linear stability of the trivial steady state for each connection using the algorithm discussed in Section 2.2.4. The dashed green on black lines indicate connections where the pathogens washout. The red connections indicate the connections where the pathogens persist and the yellow dotted lines on black denote connections that are induced by upstream connections to have a non-zero pathogen concentration as $t \rightarrow \infty$. We can see that of the eleven connections, there are eight (4-11) that would washout for non-zero pathogen concentrations for long times. Of these eight, four (4, 5, 6, 11) would persist outside of the network and the other four connections (7, 8, 9, 10) would have the pathogens washout if they were not in the network. This suggests that a possible approach to disinfection would
be to focus on the four persistent connections (4, 5, 6, 11). If the pathogens were removed from these connections, the other four connections (7, 8, 9, 10) would have the pathogens washout due to only the fluid flow. This observation suggests that connections 4, 5, 6, and 11 are possible vulnerable points within the network.

To confirm the long-time behavior predicted by the linear stability analysis, we numerically simulated the entire network until a steady state was reached. Figure (2.15) shows the steady state numerical solution for this network. The dashed black-on-green lines indicate the pathogens are washed out of that connection. Solid yellow, orange, or red lines indicate that the pathogens persist within that connection (both in the network and in isolation) with the color indicating low, intermediate, and high concentrations, respectively. The dotted black-on-yellow/orange lines indicate the connections were induced by an upstream connection and have a non-zero pathogen concentration. Again, the color indicates the relative pathogen concentration as with the solid lines. Both Figures 2.14 and 2.15 give equivalent results, even though connection 5 appears in different colors. To clarify this point, we emphasize that none of the ancestors of this connection is contaminated by pathogen which implies that the contamination is caused by persisting in isolation, and hence it appears with red color in Figure 2.14. On the other hand, the pathogens that introduced into the fluid grow and are transported within the network and, as the network achieves the equilibrium, each
Figure 2.15: Steady state numerical solution for the medium-sized network example. The lines indicate the connections (pipes) and the circles indicate the junctions, which are numbered. The triangles indicate the direction of flow. The dashed black-on-green lines denote connections where the pathogens will washout. Solid yellow, orange, or red lines indicate that the pathogens persist within that connection both in the network and in isolation. The dotted black-on-yellow/orange lines indicate the connections were induced by an upstream connection and have a non-zero pathogen concentration. The color indicates the relative pathogen concentration as with the solid lines. The yellow, orange, and red colors indicate relatively low, intermediate, and high pathogen concentrations, respectively.

connection has received different amount of pathogens. The pathogens concentration within connection 5, which is labelled persistent, is lower that the concentration is connection 9, which is labelled induced.

**Large Network Example**

We present here an example to illustrate how our methodology scales to a larger water distribution network. Consider a water distribution network with $M = 134$ junctions and $N = 167$ connections. The network parameters given in Appendix B. The topology of the network and the flow through the network is shown in Figure 2.16

We determine the linear stability of each connection independently of the network and compare it to the steady state numerical solution. Since the linear stability assumes a general perturbation from the trivial steady state, we introduce pathogens into the bulk fluid in all connections and and no pathogens within the biofilm in any connections.

Figure 2.17 shows the linear stability of the trivial steady state for each connection
Figure 2.16: Topology and flow for the large network example. The lines indicate connections (pipes) and the circles indicate the junctions. The triangles indicate the direction of the flow through the connections. The numbers indicate the index assigned to each junction with each connection indexed sequentially starting for lower junctions to larger junctions. Using the algorithm discussed in Section 2.2.4. The dashed green on black lines indicate connections where the pathogens washout. The red connections indicate the connections where the pathogens persist and the yellow dotted lines on black denote connections that are induced by upstream connections to have a non-zero pathogen concentration as $t \to \infty$.

We can see that there are 30 connections where pathogens persist that induces another 52 connections to have non-zero pathogen concentrations. If one focuses the cleanup efforts on the 30 persistent connections, the other 52 connections could then be washed clean as long as there are not pathogens coming in from upstream. This observation could greatly reduce the cleanup costs from the pathogen introduction into a water distribution network and illustrates the possible vulnerable points within the network.

To confirm the long-time behavior predicted by the linear stability analysis, we numerically simulated the entire network until a steady state was reached. Figure 2.18 shows the steady state pathogen concentration within the network.

The connections with high relative pathogen concentration are denoted by red lines. The
Figure 2.17: Linear stability of the trivial steady state for the large network example. The lines indicate the connections (pipes) and the circles indicate the junctions, which are numbered. The triangles indicate the direction of flow. The green dashed lines on black denote connections where the pathogens will washout. The solid red lines indicate the connection where the pathogens persist and dotted yellow lines on black indicate the connections where the pathogens are induced by upstream connections.
Figure 2.18: Steady state numerical solution for the large network example. The lines indicate the connections (pipes) and the circles indicate the junctions, which are numbered. The triangles indicate the direction of flow. The dashed black-on-green lines denote connections where the pathogens will washout. Solid yellow, orange, or red lines indicate that the pathogens persist within that connection both in the network and in isolation. The dotted black-on-yellow/orange lines indicate the connections were induced by an upstream connection and have a non-zero pathogen concentration. The color indicates the relative pathogen concentration as with the solid lines. The yellow, orange, and red colors indicate relatively low, intermediate, and high pathogen concentrations, respectively.
orange lines refers to connections with medium relative pathogen concentration while the yellow line indicate that a pipe has low relative pathogen concentration. The solid lines indicate that the pathogens would persist in that connection in isolation while the dotted black-on-yellow/orange lines indicated that the pathogens would wash out of the connection in isolation but are induced by an upstream connection. Finally, the dashed black-on-green line used for connections where the pathogens washout.

We can see that Figures 2.17 and 2.18 are equivalent, indicating the algorithm described in Section 2.2.4 is able to predict the steady state behaviour of the network, except for the connection between junctions 117 and 120. In the linear stability results, pathogens would persist in this connection in isolation and is labelled red but from the numerical solution, this connection is yellow. This indicates that the pathogen concentration relative to the other connections within the network is low for this connection. It does not indicate that the linear stability result. For this connection, the pathogen persists but at a lower concentration that other connections shown in red.

2.2.6 Discussion

In this chapter, we introduced a theoretical model of the dynamics of a released pathogen within an idealized water distribution system with a time-constant fluid flow through the network. For a single connection, we determined the conditions in which the pathogen persists within the connection or washes out of the system. For a network of connections, we observed that there are three distinct types of dynamics for a connection: “persist”, which corresponds to the pathogens persisting in this connection for the given flow rate independent of the pathogen concentration in the rest of the network, “washout”, which corresponds to the pathogens being removed from this connection, and “induced”, which corresponds to a non-zero pathogen concentration as $t \to \infty$ due to pathogens persisting in an upstream connection. Using these three cases, we are able to develop an efficient algorithm that decomposes the network into individual connections and applies an iteration through the
network to determine the behaviour of the pathogens within each connection of the network. This algorithm is proven to produce equivalent results when compared to Lyapunov stability results for the entire network. Results are computed for medium and large size networks that validate the analysis and demonstrate the feasibility of the described methods for predicting the dynamics of introduced pathogens within a time-constant flow network.
Chapter 3

Periodic Flow Network

3.1 Single Connection Model

In the previous chapter, we introduced and analyzed a simple model of the dynamics of introduced pathogens within a water distribution network with a time-constant flow. Here, we assume the flow within the network is periodic in time with a common period for the entire network. To begin, we examine the pathogen dynamics within a single connection.

3.1.1 Mathematical Model

We begin restating the model discussed in Section 2.1:

\[
\begin{align*}
C_f' &= -\frac{Q(t)}{V}C_f + \beta \frac{S_A}{V}C_b - \alpha C_f \\
C_b' &= \alpha \frac{V}{S_A}C_f - \beta C_b + r C_b \left(1 - \frac{C_b}{K}\right)
\end{align*}
\] (3.1)

with initial conditions \(C_f(0) = C_0\) and \(C_b(0) = 0\). Again, \(C_f\) is the pathogen concentration within the bulk fluid, \(C_b\) is the pathogen concentration within the biofilm, \(V\) is the fluid volume within the connection, \(S_A\) is the surface area of the biofilm, \(\alpha\) is the attachment rate, \(\beta\) is the detachment rate, \(r\) is the growth rate of the pathogens within the biofilm, \(K\) is the

43
the pathogen carrying capacity within the biofilm, and \( Q = Q(t) \) is the non-negative magnitude of time-varying fluid rate through the connection.

The analysis of the model of a single connection for a general \( Q(t) \) is difficult so we will initially restrict our analysis to the case where \( Q(t) \) is a continuous, bounded, non-negative function. We will specifically examine the case where

\[
Q(t) = (Q_2 - Q_1) H_\varepsilon (\sin (\omega t)) + Q_1
\]  

(3.2)

where \( Q_2 > Q_1 \) are the upper and lower bounds on the fluid flow rate, respectively, \( T = 2\pi/\omega \) is the period, and

\[
H_\varepsilon (x) = \frac{1}{\pi} \tan^{-1} \left( \frac{x}{\varepsilon} \right) + \frac{1}{2}
\]  

(3.3)

is a smoothed Heaviside function with \( H_\varepsilon (x) \to H (x) \) as \( \varepsilon \to 0 \). Using the scales \( C_f = K (S_A/V) u \), \( C_b = K v \), and \( t = r^{-1} \tau \), the dimensionless equations are

\[
\begin{cases}
  u' = \hat{\beta} v - \left( \hat{\alpha} + \hat{Q}(t) \right) u \\
  v' = \hat{\alpha} u - \hat{\beta} v + v (1 - v)
\end{cases}
\]  

(3.4)

with the initial conditions \( u(0) = u_0 \) and \( v(0) = 0 \), where \( \hat{\alpha} = \alpha/r \), \( \hat{\beta} = \beta/r \), \( \hat{Q}(t) = Q(t)/(Vr) \), and \( u_0 = C_0 V/(S_A K) \).

### 3.1.2 General Solution of Linear Systems of Differential Equations

In this section, we will review the form of the general solution of inhomogeneous linear systems of differential equations. For a detailed review the general solutions to systems of homogeneous linear systems, see Verhulst [2012].

Consider the homogeneous linear system of equations

\[
x' = A(t) x
\]  

(3.5)
where $x: \mathbb{R} \to \mathbb{R}^n$ and $A: \mathbb{R} \to \mathbb{R}^{n \times n}$ are continuous vector and matrix functions, respectively.

**Definition 3.1.1.** A fundamental matrix, $\Phi: \mathbb{R} \to \mathbb{R}^{n \times n}$, of (3.5) is a nonsingular matrix whose columns are linearly independent solutions of (3.5) and satisfies the matrix differential equation

$$\Phi' = A(t) \Phi$$

for all $t \in \mathbb{R}$.

**Definition 3.1.2.** The general solution of (3.5) that satisfies the initial condition $x(0) = x_0$ is given by

$$x(t) = \Phi(t) \Phi^{-1}(0) x_0,$$

where $\Phi$ is the fundamental matrix of the system (3.5).

Now consider the non-homogeneous version of (3.5) is given by

$$x' = A(t) x + h(t)$$

(3.7)

where $h: \mathbb{R} \to \mathbb{R}^n$ is a continuous vector function.

**Definition 3.1.3.** The general solution of (3.7) is given by

$$x(t) = \Phi(t) \Phi^{-1}(0) x_0 + \int_0^t \Phi(t) \Phi^{-1}(\tau) h(\tau) d\tau$$

(3.8)

where $\Phi$ is the fundamental matrix of the homogeneous system (3.5).

### 3.1.3 Floquet Theory

Floquet Theory is a useful tool for analyzing the stability of linear systems of differential equations with periodic coefficients. For detailed reviews of Floquet Theory, see Adrianova [1995], Barone et al. [1977], Johnson [1980], Gökçek [2004], Klausmeier [2008], Kuchment [1993], and Verhulst [2012].
Consider the homogeneous linear system of equations

\[ x' = A(t)x \tag{3.9} \]

where \( x: \mathbb{R} \rightarrow \mathbb{R}^n \) and \( A: \mathbb{R} \rightarrow \mathbb{R}^{n \times n} \) are continuous vector and matrix functions, respectively.

**Theorem 3.1.4 (Floquet’s Theorem).** If \( A(t) \) is a continuous, \( T \)-periodic matrix function, then for all \( t \in \mathbb{R} \), any fundamental solution for (3.9) can be written in the form

\[ \Phi(t) = P(t)e^{Bt}, \]

where \( P: \mathbb{R} \rightarrow \mathbb{R}^{n \times n} \) is a nonsingular, differentiable, \( T \)-periodic matrix function and \( B \in \mathbb{R}^{n \times n} \) is a constant matrix. Furthermore, if \( \Phi(0) = I \) then \( P(0) = I \).

Note that if \( \Phi(0) = I \) then,

\[
\Phi(T) = P(T)e^{BT} = P(0)e^{BT} = e^{BT}
\]

since \( P(t) = P(t + T) \).

**Definition 3.1.5 (Floquet Multipliers and Exponents).** *The eigenvalues of \( B \), \( \mu_i \in \sigma(B) \) for \( i = 1, \ldots, n \), are called characteristic or Floquet exponents and the eigenvalues of \( e^{BT} \), \( \rho_i = e^{\mu_i T} \in \sigma(e^{BT}) \) for \( i = 1, \ldots, n \), are called characteristic or Floquet multipliers.*

**Corollary 3.1.6.** The general form of the solution of (3.9) in terms of the Floquet exponents is

\[ x(t) = \sum_{i=1}^{n} c_i e^{\mu_i t} p_i(t) \tag{3.10} \]

where \( c_i \) are arbitrary constants that depend on the initial conditions and \( p_i: \mathbb{R} \rightarrow \mathbb{R}^n \) for \( i = 1, 2, \ldots, n \) are \( T \)-periodic vector functions.
Note that the solution (3.10) of (3.9) either grows or decays exponentially as \( t \to \infty \) depending on the signs of the real part of the Floquet exponents, \( \text{Re}(\mu_i) \). Thus, the stability of the trivial steady state of (3.9) is determined by the Floquet exponents \( \mu_i \), which depend on the Floquet multipliers, \( \rho_i \). Hence, the trivial steady state is stable if \( \text{Re}(\mu_i) = T^{-1} \ln(\rho_i) < 0 \) for all \( i = 1, 2, \ldots, n \) and is unstable if \( \text{Re}(\mu_i) > 0 \) for any \( i = 1, 2, \ldots, n \). In terms of the Floquet multiplies, the condition for stability of the trivial steady state is \( |\rho_i| < 1 \) for all \( i = 1, 2, \ldots, n \).

Now consider the homogeneous system with periodic coefficients

\[
x' = f(t, x) \tag{3.11}
\]

with \( f(t + T, x) = F(t, x) \). We can study the stability of a periodic solution, \( \phi \), of (3.11) by linearizing the system about \( \phi \). Since \( \phi \) is solution of (3.11), we have

\[
\phi' = f(t, \phi) \tag{3.12}
\]

Assuming a small perturbation, \( y \), to the periodic solution, we let \( y = x - \phi \) to obtain

\[
x' = y' + \phi' = f(t, y + \phi) \tag{3.13}
\]

Using (3.12) and expanding the right hand side of (3.13) using Taylor series about \( \phi \), we obtain the linearized system

\[
y' = D_x f(t, \phi(t)) y + O(|y|) \tag{3.14}
\]

where \( D_x f \) is the Jacobian of \( f \) with respect to \( x \) evaluated at the periodic solution \( \phi(t) \). Therefore, to study the stability of the solution \( \phi \) of the nonlinear system (3.11) to small perturbations, it is sufficient to study the Floquet multipliers of the linearized system.
3.1.4 Stability of the Trivial Solution

To study the linear stability of the trivial solution, we begin by linearizing system (3.4) about the trivial steady state to get a periodic linear system given by:

\[
\begin{align*}
  u' &= \hat{\beta}v - \left(\hat{\alpha} + \hat{Q}(t)\right)u \\
  v' &= \hat{\alpha}u - \left(1 - \hat{\beta}\right)v,
\end{align*}
\]

(3.15)

where \(\hat{\alpha}\) and \(\hat{\beta}\) are positive, dimensionless constants with \(\hat{Q}(t)\) a positive, continuous, \(T\)-periodic function defined by (3.2).

For the constant flow case, discussed in Section 2.1.2, the bifurcation conditions for the trivial steady state to be stable (washout) are \(\hat{\beta} > 1\) and \(\hat{Q} > Q_c\). For periodic flow, if \(\beta \leq 1\), then the trivial steady state is linearly unstable for all \(\alpha\) and \(\omega\) as well as all flow rates. When \(\beta > 1\), the stability of the steady state depends on the flow rate. If \(\min Q(t) = Q_1 > Q_c\), then the trivial steady state will be linearly stable since the flow rate is always above the critical flow rate and introduced pathogens will washout. Likewise, when \(\max Q(t) = Q_2 < Q_c\), the trivial steady state will be linearly unstable and introduced pathogens will persist. The interesting case is when \(\max Q(t) > Q_c\) and \(\min Q(t) < Q_c\). Figure 3.1 shows the bifurcation diagram when \(\hat{\alpha} = 9\), \(\hat{\beta} = 3\), and \(\omega = 1\). We can see that the trivial steady state is linearly stable for all \(Q_2, Q_1 > Q_c\) and is linearly unstable for all \(Q_1, Q_2 < Q_c\). When \(Q_2 > Q_c > Q_1\), there is a curve, \(Q^*(Q_1)\), that defines the critical value \(Q_2\) such that if \(Q_2 > Q^*(Q_1)\) then the trivial steady state is linearly stable and the pathogen will washout and, when \(Q_2 < Q^*(Q_1)\), the trivial steady state is linearly unstable and the pathogens persist. The equivalent result occurs when \(Q_1 > Q_2\).

To see how the other parameters affect the bifurcation curve \(Q_2^*(Q_1)\), we plot the bifurcation diagram for various \(\omega\) with \(\alpha\) and \(\beta\) fixed (Figure 3.2). We can see as \(\omega\) increases, the bifurcation curve approaches the \(Q_{Avg} = Q_c\) line, maximizing the region where the
When $Q_2 > Q_1 > Q_c$ the trivial steady state is linearly stable and when $Q_1 < Q_2 < Q_c$, the trivial steady state is linearly unstable. If the maximum and minimum flow rates straddle the critical flow rate, $Q_2 > Q_c > Q_1$, then the trivial steady state is linearly stable when $Q_2 > Q^*(Q_1)$. Here, $\alpha = 9$, $\beta = 3$, and $\omega = 1$ so $Q_c = 9/2$. 

Figure 3.1: Bifurcation diagram for the bifurcation parameters $Q_1$ and $Q_2$ when $\hat{\beta} > 1$. The diagram shows the transitions between persistency and washout regimes as the flow rates vary.
pathogens would washout of the system. As $\omega$ decreases, the region in the parameter space where the pathogens are predicted to washout is reduced with the bifurcation curves appears to approach a limiting curve. This suggests that higher frequency oscillations in the flow might suppress the potential for the pathogens to persist. Figure 3.3 shows the bifurcation diagram for various $\alpha$ and $\beta$ with $\omega$ and $Q_c$ fixed. As $\alpha$ and $\beta$ increase, the bifurcation curve approaches the $Q_{Avg} = Q_c$ line, suggesting that increasing attachment rates with a corresponding increase in the detachment rate reduces the potential for persistence. For $\alpha \to 0$, with $\beta \to 1$ to keep $Q_c$ fixed, the region in parameter space where the pathogens are predicted to washout is reduced. This suggests that persistence is more strongly affected by $\beta$ that $\alpha$ since decreasing the attachment rate has the opposite effect than what would be expected when the detachment rate is reduced correspondingly.

Figure 3.2: Bifurcation diagram for the bifurcation parameters $Q_1$ and $Q_2$ when $\beta > 1$ for various $\omega$. We can see as $\omega$ increases, the bifurcation curve $Q_2 = Q_2^*(Q_1)$ approaches the $Q_{Avg} = Q_c$ line. As $\omega$ decreases, the bifurcation curve changes little and appears to approach a limiting curve as $\omega \to 0$. Here, $\alpha = 9$ and $\beta = 3$ so $Q_c = 9/2$. 
Figure 3.3: Bifurcation diagram for the bifurcation parameters $Q_1$ and $Q_2$ when $\hat{\beta} > 1$ for various $\alpha$ and $\beta$ with $Q_c$ fixed. We can see that as $\alpha$ increases, with $\beta$ increasing to keep $Q_c$ fixed, the bifurcation curve approaches the $Q_{Avg} = Q_c$ line. As $\alpha \to 0$, with $\beta \to 1$ to keep $Q_c$ fixed, we see that the washout region of the parameter space shrinks. Here, $Q_c = 9/2$ and $\omega = 1$. 

\[ \text{Diagram showing bifurcation curves for various } \alpha \text{ and } \beta. \]
3.1.5 Comparison of Numerical and Theoretical Results

To validate the Floquet analysis, we performed numerical simulations on the full nonlinear system, (3.4) for various parameters. Here, we fix $\hat{\alpha} = 9, \hat{\beta} = 3, \omega = 1,$ and $Q_2 = 6 > Q_c$ while varying $Q_1$. For these parameters, $Q_c = 9/2$. The numerical solutions for $Q_1 = 2, 3, 4, 5$ are shown in Figure 3.4.

For $Q_1 = 5 > Q_c$, the lower bound of the flow rate is above critical flow rate when the flow is constant. As expected, the pathogens washout of the system, which is shown in Figure 3.4a. For $Q_1 = 4 < Q_c$, the lower bound of the flow rate is below the $Q_c$ while the upper bounds is above $Q_c$ with $Q_{Avg} = 5 > Q_c$. The numerical solution, shown in Figure 3.4b, shows that the pathogens washout of the system, confirming the result from the Floquet theory (Figure 3.1).

For $Q_1 = 3 < Q_c$, the lower bound of the flow rate is again below $Q_c$ with the upper bound above $Q_c$. Here, the average flow rate, $Q_{Avg} = 9/2 = Q_c$, is equal to the critical flow rate. For a time-constant flow, this would be the bifurcation point and the trivial steady state would be marginally stable but due to the flow varying periodically in time, the trivial steady state is linearly unstable and the numerical solution (Figure 3.4c) approaches a stable periodic solution in time. This is the result predicted by Floquet theory.

For $Q_1 = 2 < Q_c$, again the lower bound of the flow rate is above $Q_c$ with the upper bound above $Q_c$. Here, the average flow rate, $Q_{Avg} = 4 < Q_c$, is less than the critical flow rate. The results from Floquet theory predicts that the trivial steady state will be linearly unstable, which is confirmed in the numerical solution shown in Figure 3.4d.

3.2 Periodic Flow Network Model

3.2.1 Mathematical Model

While we will be restricting our study to periodic flows within the network, the derivation of the network model is the same for any time-dependent flow.
Figure 3.4: Numerical results for the four different values of $Q_1$ where the solution has reached a steady amplitude. The pathogen concentration is shown for different values of $Q_1$ with $Q_2 = 6$ and $\omega = 1$ fixed. The biofilm parameters are $\alpha = 9$, $\beta = 3$, and $r = 1$, which yield $Q_c = 9/2$. For (a), $Q_1 = 5 > Q_c$ and the pathogens washout of the system. For (b), $Q_1 = 4 < Q_c$ so from Figure 3.1 we have $Q_2 > Q_2^*(4)$ and the pathogens washout of the system. For (c), $Q_1 = 3$ so that $Q_{\text{Avg}} = 9/2 = Q_c$. For a constant flow rate, this would be the bifurcation point and the trivial steady state would be marginally stable but due to the flow varying periodically in time, the trivial steady state is linearly unstable and the pathogens persist within the system. For (d), $Q_1 = 2 < Q_c$ and $Q_{\text{Avg}} = 4 < Q_c$ so the pathogens are predicted to persist within system by the analysis, which is confirmed by the numerical solution.
For the time-dependent flow network, we follow a similar derivation as for the
time-constant flow case, discussed in Chapter 2. Again, the connectivity of the network is defined
using the junction-to-junction connectivity matrix, $C$, as well as the connection-to-junction
connectivity matrix, $H$. Both of these matrices are constant for any time-dependent flow.
Note that $H$ contains direction information about the network that is independent of the
actual flow direction, using the index numbering of the junctions to indicate a direction
from smaller index to larger. The directed connection-to-junction connectivity matrix, $H_Q$,
contains the direction of the flows through each connection. For the time-constant flow case,$H_Q$ was a constant matrix but when the flow rate can vary in time, it can become time
dependent. This occurs when the flow rate can change sign, indicated that the direction
of the flow has reversed in the connection. For simplicity, we will assume that there is no
change in the flow direction for any connection within the network and take $H_Q$ as a constant
matrix for all time. Consequently, all the connectivity matrices will be constant in time and
the ancestry/descendent relationships between connections in the network will not change
as time progresses.

Using the quasi-steady state assumption within each junction, the time-dependent flow
network model is
\[
\begin{align*}
\mathbf{u}' &= (K(t) - Q(t)) \mathbf{u} + f(\mathbf{u}, \mathbf{v}) \\
\mathbf{v}' &= g(\mathbf{u}, \mathbf{v})
\end{align*}
\]  
(3.16)

where $\mathbf{u}$ is the vector of the pathogen concentration within the bulk fluid for each connection,$\mathbf{v}$ is the vector of the pathogen concentration within the biofilm for each connection, the
connection-to-connection inflows defined by the matrix $K(t) = Z(t) (T(t) + D(t))^{-1} W(t) \in \mathbb{R}^{N \times N}$. The inflow matrix functions $Z: \mathbb{R} \to \mathbb{R}^{N \times M_I}$ and $W: \mathbb{R} \to \mathbb{R}^{M_I \times N}$ indicate the flow
into the connections and junctions, respectively, and defined as

\[
Z(t) = Q(t) (H_I < 0) \quad \text{and} \quad W(t) = (H_T^T > 0) Q(t).
\]  
(3.17)
The outflow matrix functions $T : \mathbb{R} \rightarrow \mathbb{R}^{M_i \times M_i}$ and $D : \mathbb{R} \rightarrow \mathbb{R}^{M_i \times M_i}$ indicate the flow out of junction into connections and the demands at each junction that remove fluid from the network, respectively, with $T(t) = (H^T_t < 0) Q(t) (H_t < 0)$.

The vector functions $f(u, v)$ and $g(u, v)$ describe the attachment/detachment/growth dynamics of the pathogens and were previously defined in Section 2.2.1. Again, they are

$$f(u, v) = \Gamma^{-1} B v - A u \quad (3.18)$$

and

$$g(u, v) = \Gamma A u + (R - B) v + Y^{-1} R (v \odot v) \quad (3.19)$$

where the $v \odot v$ indicates component-wise multiplication of the vector. Each matrix is a diagonal matrix of connection parameters where, for the $i$th connection, $A_{i,i} = \alpha_i$ is the pathogen attachment rate from the bulk fluid to the biofilm, $B_{i,i} = \beta_i$ is the pathogen detachment rate from the biofilm to the bulk fluid, $\Gamma_{i,i} = \gamma_i = V_i/S_{A_i}$ is the bulk fluid volume to connection surface area ratio, $Y_{i,i} = K_i$ is the carrying capacity of the pathogens within the biofilm, and $R_{i,i} = r_i$ is the pathogen growth rate in the biofilm. Each are considered to be constant and positive so each diagonal matrix is invertible.

Along with initial conditions, the time-dependent flow rate, $Q(t)$, need to be defined for well-posedness and is defined in the next section.

### 3.2.2 Determining the Periodic Flow Rate

For simplicity, we will again assume that the flow through the connections can be modeled as a linear resistance pipe as was done in Section 2.2.2. The flow rate is defined as

$$Q(t) = \frac{P_a(t) - P_b(t)}{R} \quad (3.20)$$
where \( P_a : \mathbb{R} \to \mathbb{R} \) and \( P_b : \mathbb{R} \to \mathbb{R} \) are the time-dependent pressures at the ends of the pipe and \( R \) is the hydraulic resistance of the pipe and is assumed to be constant in time. Note that we have assumed an instantaneously response in the flow rate due to changes in pressure, neglecting transient effects.

The flow rates for a network, defined by \( N \) connections and \( M \) junctions, are given by

\[
r_f \odot q(t) = H p(t) \tag{3.21}
\]

where \( H \) is a connection to junction connectivity matrix, \( q : \mathbb{R} \to \mathbb{R}^N \) is the vector function of flow rates for each connection. \( r_f \in \mathbb{R}^N \) is the time-constant vector of hydraulic resistances for each connection, \( p : \mathbb{R} \to \mathbb{R}^N \) is the vector function of pressures at the junctions, which is defined at each non-interior junction. Solving (3.21) along with the pressure conditions at the non-interior junctions, and the equations given by conservation of mass at each interior junction given by

\[
\sum_{k=1}^{N_{\text{inflow}}} q_j(t) - \sum_{k=1}^{N_{\text{outflow}}} q_k(t) = D(t), \tag{3.22}
\]

we can determine all \( N + M \) pressures and flow rates. Here, \( D : \mathbb{R} \to \mathbb{R} \) is the demand at the junction and we have again neglected transient effects by assuming the flow rates respond instantaneously to changes in the demands.

The equations for the entire network, using the connection-to-junction connectivity matrix \( H \) is

\[
\Delta H^T = d \tag{3.23}
\]

where \( \Delta \in \mathbb{R}^{M \times M} \) is a diagonal matrix with \( \Delta_{i,i} = 1 \) is the \( i \)th junction is an interior junction and \( \Delta_{i,i} = 0 \) otherwise. The vector \( d \in \mathbb{R}^M \) is the vector of the demand at each junction. For non-interior junctions, the values of \( d \) are undefined and not used. Combining (3.21), (3.23), and the conditions on the pressure at non-interior junctions, the linear system governing the
flow rates through the connections and the pressures at the junctions for time $t$ is

$$
\begin{bmatrix}
I - \Delta & \Delta H^T \\
H & R_f
\end{bmatrix}
\begin{bmatrix}
p(t) \\
q(t)
\end{bmatrix} =
\begin{bmatrix}
b(t) \\
0
\end{bmatrix},
$$

where $p(t)$ is the vector of pressures at each junction at time $t$, $I$ is the identity matrix, $R_f = \text{diag}(r_f)$ is the diagonal matrix of hydraulic resistances for each connection, and $b(t) = \Delta d_0(t) + (I - \Delta)p_0(t)$. Here, $d_0(t) = D_i(t)$ for each interior junction and $p_0 = P_i(t)$ for each non-interior junction. Solving this system provides the flow rate $q(t)$ at time $t$, which we use to define the flow rate matrix $Q(t) = \text{diag}(q(t))$.

For our study, we wish to restrict the flow rate to periodic functions in time. This can be achieved for the $i$-th connection by letting the pressures at the ends be periodic functions in time with period $T_i$ ($P(t + T_i) = P(t)$). In addition to each connection being periodic, we want the entire network to be periodic with a single period. This can be achieved by letting the demands at the $j$-th interior junction be periodic in time with period $T_j$ and requiring that all periods of the entire network be an integer multiple of all the periods for each pressure and demand, $T = n_iT_i$ and $T = n_jT_j$ for all $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, M_I$.

### 3.2.3 Theorems

We now consider the linear non-homogeneous system governing a single connection

$$
u' = A(t)u + h(t)
$$

where

$$A(t) = 
\begin{bmatrix}
-(\alpha + Q(t)) & \beta \\
\gamma & r - \beta
\end{bmatrix},
$$

$\alpha$, $\beta$, $\gamma$, and $r$ are positive constants, $Q = Q(t)$ is a piecewise-continuous, bounded, non-negative function, and $h(t) = [h(t), 0]^T$ with $h(t)$ a piecewise-continuous, non-negative
vector function. We assume that \( h(t) \) takes the form

\[
h(t) = \sum_{j=1}^{N} k_j F_j(t) e^{\nu_j t}
\]  

(3.27)

where \( k_j \) is a constant, \( F_j(t) \) is a bounded, periodic function, and \( \nu_j \in \mathbb{R} \). Note that \( h(t) \to 0 \) or \( h(t) \to \infty \) as \( t \to \infty \) depending of the signs of \( \nu_j \).

Let \( u = u_h + u_p \) be the solution vector of (3.25) for \( t \geq 0 \) where

\[
u_{\nu_j}(t) &= \int_{0}^{t} \Phi(t) \Phi^{-1}(\tau) h(\tau) d\tau
\]  

(3.28)

are the homogeneous and particular solutions of (3.25), respectively, where \( \Phi(t) \) is the fundamental matrix solution and is invertible for all \( t \geq 0 \).

Writing \( u = [u_h(t), v_h(t)]^{T} \), corresponding homogeneous system is

\[
\begin{align*}
u_{\nu_j}(t) &= \int_{0}^{t} \Phi(t) \Phi^{-1}(\tau) h(\tau) d\tau
u'_{\nu_j}(t) &= \alpha \gamma u_h + (r - \beta)v_h
\end{align*}
\]

(3.29)

with \( u_h, v_h \in C^{0} \) for all \( t \geq 0 \).

**Lemma 3.2.1.** If \( u > 0 \) at \( t = 0 \), then the solution of the homogeneous equation (3.29), \( u_h \), is non-negative for all \( t \geq 0 \).

**Proof.** Since \( u_h, v_h \in C^{0} \) and \( u > 0 \) at \( t = 0 \), the intermediate value theorem states that \( u_h \) or \( v_h \) can only become negative by passing through \( u_h = 0 \) and/or \( v_h = 0 \). Since \( u_h' = 0 \) when \( u_h = v_h = 0 \), it is not possible for \( u_h \) or \( v_h \) to become negative by passing through \( u_h = v_h = 0 \).

If we assume that \( u_h = 0 \) and \( v_h > 0 \), then \( u_h \) would become negative if \( u_h' < 0 \). In this case, \( u_h' = \beta/\gamma v_h > 0 \), so this is not possible.

Likewise, if we assume that \( v_h = 0 \) and \( u_h > 0 \), then \( v_h \) would become negative if \( v_h' < 0 \).
In this case, $v' = \alpha \gamma u_h > 0$, so this is not possible. Thus, $u_h > 0$ for all $t \geq 0$.

We next consider the behaviour of the particular solution in terms so the homogeneous solution, $u_h$, and the forcing function, $h(t)$.

**Lemma 3.2.2.** For the system (3.25), the following hold

(i) If $u_h \to 0$ and $h(t) \to 0$ as $t \to \infty$, then $u_p \to 0$.

(ii) If $u_h \to \infty$ and $h(t) \to 0$ as $t \to \infty$, then $u_p \to \infty$.

(iii) If $u_h \to 0$ and $h(t) \to \infty$ as $t \to \infty$, then $u_p \to \infty$.

(iv) If $u_h \to \infty$ and $h(t) \to \infty$ as $t \to \infty$, then $u_p \to \infty$.

**Proof.** Recall that the general form of the homogeneous solution of system (3.25) is given by Floquet Theory:

$$u_h(t) = \sum_{i=1}^{2} c_i e^{\mu_i t} p_i(t)$$

(3.30)

where $c_i$ are arbitrary constants that depend on initial conditions, $\mu_i$ are Floquet exponents, and $p_i(t)$ are piecewise-continuous, $T$-periodic vector functions for $i = 1, 2$. Additionally, by (3.6),

$$u_h = \Phi(t) \Phi^{-1}(0) u_0$$

so we have

$$\Phi(t) = P(t) e^{Bt}$$

where

$$e^{Bt} = \begin{bmatrix} e^{\mu_1 t} & 0 \\ 0 & e^{\mu_2 t} \end{bmatrix}.$$ (3.31)

Hence,

$$\Phi^{-1}(t) = e^{-Bt} P^{-1}(t).$$

Therefore,

$$u_h = P(t) e^{Bt} P^{-1}(0) u_0$$

59
and

\[ u_p(t) = \int_0^t P(t) e^{B(t-\tau)} P^{-1}(\tau) h(\tau) \, d\tau, \]

where

\[ P(t) = \begin{bmatrix} P_{11}(t) & P_{12}(t) \\ P_{21}(t) & P_{22}(t) \end{bmatrix}, \]

(3.32)

and \( h(\tau) = [h(\tau), 0]^T \). Since \( u_p = [u_p(t), v_p(t)]^T \), we can write the particular solution in terms of the individual elements:

\[ u_p(t) = \int_0^t \frac{1}{\det(P(\tau))} \left[ P_{11}(t) e^{\mu_1(t-\tau)} P_{22}(\tau) - P_{12}(t) e^{\mu_2(t-\tau)} P_{21}(\tau) \right] h(\tau) \, d\tau \]

and

\[ v_p(t) = \int_0^t \frac{1}{\det(P(\tau))} \left[ P_{21}(t) e^{\mu_1(t-\tau)} P_{22}(\tau) - P_{22}(t) e^{\mu_2(t-\tau)} P_{21}(\tau) \right] h(\tau) \, d\tau. \]

Without loss of generality, we assume that \( \mu_1 > \mu_2 \), therefore, as \( t \to \infty \),

\[ u_p(t) \approx \int_0^t \frac{1}{\det(P(\tau))} P_{11}(t) e^{\mu_1(t-\tau)} P_{22}(\tau) h(\tau) \, d\tau \]

and

\[ v_p(t) \approx \int_0^t \frac{1}{\det(P(\tau))} P_{21}(t) e^{\mu_1(t-\tau)} P_{22}(\tau) h(\tau) \, d\tau. \]

Using the assumed form of \( h(t) \) given by 3.27, there exists \( M_j > 0 \) such that \( |F_j| < M_j \) for all allowable \( j \). Since from Floquet theory, \( P \) is nonsingular and \( |\det(P)| > m > 0 \) for some positive constant \( m \). Furthermore, \( P \) is bounded so there exists \( M > 0 \) such that
\(|P_{ij}| < M\) for \(i, j = 1, 2\). Hence,

\[
u_p(t) \approx \sum_{j=1}^{N_h} \int_0^t \frac{P_{11}(t) P_{22}(\tau) F_j(\tau)}{\det(\mathbf{P}(\tau))} e^{\mu_1 t + (\nu_j - \mu_1) \tau} d\tau
\]

\[
\leq \sum_{j=1}^{N_h} \frac{k_j M^2 M_j m}{m} \int_0^t e^{\mu_1 t + (\nu_j - \mu_1) \tau} d\tau
\]

and

\[
u_p(t) \approx \sum_{j=1}^{N_h} k_j \int_0^t \frac{P_{21}(t) P_{22}(\tau) F_j(\tau)}{\det(\mathbf{P}(\tau))} e^{\mu_1 t + (\nu_j - \mu_1) \tau} d\tau
\]

\[
\leq \sum_{j=1}^{N_h} \frac{k_j M^2 M_j}{m} \int_0^t e^{\mu_1 t + (\nu_j - \mu_1) \tau} d\tau.
\]

Therefore,

\[
u_p(t), \nu_p(t) \leq \sum_{j=1}^{N_h} \frac{k_j M^2 M_j}{m} g_j(t) \tag{3.33}
\]

where

\[
g_j(t) = \begin{cases} 
\frac{e^{\nu_j t} - e^{\mu_1 t}}{\nu_j - \mu_1} & \text{if } \nu_j \neq \mu_1 \\
t e^{\mu_1 t} & \text{if } \nu_j = \mu_1
\end{cases}
\]

Hence, the following holds:

(i) If \(u_h \to 0\) and \(h(t) \to 0\), then \(\mu_1 < 0\) and \(\nu_j < 0\) for all \(j\). Hence, \(e^{\mu_1 t}, e^{\nu_j t} \to 0\) for all \(j\) as \(t \to \infty\) so \(u_p \to 0\).

(ii) If \(u_h \to \infty\) and \(h(t) \to 0\) then \(\mu_1 > 0\) and \(\nu_j < 0\) for all \(j\). Hence, \(e^{\mu_1 t} \to \infty\) and \(e^{\nu_j t} \to 0\) for all \(j\) as \(t \to \infty\) so \(u_p \to \infty\).

(iii) If \(u_h \to 0\) and \(h(t) \to \infty\) then \(\mu_1 < 0\) and \(\nu_j > 0\) for some \(j\). Hence, \(e^{\mu_1 t} \to 0\) and \(e^{\nu_j t} \to \infty\) for some \(j\) as \(t \to \infty\) so \(u_p \to \infty\).

(iv) If \(u_h \to \infty\) and \(h(t) \to \infty\) then \(\mu_1 > 0\) and \(\nu_j > 0\) for some \(j\). Hence, \(e^{\mu_1 t}, e^{\nu_j t} \to \infty\) for some \(j\) as \(t \to \infty\) so \(u_p \to \infty\).
We now consider the behavior of the system governing the entire network, given by (3.16). Linearizing (3.16) about the trivial steady state yields the linear system

\[
\begin{bmatrix}
u' \\
v'
\end{bmatrix} = \begin{bmatrix} K(t) - Q(t) - A & \Gamma^{-1}B \\
\Gamma A & R - B \end{bmatrix} \begin{bmatrix} u \\
v\end{bmatrix}.
\] (3.35)

**Theorem 3.2.3.** The long-time behavior of the elements of the solution, \( u_i = [u_i, v_i]^T \) for \( i = 1, 2, \ldots, N \), to the linearized system (3.35) is

(i) \( u_i \to \infty \) if \( \tilde{u}_i \to \infty \).

(ii) \( u_i \to 0 \) if \( \tilde{u}_i \to 0 \) and \( \tilde{u}_j \to 0 \) for all \( j \in A(i) \).

(iii) \( u_i \to \infty \) if \( \tilde{u}_i \to 0 \) and \( \tilde{u}_j \to \infty \) for any \( j \in A(i) \).

where \( \tilde{u}_i \) is the solution to the linearized system (3.15) for each connection in isolation.

**Proof.** Consider a network that consists of \( N \) connections and the connections are indexed by generations so that a higher indexed connection cannot be an ancestor of a smaller indexed connection, i.e., \( j \not\in A(i) \) for \( j = i + 1, \ldots, N \). Assume that the first \( m \) connections have no parents. Hence, \( K_{i,j} = 0 \) for \( i = 1, \ldots, m, \ j = 1, \ldots, i \).

We rewrite the system to separate the first connection from the system:

\[
\begin{bmatrix}
u_1 \\
v_1
\end{bmatrix} = \begin{bmatrix}-\hat{q}_1(t) - \alpha_1 & 0 & \gamma_1^{-1} \beta_1 & 0 \\
c^{(1,1)}(t) & K^{(1)}(t) - Q^{(1)}(t) - A^{(1)} & 0 & G^{(1)} \\
\gamma_1 \alpha_1 & 0 & r_1 - \beta_1 & 0 \\
0 & \Gamma^{(1)} A^{(1)} & 0 & R^{(1)} - B^{(1)}
\end{bmatrix} \begin{bmatrix} u_1 \\
v_1 \\
v^{(1)}
\end{bmatrix}.
\] (3.36)
where

\[
\mathbf{u}^{(p)} = [u_{p+1}, u_{p+2}, \ldots, u_N]^T \in \mathbb{R}^{N-p}
\]  

(3.37)

\[
\mathbf{v}^{(p)} = [v_{p+1}, v_{p+2}, \ldots, v_N]^T \in \mathbb{R}^{N-p}
\]  

(3.38)

\[
\mathbf{c}^{(p,q)} = [K_{q+1,p}, K_{q+2,p}, \ldots, K_{N,p}]^T \in \mathbb{R}^{N-p-q}
\]  

(3.39)

with

\[
K^{(p)} = [K_{i,j}] \text{ for } i, j = p + 1, \ldots, N
\]  

(3.40)

\[
Q^{(p)} = [Q_{i,j}] \text{ for } i, j = p + 1, \ldots, N
\]  

(3.41)

\[
A^{(p)} = [A_{i,j}] \text{ for } i, j = p + 1, \ldots, N
\]  

(3.42)

\[
B^{(p)} = [B_{i,j}] \text{ for } i, j = p + 1, \ldots, N
\]  

(3.43)

\[
R^{(p)} = [R_{i,j}] \text{ for } i, j = p + 1, \ldots, N
\]  

(3.44)

\[
\Gamma^{(p)} = [\Gamma_{i,j}] \text{ for } i, j = p + 1, \ldots, N
\]  

(3.45)

\[
G^{(p)} = [B_{i,j}/\Gamma_{i,j}] \text{ for } i, j = p + 1, \ldots, N
\]  

(3.46)

Note that \(c^{(p,q)}_i = 0\) if \(i \notin D(p)\).

We now have two systems,

\[
\begin{bmatrix}
  u_1 \\
v_1
\end{bmatrix}' =
\begin{bmatrix}
  -q_1(t) - \alpha_1 & \gamma_1^{-1} \beta_1 \\
\gamma_1 \alpha_1 & r_1 - \beta_1
\end{bmatrix}
\begin{bmatrix}
  u_1 \\
v_1
\end{bmatrix}
\]  

(3.47)

and

\[
\begin{bmatrix}
  \mathbf{u}^{(1)} \\
  \mathbf{v}^{(1)}
\end{bmatrix}' =
\begin{bmatrix}
  K^{(1)}(t) - Q^{(1)}(t) - A^{(1)} & G^{(1)} \\
\Gamma^{(1)}A^{(1)} & R^{(1)} - B^{(1)}
\end{bmatrix}
\begin{bmatrix}
  \mathbf{u}^{(1)} \\
  \mathbf{v}^{(1)}
\end{bmatrix} +
\begin{bmatrix}
  \mathbf{h}^{(1)}(t) \\
  0
\end{bmatrix}
\]  

(3.48)

where \(h^{(1)}(t) = u_1(t) c^{(1,1)}(t)\). Since connections \(i = 1, 2, \ldots, m\) are assumed to have no
parents, we have $h_i^{(1)} = 0$ for $i = 1, 2, \ldots, m - 1$.

We can repeat this for $i = 2, \ldots, m$ to obtain

$$
\begin{bmatrix}
  u_i \\
  v_i
\end{bmatrix}' =
\begin{bmatrix}
  -q_i(t) - \alpha_i & \gamma_i^{-1} \beta_i \\
  \gamma_i \alpha_i & r_i - \beta_i
\end{bmatrix}
\begin{bmatrix}
  u_i \\
  v_i
\end{bmatrix}
$$

(3.49)

for $i = 1, 2, \ldots, m$ and

$$
\begin{bmatrix}
  u^{(m)}_i \\
  v^{(m)}_i
\end{bmatrix}' =
\begin{bmatrix}
  K^{(m)}(t) - Q^{(m)}(t) - A^{(m)} & G^{(m)} \\
  \Gamma^{(m)} A^{(m)} & R^{(m)} - B^{(m)}
\end{bmatrix}
\begin{bmatrix}
  u^{(m)}_i \\
  v^{(m)}_i
\end{bmatrix} +
\begin{bmatrix}
  h^{(m)}_i(t) \\
  0
\end{bmatrix}
$$

(3.50)

where

$$
\mathbf{h}^{(m)}_i(t) = \sum_{j=1}^{m} c^{(j,m)}_i(t) u_j(t).
$$

(3.51)

The first $m$ systems are the systems that govern their corresponding connections in isolation and are independent of each other and the remaining system. The last system represents a subnetwork consisting of the remaining connections where the removed connections are represented by forcing functions.

We now note that the long-time behavior of the solutions to (3.49) can be determined using Floquet theory as described in Section 3.1.4. Since these connections have no parents, the result is proven for these connections.

For the remaining subnetwork, we again determine the connections without parents. Assuming connections $i = m + 1, \ldots, m_1$ have no parents, we can again split the subnetwork system (3.50) into

$$
\begin{bmatrix}
  u_i \\
  v_i
\end{bmatrix}' =
\begin{bmatrix}
  -q_i(t) - \alpha_i & \gamma_i^{-1} \beta_i \\
  \gamma_i \alpha_i & r_i - \beta_i
\end{bmatrix}
\begin{bmatrix}
  u_i \\
  v_i
\end{bmatrix} +
\begin{bmatrix}
  h_i^{(m)}(t) \\
  0
\end{bmatrix}
$$

(3.52)
for $i = m + 1, \ldots, m_1$ and

\[
\begin{pmatrix}
    u^{(m_1)} \\
    v^{(m_1)}
\end{pmatrix}' =
\begin{pmatrix}
    K^{(m_1)}(t) - Q^{(m_1)}(t) - A^{(m_1)} & G^{(m_1)} \\
    \Gamma^{(m_1)} A^{(m_1)} & R^{(m_1)} - B^{(m_1)}
\end{pmatrix}
\begin{pmatrix}
    u^{(m_1)} \\
    v^{(m_1)}
\end{pmatrix} +
\begin{pmatrix}
    h^{(m_1)}(t) \\
    0
\end{pmatrix}
\]

where

\[
h_i^{(m)}(t) = \sum_{j=1}^{m-1} c_{1}^{(j,m)}(t) u_j(t) = \sum_{j=1}^{m-1} K_{m+i,j}(t) u_j(t).
\]

We again have set of independent systems and a new subnetwork.

Since the elements of $K$ are bounded periodic functions of time and we have computed the long-time behavior of $u_i$ for $i = 1, 2, \ldots, m$ from Floquet theory, we can write

\[
h_i^{(m)}(t) = \sum_{j=1}^{m-1} K_{m+i,j}(t) \left( c_{j,1} p_{j,1}(t) e^{\mu_{j,1} t} + c_{j,2} p_{j,2}(t) e^{\mu_{j,2} t} \right).
\]

where $c_{j,1}$ and $c_{j,2}$ are constants, $p_{j,1}(t)$ and $p_{j,2}(t)$ are bounded periodic functions of time, and $\mu_{j,1}$ and $\mu_{j,2}$ are the Floquet exponents for the $j$-th connection in isolation.

To determine the long-time behavior of $h_i^{(m)}(t)$, we first denote the set of indices $\mathcal{J}_i^0$ whose elements correspond to connections which have Floquet exponents with positive real parts:

\[
\mathcal{J}_i^0 = \{ j \mid 0 < j < i \land \text{Re}\{\mu_{j,k}\} > 0 \text{ for } k = 1 \text{ and/or } 2 \}.
\]

We now state that $h_i^{(m)}(t) \to \infty$ as $t \to \infty$ if $K_{m+i,j} \neq 0$ for and $j \in \mathcal{J}_i^0$. This condition states that the $i$-th forcing function in the vector will grow unbounded if the $i$-th connection has a parent that grows unbounded. If $\mathcal{J}_i^0 = \emptyset$ or $K_{m+i,j} = 0$ for all $j \in \mathcal{J}_i^0$ then $h_i^{(m)}(t) \to 0$ as $t \to \infty$. These conditions state that the $i$-th forcing function goes to zero if there are no parent connections whose solution grows unbounded.

Each system in 3.52 has the form given by (3.25). With the form of the forcing function determined, the result for $i = m + 1, \ldots, m_1$ follow directly from Lemma 3.2.2.

For the remaining subnetwork, we repeat the process above by finding connections with
no parents and determine their long-time behavior based on their behavior isolated from the
network and the behavior of their ancestors. The desired result follows as per above.

3.2.4 Efficient Algorithm for Determining the Pathogen Persistence within the
Network

To obtain the long-time behavior of the entire network, we can linearize the system about
the trivial steady state and then, from Floquet theory, we solve the $2N$ matrix differential
equation for $X(t)$ numerically over the period $[0, T]$ with the initial condition given by the
identity matrix $X(0) = I$ to determine the Floquet multiplies and exponents. The Floquet
multipliers, $\rho_i$ for $i = 1, 2..., N$, are determined by computing the eigenvalues of $X(T)$ and
the Floquet exponents are computed using the relation $\mu_i = \log(\rho_i)/T$. The long-time
behavior is determined by the sign of the real part of the Floquet exponents.

To determine the computational cost of determining the Floquet multipliers for the entire
network, we examine the computational costs of the two key steps: solving for $X(T)$ and
computing its eigenvalues. To solve for $X(T)$ using an explicit ODE integrator, each step
requires a matrix-matrix multiplication, which requires $O(8N^3)$ operations, and if we assume
it takes $M$ steps in time, then computing $X(T)$ requires $O(8MN^3)$ operations. Note that
the minimum value for $M$ depends on the largest eigenvalue of the matrix function $A(t)$
for all values of $t \in [0, T)$. We need to compute all the eigenvalues of $X(T)$ so one could
use QR factorization, which is an $O(8N^3)$ operation for this problem. Hence, determining
the stability of the trivial steady state for the entire network requires $O(8(M + 1)N^3)$
operations.

Due to the high computational costs as the size of the network increases, we propose the
following algorithm, whose justifications comes from Theorem 3.2.3. We begin by identifying
all connections that have no parents. By Theorem (3.2.3), the stability of any connection
that has no parents is independent of the rest of the network. We then compute the stability
of these connections assuming they are not part of the network using Floquet theory. Deter-
mining the stability of the $i$-th connection requires $O(8(M_i + 1))$ operations, where $M_i$ is the number of time steps required to solve the matrix equation. Note that $M_i$ is independent of the dynamics of other parts of the network and will be smaller than $M$, which was the time step required for the entire network.

After the stability of the connections without parents, we can remove these connections from the network. The affect of the removed connections is replaced in each of the children with a forcing function that have the same long-time behavior. For the new network, there will be connections without parents and from Theorem 3.2.3, the stability of the trivial steady state is independent of the rest of the network. Using Lemma 3.2.2, we determine the stability of the trivial steady state. We then remove these connections and repeat the process until the stability of all connections are computed.

The above process allows us to determine the stability of the trivial steady state in an iterative manner. This process requires $O(8(M_i + 1))$ for $i = 1, 2, \ldots, N$ so the total operation code is less than $O(8(M + 1)N)$, which is significantly more efficient than computing the Floquet multipliers for the entire network at once.

In practice, it is more efficient to compute the Floquet multipliers for each connection in isolation and then performing an iteration similar to the process described in Section 2.2.4. After we compute the Floquet multipliers for each connection, we use the parent connectivity matrix, $H_p$, to iteratively to classify the connections. The algorithm is

1. Compute the numerical solution of the matrix equation

   \[ X' = A(t)X \]

   with $X(0) = I$ at $t = T$ for each single connection.

2. Calculate Floquet multiplies for each connection in isolation and classify each connection as “persistent” or “washout”.

3. Define the elements of the vector $s^{(0)}$ as $s_i^{(0)} = 1$ if the $i$th connection is “persistent”
and \( s(0)_i = 0 \) if the \( i^{th} \) connection is “washout”.

4. Define the iteration matrix \( Z = I + H_p^T \).

5. Set \( s = s^0 \)

6. Repeat

\[
\begin{align*}
(a) & \quad s = Zs \\
(b) & \quad \text{If } s^0_i = 1, \text{ then set } s_i = 1 \text{ for all } i = 1, \ldots, N. \\
(c) & \quad \text{If } s^0_i = 0 \text{ and } s_i > 0, \text{ then set } s_i = 2 \text{ for all } i = 1, \ldots, N.
\end{align*}
\]

Until \( ||s - s^0|| = 0 \)

After this algorithm is completed, the vector \( s \) indicates the connection is “persistent” when \( s_i = 1 \), “induced” when \( s_i = 2 \), and “washout” when \( s_i = 0 \). This algorithm is trivially parallelizable when computing the Floquet multipliers, making it even more efficient.

### 3.2.5 Theoretical and Numerical Results

#### Small Network with Periodic Flow Rates

Here, we introduce an example of a network consisting of two connections in serial. The dimensionless system is

\[
\begin{align*}
\dot{u}_1 &= \hat{\beta}v_1 - (\hat{q}_1(t) + \hat{\alpha})u_1 \\
\dot{u}_2 &= \hat{q}_2(t)(u_1 - u_2) + \hat{\beta}v_2 - \hat{\alpha}u_2 \\
\dot{v}_i &= \hat{\alpha}u_i - \hat{\beta}v_i + v_i(1 - v_i) \text{ for } i = 1, 2.
\end{align*}
\]

with \( \hat{q}_1, \hat{q}_2 \) are a continuous, bounded, non-negative, periodic functions with periods \( T_1, \) and \( T_2, \) respectively such that \( T_1 = nT_2 \) for some \( n \in \mathbb{Z}^+ \). We assume that \( \hat{q}_1(t) \) has the form

\[
\hat{q}_1(t) = \frac{b - a}{2} \left( \frac{2}{\pi} \tan^{-1} \left( \frac{\sin(\omega t)}{\varepsilon} \right) + 1 \right) + a
\]
and
\[ \dot{q}_2(t) = \dot{q}_1(t) - D \] (3.59)

where \( D \) is the time-constant demand at the junction, \( \varepsilon > 0 \), and the non-negative constants \( a > b \) are the minimum and maximum values of \( \dot{q}_1(t) \), respectively. Here, we have assumed that dimensional constants \( \alpha, \beta, r, \) and \( K \) are the same for each pipe.

To examine the periodic network with period \( T = \max(T_1, T_2) \), we first linearize the nonlinear system (3.57) about the trivial steady state to get the linearized system given by:

\[
\begin{align*}
\dot{u}_1' &= \hat{\beta}v_1 - (\hat{q}_1(t) + \hat{\alpha})u_1 \\
\dot{u}_2' &= \hat{q}_2(t)(u_1 - u_2) + \hat{\beta}v_2 - \hat{\alpha}u_2 \\
v_i' &= \hat{\alpha}u_i - \hat{\beta}v_i \text{ for } i = 1, 2.
\end{align*}
\] (3.60)

We can determine the linear stability of the trivial steady state by numerically computing the Floquet multipliers of the linearized system (3.60). We also compute the linear stability of the trivial steady state using Theorem 3.2.3 by computing the Floquet multipliers for each connection isolated from each other by setting \( u_1 = 0 \) in the \( u_2 \) equation. Note that the first connection has no parents in the network so the \( u_1 \)-equation in (3.60) is the equation governing \( u_1 \) in isolation. The results are summarized in Table 3.1 and confirm Theorem 3.2.3. The four cases for the two-connection example are:

<table>
<thead>
<tr>
<th>In Isolation</th>
<th>In Network</th>
<th># of Floquet Exponents</th>
<th>Form of the Eigenvectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pipe #1</td>
<td>Pipe #2</td>
<td>where Re (( \mu )) &gt; 0</td>
<td>where Re (( \mu )) &gt; 0</td>
</tr>
<tr>
<td>Washout</td>
<td>Washout</td>
<td>Washout</td>
<td>Washout</td>
</tr>
<tr>
<td>Washout</td>
<td>Persist</td>
<td>Washout</td>
<td>Persist</td>
</tr>
<tr>
<td>Persist</td>
<td>Washout</td>
<td>Persist</td>
<td>Induced</td>
</tr>
<tr>
<td>Persist</td>
<td>Persist</td>
<td>Persist</td>
<td>Persist</td>
</tr>
</tbody>
</table>

Table 3.1: Table of stable steady states in isolation and within the two-pipe, periodic network. Persisting of pathogens in either pipe in isolation, leads to persisting of pathogens within the system.

1. If the trivial steady state is linearly stable in both connections in isolation Figure 3.5a,
then all Floquet exponents for the network have negative real parts and the zero steady state is linearly stable (i.e. the pathogen will wash out of network), which is shown in Figure 3.5b.

2. If the trivial steady state is linearly stable in the first connection and linearly unstable in the second one in isolation as shown in Figure 3.6a, then there is one Floquet exponent that has a positive real part for the network with an eigenvector with non-zero entries that correspond to the second connection only. This refers to the pathogens being washed out of the first connection and persisting in the second one, as shown in Figure 3.6b.

3. If the trivial steady state is linearly unstable in the first connection and linearly stable in the second one in isolation, as shown in Figure 3.7a, then there is one Floquet exponent with positive real part for the network with non-zero values in the eigenvector that correspond to both the first and second connections. This refers to the pathogen persisting in the first connection and the second connection being induced to have a non-zero pathogen concentration by the first connection, as shown in Figure 3.7b.

4. If the trivial steady state is linearly unstable in both connections in isolation, as shown in Figure 3.7b, then there are two Floquet exponent that have positive real parts with one eigenvector have non-zero entries for the second connection only and the other eigenvector what has non-zero entries both the first and second connections. This refers to the pathogens persisting in both connections, which is shown in Figure 3.7b.

**Medium-Sized Network with Periodic Flow Rates**

In this example we investigate the stability of a periodic water distribution network consisting of $M = 16$ junctions and $N = 19$ connections. Initially, we introduce pathogens into the bulk fluid within all connection, but no pathogens have been introduced within the biofilms.
Figure 3.5: Numerical solution for the pathogen concentration within the biofilm for both connections when they are in isolation, (a), and when in the network, (b). In both connections, the analysis predicts that the pathogens will washout of the system for both in-isolation and in-network cases. This is validated by the numerical solution. The biofilm parameters are $\alpha_1 = \alpha_2 = 9$ and $\beta_1 = \beta_2 = 9$ so $Q_c = 9/2$. The flow rates are of the form $Q_i = (b_i - a_i) H_\varepsilon \sin(\omega t) + a_i$ where $a_1 = 5$, $a_2 = 4$, $b_1 = b_2 = 6$, and $\omega = 1$. 
Figure 3.6: Numerical solution for the pathogen concentration within the biofilm for both connections when they are in isolation, (a), and when in the network, (b). The analysis predicts that the pathogens will washout of the first connection and will persist in the second connection for both the in-isolation and in-network cases. This is validated by the numerical solution. The biofilm parameters are $\alpha_1 = \alpha_2 = 9$ and $\beta_1 = \beta_2 = 9$ so $Q_c = 9/2$. The flow rates are of the form $Q_i = (b_i - a_i) H \epsilon (\sin(\omega t)) + a_i$ where $a_1 = 4$, $a_2 = 3$, $b_1 = b_2 = 6$, and $\omega = 1$. 


Figure 3.7: Numerical solution for the pathogen concentration within the biofilm for both connections when they are in isolation, (a), and when in the network, (b). The analysis predicts that the pathogens will persist in both connections for both the in-isolation and in-network cases. This is validated by the numerical solution. The biofilm parameters are $\alpha_1 = \alpha_2 = 9$ and $\beta_1 = \beta_2 = 9$ so $Q_c = 9/2$. The flow rates are of the form $Q_i = (b_i - a_i) H_\varepsilon (\sin (\omega t)) + a_i$ where $a_1 = 3$, $a_2 = 2$, $b_1 = b_2 = 6$, and $\omega = 1$. 

\[
Q_c = \frac{9}{2}
\]
Figure 3.8: Numerical solution for the pathogen concentration within the biofilm for both connections when they are in isolation, (a), and when in the network, (b). The analysis predicts that the pathogens will persist in the first connection for both the in-isolation and in-network cases. The second connection is predicted to washout when in isolation but when it is in the network, the analysis predicts that it will have a periodic solution with a non-zero steady state amplitude due to the first connection. This is validated by the numerical solution. The biofilm parameters are $\alpha_1 = \alpha_2 = 9$ and $\beta_1 = \beta_2 = 9$ so $Q_c = 9/2$. The flow rates are of the form $Q_i = (b_i - a_i) H_\varepsilon (\sin (\omega t)) + a_i$ where $a_1 = 3$, $a_2 = 4$, $b_1 = b_2 = 6$, and $\omega = 1$. 

74
The fluid flow is driven periodically in time with period $T = \pi/3$. The parameters governing the biofilm and fluid dynamics are given Appendix C.

The flow within the network is computed using the method described in Section 3.2.2 and is shown in Figure 3.9. The equations governing the pathogen dynamics in the network are given by 3.16. The system is linearized about the trivial steady state and we compute the Floquet exponents assuming each connection is in isolation subject to the flow within the network. We then apply the methodology described in 3.2.4 to determine the linear stability of the trivial steady state. The stability results are shown in Figure 2.14 with the dashed green on black lines indicating the pathogens are washed out of that connection, the solid red lines indicating the pathogens persist in that connection, and the dotted yellow on black lines indicating that the connection is “induced”, meaning that the pathogens would be washed out in isolation but there is a non-zero pathogen concentration due to pathogens persisting in an upstream connection. In this example, connections 7, 10, 13, and 16 are “persist” and cause six other connections (8, 14, 15, 17, 18, 19) to be “induced”. This
suggests that if pathogens were removed from connections 7, 10, 13, and 16, the pathogens would be removed from the entire network.

To validate the results from the analysis, we numerically compute the pathogen concentration within the network using the full nonlinear system. Figure 3.11 show the pathogen concentration within the biofilm once the solution settles down to a periodic solution with a steady amplitude at the end of the period, which corresponds to the maximum concentration over the period.

**Large Network with Periodic Flow Rates**

We present here an example to illustrate how our methodology scales to a larger water distribution network with a unidirectional periodic flow. Consider a water distribution network with $M = 134$ junctions and $N = 167$ connections. The network parameters given in Appendix D. The topology of the network and the flow through the network is shown in Figure 3.11.
Figure 3.11: Steady state numerical solution for the medium network example. The lines indicate the connections (pipes) and the circles indicate the junctions with the circled numbers indicating the index for each connection and the uncircled numbers the index for each junction. The triangles indicate the direction of flow. The dashed green on black lines denote connections where the pathogens have washed out of the connection. The solid yellow, orange, and red lines indicate the connection where the pathogens persist and the dotted black on yellow lines indicating connections have a non-zero pathogen concentration due to be induced by an upstream connection. The colors (yellow, orange, and red) indicate the relative concentration of pathogens within the biofilm relative to the maximum pathogen concentration within the network with red indicating the highest concentration, orange intermediate concentration, and yellow lower concentration. Note that some “induced” connections have a pathogen concentration of the same approximate size as some “persist” connections, suggesting that relative pathogen concentration size does not correspond to the classification of the connection.
3.12.

Figure 3.12: Topology and flow for the large network example with an unidirectional periodic flow. The lines indicate connections (pipes) and the circles indicate the junctions. The triangles indicate the direction of the flow through the connections. The numbers indicate the index assigned to each junction with each connection indexed sequentially starting for lower junctions to larger junctions.

Figure 3.13 shows the stability of the trivial steady state for each connection using the algorithm discussed in Section 3.2.4. The dashed green on black lines indicate the connections where the pathogens washout. The red connections indicate the connections where the pathogens persist and the yellow dotted lines on black denote connections that are induced by upstream connections to have a non-zero pathogen concentration as $t \to \infty$. We can observe that there are 55 connections where pathogens persist that induce another 38 connections to have non-zero pathogen concentrations. If the pathogens were removed from the 55 persistent connections and no pathogens coming in from up streams, the other 38 connections would be cleaned up due to the fluid flow. This observation could reduce the cleanup costs from the pathogen introduction into a water distribution network and illustrates the possible vulnerable points within the network.
Figure 3.13: Linear stability of the trivial steady state for the large network example with an unidirectional flow. The lines indicate the connections (pipes) and the circles indicate the junctions, which are numbered. The triangles indicate the direction of flow. The green dashed lines on black denote connections where the pathogens will washout. The solid red lines indicate the connection where the pathogens persist and dotted yellow lines on black indicate the connections where the pathogens are induced by upstream connections.
Figure 3.14: Steady state numerical solution for the large network example with a unidirectional flow. The lines indicate the connections (pipes) and the circles indicate the junctions, which are numbered. The triangles indicate the direction of flow. The dashed black-on-green lines denote connections where the pathogens will washout. Solid yellow, orange, or red lines indicate that the pathogens persist within that connection both in the network and in isolation. The dotted black-on-yellow/orange lines indicate the connections were induced by an upstream connection and have a non-zero pathogen concentration. The color indicates the relative pathogen concentration as with the solid lines. The yellow, orange, and red colors indicate relatively low, intermediate, and high pathogen concentrations, respectively.
To confirm the long-time behavior predicted by the linear stability analysis, we numerically simulated the entire network until a steady state was reached. Figure 3.14 shows the steady state pathogen concentration within the network.

The connections with high relative pathogen concentration are denoted by red lines. The orange lines refer to connections with medium relative pathogen concentration while the yellow line indicate that a pipe has low relative pathogen concentration. The solid lines indicate that the pathogens would persist in that connection in isolation while the dotted black-on-yellow/orange lines indicated that the pathogens would wash out of the connection in isolation but are induced by an upstream connection. Finally, the dashed black-on-green line used for connections where the pathogens washout. We can see that Figures 3.13 and 3.14 are equivalent, indicating the algorithm described in Section 3.2.4 is able to predict the steady state behavior of the network, except for the connections between junctions 48 and 49, 49 and 41, 127 and 126, 127 and 132. In the stability analysis by Floquet Theory, we observe that pathogens would persist in these connections in isolation and are labeled red, but the numerical solution shows these connections in yellow. This means that the pathogen concentration relative to the other connections within the network is low for these connections. The stability result by Floquet Theory indicates that for these connections, the pathogen persists but at a lower concentration that other connections shown in red.

3.2.6 Discussion and Future Direction

In this chapter, we presented three examples of water distributions networks with time-periodic flows. The algorithm described in (3.2.4) has been used to determine the persistence of pathogens within the three networks. The first example network consisted of only two pipes. The stability of the trivial steady states where computed using Floquet theory for the entire network as well as by the proposed algorithm. The results are equivalent and are validated by numerical simulations. In the second example, we examined a medium-sized network consisting of seventeen connections. Again, the results from numerical simulations
validate the accuracy of the proposed algorithm. In the last example, we studied the dynamic
of a pathogen release within a larger network with 167 connections. The algorithm scales
efficiently as the size of the network increases. Numerical simulations again validate the
predictions of the proposed algorithm.

In addition, Theorem 3.2.3 indicates that the connections that would have pathogens per-
sist if the connection was isolated from the rest of the network are the critical connections
that affect the pathogen concentration in the entire network. Descendants of these critical
connections are predicted to have non-zero pathogen concentrations, possibly masking the
most vulnerable parts of the network. It suggests that when disinfecting the network, par-
ticular focus should be make on the critical connections since the pathogens are predicted
to washout of the entire network if they are removed from the “persistent” connections.

The results from this chapter and Chapter 2 only apply under time-constant and time-
periodic flows where the flow direction does not change within a connection. For realistic
water distributions networks, the flow regime can be more complicated. Future directions for
this research is to develop theory and algorithms that take into account more realistic flows.
Two specific flow regimes should be able amenable to the current approach: subnetworks
that contains connections where the flow direction changes in time but the flows in and
out the subnetwork does not and time-aperiodic flows for the entire the network but each
connection has a time-periodic flow. For the former, one could examine the subnetwork
separately from the rest of the network using Floquet theory and then treat the rest of the
network using the approach in this thesis. For the latter, it might be possible to use the
proposed Floquet theory algorithm in Section 3.2.4 to compute the stability when Floquet
theory would not be applicable for the entire network. Both of these approaches need further
examination in the future.
References


K. N. Power and L. A. Nagy. Relationship between bacterial regrowth and some physical


Appendices
Appendix A

Parameters for Medium-Sized Network Example with Time-Constant Flow

The medium network examples consists of $M = 11$ junctions and $N = 11$ connections as shown in Figure A.1. At the interior junctions, we assign a demand, which is zero. At each non-interior junction, we prescribe a pressure. For this example, we let $P_1 = 29$, $P_8 = 2$, and $P_{11} = 9$. All the hydraulic resistances are set to one. We assume the parameters governing the pathogen dynamics are the same for each connection with $r = 1$, and $K = 1$ with $V = 1$ and $S_A = 1$ except for $\alpha$ and $\beta$, which are given in Table A.1.
Table A.1: Attachment and detachment rates for each connection for the time-constant medium-sized network example.

<table>
<thead>
<tr>
<th>i</th>
<th>$\alpha_i$</th>
<th>$\beta_i$</th>
<th>i</th>
<th>$\alpha_i$</th>
<th>$\beta$</th>
<th>i</th>
<th>$\alpha_i$</th>
<th>$\beta_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>6.0</td>
<td>5</td>
<td>0.01</td>
<td>0.8</td>
<td>9</td>
<td>5.00</td>
<td>5.0</td>
</tr>
<tr>
<td>2</td>
<td>0.01</td>
<td>11.0</td>
<td>6</td>
<td>1.00</td>
<td>0.3</td>
<td>10</td>
<td>1.00</td>
<td>2.0</td>
</tr>
<tr>
<td>3</td>
<td>1.40</td>
<td>9.0</td>
<td>7</td>
<td>1.00</td>
<td>2.0</td>
<td>11</td>
<td>1.00</td>
<td>2.0</td>
</tr>
<tr>
<td>4</td>
<td>0.01</td>
<td>0.2</td>
<td>8</td>
<td>1.00</td>
<td>2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Appendix B

Parameters for Large Network Example with Time-Constant Flow

The large network examples consists of $M = 134$ junctions and $N = 167$ connections as shown in Figure B.1.

At the interior junctions, we assign a demand, which is zero except for the ones given in Table B.1. At each non-interior junction, we prescribe a pressure, which is given in Table B.2. All the hydraulic resistances are set to one. We assume the parameters governing the
pathogen dynamics are the same for each connection and are $\alpha = 1$, $\beta = 2$, $r = 1$, and $K = 1$ with $V = 1$ and $S_A = 1$. This gives a critical flow rate of $Q_c = 1$.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$D_i$</th>
<th>$i$</th>
<th>$D_i$</th>
<th>$i$</th>
<th>$D_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>20</td>
<td>30</td>
<td>66</td>
<td>34</td>
<td>29</td>
</tr>
<tr>
<td>88</td>
<td>99</td>
<td>90</td>
<td>78</td>
<td>107</td>
<td>44</td>
</tr>
<tr>
<td>122</td>
<td>18</td>
<td>124</td>
<td>67</td>
<td>125</td>
<td>56</td>
</tr>
<tr>
<td>43</td>
<td>66</td>
<td>117</td>
<td>90</td>
<td>132</td>
<td>44</td>
</tr>
<tr>
<td>86</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.1: Demands, $D_i$, at the interior junctions in the large network example. Interior junctions without an entry have a zero demand and non-interior junctions are prescribed a pressure, which is given in Table B.2.

<table>
<thead>
<tr>
<th>$i$</th>
<th>$P_i$</th>
<th>$i$</th>
<th>$P_i$</th>
<th>$i$</th>
<th>$P_i$</th>
<th>$i$</th>
<th>$P_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>25</td>
<td>0</td>
<td>60</td>
<td>1</td>
<td>66</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>32</td>
<td>0</td>
<td>61</td>
<td>1</td>
<td>91</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>42</td>
<td>1</td>
<td>63</td>
<td>0</td>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>64</td>
<td>0</td>
<td>123</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>53</td>
<td>0</td>
<td>65</td>
<td>0</td>
<td>133</td>
<td>1</td>
</tr>
</tbody>
</table>

Table B.2: Pressures at the non-interior junctions in the large network example. Interior junctions do not have a prescribed pressure but have demands, which are given in Table B.1.
Appendix C

Parameters for Medium-Sized Network Example with Periodic Flow

The medium periodic network example with time-periodic flow consists of $M = 16$ junctions and $N = 19$ connections as shown in Figure C.1. At the interior junctions, we assign a zero demand throughout. At each non-interior junction, we prescribe a pressure, which is a constant zero at junctions 12, 15 and 16. At junction 1, we prescribe a periodic pressure

$$p_1(t) = (b - a) H_\varepsilon(\sin(\omega)) + a,$$  \hspace{1cm} (C.1)

where $a_1 = 1$, $b_1 = 50$, $\omega = 6$, and $H_\varepsilon(t)$ is the smoothed Heaviside function defined by (3.3). All the hydraulic resistances are set to one. We assume the parameters governing the

Figure C.1: Connectivity of the medium-sized network example with time-periodic flow.
pathogen dynamics are the same for each connection: $\alpha = 1$, $\beta = 2$, $r = 1$, and $K = 1$ with $V = 1$ and $S_A = 1$. 
Appendix D

Parameters for Large Network Example with Time-Periodic Flow

The large periodic network examples has the same topology as the large time-constant flow example described in Appendix B. At the interior junctions, we assign a demand, which is zero except for the following: \(D_{29} = D_{34} = 0.1\) and \(D_{101} = D_{126} = 1\). At each non-interior junction, we prescribe a pressure, which is given in Table D.1. All the hydraulic resistances are set to one. We assume the parameters governing the pathogen dynamics are the same for each connection and are \(\alpha = 1/2\), \(\beta = 2\), \(r = 1\), and \(K = 1\) with \(V = 1\) and \(S_A = 1\). This gives a critical flow rate of \(Q_c = 1/2\).

<table>
<thead>
<tr>
<th>(i)</th>
<th>(P_i)</th>
<th>(i)</th>
<th>(P_i)</th>
<th>(i)</th>
<th>(P_i)</th>
<th>(i)</th>
<th>(P_i)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(p_1(t))</td>
<td>25</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>66</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>32</td>
<td>0</td>
<td>61</td>
<td>0</td>
<td>91</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>42</td>
<td>0</td>
<td>63</td>
<td>15</td>
<td>110</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>64</td>
<td>15</td>
<td>123</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>53</td>
<td>9</td>
<td>65</td>
<td>15</td>
<td>133</td>
<td>(p_{133}(t))</td>
</tr>
</tbody>
</table>

Table D.1: Pressures at the non-interior junctions in the large network example with time-periodic flow. There are two junctions with pressures that are periodic in time: \(p_1(t) = (b_1 - a_1) H_\varepsilon (\sin(\omega t)) + a_1\) and \(p_{133}(t) = (b_{133} - a_{133}) H_\varepsilon (\sin(\omega t)) + a_{133}\). Here, \(a_1 = 50, b_1 = 100, \omega = 1, a_{133} = 20, b_{133} = 35, \omega_{133} = 3\) and \(H_\varepsilon (t)\) is the smoothed Heaviside function defined by (3.3).
Appendix E

Proof for Eigenvalues with Multiplicity Greater Than One

Here, we examine the case where the algebraic multiplicity of the eigenvalue is greater than one in the proof of Theorem 2.2.3. We start by assuming that there is an eigenvalue \( \lambda = \lambda_p \) for some integer \( p < N \) that has an algebraic multiplicity that is greater than one, implying that there is another eigenvalue \( \lambda_m \) for some integer \( m \) such that \( \lambda_m = \lambda_p \) with \( p < m < N \).

From the first row in the eigenvalue problem, we have \( D_{1,1}v_1 = 0 \). Since \( \lambda_p \neq \lambda_1 \), we have \( D_{1,1} \neq 0 \) so \( v_1 = 0 \). The second row in the eigenvalue problem, using \( v_1 = 0 \), is \( D_{2,2}v_2 = 0 \), which implies \( v_2 = 0 \) since \( D_{2,2} \neq 0 \). Assuming \( v_j = 0 \) for \( j = 1, 2, \ldots, i \), the \( i \)-th equation in the eigenvalue problem is

\[
\sum_{j=1}^{i-1} K_{i,j}v_j + D_{i,i}v_i = D_{i,i}v_i = 0 \quad (E.1)
\]

since \( v_j = 0 \) for \( j = 1, 2, \ldots, i - 1 \). Hence, \( v_i = 0 \) since \( D_{i,i} \neq 0 \) for \( i < p \). By induction, we obtain

\[
v_i = 0 \text{ for } i = 1, 2, \ldots, p - 1. \quad (E.2)
\]

We now examine \( p \)-th row of the eigenvalue problem. Since \( v_j = 0 \) for \( j = 1, 2, \ldots, p - 1 \), (E.1) reduces to \( D_{p,p}v_p = 0 \). Here, \( D_{p,p} = 0 \) since \( \lambda = \lambda_p \) so \( v_p \) is an arbitrary constant,
which is assume to be \( v_p = c_p \).

For the row \( p + 1 \), we can solve (E.1) for \( v_{p+1} \), using (E.2) and \( v_p = c_p \), to obtain

\[
v_{p+1} = -\frac{K_{p+1,p}}{D_{p+1,p+1}} c_p = a_{p+1} c_p \tag{E.3}
\]

since \( D_{p+1,p+1} \neq 0 \). If \( p \in \mathcal{P}(p + 1) \), then \( K_{p+1,p} \neq 0 \) and \( v_{p+1} \neq 0 \). If \( p \not\in \mathcal{P}(p + 1) \), then \( K_{p+1,p} = 0 \) and \( v_{p+1} = 0 \). For the row \( p + 2 \), we obtain

\[
v_{p+2} = -\frac{1}{D_{p+2,p+2}} \left( K_{p+2,p} v_p + K_{p+2,p+1} v_{p+1} \right) = -\frac{1}{D_{p+2,p+2}} \left( K_{p+2,p} - K_{p+2,p+1} \right) \frac{K_{p+1,p}}{D_{p+1,p+1}} c_p = a_{p+2} c_p. \tag{E.4}
\]

If \( p \in \mathcal{P}(p + 2) \) or both \( p + 1 \in \mathcal{P}(p + 2) \) and \( p \in \mathcal{P}(p + 1) \), then \( v_{p+2} \neq 0 \). These conditions reduce to \( p \in \mathcal{A}(p + 2) \) as the necessary condition for \( v_{p+2} \neq 0 \).

We assume \( v_k = 0 \) if \( p \not\in \mathcal{A}(k) \) and examine the \( i \)-th row of the eigenvalue problem where \( i = p + 3, p + 4, \ldots, m - 1 \). We solve (E.1) for \( v_i \), noting that \( D_{i,i} \neq 0 \). This yields

\[
v_i = -\frac{1}{D_{i,i}} \sum_{j=1}^{i-1} K_{i,j} v_j \tag{E.5}
\]

\[
= -\frac{1}{D_{i,i}} \left( K_{i,p} v_p + \sum_{j=p+1}^{i-1} K_{i,j} v_j \right) \text{ from (E.2).} \tag{E.6}
\]

\[
= -\frac{1}{D_{i,i}} \left( K_{i,p} c_p + \sum_{j=p+1}^{i-1} K_{i,j} v_j \right) \text{ since } K_{i,j} = 0 \text{ if } j \not\in \mathcal{P}(i) . \tag{E.7}
\]

If \( p \not\in \mathcal{A}(j) \) for some \( j = p + 1, \ldots, i - 1 \) where \( j \in \mathcal{P}(i) \), then \( v_j = 0 \) and the sum is zero in (E.7). If \( p \not\in \mathcal{P}(i) \), then \( K_{i,p} = 0 \). Hence, if \( p \not\in \mathcal{A}(i) \) then \( v_i = 0 \). Likewise, if \( p \in \mathcal{A}(i) \) then on or more terms in the parenthesis of (E.7) is non-zero so \( v_i \neq 0 \). So by induction,

\[
v_i \neq 0 \text{ if } p \in \mathcal{A}(i) \text{ for } i = p + 1, p + 2, \ldots, m - 1. \tag{E.9}
\]
For $i = m$, we have $D_{m,m} = 0$ so the $m$-th equation, using (E.2) and (E.9), is

$$
\left( K_{m,p} + \sum_{j=p+1}^{i-1} K_{i,j}a_j \right) c_p = 0 \quad \text{(E.10)}
$$

If $p \notin A(m)$, then (E.10) is automatically satisfied and we have a second free constant in the eigenvector, $v_m = c_m$. This allows us to generate two linearly independent eigenvectors. This results is scales with the algebraic multiplicity of the eigenvalue if none of the corresponding connections are ancestors/descendants of each other. This allows us to deal with eigenvalues of geometric multiplicity one by treating the corresponding connections as parents of distinct subnetworks, on which we can apply the proof for simple eigenvalues or for the following proof for eigenvalues with a geometric multiplicity greater than one.

If $p \in A(m)$, then (E.10) is satisfied by setting $c_p = 0$ and we are not able to generate two linearly independent eigenvectors. For simplicity, we will now assume that the algebraic multiplicity of the eigenvalue is two and $p \in A(m)$. The proof for higher algebraic multiplicity is similar.

We complete the computation of the eigenvector by noting that, for $i = m + 1, \ldots, N$, we have

$$
v_i = -\frac{c_m}{D_{i,i}} \left( K_{i,m} + \sum_{j=p+1}^{i-1} K_{i,j}a_j \right) \quad \text{(E.11)}
$$

If $m \notin A(i)$, then $K_{i,m}$ and the sum are both zero so $v_i = 0$. If $m \in A(i)$, then $v_i \neq 0$, which is the necessary condition for a non-zero element of the eigenvalue. This completes the proof for the eigenvector corresponding to $\lambda_m$.

To determine the form of the eigenvalue that corresponds to $\lambda_p$, we need to compute the
set of generalized eigenvectors. We start with the eigenvector that corresponds to $\lambda_m$:

$$v_i = \begin{cases} 
0 & \text{for } 0 < i < m \\
1 & \text{for } i = m \\
0 & \text{for } m < i \leq N, m \notin A(i) \\
a_i & \text{for } m < i \leq N, m \in A(i) 
\end{cases}$$

(E.12)

To compute the generalized eigenvector for $\lambda_p$, we compute the iteration $(K + D) v^* = v$ where $v$ is the eigenvector corresponding to $\lambda_m$ and $v^*$ is the eigenvector that corresponds to $\lambda_p$.

For $i = 1, 2, \ldots, p - 1$, using the same arguments as before, we obtain $v^*_i = 0$. For $i = p$, we have $D_{p,p} v^*_p = 0$ and since $D_{p,p} = 0$, we have $v^*_p = c_p^*$. For $i = p + 1, p + 2, \ldots, m - 1$, we have

$$K_{i,p} v^*_p + \sum_{j=p+1}^{i-1} K_{i,j} v^*_j + D_{i,i} v^*_i = 0$$

(E.13)

Using the same arguments as before, $p \in A(i)$ is the necessary condition for $v^*_i \neq 0$ for $i = p + 1, p + 2, \ldots, m - 1$. When $i = m$, we have

$$K_{m,p} c_p^* + \sum_{j=p+1}^{i-1} K_{i,j} v^*_j = 1.$$  

(E.14)

For each $j$ where $p \in A(j)$, we have $v^*_j = a_j c_p^*$ and we can solve for a non-zero $c_p^*$:

$$c_p^* = \left( K_{m,p} + \sum_{j=p+1}^{i-1} K_{i,j} a_j \right)^{-1}.$$  

(E.15)

We also have $v^*_m = c_m^*$. For $i = m + 1, m + 1, \ldots, N$, we have $v^*_i \neq 0$ if $p \in A(i)$ by using the same arguments as before. Hence, the non-zero entries of the second eigenvector are all...
descendants of the $p$-th connection, which completes the proof.