The Human Immune System: A Challenging Control Problem

by

Julie Racquel Vale

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2004

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This work deals with the control of the human immune system. A standard immune system model is modified by introducing control signals corresponding to drug cocktail and immune suppressor treatments. The ultimate objective is to use these control signals to 'cure' a chronically-ill patient. Control is challenging for this system due to nonlinearities and time delays. In fact, it is shown that fundamental aspects of the system dynamics are lost when the system is linearised; hence, control approaches involving linearisation are fruitless. Feedback linearisation and some optimal control methods are also investigated and shown to be infeasible. However, it is shown that, for certain parameter values and initial conditions related to the virus and patient, a specific open-loop control scheme using only the drug cocktail achieves the objective. It is also proven that, unfortunately, this control scheme fails for other parameter values and initial conditions. A two-stage open-loop controller that uses both control inputs is then proposed. It is shown in simulation that the two-stage controller works over a larger set of parameters and initial conditions than the single-stage controller, but a rigorous analysis of the two-stage controller remains elusive.

Acknowledgments

I would like to thank my supervisor Professor Daniel Davison, without whom this research would not have been possible. Also, to Professors Daniel Miller and Andrew Huenis for their patience and unending willingness to help. My heartfelt gratitude goes to Leslie Toews for putting up with me knocking on her door every 20 minutes to ask obtuse questions and for her ability to magically ask the right question, even though she usually had no idea what I was talking about. I would also like to acknowledge Tan Wang, for his assistance in understanding the biology of the immune system. Finally, thanks to my husband Adrian Ennis and to my parents for their continued support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The motivation for this masters thesis came from the field of immunology. Immunology is the study of the immune system, which is a highly complex collection of processes whose main purpose is to remove or destroy hostile substances, such as harmful bacteria and viruses, from the body. By ignoring factors such as swelling and temperature increase, the fundamental cell based process can be examined, specifically the humoral and cell-mediated immune responses. These processes are both part of a system that the layman thinks of as white blood cells.

For the purposes of this thesis only the humoral response will be investigated. It suffices to explain this response in the following way: A hostile virus or bacteria, henceforth referred to as the antigen or virus (V), causes the system to increase the rate of production of antigen specific antibodies (F) from a source of plasma cells (P). These antibodies then destroy the antigen. However, a certain number of antibodies are 'used up' for each viral cell that is killed, so antibody production must be ongoing. It is possible that healthy cells can be damaged or destroyed by the antigen or, as collateral damage, by the antibody. This damage can affect the ability of the immune system to function, which is accounted for through a concept called mass damage (m). Once the virus has been eliminated from the body, the immune system can recover and return to normal operation.

The variables V, P, and F have units of moles/cm^3, while m is dimensionless. It is important to note that, while the equations that will be used in this thesis are written as continuous functions of concentration, cells are discrete entities; therefore, due to quantisation error, the models break down for small concentrations.

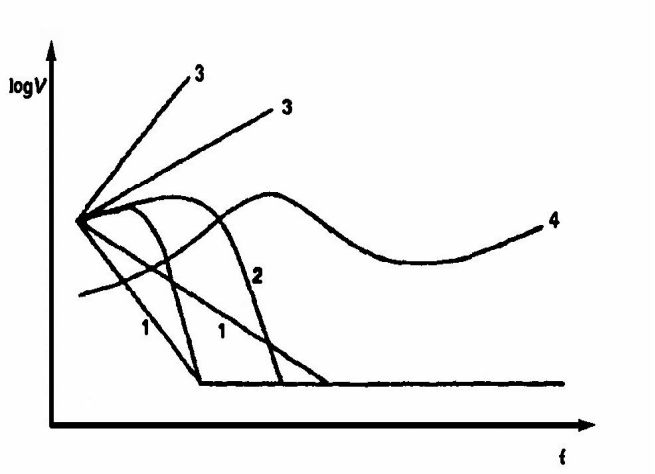

Figure 1.1: Possible forms of disease: 1-Subclinical; 2-Acute; 3-Lethal; 4-Chronic [1]

Another difficulty that appears at low concentrations, from the point of view of physical application, is that there is a lower limit to the concentration that medical equipment can detect.

Throughout this thesis, two pairs of terms will be used that are somewhat similar but that have very different meanings; therefore, we will clearly define them now to avoid confusion in the future. To discuss one of these pairs, a threshold value, V^*, must first be defined to be the larger of the following two concentrations: the level below which the model breaks down due to quantization error, and the minimum measurable concentration. Below V^* we are unsure of the physical system's true behaviour; indeed, for model concentrations below V^*, the actual physical viral concentration may in fact be zero! Using the notion of V^*, we can now define the terms 'healthy' and 'sick', which refer to viral levels. If, at a given time t, $V(t) \leq V^*$, then the person is healthy at that time instant; otherwise, they are referred to as being sick. 'Dead' and 'alive' refer to damage. A patient is considered dead if, at any time during the course of treatment, $m(t) \geq 1$. The patient is considered to be alive if, for the entire duration of the treatment, $m(t) < 1$. So, it is possible for a patient to be referred to as healthy but dead! The important distinction to make here is that the terms alive and dead do not refer to the entire patient, but rather to the concept of mass damage and the ability of the immune system to continue producing plasma cells. In this thesis death will be considered unacceptable.

There are four standard descriptions of a person's reaction to infection: subclin-

ical, acute, chronic, and lethal (see Figure 1.1). The particular behaviour exhibited by a patient depends on the initial conditions and the model parameters, determined by properties of the patient's immune system and the particular virus. The subclinical case is characterised by a constantly decreasing amount of virus in the body until the virus is eradicated, while the acute response is characterised by an initial rise in virus followed by swift elimination due to a strong immune response. Both of these cases result in a healthy patient. In the chronic case, the virus maintains a presence in the body that is balanced by the antibody and hence results in neither death nor health. In the lethal case, the virus maintains unlimited growth resulting in death and therefore an inability to combat the proliferation of viral cells [1]. Obviously, the lethal case is less desirable than the chronic case, which is less desirable than the acute case, which in turn is less desirable than the subclinical case.

The previous discussion indicates that an appropriate model will have at least one equilibrium point corresponding to the healthy case and one limit cycle or equilibrium point corresponding to the chronic case.

1.2 Objectives

In laymen's terms, the objective of this thesis is to take a patient that is chronically ill and make them healthy in finite time, t_f. Typically t_f is on the order of many months; however, it will vary according to the patient and the virus. To accomplish this, two objectives must be satisfied:

1. Find a mathematical model of the immune system that is suitable for control purposes, and
2. Find a controller that will take an uncontrolled system whose initial conditions lie within the region of convergence of a stable chronic equilibrium point and control it such that the trajectory described by the system under control ends with $V(t_f) \leq V^*$ without causing death in the process.

Inherent in Objective 2 is the assumption that the chronic equilibrium point is stable. Equilibrium points of nonlinear systems are either stable or unstable, possibly asymptotically. If they are unstable they may exhibit the properties of a saddle point, where, even though they move away from the equilibrium for most $x(t_0)$, there exists a non-empty 'stable submanifold' which is defined to be the set

of all $x(t_0)$ such that $x(t)$ with initial condition $x(t_0)$, approaches the equilibrium point asymptotically.

The ideal situation for the achievement of Objective 2 is to have the healthy equilibrium be asymptotically stable; in this case, the objective would be met by simply forcing the system into the region of convergence of that point, then removing the control signal. If this is not possible, then the next best situation would be to have the healthy equilibrium point exhibit the properties of a saddle; in this case, it would be sufficient to force the system to the stable submanifold. In this thesis, it will be shown that if the chronic equilibrium point is stable, then the healthy equilibrium point is an unstable saddle whose stable submanifold includes all $x(t_0)$ where $V(t_0) = 0$.

Recall that below V^* the physical value of the virus may be zero. We conclude from the preceding paragraph that, if the true physical value is indeed zero, then satisfying Objective 2 will have the added benefit of forcing the entire system to settle to the healthy equilibrium point. Unfortunately, if the true physical value is non-zero then, since the healthy point is unstable, the viral concentration will increase until, eventually, it is above V^*. At this point, additional control would be required; however, this situation is beyond the scope of this thesis.

It is assumed that all states of the chosen model are acessible.

1.3 Overview of Thesis

The thesis will begin by finding a model of the immune system and adjusting it such that it is suitable for control. Properties of the model, such as equilibrium points and stability, bounds, and existence of solutions will be investigated. A list of important terms and ideas can be found in Appendix B.

In Chapter 3, previous work by other researchers will be investigated. Furthermore, standard control methods such as linearisation and optimal control though Model Predictive Control will be investigated. These methods are shown to be either infeasible or problematic for this system.

Finally, in Chapter 4, a simple open-loop 'step' controller, using only one input, is investigated and shown to work under certain conditions, but fail under others. A two-stage controller is then proposed and shown to satisfy the thesis objectives for a wider range of model parameters and initial conditions.

This thesis is purely theoretical. Neither the physical realisation of the control scheme, nor the feasibility of such a realisation, will be discussed.

Chapter 2

The Model

2.1 Finding a Model

Two papers, [1] and [2], were found which provide sets of equations that model the humoral response of the human immune system. The equations in [1] describe the reaction to viruses, while those in [2] describe the reaction to bacteria. A thorough discussion of the physical interpretation of the equations and parameters for the chosen model will occur in Section 2.1.2; here we are only interested in the form of the equations. Given that the relationship between $\zeta(m)$ and m is as shown in Figure 2.1 and that V, P, F, and m are as defined in Chapter 1, the equations from [1] are:

$$\begin{aligned}
\dot{V}(t) &= (\beta - \gamma F(t))V(t), && (2.1)\\
\dot{P}(t) &= \zeta(m(t-\tau))\alpha V(t-\tau)F(t-\tau) - \mu_P(P(t) - P^*), && (2.2)\\
\dot{F}(t) &= \rho P(t) - (\mu_f + \eta\gamma V(t))F(t), && (2.3)\\
\dot{m}(t) &= \sigma V(t) - \mu_m m(t) + \sigma_f F(t). && (2.4)
\end{aligned}$$

The following set of equations, from [2], describe the interaction between the bacteria (B) and the antibody (A) and are based on a predator-prey model. Given

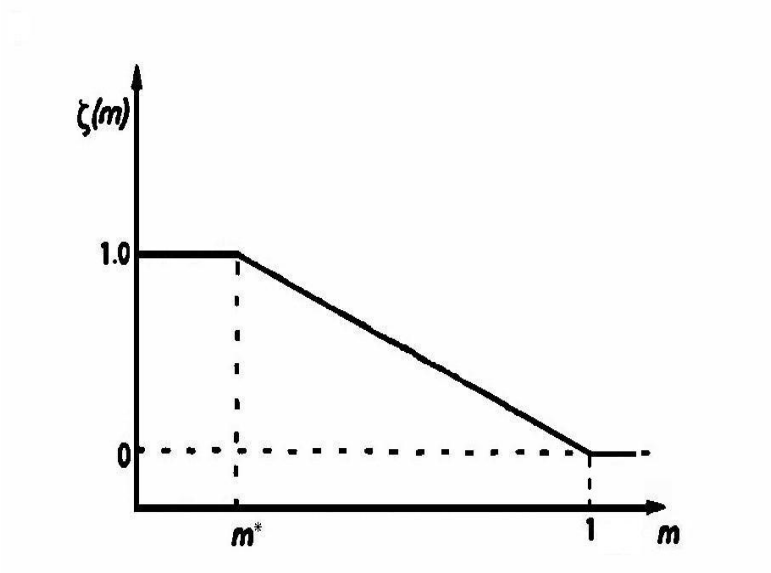

Figure 2.1: Relationship between mass damage and $\zeta(m)$ [1]

that $1^+(t)$ is the step function, the equations from [2] are:

$$\begin{aligned}
\dot{B}(t) &= a_1 - \omega \frac{(A(t) - A_{eq})B(t)}{d + B(t)/\eta + A(t) - A_{eq}} \\
&\quad -bB(t)^2 + \alpha B(t)\hat{u}(t), \qquad (2.5) \\
\dot{A}(t) &= \rho \frac{A(t-\tau)B(t-\tau)}{\eta(\gamma + A(t-\tau) + B(t-\tau)/\eta)} \left[1 - \frac{A(t)}{A^*}\right] \\
&\quad \times 1^+(t-\tau) - \omega \frac{(A(t) - A_{eq})B(t)}{\eta(d + B(t)/\eta + A(t) - A_{eq})} - a_2(A(t) - A_{eq}).
\end{aligned}$$

The equations in (2.1)-(2.4) are slightly more general than (2.5). Perhaps more importantly from a control standpoint, the nonlinearities in (2.1)-(2.4) are 'simpler' than those in (2.5). Due to these considerations, it was decided to use [1], and therefore (2.1)-(2.4), as the basis for this work. There exist further extensions to (2.1)-(2.4) in [3] that also take into account the cell mediated response; however, it was felt that it would be prudent to control the simpler model first, then move on

to the extended one. Henceforth, the system state will be denoted:

$$x(t) \triangleq \begin{bmatrix} V(t) \\ P(t) \\ F(t) \\ m(t) \end{bmatrix}. \tag{2.6}$$

2.1.1 Initial Conditions

It will be established in Section 2.4 that the system in (2.1)-(2.4) has a healthy equilibrium point, x_{0h}, corresponding to $V_{0h} = 0$. For the purposes of this work it will be assumed that the system has been at the healthy equilibrium point for an indefinite time and that, at $t = 0$, the patient encounters a specified amount of virus, in an impulse type interaction. Hence, the initial conditions of all variables will correspond to the healthy equilibrium point for the given set of parameters, with the exception of V, which will take on a value equivalent to the initial dose of germs. Furthermore, for all $t \in [-\tau, 0)$, $V(t) = 0$ and, if F_{0h} denotes the healthy equilibrium value of F, $F(t) = F_{0h}$. For the remainder of the thesis define $x_0 \triangleq x(0)$ to be the initial condition.

2.1.2 Physical Description of Model

In this section, a high-level description of each equation in (2.1)-(2.4) will be presented, followed by a list summarising descriptions of each parameter.

The first equation in the system, (2.1), describes the rate of growth of the antigen cells. These cells replicate exponentially and are destroyed only though encounters with antibody cells.

Plasma cell replication, modeled in (2.2) is influenced by two factors: prior antigen/antibody encounters and the current concentration of plasma cells. There is a probability, affected by the amount of damage sustained, that each antigen/antibody encounter will result in the creation of a new plasma cell. This creation will happen only after a delay, associated with the generation of memory cells and other dynamic components otherwise ignored in this model. The body will also create enough plasma cells to ensure that the concentration of cells remains above a threshold level, P^*. Plasma cells are not 'used up' in the creation of antibodies, nor are they killed by virus cells. Instead, they die off exponentially.

As indicated in (2.3), the only source of antibody cells is plasma. Plasma cells produce antibodies at a specific rate. For each antigen cell destroyed in an

antigen/antibody reaction, η antibody cells are 'used up' or destroyed. Antibodies are also lost exponentially due to natural attrition.

In contrast to (2.1)-(2.3), (2.4) reflects an attempt to model an idea, rather than specific cell interactions. Damage can be caused by the virus or, as collateral damage, by the antibody; in addition, the immune system is a self-healing mechanism. The easiest way to model these concepts is in a linear fashion, with both damage and repair occurring at exponential rates relative to the appropriate factor.

The following list describes each parameter individually:

- β - Inverse time constant of virus (or antigen)
- γ - Probability of an antigen/antibody encounter that results in antigen neutralization
- α - Probability of an antibody encountering a virus such that production of plasma cells is stimulated
- τ - Approximates delay associated with generation of memory cells and other dynamic components otherwise ignored in this model
- μ_p - Inverse of average life span of plasma cells
- P^* - Minimum number of plasma cells in the body
- ρ - Rate of production of antibodies due to one plasma cell
- μ_f - Inverse of average life span of antibody cells
- η - Number of antibody cells required to neutralize one antigen cell; this is a natural number
- σ - Rate of damage due to one viral cell
- μ_m - Inverse of recuperation period of m
- σ_f - Rate of damage due to one antibody cell.

2.2 Control Input Placement

The researchers who developed (2.1)-(2.4) did not address the issue of how a control input would be included. By consulting medical literature, it was found that there are two main types of drugs designed to assist the humoral immune system in fighting viruses. These drugs take the form of immune response inhibitors and drug cocktails.

The drug cocktail acts in a similar manner to the antibody, reducing the rate of replication of the viral cells. The biological mechanisms responsible for the interaction between the drug cocktail and the virus indicate that a naive way to

introduce the drug cocktail input, u, would be as follows:

$$\begin{aligned}
\dot{V}(t) &= (\beta - u(t) - \gamma F(t))V(t), \\
\dot{P}(t) &= \xi(m(t-\tau))\alpha V(t-\tau)F(t-\tau) - \mu_P(P(t) - P^*), \\
\dot{F}(t) &= \rho P(t) - (\mu_f + \eta\gamma V(t))F(t), \\
\dot{m}(t) &= \sigma V(t) - \mu_m m(t) + \sigma_f F(t).
\end{aligned} \tag{2.7}$$

This system has a trivial control scheme: simply choosing an input large enough will force the viral level below V^* within the desired time. The equations do not reflect any detrimental aspect to this treatment, so, assuming that the financial cost of the drug is of no consequence, our problem is solved. In reality, the drug cocktail can increase the damage. This is due to the fact that the drug cocktail works by killing infected cells, so healthy cells can be damaged or killed inadvertently; hence, the drug cocktail causes collateral damage. To account for this, a u term is introduced linearly in the damage equation.

$$\begin{aligned}
\dot{V}(t) &= (\beta - u(t) - \gamma F(t))V(t), \\
\dot{P}(t) &= \xi(m(t-\tau))\alpha V(t-\tau)F(t-\tau) - \mu_P(P(t) - P^*), \\
\dot{F}(t) &= \rho P(t) - (\mu_f + \eta\gamma V(t))F(t), \\
\dot{m}(t) &= \sigma V(t) - \mu_m m(t) + \sigma_f F(t) + \sigma_c u(t).
\end{aligned} \tag{2.8}$$

This choice of a linear operation is somewhat arbitrary, but is sufficient to capture the desired concept.

The immune inhibitor drug, v, does not stop existing antibodies from attacking and killing viral cells; it simply blocks the production of plasma cells. This indicates that a logical way to introduce v is as a multiplicitive input; therefore, when $v = 0$ it is 'on', actively stopping production of plasma due to antigen/antibody interactions. Conversely, when $v = 1$ it is 'off', since multiplication by 1 does not affect a product. Using v in this way, combined with u as discussed above, results in the final system to be controlled:

$$\dot{V}(t) = (\beta - u(t) - \gamma F(t))V(t), \tag{2.9}$$

$$\dot{P}(t) = \xi(m(t-\tau))\alpha V(t-\tau)F(t-\tau)v(t-\tau) - \mu_P(P(t) - P^*), \tag{2.10}$$

$$\dot{F}(t) = \rho P(t) - (\mu_f + \eta\gamma V(t))F(t), \tag{2.11}$$

$$\dot{m}(t) = \sigma V(t) - \mu_m m(t) + \sigma_f F(t) + \sigma_c u(t). \tag{2.12}$$

It is important to note that having both inputs active at the same time is counterproductive, as u's function is to decrease the viral levels, while v increases V while it is active. Indeed, this discussion also implies that, if v is used, it should be active before u.

2.3 Existence and Continuity

Since (2.9)-(2.12) are continuous in the state variables then a solution exists and is locally continuous [4].

2.4 Equilibrium Points and Stability

The controlled system (2.9)-(2.12) has two equilibrium points, healthy and chronic, whose stability can be affected through the choice of parameter values. Subsequently '0h' and '0c' will be used to denote the healthy and mild chronic equilibrium points, respectively.

The healthy equilibrium point is characterized by the following parameter-to-variable relationships:

$$\begin{aligned} V_{0h} &= 0, \\ F_{0h} &= \rho P^*/\mu_f, \\ P_{0h} &= P^*, \\ m_{0h} &= \frac{\sigma_f \rho P^*}{\mu_m \mu_f}. \end{aligned} \tag{2.13}$$

The chronic equilibrium point has two main forms corresponding to different values of m_{oc}. In the first case, which will be called the 'mild' case, $m_{oc} < m^*$, while in the 'severe' case, $m_{oc} \geq m^*$, where m^* is as shown in Figure 2.1. Only one of these cases will occur for a given choice of parameters. In particular, the mild case arises when $\mu_p \tau \leq 1$ [1]. This inequality indicates that there is a critical relationship between the delay in creation of plasma cells, τ, and the rate at which they die, μ_p. In this thesis we will only be concerned with the mild chronic case, whose equilibrium point is the following:

$$\begin{aligned} V_{0c} &= \frac{\mu_P[\mu_f(\beta - u_{0c}) - \gamma\rho P^*]}{(\beta - u_{0c})(\alpha v_{0c}\rho - \mu_P \eta\gamma)} > 0, \\ F_{0c} &= (\beta - u_{0c})/\gamma, \\ P_{0c} &= \frac{\alpha v_{0c}\mu_f(\beta - u_{0c}) - \eta\mu_P\gamma^2 P^*}{\gamma(\alpha v_{0c}\rho - \mu_P\eta\gamma)}, \\ m_{0c} &= \frac{\sigma V_{0c} + \sigma_f F_{0c} + \sigma_c u_{0c}}{\mu_m} < m^*. \end{aligned} \tag{2.14}$$

2.4.1 Linearisation

Perhaps the simplest way to gain insight into the qualitative properties of a non-linear system is to linearise it about an equilibrium point and to investigate the properties of that linearised system. Using Taylor series techniques, denoting x_e to be the equilibrium point, and defining $\Delta x \triangleq x - x_e$, the linearisation for (2.9)-(2.12) is:

$$\begin{aligned}
\Delta\dot{V}(t) &= (\beta - u_e - \gamma F_e)\Delta V(t) - V_e\Delta u(t) - \gamma V_e \Delta F(t), \\
\Delta\dot{P}(t) &= \alpha V_e v_e \Delta F(t-\tau) + \alpha F_e v_e \Delta V(t-\tau) + \\
&\quad \alpha F_e V_e \Delta v(t-\tau) - \mu_P \Delta P(t), \\
\Delta\dot{F}(t) &= \rho\Delta P(t) - (\mu_f + \eta\gamma V_e)\Delta F(t) - \eta\gamma F_e \Delta V(t), \\
\Delta\dot{m}(t) &= \sigma\Delta V(t) - \mu_m \Delta m(t) + \sigma_f \Delta F(t) + \sigma_c \Delta u(t).
\end{aligned} \tag{2.15}$$

In particular, linearising about the healthy equilibrium point results in:

$$\begin{aligned}
\Delta\dot{V}(t) &= (\beta - u_{0h} - \gamma\rho P^*/\mu_f)\Delta V(t), \\
\Delta\dot{P}(t) &= \alpha v_{oh}\rho P^*/\mu_f \Delta V(t-\tau) - \mu_P \Delta P(t), \\
\Delta\dot{F}(t) &= \rho\Delta P(t) - \mu_f \Delta F(t) - \eta\gamma\rho P^*/\mu_f \Delta V(t), \\
\Delta\dot{m}(t) &= \sigma\Delta V(t) - \mu_m \Delta m(t) + \sigma_f \Delta F(t) + \sigma_c \Delta u(t).
\end{aligned} \tag{2.16}$$

Similarly, linearising about the chronic equilibrium point results in:

$$\Delta\dot{V}(t) = \frac{\mu_P[\mu_f(\beta - u_{0c}) - \gamma\rho P^*]}{(\beta - u_{0c})(\alpha v_{0c}\rho - \mu_P\eta\gamma)}(-\Delta u - \gamma\Delta F(t)), \tag{2.17}$$

$$\begin{aligned}
\Delta\dot{P}(t) &= \alpha v_{0c}\frac{\mu_P[\mu_f(\beta - u_{0c}) - \gamma\rho P^*]}{(\beta - u_{0c})(\alpha v_{0c}\rho - \mu_P\eta\gamma)}\Delta F(t-\tau) \\
&\quad + \frac{\alpha\mu_P[\mu_f(\beta - u_{0c}) - \gamma\rho P^*]}{\gamma(\alpha v_{0c}\rho - \mu_P\eta\gamma)}\Delta v(t-\tau) \\
&\quad + \frac{\alpha v_{0c}(\beta - u_0)}{\gamma}\Delta V(t-\tau) - \mu_P\Delta P(t),
\end{aligned} \tag{2.18}$$

$$\begin{aligned}
\Delta\dot{F}(t) &= \rho\Delta P(t) - \left(\mu_f + \eta\gamma\frac{\mu_P(\mu_f(\beta - u_{0c}) - \gamma\rho P^*)}{(\beta - u_{0c})(\alpha v_{oc}\rho - \mu_P\eta\gamma)}\right)\Delta F(t) \\
&\quad - \frac{\eta\gamma(\beta - u_{0c})}{\gamma}\Delta V(t),
\end{aligned} \tag{2.19}$$

$$\Delta\dot{m}(t) = \sigma\Delta V(t) - \mu_m\Delta m(t) + \sigma_f\Delta F(t) + \sigma_c\Delta u(t). \tag{2.20}$$

2.4.2 Stability

This subsection will deal with the uncontrolled system; however, we would like to use the linearisation results determined above. Under the substitutions $\Delta u = \Delta v = 0$, $u_0 = 0$, and $v_0 = 1$, (2.16) and (2.17)-(2.20) correspond to the uncontrolled case. Therefore, under these substitutions, (2.16) and (2.17)-(2.20) can be used to evaluate the stability of the uncontrolled equilibrium points.

The healthy equilibrium point will be stable if and only if the rate of increase of the virus, β, is smaller than the equilibrium rate at which the antibody is able to destroy the virus, γF_{0h}:

Theorem 2.1 *The healthy equilibrium point is asymptotically stable iff* $\beta < \gamma F_{0h}$.

Proof

The characteristic polynomial of (2.16) under no control is:

$$det \begin{bmatrix} s-(\beta-\gamma F_{0h}) & 0 & 0 & 0 \\ -\alpha F_{0h} e^{s\tau} & s+\mu_p & 0 & 0 \\ \eta\gamma F_{0h} & -\rho & s+\mu_f & 0 \\ -\sigma & 0 & -\sigma_f & s+\mu_m \end{bmatrix},$$

which corresponds to:

$$(s-(\beta-\gamma F_{0h}))(s+\mu_p)(s+\mu_f)(s+\mu_m). \tag{2.21}$$

Recall that all parameters are positive; therefore, the only term of interest here is $(s-(\beta-\gamma F_{0h}))$, since all other roots are guaranteed to lie in the Open Left Half Plane (OLHP).

If $\beta < \gamma F_{0h}$ then $\beta - \gamma F_{0h} < 0$; consequently, all roots are in the OLHP, thus, the system is stable. Similarly, if the system is stable, then all roots of the characteristic polynomial must lie in the OLHP; hence $\beta - \gamma F_{0h} < 0$ and therefore $\beta < \gamma F_{0h}$. □

It is interesting to note that the time delay in (2.16) does not enter into the characteristic polynomial and hence does not affect the stability of the system.

In order to state the conditions for the stability of the chronic equilibrium point, some terms must first be defined:

$$a \triangleq \mu_p + \mu_f + \eta\gamma V_{0c}, \tag{2.22}$$
$$b \triangleq \mu_p(\mu_f + \eta\gamma V_{0c}) - \eta\gamma\beta V_{0c}, \tag{2.23}$$
$$d \triangleq \eta\gamma\beta\mu_p V_{0c}, \tag{2.24}$$
$$f \triangleq \alpha\beta\rho V_{0c}, \tag{2.25}$$
$$g \triangleq \alpha\rho V_{0c}. \tag{2.26}$$

From (2.22)-(2.26) and the fact that all parameters in (2.9)-(2.12) are positive, it follows that each of a, d, f, and g is positive, but no conclusions can be made about the sign of b.

A set of conservative sufficient conditions for the stability of the mild chronic equilibrium point is stated in the following theorem:

Theorem 2.2 *The mild chronic equilibrium point (2.14), where $\mu_p\tau \leq 1$, is stable if $0 < \frac{f-d}{a} < b - g - f\tau$ and $a - g\tau > 0$.*

Proof Please see Appendix A.1.

Finally, the chronic point is stable only if the combination of the creation rates of the plasma cells, α, and antibody cells, ρ, is larger than the combination of the rate of destruction of plasma cells due to natural attrition, μ_p, and the rate at which antibody cells are 'used up' to fight viruses, $\eta\gamma$. This is equivalent to the condition $f - d > 0$, which is the notation used in the following theorem:

Theorem 2.3 *If the mild chronic equilibrium point is stable then $f - d > 0$*

Proof Please see Appendix A.3.

2.4.3 Simultaneous Stability of Equilibrium Points

It is worthwhile to determine if both equilibrium points can be stable simultaneously. As discussed in Section 1.2, if this were true, a chronic patient could be 'cured' by using a control scheme that simply forces the system into the region of attraction of the healthy equilibrium point. Unfortunately, in this model, both equilibrium points cannot be stable at the same time:

Theorem 2.4 *It is impossible for both equilibrium points to be stable simultaneously.*

Proof

We will use proof by contradiction, so assume that both equilibrium points exist and are stable. This implies that:

1. $\beta < \gamma\rho P^*/\mu_f$ (see Theorem 2.1 and (2.13)) and
2. $f - d > 0$ (see Theorem 2.3).

Item 1 is equivalent to:

$$\mu_f\beta - \gamma\rho P^* < 0. \tag{2.27}$$

Investigating item 2:

$$\begin{array}{rrcl} & 0 & < & f - d, \\ \Leftrightarrow & 0 & < & \beta\alpha\rho V_{0c} - \eta\gamma\beta\mu_p V_{0c}, \\ \Leftrightarrow & 0 & < & \alpha\rho - \eta\gamma\mu_p. \end{array} \tag{2.28}$$

Now, recall from (2.14) that

$$V_{0c} = \frac{\mu_P[\mu_f(\beta) - \gamma\rho P^*]}{(\beta)(\alpha\rho - \mu_P\eta\gamma)} > 0. \tag{2.29}$$

Inequality (2.27) implies that the numerator of the expression in (2.29) is negative, and inequality (2.28) implies that the denominator of the expression in (2.29) is positive. Hence, V_{0c} is negative. This contradicts (2.29). □

So we have shown that if the chronic point is stable, then the healthy point must be unstable. It is worthwhile to investigate whether the healthy point is a saddle point and, if it is, whether $\{x(t_0) : V(t_0) = 0\}$ is on the stable submanifold.

2.4.4 System Behaviour when $V = 0$

Recall that in the discussion of the second objective in Section 1.2, it was stated that, if the chronic equilibrium point is stable then the healthy equilibrium is a saddle such that $\{x(t_0) : V(t_0) = 0\}$ is in the stable submanifold. We have already shown that the healthy equilibrium is unstable. The following theorem establishes that there is a stable submanifold that contains $\{x(t_0) : V(t_0) = 0\}$ and that the equilibrium point is a saddle (see Figure 2.2).

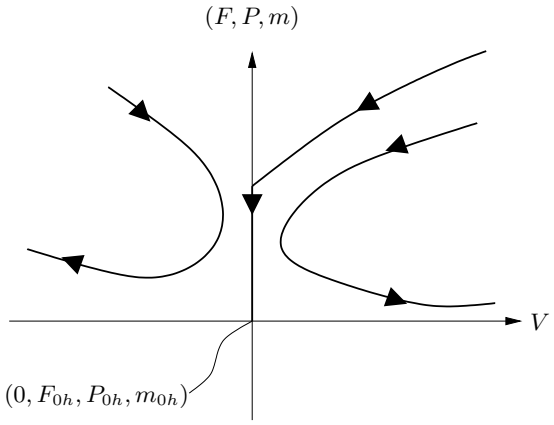

Figure 2.2: Example of a saddle point

Theorem 2.5 *In the uncontrolled system, (2.1)-(2.4), if* $V(t_f) = 0$ *then* $V(t) = 0$ *for all* $t \geq t_f$ *and the system asymptotically approaches the healthy equilibrium point (2.13).*

Proof Please see Appendix A.4

This theorem implies that F and P are guaranteed to tend to their healthy equilibrium values after control is released (assuming $V(t_f) = 0$ is achieved), and therefore justifies the appropriateness of the emphasis placed on V and m in this work.

2.5 Bounds and Finite Escape Time

To confirm that the equations reflect reality appropriately, bounds on variables were investigated. In order to ensure that an upper bound exists for each variable over a finite time interval, it is necessary to establish that there are no conditions that result in finite escape time. The proof regarding finite escape time requires lower bounds on the variables, so these bounds will be investigated first.

2.5.1 Lower Bounds

Begin with the basic nonlinear controlled equations (2.9)-(2.12) and the following assumptions:

1. All initial conditions are positive and $P_0 \geq P^*$.
2. Both $F(t)$ and $V(t)$ are positive for all $t \in [-\tau, 0)$.
3. All model parameters are positive.
4. The state variables are locally continuous (Section 2.3).

Then the following lower bounds hold:

Theorem 2.6 *All variables in the system (2.9)-(2.12) have a lower bound of* 0. *Furthermore,* $P(t)$ *has a lower bound of* P^*.

Proof Please see Appendix A.5.

Henceforth, all lower bounds will be denoted with a check, for example, $\check{x}$.

2.5.2 Finite Escape Time

To show that (2.9)-(2.12) does not exhibit finite escape time in our region of interest, we will show that each component of the state is finite for all time $t \in [0, \infty)$.

Theorem 2.7 *The signals* $V(t)$, $P(t)$, $F(t)$, *and* $m(t)$ *in (2.9)-(2.12) are finite for all time* $t \in [0, \infty)$.

Proof Please see Appendix A.6

It follows directly from Theorem 2.7 that the system (2.9)-(2.12) does not exhibit finite escape time.

2.5.3 Upper Bounds

Since the system does not exhibit finite escape time (Theorem 2.7) it follows that, over any finite time interval, there exists a maximum value for each variable. Henceforth, all upper bounds over the interval $t \in [0, t_f]$ will be denoted with a hat, for example, $\hat{x}$.

Table 2.1: Parameter Values for Healthy, Chronic, and Unstable Cases

Parameter	Chronic	Acute/Subclinical	Lethal
P^*	2×10^{-21}	2×10^{-21}	2×10^{-21}
ρ	2×10^{8}	2×10^{8}	2×10^{8}
μ_F	0.043	0.043	0.043
μ_P	0.32	0.4	0.33
η	220	220	1×10^{10}
γ	4.09091×10^{9}	4.09091×10^{10}	90
β	0.1	0.5	1
α	10^{10}	10^{10}	100
σ	1×10^{11}	1×10^{11}	1×10^{11}
σ_f	10	10	10
μ_m	1×10^{-4}	1×10^{-4}	0.01
τ	3	3	3

2.6 Simulation Setup

In order to simulate (2.9)-(2.12), numerical values for the parameters must be determined, and the equations must be normalised.

2.6.1 Parameter Value Selection

In [3] a range of acceptable values for the parameters of a more complex system were found. These parameters were mapped to the equivalent terms in the simplified equations where possible. In this thesis, the focus is on the acute and chronic cases; these two cases require that the healthy and mild chronic equilibrium points be stable, respectively. The parameters for the lethal case, which is of little interest in this work, can be determined by forcing both equilibrium points to be unstable. Parameters corresponding to the three cases are provided in Table 2.1.

2.6.2 Normalization

The variables associated with (2.9)-(2.12) vary drastically in size; therefore, to assist with the reduction of numerical errors in simulation, (2.9)-(2.12) was normalized using $\bar{x} = x/x_0$. The inhibitor input, v, and the damage, m, were not normalised

because the range of interest is already $[0, 1]$. Note that $P^* = P_0$ is always true. The normalized equations are:

$$\begin{aligned}
\dot{\bar{V}}(t) &= (\beta - u_0\bar{u}(t) - \gamma\bar{F}(t)F_0)\bar{V}(t), \\
\dot{\bar{P}}(t) &= \zeta(m(t-\tau))\frac{F_0V_0}{P_0}\alpha\bar{V}(t-\tau)\bar{F}(t-\tau)v(t-\tau) - \mu_P(\bar{P}(t)-1), \quad (2.30) \\
\dot{\bar{F}}(t) &= \frac{P_0}{F_0}\rho\bar{P}(t) - (\mu_f + \eta\gamma\bar{V}(t)V_0)\bar{F}(t), \\
\dot{m}(t) &= \sigma\bar{V}(t)V_0 - \mu_m m(t) + \sigma_f\bar{F}(t)F_0 + \sigma_c u_0 u(t).
\end{aligned}$$

Because the normalisation is only used for simulations it is acceptable to use the non-normalised equations for the analysis in the remainder of this thesis.

2.7 Restatement of Objectives

At this point we have satisfied Objective 1; that is, we have found a model suitable for control. Now that we have chosen the model and defined the notation that will be used in this thesis, it is possible to state the second objective in clear mathematical terminology:

1. Fix all parameters (listed in Section 2.1.2) and fix $P_0 = P_{0h}$, $F_0 = F_{0h}$, and $m_0 = m_{0h}$.
2. Fix V_0 such that (2.1)-(2.4) settles to the chronic equilibrium point.
3. Fix any restrictions on the inputs such as maximum or minimum values, switching times, and t_f. (This point will be clearer when the class of controllers is introduced.)
4. Determine if these selections result in a system of the form (2.9)-(2.12) that, for an appropriate choice of controller, has the properties $V(t_f) < V^*$ and $m(t) < 1$ for all $t \in [0, t_f]$.

Before investigating the final two-stage controller, we digress to explore some previous work and other standard control techniques, such as linearisation.

Chapter 3

Preliminary Control Attempts

In this chapter, initial attempts at control will be briefly discussed, including previous work by the authors of [2] and some interesting AIDS-related work from [5]. Methods of control such as linearisation and optimal control will be investigated and shown to be infeasible or problematic for this problem.

3.1 Previous Work

When originally investigating the control problem posed in this thesis, various control schemes were explored, including: methods of controlling bilinear systems [6], converting quadratic systems to bilinear with linear feedback [7], and optimal control. Bilinear control methods are not applicable to our system, since, even when combining them with the transformation outlined in [7], the resultant system was fundamentally nonlinear. Most of the literature regarding optimal control as applied to the immune system includes discussions of how the immune system itself is optimal; however, these documents do not address control through drug inputs [8].

Now we will investigate the control approaches taken in [2] and [5]. Note that there was no control discussion in [1]; that paper was concerned with analysis of the model, not control.

3.1.1 Linear Matrix Inequalities and the 'Epsilon Trick'

In [2] the authors took advantage of the quotient nature of the equations, which our system does not exhibit, to perform an input-output linearisation. This lin-

earisation resulted in a system with a linear forward path and nonlinear feedback. The nonlinear feedback was then thought of as a set of structured, bounded uncertainties. This allowed standard Linear Matrix Inequality (LMI) techniques from robust control to be applied. It was then shown that to minimise the total drug administered over the period of treatment, it is sufficient to minimize the upper bound on one of the internal interactions in the model.

This LMI approach is not feasible for this thesis since we were unable to perform the initial step of input-output linearisation. There was, however, a technique used in [2] that was thought, for a time, to be useful to us, we refer to the the technique as the 'Epsilon Trick'. In both systems, (2.5) and (2.1)-(2.4), numerical errors begin to occur when the antigen concentration gets very small, so the authors of [2] implemented the desired drug concentration as:

$$\hat{u}(t) = \begin{cases} 0 & \text{if } \bar{u}(t) < 0 \\ \frac{\bar{u}(t)}{B(t)+\varepsilon} & \text{otherwise,} \end{cases} \tag{3.1}$$

where $\bar{u}$ is the control signal that is determined by the LMI analysis and ε was chosen to represent the concentration below which drug dosages would be stopped. This value was chosen to correspond to the concentration below which the authors consider the antigen to be eliminated. This concept was the inspiration for investigating and using the V^* term in this thesis.

3.1.2 AIDS and Model Predictive Control

In [5] the researchers use MPC to determine the optimal drug dosage for a patient suffering from HIV. The system under study in this paper has the benefit of very large sampling times, on the order of weeks. This means that computation times on the order of weeks for one control step are acceptable. The researchers also restrict the controller to having only two possible values, 0 or 1, which reduces the required computation time. The choice of stage and terminal cost are straightforward and lead to a reasonable controller for the system.

3.2 New Attempts

3.2.1 Linearisation

A simple approach for controlling a nonlinear system is through linearisation. First, the system is linearised via Taylor series about the operating point of interest.

This linear plant is then used in place of the nonlinear plant to design a linear controller that results in the desired performance. From [9] we know that the linear controller, when combined with the nonlinear system, will exhibit behaviour in a similar manner as to when it is combined with the linear plant, as long as the operating region is within a neighbourhood of the operating point. In [9] the system under study has no delay; none-the-less, this is still an accepted control approach for systems with delay. The two linearisations from the previous chapter will be investigated as to their usefulness for control purposes.

Recall that the system when linearised about the healthy equilibrium point is:

$$\begin{aligned}
\Delta\dot{V}(t) &= (\beta - u_{0h} - \gamma\rho P^*/\mu_f)\Delta V(t), \\
\Delta\dot{P}(t) &= \alpha v_{oh}\rho P^*/\mu_f \Delta V(t-\tau) - \mu_P \Delta P(t), \\
\Delta\dot{F}(t) &= \rho\Delta P(t) - \mu_f \Delta F(t) - \eta\gamma\rho P^*/\mu_f \Delta V(t), \\
\Delta\dot{m}(t) &= \sigma\Delta V(t) - \mu_m \Delta m(t) + \sigma_f \Delta F(t) + \sigma_c \Delta u(t).
\end{aligned} \tag{3.2}$$

Clearly, (3.2) is not useful, since the critical variable, V, depends only on itself and hence is an uncontrollable mode. Indeed, both inputs are eliminated everywhere that they appear, other than in the damage equation, where u has a negative impact and hence is not useful. This linearisation has eliminated a fundamental part of the dynamics in the system; therefore, this approach cannot be used to design a controller for operation near the healthy operating point.

The mild chronic linearisation,

$$\Delta\dot{V}(t) = \frac{\mu_P[\mu_f(\beta - u_{0c}) - \gamma\rho P^*]}{(\beta - u_{0c})(\alpha v_{0c}\rho - \mu_P\eta\gamma)}(-\Delta u - \gamma\Delta F(t)), \tag{3.3}$$

$$\begin{aligned}
\Delta\dot{P}(t) &= \alpha v_{0c}\frac{\mu_P[\mu_f(\beta - u_{0c}) - \gamma\rho P^*]}{(\beta - u_{0c})(\alpha v_{0c}\rho - \mu_P\eta\gamma)}\Delta F(t-\tau) \\
&\quad + \frac{\alpha\mu_P[\mu_f(\beta - u_{0c}) - \gamma\rho P^*]}{\gamma(\alpha v_{0c}\rho - \mu_P\eta\gamma)}\Delta v(t-\tau) \\
&\quad + \frac{\alpha v_{0c}(\beta - u_0)}{\gamma}\Delta V(t-\tau) - \mu_P\Delta P(t),
\end{aligned} \tag{3.4}$$

$$\begin{aligned}
\Delta\dot{F}(t) &= \rho\Delta P(t) - \left(\mu_f + \eta\gamma\frac{\mu_P(\mu_f(\beta - u_{0c}) - \gamma\rho P^*)}{(\beta - u_{0c})(\alpha v_{oc}\rho - \mu_P\eta\gamma)}\right)\Delta F(t) \\
&\quad - \frac{\eta\gamma(\beta - u_{0c})}{\gamma}\Delta V(t),
\end{aligned} \tag{3.5}$$

$$\Delta\dot{m}(t) = \sigma\Delta V(t) - \mu_m\Delta m(t) + \sigma_f\Delta F(t) + \sigma_c\Delta u(t), \tag{3.6}$$

is no more useful for control than the healthy one. From (2.17), it is clear that the virus no longer affects itself: there is replication of the virus only through

an unstable parameter selection or through the inputs, but not due to its own presence. So a fundamental part of the dynamics of the system are missing yet again. Therefore, as in the healthy case, this control method cannot be used to attempt to control (2.9)-(2.12) near the mild chronic equilibrium point.

3.2.2 Linearisation After Eliminating the Bilinear Term

The following method will be used only on the u input.

In both the general linearised equations (2.15) and the original equations (2.9)-(2.12), it is easy to see that driving V to zero becomes problematic, since the effect of any control input will be eliminated as $V \to 0$. As discussed in Section 3.1.1, the authors of [2] introduced an ε term and a variable substitution to attempt to avoid numerical problems. A similar approach was taken here, introducing a new input term, $\hat{u}$, under the following substitution:

$$u(t) = \begin{cases} \hat{u}(t)/V(t) & V(t) > V^* \\ 0 & V(t) \leq V^*. \end{cases} \tag{3.7}$$

The difficulty with this method appears when trying to introduce this new input into (2.12). Since u appears linearly in this equation, the introduction of $\hat{u}$ introduces a division, so when (2.9)-(2.12) is linearised about the healthy equilibrium point under the substitution (3.7), the resulting system has a zero division, which is unacceptable.

3.2.3 Feedback Linearisation

There are two main types of feedback linearisation, full and input-output. Both require the choice of an output. In this case, the natural choice is V since that is the variable of most interest. It is easiest to use these methods on systems with only one input, so they were applied to one input at a time under the logic that, if it does not work for them individually then it likely will not work for both together.

There is a standard test, outlined in [10], that can be used to determine whether full feedback linearisation is possible. The system fails this test.

Input-output feedback linearisation is somewhat more complicated. This method involves a change of variables that splits the dynamics of a nonlinear system into two parts, linear and nonlinear, such that the input-output characteristics are fully expressed in the linear part. The transformed state variables of the linear part

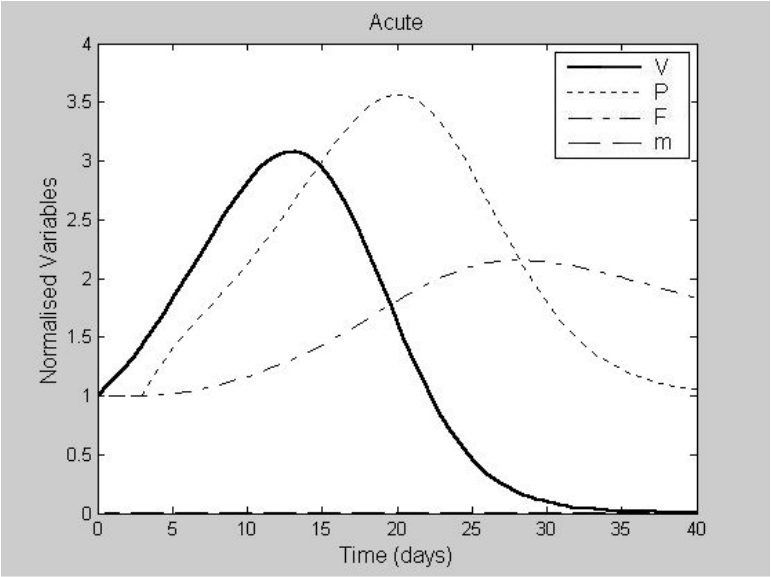

Figure 3.1: Simulation of an uncontrolled acute response

act as an input to the remainder of the system dynamics, which are nonlinear. A transformation was found that, at first, seemed viable for this system, however the diffeomorphism broke down at $V = 0$.

3.2.4 Optimal Control - Model Predictive Control

Optimal Control through Model Predictive Control (MPC) was investigated as a possible avenue, motivated by the work in [5]. The main idea of MPC is to find the optimal control signal over a finite horizon, then apply that signal at the current time step. This cycle is repeated until the system reaches the desired region or point. Optimality can be determined through a combination of state cost and final value cost. It can be shown that stability is guaranteed if a specific set of assumptions is satisfied [11].

One of the major drawbacks of this scheme is that it is calculation intense. In simulations of an acute case it can easily be seen that in order to capture the upswing and downturn characteristic of the response, it is necessary to have a horizon of at least 20 days (see Figure 3.1); however, in order to capture the fast downturn, the step size must be on the order of 1/4 days, so there are easily 80 steps in the horizon. In the simplest case, where the input has only two possible values, this

results in 2^{80} calculations if every path is investigated exhaustively. Even Dynamic Programming [12] is not able to reduce this value to a tractable number, since we run into quantization problems and numerous situations where paths do not cross and, hence, there is little to no simplification. Our conclusion is that MPC is not useful in determining the optimal time in which to apply v, essentially because of the curse of dimensionality.

At the time that we were investigating MPC the drug cocktail input information had not yet been found, so MPC was not attempted on this input. This is definitely an avenue to pursue for future work.

3.3 Summary

From the results in this section it quickly becomes clear that the system's dynamics are fundamentally non-linear. Any attempt at linearisation, even with certain substitutions, results in a loss of critical dynamics. Furthermore, optimal control via MPC is infeasible for determining a useful v since that aspect of the system suffers from the curse of dimensionality. We conclude that, for a control approach to be succesful, it must take advantage of the particular nonlinear structure in (2.9)-(2.12). Such a control scheme is considered in the next chapter.

Chapter 4

Two-Stage Control

In this chapter, we first investigate the use of the cocktail drug (u) to achieve Objective 2. It is shown that the only-u scheme works sometimes, but there exist parameter and initial condition choices which result in death if health is achieved, which is possible due to the definitions of death and health. This motivates the use of the second control input, the inhibitor drug (v), to increase the range of parameters and initial conditions for which Objective 2 can be satisfied. The final controller proposed in this work is a two-stage open-loop controller, in which application of v is followed by application of u. This is similar to multi-stage treatments used in chemotherapy, but we are not aware of multi-stage controllers being used for this type of application.

In Section 2.5, it was established that $F \in [\check{F}, \hat{F}]$, where a conservative choice of $\check{F}$ is zero, and $V \in [0, \hat{V}]$. Note that most variable bounds will depend on plant parameters, initial conditions, and the control inputs. This dependence can easily introduce difficulties in the mathematical analysis. We avoid such problems by considering only the use of open-loop control applied over a finite time interval, $[0, t_f]$. Throughout, the reader should consider everything to be fixed, that is, we will determine results for a given set of parameters, initial conditions, t_1 (to be introduced soon), t_f, and either U or $\check{U}$ and $\hat{U}$.

In the theorem statements in this chapter, $\hat{F}$, $\check{F}$ and $\hat{V}$ are, as usual, bounds on the signals F and V; they can be (typically conservative) bounds computed *a priori*, or they can be the actual maximum and minimum values of F and V in the range $t \in [0, t_f]$,

Throughout this chapter, the following assumptions are made:

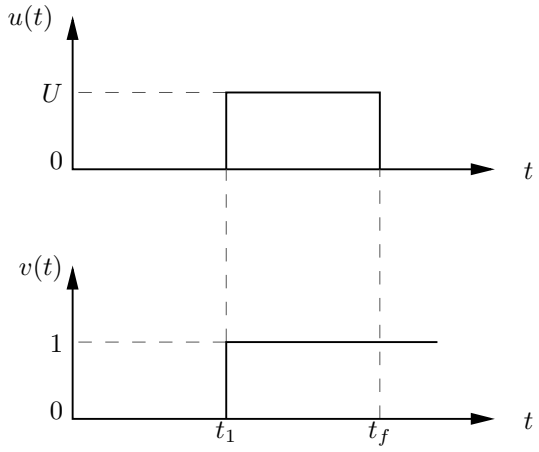

Figure 4.1: Timing diagram for control signals v and u.

- $V_0 > V^*$, since otherwise the patient is already healthy and control is not needed
- $m_0 = m_{0h} < 1$, $F_0 = F_{0h}$, and $P_0 = P_{0h}$ as outlined in Section 2.1.1
- $F(t) = F_0$ and $V(t) = 0$ for all $t \in [-\tau, 0)$ as outlined in Section 2.1.1
- $\check{F}$, $\hat{F}$, and $\hat{V}$ all exist and lie in $[0, \infty)$, as discussed above.

4.1 Defining the Controller Class

This chapter deals with open-loop controllers consisting of a single switch between application of v and application of u, as shown in Figure (4.1); recall that v is 'on' when $v = 0$. Mathematically, inputs of the following form, where $t_1 \in [0, t_f]$, will be considered:

$$v(t) = \begin{cases} 0 & t \in [0, t_1] \\ 1 & \text{otherwise} \end{cases}, \tag{4.1}$$

$$u(t) = \begin{cases} U & t \in [t_1, t_f] \\ 0 & \text{otherwise} \end{cases}. \tag{4.2}$$

This choice of controller class is reasonable both mathematically and physically; step-type signals make analysis feasible while modelling drug dosage through an intravenous drip. It is also important to note that there are upper and lower

physical limits to the amount of u that can be used (i.e., you can only put so much or so little drug into a person!); therefore, we make the reasonable assumption that $U \in [\check{U}, \hat{U}]$, where $\check{U} > 0$ and $\hat{U} < \infty$ are known.

The following theorem gives a condition under which the patient is guaranteed to stay alive; recall that the term 'alive' was defined in Section 1.1. The result will prove to be useful throughout the chapter.

Theorem 4.1 *If the condition*

$$\frac{\sigma\hat{V} + \sigma_f\hat{F} + \sigma_C U}{\mu_m} \leq 1$$

holds, then the patient is alive.

Proof:

Begin with (2.12), where $t \in [0, t_f]$:

$$\begin{aligned} \dot{m}(t) &= \sigma V(t) - \mu_m m(t) + \sigma_f F(t) + \sigma_c u(t) \\ &\leq \sigma\hat{V} - \mu_m m(t) + \sigma_f\hat{F} + \sigma_c U. \end{aligned}$$

This implies that

$$\begin{aligned} m(t) &\leq e^{-\mu_m t} m_0 + \frac{\sigma\hat{V} + \sigma_f\hat{F} + \sigma_c U}{\mu_m}(1 - e^{-\mu_m t}) \\ &< \max\left\{ m_0, \frac{\sigma\hat{V} + \sigma_f\hat{F} + \sigma_c U}{\mu_m} \right\} \\ &\leq 1, \end{aligned}$$

where the second inequality follows since t is restricted to $[0, t_f]$. The final inequality follows from the theorem condition and $m_0 < 1$. □

We now turn to the controller analysis, starting with the situation where only the cocktail drug, u, is applied, that is, $t_1 = 0$. Before investigating the true controller, the ideal situation in which U can be made arbitrarily large will be considered.

4.2 Control Using Only u: The Ideal Case, $\hat{U} = \infty$

In this subsection, we examine the situation when U is not bounded above since, it is conjectured that using a large enough U for a short enough time guarantees that the patient will be both alive and healthy. It will be shown that this conjecture is true only sometimes.

Since we are focusing on application of control for a short time interval, it is assumed in this section that the region of operation satisfies $t_f \leq \tau$. The equations (2.9)-(2.12) simplify to the following (for $t \in [0, t_f]$):

$$\begin{aligned}
\dot{V}(t) &= (\beta - U - \gamma F)V(t) && (4.3)\\
\dot{P}(t) &= \mu_P(P(t) - P^*) = 0 && (4.4)\\
\dot{F}(t) &= \rho P_0 - (\mu_f + \eta\gamma V(t))F(t) && (4.5)\\
\dot{m}(t) &= \sigma V(t) + \sigma_f F(t) + \sigma_C U - \mu_m m(t). && (4.6)
\end{aligned}$$

Since all variables are positive, (4.5) implies that $F(t) \leq F_0$, so a (nonconservative) choice for $\hat{F}$ is F_0. Furthermore, for $U > \beta$ it follows that $V(t) \leq V_0$, so a (nonconservative) choice for $\hat{V}$ is V_0.

Define

$$w \triangleq \frac{\sigma V_0 + \sigma_f F_0 + \sigma_C U}{\mu_m}.$$

Fix t_f. From (4.3), it is clear that health can be achieved by choosing sufficiently large U. For any such U, if $w \leq 1$, then (using Theorem 4.1) the patient will also be alive and Objective 2 is satisfied. An obvious question at this point is, what happens when $w > 1$? The following theorem addresses this case:

Theorem 4.2 *Suppose that* $U > \max\left\{\beta, \frac{\sigma_c U}{\mu_m}\right\}$ *and that* $\frac{\sigma_c U}{\mu_m} > 1$. *Define:*

$$\begin{aligned}
T_1 &\triangleq \frac{\ln \frac{V^*}{V_0}}{\beta - U},\\
T_2 &\triangleq \frac{\ln \frac{V^*}{V_0}}{\beta - U - \gamma\hat{F}},\\
T_3 &\triangleq \frac{\ln \frac{1-w}{m_0-w}}{-\mu_m},\\
T_4 &\triangleq \frac{\ln \frac{1-\frac{\sigma_C U}{\mu_m}}{m_0-\frac{\sigma_C U}{\mu_m}}}{-\mu_m}.
\end{aligned}$$

Then the following relationships hold:

1. *If $t_f > T_1$ then the patient is healthy.*
2. *If the patient is healthy, then $t_f > T_2$.*
3. *If $t_f < T_3$ then the patient is alive.*
4. *If the patient is alive, then $t_f < T_4$.*

In the limit as U approaches infinity, T_1 approaches T_2 and T_3 approaches T_4. Furthermore, in this limit, the patient is alive and healthy iff $1 - m_0 > \sigma_c \ln(V_0/V^*)$.

Proof

This proof will begin by investigating Items 1 through 4.

Proof of Item 1:

From (4.3),

$$\begin{aligned} \dot{V}(t) &= (\beta - U - \gamma F(t))V(t) \\ &\leq (\beta - U)V(t), \end{aligned}$$

which implies

$$V(t) \leq e^{(\beta-U)t}V_0,$$

which in turn implies

$$\begin{aligned} V(t_f) &\leq e^{(\beta-U)t_f}V_0, \\ &< e^{(\beta-U)T_1}V_0, \\ &= e^{(\beta-U)\frac{\ln \frac{V^*}{V_0}}{\beta-U}}V_0 \\ &= V^*, \end{aligned}$$

where the second inequality follows since $U > \beta$. □

Investigating the form of T_1 reveals that $T_1 \to 0$ as $U \to \infty$. Hence, for large enough U it is possible to restrict t_f to the range $[0, \tau]$ while still guaranteeing health.

Proof of Item 2:

From (4.3),

$$\begin{aligned}\dot{V}(t) &= (\beta - U - \gamma F(t))V(t) \\ &\geq (\beta - U - \gamma \hat{F})V(t),\end{aligned}$$

which implies

$$V(t) \geq e^{(\beta - U - \gamma \hat{F})t} V_0,$$

which in turn implies

$$V(t_f) \geq e^{(\beta - U - \gamma \hat{F})t_f} V_0.$$

Use the fact that $V(t_f) < V^*$ to conclude

$$V^* > e^{(\beta - U - \gamma \hat{F})t_f} V_0,$$

which, since $U > \beta$, implies that

$$t_f > \frac{\ln(V^*/V_0)}{\beta - U - \gamma \hat{F}} = T_2.$$

□

Proof of Item 3:

Recall that the region of interest satisfies $t_f \leq \tau$, so we can use $\hat{V} = V_0$ and $\hat{F} = F_0$, as discussed above. To verify that the patient is alive, we must examine m. To this end, start with

$$\dot{m}(t) \leq \sigma V_0 + \sigma_f F_0 + \sigma_C u(t) - \mu_m m(t),$$

where $t \in [0, t_f]$. It follows that

$$\begin{aligned}m(t) &\leq e^{-\mu_m t} m_0 + \frac{\sigma V_0 + \sigma_f F_0 + \sigma_C U}{\mu_m}(1 - e^{\mu_m t}) \\ &= e^{-\mu_m t} m_0 + w(1 - e^{\mu_m t}), \\ &= w + e^{-\mu_m t}(m_0 - w) \\ &< w + e^{-\mu_m T_3}(m_0 - w) \\ &= w + e^{-\mu_m \frac{\ln \frac{1-w}{m_0 - w}}{-\mu_m}}(m_0 - w) \\ &= 1,\end{aligned}$$

where the third inequality inequality follows from the facts that $t < T_3$ and $m_0 - w < 0$. □

Proof of Item 4:

We will show that, if $t_f \geq T_4$, then the patient is dead. We proceed by finding a lower bound on m. Starting with

$$\dot{m}(t) \geq \sigma_C u(t) - \mu_m m(t)$$

we see that

$$m(t) \;\; \geq \;\; e^{-\mu_m t} m_0 + \frac{\sigma_C U}{\mu_m}(1 - e^{\mu_m t}),$$

which implies

$$\begin{aligned} m(t_f) &\geq e^{-\mu_m t_f} m_0 + \frac{\sigma_C U}{\mu_m}(1 - e^{\mu_m t_f}) \\ &= w + e^{-\mu_m t_f}(m_0 - \frac{\sigma_C U}{\mu_m}) \\ &\geq w + e^{-\mu_m T_4}(m_0 - \frac{\sigma_C U}{\mu_m}) \\ &= w + e^{-\mu_m \frac{\ln \frac{1 - \frac{\sigma_C U}{\mu_m}}{m_0 - \frac{\sigma_C U}{\mu_m}}}{-\mu_m}}(m_0 - \frac{\sigma_C U}{\mu_m}) \\ &= 1, \end{aligned}$$

where the second inequality follows from the fact that $m_0 - \frac{\sigma_C U}{\mu_m}$ is negative. □

Now we must investigate what happens in the limit as U approaches infinity. It is desired to obtain the following relationships:

- $T_1 = T_2$,
- $T_3 = T_4$, and
- $T_3 > T_1$.

First we will investigate T_1 and T_2:

$$\begin{aligned} \lim_{U \to \infty} \frac{T_2}{T_1} &= \frac{\frac{\ln \frac{V^*}{V_0}}{\beta - U - \gamma \hat{F}}}{\frac{\ln \frac{V^*}{V_0}}{\beta - U}} \\ &= 1. \end{aligned}$$

Similarly, for T_3 and T_4:

$$\begin{aligned}
\lim_{U\to\infty}\frac{T_3}{T_4} &= \lim_{U\to\infty}\frac{\frac{\ln\frac{1-w}{m_0-w}}{-\mu_m}}{\frac{\ln\frac{1-\frac{\sigma_C U}{\mu_m}}{m_0-\frac{\sigma_C U}{\mu_m}}}{-\mu_m}} \\
&= \lim_{U\to\infty}\left(\frac{\frac{\sigma_c(m_0-1)}{\mu_m^2(m_0-w)(1-w)}}{\frac{\sigma_c(m_0-1)}{\mu_m^2(m_0-\frac{\sigma_C U}{\mu_m})(1-\frac{\sigma_C U}{\mu_m})}}\right) \\
&= \frac{\frac{m_0-1}{\sigma_c}}{\frac{\sigma_c}{m_0-1}} \\
&= 1,
\end{aligned}$$

where the second equality follows from the definition of w and L'Hopital's rule.

Finally, we must show that the condition outlined in the theorem is the only way for $T_3 > T_1$ to be true in the limit $U \to \infty$. In other words, we must show that

$$\lim_{U\to\infty}\frac{T_3}{T_1} > 1 \tag{4.7}$$

holds iff $1 - m_0 > \sigma_c \ln(V_0/V^*)$. Towards this end, compute:

$$\begin{aligned}
\lim_{U\to\infty}\frac{T_3}{T_1} &= \lim_{U\to\infty}\frac{\frac{\ln\frac{1-w}{m_0-w}}{-\mu_m}}{\frac{\ln\frac{V^*}{V_0}}{\beta-U}} \\
&= \lim_{U\to\infty}\left(\frac{\frac{\sigma_c(m_0-1)}{\mu_m^2(m_0-w)(1-w)}}{\frac{-\ln\frac{V^*}{V_0}}{(\beta-U)^2}}\right) \\
&= \frac{m_0-1}{\sigma_c\ln\frac{V^*}{V_0}} \\
&= \frac{1-m_0}{\sigma_c\ln\frac{V_0}{V^*}},
\end{aligned}$$

where the second equality follows from the definition of w and L'Hopital's rule and where the third equality follows from the definition of w. We conclude that (4.7) holds iff $1 - m_0 > \sigma_c \ln(V_0/V^*)$. □

A conclusion from Theorem 4.2 is that, even if the step size of u is unbounded, there are some cases where a drug cocktail treatment alone cannot satisfy Objective 2.

4.3 Control Using Only u: The Case Where $\hat{U} < \infty$

As discussed in the previous section, there are cases where even an unbounded u cannot achieve the control goals. One might expect that placing a finite upper bound on U only worsens the situation. The purpose of this section is to evaluate conditions under which the control objectives can be achieved and those under which they cannot be achieved, when $\hat{U} < \infty$. Our approach is to derive conditions for health and conditions for death. Once those conditions are found, it will be shown that there are situations where it is possible to achieve health before causing death and where, to achieve health, the patient must first die.

4.3.1 Conditions for Health

Recall that health is defined as $V(t_f) \leq V^*$. It is worthwhile to note that there are cases where, irrespective of the length of time that u is applied, the patient will be sick.

Theorem 4.3 *If $\hat{U} \leq \beta - \gamma\hat{F}$ then $V(t_f) > V^*$.*

Proof

From (2.9),

$$\begin{aligned} \dot{V}(t) &= (\beta - u(t) - \gamma F(t))V(t) \\ &\geq (\beta - \hat{U} - \gamma\hat{F})V(t). \end{aligned}$$

The assumption $\hat{U} \leq \beta - \gamma\hat{F}$ gives $0 \leq \beta - \gamma\hat{F} - \hat{U}$, implying that $\dot{V} \geq 0$; therefore, V is monotonically increasing. Since $V_0 > V^*$, this implies that $V(t) > V^* \ \forall \ t \in [0, t_f]$. □

To state necessary conditions for health the following definition will be used:

$$T_{h1} \triangleq \frac{\ln(V^*/V_0)}{\beta - \hat{U} - \gamma\hat{F}}. \tag{4.8}$$

Note that T_{h1} is well defined if the patient is healthy since, by the contrapositive statement of Theorem 4.3, $\hat{U} > \beta - \gamma\hat{F}$. Necessary conditions for health are outlined in the following theorem.

Theorem 4.4 *If $V(t_f) \leq V^*$ then $t_f \geq T_{h1} > 0$.*

Proof

From (2.9),

$$\begin{aligned} \dot{V}(t) &= (\beta - u(t) - \gamma F(t))V(t) \\ &\geq (\beta - \hat{U} - \gamma \hat{F})V(t), \end{aligned}$$

which implies

$$V(t) \geq e^{(\beta - \hat{U} - \gamma \hat{F})(t)} V_0$$

and

$$V(t_f) \geq e^{(\beta - \hat{U} - \gamma \hat{F})t_f} V_0.$$

The assumption is that $V^* \geq V(t_f)$, so

$$V^* \geq e^{(\beta - \hat{U} - \gamma \hat{F})t_f} V_0,$$

which implies

$$\ln(V^*/V_0) \geq (\beta - \hat{U} - \gamma \hat{F})t_f$$

and, using the contrapositive of Theorem 4.3,

$$t_f \geq \frac{\ln(V^*/V_0)}{\beta - \hat{U} - \gamma \hat{F}} = T_{h1}.$$

Finally, since $V_0 > V^*$, it follows that $\ln(V^*/V_0) < 0$; hence, T_{h1} is positive. □

The following time (which is well-defined when $\check{U} > \beta - \check{F}$) is required to state sufficient conditions for health:

$$T_{h2} \triangleq \frac{\ln(V^*/V_2)}{\beta - \check{U} - \gamma \check{F}}. \tag{4.9}$$

These conditions are outlined in the following theorem:

Theorem 4.5 *If $\check{U} > \beta - \check{F}$ and $t_f \geq T_{h2}$ then $V(t_f) \leq V^*$.*

Proof This proof is similar to the proof for Theorem 4.4 and is given in Appendix A.7.

4.3.2 Conditions for Death

The following definitions will be useful in stating the next theorem, which outlines necessary conditions for death:

$$w_{d1} \triangleq \frac{\sigma\hat{V} + \sigma_f\hat{F} + \sigma_c\hat{U}}{\mu_m}, \tag{4.10}$$

$$T_{d1} \triangleq \frac{-1}{\mu_m} \ln\left(\frac{1 - w_{d1}}{m_0 - w_{d1}}\right). \tag{4.11}$$

Note that T_{d1} is well-defined in the case where $w_{d1} > 1$, which is the only area of interest, since by Theorem 4.1, if $w_{d1} \leq 1$ the patient is guaranteed to be alive.

Theorem 4.6 *If $m(t_f) \geq 1$ then $t_f \geq T_{d1}$.*

Proof This proof is similar to the proof for Theorem 4.4 and is given in Appendix A.8

Finally, the following definition will be useful in stating the next theorem. which gives sufficient conditions for death:

$$T_{d2} \triangleq \frac{-1}{\mu_m} \ln\left(\frac{1 - \frac{\sigma_c\check{U}}{\mu_m}}{m_0 - \frac{\sigma_c\check{U}}{\mu_m}}\right). \tag{4.12}$$

This term is well defined when $\frac{\sigma_c\check{U}}{\mu_m} > 1$.

Theorem 4.7 *If $\frac{\sigma_c\check{U}}{\mu_m} > 1$ and $t_f \geq T_{d2}$ then $m(t_f) \geq 1$.*

Proof This proof is similar to the proof for Theorem 4.4 and is given in Appendix A.8

4.3.3 Healthy and Alive

The only way to ensure that the patient is both alive and healthy is to choose control variables, initial conditions, and parameters such that $t_f \in [T_{h2}, T_{d1})$. This follows directly from Theorems 4.5 and 4.6. To achieve this requirement it is first necessary to have $T_{h2} < T_{d1}$. Note that satisfying the requirement $t_f < T_{d1}$ may be difficult, since T_{d1} depends on t_f. Using $V_0 = 5.5 \times 10^{-21}$, $V^* = 10^{-40}$, $\check{U} = 0.0001$, $\hat{U} = 1$, $\sigma_c = 0.5$, $t_1 = 0$, and $t_f = 85$ satisfies the requirements and results in Figure 4.2. Observe that $V(t_f) < V^*$ and $m(t) < 1$ for all $t \in [0, t_f]$, so the patient is both healthy and alive.

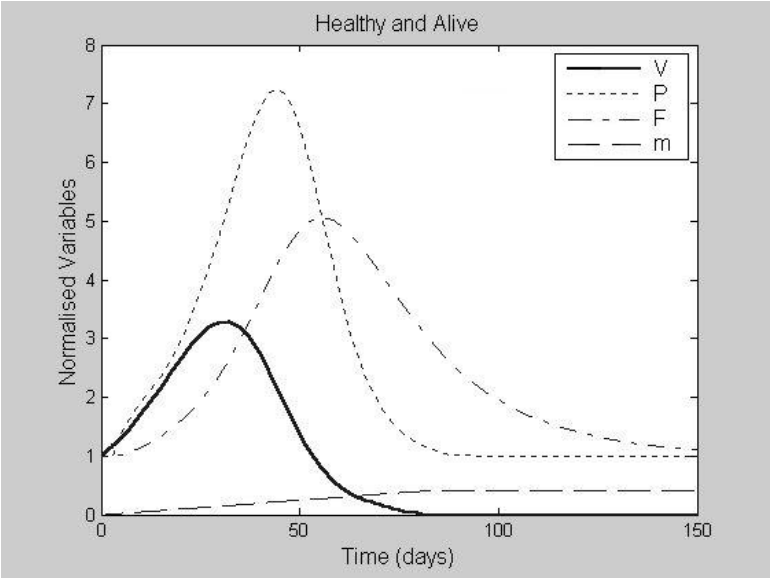

Figure 4.2: Patient is healthy and alive, $U = 0.01$

4.3.4 Healthy but Dead

A particularly important result of the theorems outlined above is that, if $T_{h1} > T_{d2}$, then in every case where the patient is healthy they are also dead. This is a direct result of Theorems 4.4 and 4.7. An example of when this would occur is with the choices $V_0 = 5.5 \times 10^{-21}$, $V^* = 10^{-50}$, $\check{U} = 0.02$, $\hat{U} = 1$, $\sigma_c = 1$, $t_1 = 0$, and $t_f = 16$, shown in Figure 4.3. Observe that, although health is achieved at approximately 15 days, the patient died long before, at approximately 11 days.

4.4 Control using u and v

To try to relax the conditions under which the controller is guaranteed to fail, we now explore the possibility of using the second control input, v, in conjunction with control input u, as described at the beginning of this chapter. We will establish some notation before progressing: define x_u to be the state under u-only control, x_v to be the state under v-only control, and x_n to be the state under no control.

Before discussing the actual two-stage controller, we first attempt to gain some insight into how the inhibitor input affects the system when it is used with no other inputs.

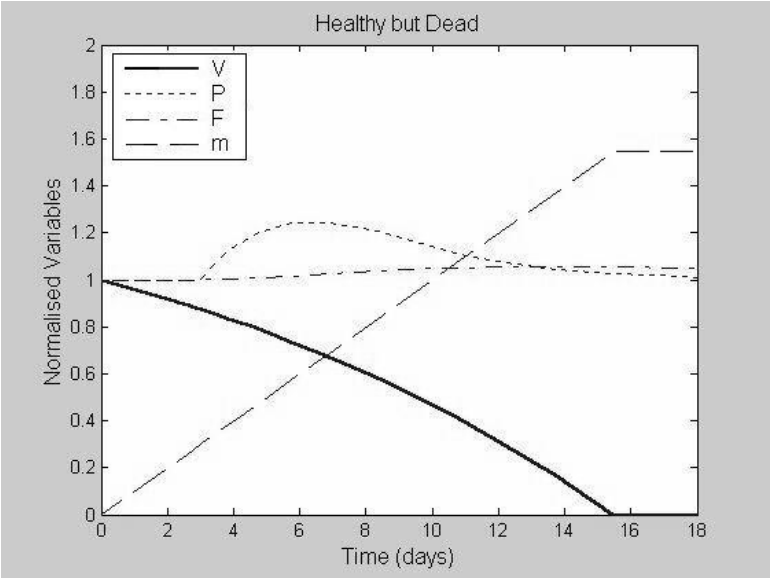

Figure 4.3: Patient is healthy but dead, $u = 0.1$

4.4.1 The Mechanism Behind the Inhibitor Drug

The problem with the u-only control method is that there are cases where applying enough u to make the patient healthy always results in death. A way around this problem is to find a method of decreasing V while using less u; this is the ultimate purpose of the inhibitor drug, v.

It is worthwhile to return to the original equations to obtain some insight into the mechanism underlying the inhibitor drug's performance. Recall that the controlled equations are as follows:

$$\dot{V}(t) = (\beta - u(t) - \gamma F(t))V(t), \tag{4.13}$$
$$\dot{P}(t) = \xi(m(t-\tau))\alpha V(t-\tau)F(t-\tau)v(t-\tau) - \mu_P(P(t) - P^*), \tag{4.14}$$
$$\dot{F}(t) = \rho P(t) - (\mu_f + \eta\gamma V(t))F(t), \tag{4.15}$$
$$\dot{m}(t) = \sigma V(t) - \mu_m m(t) + \sigma_f F(t) + \sigma_c u(t). \tag{4.16}$$

Consider (4.13). Since β is a constant, then clearly the only way to decrease the virus, other than by increasing u, is by increasing F. Moving to (4.15) we see that the only term that increases F is P. Finally, investigating (4.14), we see that P increases through increases in V and F. So we conclude that an *increase* in V is

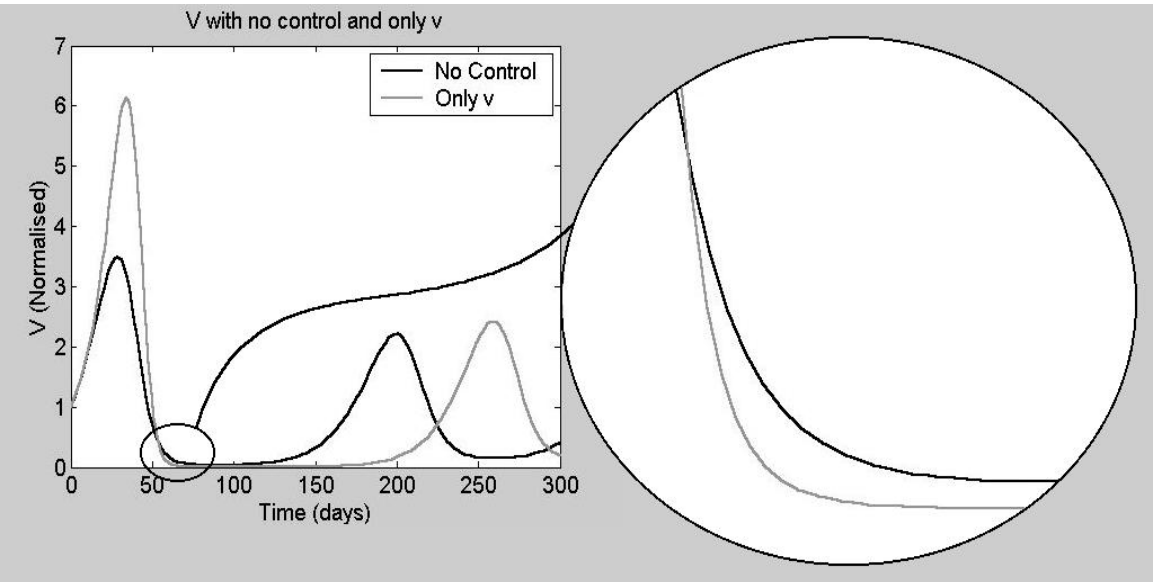

Figure 4.4: V with no Control and Only v, $t_1 = 20$, $V_0 = 5.5 \times 10^{-21}$

needed. At this point we have an obvious paradox: we are seeking to decrease V through a mechanism that requires an increase in V.

The paradox can be resolved by using the inhibitor drug, v, in a somewhat counter-intuitive manner. Introducing v allows the virus to increase *in the short term* to create an increase in the antibody over the long term, which will, in turn, result in a rapid decrease in the virus. In more detail, the mechanism by which the inhibitor drug is intended to increase the antibody concentration is described as follows:

1. At $t = 0$, set $v = 0$, causing plasma production to stop, even though there is virus in the body. This will stop the generation of antibody and hence allow the virus to increase (in the short term) more than it would in the uncontrolled case.

2. At $t = t_1$, set $v = 1$, allowing the plasma cells to again 'see' the virus. Since the virus level is higher now than it would have been in the uncontrolled case, the immune system will have a much stronger reaction; plasma production, and hence antibody production, will occur at a much higher rate than in the uncontrolled case.

3. The end result of Item 2 is that, after a certain time, $F_v > F_n$.

Of course the hope is that the increase in F is enough to offset the increase in V so that, in the long term, $V_v < V_n$. Simulations verify that this does, in fact, happen. For example, a comparison of V_v and V_n can be seen in Figure 4.4. Observe that the V_v reaches a minimum faster than V_n and that the minimum is smaller. Note that, throughout the interval shown in Figure 4.4, $u = 0$. In the next section, the full two-stage controller is considered.

4.4.2 Demonstration of the Two-Stage Controller

In this section the full two-stage controller is considered. We provide further justification for proposing this control scheme and present simulation results showing a situation in which the scheme is successful.

Let us assume that v has been applied for some time; hence, the peak value of V_v will be larger than the peak value of V_n, as explained above. From (4.13) we see that every peak in V, in particular the first peak, occurs when $F(t) = \beta/\gamma$. From simulation we have found that the first peaks of the controlled and uncontrolled V occur at almost the same time instant; call this time T^*. As a result, in both cases it is approximately true that $F(T^*) = \beta/\gamma$. This implies that, for $t \geq T^*$ until an unknown time, $P_v(t) > P_n(t)$; hence we have achieved our objective of increasing P (and, thus, F). One way of thinking about the effects of this input is to imagine that we are simply increasing the initial condition of V, but that the time it takes to accomplish this increase is such that the peaks of the two systems coincide in time. Note that the switch in v occurs well before V reaches its first peak and that v has no adverse affect on m aside from the increase in V.

It is important to note that the final effect of control with only v is to slightly decrease the time to the first minimum in V and to decrease the value of that minimum (Figure 4.4). If that minimum is decreased to such an extent that it is less than V^*, then v-only has made the patient healthy! Unfortunately, the increase required in V to achieve health for a reasonable sized V^* is very large. Therefore, the true use of this input is to decrease the value of V to a point where the u input is able to quickly make the patient healthy without causing death. An example that uses the same settings as in the healthy-but-dead example, i.e., $V_0 = 5.5 \times 10^{-21}$, $V^* = 10^{-50}$, $\check{U} = 0.02$, $\hat{U} = 1$, $\sigma_c = 1$, is given in Figure 4.5. The critical result here is that it was possible to use a much lower value of U while still achieving health. This U was sufficiently small and applied for a sufficiently short period of time such that it allowed the patient to be made healthy before they died, as opposed to the size of U required to make the patient healthy in prior sections (Figure 4.3). Note that m is close to 1 but does not actually reach that value.

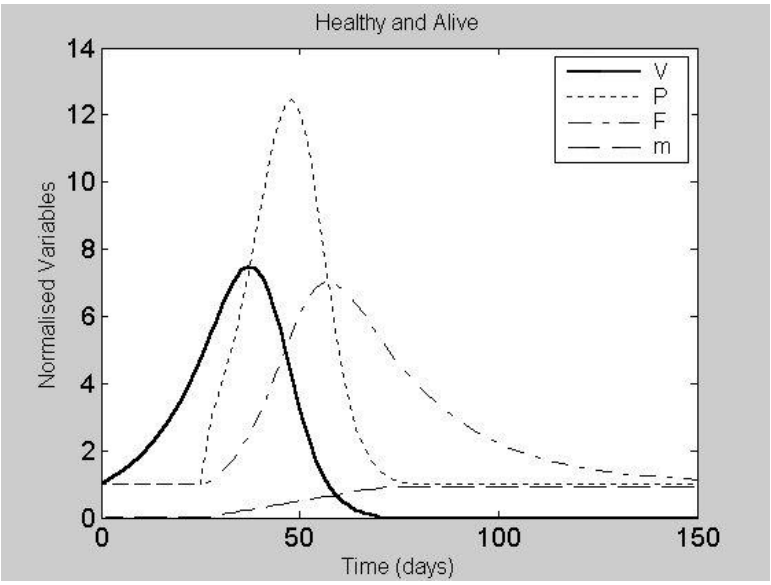

Figure 4.5: Results of Two-Stage Controller, $t_1 = 25$, $t_f = 71$, $U = 0.02$

We are well aware that there are numerous questions remaining regarding this control scheme, such as how to choose t_1 and how to guarantee that F increases enough to offset the increase in V. These questions are very difficult to answer due to the nonlinearities and the time delay exhibited by the system; rigorous analysis remain elusive.

Chapter 5

Summary and Future Work

In this thesis a model of the human immune system was found and modified such that it is suitable for control purposes. This model is nonlinear with a time delay and has proven to be very difficult to control. Early on it was shown that certain standard control techniques, such as linearisation, were either problematic or infeasible for this model. The root of the problem is that critical system dynamics are nonlinear in nature. Specifically, problems tended to arise at $V = 0$. Hence, it was concluded that the system is fundamentally nonlinear.

Some open-loop controllers were then investigated. A single-stage controller, using only the u input, was investigated and it was concluded that the u input alone would only work under certain choices of parameters and initial conditions. Finally, a two-stage controller was shown, in simulation, to satisfy the control objective for a wider range of parameters and initial conditions than the u-only controller.

Avenues for further investigation include:

1. Using MPC for the u only input while not restricting u to having the form of a step.
2. Finding a rigorous analysis for the usefulness of the two-stage controller.
3. Investigating a two-stage controller where u and v are not in the form of a step.
4. Attempting to implement control of this type on a more complicated model.

Appendix A

Proofs

A.1 Proof of Theorem 2.2 – Sufficient Conditions for Stability of the Mild Chronic Equilibrium Point

The goal here is to show that the mild chronic equilibrium point (2.14) is stable if $0 < \frac{f-d}{a} < b - g - f\tau$ and $a - g\tau > 0$.

Throughout this discussion, the term 'quasipolynomial' will be used to refer to systems of the form

$$f(s) = a_n s^n + a_{n-1} s^{n-1} + \cdots + a_0 + (b_n s^q + b_{q-1} s^{q-1} + \cdots + b_0) e^{-s\tau}, \qquad \text{(A.1)}$$

where a_i, b_i and τ are constants. Also, throughout this proof the functions $\Re(x)$ and $\Im(x)$ will be used to denote the real and imaginary parts of x.

This proof is not straightforward. Before starting the formal proof, an outline will be presented.

A.1.1 Outline of the Proof

The proof begins by using the fact that the stability of a nonlinear equilibrium point can be assessed via the stability of the linearisation about that point, as outlined in Section 2.4.1. To evaluate the stability of a linear system with delay, the characteristic quasipolynomial is evaluated. If all the roots of that quasipolynomial

lie in the open left half plane, then the system is stable [13]. Therefore, the first step of the proof is to find the characteristic quasipolynomial of (2.17)-(2.20) so that the location of its roots can be determined.

Due to the $e^{-s\tau}$ terms in the quasipolynomial, standard polynomial based methods cannot be used to determine root locations. The Mikhailov Criterion, however, is ideally suited to this task. For readers not familiar with this criterion, it is presented in Appendix A.1.2. Using this criterion, it is possible to conclude that, if a specific curve never enters the fourth quadrant of the complex plane, then the system will be stable. It is shown that the assumptions in Theorem 2.2 are sufficient to ensure that this quadrant exclusion will occur.

A.1.2 A Useful Theorem

The Mikhailov Criterion is ideal for evaluating the stability of a delay system; however, it is not well-known in North America and a proper statement of the theorem, let alone a proof, is difficult to find.

Before stating the theorem, we will define a specific type of quasipolynomial, denoted $Q(s)$. Let $P_1(s)$ and $P_2(s)$ be real-valued polynomials such that $\deg(P_1) > \deg(P_2)$, $P_1(s)$ is monic, and

$$Q(s) \triangleq P_1(s) + P_2(s)e^{-s\tau}$$

has no roots on the imaginary axis. Also define $n \triangleq \deg(P_1(s))$.

Theorem A.1 (The Mikhailov Criterion) *All roots of $Q(s)$ have negative real parts iff the increase in the argument of $Q(j\omega)$ equals $n\pi/2$ as ω increases from zero to infinity.*

Proof Please see Appendix A.2.

A.1.3 Main Proof

Assume that

$$0 < \frac{f-d}{a} < b - g - f\tau \tag{A.2}$$

and

$$a - g\tau > 0. \tag{A.3}$$

Recall that if the system, when linearised about the equilibrium point in question, is stable, then so is the nonlinear equilibrium point (see Section 2.4.1). Also recall that the mild chronic equilibrium point relates to the case where $\mu_p \tau \leq 1$ and that we are assessing stability when there is no input to the system.

To evaluate stability of a linear system, the characteristic quasipolynomial must be investigated; therefore, the first step in the proof is to find the characteristic quasipolynomial, which can easily be determined via the chronic linearisation (2.17)-(2.20) to be

$$\det \begin{bmatrix} s & 0 & \gamma V_{0c} & 0 \\ -\alpha\frac{\beta}{\gamma}e^{-s\tau} & s+\mu_p P & -\alpha V_{0c}e^{-s\tau} & 0 \\ \eta\beta & -\rho & s+(\mu_f+\eta\gamma V_{0c}) & 0 \\ -\sigma & 0 & -\sigma_f & s+\mu_m \end{bmatrix} \tag{A.4}$$

or

$$\begin{aligned}&(s+\mu_m)\{s^3+(\mu_p+\mu_f+\eta\gamma V_{0c})s^2+\\ &[\mu_p(\mu_f+\eta\gamma V_{0c})-\eta\gamma\beta V_{0c}]s-\eta\gamma\beta\mu_p V_{0c}+(\alpha\rho V_{0c}s-\alpha\beta\rho V_{0c})e^{-s\tau}\}.\end{aligned} \tag{A.5}$$

By [13], (2.17)-(2.20) is stable iff all roots of (A.5) lie in the Open Left Half Plane (OLHP). Using (2.22)-(2.26), define

$$Z(s) \triangleq s^3+as^2+bs-d+(gs-f)e^{-s\tau}; \tag{A.6}$$

hence, the characteristic quasipolynomial is $(s+\mu_m)Z(s)$. Since all model parameters are positive, the system pole at $-\mu_m$ is stable and can be ignored for the remainder of the discussion. For stability, it is required that all of the roots of $Z(s)$ lie in the OLHP. The Mikhailov Criterion will be used to find sufficient criteria for this to occur; therefore, $Z(s)$ must be shown to be a quasipolynomial of the form of $Q(s)$, i.e.:

1. s^3+as^2+bs-d is monic,
2. $\deg(s^3+as^2+bs-d) > \deg(gs-f)$, and
3. $Z(s)$ has no roots on the imaginary axis.

Clearly, Item 1 is satisfied. Item 2 is satisfied since $\deg(s^3+as^2+bs-d)=3$ and $\deg(gs-f)=1$. Note that this implies that $n=3$. Item 3 is more difficult to satisfy, but it will be shown later in the proof that it is indeed fulfilled, so we will

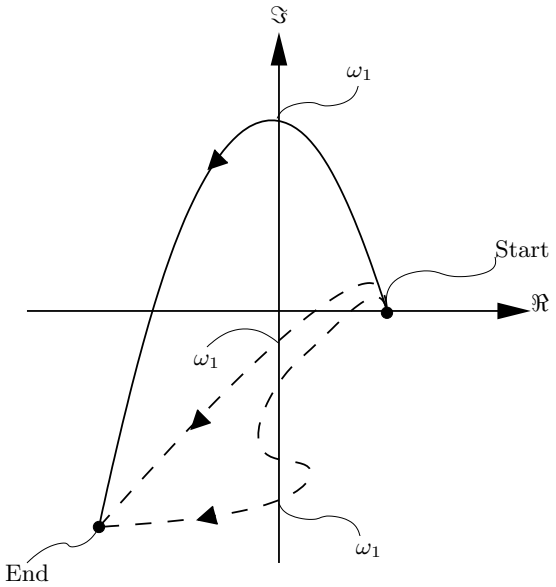

Figure A.1: Path of Z through the Imaginary plane. Solid line indicates desirable path, dashed line indicates undesirable paths.

proceed with the investigation of the properties of $Z(j\omega)$ as ω increases from zero to infinity. As a result, throughout the remainder of the proof, assume that $\omega \geq 0$ unless stated otherwise.

The basis for the remainder of the proof will be outlined before we continue. First it be will shown that $\arg(Z(0)) = 0$ and $\arg(Z(j\omega))$ for large ω approaches $3\pi/2 + 2\pi k$, for some $k \in \mathbf{Z}$. Therefore, if $Z(j\omega)$ is restricted from entering the fourth quadrant and from passing through the origin then $k = 0$, consistent with the requirement from Theorem A.1 (see Figure A.1). One way to ensure that $Z(j\omega)$ neither enters the fourth quadrant nor passes through the origin is to require $\Im(Z(j\omega)) > 0$ for all ω where $\Re(Z(j\omega)) \geq 0$. Note that restricting $Z(j\omega)$ from passing though the origin is sufficient to ensure that $Z(s)$ has no zeros on the imaginary axis, as required by the third restriction on quasipolynomials of the form of $Q(s)$.

The following technical result will be required:

Lemma A.2 $\mu_p \tau \leq 1 \Rightarrow a > b\tau$.

Proof:

Assume $\mu_p\tau \leq 1$ and begin with (2.22):

$$\begin{aligned}
a &= \eta\gamma V_{c0} + \mu_f + \mu_p \\
&= (\eta\gamma V_{c0}\mu_p + \mu_f\mu_p + \mu_p^2)/\mu_p \\
&= \frac{\mu_p(\eta\gamma V_{c0}+\mu_f)-\beta\eta\gamma V_{co}+\beta\eta\gamma V_{co}+\mu_p^2}{\mu_p} \\
&= \frac{b+\beta\eta\gamma V_{c0}+\mu_p^2}{\mu_p} && \text{(from (2.23))} \\
&> \frac{b}{\mu_p} && (\because \text{ all parameters} > 0) \\
&\geq \frac{b(\mu_p\tau)}{\mu_p} && (\because\ \mu_p\tau \leq 1) \\
&= b\tau.
\end{aligned}$$

□

The following facts are used throughout the remainder of the proof:

- $\sin(x) \leq x\ \forall\ x \geq 0$ and
- $\cos(y) \leq 1\ \forall\ y$.

This choice of bound for sine is being used instead of $\sin(x) \leq 1$ since it makes the math cleaner.

Continuing with the formal proof, directly from (A.6) we have

$$\begin{aligned}
\Re(Z(j\omega)) &= -a\omega^2 - d - g\omega\sin(\omega\tau) + f\cos(\omega\tau), && \text{(A.7)} \\
\Im(Z(j\omega)) &= -\omega^3 + b\omega - g\omega\cos(\omega\tau) - f\sin(\omega\tau), && \text{(A.8)} \\
\frac{d\Re(Z(j\omega))}{d\omega} &= -2a\omega - g\omega\tau\cos(\omega\tau) - (g + f\tau)\sin(\omega\tau), && \text{(A.9)} \\
\frac{d\Im(Z(j\omega))}{d\omega} &= -3\omega^2 + b + g\omega\tau\sin(\omega\tau) - (g + f\tau)\cos(\omega\tau). && \text{(A.10)}
\end{aligned}$$

From (A.2), we have $0 < (f - d)/a$. Since $a > 0$, then $f - d > 0$. Use this fact, (A.7), and (A.8) to obtain

$$\begin{aligned}
\Re(Z(0)) &= f - d > 0, && \text{(A.11)} \\
\Im(Z(0)) &= 0, && \text{(A.12)}
\end{aligned}$$

and, for all large ω,

$$\begin{aligned}
\Re(Z(j\omega)) &= -a\omega^2 + O(\omega^2) < 0, && \text{(A.13)} \\
\Im(Z(j\omega)) &= -\omega^3 + O(\omega^2) < 0. && \text{(A.14)}
\end{aligned}$$

Note that (A.13) and (A.14) imply that, as $\omega \to \infty$, $\arg(Z(j\omega)) \to 3\pi/2 + 2\pi k$, $k \in \mathbf{Z}$.

Since the curve starts with a positive real part, ends with a negative real part, and is continuous, there must exist some ω_1 such that $\Re(Z(j\omega_1)) = 0$ and $\Re(Z(j\omega)) < 0 \ \forall \ \omega > \omega_1$. That is, ω_1 represents the largest ω such that the curve crosses the imaginary axis (see Figure A.1). Furthermore, it can be shown that ω_1 is the only crossing.

Claim 1: $Z(j\omega)$ crosses the imaginary axis only once, at $\omega = \omega_1$.

Proof of Claim 1:

Starting from (A.9):

$$\begin{aligned} \frac{d\Re(Z(j\omega))}{d\omega} &= -2a\omega - g\omega\tau\cos(\omega\tau) - (g + f\tau)\sin(\omega\tau) \\ &\leq (-2a + g\tau + (g\tau + f\tau^2))\omega \\ &= [(-a + g\tau) + (-a + g\tau + f\tau^2)]\omega \\ &\leq [-a + g\tau + (-b + g + f\tau)\tau]\omega. \qquad \text{(by Lemma A.2)} \end{aligned}$$

If we combine this result with (A.2), (A.3), and the fact that τ and ω are non-negative we can conclude that

$$\frac{d\Re(Z(j\omega))}{d\omega} < 0, \tag{A.15}$$

which in turn implies that there can be only one imaginary axis crossing, specifically at ω_1. □

The following additional technical result is required for the main proof.

Lemma A.3 *If $\omega \leq \omega_1$, then $\omega_1^2 \leq (f - d)/a$.*

Proof

Since a and d are positive and ω is non-negative, (A.7) gives

$$\Re(Z(j\omega)) < -g\omega\sin(\omega\tau) + f\cos(\omega\tau). \tag{A.16}$$

Since ω_1 is the only imaginary axis crossing (Claim 1), it follows that once $\Re(Z(j\omega))$ becomes negative it must remain negative. It is useful to determine conditions on ω that result in $\Re(Z(j\omega)) < 0$ since these conditions will give a bound on ω_1.

Both $\sin(\omega\tau)$ and $\cos(\omega\tau)$ can be positive or negative; however, it is known that ω follows a path that begins at zero and increases. Therefore, we can conclude that $\sin(\omega\tau)$ and $\cos(\omega\tau)$ begin at 0 and 1 respectively and, as ω increases, $\sin(\omega\tau)$ and $\cos(\omega\tau)$ are positive until $\omega\tau = \pi/2$, after which, until $\omega\tau = \pi$, $\sin(\omega\tau)$ is positive and $\cos(\omega\tau)$ is negative. The critical piece of information here is that $\cos(\omega\tau)$ switches signs before $\sin(\omega\tau)$; therefore, for $\omega\tau \in (\pi/2, \pi)$,

$$\begin{aligned} \Re(Z(j\omega)) &< -g\omega\sin(\omega\tau) + f\cos(\omega\tau) \\ &< 0. \end{aligned}$$

As previously mentioned, $Z(j\omega)$ can only cross the imaginary axis once, so it follows from the above inequality that $\omega_1 < \frac{\pi}{2\tau}$, which in turn implies that $\cos(\omega\tau) \geq 0$ and $-\sin(\omega\tau) \leq 0$ for all $\omega \leq \omega_1$; therefore, from (A.16),

$$\Re(Z(j\omega)) < -a\omega^2 - d + f \quad \forall\ \omega \leq \omega_1. \tag{A.17}$$

Finally,

$$\begin{aligned} & 0 = \Re(Z(j\omega_1)), && \text{(definition of } \omega_1) \\ \Rightarrow\ & 0 \leq -a\omega_1^2 - d + f, && \text{(from (A.17))} \\ \Leftrightarrow\ & \omega_1^2 \leq (f-d)/a. && (\because\ a > 0) \end{aligned}$$

□

To complete the proof, we show that for any $\omega \leq \omega_1$, which includes all ω for which the real part of $Z(j\omega)$ is non-negative, the imaginary part of $Z(j\omega)$ is positive. Begin with (A.8) and the assumption that $\omega \leq \omega_1$:

$$\begin{aligned} \Im(Z(j\omega)) &= -\omega^3 + b\omega - g\omega\cos(\omega\tau) - f\sin(\omega\tau), \\ &\geq \omega[-\omega^2 + (b - g - f\tau)], \\ &> \omega[-\omega^2 + (f-d)/a], && \text{(from (A.2))} \\ &\geq \omega(-\omega^2 + \omega_1^2), && \text{(by Lemma A.3)} \\ &\geq 0. && (\because\ \omega \leq \omega_1) \end{aligned}$$

Therefore, for all $\omega \geq 0$ such that $\Re(Z(j\omega)) \geq 0$ it is also true that $\Im(Z(j\omega)) > 0$, so $Z(j\omega)$ will never enter the fourth quadrant, nor will it pass through the origin, hence $\arg(Z(j\omega)) \to 3\pi/2$ as $\omega \to \infty$. Combining this with the fact that $\arg(Z(0)) = 0$ results in the conclusion that $Z(j\omega)$ increases by $3\pi/2$ as ω increases from zero to infinity. Applying the Mikhailov Criterion (Theorem A.1) gives the result that all roots of $Z(s)$ have negative real parts; therefore, the chronic equilibrium point is stable. □

A.2 Proof of Theorem A.1 - The Mikhailov Criterion

To prove this theorem we first need to present some definitions, introduce the Principle of the Argument, and state some useful lemmas.

A.2.1 Definitions

A complex function f is 'analytic' on a domain if it has a derivative at every point in the domain. An equivalent definition is that f is analytic if it has a Taylor series about each point, x, in the domain such that the series converges to f in an open neighbourhood of x.

A function is 'meromorphic' in a domain if it is analytic throughout that domain except possibly for poles. A function that is analytic on a domain is clearly also meromorphic in that domain.

A 'path' is a curve in the complex plane that has a direction. A path is 'closed' if it ends where it starts and is 'simple' if it never crosses itself.

A point is 'enclosed' by a simple closed path if the point always lies to the right of the path when that path is traversed in its prescribed direction.

The 'net number of counter clockwise (CCW) encirclements' of a point by a curve travelling in the CCW direction is determined by drawing a ray from that point to infinity, then determining the number of CCW crossings of the ray minus the number of CW crossings of the ray. Note that one CCW encirclement is equivalent to an increase of 2π in the argument of the function.

Define Γ to be a simple closed and bounded curve in the complex plane. Furthermore, define $\mathbf{D}$ to be the domain enclosed by Γ. Then $\mathbf{B} \triangleq \mathbf{D} \cup \Gamma$ is closed and bounded.

Denote $G(s),\ s \in \mathbf{C}$, to be some function that is meromorphic on $\mathbf{D}$ and analytic and nonzero on Γ. Define N to be the (finite) number of zeros of $G(s)$ in $\mathbf{D}$ and D to be the (finite) number of poles of $G(s)$ in $\mathbf{D}$. Both N and D count multiplicities. Define $G(\Gamma)$ to be the mapped path determined by evaluating $G(s)$ as s traverses Γ. $G(\Gamma)$ lies in the 'mapped plane'.

A.2.2 The Principle of the Argument

The Principle of the Argument is stated as follows:

Theorem A.4 (Principle of the Argument (POA)) *$G(\Gamma)$ is a simple closed curve whose net number of clockwise encirclements of the origin is exactly $N - D$.*

Proof

The proof for the case where $G(s)$ is rational is presented in [14], but also holds for any meromorphic function under the assumption that N and D are finite. The proof follows directly from Cauchy's theorem. □

A.2.3 Some Useful Lemmas

The following lemmas and corollary will be needed. The first lemma is a standard result:

Lemma A.5 *Any non-constant analytic function satisfies the following two properties:*

1. *Every zero is isolated.*
2. *The set of all zeros is closed.*

Proof

Let's start with Property 1. Let f be a function that is analytic on some domain $\mathbf{D}_1$ and let z be a zero of f with multiplicity k. Therefore $f(z) = 0$ and, by the definition of analyticity, there exists a Taylor series expansion for f around z which converges on an open disk $|x - z| < r$, for some $x \in \mathbf{D}_1$ and some $r > 0$. Write this expansion as $f(x) = (x - z)^k \sum_{n=0}^{\infty} a_{n+k}(x - z)^n$, with $a_k \neq 0$ and $k > 0$.

Observe that $g(x) \triangleq \sum_{n=0}^{\infty} a_{n+k}(x - z)^n$ is analytic on $|x - z| < r$ and, furthermore, $g(z) = a_k \neq 0$, since we have removed all of the zeros of f that lie at z. Since $g(x)$ is analytic on $|x - z| < r$, it is continuous at z. Use this fact and the fact that $g(z) \neq 0$ to conclude that there exists an $\varepsilon > 0$ such that, for all x where $|x - z| < \varepsilon$, we have $|g(x) - a_k| < \frac{a_k}{2}$, implying that $g(x)$ is nonzero on that set. Consequently, the zero z is isolated.

Property 2 follows directly from Property 1. □

It is possible to extend this result to establish that the number of zeros on $\mathbf{D}$ is finite:

Corollary A.6 *If a complex function, f, is analytic on $\boldsymbol{B}$ and nonzero everywhere on Γ, then the number of zeros of f in $\boldsymbol{D}$ is finite.*

Proof

Property 1 from Lemma A.5, combined with the fact that $\mathbf{B}$ is closed and bounded implies that the limit point of any convergent sequence of zeros of f in $\mathbf{B}$ must lie in the set of zeros. However, limit points are not isolated, so from property 2, f cannot have a sequence of zeros in $\mathbf{B}$ that also forms a limit point; hence, there will be a finite number of zeros in $\mathbf{B}$. Furthermore, since $\mathbf{D} \subset \mathbf{B}$, the number of zeros in $\mathbf{D}$ is finite. □

Define $z \in \{x \in \mathbf{C} | Q(x) = 0 \text{ and } \Re(x) \geq 0\}$, i.e., z is a zero of $G(s)$ that lies in the closed right half plane (CRHP). Also, write

$$\begin{aligned} P_1(s) &= s^n + \sum_{i=0}^{n-1} a_i s^i, \\ P_2(s) &= \sum_{i=0}^{q} b_i s^i, \end{aligned}$$

where $q < n$. A useful technical result regarding quasipolynomials of the form $Q(s)$, is:

Lemma A.7 *For a given quasipolynomial, $Q(s)$, there exists an $r \in [0, \infty)$ such that every zero of $Q(s)$ that lies in the ORHP satisfies $|z| < r$.*

Proof

Before we begin the proof note the following:

Fact 1: There exists an $R > 0$ such that, for all $\bar{r} \geq R$,

$$\bar{r}^n - \sum_{i=0}^{n-1} (|a_i| \bar{r}^i) - \sum_{i=0}^{q} (|b_i| \bar{r}^i) > 0.$$

This is a proof by contradiction; hence, assume that, for all $r \in [0, \infty)$, there exists a z such that $|z| \geq r$. Choose $r = R$. This implies that there exists a $z = \bar{r} e^{j\theta}$ such that $|z| = \bar{r} \geq R$. Recall that, since z is a zero of $Q(s)$, $0 = Q(z) = z^n + \sum_{i=0}^{n-1} a_i z^i + (\sum_{i=0}^{q} b_i z^i) e^{-z\tau}$. Since we are only concerned with zeros in the ORHP,

we can choose $\theta \in [-\pi/2, \pi/2]$. Using the fact that $\cos\theta \in [0,1]\ \forall\ \theta \in [-\pi/2, \pi/2]$, the last term in $Q(z)$ can be bounded as follows:

$$\begin{aligned}
|e^{-z\tau}| &= |e^{-(\bar{r}e^{j\theta})\tau}| \\
&= |e^{-\tau\bar{r}(\cos\theta + j\sin\theta)}| \\
&= |e^{-\tau\bar{r}\cos\theta} \cdot e^{-\tau\bar{r}j\sin\theta}| \\
&\leq |e^{-\tau\bar{r}\cdot 0} e^{-\tau\bar{r}j\sin\theta}| \\
&= |e^{-\tau\bar{r}j\sin\theta}| \\
&= 1.
\end{aligned}$$

Using this result and the triangle inequality, we compute

$$\begin{aligned}
0 = |Q(z)| &= \left| z^n + \sum_{i=0}^{n-1} a_i z^i + \left[\sum_{i=0}^{q} b_i z^i\right] e^{-\bar{r}e^{j\omega}\tau} \right| \\
&\geq |z^n| - \left|\sum_{i=0}^{n-1} a_i z^i\right| - \left|\sum_{i=0}^{q} b_i z^i\right| |e^{-\bar{r}e^{j\omega}\tau}| \\
&\geq |z^n| - \left|\sum_{i=0}^{n-1} a_i z^i\right| - \left|\sum_{i=0}^{q} b_i z^i\right| \\
&\geq \bar{r}^n - \sum_{i=0}^{n-1} (|a_i|\bar{r})^i - \sum_{i=0}^{q} (|b_i|\bar{r})^i \\
&> 0,
\end{aligned}$$

which is a contradiction. Note that the final inequality follows from Fact 1. □

This Lemma is useful since it shows that it is possible to enclose all of the CRHP zeros of $Q(s)$ with a bounded Γ.

A.2.4 Main Proof

Proceeding to the main proof of Theorem A.1, recall that the theorem states the following: all roots of $Q(s)$ have negative real parts iff the increase in the argument of $Q(j\omega)$ equals $n\pi/2$ as ω increases from zero to infinity. Note that $Q(s)$ is analytic and therefore has no poles, so $D = 0$.

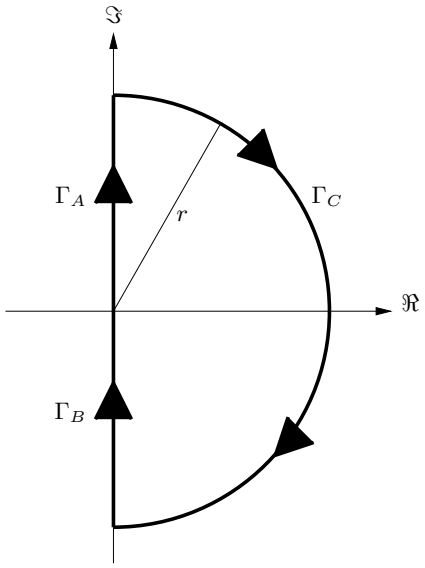

Figure A.2: The path described by Γ, where Γ_A and Γ_B are the positive and negative imaginary axies respectively and Γ_C describes a semicircle with infinite radius

Define the path Γ via Figure A.2. Choose a Γ that encloses all of the zeros of $Q(s)$ that lie in the ORHP (Lemma A.7). Recall that Γ is closed and bounded by definition. Note that Γ, **B**, and **D** fulfil the requirements of Corollary A.6, hence N is finite.

At this point we have satisfied all of the requirements for applying the POA (Theorem A.4) to the function $Q(s)$ and the curve Γ. Recall that it is already known that $D = 0$; hence, from the POA, the net number of CW encirclements of the origin is exactly N. Since N is the number of zeros of $Q(s)$ that lie on the open right half plane, then $N = 0$ iff there are no zeros of $Q(s)$ with non-negative real parts. To achieve no net encirclements of the origin, the total change in the argument of $Q(s)$ as it traverses Γ must equal zero.

It follows from Lemma A.7 and the definition of Γ_C that the change in the argument of $Q(\Gamma_C)$ as θ decreases from π to $-\pi$ is $n\pi$. Therefore, $N = 0$ iff $Q(\Gamma_A)$ and $Q(\Gamma_B)$ together have an $n\pi$ increase of argument as ω increased from negative infinity to positive infinity. Since all coefficients in $Q(s)$ are real, $Q(\Gamma_A)$ and $Q(\Gamma_B)$ are complex conjugates of each other; therefore, all roots of $Q(s)$ have negative real parts iff the total change in the argument of $Q(j\omega)$ as ω increases from zero to infinity equals $n\pi/2$. □

A.3 Proof of Theorem 2.3 - Necessary Conditions for Stability of Chronic Case

Recall that the theorem states that, if the mild chronic case is stable, then $f-d>0$. We will prove the contrapositive statement; hence, assume $f-d\leq 0$.

The beginning of the proof for sufficient conditions (Appendix A.1), up to and including (A.14), with the exception of (A.13), holds in this proof. Consequently, from (A.7) we have

$$\Re(Z(j\omega)) = -a\omega^2 - d - g\omega\sin(\omega\tau) + f\cos(\omega\tau).$$

Therefore,

$$\Re(Z(0)) = f - d \leq 0.$$

Since the equilibrium point does not exist for $f-d=0$ (see (2.28) and the definition of V_{0c}) it cannot be 'stable' there. Hence, restrict the remainder of the discussion to $f-d<0$. Under this restriction, the mapping of $Z(j\omega)$ for $\omega\geq 0$ begins on the negative real line and ends in the third quadrant, and it is therefore impossible to obtain the necessary $3\pi/2$ increase in the argument. We conclude by Mikhailov's Criterion (Theorem A.1) that at least one root of $Z(s)$ does not lie in the OLHP; hence, the equilibrium point is unstable. □

A.4 Proof of Theorem 2.5 - $V=0\subset$ Stable Submanifold

Begin by setting $V(t_f)=0$.

Claim 1: $V(t)=0$ for all $t\geq t_f$.

Proof of Claim 1:

From (2.1),

$$\begin{aligned}\dot{V}(t_f) &= (\beta-\gamma F(t_f))V(t_f)\\ &= (\beta-\gamma F(t_f))\cdot 0\\ &= 0,\end{aligned}$$

which implies that $V=0$ is an equilibrium point of (2.1), independent of the value of $F(t)$, hence $V(t)=0$ for all $t\geq t_f$. □

Note that Claim 1 is logical from a physical standpoint, since the only way that viral levels in a patient can move away from zero is due to an external viral source.

Let's investigate the remaining variables, starting with (2.2). For all $t \geq t_f + \tau$,

$$\begin{aligned} \dot{P}(t) &= \alpha\xi(m(t))V(t-\tau)F(t-\tau) - \mu_p(P(t) - P^*) \\ &= -\mu_p(P(t) - P^*), \end{aligned}$$

where the second equality follows from Claim 1. This is an affine, stable, differential equation in $P(t)$; hence, P will asymptotically approach P^*, its healthy equilibrium value.

Next, move on to F, using (2.3) and Claim 1. For all $t \geq t_f$,

$$\begin{aligned} \dot{F}(t) &= \rho P(t) - (\mu_f + \eta\gamma V(t))F(t) \\ &= \rho P(t) - \mu_f F(t). \end{aligned}$$

This is a stable, linear, forced, differential equation in $F(t)$. Since it is already known that $P(t) \to P^*$ asymptotically, it follows that $F(t) \to \rho P^*/\mu_f$ asymptotically.

Again using Claim 1, (2.4) gives (for $t \geq t_f$)

$$\begin{aligned} \dot{m}(t) &= \sigma V(t) + \sigma_f F(t) - \mu_m m(t) \\ &= \sigma_f F(t) - \mu_m m(t), \end{aligned}$$

which is a stable, linear, forced, differential equation in m. Since it is already known that $F(t) \to \rho P^*/\mu_f$ asymptotically, it follows that $m(t) \to \frac{\sigma_f \rho P^*}{\mu_f \mu_m}$ asymptotically. □

A.5 Proof of Theorem 2.6 - Lower Bounds

Recall that this proof applies to the controlled equations

$$\dot{V}(t) = (\beta - u(t) - \gamma F(t))V(t), \tag{A.18}$$

$$\dot{P}(t) = \xi(m)\alpha V(t-\tau)F(t-\tau)v(t) - \mu_P(P(t) - P^*), \tag{A.19}$$

$$\dot{F}(t) = \rho P(t) - (\mu_f + \eta\gamma V(t))F(t), \tag{A.20}$$

$$\dot{m}(t) = \sigma V(t) - \mu_m m(t) + \sigma_f F(t) + \sigma_c u(t), \tag{A.21}$$

and that the control signals satisfy $u(t) \geq 0$ and $v(t) \in [0, 1]$.

We begin by investigating V:

Claim 1: $V(t) \geq 0 \; \forall \; t \geq 0$.

Proof:

This is a proof by contradiction; hence, assume that there exists t_1 such that $V(t_1) < 0$. Since $V_0 \geq 0$ and V is continuous, there exists a $t_2 \in (0, t_1)$ such that $V(t_2) = 0$. From (A.18) and the definition of t_2 we know that $\dot{V}(t_2) = 0$ and, therefore, $V(t) = 0$ for all $t \geq t_2$. In particular, $V(t_1) = 0$. This is a contradiction. □

The variables F and P are closely linked, and as such will be investigated together, in the following two claims:

Claim 2: If $F(t_a) < 0$ for some $t_a > 0$, then there exists a $t_b \in (0, t_a)$ and a $t_d \in (t_b, t_a)$ such that:

1. $F(t_b) = 0$,
2. $F(t) \geq 0$ for all $t \in [-\tau, t_b)$,
3. $F(t) < 0$ for all $t \in (t_b, t_d]$, and
4. $P(t_b) < 0$.

Proof:

Refer to Figure A.3 for the placement of t_a and t_b. Assume $F(t_a) < 0$, $t_a > 0$. The continuity of F and the fact that $F_0 > 0$ guarantee the satisfaction of Items 1 through 3. Then, from (A.20),

$$0 > \dot{F}(t_b) = \rho P(t_b) - (\mu_f + \eta\gamma V(t_b)) \cdot 0. \tag{A.22}$$

Since it is known that $\rho > 0$, we conclude that $0 > P(t_b)$. □

Claim 3: If $P(t_b) < P^*$ for some $t_b > 0$, then there exists a $t_c \in (0, t_b)$ and a $t_e \in (t_c, t_b)$ such that:

1. $P(t_c) = P^*$,
2. $P(t) \geq P^*$ for all $t \in [-\tau, t_c)$,
3. $P(t) < P^*$ for all $t \in (t_c, t_e]$, and

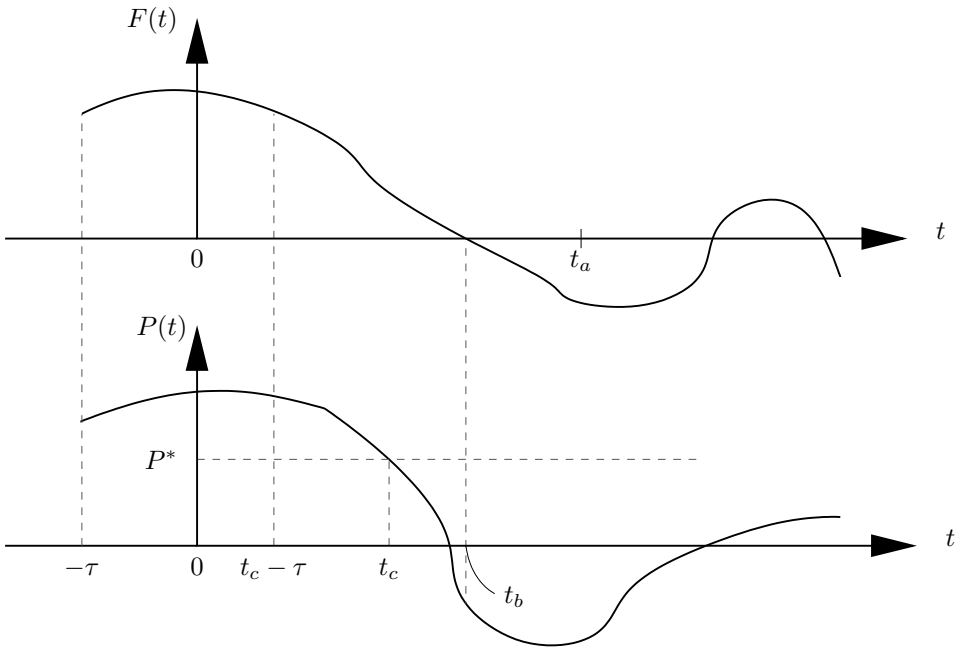

Figure A.3: Relationship Between P and F

4. $F(t_c - \tau) < 0$.

Proof:

Assume there exists a $t_b > 0$ such that $P(t_b) < P^*$. The continuity of P and the fact that $P_0 \geq P^*$ guarantee the satisfaction of Items 1 through 3. Then, from (A.19),

$$0 > \dot{P}(t_c) = \alpha\xi(m(t_c))V(t_c - \tau)F(t_c - \tau)v + \mu_P(P^* - P^*). \tag{A.23}$$

From Claim 1 we know that $V(t_c - \tau) \geq 0$ and (by definition) α and $\xi(m)$ are non-negative, so the above inequality implies that $F(t_c - \tau) < 0$. □

Claims 2 and 3 can be used to prove that, for all $t \geq 0$, $F(t) \geq 0$ and $P(t) \geq P^*$. For example, the following contradiction argument is used to show that $F(t) > 0$ (see Figure A.3):

- Assume that there exists a $t_a > 0$ such that $F(t_a) < 0$.
- From Claim 2, there exists a t_b that corresponds to the first time that F crosses the axis to become negative and, furthermore, $P(t_b) < 0 < P^*$.

- From Claim 3, there exists a $t_c < t_b$ that corresponds to the first time that P crosses the P^* line to become less than P^* and, furthermore, $F(t_c - \tau) < 0$.
- The contradiction is as follows: $t_c < t_b$, so $t_c - \tau < t_b$, but t_b is the *first* time that F crosses the axis to become negative, so F cannot be negative prior to t_b.
- We conclude that $F(t) \geq 0$ for all $t \geq 0$.

A similar argument can be made by starting with $P(t_a) < P^*$ to show that $P(t) > P^*$ for $t \geq 0$.

Finally, note that (A.21) is a linear differential equation whose only negative sign is associated with the variable m. Since all variables other than m are known to be positive and $m_0 < 1$, then it follows directly that $m(t) > 0$ for all $t \geq 0$. $\square$

A.6 Proof of Theorem 2.7 - Finite Escape Time

This proof is divided into three parts: investigating V, then F and P together, and finally m. Throughout this proof it is critical to note that all parameters and variables are nonnegative (Theorem 2.6).

The following Lemma will be required in the proof:

Lemma A.8 *For some vector* $z(t) = \left\| \begin{array}{c} z_1(t) \\ z_2(t) \end{array} \right\|$, *and any function* $a(t)$ *and constant* b, *the following inequality holds:*

$$\left\| \left[\begin{array}{c} a(t)z_2(\theta - \tau) \\ bz_1(\theta) \end{array} \right] \right\| \leq (a(t) + b)(2 \max_{\phi \in [0,\theta]} \|z(\phi)\| + \max_{\phi \in [0-\tau,0]} |z_2(\phi)|).$$

Proof

The following norm rule will be used throughout this proof:

$$|z_i| \leq \|z\| \leq \sum_i |z_i|.$$

Begin with the left-hand side of the inequality in the Lemma:

$$\begin{aligned}\left\|\begin{bmatrix} a(t)z_2(\theta-\tau) \\ bz_1(\theta)\end{bmatrix}\right\| &\leq \left\|\begin{bmatrix} a(t) & 0 \\ 0 & b\end{bmatrix}\right\|\left\|\begin{bmatrix} z_2(\theta-\tau) \\ z_1(\theta)\end{bmatrix}\right\| \\ &\leq (a(t)+b)\left\|\begin{bmatrix} z_2(\theta-\tau) \\ z_1(\theta)\end{bmatrix}\right\|.\end{aligned}$$

Using the norm rule outlined above it is possible to perform the following steps:

$$\begin{aligned}\left\|\begin{bmatrix} z_2(\theta-\tau) \\ z_1(\theta)\end{bmatrix}\right\| &\leq |z_1(\theta)| + |z_2(\theta-\tau)| \\ &\leq \max_{\phi\in[0,\theta]}|z_1(\phi)| + \max_{\phi\in[0,\theta]}|z_2(\phi-\tau)| \\ &\leq \max_{\phi\in[0,\theta]}\|z(\phi)\| + \max_{\phi\in[0,\theta]}|z_2(\phi-\tau)| \\ &\leq \max_{\phi\in[0,\theta]}\|z(\phi)\| + \max_{\psi\in[-\tau,\theta-\tau]}|z_2(\psi)| \\ &\leq \max_{\phi\in[0,\theta]}\|z(\phi)\| + \max_{\psi\in[-\tau,0]}|z_2(\psi)| + \max_{\psi\in[0,\theta-\tau]}|z_2(\psi)| \\ &\leq \max_{\phi\in[0,\theta]}\|z(\phi)\| + \max_{\psi\in[-\tau,0]}|z_2(\psi)| + \max_{\psi\in[0,\theta]}|z_2(\psi)| \\ &\leq 2\max_{\phi\in[0,\theta]}\|z(\phi)\| + \max_{\psi\in[-\tau,0]}|z_2(\psi)| \\ &= 2\max_{\phi\in[0,\theta]}\|z(\phi)\| + \max_{\phi\in[-\tau,0]}|z_2(\phi)|.\end{aligned}$$

Therefore,

$$\left\|\begin{bmatrix} a(t)z_2(\theta-\tau) \\ bz_1(\theta)\end{bmatrix}\right\| \leq (a(t)+b)(2\max_{\phi\in[0,\theta]}\|z(\phi)\| + \max_{\phi\in[-\tau,0]}|z_2(\phi)|).$$

□

Now investigate V:

Claim 1: $V(t) \leq e^{\beta t}V_0$ and is finite for all $t \in [0,\infty)$.

Proof of Claim 1:

From (2.9),

$$\begin{aligned}\dot{V}(t) &= (\beta - u(t) - \gamma F(t))V(t) \\ &\leq \beta V(t).\end{aligned}$$

It follows directly that

$$V(t) \leq e^{\beta t}V_0,\ t \in [0,\infty). \tag{A.24}$$

Since V is bounded above by a continuous function in the range $t \in [0, \infty)$ it must be finite in that range. $\square$

Claim 2: $P(t)$ and $F(t)$ are finite for all $t \in [0, \infty)$.

Proof of Claim 2:

From (2.10) and using Theorem 2.6, $\xi(m) \leq 1$, $v(t) \leq 1$, and Claim 1,

$$\begin{aligned} \dot{P}(t) &= \xi(m)\alpha V(t-\tau)F(t-\tau)v(t-\tau) - \mu_P(P(t) - P^*) \\ &\leq \alpha V_0 e^{\beta(t-\tau)} F(t-\tau) \\ &= \alpha V_0 e^{-\beta\tau} e^{\beta t} F(t-\tau). \end{aligned} \tag{A.25}$$

Similarly, from (2.11),

$$\begin{aligned} \dot{F}(t) &= \rho P(t) - (\mu_f + \eta\gamma V(t))F(t) \\ &\leq \rho P(t). \end{aligned} \tag{A.26}$$

Integrate (A.25) and (A.26) to obtain

$$P(t) = P_0 + \int_0^t \alpha V - 0 e^{-\beta\tau} e^{\beta t} F(\theta - \tau) d\theta, \tag{A.27}$$

$$F(t) = F_0 + \int_0^t \rho P(\theta) d\theta. \tag{A.28}$$

Defining

$$z(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} \triangleq \begin{bmatrix} P(t) \\ F(t) \end{bmatrix},$$

substituting (A.27) and (A.28) into z, and taking a p-norm yields

$$\|z(t)\| \leq \|z_0\| + \int_0^t \left\| \begin{bmatrix} \alpha V_0 e^{-\beta\tau} e^{\beta t} z_2(\theta - \tau) \\ \rho z_1(\theta) \end{bmatrix} \right\| d\theta. \tag{A.29}$$

If it were possible to solve this implicit inequality then we would be able to find a bound on $\|z\|$ and hence on P and F. We will apply the Gronwall-Bellman Inequality (GBI) [15] instead. The GBI requires that the inequality be of the form

$$y(t) \leq \lambda(t) + \int_{t_0}^t \nu(\theta) y(\theta) d\theta \tag{A.30}$$

for a constant a and b such that λ and $y : [a, b] \to \mathbf{R}$ are continuous, $\nu : [a, b] \to \mathbf{R}$ is continuous and nonnegative, and $a \leq t \leq b$.

The form of (A.29) is similar to (A.30), but is not quite the same, due primarily to the time delay in the z_2 term of the integrand. It will be shown that (A.29) can be manipulated to find an inequality in the form of (A.30) if the following substitutions are made:

$$\begin{aligned} y(t) &= \max_{\theta\in[0,t]} \|z(\theta)\|, \\ \lambda(t) &= \|z_0\| + \int_0^t (\alpha V_0 e^{-\beta\tau} e^{\beta\theta} + \rho) \max_{\theta\in[-\tau,0]} |z_2(\theta)| d\theta, \\ \nu(\theta) &= 2(\alpha V_0 e^{-\beta\tau} e^{\beta\theta} + \rho), \\ t_0 &= 0. \end{aligned}$$

Applying Lemma A.8 to the integrand of (A.29) results in

$$\|z(t)\| \leq \|z_0\| + \int_0^t (\alpha V_0 e^{-\beta\tau} e^{\beta\theta} + \rho)(2 \max_{\phi\in[0,\theta]} \|z(\phi)\| + \max_{\phi\in[-\tau,0]} |z_2(\phi)|) d\theta.$$

Note that the integrand is positive. Hence, the right hand side is strictly increasing and it is therefore possible to write

$$\max_{\phi\in[0,t]} \|z(\phi)\| \leq \|z_0\| + \int_0^t (\alpha V_0 e^{-\beta\tau} e^{\beta\theta} + \rho)(2 \max_{\phi\in[0,\theta]} \|z(\phi)\| + \max_{\phi\in[-\tau,0]} |z_2(\phi)|) d\theta, \quad \text{(A.31)}$$

which is in the form required to apply the GBI under the substitutions outlined in (A.31). When the GBI is applied, the following inequality is obtained

$$y(t) \leq \lambda(t) + \int_0^t \lambda(\theta)\nu(\theta) \exp\left(\int_\theta^t \nu(\phi) d\phi\right) d\theta, \quad \text{(A.32)}$$

which, from the definitions of λ and ν, is finite for each $t \geq 0$. Therefore, for each $t \geq 0$,

$$\max_{\phi\in[0,t]} \|z(\phi)\| = y(t) < \infty.$$

Hence, an upper bound on $\|z(t)\|$ exists for all $t \in [0, \infty)$ which implies that both $P(t)$ and $F(t)$ are finite for all $t \in [0, \infty)$. □

Finally, we investigate $m(t)$. The final equation in the system, (2.12), is linear time-invariant, as can be seen below:

$$m(t) = \sigma V(t) + \sigma_f F(t) + \sigma_c u(t) - \mu_m m(t).$$

A property of linear time-invariant differential equations is that, assuming the forcing function is finite for $t \in [0, \infty)$, the state is finite for all $t \in [0, \infty)$. From Claims 1 and 2 it is known that V and F are finite in this time range. The control input, u, is also finite. Hence, $m(t)$ is finite for all $t \in [0, \infty)$. □

A.7 Proof of Theorem 4.5 - Sufficient Conditions for Health

From (2.9),

$$\begin{aligned} \dot{V}(t) &= (\beta - u(t) - \gamma F(t))V(t) \\ &\leq (\beta - \check{U} - \gamma \check{F})V(t), \end{aligned}$$

which implies

$$V(t) \leq e^{(\beta - \check{U} - \gamma \check{F})(t)} V_0,$$

and

$$\begin{aligned} V(t_f) &\leq e^{(\beta - \check{U} - \gamma \check{F})t_f} V_0 \\ &\leq e^{(\beta - \check{U} - \gamma \check{F})T_{h2}} V_0 \\ &= e^{(\beta - \check{U} - \gamma \check{F})\frac{\ln(V^*/V_0)}{\beta - \check{U} - \gamma \check{F}}} V_0 \\ &= V^*, \end{aligned}$$

where the second inequality follows from $\check{U} > \beta - \gamma \check{F}$ and $t_f > T_{h2}$. □

A.8 Proof of Theorem 4.6 - Necessary Conditions for Death

From (2.12),

$$\begin{aligned} \dot{m}(t) &= \sigma V(t) - \mu_m m(t) + \sigma_f F(t) + \sigma_c u(t) \\ &\leq \sigma \hat{V} + \sigma_f \hat{F} + \sigma_c \hat{U} - \mu_m m(t), \end{aligned}$$

which implies

$$1 < m(t) \leq w + (m_0 - w_{d1}) e^{-\mu_m(t)}.$$

Since $m_0 < 1$ and $w_{d1} > 1$ (Theorem 4.1), this implies

$$t_f \geq \frac{-1}{\mu_m} \ln\left(\frac{1 - w_{d1}}{m_0 - w_{d1}}\right). \tag{A.33}$$

□

A.9 Proof of Theorem 4.7 - Sufficient Conditions for Death

From (2.12),

$$\begin{aligned} \dot{m}(t) &= \sigma V(t) - \mu_m m(t) + \sigma_f F(t) + \sigma_c u(t) \\ &\geq \sigma_c \check{U} - \mu_m m(t), \end{aligned}$$

which implies

$$m(t) \geq \frac{\sigma_c \check{U}}{\mu_m} + \left(m_0 - \frac{\sigma_c \check{U}}{\mu_m}\right) e^{-\mu_m(t)};$$

therefore,

$$\begin{aligned} m(t_f) &\geq \frac{\sigma_c \check{U}}{\mu_m} + \left(m_0 - \frac{\sigma_c \check{U}}{\mu_m}\right) e^{-\mu_m(t_f)} \\ &\geq \frac{\sigma_c \check{U}}{\mu_m} + \left(m_0 - \frac{\sigma_c \check{U}}{\mu_m}\right) e^{-\mu_m(T_{d2})} \\ &\geq \frac{\sigma_c \check{U}}{\mu_m} + \left(m_0 - \frac{\sigma_c \check{U}}{\mu_m}\right) \exp\left\{-\mu_m \left[\frac{-\ln\left(\frac{1-\frac{\sigma_c \check{U}}{\mu_m}}{m_0 - \frac{\sigma_c \check{U}}{\mu_m}}\right)}{\mu_m}\right]\right\} \\ &= \frac{\sigma_c \check{U}}{\mu_m} + \left(m_0 - \frac{\sigma_c \check{U}}{\mu_m}\right)\left(\frac{1 - \frac{\sigma_c \check{U}}{\mu_m}}{m_0 - \frac{\sigma_c \check{U}}{\mu_m}}\right) \\ &= \frac{\sigma_c \check{U}}{\mu_m} + \left(1 - \frac{\sigma_c \check{U}}{\mu_m}\right) \\ &= 1, \end{aligned}$$

where the second inequality follows from $t_f \geq T_{d2}$ and the third inequality follows from the assumption $\frac{\sigma_c u}{\mu_m} > 1$. □

Appendix B

Important Terms

Important Words and Concepts

Alive - Refers to mass damage, occurs if $m(t) < 1$ for all $t \in [0, t_f]$, Pg. 2
'check' - $\check{x}$ - Indicates a lower bound, Pg. 16
Dead - Not alive, Pg. 2
'hat' - $\hat{x}$ - Indicates an upper bound, Pg. 16
Healthy - Refers to viral levels, occurs at time t if $V(t) < V^*$, Pg. 2
Sick - Refers to viral levels, occurs at time t if $V(t) \geq V^*$, Pg. 2.

Important Variables and Parameters

Note: All Greek letters first appear in (2.1)-(2.4) and are described on pages 7-8 unless stated otherwise.

$a - g$ - Parameter substitutions for chronic stability theorems, Pg. 13 and (2.22)-(2.26)
F - Antibody concentration, Pg. 1 and (2.3)
m - Mass damage term, relates to the ability of the body to produce plasma, Pg. 1 and (2.4)
P - Plasma concentration, Pg. 1 and (2.2)
P^* - Minimum number of plasma cells in the body, Pg. 8 and (2.2)
t_1 - Switching time, Pg. 26 and (4.1)-(4.2)
t_f - Time at which control is stopped, Pg. 3

U - Magnitude of step drug cocktail input, Pg. 26 and (4.2)
u - Drug cocktail input, Pg. 8-9 and (2.7)-(2.8)
V - Virus concentration, Pg. 1 and (2.1)
V^* - max {level below which model breaks down due to quantization error, minimum measurable concentration of V}, Pg. 2
v - Immune inhibitor input, Pg. 9 and (2.10)
α - Probability of an antibody encountering a virus such that production of plasma cells is stimulated
β - Inverse time constant of virus (or antigen)
γ - Probability of an antigen/antibody encounter that results in antigen neutralization
η - Number of antibody cells required to neutralize one antigen cell; this is a natural number
μ_f - Inverse of average life span of antibody cells
μ_m - Inverse of recuperation period of m
μ_p - Inverse of average life span of plasma cells
ρ - Rate of production of antibodies due to one plasma cell
σ - Rate of damage due to one viral cell
σ_c - Scalar effect of damage due to u, Pg. 9 and (2.8)
σ_f - Rate of damage due to one antibody cell
τ - Approximates delay associated with generation of memory cells and other dynamic components otherwise ignored in this model.

Bibliography

[1] A. L. Asachenkov, G. I. Marchik, R. R. Mohler, and S. M. Zuev, "Immunology and disease control: A systems approach," *IEEE Transactions on Biomedical Engineering*, vol. 41, pp. 943–953, Oct. 1994.

[2] A. E. Rundell, R. DeCarlo, V. Balakrishnan, and H. HogenEsch, "Systematic method for determining intravenous drug treatment strategies aiding the humoral immune response," *IEEE Transactions on Biomedical Engineering*, vol. 45, pp. 429–439, Apr. 1998.

[3] A. L. Asachenkov, G. I. Marchik, R. R. Mohler, and S. M. Zuev, *Disease Dynamics*, ch. 4. Birkhauser, 1994.

[4] J. K. Hale, *Functional Differential Equations*, pp. 3–4. Springer-Verlag, 1971.

[5] R. Zurakowski and A. Teel, "Enhancing immune response to HIV infection using MPC-based treatment scheduling," *Proceedings of the 2003 American Control Conference*, vol. 2, pp. 1182–1187, June 2003.

[6] C. Bruni, G. DiPillo, and G. Koch, "Bilinear systems: An appealing class of 'nearly linear' systmes in theory and applications," *IEEE Transactions on Automatic Control*, vol. AC-19, pp. 334–348, Aug. 1974.

[7] M. Frayman, "On the relationship between bilinear and quadratic systems," *IEEE Transactions on Automatic Control*, pp. 567–568, Aug. 1975.

[8] A. S. Perelson, M. Mirmirani, and G. F. Oster, "Optimal strategies in immunology, i: B-cell differentiation and proliferation," *Journal of Math and Biology*, vol. 3, pp. 325–367, 1976.

[9] H. K. Khalil, *Nonlinear Systems*, ch. 12. Prentice Hall, third ed., 2002.

[10] H. K. Khalil, *Nonlinear Systems*, ch. 13. Prentice Hall, third ed., 2002.

[11] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, pp. 789–814, June 2000.

[12] R. Bellman, *Dynamic Programming*. Princeton University Press, 1957.

[13] K. Gu, V. L. Kharitonov, and J. Chen, *Stability of Time-Delay Systems*, pp. 18–19. Birkhauser, 2003.

[14] M. Gopal, *Control Systems Principles and Design*, pp. 943 – 945. Tata McGraw-Hill, 1997.

[15] H. K. Khalil, *Nonlinear Systems*, pp. 651–652. Prentice Hall, third ed., 2002.