2007-03-20

The Importance of the Riemann-Hilbert Problem to Solve a Class of Optimal Control Problems

Nicholas Dewaal
Brigham Young University - Provo

THE IMPORTANCE OF THE RIEMANN-HILBERT PROBLEM TO SOLVE A CLASS OF OPTIMAL CONTROL PROBLEMS

by

Nicholas DeWaal

A project submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Science

Department of Mathematics

Brigham Young University

April 2007

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a project submitted by

Nicholas DeWaal

This project has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	Scott A. Glasgow, Chair
Date	Christopher P. Grant
Date	David Glen Whiting

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the project of Nicholas DeWaal in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

Scott A. Glasgow
Chair, Graduate Committee

Accepted for the Department

Gregory R. Conner
Graduate Coordinator

Accepted for the College

Thomas Sederberg, Associate Dean
College of Physical and Mathematical Sciences

ABSTRACT

THE IMPORTANCE OF THE RIEMANN-HILBERT PROBLEM
TO SOLVE A CLASS OF OPTIMAL CONTROL PROBLEMS

Nicholas DeWaal

Department of Mathematics

Master of Science

Optimal control problems can in many cases become complicated and difficult to solve. One particular class of difficult control problems to solve are singular control problems. Standard methods for solving optimal control are discussed showing why those methods are difficult to apply to singular control problems. Then standard methods for solving singular control problems are discussed including why the standard methods can be difficult and often impossible to apply without having to resort to numerical techniques. Finally, an alternative method to solving a class of singular optimal control problems is given for a specific class of problems.

Contents

The Importance of the Riemann-Hilbert Problem to Solve a Class of Optimal Control Problems

A Riemann-Hilbert problem is a type of mathematics problem in the complex plane that is often useful for various applications in applied mathematics. Details of this problem will be presented later. This paper will show how a certain type of Riemann-Hilbert problem is useful for solving a class of problems in control theory that tend to be difficult or even impossible to solve using standard control theory techniques. Examples are shown why standard control theory techniques can be very challenging to use on the class of control problems to be discussed.

By assuming that $x : \mathbb{R} \to \mathbb{R}^n$ (or possibly $\mathbb{C}^n$), the following notation will be used in this document: $\dot{x} := \frac{dx(t)}{dt}$, x' denotes the transpose of x, $\bar{x}$ is the complex conjugate of x, x^* is the conjugate transpose of x, and x° is a function that is optimal in the context of the given situation. The Fourier Transform of a function f will be denoted $\widehat{f}$ where the Fourier transform used is $\widehat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{i\omega t}dt$, and the inverse fourier transform is denoted $f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{f}(\omega)e^{-i\omega t}d\omega$.

One important optimal control problem is that of applying electric fields to controlling various states of the changing polarization fields of different types of dielectrics. The goal is to choose the electric field as a function of time that is applied to the dielectric such that the energy to go from the initial state to the final desired state is minimized.

The single Lorentz oscillator with the controller being the choice of a function of an electric field applied to a dielectric satisfies the differential equation

$$\ddot{P} + \gamma \dot{P} + \omega_0^2 P = \omega_p E(t), \tag{1}$$

where $E(t)$ is the control function. In order to write this equation in the typical notation used in control theory, define

$$x_1 := P, x_2 := \dot{P}, \text{ and } u(t) := E(t).$$

This allows equation (1), to be written as a system of first order differential equations

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_p \end{bmatrix} u. \tag{2}$$

By choosing the control $u(t)$ and starting at the initial state $x_1(0), x_2(0)$, the goal may be to arrive at a prescribed state at time T namely $x_1(T), x_2(T)$. Of course, for some control problems governed by $\dot{x} = f(x, u, t)$, achieving that objective may not even be possible.

Often the set of allowed controls for $u(t)$ are restricted to a set of allowed values defined by some set U. Given a target set Υ and a control constrained to $u(t) \in U$ for all t, the controllable set for Υ is defined as the set of all initial points C such that Υ can be reached in some finite time. A reachable set from any given starting set X is defined as the set of all points that can be achieved in finite time when starting in set X. It would be ideal if every point were in the reachable set of every given point. For equations of the form $\dot{x} = Ax + Bu$ where A and B are t independent constant matrices and $u(t)$ is the control, the following theorem holds [1]:

Theorem 1. *[2] Given that $U = \mathbb{R}^{n_u}$, and given any initial starting point $x \in \mathbb{R}^n$, then the controllable set of x is $\mathbb{R}^n$ if and only if $C := [B|AB|A^2B \dots A^nB]$ is full rank.*

Hence, the controllable set for all $x \in \mathbb{R}^2$ is $\mathbb{R}^2$ for a system governed by equation

(2) *because*

$$C = \left[\begin{bmatrix} 0 \\ \omega_p \end{bmatrix} \Big| \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{bmatrix} \begin{bmatrix} 0 \\ \omega_p \end{bmatrix} \Big| \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{bmatrix}^2 \begin{bmatrix} 0 \\ \omega_p \end{bmatrix}\right]$$
$$= \begin{bmatrix} 0 & \omega_p & -\gamma\omega_p \\ \omega_p & -\gamma\omega_p & (-\omega_0^2 + \gamma^2)\omega_p \end{bmatrix}.$$

This matrix can easily be checked to have a rank of two for nonzero γ, ω_p, and ω_0 making the controllable set for all $x \in \mathbb{R}^2$ be $\mathbb{R}^2$ for (2).

Due to the fact that this system is controllable, and because the equation is linear, calculating any needed control for it can usually be done. However, sometimes serious problems can arise when trying to find a control that minimizes certain cost functions. If your goal is to perform the control on a matrix differential equation of the form

$$\dot{x} = Ax + Bu \tag{3}$$

where $[B|AB|A^2B \ldots A^nB]$ is full rank, while trying to minimize

$$J[u] = \int_0^T (x'Cx + u'Du)\, dt + x(T)'Mx(T) \tag{4}$$

where C, D, and M are positive definite matrices, then the optimal control can be found that gives the desired final state via the Hamilton-Jacobi-Bellman equation which is discussed in detail later.

More generally speaking, the problem can be defined by having the goal to minimize the cost function

$$J[u] = \int_{t_0}^T l(x(\tau), u(\tau), \tau)d\tau + m[x(T)]$$

constrained to the initial value problem $\dot{x} = f(x, u(t), t)$ and $x(t_0) = a$, where $u(t)$ is the control.

Deciding how close to a prescribed point at time T you want to arrive is decided by your choice of the function $m[x(T)]$, which is a penalty function that adds more to the cost function when final states $x(T)$ are undesirable. More on this will be discussed later. By defining the function $V^\circ(x,t)$ as the minimal possible cost to go, that is the cost from starting at state $x(t) \in \mathbb{R}^n$ and time t, and ending at time T, then

$$V^\circ(x,t) = \int_t^T l(x(\tau), u^\circ(\tau), \tau) d\tau + m[x(T)].$$

It can be shown [3] that the minimum cost to go function $V^\circ(x,t)$ with the given differential equation as the constraint satisfies the Hamilton-Jacobi-Bellman equation

$$\frac{-\partial V^\circ(x,t)}{\partial t} = \min_{u \in \mathbb{R}^{n_u}} \left\{ l(x,u,t) + \left[\frac{\partial V^\circ}{\partial x}\right]' f(x,u,t) \right\},$$

and final condition

$$V^\circ(x,T) = m(x)$$

for all x.

However, the goal is to find the optimal control $u^\circ(t)$ that gives us the minimum cost to go: $V^\circ(x,t)$. It has also been shown [3] that the $u(t)$ that minimizes the right hand side of the Hamilton-Jacobi-Bellman equation is also the optimal control that minimizes $J[u]$. Therefore, the first step to solving this problem is to first do the minimization with respect to u which leads to u being some function in the following form:

$$u^\circ = \psi\left(\frac{\partial V^\circ}{\partial x}, x, t\right). \tag{5}$$

Then after finding this form of u°, it can be substituted back into the Hamilton-Jacobi-Bellman equation making it then possible to solve for V°. Then after solving for V°, you can then easily find $\frac{\partial V^\circ}{\partial x}$, and then substitute that to find

$$u^\circ(x,t) = \psi\left(\frac{\partial V^\circ(x,t)}{\partial x}, x, t\right) =: \phi(x,t).$$

You can now solve $\dot{x} = f(x,u,t)$ as $\dot{x} = f(x, \phi(x,t), t)$ which then allows the representation $u^\circ(t) = \phi(x(t), t)$.

For example, assume that our goal is to minimize the integral in (4) constrained to (3), then we begin finding the optimal control using the Hamilton-Jacobi-Bellman equation.

This long process can easily introduce many difficult nonlinearities for both ordinary differential equations and partial differential equations that can make the problem difficult and lead to a need for numerical techniques. One type of problem that is commonly used in control applications is the specific problem mentioned earlier that specifies

$$l(x,u) = x'Qx + u'Ru, \tag{6}$$

$$m(x) = x'Mx, \tag{7}$$

and

$$f(x,u) = Ax + Bu.$$

The first step to solving this control problem is to find the minimization of

$$x'Qx + u'Ru + \left[\frac{\partial V^\circ}{\partial x}\right]'(Ax + Bu) \tag{8}$$

with respect to u just as in equation 5. By taking the derivative of (8) with respect to u, and setting it 0 to get

$$\left(\left[\frac{\partial V^\circ}{\partial x}\right]' B\right)' + 2Ru = 0. \tag{9}$$

We can now solve for u° to get

$$u^\circ = -\frac{1}{2}R^{-1}B'\frac{\partial V^\circ}{\partial x}. \tag{10}$$

Now we need to find the form of V° to find u°. Substituting (10) into the Hamilton-Jacboi-Bellman equation we get the PDE and final condition

$$-\frac{\partial V^\circ}{\partial t} = x'Qx + \frac{1}{4}\left(R^{-1}B'\frac{\partial V^\circ}{\partial x}\right)' RR^{-1}B'\frac{\partial V^\circ}{\partial x} + \left[\frac{\partial V^\circ}{\partial x}\right]' (Ax - \frac{1}{2}BR^{-1}B'\frac{\partial V^\circ}{\partial x}), \tag{11}$$

and

$$V^\circ(x, T) = x'Mx. \tag{12}$$

Solving complicated PDE's is outside the scope of this project, but it can be shown [4] that after simplification and other analysis, the solution to (11) and (12) is

$$V^\circ(x, t) = x'P(t)x, \tag{13}$$

where $P(t)$ is the solution to the matrix ODE

$$-P' = A'P + PA + Q - PBR^{-1}B'P, \tag{14}$$

$$P(T) = M. \tag{15}$$

This makes the optimal control $u^\circ(t) = -\frac{1}{2}R^{-1}B'\frac{\partial V^\circ}{\partial x} = -\frac{1}{2}R^{-1}B'(2P(t)x) = -R^{-1}B'P(t)x$.

However, for the optimal control problem involving equation (2) the goal is to start at the given initial state

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}$$

and end at the prescribed state $\begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$ at time T, that is

$$\begin{bmatrix} x_1(T) \\ x_2(T) \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$$

all while minimizing the integral

$$J[u] = \int_0^T u x_2 dt.$$

Sometimes it is possible to apply the Hamilton-Jacobi-Bellman method to solve optimal control problems that require arrival at a prescribed final state at time T. Of course the process for Hamilton-Jacobi-Bellman requires the minimization of a cost function of the form:

$$V = \int_t^T l(x, u, \tau) d\tau + m[x(T)],$$

for a nonzero function $m : \mathbb{R}^{n_x} \to \mathbb{R}$. The function $m[x(T)]$ is put into the cost function to be a positive function that is larger (greater cost) for final state values that are further away from the desired set of states. This penalizes controls $u(t)$ that give final states far from desired states. However, the class of control problems we are interested in solving in this project involves no leniency for approximate arrival to a final state and so $x(T)$ must be the prescribed final state. Such is the case for the example problem specified earlier to control equation (2) which we recall has the cost function

$$J[u] = \int_0^T ux_2 dt + m[x(T)]$$

where m is the zero function. This is because the cost functions we want does not incorporate a function $m[x(T)]$. However, if you let $m_j[x(T)]$ be a sequence of smooth functions that are zero for the prescribed final state, and approach infinity everywhere else as j goes to infinity, then the resulting control for each j that minimizes the cost to go function V is the same in the limit as for a cost to go function V with no $m[x(T)]$ and a prescribed final state. The main idea behind this is that the cost of being anywhere but the prescribed final state is infinity which forces the solution to the prescribed state while minimizing the integral part of the cost function. A good choice of m that satisfies this property is $m_j[x(T)] = j(b-x(T)')Q(b-x(T))$ where Q is a positive definite matrix, and b is the desired final state. If the problem is solved with this choice of $m_j[x(T)]$, then the limit can be taken as j approaches infinity. This forces the cost function to approach infinity except for at the exact destination that the final condition prescribes at time T. Hence in the limit, the control should converge.

The Hamilton-Jacobi-Bellman equation and final condition for problem (2) with the choice that $m_j[x(T)] = j(b - x(T)')Q(b - x(T))$ is therefore

$$\frac{-\partial V^\circ(x,t)}{\partial t} = \min_{u \in \mathbb{R}^{n_u}} \left\{ ux_2 + \left[\frac{\partial V^\circ}{\partial x}\right]' \left(\begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_p \end{bmatrix} u \right) \right\}$$

and

$$V^\circ(x,T) = x' jQx$$

for all x.

This integral presents a nonlinearity between u and x that complicates matters significantly for finding an optimal control by using the Hamilton-Jacobi-Bellman equation. The reason that it is a difficulty is because on the right hand side of the equation, the minimum for $u(t)$ is an infinite value. As will be shown later, this is a singular control problem. I tried to overcome this barrier by taking the various parts of the Hamilton-Jacobi-Bellman equation, and inserting some sort of nonlinearity times ϵ that could be taken in the limit to go to zero when everything was done. However, no such nonlinearity gave a Hamilton-Jacobi-Bellman equation that wouldn't have had to be solved numerically.

Here singular optimal control problems will be approached using other techniques commonly used in control theory, and discussion will include why those methods can be difficult or impossible for solving many singular control problem. This then motivates the need for control theorists to consider more seriously the importance of the Riemann-Hilbert problem, which will been shown to solve an important class of problems that include an important class of singular control problems.

1 Approach Using The Hamiltonian

Another approach can be applied to the optimal control for $\dot{x} = f(x, u, t)$ while minimizing the cost J in the different general form of

$$J[u] = \int_{t_0}^{T} l(x(\tau), u(\tau), \tau)d\tau + m[x(T), T]$$

with the final state constraint that defines the target set as

$$\psi(x(T), T) = 0. \tag{16}$$

Equation (16) is usually represented in the form $x(T) - b = 0$ meaning that the final state is b.

If the Hamiltonian is defined as the function $H : \mathbb{R}^{n_x+n_u+1} \to \mathbb{R}$ where

$$H(x,u,t) := l(x,u,t) + \lambda(t)' f(x,u,t)$$

$$t \geq t_0,$$

where

$$-\dot{\lambda} = \frac{\partial H(x,u,t)}{\partial x}$$

$$t \leq T,$$

then the optimal control $u : \mathbb{R} \to \mathbb{R}^{n_u}$ that minimizes $J[u]$ can be shown [5] to be the $u(t)$ that when substituted into $\frac{\partial H(x,u,t)}{\partial u}$ along with the optimally controlled $x^{\circ}(t)$ satisfies the stationary condition on u:

$$0 = \frac{\partial H(x,u,t)}{\partial u}.$$

Solving this stationary condition for u will give a relationship $u = \phi(x,t)$. Then by plugging $\phi(x,t)$ in the place of u in $\dot{x} = f(x,u,t) = f(x,\phi(x,t),t)$, $x(t)$ can be found and substituted to find $u^{\circ}(t) = \phi(x(t),t)$. And of course using the restriction of a prescribed starting state: $x(t_0) = a$.

For example [6], suppose that the goal is to find the shortest distance between two points $x(a) = A$, $x(b) = B \in \mathbb{R}^2$. From calculus it is known that the length of a curve is

$$L = \int_a^b \sqrt{1 + \dot{x}(t)^2} dt.$$

In order to make this an optimal control problem, it is written in the simple form $\dot{x} = u$.

This then makes

$$J[u] = \int_a^b \sqrt{1+u^2}dt,$$

the Hamiltonian being

$$H = \sqrt{1+u^2} + \lambda u.$$

So now

$$0 = \frac{\partial H}{\partial u} = \lambda + \frac{u}{\sqrt{1+u^2}},$$

and hence

$$-\dot{\lambda} = \frac{\partial H}{\partial x} = 0.$$

Therefore $\lambda(t)$ is a constant which then after solving for u in terms of λ forces u to also be a constant. Because $\dot{x} = u$, $x(t) = c_1 t + c_2$. By using the boundary condition it is easily shown that

$$x(t) = \frac{(A-B)t + (aB - bA)}{a-b},$$

which of course is a straight line as it ought to be.

Now it is possible to try and apply these equations to solve the optimal control problem that applies to equation (2).

Again for the optimal control to be performed on problem (2),

$$l(x,u,t) = ux_2,$$

$$m(x(T),T) = 0,$$

and

$$\psi(x(T),T) = x(T) - \begin{bmatrix} a \\ b \end{bmatrix} = 0$$

where $x(T) = \begin{bmatrix} a \\ b \end{bmatrix}$ is the desired final state of the system at T.

In this case the Hamiltonian is

$$H = ux_2 + \begin{bmatrix} \lambda_1(t) & \lambda_2(t) \end{bmatrix} \left(\begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\gamma \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \omega_p \end{bmatrix} u \right)$$

$$= ux_2 + \begin{bmatrix} \lambda_1(t) & \lambda_2(t) \end{bmatrix} \begin{bmatrix} x_2 \\ -\omega_0^2 x_1 - \gamma x_2 + \omega_p u \end{bmatrix}$$

$$= ux_2 + \lambda_1 x_2 + \lambda_2(-\omega_0^2 x_1 - \gamma x_2 + \omega_p u).$$

The derivation of the equations that are used to find $u(t)$ was done under the assumption that the goal is to minimize the Hamiltonian. However, the Hamiltonian is linear in u which means that when calculating $\frac{\partial H}{\partial u}$, u drops out making the method outlined earlier for finding the optimal u break down.

Singular control is defined as a control problem where the derivative of the Hamiltonian with respect to u does not depend on u. Just as with the Hamilton-Jacobi-Bellman equation, this example singular control problem has caused issues in finding an optimal control using the Hamiltonian approach.

2 Bang-Bang and Bang-off-Bang Control

The approach that is commonly used for singular control problems is bang-bang or bang-off-bang control. Most practical problems have a bound on the range that is allowed for the control. However, sometimes it is useful to solve the problem without a constraint on u. As will be seen, methods for bang-off-bang control are centered on some type of bounds on u. In order to solve a problem without bounds, a problem should be solved using bang-off-bang first where bounds are defined by $-a \leq u(t) \leq a$. Then in the end after finding $u_a(t)$, the limit as a approaches ∞ of $u_a(t)$ should converge to the control that has no constraint. In order to learn how to use bang bang and bang-off-bang control, some important theorems and definitions need to first be presented.

The set of allowed values for the control is defined by the set

$$U := \{u \in \mathbb{R}^m : h(u) \geq 0\},$$

where $h : \mathbb{R}^m \to \mathbb{R}^{n_h}$ for some n_h. Here stating that a vector is greater than 0 means that each row is greater than 0.

The terminal set is the set of any final states that would achieve the desired objective which for some g satisfies will be restricted to the following set

$$x(T) \in X := \{x \in \mathbb{R}^n : g(x) = 0\}$$

for a given function $g : \mathbb{R}^n \to \mathbb{R}^{n_g}$.

Here the objective is to minimize a less general form of integral as

$$J[u] = \int_0^T l(x(\tau), u(x(\tau)))d\tau$$

constrained to $\dot{x} = f(x, u)$, and some initial state x_0. The theory to be presented requires an autonomous ODE $\dot{x} = f(x, u)$, instead of $\dot{x} = f(x, u, t)$. Being autonomous and having l depend only on x and u will allow the optimal control to only depend on the state x making u strictly a function of x, namely $u(x)$ which because x is a function of t, u can also be considered a function of t after substitution. That is, after solving for $x(t)$ from $\dot{x} = f(x, u(x))$, substitution gives $u(x(t))$.

Because the goal is to minimize the cost $J[u]$, the problem can be transformed to minimizing the accumulated cost up to time t being

$$V_t\,(u) := \int_0^t l(x(\tau), u(x(\tau)))d\tau.$$

Minimizing $V_t\,(u)$ for each t is equivalent to minimizing $J[u]$ because l does not depend on t, and the system is autonomous. One step further can be taken to say that $\widetilde{x}(t) := \widetilde{V}_t(u(x(t)))$ in order to write

$$\dot{\widetilde{x}}(t) = l(x(t), u(x(t))).$$

This then allows for the system to be augmented to

$$\frac{d}{dt}\begin{bmatrix}\widetilde{x}(t)\\ x(t)\end{bmatrix} = \begin{bmatrix}l(x,u)\\ f(x,u)\end{bmatrix} =: F(X,u).$$

Here $\widetilde{x}(t)$ is the accumulation of cost over time, and $X = \{\widetilde{x}, x_1, x_2 \ldots\}$ making an $n+1$ dimensional system of equations. It is legitimate to assume that $V_{t=0}(u) = 0$ due to the fact that integration over an empty interval is 0. The goal can now be stated as choosing $u(x)$ that minimizes $\widetilde{x}(T)$ where $x(T) \in U$.

Of course, the point-wise choice of $u(x)$ is restricted not only by U, but it should also be piecewise continuous, and piecewise differentiable that yields a unique forward time solution to $\dot{x} = f(x,u)$. It must also be possible to reach the objective target before considering an optimal control. Note that $J[u]$ is minimized when $J[u^\circ(x)] \leq J[u(x)]$ for all admissible and feasible u.

Assuming that the function H is defined similarly to the Hamiltonian earlier, but assuming a bit more generality, H is defined as

$$H(x, u, \lambda, \lambda_0) := \lambda_0 l(x,u) + \lambda' f(x,u),$$

where for some existing $\rho \in \mathbb{R}^{n_g}$, $\lambda(t)$ satisfies

$$\dot{\lambda} = -\frac{\partial H}{\partial x} - \frac{\partial H}{\partial u}\frac{\partial u}{\partial x},$$

and

$$\lambda(T) = \rho' \frac{\partial g}{\partial x}|_{x=x(T)}.$$

It has been shown [7] that the function $u(x)$ that minimizes H point-wise is equivalent to minimizing $J(u(x))$ point-wise, which can be shown to minimize the final objective. This technique differs from techniques used in variational calculus. This information prepares for the following theorems:

Theorem 2. *[8] If $H(u^\circ(x)) \leq H(u(x))$ point-wise in x $\forall$ u admissible, then $\exists \gamma \in \mathbb{R}^{n_h}$ such that for*

$$L(x, u, \lambda, \lambda_0, \gamma) := H(x, u, \lambda, \lambda_0) - \gamma' h(u)$$

the following properties hold:

$$\frac{\partial L}{\partial u}|_{u=u^\circ} = 0$$

$$h_i(u^\circ) \geq 0 \quad \forall\ i = 1, .., n_h$$

$$\gamma' h(u^\circ) = 0$$

$$\gamma_j \geq 0 \quad \forall\ j$$

Theorem 3. *[8] If u° is an optimal control with respect to minimizing $J[u]$, and u° is admissible, satisfying the constraint U point-wise, then $\exists$ $\lambda : \mathbb{R} \rightarrow \mathbb{R}^n$ piecewise differentiable and continuous, and $\exists$ $\lambda_0 \geq 0$ (1 or 0) such that $(\lambda', \lambda_0) \neq 0$ and*

$\exists\ \rho \in \mathbb{R}^{n_g}$ such that λ satisfies

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial x}|_{u=u^\circ},$$

such that at the terminal set

$$g(x(T)) = 0,$$

$\lambda(T)$ satisfies the condition

$$\lambda(T) = \rho' \frac{\partial g}{\partial x}|_{x=x(T)},$$

and such that H has a global minimum at every point $x(t)$ caused by u° with respect to all admissible u. Finally,

$$0 = H[(x(t), u^\circ[x(t)], \lambda(t), \lambda_0] = \min_{u \in U} H[(x(t), u, \lambda(t), \lambda_0]$$

for all $t \in [0, T]$.

Note that in Theorem 2, γ can be thought of as a vector of lagrange multipliers.

3 Example of Bang-off-Bang Control [9]

The amount of a crop x in a farmers field satisfies

$$\dot{x} = \frac{r}{K}x(K - x) - ux$$

where K is the carrying capacity, r is the intrinsic growth rate, and $u(t)$ is the rate of harvesting, making it the control. The goal is to optimize the amount of harvest over a given time period $[0, T]$. In other words, the goal is to maximize

$$\int_0^T u(t)x(t)\, dt.$$

However, Theorems 2 and 3 are centered on minimization which means that the goal is to minimize the negative of the integral we want to maximize. Hence $J[u]$ is then

$$J[u] = -\int_0^T u(t)x(t)\,dt.$$

The harvest effort can never be less than 0, and can never be greater than some maximum effort= u_{max}.

The control constraint function $h(x)$ must be defined such that the allowed controls are restricted by the set

$$U := \{u \in \mathbb{R}^m : h(u) \geq 0\}.$$

Hence, $h(x)$ is defined as

$$h_1(u) = u \geq 0$$

$$h_2(u) = -u + u_{max} \geq 0.$$

In this problem, the Hamiltonian is

$$H = -ux + \lambda \left[\frac{r}{K} x(K - x) - ux \right],$$

which means that

$$L = H - \gamma_1 u - \gamma_2(u_{max} - u)$$

for some $\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}$ satisfying the conditions of Theorem 2.

Because of Theorem 2,

$$\frac{\partial L}{\partial u} = -x(1 + \lambda) - \gamma_1 + \gamma_2 = 0.$$

For γ, which varies given the value of u, Theorem 2 says that $\gamma' h(u^\circ) = 0$, and $\gamma_1, \gamma_2 \geq 0$.

In order to maintain the equality $-x(1+\lambda) - \gamma_1 + \gamma_2 = 0$, γ_1 must be 0 and $\gamma_2 > 0$ whenever $-x(1+\lambda)$ is negative. Likewise, γ_2 must be 0 and $\gamma_1 > 0$ whenever $-x(1+\lambda)$ is positive. Now, because $\gamma_1 u^\circ = 0$ and $\gamma_2(u^\circ - u_{max}) = 0$, then if $\gamma_1 = 0$ and $\gamma_2 > 0$, then it must be that $u^\circ = u_{max}$. Likewise, if $\gamma_2 = 0$ and $\gamma_1 > 0$, then it must be that $u^\circ = 0$.

In other words, there should be a maximum harvest effort as long as $-x(1+\lambda) < 0$, and employ no harvest effort when $-x(1+\lambda) > 0$. Because the value of the function $-x(1+\lambda)$ being positive or negative decides if the control is one extreme value or the other (0 or u_{max}), $-x(1+\lambda)$ is called the switching function. Whenever bang-bang or bang-off-bang control is used, it is easy to see that when using Theorem 2, a type of switching function will be involved. The only problem left to decide is what to do when the switching function $-x(1+\lambda) = 0$.

According to Theorem 3,

$$\dot{\lambda} = -\frac{\partial H}{\partial x}|_{u=u^\circ} = u - \lambda \left(r - \frac{2r}{K} x - u \right)$$

and

$$\lambda(T) = \rho \frac{\partial g(x)}{\partial x}|_{x=x(T)}.$$

However, this information is of little use because the final point of arrival is not specified because of not knowing ρ. The question that needs to be answered is what to do when $-x(1+\lambda) = 0$. Well, if it is zero for a significant amount of time, then, so will its derivatives w.r.t t be zero. That is

$$-\dot{x}(1+\lambda) - x\dot{\lambda} = 0. \tag{17}$$

Substituting $\dot{\lambda}$, and $\dot{x}$, (17) simplifies to

$$\frac{r}{K}x(-K+x-\lambda x) = \frac{r}{K}x(-K+x+x-x-\lambda x) = \frac{r}{K}x[-K+x+x-x(1+\lambda)] = 0.$$

Because $-x(1+\lambda) = 0$, the expression simplifies to

$$\frac{r}{K}x[-K+2x] = 0 \Rightarrow x = \frac{K}{2}.$$

Now, the second derivative of $-x(x+\lambda)$ must also be 0, that is

$$0 = \ddot{x}(1+\lambda) - 2\dot{x}\dot{\lambda} - x\ddot{\lambda} = \frac{r}{K}\dot{x}[4x - K + -x(1+\lambda)] + \frac{r}{K}x[-\dot{x}(1+\lambda) - x\dot{\lambda}].$$

Because $x = \frac{K}{2}$, its derivative is 0. Also using that $-x(x+\lambda) = 0$, and $-\dot{x}(1+\lambda) - x\dot{\lambda} = 0$, the second derivative expression then simplifies to

$$\frac{rK}{2}\left[\frac{r}{2} - u\right] = 0 \quad \Rightarrow \quad u = \frac{r}{2}.$$

This then completes the control for any value we need. This example serves for an understanding of how to use these techniques to solve control problems. It is easy to see how many of the processes involved can become complicated and often un-doable–specifically that of finding the control on the intervals where a switching function is identically zero.

4 Optimal Control Using the Riemann-Hilbert Problem [11]

The Riemann-Hilbert problem generally speaking is to find a function $w(x,y) = u(x,y) + iv(x,y)$ where u and v are real, and w is analytic inside of a region enclosed by a contour C, such that

$$\alpha(t)u(t) + \beta(t)v(t) = \gamma(t)$$

for every t in C where α, β, and γ are given real functions. The specific type of Riemann-Hilbert problem that will later be shown to apply to control theory has the contour C as the real axis.

In order to derive a Riemann-Hilbert problem from an optimal control problem, the following important theorem needs to be used:

Theorem 4. *[10] If the Fourier transforms exist for maps $f, g : \mathbb{R} \to \mathbb{R}$, and both $f, g \in L^2(\mathbb{R})$, then $\langle f, g\rangle = \langle \widehat{f}, \widehat{g}\rangle$ where $\widehat{f}$ is the Fourier transform of f, and the inner product is defined as*

$$\langle f, g\rangle := \int_{\mathbb{R}} f^*(t)g(t)dt.$$

This theorem will be important in deriving a Riemann-Hilbert problem from a given control problem. Deriving the Riemann-Hilbert problem starts with the goal of solving the optimal control problem by trying to simplify the control problem to working in the complex plane by taking the Fourier transform of the control and state. Then variational calculus is applied to finding an optimal control. It is almost always impossible to find simple representations of Fourier transforms of non-linear differential equations in order to analyze controls in phase space. The theory being discussed now is restricted to control problems that involve linear differential equations. The derivation will be shown using the following simple example.

Consider the control objective to minimize $J[u] = \int_0^\infty ux[u]dt$ subject to $\dot{x} = -x + u$ with initial state $x(0) = a$. Then if u_+ represent the control supported on positive times $t \in [0, \infty)$, then $J[u] = \int_{-\infty}^{\infty} u_+x[u_- + u_+]dt$. Here we can suppose some past control u_- which is only supported on negative times, has put the state x of the system in the current state $x(0) = a$. Because $u_+^*(t) = u_+(t)$, and by Theorem 4 if

$x, u \in L^2$,

$$J[u] = \int_{-\infty}^{\infty} u_+(t)x[u_- + u_+](t)dt = \int_{-\infty}^{\infty} u_+^*(t)x[u_- + u_+](t)dt \tag{18}$$
$$= \int_{-\infty}^{\infty} \widehat{u}_+^*(\omega)\widehat{x}[\widehat{u}_- + \widehat{u}_+](\omega)d\omega.$$

The following important identity holds for $f(t)$ being a real valued function: $\overline{\widehat{f}(\omega)} = \widehat{f}(-\overline{\omega})$. This is because

$$\overline{\widehat{f}(\omega)} = \int_{-\infty}^{\infty} \overline{e^{i\omega t}} f(t)dt, \tag{19}$$

and

$$\overline{e^{iwt}} = \overline{e^{i(a+bi)t}} = \overline{e^{(ai-b)t}} = e^{-bt}\overline{[\cos(at) + i\sin(at)]}$$

$$= e^{-bt}[\cos(at) - i\sin(at)] = e^{-bt}[\cos(-at) + i\sin(-at)] = e^{(-b-ai)t} = e^{i(-a+bi)t} = e^{i(-\overline{\omega})t}.$$

This makes equation (19)

$$= \int_{-\infty}^{\infty} [e^{i(-\overline{\omega})t}]f(t)dt = \widehat{f}(-\overline{\omega}).$$

Therefore, because the integral (18) is over the real line, and by applying $\overline{\widehat{u}_+}(\omega) = \widehat{u}_+(-\overline{\omega})$,

$$J[u] = \int_{-\infty}^{\infty} \widehat{u}_+(-\omega)\widehat{x}[\widehat{u}_-(\omega) + \widehat{u}_+(\omega)]d\omega$$

Now, by taking the Fourier transform of $\dot{x} = -x + u$, we get $-i\omega\widehat{x} = -\widehat{x} + \widehat{u}$. Solving for $\widehat{x}$ gives

$$\widehat{x} = \frac{\widehat{u}}{-i\omega + 1} = \chi(\omega)\widehat{u},$$

where

$$\chi(\omega) = \frac{1}{-i\omega + 1}.$$

Therefore by substituting $\widehat{x}$

$$J[u] = \int_{-\infty}^{\infty} \widehat{u}_+(-\omega) \frac{\widehat{u}_-(\omega) + \widehat{u}_+(\omega)}{-i\omega + 1} d\omega. \tag{20}$$

We want to be able to choose $\widehat{u}_+$ as generally as possible while maintaining that (20) makes sense and is not infinite. The integrand needs to be $O(\frac{1}{\omega^{1+\epsilon}})$ as $\omega \to \infty$ in order for the integral to not be infinite. This requires $\widehat{u}_+ \sim \frac{1}{\omega}$ as $\omega \to \infty$ which is not a legitimate requirement because the optimal control may involve delta functions whose Fourier transforms $\sim O(\omega^0)$ as $\omega \to \infty$. However we can widen the space of allowed functions $\widehat{u}_+$ to include the class of possibly needed controls by noting that if you replace ω with $-\omega$ in an integrand where the integral is over all the reals, then the value of an integral does not change. Using this fact

$$\begin{aligned} J[u] = \int_{-\infty}^{\infty} \widehat{u}_+(-\omega) \frac{\widehat{u}_-(\omega) + \widehat{u}_+(\omega)}{-i\omega + 1} d\omega &= \int_{-\infty}^{\infty} \widehat{u}_+(-\omega) \frac{\widehat{u}_-(\omega)}{-i\omega + 1} d\omega \\ &+ 2\frac{1}{2} \int_{-\infty}^{\infty} \frac{\widehat{u}_+(-\omega)\widehat{u}_+(\omega)}{-i\omega + 1} d\omega \end{aligned} \tag{21}$$

$$\begin{aligned} &= \int_{-\infty}^{\infty} \widehat{u}_+(-\omega) \frac{\widehat{u}_-(\omega)}{-i\omega + 1} d\omega + \frac{1}{2} \int_{-\infty}^{\infty} \frac{\widehat{u}_+(-\omega)\widehat{u}_+(\omega)}{-i\omega + 1} d\omega \\ &+ \frac{1}{2} \int_{-\infty}^{\infty} \frac{\widehat{u}_+(\omega)\widehat{u}_+(-\omega)}{i\omega + 1} d\omega \end{aligned} \tag{22}$$

$$= \int_{-\infty}^{\infty} \widehat{u}_+(-\omega) \frac{\widehat{u}_-(\omega)}{-i\omega + 1} d\omega + \frac{1}{2} \int_{-\infty}^{\infty} \widehat{u}_+(-\omega)\widehat{u}_+(\omega) \left(\frac{1}{-i\omega + 1} + \frac{1}{i\omega + 1} \right) d\omega \tag{23}$$

$$= \int_{-\infty}^{\infty} \widehat{u}_+(-\omega) \frac{\widehat{u}_-(\omega)}{-i\omega + 1} d\omega + \int_{-\infty}^{\infty} \frac{\widehat{u}_+(-\omega)\widehat{u}_+(\omega)}{\omega^2 + 1} d\omega. \tag{24}$$

By using J in the form (24), we can now allow for $\widehat{u}_+$ to go as a constant as long as $\widehat{u}_- \sim O(\frac{1}{\omega^\epsilon})$. This makes sense because it is a legitimate to put the restriction on an assumed past control.

Now by taking the variational δ of J in it's form (24) with respect to $\widehat{u}_+$,

$$\delta_{\widehat{u}_+} J = \int_{-\infty}^{\infty} \left[\frac{\delta\widehat{u}_+(-\omega)\widehat{u}_-(\omega)}{-i\omega + 1} + \frac{\delta\widehat{u}_+(-\omega)\widehat{u}_+(\omega)}{\omega^2 + 1} + \frac{\widehat{u}_+(-\omega)\delta\widehat{u}_+(\omega)}{\omega^2 + 1} \right] d\omega. \tag{25}$$

The variational is on both $u_+(-\omega)$ and $u_+(\omega)$, and it would be convenient if the variational were on the same term. Noticing that the integral is over all reals, we can replace ω with $-\omega$ everywhere in the second term without changing the value of the integral giving us that (25) is

$$\delta_{\widehat{u}_+} J = \int_{-\infty}^{\infty} \left[\frac{\delta\widehat{u}_+(-\omega)\widehat{u}_-(\omega)}{-i\omega + 1} + \frac{\delta\widehat{u}_+(-\omega)\widehat{u}_+(\omega)}{\omega^2 + 1} + \frac{\widehat{u}_+(\omega)\delta\widehat{u}_+(-\omega)}{(-\omega)^2 + 1} \right] d\omega \tag{26}$$

$$= \int_{-\infty}^{\infty} \delta\widehat{u}_+(-\omega) \left[\frac{\widehat{u}_-(\omega)}{-i\omega + 1} + \frac{2\widehat{u}_+(\omega)}{\omega^2 + 1} \right] d\omega. \tag{27}$$

Now because J is minimized when the variational $\delta_{\widehat{u}_+} J = 0$, the optimal control occurs when

$$\delta_{\widehat{u}_+} J = \int_{-\infty}^{\infty} \delta\widehat{u}_+(-\omega) \left[\frac{\widehat{u}_-(\omega)}{-i\omega + 1} + \frac{2\widehat{u}_+(\omega)}{\omega^2 + 1} \right] d\omega = 0.$$

This integral would be zero if the integrand were analytic and decaying as $\omega \to \infty$ in at least one of the half-planes to take the integral as the limit of half circle paths whose interior has no poles making the sum of residues zero. In order to have the integral make sense (be non-infinite), the integrand needs to be decaying on an order of $O(\frac{1}{\omega^{1+\epsilon}})$ as $\omega \to \infty$ just as for the previous integral, which again constrains the possible functions u_+. The terms in the variation of the integral thankfully do not add new constraints to $\widehat{u}$ that were previously needed.

Notice that $\delta\widehat{u}_+(\omega)$ is analytic in the upper-half plane which makes $\delta\widehat{u}_+(-\omega)$ analytic in the lower-half plane. Therefore, if

$$\frac{\widehat{u}_-(\omega)}{-i\omega+1} + \frac{2\widehat{u}_+(\omega)}{\omega^2+1} = Z_-(\omega) \tag{28}$$

for some $Z_-(\omega)$ analytic in the lower half plane and decaying as $\omega \to \infty$, then the variational of J is zero by integrating over the limit of half circles in the lower half plane. This is because the integral is over a set with no poles whose sum of residues is zero. The control $\widehat{u}_+$ that satisfies (28) therefore minimizes J. Finding the function $\widehat{u}_+$ that satisfies (28) is a form of the Riemann-Hilbert problem. In solving such a problem the following terminology is used: A plus function is a function that is analytic and decaying in the upper-half plane. A minus function is defined similarly.

The plan to solving the Riemann-Hilbert problem starts by trying to find some way to separate the parts of the Riemann-Hilbert problem into a plus function on one side of the equation, and a minus function on the other side of the equation, then that would force both sides to be analytic and decaying. By Liouville's theorem, a bounded entire function is a constant which would force the plus side of the equation equal to zero. This then would allow us to solve for $\widehat{u}_+$.

The first step to separating parts of this Riemann-Hilbert problem into plus and minus parts is using separation of fractions to get

$$\frac{\widehat{u}_+(\omega)}{i\omega+1} + \frac{\widehat{u}_+(\omega)}{-i\omega+1} + \frac{\widehat{u}_-(\omega)}{-i\omega+1} = Z_-(\omega). \tag{29}$$

Multiplying (29) by $\frac{1}{-i}$ to simplify we get

$$\frac{\widehat{u}_+(\omega)}{\omega+i} - \frac{\widehat{u}_+(\omega)}{\omega-i} + \frac{\widehat{u}_-(\omega)}{\omega+i} = Z_-(\omega). \tag{30}$$

However for the second term, $\widehat{u}_+(\omega)$ is a plus function, but the denominator is

$\omega - i$ putting a pole at i, making the middle term neither a plus nor a minus function. But notice that the middle term is

$$-\frac{\widehat{u}_+(\omega)}{(\omega - i)} = -\frac{\widehat{u}_+(\omega) - \widehat{u}_+(i)}{(\omega - i)} - \frac{\widehat{u}_+(i)}{(\omega - i)}.$$

The limit as $\omega \to i$ of

$$\frac{\widehat{u}_+(\omega) - \widehat{u}_+(i)}{(\omega - i)} \tag{31}$$

approaches the derivative of u_+ at i due to the fact that u_+ is analytic in the upper-half-plane. This means that (31) has a removable singularity at i making it a plus function. Using this trick on both the second and third terms,

$$\frac{\widehat{u}_+(\omega)}{(\omega + i)} - \frac{\widehat{u}_+(\omega) - \widehat{u}_+(i)}{(\omega - i)} - \frac{\widehat{u}_+(i)}{(\omega - i)} + \frac{\widehat{u}_-(\omega) - \widehat{u}_-(-i)}{\omega + i} + \frac{\widehat{u}_-(-i)}{\omega + i} = Z_-(\omega).$$

This makes the first, second, and fifth term plus functions, and all other terms minus functions. Now by putting the plus functions on one side of the equation and the minus functions on the other side of the equation we get

$$\frac{\widehat{u}_+(\omega)}{(\omega + i)} - \frac{\widehat{u}_+(\omega) - \widehat{u}_+(i)}{(\omega - i)} + \frac{\widehat{u}_-(-i)}{\omega + i} = Z_-(\omega) + \frac{\widehat{u}_+(i)}{(\omega - i)} - \frac{\widehat{u}_-(\omega) - \widehat{u}_-(-i)}{\omega + i} = 0.$$

The reason this expression is equal to zero is because one side is a plus function and the other side is a minus function making both sides entire functions decaying to zero and bounded, which by Liouville's theorem makes the function a constant which is zero. Now solving for $\widehat{u}_+(\omega)$ we get

$$\widehat{u}_+(\omega) = \frac{-1}{2i}[-(\omega - i)\widehat{u}_-(-i) - (\omega + i)\widehat{u}_+(i)].$$

Due to the relationship above and due to the fact that $\widehat{u}_+(\omega)$ must be converge to a constant at infinity, it must be true that $\widehat{u}_+(i) = \widehat{u}_-(-i)$ forcing $\widehat{u}_+(\omega) = \widehat{u}_-(-i)$

which solves this Riemann-Hilbert problem when $\widehat{u}_-(-i)$ is found. Now in order to find out what the control is, remember that $x(0) = a$. By taking the Fourier transform of $\dot{x} = -x + u$ we get

$$-i\omega\widehat{x}(\omega) = -\widehat{x}(\omega) + \widehat{u}(\omega) = -\widehat{x}(\omega) + \widehat{u}_-(\omega) + \widehat{u}_+(\omega) = -\widehat{x}(\omega) + \widehat{u}_-(\omega) + \widehat{u}_-(-i).$$

Solving for $\widehat{x}(\omega)$ we get

$$\widehat{x}(\omega) = \frac{\widehat{u}_-(\omega) + \widehat{u}_-(-i)}{-i(\omega + i)}.$$

Solving the differential equation $\dot{x} = -x + u$ gives us

$$x(t) = \int_{-\infty}^{t} e^{\tau - t} u(\tau) d\tau.$$

Taking the limit from the left,

$$a = x(0_-) = \int_{-\infty}^{0_-} e^{\tau - t} u(\tau) d\tau = \sqrt{2\pi} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\omega\tau} u_-(\tau) d\tau|_{\omega=-i} = \sqrt{2\pi}\widehat{u}_-(-i). \tag{32}$$

This then shows that the optimal control is

$$\widehat{u}_+(\omega) = \widehat{u}_-(-i) = \frac{-a}{\sqrt{2\pi}}.$$

5 A Generalization of the Riemann-Hilbert Problem for Optimal Control

For the general problem, suppose that the goal is to minimize the integral

$$J[u] = \int_{0}^{\infty} l(x[u(t)], u(t)) dt, \tag{33}$$

constrained to the dynamics

$$\dot{x} = Ax + Bu \tag{34}$$

and initial state $x(0) = a$ where the eigenvalues of A have negative real parts. It is necessary to assume that all eigenvalues of A have a negative real part. This is a needed condition to solve the optimal control problem using the Riemann-Hilbert problem on the real line. Further development of applying the Riemann-Hilbert problem may help us to remove the requirement of A having all eigenvalues with negative real part by moving the contour C to be above all eigenvalues of A and modifying the Riemann-Hilbert problem derivations accordingly.

The solution to equation (34) is then

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau = \int_{-\infty}^0 e^{A(t-\tau)}Bu_-(\tau)d\tau + \int_0^t e^{A(t-\tau)}Bu_+(\tau)d\tau. \tag{35}$$

$$= \int_{-\infty}^t e^{A(t-\tau)}B[u_-(\tau) + u_+(\tau)]d\tau,$$

where u_- is an assumed past control that put the initial state to $x(0) = a$. Assuming no two eigenvalues are the same, from this we get

$$x(0) = \int_{-\infty}^{0_-} e^{-A\tau}Bu_-(\tau)d\tau = \int_{-\infty}^{0_-} Ge^{-D\tau}G^{-1}Bu_-(\tau)d\tau$$

$$= \int_{-\infty}^{0_-} \begin{bmatrix} \sum_j \sum_k c_{1jk}e^{-\lambda_j\tau}u_{-,k}(\tau) \\ \sum_j \sum_k c_{2jk}e^{-\lambda_j\tau}u_{-,k}(\tau) \\ \vdots \end{bmatrix} d\tau = \begin{bmatrix} \sum_j \sum_k c_{1jk}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0_-} e^{-\lambda_j\tau}u_{-,k}(\tau)d\tau \\ \sum_j \sum_k c_{2jk}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{0_-} e^{-\lambda_j\tau}u_{-,k}(\tau)d\tau \\ \vdots \end{bmatrix}$$

$$= \begin{bmatrix} \sum_j \sum_k c_{1jk}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega\tau}u_{-,k}(\tau)d\tau|_{\omega=i\lambda_j} \\ \sum_j \sum_k c_{2jk}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega\tau}u_{-,k}(\tau)d\tau|_{\omega=i\lambda_j} \\ \vdots \end{bmatrix} = \begin{bmatrix} \sum_j \sum_k c_{1jk}\sqrt{2\pi}\widehat{u}_{-,k}(i\lambda_j) \\ \sum_j \sum_k c_{2jk}\sqrt{2\pi}\widehat{u}_{-,k}(i\lambda_j) \\ \vdots \end{bmatrix}. \tag{36}$$

For this general control problem, the relationship for $x(0)$ in (36) can later be used for finding $\widehat{x}_+(\omega)$ just as the simpler example problem used a simpler form of (87) in the example for (32).

In order to use Theorem 4 to derive the Riemann-Hilbert problem for minimizing (33) subject to constraint (34), we need to assume also that $l(x,u) = L_1u + L_2x + x'L_3x + u'L_4u + x'L_5u$. We are most interested in the case where $L_4 = 0$ making the control problem a singular control problem (again meaning that $D_uH(x,u)$ does not depend on u where H is the Hamiltonian). If L_4 is non-zero, then the control problem is no longer a singular control problem and the problem can be easily solved using the Hamilton-Jacobi-Bellman equation assuming the dynamics are controllable. Also it would be rare for there to be an interest in minimizing an integral with terms $L_1u + L_2x$. We will then continue assuming L_1, L_2, and L_4 are zero matrices. Then the control u constrained to the dynamics and final condition that minimizes

$$J[u] = \int_0^{\infty} [x'L_3x + x'L_5u]dt, \tag{37}$$

is the same function u that when constrained to the dynamics minimizes

$$\widetilde{J}[u] := \int_{-\infty}^{\infty} [x'L_3x + x'L_5u_+]dt$$

where u_+ is the part of u that is supported on positive times, and some assumed past control u_- supported on negative times has put the state at $x(0)$. The function u is then the sum of it's plus and minus part $u = u_- + u_+$.

By taking the Fourier transform of the dynamics we get

$$-i\omega\widehat{x} = A\widehat{x} + B\widehat{u}$$

$$\widehat{x} = (-i\omega I - A)^{-1}B\widehat{u}. \tag{38}$$

By defining $\chi(\omega) := (-i\omega I - A)^{-1}B$, expression (38) can be simplified to $\widehat{x} = \chi(\omega)\widehat{u}$.

Assuming that u, and $x \in L^2$, then by Parseval's theorem (Theorem 4),

$$\widetilde{J}[u] = \int_{-\infty}^{\infty} [x'[u]L_3x[u] + x'[u]L_5u_+]dt \tag{39}$$

$$= \int_{-\infty}^{\infty} [\widehat{x}^*[u](\omega)L_3\widehat{x}[u](\omega) + \widehat{x}^*[u](\omega)L_5\widehat{u}_+(\omega)]d\omega.$$

Just as was done in the earlier example, we can replace $\widehat{x}^*[\widehat{u}](\omega)$ with $\widehat{x}'[\widehat{u}](-\overline{\omega})$ which were shown to be equal earlier. Also noting that $\widehat{x} = \chi(\omega)\widehat{u}$ we can simplify equation (39) to no longer depend on $\widehat{x}$ as

$$= \int_{-\infty}^{\infty} [\widehat{u}'(-\overline{\omega})\chi'(-\overline{\omega})L_3\chi(\omega)\widehat{u}(\omega) + \widehat{u}'(-\overline{\omega})\chi'(-\overline{\omega})L_5\widehat{u}_+(\omega)]d\omega. \tag{40}$$

Also, because the integral is over the real line, $\overline{\omega}$ can be replaced with ω giving us that (40) is

$$= \int_{-\infty}^{\infty} [\widehat{u}'(-\omega)\chi'(-\omega)L_3\chi(\omega)\widehat{u}(\omega) + \widehat{u}'(-\omega)\chi'(-\omega)L_5\widehat{u}_+(\omega)]d\omega. \tag{41}$$

Substituting $\widehat{u}(\omega) = \widehat{u}_+(\omega) + \widehat{u}_-(\omega)$ and factoring, (41) becomes

$$= \int_{-\infty}^{\infty} [(\widehat{u}'_+(-\omega) + \widehat{u}'_-(-\omega))\chi'(-\omega)L_3\chi(\omega)(\widehat{u}_+(\omega) + \widehat{u}_-(\omega)) + \widehat{u}'(-\omega)\chi'(-\omega)L_5\widehat{u}_+(\omega)]d\omega \tag{42}$$

$$= \int_{-\infty}^{\infty} [\widehat{u}'_-(-\omega)(\chi'(-\omega)L_3\chi)\widehat{u}_-(\omega) + \widehat{u}'_-(-\omega)(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\widehat{u}_+(\omega)$$

$$+ \widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi)\widehat{u}_-(\omega) + \widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\widehat{u}_+(\omega)]d\omega. \tag{43}$$

Assuming that (43) can be transformed to allow for controls $\widehat{u}_+$, and assuming that there is a minimum of (43), then we can find that minimum by taking the variation of (43) with respect to $\widehat{u}_+$. (43) will have a minimum if we restrict L_3 and $\chi'(-\omega)L_5$ to be positive semi-definite matrices. Now by taking the variation of $\widetilde{J}$ from (43) and setting it equal to zero in order to minimize $\widetilde{J}$ we get

$$\delta_{\widehat{u}_+}\widetilde{J} = \int_{-\infty}^{\infty} [\widehat{u}'_-(-\omega)(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\delta\widehat{u}_+(\omega) + \delta\widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi)\widehat{u}_-(\omega)$$

$$+ \delta\widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\widehat{u}_+(\omega)$$

$$+ \widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\delta\widehat{u}_+(\omega)]d\omega. \tag{44}$$

The integral in (44) is over all reals, so a term in the integral could have ω replaced by $-\omega$ without changing the value of the integral. We want to have the variational on only one function. By replacing ω by $-\omega$ on terms where the variational is on $\widehat{u}_+(\omega)$, we then get the variational on one term only with (44) as

$$\delta_{\widehat{u}_+}\widetilde{J} = \int_{-\infty}^{\infty}[\widehat{u}'_-(\omega)(\chi' L_3\chi(-\omega) + \chi' L_5)\delta\widehat{u}_+(-\omega) + \delta\widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi)\widehat{u}_-(\omega)$$
$$+ \delta\widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\widehat{u}_+(\omega)$$
$$+ \widehat{u}'_+(\omega)(\chi' L_3\chi(-\omega) + \chi' L_5)\delta\widehat{u}_+(-\omega)]d\omega. \tag{45}$$

The variational is still on the transpose of $\widehat{u}_+(\omega)$ in some places, whereas on other terms it is not transposed. It would also be helpful to have the variational be only on the transpose of $\widehat{u}_+(\omega)$. Transposing a term by itself is simply the transpose of a scalar which does not effect the integral. Taking advantage of this fact (45) is

$$\delta_{\widehat{u}_+}\widetilde{J} = \int_{-\infty}^{\infty}[\delta\widehat{u}'_+(-\omega)(\chi'(-\omega)L'_3\chi + L'_5\chi)\widehat{u}_-(\omega) + \delta\widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi)\widehat{u}_-(\omega)$$
$$+ \delta\widehat{u}'_+(-\omega)(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\widehat{u}_+(\omega) + \delta\widehat{u}'_+(-\omega)(\chi'(-\omega)L'_3\chi$$
$$+ L'_5\chi)\widehat{u}_+(\omega)]d\omega \tag{46}$$
$$= \int_{-\infty}^{\infty}\delta\widehat{u}'_+(-\omega)[(\chi'(-\omega)L'_3\chi + L'_5\chi)\widehat{u}_-(\omega) + (\chi'(-\omega)L_3\chi)\widehat{u}_-(\omega)$$
$$+ (\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\widehat{u}_+(\omega) + (\chi'(-\omega)L'_3\chi + L'_5\chi)\widehat{u}_+(\omega)]d\omega = 0. \tag{47}$$

Now $\delta\widehat{u}'_+(\omega)$ is a plus function componentwise making $\delta\widehat{u}'_+(-\omega)$ a minus function componentwise. The integral (45) can be forced to zero by requiring the part in (45) in brackets to be some unknown componentwise minus function $Z_-(\omega)$. This then makes the integral zero because integrating as the limit of half circles in the lower half plane shows that the integral is zero. From this we get a Riemann-Hilbert problem for the general control problem as:

$$(\chi'(-\omega)L_3'\chi + L_5'\chi)\widehat{u}_-(\omega) + (\chi'(-\omega)L_3\chi)\widehat{u}_-(\omega)$$

$$+(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5)\widehat{u}_+(\omega) + (\chi'(-\omega)L_3'\chi + L_5'\chi)\widehat{u}_+(\omega) = Z_-(\omega). \qquad (48)$$

Simplifying, (48) becomes

$$(\chi'(-\omega)L_3'\chi + L_5'\chi + \chi'(-\omega)L_3\chi)\widehat{u}_-(\omega)$$

$$+(\chi'(-\omega)L_3\chi + \chi'(-\omega)L_5 + \chi'(-\omega)L_3'\chi + L_5'\chi)\widehat{u}_+(\omega) = Z_-(\omega). \qquad (49)$$

The solution to (49) is the optimal control because it is the control that minimizes the variation $\delta_{\widehat{u}_+}\widetilde{J}$. This problem can then be solved to find u_+ using the same methods in the example problem earlier.

Notice that

$$\chi(\omega) = (-i\omega I - A)^{-1}B = \frac{adj(-i\omega I - A)}{det(-i\omega I - A)}B$$

is a plus function with poles at $\omega = i\lambda_j$ where λ_j are eigenvalues of A which were assumed to be negative. This then makes $\chi(-\omega)$ a minus function. From this you can see that some of the terms in the sum for (49) are plus functions, others minus functions, and others neither. For those that are neither plus nor minus functions we can use the same trick used earlier to do separation of fractions component-wise on each vector, and then use the trick of creating removable singularities componentwise. Then by putting the plus functions on one side of the equation, and minus functions on the other side, we can then use Liouville's theorem and continue using the same methods in the example problems.

Another possible method to solve (49) is to try and factor the matrix to the left of $\widehat{u}_+$ into two matrices $= \rho_-\rho_+$ where ρ_- and ρ_-^{-1} are analytic in the lower half plane, and ρ_+ is analytic in the upper half plane. This could be used as a step for separating the Riemann-Hilbert problem into plus and minus parts eventually allowing you to multiply by ρ_-^{-1} to both sides. However, finding such a factorization is a challenging problem. When a factorization is hard to find, then using the trick to create removable singularities would be the better option to solving (48).

6 Example Riemann-Hilbert Problem for Deadline Optimal Control

Now assume that the goal is to derive a Riemann-Hilbert problem to solve the control problem of minimizing

$$J[u] = \int_0^T uxdt \tag{50}$$

constrained to the dynamics $\dot{x} = -x + u$, $x(0) = a$, and the final destination $x(T) = b$. The difference between this problem and the example earlier is that the final state must be at b at time T which is a harder problem than the earlier example which does not have the final constraint $x(T) = b$. Before solving the problem with $x(T) = b$, we will solve a simpler problem of simply minimizing (50) within $[0, T]$ without the requirement that $x(T) = b$ which will help us to develop the theory to where we can include the constraint $x(T) = b$.

Notice that,

$$\int_0^T x[u]udt = \int_0^\infty x[u]u_T dt = \int_{-\infty}^\infty x[u]u_T dt = \int_{-\infty}^\infty x[u_- + u_+]u_T dt = \int_{-\infty}^\infty x[u_- + u_T]u_T dt \tag{51}$$

where u_T is the control supported only on the interval $[0, T]$. As in the earlier example, by Parseval's theorem (requiring that x and $u \in L^2$), the fact that $\overline{\widehat{f}(\omega)} = \widehat{f}(-\overline{\omega})$ for $f(t)$ real, and the fact that the integral is over the real line, (51) is

$$= \int_{-\infty}^{\infty} \widehat{x}[\widehat{u}_- + \widehat{u}_T](-\omega)\widehat{u}_T(\omega)dt. \tag{52}$$

Due to the dynamics $\dot{x} = -x + u$, the Fourier transform gives us $\widehat{x} = \frac{\widehat{u}}{-i\omega+1}$, then making (52)

$$= \int_{-\infty}^{\infty} \frac{\widehat{u}_-(-\omega) + \widehat{u}_T(-\omega)}{i\omega + 1}\widehat{u}_T(\omega)d\omega. \tag{53}$$

It is important to make sure that the integral in 53 makes sense and converges. Hence the integrand must be at least on the order $O(\frac{1}{\omega^{1+\epsilon}})$. The set of possible functions u_T and u_- are then restricted in order for the integral to converge. 53 is the same integral back in (20) with $\widehat{u}_T$ in the place of $\widehat{u}_+$. 53 has the same problem with $\widehat{u}_T$ as (20) in the manner that it restricted the allowed possible functions $\widehat{u}_+$. Earlier we converted (20) to (24) which allows the needed space of functions for $\widehat{u}_+$. Using the the integral (24) replacing $\widehat{u}_+$ with $\widehat{u}_T$, we can allow for $\widehat{u}_T$ to go as a constant $\sim O(\omega^0)$ as $\omega \to \infty$. This then allows the time domain for u_T to have delta functions. Now we can take the variation of 53 or (24) either way giving us the same Riemann-Hilbert problem. Basically (24) shows us that the allowed values for u_T will not be restricted beyond what is necessary.

Taking the variational of 53 gives us

$$\delta_{\widehat{u}_T} J[u] = \int_{-\infty}^{\infty} \frac{(\widehat{u}_-(-\omega) + \widehat{u}_T(-\omega))\delta\widehat{u}_T(\omega) + \widehat{u}_T(\omega)\delta\widehat{u}_T(-\omega)}{i\omega + 1}d\omega.$$

It would be useful to have the variational on the same function in order to simplify. Because the integral is over all reals, ω can be replaced by $-\omega$ in the first integrand without changing the value of the integral which then after factoring gives us

$$\delta_{\widehat{u}_T} J[u] = \int_{-\infty}^{\infty} \left[\frac{\widehat{u}_-(\omega) + \widehat{u}_T(\omega)}{-i\omega + 1} + \frac{\widehat{u}_T(\omega)}{i\omega + 1} \right] \delta\widehat{u}_T(-\omega) d\omega$$
$$= \int_{-\infty}^{\infty} \left[\frac{2\widehat{u}_T(\omega)}{(i\omega + 1)(-i\omega + 1)} + \frac{\widehat{u}_-(\omega)}{-i\omega + 1} \right] \delta\widehat{u}_T(-\omega) d\omega = 0. \quad (54)$$

Here the variational is set to zero in order to find the control that minimizes J.

Note that multiplying a function $\widehat{f}$ in the frequency domain by $e^{i\omega T}$ gives a function whose inverse Fourier transform in the time domain is a shifted $f(t)$ being $f(t-T)$. Therefore by defining $\widehat{u}_{-T}(\omega) := e^{-i\omega T}\widehat{u}_T(\omega)$, we can note that $\widehat{u}_{-T}(\omega)$ is then analytic in the lower half plane. Similarly we define $\delta\widehat{u}_{-T}(\omega) = e^{-i\omega T}\delta\widehat{u}_T(\omega)$. Using $\delta\widehat{u}_T(\omega) = e^{i\omega T}\delta\widehat{u}_{-T}(\omega)$, we then get $\delta\widehat{u}_T(-\omega) = e^{-i\omega T}\delta\widehat{u}_{-T}(-\omega)$ giving another possible representation for (54) as

$$= \int_{-\infty}^{\infty} \left[\left(\frac{2\widehat{u}_T(\omega)}{(i\omega + 1)(-i\omega + 1)} + \frac{\widehat{u}_-(\omega)}{-i\omega + 1} \right) e^{-i\omega T} \right] \delta\widehat{u}_{-T}(-\omega) d\omega = 0. \quad (55)$$

$\delta\widehat{u}_T(\omega)$ is a plus function making $\delta\widehat{u}_T(-\omega)$ a minus function. The integral 55 can be forced to zero by requiring the part in brackets to be $Z_-(\omega) + e^{i\omega T}Z_+(\omega)$ for some unknown minus function $Z_-(\omega)$, and plus function $Z_+(\omega)$. This then makes the integral zero because then (54) is

$$\int_{-\infty}^{\infty} \left[\frac{2\widehat{u}_T(\omega)}{(i\omega + 1)(-i\omega + 1)} + \frac{\widehat{u}_-(\omega)}{-i\omega + 1} \right] \delta\widehat{u}_T(-\omega) d\omega$$
$$= \int_{-\infty}^{\infty} \left[Z_-(\omega) + e^{i\omega T} Z_+(\omega) \right] \delta\widehat{u}_T(-\omega) d\omega = 0. \quad (56)$$

The integral is zero because $Z_-(\omega)\delta\widehat{u}_T(-\omega)$ is a minus function, making it's integral as the limit of half circles in the lower half plane. This forces the integral of the term to be zero. Similarly $Z_+(\omega)e^{i\omega t}\delta\widehat{u}_T(-\omega) = Z_+(\omega)\delta\widehat{u}_{-T}(-\omega)$, making it a plus function whose integral as the limit of half circles in the upper half plane forces the integral of this term to also be zero. Here the reason why $Z_{-}(\omega) + e^{i\omega T} Z_{+}(\omega)$ was chosen is because we want to allow the right hand side of the Riemann-Hilbert problem to be general enough to solve and find the optimal control u_T.

From this we get a Riemann-Hilbert problem for the general control problem as

$$\frac{2\widehat{u}_T(\omega)}{(i\omega+1)(-i\omega+1)} + \frac{\widehat{u}_{-}(\omega)}{-i\omega+1} = Z_{-}(\omega) + e^{i\omega T} Z_{+}(\omega). \tag{57}$$

Together equation (57) and the relationship $\widehat{u}_{-T}(\omega) := e^{-i\omega T}\widehat{u}_T(\omega)$ define a vector (multivariate) Riemann-Hilbert problem.

Again as before we want to try and separate the problem into minus functions and plus functions to take advantage of Liouville's theorem. Doing separation of fractions on the first term we get

$$\frac{\widehat{u}_T(\omega)}{-i\omega+1} + \frac{\widehat{u}_T(\omega)}{i\omega+1} + \frac{\widehat{u}_{-}(\omega)}{-i\omega+1} = Z_{-}(\omega) + e^{i\omega T} Z_{+}(\omega). \tag{58}$$

The first term is a minus function, but the second and third are mixed which requires using the trick used in the earlier example. Note that

$$\frac{\widehat{u}_T(\omega)}{i\omega+1} = \frac{\widehat{u}_T(\omega) - \widehat{u}_T(i)}{i\omega+1} + \frac{\widehat{u}_T(i)}{i\omega+1}.$$

In the limit as $\omega \to i$, the first term on the right side of the equation converges to the derivative of $\widehat{u}_T$ which exists due to $\widehat{u}_T$ being analytic in the upper half plane. This means the term has a removable singularity. The last term is a minus function. Using this trick we can separate the entire problem piece by piece into plus and minus functions to allow us to eventually use Liouville's theorem. Now (58) becomes

$$\frac{\widehat{u}_T(\omega)}{-i\omega+1}+\frac{\widehat{u}_T(\omega)-\widehat{u}_T(i)}{i\omega+1}+\frac{\widehat{u}_T(i)}{i\omega+1}+\frac{\widehat{u}_-(\omega)-\widehat{u}_-(-i)}{-i\omega+1}+\frac{\widehat{u}_-(-i)}{-i\omega+1}$$
$$=Z_-(\omega)+e^{i\omega T}Z_+(\omega). \tag{59}$$

In equation (59), the first, second, and fifth terms are now minus functions, and the other terms are plus functions. Now by putting the plus functions on the left hand side of the equation, and the minus functions on the right we get

$$\frac{\widehat{u}_T(\omega)}{-i\omega+1}+\frac{\widehat{u}_T(\omega)-\widehat{u}_T(i)}{i\omega+1}+\frac{\widehat{u}_-(-i)}{-i\omega+1}-e^{i\omega T}Z_+(\omega)$$
$$=Z_-(\omega)-\frac{\widehat{u}_T(i)}{i\omega+1}-\frac{\widehat{u}_-(\omega)-\widehat{u}_-(-i)}{-i\omega+1}. \tag{60}$$

Both sided of this equation are then analytic and bounded (assuming the removable singularities are replaced) and going to zero as $\omega \to \infty$. Thus by Liouville's theorem both sides are a constant, that constant being zero making the right side

$$\frac{\widehat{u}_T(\omega)}{-i\omega+1}+\frac{\widehat{u}_T(\omega)-\widehat{u}_T(i)}{i\omega+1}+\frac{\widehat{u}_-(-i)}{-i\omega+1}-e^{i\omega T}Z_+(\omega)=0. \tag{61}$$

Now by multiplying both sides of (61) by $e^{-i\omega T}$ and substituting $\widehat{u}_{-T}(\omega) = e^{-i\omega T}\widehat{u}_T(\omega)$ we get

$$\frac{\widehat{u}_{-T}(\omega)}{-i\omega+1}+\frac{\widehat{u}_{-T}(\omega)-e^{-i\omega T}\widehat{u}_T(i)}{i\omega+1}+\frac{e^{-i\omega T}\widehat{u}_-(-i)}{-i\omega+1}=Z_+(\omega). \tag{62}$$

The importance of $\widehat{u}_{-T}(\omega) = e^{-i\omega T}\widehat{u}_T(\omega)$ is seen in that the shift is needed for preparing to solve the next step. Now we can take (62) and use the trick of creating removable singularities to create plus and minus functions. Notice that the third term has a pole at $-i$ and is analytic in the upper-half-plane. However the exponential in that term is decaying in the lower-half-plane and exploding in the upper-half-plane. For this reason the trick of creating removable singularities is also used for the third term. (62) is now

$$\frac{\widehat{u}_{-T}(\omega)-\widehat{u}_{-T}(-i)}{-i\omega+1}+\frac{\widehat{u}_{-T}(-i)}{-i\omega+1}+\frac{\widehat{u}_{-T}(\omega)-e^{-i\omega T}\widehat{u}_{T}(i)}{i\omega+1} \\ +\widehat{u}_{-}(-i)\frac{e^{-i\omega T}-e^{-T}}{-i\omega+1}+\frac{\widehat{u}_{-}(-i)e^{-T}}{-i\omega+1}=Z_{+}(\omega)$$

with removable singularities. Now by putting the minus functions on the left, and the plus function on the right and again using the convergence to 0 at infinity, we use Liouville's theorem to get that

$$\frac{\widehat{u}_{-T}(\omega)-\widehat{u}_{-T}(-i)}{-i\omega+1}+\frac{\widehat{u}_{-T}(\omega)-e^{-i\omega T}\widehat{u}_{T}(i)}{i\omega+1}+\widehat{u}_{-}(-i)\frac{e^{-i\omega T}-e^{-T}}{-i\omega+1} \\ =Z_{+}(\omega)-\frac{\widehat{u}_{-T}(-i)}{-i\omega+1}-\frac{\widehat{u}_{-}(-i)e^{-T}}{-i\omega+1},$$

which then implies

$$\frac{\widehat{u}_{-T}(\omega)-\widehat{u}_{-T}(-i)}{-i\omega+1}+\frac{\widehat{u}_{-T}(\omega)-e^{-i\omega T}\widehat{u}_{T}(i)}{i\omega+1}+\widehat{u}_{-}(-i)\frac{e^{-i\omega T}-e^{-T}}{-i\omega+1}=0. \tag{63}$$

Simplifying (63) gives

$$\frac{2}{(-i\omega+1)(i\omega+1)}\widehat{u}_{-T}(\omega)=\frac{\widehat{u}_{-T}(-i)}{-i\omega+1}+\frac{e^{-i\omega T}\widehat{u}_{T}(i)}{i\omega+1}-\widehat{u}_{-}(-i)\frac{e^{-i\omega T}-e^{-T}}{-i\omega+1}. \tag{64}$$

Then multiplying both sides of 64 by $(-i\omega+1)(i\omega+1)$,

$$2\widehat{u}_{-T}(\omega)=(i\omega+1)\widehat{u}_{-T}(-i)+(-i\omega+1)e^{-i\omega T}\widehat{u}_{T}(i)-\widehat{u}_{-}(-i)(i\omega+1)(e^{-i\omega T}-e^{-T}). \tag{65}$$

It is useful here to remember that $\widehat{u}_{T}(\omega)$ is analytic and bounded in the upper half plane also making $\widehat{u}_{-T}(\omega)$ analytic and bounded in the lower half plane. In order to guarantee that these facts remain true, notice that as $\omega \to -i\infty$,

$$2\widehat{u}_{-T}(\omega) \sim i\omega(\widehat{u}_{-T}(-i) + \widehat{u}_{-}(-i)e^{-T}).$$

But $\widehat{u}_{-T}(\omega)$ must converge to a constant as $\omega \to -i\infty$ forcing $\widehat{u}_{-T}(-i) = -\widehat{u}_{-}(-i)e^{-T}$ in order for the convergence to hold.

Multiplying (65) by $e^{i\omega T}$ we get

$$2\widehat{u}_{T}(\omega) = e^{i\omega T}(i\omega+1)\widehat{u}_{-T}(-i) + (-i\omega+1)\widehat{u}_{T}(i) - \widehat{u}_{-}(-i)(i\omega+1)(1-e^{i\omega T-T}). \quad (66)$$

From (66) we get that as $\omega \to +i\infty$,

$$2\widehat{u}_{T}(\omega) \sim i\omega(-\widehat{u}_{T}(i) - \widehat{u}_{-}(-i)).$$

But $\widehat{u}_{T}(\omega)$ must converge to a constant as $\omega \to +i\infty$ forcing $\widehat{u}_{T}(i) = -\widehat{u}_{-}(-i)$ in order for the convergence to hold.

We can plug these relationships into (65) to get

$$\begin{aligned}\widehat{u}_{-T}(\omega) = \frac{1}{2}[&-(i\omega+1)\widehat{u}_{-}(-i)e^{-T} - (-i\omega+1)e^{-i\omega T}\widehat{u}_{-}(-i) \\ &- \widehat{u}_{-}(-i)(i\omega+1)(e^{-i\omega T} - e^{-T})]. \end{aligned} \quad (67)$$

Using the fact that $\widehat{u}_{-}(-i) = \frac{a}{\sqrt{2\pi}}$, (67) is then

$$\begin{aligned}\widehat{u}_{-T}(\omega) = \frac{1}{2}\Big[&-(i\omega+1)\frac{a}{\sqrt{2\pi}}e^{-T} - (-i\omega+1)e^{-i\omega T}\frac{a}{\sqrt{2\pi}} \\ &- \frac{a}{\sqrt{2\pi}}(i\omega+1)(e^{-i\omega T} - e^{-T})\Big]. \end{aligned} \quad (68)$$

Finally we can multiply (68) by $e^{i\omega T}$ to get the final optimal control as

$$\begin{aligned}\widehat{u}_T(\omega) &= \frac{a}{2\sqrt{2\pi}}[-(i\omega+1)e^{i\omega T-T} - (-i\omega+1) - (i\omega+1)(1-e^{i\omega T-T})] \\ &= \frac{-a}{\sqrt{2\pi}}. \end{aligned} \tag{69}$$

We have solved the problem of minimizing $J[u]$ with the extra requirement of doing so by time T. But we still haven't done so with the constraint that $x(T) = b$. By now including the constraint $x(T) = b$ we get that

$$\begin{aligned} x(T) = b &= \int_{-\infty}^{T} e^{-(T-\tau)}u(\tau)d\tau = e^{-T}\int_{-\infty}^{0} e^{\tau}u_-(\tau)d\tau + e^{-T}\int_{0}^{T} e^{\tau}u_T(\tau)d\tau \\ &= e^{-T}a + e^{-T}\int_{-\infty}^{\infty} e^{\tau}u_T(\tau)d\tau = e^{-T}a + e^{-T}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega\tau}u_T(\tau)d\tau|_{\omega=-i} \\ &= e^{-T}a + \sqrt{2\pi}e^{-T}\widehat{u}_T(-i). \end{aligned}$$

Solving for $\widehat{u}_T(-i)$ we get

$$\widehat{u}_T(-i) = \frac{e^T b - a}{\sqrt{2\pi}}. \tag{70}$$

The variation $\delta\widehat{u}_T(-\omega)$ back in 55 must be zero at $\omega = i$ because any control $\widehat{u}_T$ must satisfy the constraint (70). Also notice that back in (56) it was shown that the part of the integrand in brackets could be chosen as general as possible to make the variation of the integral be zero. We chose $Z_-(\omega) + e^{i\omega t}Z_+(\omega)$ which was a sufficiently general assumption in order to solve the control problem. However with the constraint that $\delta\widehat{u}_T(-i) = 0$, the integrand can be made more general to be $Z_-(\omega) + e^{i\omega T}(Z_+(\omega) + \frac{\beta}{\omega - i})$ for some β. This is because

$$\int_{-\infty}^{\infty}\left[Z_-(\omega)+e^{i\omega T}(Z_+(\omega)+\frac{\beta}{\omega-i})\right]\delta\widehat{u}_T(-\omega)d\omega=0+$$

$$\int_{-\infty}^{\infty}\frac{e^{i\omega T}\beta\delta\widehat{u}_T(-\omega)}{\omega-i}d\omega=0 \tag{71}$$

due to the fact that the integrand has a removable singularity at i. In fact this generality will be seen later as needed in order to solve the control problem that includes the constraint $x(T)=b$. As was done in (57), we can assume a Riemann-Hilbert problem, here more generally as

$$\frac{2\widehat{u}_T(\omega)}{(i\omega+1)(-i\omega+1)}+\frac{\widehat{u}_-(\omega)}{-i\omega+1}=Z_-(\omega)+e^{i\omega T}\left(Z_+(\omega)+\frac{\beta}{\omega-i}\right). \tag{72}$$

As was done earlier on (57) we need to separate this into plus and minus functions getting

$$\frac{\widehat{u}_T(\omega)}{-i\omega+1}+\frac{\widehat{u}_T(\omega)-\widehat{u}_T(i)}{i\omega+1}+\frac{\widehat{u}_-(-i)}{-i\omega+1}-e^{i\omega T}Z_+(\omega)-\beta\frac{e^{i\omega T}-e^{-T}}{\omega-i}$$
$$=Z_-(\omega)+\frac{\beta e^{-T}}{\omega-i}-\frac{\widehat{u}_T(i)}{i\omega+1}-\frac{\widehat{u}_-(\omega)-\widehat{u}_-(-i)}{-i\omega+1}. \tag{73}$$

This implies that

$$\frac{\widehat{u}_T(\omega)}{-i\omega+1}+\frac{\widehat{u}_T(\omega)-\widehat{u}_T(i)}{i\omega+1}+\frac{\widehat{u}_-(-i)}{-i\omega+1}-e^{i\omega T}Z_+(\omega)-\beta\frac{e^{i\omega T}-e^{-T}}{\omega-i}=0. \tag{74}$$

Multiplying (74) by $e^{-i\omega T}$ we get

$$\frac{\widehat{u}_{-T}(\omega)}{-i\omega+1}+\frac{\widehat{u}_{-T}(\omega)-e^{-i\omega T}\widehat{u}_T(i)}{i\omega+1}+\frac{e^{-i\omega T}\widehat{u}_-(-i)}{-i\omega+1}-Z_+(\omega)-\beta\frac{1-e^{-i\omega T-T}}{\omega-i}=0. \tag{75}$$

Then separating parts into plus and minus function we get

$$Z_+(\omega) = \frac{\widehat{u}_{-T}(\omega) - \widehat{u}_{-T}(-i)}{-i\omega + 1} + \frac{\widehat{u}_{-T}(-i)}{-i\omega + 1} + \frac{\widehat{u}_{-T}(\omega) - e^{-i\omega T}\widehat{u}_T(i)}{i\omega + 1}$$

$$+\widehat{u}_-(-i)\frac{e^{-i\omega T} - e^{-T}}{-i\omega + 1} + \widehat{u}_-(-i)\frac{e^{-T}}{-i\omega + 1} - \beta\frac{1 - e^{-i\omega T - T}}{\omega - i}. \tag{76}$$

Moving the minus functions to one side and applying Liouville's theorem we get

$$0 = \frac{\widehat{u}_{-T}(\omega) - \widehat{u}_{-T}(-i)}{-i\omega + 1} + \frac{\widehat{u}_{-T}(\omega) - e^{-i\omega T}\widehat{u}_T(i)}{i\omega + 1}$$

$$+ \widehat{u}_-(-i)\frac{e^{-i\omega T} - e^{-T}}{-i\omega + 1} - \beta\frac{1 - e^{-i\omega T - T}}{\omega - i}. \tag{77}$$

Solving for $\widehat{u}_{-T}(\omega)$,

$$0 = \frac{\widehat{u}_{-T}(\omega)}{-i\omega + 1} - \frac{\widehat{u}_{-T}(-i)}{-i\omega + 1} + \frac{\widehat{u}_{-T}(\omega)}{i\omega + 1} - \frac{e^{-i\omega T}\widehat{u}_T(i)}{i\omega + 1}$$

$$+ \widehat{u}_-(-i)\frac{e^{-i\omega T} - e^{-T}}{-i\omega + 1} - \beta\frac{1 - e^{-i\omega T - T}}{\omega - i}.$$

$$0 = \frac{2\widehat{u}_{-T}(\omega)}{(-i\omega + 1)(i\omega + 1)} - \frac{\widehat{u}_{-T}(-i)}{-i\omega + 1} - \frac{e^{-i\omega T}\widehat{u}_T(i)}{i\omega + 1}$$

$$+ \widehat{u}_-(-i)\frac{e^{-i\omega T} - e^{-T}}{-i\omega + 1} - \beta\frac{1 - e^{-i\omega T - T}}{\omega - i}.$$

$$-2\widehat{u}_{-T}(\omega) = -\widehat{u}_{-T}(-i)(i\omega + 1) - e^{-i\omega T}\widehat{u}_T(i)(-i\omega + 1) + \widehat{u}_-(-i)(e^{-i\omega T} - e^{-T})(i\omega + 1)$$

$$- (-i\omega + 1)(i\omega + 1)\beta\frac{1 - e^{-i\omega T - T}}{\omega - i}.$$

$$2\widehat{u}_{-T}(\omega) = \widehat{u}_{-T}(-i)(i\omega + 1) + e^{-i\omega T}\widehat{u}_T(i)(-i\omega + 1)$$

$$- \widehat{u}_-(-i)(e^{-i\omega T} - e^{-T})(i\omega + 1) + (\omega + i)\beta(1 - e^{-i\omega T - T}). \tag{78}$$

As $\omega \to -i\infty$

$$2\widehat{u}_{-T}(\omega) \sim i\omega[\widehat{u}_{-T}(-i) + \widehat{u}_{-}(-i)e^{-T} - i\beta].$$

However in order for $\widehat{u}_{-T}(\omega)$ to go as a constant $O(\omega^0)$ as $\omega \to -i\infty$, it is necessary that

$$\widehat{u}_{-T}(-i) = -\widehat{u}_{-}(-i)e^{-T} + i\beta. \tag{79}$$

Multiplying (78) by $e^{i\omega T}$ gives us

$$2\widehat{u}_T(\omega) = e^{i\omega T}\widehat{u}_{-T}(-i)(i\omega + 1) + \widehat{u}_T(i)(-i\omega + 1)$$

$$-\widehat{u}_{-}(-i)(1 - e^{i\omega T - T})(i\omega + 1) + (\omega + i)\beta(e^{i\omega T} - e^{-T}). \tag{80}$$

As $\omega \to i\infty$

$$2\widehat{u}_T(\omega) \sim i\omega[-\widehat{u}_T(i) - \widehat{u}_{-}(-i) - (-i)\beta e^{-T}].$$

However in order for $\widehat{u}_T(\omega)$ to go as a constant $O(\omega^0)$ as $\omega \to i\infty$, it is necessary that

$$\widehat{u}_T(i) = -\widehat{u}_{-}(-i) + i\beta e^{-T}.$$

Substituting these relationships into (80), and using the fact that $u_{-}(-i) = \frac{a}{\sqrt{2\pi}}$ we get

$$\begin{aligned}
2\widehat{u}_T(\omega) &= e^{i\omega T}(-\widehat{u}_{-}(-i)e^{-T} + i\beta)(i\omega + 1) + (-\widehat{u}_{-}(-i) + i\beta e^{-T})(-i\omega + 1) \\
&\quad - \widehat{u}_{-}(-i)(1 - e^{i\omega T - T})(i\omega + 1) + (\omega + i)\beta(e^{i\omega T} - e^{-T}) \\
&= e^{i\omega T}\left(-\frac{a}{\sqrt{2\pi}}e^{-T} + i\beta\right)(i\omega + 1) + \left(-\frac{a}{\sqrt{2\pi}} + i\beta e^{-T}\right)(-i\omega + 1) \\
&\quad - \frac{a}{\sqrt{2\pi}}(1 - e^{i\omega T - T})(i\omega + 1) + (\omega + i)\beta(e^{i\omega T} - e^{-T}).
\end{aligned} \tag{81}$$

The only unknown left to find to get our optimal control is β. Using the relationship $u_T(\omega) = e^{i\omega T} u_{-T}(\omega)$, (79) becomes

$$\widehat{u}_T(-i) = -\widehat{u}_-(-i)e^{i(-i)T-T} + e^{i(-i)T} i\beta = -\frac{a}{\sqrt{2\pi}} + e^T i\beta. \tag{82}$$

Now using $\widehat{u}_T(-i) = \frac{e^T b - a}{\sqrt{2\pi}}$ coming from (70) together with (82) we get

$$\frac{e^T b - a}{\sqrt{2\pi}} = -\frac{a}{\sqrt{2\pi}} + e^T i\beta.$$

Solving for β we get

$$\beta = \frac{-ib}{\sqrt{2\pi}}.$$

Now we can solve for our optimal control from (81)

$$\begin{aligned}
\widehat{u}_T(\omega) = &\frac{1}{2}[e^{i\omega T}(-\frac{a}{\sqrt{2\pi}}e^{-T} + i\frac{-ib}{\sqrt{2\pi}})(i\omega+1) + (-\frac{a}{\sqrt{2\pi}} + i\frac{-ib}{\sqrt{2\pi}}e^{-T})(-i\omega+1) \\
&- \frac{a}{\sqrt{2\pi}}(1 - e^{i\omega T - T})(i\omega+1) + (\omega+i)\frac{-ib}{\sqrt{2\pi}}(e^{i\omega T} - e^{-T})] \qquad (83) \\
= &\frac{1}{2\sqrt{2\pi}}[e^{i\omega T}(-ae^{-T} + b)(i\omega+1) + (-a + be^{-T})(-i\omega+1) \\
&- a(1 - e^{i\omega T - T})(i\omega+1) + (-i\omega+1)b(e^{i\omega T} - e^{-T})] \qquad (84) \\
= &\frac{1}{2\sqrt{2\pi}}[(e^{i\omega T}b - a)(i\omega+1) + (-a + be^{i\omega T})(-i\omega+1)] \\
= &\frac{1}{\sqrt{2\pi}}[e^{i\omega T}b - a].
\end{aligned}$$

This finally gives us the optimal control for the deadline control example.

7 The Riemann-Hilbert Problem for General Optimal Control With $x(T) = b$

As was done for the general problem without the final constraint $x(T) = b$, we will need $J[u]$ to be (37) for the exact same reasons as before. Again as before the dynamics are $\dot{x} = Ax + Bu$, $x(0) = a$ making $\widehat{x} = \chi\widehat{u}$ where $\chi(\omega) := (-i\omega I - A)^{-1}B$. Let $u_T(t)$ be the control supported on the interval $[0, T]$. Then following the same steps in taking the variational of $\widetilde{J}$ as was done to get (45) with $\widehat{u}_T$ in the place of $\widehat{u}_+$ we get

$$\begin{aligned}
\delta_{\widehat{u}_T}\widetilde{J} &= \int_{-\infty}^{\infty} \delta\widehat{u}_T'(-\omega)[(\chi'(-\omega)L_3'\chi + L_5'\chi)\widehat{u}_-(\omega) + (\chi'(-\omega)L_3\chi)\widehat{u}_-(\omega) \\
&\quad + (\chi'(-\omega)L_3\chi + L_4 + \chi'(-\omega)L_5)\widehat{u}_T(\omega) + (\chi'(-\omega)L_3'\chi + L_4' + L_5'\chi)\widehat{u}_T(\omega)]d\omega \\
&= \int_{-\infty}^{\infty} \delta\widehat{u}_T'(-\omega)[(\chi'(-\omega)L_3'\chi + L_5'\chi + \chi'(-\omega)L_3\chi)\widehat{u}_-(\omega) \\
&\quad + (\chi'(-\omega)L_3\chi + L_4 + \chi'(-\omega)L_5\chi'(-\omega)L_3'\chi + L_4' + L_5'\chi)\widehat{u}_T(\omega)]d\omega. \qquad (85)
\end{aligned}$$

This then gives us the Riemann-Hilbert problem

$$\begin{aligned}
&(\chi'(-\omega)L_3'\chi + L_5'\chi + \chi'(-\omega)L_3\chi)\widehat{u}_-(\omega) \\
&+(\chi'(-\omega)L_3\chi + L_4 + \chi'(-\omega)L_5\chi'(-\omega)L_3'\chi + L_4' + L_5'\chi)\widehat{u}_T(\omega) \\
&= Z_- + e^{i\omega T}\left(Z_+ + \sum_j \frac{\vec{\beta}_j}{\omega + i\lambda_j}\right), \qquad (86)
\end{aligned}$$

along with

$$e^{-i\omega T}\widehat{u}_T(\omega) = \widehat{u}_{-T}(\omega).$$

With λ_j being the eigenvalues of A, the term

$$\sum_j \frac{\vec{\beta}_j}{\omega + i\lambda_j}$$

is allowed for the same reason a similar term was allowed in the example deadline problem–that being that the variation $\delta\widehat{u}'_T(-\omega)$ must have zeros at $-i\lambda_j$. The variation has these zeros due to the fact that the final control $\widehat{u}_T$ must be fixed by

$$\begin{aligned}
x(T) &= e^{AT}x_0 + e^{AT}\int_0^T e^{-A\tau}Bu_T(\tau)d\tau = e^{AT}x_0 + e^{AT}\int_{-\infty}^{\infty} Ge^{-D\tau}G^{-1}Bu_T(\tau)d\tau \\
&= e^{AT}x_0 + \int_{-\infty}^{\infty}\begin{bmatrix} \sum_j\sum_k c_{1jk}e^{-\lambda_j\tau}u_{T,k}(\tau) \\ \sum_j\sum_k c_{2jk}e^{-\lambda_j\tau}u_{T,k}(\tau) \\ \vdots \end{bmatrix} d\tau \\
&= e^{AT}x_0 + \begin{bmatrix} \sum_j\sum_k c_{1jk}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\lambda_j\tau}u_{T,k}(\tau)d\tau \\ \sum_j\sum_k c_{2jk}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\lambda_j\tau}u_{T,k}(\tau)d\tau \\ \vdots \end{bmatrix} \\
&= e^{AT}x_0 + \begin{bmatrix} \sum_j\sum_k c_{1jk}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega\tau}u_{-,k}(\tau)d\tau|_{\omega=i\lambda_j} \\ \sum_j\sum_k c_{2jk}\sqrt{2\pi}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{i\omega\tau}u_{-,k}(\tau)d\tau|_{\omega=i\lambda_j} \\ \vdots \end{bmatrix} \\
&= e^{AT}x_0 + \begin{bmatrix} \sum_j\sum_k c_{1jk}\sqrt{2\pi}\widehat{u}_{T,k}(i\lambda_j) \\ \sum_j\sum_k c_{2jk}\sqrt{2\pi}\widehat{u}_{T,k}(i\lambda_j) \\ \vdots \end{bmatrix}.
\end{aligned} \tag{87}$$

We can now use the relationship

$$x(T) = e^{AT}x_0 + \begin{bmatrix} \sum_j \sum_k c_{1jk}\sqrt{2\pi}\widehat{u}_{T,k}(i\lambda_j) \\ \sum_j \sum_k c_{2jk}\sqrt{2\pi}\widehat{u}_{T,k}(i\lambda_j) \\ \vdots \end{bmatrix}$$

to solve for each β_j after solving the Riemann-Hilbert problem (86). Problem (86) is solved using the same trick to create removable singularities to put plus functions on one side and minus functions on the other side to use Liouville's theorem componentwise.

Finally, it should be mentioned that there are possibilities to derive a Riemann-Hilbert problem that takes into account the possibility for restrictions on possible values for $u(t)$ having allowed values in

$$U := \{u \in \mathbb{R}^m : h(u) \geq 0\}.$$

Accounting for this restriction would entail putting restrictions on the variation $\delta\widehat{u}'_T(-\omega)$ which allows (and actually would require) more freedom on the right hand side of the Riemann-Hilbert problem in (86) to solve the optimal control problem.

8 Overall Conclusion

We have seen the difficulties that arise in solving singular optimal control problems using standard methods in control theory. Here the beginning to an alternative method using the Riemann-Hilbert problem has been given which can hopefully be generalized to include control problems where A can have eigenvalues with positive real part. Given that singular control problems are difficult to solve using modern control theory, then typical techniques for solving Riemann-Hilbert problems could be useful in control theory applications.

References

1. Vincent and Grantham, "Non-Linear and Optimal Control" p. 317.

2. Rugh, "Linear System Theory" p. 204.

3. Lawrence C. Evans, "Partial Differential Equations" p. 560.

4. Dorato, Abdallah, and Cerone, "Linear-Quadratic Control" p. 14.

5. Lewis and Syrmose, "Optimal Control 2e" p. 131.

6. Lewis and Syrmose, "Optimal Control 2e" p. 139.

7. Vincent and Grantham, "Non-Linear and Optimal Control" p. 373.

8. Vincent and Grantham, "Non-Linear and Optimal Control" pp. 374–375.

9. Vincent and Grantham, "Non-Linear and Optimal Control" p. 404.

10. Kenneth Kuttler "Basic Analysis" p. 307

11. S. Glasgow, M. Meilstrup, J. Peatross, and M. Ware, "Real-time recoverable energy allocation in dispersive dielectrics," Phys. Rev. E 75, 016616 (2007).