AN ABSTRACT OF THE THESIS OF

Mehmet Emin ALPAY for the degree of Master of Science in Electrical and Computer Engineering presented on May 30, 1995.

Title: Existence and Uniqueness Results for l^1-optimal Linear Dynamic Controllers for Discrete Time SISO Systems

Redacted for Privacy

Abstract approved: ______________________

Molly H. Shor

The l^1-optimality problem which addresses the ability of a system to reject persistent disturbance signals has been of interest to many researchers during the last decade. Various designs involving linear and nonlinear controllers have been proposed and shown to operate with satisfactory performances. This study is mainly about the construction and properties of linear dynamic l^1-optimal controllers for SISO discrete time systems. We will start our discussion by obtaining an alternative formulation of the problem from a system-level point of view and show that some simple arguments lead to the same dual linear programming problem proposed by Dahleh and Pearson in their early work. Next we will show that the primary problem always has a solution; i.e. an l^1-optimal linear dynamic controller always exists. In particular the degenerate case where the dual linear programming problem has multiple solutions will be thoroughly investigated and it will be shown that the l^1-optimal controller design problem has a trivial solution in such cases. Furthermore we will prove that any secondary optimization problem imposed on the l^1-optimal controllers - presuming that a multitude of them exists - admits a solution corresponding to a corner of the constraint set restricting the parameters of the dual linear programming problem regardless of the nature of this secondary problem. The missing link between these arguments will be established in the following section where it will be shown that existence of multiple solutions for the l^1-optimal controller is a pathological case which is highly unlikely to occur in practice even in the degenerate case where the dual linear programming problem has multiple solutions. Indeed,

we will show that the probability of obtaining multiple l^1-optimal controllers is zero under a Lebesgue-type continuous probability measure defined on $R^{n \times n}$. Still we will discuss solution techniques for H_2-optimization problem imposed as a secondary performance objective on the linear dynamic l^1-optimal controllers in the off-chance that there is a multitude of them and show how the previous arguments can be used to convert this problem to an ordinary quadratic programming problem solvable by available techniques.

Existence and Uniqueness Results for l^1-optimal Linear Dynamic Controllers for Discrete Time SISO Systems

by

Mehmet Emin ALPAY

A THESIS

submitted to

Oregon State University

in partial fulfillment of the

requirements for the degree of

Master of Science

Completed May 30, 1995

Commencement June 1996

Master of Science thesis of Mehmet Emin ALPAY presented on May 30, 1995

APPROVED:

Redacted for Privacy

Major Professor, representing Electrical and Computer Engineering

Redacted for Privacy

Chair of Electrical and Computer Engineering Department

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Mehmet Emin ALPAY, Author

To my mother

ACKNOWLEDGEMENTS

I am grateful to all the people whose support made it possible for me to finish this thesis. In particular I would like thank my major professor Molly H. Shor whose unfailing support and motivation helped me through many phases of this study, Prof. Bent Petersen of the Math. Dept. who introduced me to the exciting field of functional analysis, and all the faculty members and fellow students whose ideas and suggestions became a valuable source of inspiration for me.

TABLE OF CONTENTS

LIST OF FIGURES

Page

Existence and Uniqueness Results for Linear Dynamic l^1-optimal Controllers for Discrete Time SISO Systems

INTRODUCTION

The l^1-optimization problem involves the construction of feedback controllers in such a way that the ability of the closed-loop system to reject persistent bounded disturbances is maximized. Even though it is a relatively new topic in the control area, it has been extensively studied since it was first formulated by Vidyasagar [2] in 1986. Various researchers have proposed different design techniques through a variety of mathematical approaches and many important results regarding different types of systems and controllers have emerged (see [3]-[15]). Nevertheless the topic is far from being exhausted yet and the studies in the area continue to uncover new properties of the l^1-optimization problem.

The main reason for this undiminishing interest in l^1-optimization problem is the fact that having modeled the disturbances as bounded but otherwise undetermined signals, this approach is more realistic than the still popular H_∞-optimization problem where the disturbances are restricted to be finite energy signals. Although the H_∞-optimization problem has been extensively studied by many researchers for a long time, it is inherently inadequate to handle the persistent disturbances which occur quite often in practice. On the other hand, the very nature of the l^1-optimization makes it the perfect approach for such problems. Moreover the l^1 space is the natural domain for solving any design problem demanding a uniform performance over all time; e.g., we can never afford to let the altitude of an airplane fall below zero even for the briefest period of time; or we would never like to see a robot arm move outside a certain region and possibly damage itself or the equipment around it. The list can be made longer, confirming that such problems are abundant in practical life, which makes the l^1-optimization problem all the more important for control engineers.

This study has been largely inspired by the works of Dahleh and Pearson ([3],[4]) who showed that the problem of constructing l^1-optimal linear dynamic controllers for discrete time systems was equivalent to the solution of a linear programming problem. The order of the thesis is as follows: after a brief review of the literature in the area, we will first obtain an alternative formulation of the problem using system-level arguments and show that these arguments lead to the identical linear programming problem addressed in [3]. Then we will give a brief overview of the solution of the l^1-optimization problem and the dual linear programming problem and show that the l^1-optimal linear dynamic controller can always be constructed; i.e., the primary problem always has a solution. The fourth chapter addresses the degenerate case where the linear programming problem associated with the l^1-optimization problem has multiple solutions. We will show that there exists a simple solution to the primary problem in such cases and we will also show that any further optimization objectives imposed on the l^1-optimal controllers - presuming that a multitude of them exists in the degenerate case - will always lead to solutions at the corners of the constraint hyper-polygon. Then we will show that the l^1-optimal linear dynamic controllers are almost always unique - even in the degenerate case - and discuss the consequences of this finding in the degenerate case where, now, the simple solution is known to be almost always the unique solution. In the following section are developed solution techniques for the H_2-optimization problem imposed on the l^1-optimal linear dynamic controllers in the off-chance that a multitude of them exists, which also illustrates how the previous arguments can be used to transform such a problem into an ordinary quadratic programming problem solvable by readily available techniques. Chapter five contains examples of the degenerate cases for both the primary and the dual problem and illustrates the ideas developed in the previous parts. Final remarks and suggestions for further studies in the area are in the conclusion.

LITERATURE REVIEW

As mentioned earlier, the l^1-optimization problem was first formulated by M. Vidyasagar ([2]) where he obtained solutions for SISO discrete time minimum-phase plants and plants with a single unstable zero and also proposed a procedure for obtaining bounds on the attainable performances and constructing suboptimal solutions. Then M.A. Dahleh and J.B. Pearson showed in [3] that the problem could be converted to a dual linear programming problem from which the optimal linear dynamic controller could be constructed by invoking alignment conditions. One of the most important results of this study was the observation that the optimal controller always had a rational transfer function and hence easily implementable, giving the problem much practical value and significance. These results were extended to MIMO systems by the same researchers ([4]), who then proceeded to investigate the continuous time equivalent of the problem, namely the L^1-optimization problem ([5]). It was shown in this work that the continuous time case had quite different features than its discrete time counterpart, yielding irrational transfer functions for the control block even in the cases where all plant transfer functions were rational. Still, Dahleh and Pearson were able to propose two different methods to construct appropriate controllers. The former method involved solving a finite dimensional nonlinear programming problem leading to the exact solution, whereas the latter involved solving a finite dimensional linear programming problem leading to an approximate solution. They further refined their technique in [6] where they obtained a more concise set of conditions for the solvability of the problem by dropping certain matrix rank conditions and proposed an iterative algorithm to construct the optimal controller, each step of which involved solving a finite dimensional linear programming problem yielding a suboptimal controller.

It was after this point when the topic started attracting many researchers who consequently uncovered various features of the problem and proposed different approaches and solution techniques. Meyer pointed out in [7] that the l^1-optimal controller could be of arbitrarily high order and need not be unique. G. Deodhare and M. Vidyasagar investigated when a stabilizing controller is l^1-optimal for some stable weighting function and came up with the interesting result that unlike H_∞-optimality, l^1-optimality of a controller is preserved for a variety of weighting functions which are not trivially related ([8]). In [9], M. Dahleh and M.A. Dahleh devised an adaptive controller using l^1-optimal control law in conjunction with a constrained projection estimator with a dead zone and studied its convergence properties. Meanwhile, M. Vidyasagar extended the results of [4] and [5] to include the cases where the plant has zeros or poles on the stability boundary (the extended $j\omega$-axis in continuous time, the unit circle in discrete time) and showed that the optimal controller might not exist in such cases ([10]). He also proposed an iterative algorithm of constructing suboptimal controllers with performances converging to the unattainable optimal result.

Until then, the research in the area had been mainly focused on the linear time-invariant controllers. In [11] J. Shamma and M.A. Dahleh compared time-invariant and time-variant compensators; but their results were not so encouraging: they proved that the time-variant compensators could neither improve the optimal rejection - regardless of the system and/or signal norm used - nor help achieving BIBO robust stability with unstructured plant uncertainty. They obtained similar results in [12] where the nonlinear time-variant controllers were studied. Nevertheless, Shamma showed in [13] that under full state feedback, memoryless nonlinear state feedback controllers yielding the same performance as linear dynamic controllers could be constructed, which proved that nonlinear controllers had an important advantage over the linear ones which could be of arbitrarily high order and too complicated to implement.

As examples of the recent studies in the area, M.A. Dahleh and I.J. Diaz-Bobillo have addressed the multiblock problem and have obtained a concise representation of the interpolation conditions by tying them to the matrix theory ([14]). Furthermore, they have studied the structure of optimal solutions to distinguish between classes of multiblock problems and have proposed a new method of iteratively computing suboptimal solutions avoiding order inflation. F. Blanchini and M. Sznaier, on the other hand, have devised an algorithm of computing suboptimal controllers with guaranteed cost bounds using discrete Euler approximation for continuous time case ([15]). The research in the area regarding both continuous and discrete time systems continues intensively, but the focus seems to have shifted from the theoretical aspects to the practical concerns like computational complexity, order restrictions, etc., which forces the researchers to concentrate on suboptimal rather than optimal controller design schemes.

AN ALTERNATIVE FORMULATION

In this chapter we will gain further insight to the l^1-optimization problem and the dual linear programming problem associated with it using a few basic system-level arguments. However, to define the problem in a mathematically concise way, we first give some definitions and prove a lemma that justifies our calling this problem the l^1*-optimization* problem.

Mathematical Preliminaries - I

Definition A causal sequence $\{x(k)\}_{k=0}^{\infty}$ is called an l^1-sequence if it is absolutely summable; i.e.,

$$\sum_{k=0}^{\infty} |x(k)| < \infty$$

and the l^1-norm of such a sequence is defined as

$$||x||_1 = \sum_{k=0}^{\infty} |x(k)|$$

Definition A causal sequence $\{x(k)\}_{k=0}^{\infty}$ is called an l^∞-sequence if it is bounded; i.e.,

$$\sup_{k\geq 0} |x(k)| < \infty$$

and the l^∞-norm of such a sequence is defined as

$$||x||_\infty = \sup_{k\geq 0} |x(k)|$$

Lemma 3.1 *Let $x, y \in l^\infty$, $h \in l^1$ and $y = x * h$; then*

1. $||y||_\infty \leq ||h||_1 ||x||_\infty$, *and*

2. $\sup_{||x||_\infty \neq 0} \frac{||y||_\infty}{||x||_\infty} = ||h||_1$

Proof

1. Since $y = x * h$, we have

$$y(n) = \sum_{k=0}^{\infty} h(k)x(n-k)$$

hence

$$\begin{aligned} |y(n)| &\leq \textstyle\sum_{k=0}^{\infty} |h(k)x(n-k)| \\ &\leq \sup_k |x(n-k)| \textstyle\sum_{k=0}^{\infty} |h(k)| \\ &= ||x||_\infty ||h||_1 \end{aligned}$$

for all $n < \infty$, which implies

$$||y||_\infty \leq ||x||_\infty ||h||_1$$

2. For any given n, we can choose a sequence $\{x(k)\}_{k=0}^{\infty}$ such that

$$x(k) = \begin{cases} sgn[h(n-k)] & \text{for } k = 0, \ldots, n \\ 0 & \text{otherwise} \end{cases}$$

With this choice for x obviously $||x||_\infty = 1$ and

$$y(n) = \sum_{k=0}^{n} |h(k)|$$

Since $\{h(k)\}_{k=0}^{\infty}$ is an absolutely convergent sequence, for any $\epsilon > 0$ we can choose n sufficiently large to satisfy

$$y(n) = \sum_{k=0}^{n} |h(k)| > \sum_{k=0}^{\infty} |h(k)| - \epsilon = ||h||_1 - \epsilon$$

$$\Rightarrow \sup_{n \geq 0} |y(n)| = ||y||_\infty \geq ||h||_1 = ||h||_1 ||x||_\infty$$

Combining with the result of (1) we have

$$\sup_{||x||_\infty \neq 0} \frac{||y||_\infty}{||x||_\infty} = \sup_{||x||_\infty = 1} ||y||_\infty = ||h||_1$$

Q.E.D.

The above lemma justifies the name l^1-*optimization*; for the rest of this thesis we will be dealing with bounded (l^∞) type signals and try to minimize the norm induced by output/input ratio of their l^∞-norms, which is equivalent to the l^1-norm of the impulse response corresponding to the transfer function between such signals, as we have proved above. Note that each system can be viewed as a transformation between signals and had we not restricted these transformations to convolution operations, though the original l^∞-induced norm would always exist, the phrase l^1-norm would either make no sense or would not be equal to the induced norm except for in some rare cases.

An Alternative Formulation for the l^1-optimization Problem

For the rest of this thesis, we will deal with causal discrete time linear time-invariant systems with stable real rational transfer functions as illustrated below and refer to the linear dynamic l^1-optimal controller as the l^1-optimal controller for convenience.

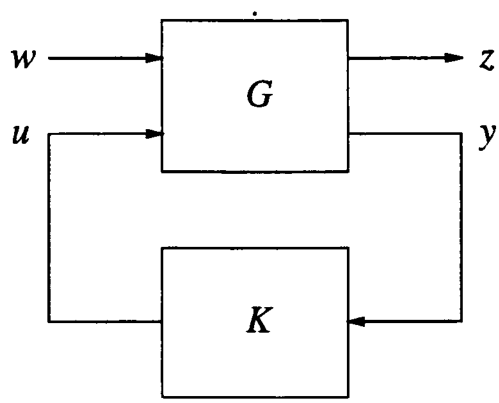

Figure 3.1: Block diagram of the SISO discrete time system studied in this thesis

In the diagram above w is the disturbance, u is the input signal, y is the measured output and z is the regulated output. With the convention that the capital letters denote the unilateral z-transforms of the signals as defined below

$$X(z) = X = \sum_{k=0}^{\infty} x(k) z^k$$

the system equations and feedback are as follows

$$\begin{aligned} Z &= G_{11}W + G_{12}U \\ Y &= G_{21}W + G_{22}U \\ U &= KY \end{aligned} \tag{3.1}$$

where G_{ij} are all real rational stable transfer functions. From above equations we can obtain the transfer function from the disturbance w to the regulated output z in a straightforward manner as

$$\Phi = G_{11} + \frac{KG_{21}G_{12}}{1 - G_{22}K} \tag{3.2}$$

Our aim is to minimize the l^∞-induced norm

$$\sup_{||w||_\infty \leq 1} \frac{||z||_\infty}{||w||_\infty} \tag{3.3}$$

and hence minimize the effect of the disturbance on the output. But from Lemma 3.1 we already know that this induced norm is identical to the l^1-norm of the impulse response corresponding to this transfer function. Thus we can state the problem in a much more concise way as finding the appropriate dynamic feedback that minimizes $||\phi||_1$ where

$$\phi = Z^{-1}\{\Phi\} = Z^{-1}\{G_{11} + \frac{KG_{21}G_{12}}{1 - G_{22}K}\}$$

For all practical purposes, the controller K should also be a stabilizing controller and itself stable. By using the YJBK parameterization of such controllers ([16],[17]), it can be shown ([3],[4]) that Φ can be expressed as

$$\Phi = H - FQ \tag{3.4}$$

where Q is any stable transfer function and H and F can be obtained from the system transfer functions G_{ij}, $i, j = 1, 2$, in the following manner:

$$\begin{aligned} H &= G_{11} + G_{12}G_{21}YM \\ F &= G_{12}G_{21}M^2 \end{aligned} \tag{3.5}$$

where $G_{22} = N/M$ such that M and N are stable, coprime, and satisfy the Bezout's identity

$$MX - NY = 1 \tag{3.6}$$

for some stable X and Y.

Now the question becomes that of finding the appropriate stable transfer function Q that minimizes $||\phi||_1$ where ϕ is the inverse z-transform of Φ which is depicted above. Now let us define

$$R = FQ \tag{3.7}$$

There is no restriction on the transfer function Q other than being a stable transfer function, which is equivalent to having all its poles outside the unit circle. A direct consequence of this restriction is that any zeros of F inside the unit circle cannot be canceled by the poles of Q. Let a_i, $i = 1, \ldots, n$ denote these zeros; for convenience we will assume that all of these zeros have single multiplicities, though the results can be extended to cover cases involving zeros with higher multiplicities as well. Then there is the extra constraint on R such that other than being a stable transfer function, it also must satisfy

$$R(a_i) = 0$$

for all $i = 1, \ldots, n$.

The condition of stability is equivalent to $r = Z^{-1}\{R\}$ being an l^1-sequence; hence if we can find a way of converting the second set of constraints on R, i.e., $R(a_i) = 0$, $i = 1, \ldots, n$, to a constraint in the time-domain, then the whole problem could be easily solved as an ordinary minimum norm problem in the l^1 space with these constraints defining a subspace in l^1 and we would be looking for the minimum distance from h to this subspace and the element for which it is achieved - if it exists.

Indeed, the above approach is exactly what was done in [3] and [4] to solve the problem. But to get a better understanding of the problem at system-level, we will try a different approach and show that some simple arguments almost lead us to the same conclusions. To do this, let us consider the following discussion: the a_i are the zeros of

R, which means had we been dealing with unilateral sequences $\{a_i^{-k}\}_{k=0}^{\infty}$, the response of the closed-loop system to such signals would have been $\{[H(a_i) - R(a_i)]a_i^{-k}\}_{k=0}^{\infty}$, which, by hypothesis, would be equal to $\{H(a_i)a_i^{-k}\}_{k=0}^{\infty}$ alone, yielding a system gain of $|H(a_i)|$. In other words, such an input would, figuratively, tie the hands of our controller completely. Obviously the same reasoning would apply to any linear combination of such signals in the sense that the controller would be equally helpless in handling such inputs, which immediately suggests that the combination leading to the highest system gain would give a lower bound on the optimal l^1-norm attainable.

Before jumping to such conclusions, however, the crude philosophy above should be further refined and put in the appropriate mathematical setting. One major technical problem that needs to be overcome is the fact that the sequences $\{a_i^{-k}\}_{k=0}^{\infty}$ are not bounded since all a_i are within the unit circle by hypothesis; consequently, such signals should not be a part of our discussion at the first place! But what if we suppress their divergent "tails" and use these new bounded signals as inputs? As we will show shortly, the essence of the reasoning does not change and the discussion above proves to be an accurate explanation of the dual linear programming problem associated with the l^1-optimization problem.

Theorem 3.1 *Let a_i, $i = 1, \ldots, n$ be the zeros of F inside the unit circle, $R = FQ$, where Q is a stable, real, rational transfer function and F be obtained from the YJBK parameterization of the transfer function Φ (3.2) as in (3.5). Then the solution of the following linear programming problem provides a lower bound for the l^1-optimal performance of the system described by (3.1):*

$$\max_{\alpha_i} \sum_{i=1}^{m} \alpha_i H(a_i) + \sum_{i=m+1}^{m+l} \alpha_i Re(H(a_i)) + \sum_{i=m+l+1}^{n} \alpha_i Im(H(a_{i-l}))$$

subject to the constraints

$$\left| \sum_{i=1}^{m} \alpha_i a_i^k + \sum_{i=m+1}^{m+l} \alpha_i Re(a_i^k) + \sum_{i=m+l+1}^{n} \alpha_i Im(a_{i-l}^k) \right| \leq 1$$

for all $k = 0, 1, 2, \ldots$.

Proof To prove this theorem, we will construct disturbance signals of the form discussed above and show that such signals yield an l^∞-induced norm which is greater than or equal to the result of the linear programming problem stated in the theorem.

By hypothesis, all a_i are within the unit circle, and since F is real and rational, they appear as either real numbers or complex conjugate pairs. Thus, let these zeros be ordered as:

$$a_i \in R \qquad i = 1, \ldots, m$$
$$\bar{a}_i = a_{i+l} \quad i = m+1, \ldots, m+l$$

where $m + 2l = n$. Now let us define the disturbance sequences $\{w_i(k)\}_{k=0}^{\infty}$ in the following manner:

for $1 \leq i \leq m$:

$$w_i(k) = \begin{cases} a_i^{N-k} & \text{for } k = 0, \ldots, N \\ 0 & \text{for } k > N \end{cases}$$

for $m+1 \leq i \leq m+l$:

$$w_i(k) = \begin{cases} Re(a_i^{N-k}) & \text{for } k = 0, \ldots, N \\ 0 & \text{for } k > N \end{cases}$$

for $m+l+1 \leq i \leq n$:

$$w_i(k) = \begin{cases} Im(a_{i-l}^{N-k}) & \text{for } k = 0, \ldots, N \\ 0 & \text{for } k > N \end{cases}$$

Obviously each w_i is a sequence of real numbers with $||w_i||_\infty \leq 1$. Now define w as a linear combination of such disturbances so that $||w||_\infty \leq 1$. This can be done by setting

$$w = \alpha_1 w_1 + \cdots + \alpha_n w_n$$

such that

$$|w(k)| \leq 1$$

for all $k = 0, 1, 2, \ldots$. By substituting in the open forms of w_i, this condition can be re-written as

$$\left| \sum_{i=1}^{m} \alpha_i a_i^{N-k} + \sum_{i=m+1}^{m+l} \alpha_i Re(a_i^{N-k}) + \sum_{i=m+l+1}^{n} \alpha_i Im(a_{i-l}^{N-k}) \right| \leq 1$$

for $k = 0, 1, \ldots, N$. Now let $y_0 = (h-r)*w$; then it can be shown in a straightforward manner that

$$\begin{aligned} y_0(N) &= \left[\sum_{i=1}^{m}\alpha_i\sum_{q=0}^{N}h(q)a_i^q + \sum_{i=m+1}^{m+l}\alpha_i\sum_{q=0}^{N}h(q)Re(a_i^q) + \sum_{i=m+l+1}^{n}\alpha_i\sum_{q=0}^{N}h(q)Im(a_{i-l}^q)\right] \\ &\quad - \left[\sum_{i=1}^{m}\alpha_i\sum_{q=0}^{N}r(q)a_i^q + \sum_{i=m+1}^{m+l}\alpha_i\sum_{q=0}^{N}r(q)Re(a_i^q) + \sum_{i=m+l+1}^{n}\alpha_i\sum_{q=0}^{N}r(q)Im(a_{i-l}^q)\right] \end{aligned}$$

Recalling that

$$\sum_{q=0}^{\infty}h(q)a_i^q = H(a_i)$$

and

$$\sum_{q=0}^{\infty}r(q)a_i^q = R(a_i) = 0$$

i.e., all the partial sums involved in the above equation are convergent, for any $\epsilon > 0$ and any set of α_i we can choose N such that

$$\begin{aligned} &\left|\sum_{i=1}^{m}\alpha_i\sum_{q=0}^{N}h(q)a_i^q + \sum_{i=m+1}^{m+l}\alpha_i\sum_{q=0}^{N}h(q)Re(a_i^q) + \sum_{i=m+l+1}^{n}\alpha_i\sum_{q=0}^{N}h(q)Im(a_{i-l}^q)\right| \geq \\ &\qquad \left|\sum_{i=1}^{m}\alpha_i H(a_i) + \sum_{i=m+1}^{m+l}\alpha_i Re(H(a_i)) + \sum_{i=m+l+1}^{n}\alpha_i Im(H(a_{i-l}))\right| - \epsilon/2 \end{aligned}$$

and

$$\left|\sum_{i=1}^{m}\alpha_i\sum_{q=0}^{N}r(q)a_i^q + \sum_{i=m+1}^{m+l}\alpha_i\sum_{q=0}^{N}r(q)Re(a_i^q) + \sum_{i=m+l+1}^{n}\alpha_i\sum_{q=0}^{N}r(q)Im(a_{i-l}^q)\right| \leq \epsilon/2$$

which immediately yields the following inequality

$$|y_0(N)| \geq \left|\sum_{i=1}^{m}\alpha_i H(a_i) + \sum_{i=m+1}^{m+l}\alpha_i Re(H(a_i)) + \sum_{i=m+l+1}^{n}\alpha_i Im(H(a_{i-l}))\right| - \epsilon$$

Since $||y_0||_\infty \geq |y_0(N)|$ and the above result is true for any $\epsilon > 0$ as $N \to \infty$ we have

$$||y_0||_\infty \geq \max_{\alpha_i}\left|\sum_{i=1}^{m}\alpha_i H(a_i) + \sum_{i=m+1}^{m+l}\alpha_i Re(H(a_i)) + \sum_{i=m+l+1}^{n}\alpha_i Im(H(a_{i-l}))\right|$$

for all α_i satisfying the constraints as $N \to \infty$. Moreover, since the constraints constitute a region completely symmetric about the origin the maximization of the absolute value of the expression above is equivalent to the maximization of the actual expression. Combining this with the constraint inequalities we obtain

$$\begin{aligned} ||h - r||_1 &= sup_{||w||_\infty \leq 1} ||y||_\infty \\ &\geq ||y_0||_\infty \\ &\geq \max_{\alpha_i} \sum_{i=1}^{m} \alpha_i H(a_i) + \sum_{i=m+1}^{m+l} \alpha_i Re(H(a_i)) + \sum_{i=m+l+1}^{n} \alpha_i Im(H(a_{i-l})) \end{aligned}$$

where α_i are subject to

$$\left| \sum_{i=1}^{m} \alpha_i a_i^k + \sum_{i=m+1}^{m+l} \alpha_i Re(a_i^k) + \sum_{i=m+l+1}^{n} \alpha_i Im(a_{i-l}^k) \right| \leq 1$$

for all $k = 0, 1, 2, \dots$ and any choice of r satisfying $R(a_i) = 0$, $i = 1, \dots, n$. Taking the infimum of the left side over r will not affect the inequality and thus we have proved our claim that the result of the linear programming problem stated in the hypothesis imposes a lower bound on the l^1-optimal performance.
Q.E.D.

Not surprisingly, the linear programming problem whose solution has been proved to provide a lower bound for the l^1-optimal performance is identical to the one in [3]. As we claimed at the beginning, some simple system-level arguments have thus lead us almost to the same conclusions about the solution of the l^1-optimization problem: *almost*, because we have not solved the problem completely but merely obtained a lower bound for the attainable l^1-optimal performance through the solution of a linear programming problem. The fact that the solution of this linear programming problem not only provides a lower bound for the l^1-optimal performance but is actually equal to this value can be established by using the duality theorem in much the same way as in [3]. Hence, it will not be repeated here, since the main objective of this section was not to solve the problem but give the reader an idea of the motivation behind converting the l^1-optimization problem to a dual linear programming problem using some basic system-level arguments.

SOLUTION OF THE l^1-OPTIMIZATION PROBLEM

In this chapter we will give a brief overview of the solution for the l^1-optimal linear dynamic feedback controller for the system described in the previous section. For completeness, however, we will start our discussion with some mathematical preliminaries which will be often referred to during future discussions.

Mathematical Preliminaries - II

Definition Let X be a linear normed space. Then its dual space is defined as the space of linear continuous functionals on X and is denoted by X^*.

Definition The norm of a functional $x^* \in X^*$ is defined as

$$||x^*|| = \sup_{||x|| \leq 1} |x^*(x)| = \sup_{||x|| \leq 1} | < x, x* > |$$

Definition A vector $x^* \in X^*$ is said to be aligned with a vector $x \in X$ if $< x, x^* >= ||x||\, ||x^*||$.

Proposition 4.1 *Any continuous linear functional f on l^1 can be represented by an l^∞-sequence y such that*

$$f(x) = \sum_{k=0}^{\infty} x(k)y(k)$$

for all $x \in l^1$. In this sense, l^∞ is the dual space of l^1.

Theorem 4.1 *(Duality Theorem) Let x be an element in a real normed linear space X and let μ denote its distance from a subspace S of X. Then*

$$\mu = \inf_{k \in S} ||x - k|| = \max_{r \in BS^\perp} < x, r >$$

where

$$BS^\perp = \{x^* \in X^* | < x, x^* >= 0 \text{ for all } x \in S \text{ and } ||x^*|| \leq 1 \}$$

The maximum on the right is achieved for some $r_0 \in BS^\perp$ with $||r_0|| = 1$. Moreover, if the infimum on the left is achieved for some $x_0 \in S$, then r_0 is aligned with $x - x_0$.

The proof of Proposition 4.1 and the duality theorem can be found in almost any text on functional analysis (see [18] for example). The major contribution of the duality theorem is that it gives us the means of converting the original (primary) optimization problem to another (dual) problem for which the solution is always known to exist and in many cases easier to find. This technique of solving the dual problem and constructing the solution to the primary problem by invoking the alignment condition together with the constraints from the hypothesis of the original problem is widely used in many optimization schemes.

Of particular interest to us are the consequences of the alignment condition where the primary space is l^1 with dual space l^∞. The following lemma gives the necessary and sufficient conditions for alignment in this case.

Lemma 4.1 *Let $x \in l^1$ and $r \in l^\infty$. Then x and r are aligned if and only if both following conditions hold:*

1. $x(k)r(k) \geq 0$ *for all* k
2. $x(k) = 0$ *whenever* $|r(k)| \neq ||r||_\infty$

Proof

1. (if) Assume (1) and (2) hold; then

$$\begin{aligned} < x, r > \; = \; \textstyle\sum_{k=0}^{\infty} x(k)r(k) \; &= \; \textstyle\sum_{k=0}^{\infty} |x(k)||r(k)| \\ &= \; ||r||_\infty \textstyle\sum_{k=0}^{\infty} |x(k)| \\ &= \; ||r||_\infty ||x||_1 \end{aligned}$$

Hence, x and r are aligned by definition.

2. (only if) Assume (1), (2) or both do not hold. Then we have the following:

If (1) does not hold, there exists a k such that $x(k)r(k) < 0$. Then

$$< x, r > = \sum_{k=0}^{\infty} x(k)r(k) < \sum_{k=0}^{\infty} |x(k)||r(k)| \leq ||r||_\infty ||x||_1$$

Hence x and r are not aligned.

If (2) does not hold, there exists a k such that $x(k) \neq 0$ and $|r(k)| = \xi < ||r||_\infty$. Then

$$< x, r > = \sum_{k=0}^{\infty} x(k)r(k) \leq \sum_{k=0}^{\infty} |x(k)||r(k)| < ||r||_\infty ||x||_1$$

Hence x and r are not aligned.

If both (1) and (2) do not hold, from the above inequalities we immediately obtain

$$<x,r>=\sum_{k=0}^{\infty}x(k)r(k)<\sum_{k=0}^{\infty}|x(k)||r(k)|<||r||_{\infty}||x||_1$$

and hence x and r are not aligned.

Q.E.D.

The Linear Programming Problem Associated with the l^1-optimization Problem

It has been shown in [3] that the dual problem to the original l^1-optimization problem can be set as a linear programming problem with infinitely many constraints as follows

$$\begin{aligned}\mu &= \inf_{R\in S}||h-r||_1\\ &= \max_{\alpha_i}\left\{\sum_{i=1}^{m}\alpha_i H(a_i)+\sum_{i=m+1}^{m+l}\alpha_i Re(H(a_i))+\sum_{i=m+l}^{n}\alpha_i Im(H(a_{i-l}))\right\}\end{aligned} \quad (4.1)$$

subject to the constraints

$$-1\le\sum_{i=1}^{m}\alpha_i a_i^k+\sum_{i=m+1}^{m+l}\alpha_i Re(a_i^k)+\sum_{i=m+l}^{n}\alpha_i Im(a_{i-l}^k)\le 1 \quad (4.2)$$

where $k=0,1,2,\ldots$. S denotes the set of stable transfer functions that satisfy $R(a_i)=0$ for $i=1,\ldots,n$ and a_i are the zeros of F (3.5) inside the unit circle.

Although the dual problem is a linear programming problem for which various computational solution algorithms are available, the fact that we have infinitely many constraints to satisfy is quite a distressing one since we possibly cannot deal with all of them using a computer. A closer look at the constraint inequalities, however, reveals that the coefficients of α_i are all in the form a_i^k and given all a_i lie within the unit circle, they would all tend to zero, leaving us with only a finite number of relevant constraints which, when satisfied, guarantee that all the remaining constraints are also satisfied. The next theorem shows us that this is indeed the case.

Theorem 4.2 *Only a finite number of the constraint inequalities are relevant.*

Proof Define

$$b_i = \left[\begin{array}{cccccccccc} a_1^i & \cdots & a_m^i & Re(a_{m+1}^i) & \cdots & Re(a_{m+l}^i) & Im(a_{m+1}^i) & \cdots & Im(a_{m+l}^i) \end{array} \right]'$$

$$B = [b_1 \cdots b_n]$$

and

$$\alpha = [\alpha_1 \cdots \alpha_n]'$$

By elementary column operations, it can easily be shown that the $n \times n$ matrix B is equivalent to a Van der Monde matrix and since all a_i are distinct by hypothesis, B is non-singular. Hence the first n constraints form a basis for R^n.

Now let $k > n$ and assume that the first n constraint inequalities are satisfied by a particular α; i.e., it is known that

$$|\alpha' b_i| \leq 1$$

for $i = 1, \ldots, n$. We want to show that the additional constraint $|\alpha' b_k| \leq 1$ is automatically satisfied for sufficiently large values of k, rendering this constraint irrelevant.

From previous discussion we know that B is nonsingular, hence

$$b_k = B\beta \Rightarrow \beta = B^{-1} b_k$$

Decomposing B^{-1} in terms of its row vectors;

$$B^{-1} = \left[\begin{array}{c} v_1' \\ \vdots \\ v_n' \end{array} \right]$$

we immediately obtain the following inequality

$$|\beta_i| \leq ||v_i||_1 \xi^k$$

where $||v_i||_1$ is the sum of the absolute values of the entries of vector v_i and ξ is defined as

$$\xi = \max_{1 \leq i \leq n} |a_i|$$

Then

$$|\sum_{i=1}^{m} \alpha_i a_i^k + \sum_{i=m+1}^{m+l} \alpha_i Re(a_i^k) + \sum_{i=m+l}^{n} \alpha_i Im(a_{i-l}^k)|$$

$$
\begin{aligned}
&= |\alpha' B \beta| \\
&= |[\alpha' b_1 \cdots \alpha' b_n] \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}| \\
&\le \sum_{i=1}^{n} |\beta_i| \\
&\le (\sum_{i=1}^{n} ||v_i||_1)\xi^k
\end{aligned}
$$

Let

$$M = \sum_{i=1}^{n} ||v_i||_1$$

Clearly M is a finite positive number. Recalling that all a_i are within the unit circle, we immediately conclude that $\xi < 1$; hence if k is sufficiently large, namely if

$$k > \frac{\log 1/M}{\log \xi} = N$$

where $N < \infty$, then $M\xi^k < 1$ and for all $k > N$, the constraint inequalities are automatically satisfied and hence irrelevant.

Q.E.D.

Having proved this important theorem, we now know that we can solve the linear programming problem associated with the primary l^1-optimization problem once we obtain the relevant constraint inequalities of which we know there are finitely many. At this point, what we need is a criterion which we can use to distinguish the relevant inequalities from the irrelevant ones and the following lemma provides an answer to this question.

Definition Let $b_i \in R^n$, $i = 1, \ldots, m$ be n-dimensional real vectors. A sub-convex combination of these vectors is defined as a linear combination of the form

$$\beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_m b_m$$

where β_i satisfy the following condition

$$\sum_{i=1}^{m} |\beta_i| \le 1$$

Lemma 4.2 *Let $|b_i' x| \le 1$ for $i = 1, \ldots, m$ with $b_i \in R^n$. An additional inequality $|b_{m+1}' x| \le 1$ is irrelevant if and only if b_{m+1} can be expressed as a sub-convex linear combination of b_i.*

Proof

1. (if) Assume b_{m+1} can be expressed as a sub-convex combination of b_i, $i = 1, \ldots, m$, i.e.,

$$b_{m+1} = \sum_{i=1}^{m} \beta_i b_i$$

where

$$\sum_{i=1}^{m} |\beta_i| \leq 1$$

Let $x \in R^n$ satisfy all first m inequalities; i.e., $|b_i'x| \leq 1$ for $i = 1, \ldots, m$. Then

$$\begin{aligned} |b_{m+1}'x| &= |\textstyle\sum_{i=1}^{m} \beta_i b_i' x| \\ &\leq \textstyle\sum_{i=1}^{m} |\beta_i||b_i'x| \\ &\leq \textstyle\sum_{i=1}^{m} |\beta_i| \\ &\leq 1 \end{aligned}$$

Hence any $x \in R^n$ that satisfies the first m constraints automatically satisfies the $m+1$st one as well and this last inequality is irrelevant.

2. (only if) Assume $|b_{m+1}'x| \leq 1$ is irrelevant and b_{m+1} cannot be expressed as a sub-convex combination. There are two possibilities.

case 1 b_{m+1} cannot be expressed as a linear combination of the remaining b_i. In such a case, we know from linear algebra that there exists a vector $x \in R^n$ such that

$$b_i'x = \begin{cases} 0 & \text{if } i = 1, \ldots, m \\ 1 & \text{if } i = m+1 \end{cases}$$

and by scaling this x by $\lambda > 1$, it is evident that we can construct a vector $y = \lambda x$ that satisfies all the first m inequalities but not the $m+1$st one, which means the last inequality is relevant (contradiction).

case 2 b_{m+1} can be expressed as

$$b_{m+1} = \beta_1 b_1 + \cdots + \beta_m b_m$$

where by hypothesis

$$\sum_{k=1}^{m} |\beta_k| > 1$$

Without loss of generality we may assume that b_i, $i = 1, \ldots, m$ are linearly independent, (if not, we can use any linearly independent subset of b_i to obtain the above representation).

Then the equation

$$\begin{bmatrix} b_1' \\ \vdots \\ b_m' \end{bmatrix} x = \begin{bmatrix} sgn(\beta_1) \\ \vdots \\ sgn(\beta_m) \end{bmatrix}$$

has a solution since the matrix on the left has full row rank. Let x be this solution; then

$$|b_i'x| = |sgn(\beta_i)| = 1$$

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

$$b_{m+1}'x = \sum_{k=1}^{m} |\beta_k| > 1$$

by hypothesis. Hence x satisfies all inequalities except the last, which implies that the last inequality is relevant (contradiction).
Q.E.D.

At this point, we should emphasize once more that the term *irrelevant* refers to those inequalities which are automatically satisfied once the relevant ones are known to be satisfied. Unfortunately, the irrelevance of an inequality in above terms does not guarantee that its presence will have no impact on the construction of the l^1-optimal controller; there are some pathological cases where the constraint region denoted by an irrelevant inequality intersects the region denoted by the relevant ones, and if the solution of the linear programming problem happens to lie within this intersection, an extra degree of freedom in the construction of the l^1-optimal controller is introduced. As a consequence, an inequality might affect the l^1-optimal controller even when it is an irrelevant one; a topic which will be investigated thoroughly in Chapter 5. All the inequalities that affect the resulting l^1-optimal controller are called *active* inequalities, and it should be clear from the above discussion that although all the relevant inequalities are active ones, the converse may not be true. Still, since the solution for the dual linear programming problem and the l^1-optimal performance index can be obtained using relevant inequalities alone, these will be of main interest in the upcoming discussion.

Since the first n inequalities are known to be independent (see the proof of Theorem 4.2), they are all relevant within themselves and the above lemma gives us a way of determining whether or not each new inequality added to these ones is relevant. Of course there is the possibility that a new inequality might eliminate some of the previous ones and this also has to be checked. Since we need at least n linearly independent inequalities to define a compact constraint set in R^n, however, it would be quite distressing if it were possible for a new inequality to eliminate so many previous ones that less than n linearly independent inequalities remain relevant with the addition of this final one. Fortunately, the lemma below assures us that this is never the case.

Lemma 4.3 *Let $|b_i'x| \leq 1$ for $i = 1, \ldots, N$ where $b_i, x \in R^n$, $N \geq n$ and $\{b_i\}_{i=1}^n$ form a linearly independent set. Then the number of relevant linearly independent inequalities is n.*

Proof (by induction)

1. Since the first n inequalities are linearly independent we already know them to be relevant within themselves. Now let $|b_{n+1}'x| \leq 1$ be a new inequality added to the first n ones. If b_{n+1} is a sub-convex combination of b_i, $i = 1, \ldots, n$, then it is an irrelevant inequality itself and cannot possibly eliminate any of the previous ones; so we are done. Else there are two possibilities:

case 1 Assume the last inequality has eliminated a single one of the original ones and is itself a linear combination of the remaining inequalities; leaving us with $n-1$ linearly independent relevant inequalities only. We will show that this assumption leads to a contradiction.

Without loss of generality we may assume $|b_1'x| \leq 1$ is the inequality eliminated by the addition of the last one. Then, from Lemma 4.2, we should be able to express b_1 as a sub-convex combination of b_i, $i = 2, \ldots, n+1$. But since b_{n+1} is a linear combination of b_i, $i = 2, \ldots, n$ by hypothesis, this means we should be able to express b_1 as a linear combination of b_i, $i = 2, \ldots, n$ which contradicts the initial assumption that $\{b_i\}_{i=1}^n$ is set of linearly independent vectors.

case 2 Assume that the last inequality has eliminated more than one of the original inequalities. Without loss of generality, we may assume that $|b_1'x| \leq 1$ and $|b_2'x| \leq 1$ are among the eliminated ones. Since $\{b_i\}_{i=1}^n$ form a linearly independent set, we can represent b_{n+1} uniquely as a linear combination of these vectors as follows:

$$b_{n+1} = \beta_1 b_1 + \beta_2 b_2 + \cdots + \beta_n b_n$$

where

$$\sum_{i=1}^{n} |\beta_i| > 1$$

from Lemma 4.2. Since we assumed that the new inequality eliminates $|b_1'x| \leq 1$ and $|b_2'x| \leq 1$, again from the same lemma, we know that both b_1 and b_2 should be representable as a sub-convex combination of the remaining b_i; i.e., we should be able to find α_i such that

$$b_1 = \alpha_3 b_3 + \cdots + \alpha_{n+1} b_{n+1}$$

and

$$\sum_{i=3}^{n+1} |\alpha_i| \leq 1$$

But from previous arguments we already know that

$$b_1 = -\frac{\beta_2}{\beta_1} b_2 - \cdots - \frac{\beta_n}{\beta_1} b_n + \frac{1}{\beta_1} b_{n+1}$$

and this representation is unique. For it to be identical to the previous expression, obviously a necessary condition is that $\beta_2 = 0$; but this is obviously a contradiction since if $\beta_2 = 0$, then there is no way that b_2 can be represented as a linear combination of b_i, $i = 3, \ldots, n+1$ at the first place.

Hence we have proved by contradiction that any single inequality added to the first n linearly independent ones cannot eliminate the original ones in a manner such that the number of linearly independent inequalities will decline below n.

2. Assume that the hypothesis is true for $n + m$ inequalities; i.e., after the addition of m extra inequalities, there are still at least n independent ones among the remaining $n + q$ relevant inequalities. Now assume that the $n + m + 1$st inequality added to the set of original relevant inequalities eliminates some of them in such away that less then n linearly independent ones are left in the resulting set of relevant inequalities. However, from Lemma 4.2, we know that all the eliminated inequalities should be expressed as sub-convex linear combinations of the remaining ones; hence the set containing all the original inequalities and the last $n+m+1$st inequality cannot contain n linearly independent elements. Consequently, the set of the original $n+q$ inequalities, which is a smaller set, could not have had n linearly independent elements at the first place, which contradicts our hypothesis.

Hence, by induction we have proved that given first n inequalities linearly independent, we will have at least n linearly independent - hence relevant - inequalities among the resulting relevant ones.

Q.E.D.

This result, combined with the fact that each inequality constraint defines a region between two parallel hyperplanes symmetric about the origin, guarantees that the constraint set for the linear programming problem associated with the l^1-optimization problem is always compact and hence the linear programming problem always has a solution yielding a finite value. Obviously, this finding is consistent with the duality theorem which assures the existence of a solution to the dual problem, and the linear programming problem in our case is nothing but the dual problem associated with the original l^1-optimization problem.

Construction of the l^1-optimal Controller

In the previous section we have verified that the linear programming problem associated with the l^1-optimization problem always has a solution. Though a useful result on its own, this finding does not immediately tell if a solution to the primary problem exists and if it does, how it can be constructed. In this section we will first give a brief overview of the conditions Φ has to satisfy to give the l^1-optimal performance and prove that a controller yielding this result can always be constructed.

For convenience let us recall the primary and dual problems once more

$$\begin{aligned}\mu &= \inf_{R\in S} \|h-r\|_1 \\ &= \max_{\alpha_i} \sum_{i=1}^{m} \alpha_i H(a_i) + \sum_{i=m+1}^{m+l} \alpha_i Re(H(a_i)) + \sum_{i=m+l}^{n} \alpha_i Im(H(a_{i-l}))\end{aligned}$$

subject to the constraints

$$-1 \leq \sum_{i=1}^{m} \alpha_i a_i^k + \sum_{i=m+1}^{m+l} \alpha_i Re(a_i^k) + \sum_{i=m+l}^{n} \alpha_i Im(a_{i-l}^k) \leq 1$$

where $k = 0, 1, 2, \ldots$ and S denotes the set of stable transfer functions that satisfy $R(a_i) = 0$ for $i = 1, \ldots, n$.

Let $[\alpha_1^*, \ldots, \alpha_n^*]'$ be a solution for the linear programming problem. Define

$$w_0(k) = \sum_{i=1}^{m} \alpha_i^* a_i^k + \sum_{i=m+1}^{m+l} \alpha_i^* Re(a_i^k) + \sum_{i=m+l+1}^{n} \alpha_i^* Im(a_{i-l}^k) \tag{4.3}$$

Now assume the infimum on the left is actually achieved for some r_0; then defining

$$\phi_0 = h - r_0 \tag{4.4}$$

it can be shown (Theorem 4.1 and Lemma 4.1) that the necessary and sufficient conditions for the existence of an optimal solution are

1. $\phi_0(k)w_0(k) \geq 0$
2. $\phi_0(k) = 0$ whenever $|w_0(k)| \neq ||w||_\infty = 1$
3. $\sum_{k=0}^{\infty} |\phi_0(k)| = \mu$
4. $\Phi_0(a_i) = \sum_{k=0}^{\infty} \phi_0(k) a_i^k = H(a_i)$ for $i = 1, \ldots, n$

The first two conditions are alignment requirements; the third one assures ϕ achieves the infimum norm and the last condition arises from the hypothesis $\Phi(a_i) = H(a_i) - R(a_i) = H(a_i)$.

The condition that immediately catches our attention is the second one which restricts the indices for which $\phi_0(k)$ can be nonzero. Recalling our definition of w_0, it is evident that $\phi_0(k)$ can be nonzero only on the parallel constraint surfaces denoted by the equations

$$\sum_{i=1}^{m} \alpha_i^* a_i^k + \sum_{i=m+1}^{m+l} \alpha_i^* Re(a_i^k) + \sum_{i=m+l+1}^{n} \alpha_i^* Im(a_{i-l}^k) = \pm 1 \tag{4.5}$$

for those indices k corresponding to active constraints. Since the corner $[\alpha_1^*, \ldots, \alpha_n^*]'$ is specified by the intersection of $q \geq n$ such surfaces obtained from different values of k, we will have q possibly nonzero entries of ϕ_0 which should be adjusted to satisfy the remaining conditions. The next theorem proves that a solution ϕ_0 to this problem satisfying all the conditions always exists, which immediately implies that an l^1-optimal linear dynamic controller can always be constructed.

Theorem 4.3 *At the corner $[\alpha_1^*, \ldots, \alpha_n^*]'$ which maximizes the cost functional of the linear programming problem associated with the l^1-optimization problem, a solution ϕ_0 satisfying existence conditions for an optimal solution, (1-4) above, can always be constructed.*

Corollary 4.3.1 *For a discrete-time SISO system, an l^1-optimal linear dynamic controller always exists.*

Proof The proof of the theorem consists of three parts: first we will show that any solution that satisfies conditions 1,2 and 4 automatically satisfies condition 3 as well. Next, we will show that a solution satisfying conditions 2 and 4 can always be found. Finally, we will show that if a solution satisfies conditions 2 and 4 and is unique, then it automatically satisfies condition 1, too, and if there are multiple solutions for 2 and 4, we will show that we can still prove at least one of them satisfies condition 1 as well.

1. Let us assume that the solution to the linear programming is at $[\alpha_1^*, \ldots, \alpha_n^*]$ satisfying

$$\sum_{i=1}^{m} \alpha_i^* a_i^{k_j} + \sum_{i=m+1}^{m+l} \alpha_i^* Re(a_i^{k_j}) + \sum_{i=m+l+1}^{n} \alpha_i^* Im(a_{i-l}^{k_j}) = s_j$$

for $j = 1, \ldots, q$, where $s_j = \pm 1$. From previous work we know that $q \geq n$ and by hypothesis, if $\alpha_1^*, \ldots, \alpha_n^*$ satisfy the above system of equations, then the optimum performance index is given by

$$\mu = \sum_{i=1}^{m} \alpha_i^* H(a_i) + \sum_{i=m+1}^{m+l} \alpha_i^* Re(H(a_i)) + \sum_{i=m+l+1}^{n} \alpha_i^* Im(H(a_{i-l}))$$

Assuming the existence condition (1) holds, the conditions (2-4) can be summarized in the matrix notation as follows

$$\underbrace{\begin{bmatrix} H(a_1) \\ \vdots \\ H(a_m) \\ Re(H(a_{m+1})) \\ \vdots \\ Re(H(a_{m+l})) \\ Im(H(a_{m+1})) \\ \vdots \\ Im(H(a_{m+l})) \\ \mu \end{bmatrix}}_{H_\mu} = \underbrace{\begin{bmatrix} a_1^{k_1} & \cdots & a_1^{k_q} \\ \vdots & \cdots & \vdots \\ a_m^{k_1} & \cdots & a_m^{k_q} \\ Re(a_{m+1}^{k_1}) & \cdots & Re(a_{m+1}^{k_q}) \\ \vdots & \cdots & \vdots \\ Re(a_{m+l}^{k_1}) & \cdots & Re(a_{m+l}^{k_q}) \\ Im(a_{m+1}^{k_1}) & \cdots & Im(a_{m+1}^{k_q}) \\ \vdots & \cdots & \vdots \\ Im(a_{m+l}^{k_1}) & \cdots & Im(a_{m+l}^{k_q}) \\ s_1 & \cdots & s_q \end{bmatrix}}_{A_\mu} \underbrace{\begin{bmatrix} \phi(k_1) \\ \vdots \\ \phi(k_q) \end{bmatrix}}_{P}$$

Now, multiplying the ith row of A_μ by α_i^* for $i = 1, \ldots, n$ and subtracting from the last row, it is clear from the previous work that we obtain a row of zeros in the last row. Moreover, performing the same row operations on H_μ would yield a last element

$$\mu - \left(\sum_{i=1}^{m} \alpha_i^* H(a_i) + \sum_{i=m+1}^{m+l} \alpha_i^* Re(H(a_i)) + \sum_{i=m+l+1}^{n} \alpha_i^* Im(H(a_{i-l})) \right)$$

which, by hypothesis, is also zero. Hence, we have shown that the existence condition (3) is irrelevant and will be automatically satisfied if the remaining conditions summarized in matrix notation as

$$\underbrace{\begin{bmatrix} H(a_1) \\ \vdots \\ H(a_m) \\ Re(H(a_{m+1})) \\ \vdots \\ Re(H(a_{m+l})) \\ Im(H(a_{m+1})) \\ \vdots \\ Im(H(a_{m+l})) \end{bmatrix}}_{H} = \underbrace{\begin{bmatrix} a_1^{k_1} & \cdots & a_1^{k_q} \\ \vdots & \cdots & \vdots \\ a_m^{k_1} & \cdots & a_m^{k_q} \\ Re(a_{m+1}^{k_1}) & \cdots & Re(a_{m+1}^{k_q}) \\ \vdots & \cdots & \vdots \\ Re(a_{m+l}^{k_1}) & \cdots & Re(a_{m+l}^{k_q}) \\ Im(a_{m+1}^{k_1}) & \cdots & Im(a_{m+1}^{k_q}) \\ \vdots & \cdots & \vdots \\ Im(a_{m+l}^{k_1}) & \cdots & Im(a_{m+l}^{k_q}) \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \phi(k_1) \\ \vdots \\ \phi(k_q) \end{bmatrix}}_{P}$$

are satisfied, assuming that the sign restrictions of condition (1); i.e., $\phi(k_i)s_i \geq 0$ are also satisfied.

2. Now we must prove that the above set of linear equations is always consistent. Our reasoning is as follows: any corner in R^n can only be specified by a minimum of n linearly independent linear equations. Since $[\alpha_1^*, \ldots, \alpha_n^*]'$ is a corner specified by

$$A'\alpha^* = s$$

where

$$\alpha^* = \begin{bmatrix} \alpha_1^* \\ \vdots \\ \alpha_n^* \end{bmatrix} \text{ and } s = \begin{bmatrix} s_1 \\ \vdots \\ s_q \end{bmatrix}$$

and $q \geq n$, the $q \times n$ matrix A' should have full column-rank n. Consequently, the matrix A has full row-rank. However, it is a well-known algebraic fact that a system of equations denoted by

$$AP = H$$

where A is an $n \times q$ matrix and P and H are q- and n-vectors respectively always has solution if A has full row-rank. Hence the equations that must be satisfied by the entries of P are consistent and can always be solved to obtain a solution $\phi_0(k)$. The only adverse argument at this point would be one about the existence of at least n relevant equations; but we already know this to be the case from Lemma 4.3

3. All that remains to be checked are the sign constraints of the existence condition (1); i.e., we need to prove that one of the solutions to the matrix equation

$$AP = H$$

satisfies

$$P_i s_i \geq 0$$

where P_i and s_i correspond to the i th entries of these vectors respectively, given

$$(\alpha^*)' H = \mu$$

$$A' \alpha^* = s$$

and $\alpha' H$ reaches its maximum at α^*.

From previous part, we know that the matrix A has full row rank and hence the matrix equation is always consistent. Now there are two possibilities:

case 1 A is a square matrix, and since it has full row rank it will be nonsingular. Then the equation

$$AP = H$$

has the unique solution

$$P = A^{-1} H$$

Now, all we need to consider are the signs of the entries of P, i.e., $P_i = \phi_0(k_i)$. To show that the entries satisfy these restrictions consider the following equations.

Let A^j denote the jth column of A and A_i^{-1} denote the ith row of A^{-1}; with this convention we have

$$P_i s_i = s_i A_i^{-1} H$$

Now define

$$\gamma_i' = s_i A_i^{-1}$$

then obviously

$$\gamma_i' A^j = \begin{cases} 0 & \text{if } i \neq j \\ s_i & \text{if } i = j \end{cases}$$

Now assume γ_i originates from a point on the surface denoted by

$$\alpha' A^i = s_i$$

It can be shown in a straightforward manner that γ_i is directed outward from the constraint set by incrementing a vector on this surface in the direction of γ_i by a small amount and observing that the new vector lies outside the constraint set. Now let us assume that γ_i originates from the corner α^*, which is possible since the surface from which γ_i originates is one of the surfaces intersecting at this point. Since γ_i points outward from the constraint set and it is parallel to all the constraint surfaces except for the one denoted by $a_i'\alpha = s_i$, $-\gamma_i$ points inward and actually lies along the edge formed by the intersection of the remaining surfaces. Since $\alpha' H$ reaches its maximum at this point, the inner product of H with any vector originating from $(\alpha^*)'$ and pointing inward towards the constraint set is non-positive; in particular

$$\begin{aligned} -\gamma_i' H \leq 0 \quad &\Rightarrow \quad \gamma_i' H \geq 0 \\ &\Rightarrow \quad P_i s_i \geq 0 \end{aligned}$$

as we wanted to show.

case 2 Assume $q > n$; i.e., A is a rectangular matrix with full row rank. Still, we could find n surfaces intersecting at the point α^* such that $\alpha' H$ would be found to achieve its maximum at this point had these n surfaces been the only surfaces intersecting at that point. Now setting all the entries of ϕ except the ones corresponding to these surfaces to zero and following the same arguments as in case 1 with the remaining possibly nonzero entries, we can always construct the appropriate solution ϕ that satisfies the sign constraints as well as the equation

$$AP = H$$

Q.E.D.

So we have proved that a solution ϕ to the above problem always exists. The proof for the corollary is now by simply recalling

$$\phi_0 = h - r_0 \Rightarrow r_0 = h - \phi_0 \tag{4.6}$$

from which the construction of the l^1-optimal controller is straightforward.

DEGENERACY AND UNIQUENESS OF OPTIMAL SOLUTIONS

An interesting question is what happens in the degenerate case when the linear programming problem associated with the l^1-optimization problem has multiple solutions and how, if in any way, this extra feature can be exploited to achieve extra performance goals. In this section we will have a closer look at this subject and prove two important theorems regarding the questions posed above; the former telling us that even in the presence of multiple solutions to the linear programming problem, any secondary optimization problem imposed on the l^1-optimal controllers - presuming that there are a multitude of them - still has its solutions at one of the corners of the constraint set; and the latter showing that in the case of multiple solutions to the linear programming problem, the l^1-optimization problem has a simple solution. Then we will try to establish the relationship between the degeneracies - i.e., existence of multiple solutions - of the primary and dual problems and prove that the primary problem almost always has a unique solution and the existence of multiple solutions for one problem does not affect the other one. But let us first investigate when it is possible for the linear programming problem to have multiple solutions.

Multiple Solutions for the Linear Programming Problem

If α^* denotes a corner which is the intersection of $q \geq n$ constraint hyperplanes of the form

$$\begin{aligned} & a_1^{k_j}\alpha_1^* + \cdots + a_m^{k_j}\alpha_m^* + \\ & Re(a_{m+1}^{k_j})\alpha_{m+1}^* + \cdots + Re(a_{m+l}^{k_j})\alpha_{m+l}^* + \\ & Im(a_{m+1}^{k_j})\alpha_{m+l+1}^* + \cdots + Im(a_{m+l}^{k_j})\alpha_n^* \quad = \quad s_j \end{aligned}$$

for $j = 1, \ldots, q$ and is a solution for the linear programming problem associated with the l^1-optimization problem, a necessary and sufficient condition for the existence of multiple solutions to the linear programming problem is H being a linear combination of $p < n$ the normals to these hyperplanes; i.e., recalling

$$b_j = [a_1^{k_j} \ldots a_m^{k_j}\ Re(a_{m+1}^{k_j}) \ldots Re(a_{m+l}^{k_j})\ Im(a_{m+1}^{k_j}) \ldots Im(a_{m+l}^{k_j})]' \tag{5.1}$$

we should have

$$H = c_1 b_1 + \cdots + c_p b_p \tag{5.2}$$

Then

$$H'\alpha^* = c_1 b_1' \alpha^* + \cdots + c_p \alpha^* = \sum_{i=1}^{p} c_i s_i$$

from above. Furthermore, since the cost functional $\alpha' H$ achieves its maximum at α^*, the c_i have the same signs as s_i, yielding the l^1-optimal performance index

$$\mu = \sum_{j=1}^{p} |c_j| \tag{5.3}$$

In the next section, we will show that in the presence of such multiple solutions, a simple solution for the l^1-optimal closed loop transfer function which can be constructed in a straightforward manner exists.

The Simple Solution and Extra Performance Constraints

Theorem 5.1 *In the presence of multiple solutions to the linear programming problem, there is a simple solution to the l^1-optimization problem.*

Proof Consider the following closed-loop impulse response

$$\phi(k) = \begin{cases} c_j & k = k_j \\ 0 & \text{otherwise} \end{cases}$$

From previous arguments, we also know that

$$w_0(k_j) = \sum_{i=1}^{m} \alpha_i^* a_i^{k_j} + \sum_{i=m+1}^{m+l} \alpha_i^* Re(a_i^{k_j}) + \sum_{i=m+l+1}^{n} \alpha_i^* Im(a_{i-l}^{k_j}) = s_j = sgn(c_j)$$

Then we have the following results:

1.

$$\phi(k) w_0(k) = \begin{cases} c_j.sgn(c_j) = |c_j| & k = k_j \\ 0 & \text{otherwise} \end{cases}$$

Hence first alignment condition is satisfied.

2. $|w_0(k_j)| = |sgn(c_j)| = 1 \Rightarrow \phi(k_j) = c \neq 0$ is in consistency with the second alignment condition.

3. $\sum_{k=0}^{\infty} |\phi(k)| = \sum_{j=1}^{p} |c_j| = \mu$

4. $\sum_{k=0}^{\infty} \phi(k) a_i^k = \sum_{j=1}^{p} c_j a_i^{k_j} = H(a_i)$ for all $i = 1, \ldots, n$.

Hence the simple $\{\phi(k)\}_{k=0}^{\infty}$ we have constructed above satisfies all the necessary and sufficient conditions for being the l^1-optimal closed-loop transfer function. Simply setting

$$r = h - \phi$$

and recalling $R = FQ$, where Q parameterizes the controller we are seeking, it is obvious that a minimizing controller always exists under the presence of non-unique solutions to the associated linear programming problem.
Q.E.D.

At this point, it is quite tempting to assume that multiple solutions to the linear programming problem would probably lead to multiple solutions to the original l^1-optimization problem each of which would correspond to a different set of parameters that solve the linear programming problem. As a direct consequence, one might think of imposing further performance objectives on these l^1-optimal controllers which can be transformed to an additional optimization problem on the set of parameters that solve the linear programming problem. However, a closer look at the problem shows us that this is not the case and in the presence of multiple solutions to the linear programming problem, the conditions for which were summarized in the previous section, a secondary optimization problem imposed on the l^1-optimal controllers - presuming there is a multitude of them - always leads to a solution corresponding to a corner of the constraint set for the linear programming problem.

Theorem 5.2 *Assume there are multiple solutions to the linear programming problem associated with an l^1-optimization problem. Then the linear programming problem parameters (α_i) corresponding to an l^1-optimal controller optimizing any secondary performance criterion imposed on the closed-loop transfer function Φ can be found at a corner of the solution set for the linear programming problem.*

Proof The proof is merely argumentative: recalling the alignment conditions and the fact that $\phi(k)$ can be nonzero only if the parameters - α's - of the linear programming problem lie on the constraint surface denoted by

$$\sum_{i=1}^{m} \alpha_i a_i^k + \sum_{i=m+1}^{m+l} \alpha_i Re(a_i^k) + \sum_{i=m+l}^{n} \alpha_i Im(a_{i-l}^k) = \pm 1$$

Now assume that there are multiple solutions to the linear programming problem; i.e., the α parameters that solve the problem need not construct a single point but rather a compact subset of some affine space. Still, a vector $[\alpha_1, \ldots, \alpha_n]'$ at any point in this set other than a corner will not satisfy at least one of the equations of the above type that a vector picked on one of the corners satisfies - as well as the other equations satisfied by the former vector. Since both vectors would lead to identical sign constraints on the entries of ϕ, the result is an effective loss of degree of freedom in choosing the entries of ϕ, which immediately leads us to the conclusion that any secondary optimization problem imposed on the l^1-optimal controllers, when transformed into the domain of the parameters of the linear programming problem, should have its solution at one of the corners of the set throughout which the l^1-optimal performance is obtained.

Q.E.D.

Uniqueness of l^1-optimal Controllers

So far we have shown that the linear dynamic l^1-optimal controller design problem always has a solution and we have proved that in the degenerate case where the linear programming problem has multiple solutions a simple solution for the original problem exists. We have also arrived at the interesting conclusion that in the degenerate case a secondary optimization problem imposed on the l^1-optimal controllers, presuming that there is a multitude of them to select from, would have its solution at one of the corners of the constraint set in the space of the parameters of the linear programming problem associated with the l^1-optimization problem regardless of the nature of this secondary problem. The missing link between these results is the answer to the question of whether or not multiple solutions to the l^1-optimization problem exist and what, if any, is the connection between such cases and the degeneracy of the linear programming problem. This section addresses these problems and the theorem we will prove shortly tells us that the l^1-optimal controller is almost always unique, even under the presence of multiple solutions to the associated linear programming problem, where the term *almost always* is used analogously to the term *almost everywhere* in real analysis and refers to a Lebesgue-type continuous probability measure defined on R^n.

Theorem 5.3 *The l^1-optimal controller is almost always unique.*

Corollary 5.3.1 *In the degenerate case where the linear programming problem has multiple solutions, the simple solution to the l^1-optimization problem is almost always the only solution.*

Proof Let α^* be a point that maximizes the cost functional $\alpha' H$ of the dual linear programming problem. There are two possibilities:

case 1 If the dual linear programming problem is non-degenerate, then α^* denotes a corner of the constraint set. From Theorem 4.3, we know that if the number of constraint surfaces intersecting at α^* is n, then there exists a unique solution obtained from

$$P = A^{-1} H$$

where A, P and H are as in Theorem 4.3. Hence a necessary (but not sufficient, as will be exemplified in Chapter 5) condition for the existence of multiple solutions for the l^1-optimal controller is that $q > n$ constraint surfaces intersect at α^*. But this condition is equivalent to an $n \times n$ real matrix being singular, which has a probability of zero under a Lebesgue-type continuous probability measure defined on $R^{n+1 \times n+1}$. Hence the l^1-optimal controller is almost always unique when the dual linear programming problem is non-degenerate.

case 2 If the dual linear programming problem is degenerate, its solution lies on a set rather than being a single point. From the discussion in case 1, however, the corners of this set are almost always obtained from the intersection of n hyperplanes, each admitting a single solution for the l^1-optimal controller. From the proof of Theorem 5.2, we also know that the corners offer more degrees of freedom in constructing the l^1-optimal controller compared to the non-corner points. Consequently, since each corner almost always admits a single optimal controller, there cannot exist multiple solutions for the optimal controller corresponding to non-corner points. Thus, every point (corner/non-corner) in the solution set for the linear programming problem almost always admits at most a single optimal controller. However, from the proof of Theorem 5.1, it is also clear that the simple solution for the l^1-optimal controller can be constructed using any point in the solution set for the linear programming problem. Hence, the simple solution should almost always be the unique solution for the l^1-optimal controller when the dual linear programming problem is degenerate.

Q.E.D.

A careful interpretation of the arguments in the proof above reveals the interesting fact that the degeneracy of the l^1-optimization problem is completely independent from that of the linear programming problem associated with it, where the term *degeneracy* refers to the existence of multiple solutions for the respective problem. Indeed, as will be exemplified in the following section, there are cases where the linear programming problem has multiple solutions whereas the l^1-optimization problem has a unique solution, and vice-versa. This result, however, does not invalidate the arguments of Theorem 5.2; i.e., if both the primary and the dual problems do have multiple solutions - which is possible, though, as explained above, one does not imply the other - any secondary optimization problem imposed on the l^1-optimal controllers, when interpreted in the space of the parameters of the linear programming problem, has its solution(s) at one of the corners of the set throughout which the l^1-optimal performance is achieved.

An Example: H_2-optimization Imposed on l^1-optimal Controllers

Now let us assume that there are multiple solutions for the l^1-optimal linear dynamic controller and a secondary optimization problem has been imposed on the resulting closed-loop transfer functions in terms of selecting the l^1-optimal controller that leads to the minimum H_2-norm for this transfer function. We have already proved in Theorem 5.2 that even in the degenerate case where the associated linear programming problem has multiple solutions, any secondary optimization problem imposed on the l^1-optimal controllers will have a solution corresponding to a corner of the solution set for the linear programming problem. As a consequence, the H_2-optimization problem can always be expressed as minimize

$$||\phi||_2 = \left(\sum_{k=0}^{\infty} |\phi(k)|^2\right)^{1/2} = P'P$$

subject to the constraints

$$AP = H \text{ and } P_i s_i \geq 0$$

where A, P and H are as defined in the previous sections; hence A is an n by q matrix with full row-rank. Moreover, we know that the constraint equations and sign restrictions are always consistent; i.e., a solution P satisfying all constraints can always be constructed.

The problem as formulated above can easily be solved by applying any of the quadratic programming algorithms available in the literature. An example of such algorithms is Wolfe's algorithm (see [12]) which solves problems of the type

$$\text{minimize } \frac{1}{2}x'Qx + x'p$$

$$\text{subject to } Fx = b \text{ and } x \geq 0$$

Let us define the diagonal matrix S as

$$S = \text{diag}\{s_i \ , i = 1, \ldots, q \}$$

Obviously $S^{-1} = S$. Now defining $\tilde{P} = SP$ and $\tilde{A} = AS$, our problem becomes identical to the problems solved by using Wolfe's method under the following setting:

$$F \leftrightarrow \tilde{A}$$

$$x \leftrightarrow \tilde{P}$$

$$b \leftrightarrow H$$

$$Q \leftrightarrow I$$

$$p \leftrightarrow 0$$

One thing that should be noted at this point is that in cases where both the primary and dual problems have multiple solutions, we have to solve the above problem at each corner of the set throughout which the l^1-optimal performance is achieved, using the different A matrices corresponding to these corners. Though a bit more tedious, the problem will still be solvable since there will always be a finite number of such corners. The main contribution of Theorem 5.2 becomes evident at this point: without this result we would have to try infinitely many different settings for various choices of α vectors all of which yield the l^1-optimal performance and possibly could not solve the problem using any such computational algorithm. The situation is exemplified in the next section.

EXAMPLES

To be able to visualize the ideas developed so far, consider the following simple 2-dimensional example: let

$$a_1 = 1/2$$

$$a_2 = 1/3$$

be the two zeros; then it can easily be shown that the first two inequalities

$$-1 \leq \alpha_1 + \alpha_2 \leq 1$$

$$-1 \leq 1/2\alpha_1 + 1/3\alpha_2 \leq 1$$

are the only relevant ones, whereas the third inequality

$$-1 \leq 1/4\alpha_1 + 1/9\alpha_2 \leq 1$$

is an active one though it is not relevant, since the region depicted by the first two inequalities completely lies within the one denoted by the third one but there are intersection points nevertheless. The situation is illustrated in the diagram on the next page. Now let us try to solve this problem for various values of $H(a_i)$'s:

1. Let

$$H(a_1) = 1/4$$

$$H(a_2) = 1/9$$

Then from the figure on the next page it is obvious that the solution is at $\alpha_1 = 8$, $\alpha_2 = -9$. Note that even though the vector $H = [1/4\ 1/9]'$ is parallel to the active - but irrelevant - inequality border, there is still a unique solution to the linear programming problem which yields the following sign restrictions on the entries of the closed-loop impulse response:

$$\phi(0) \leq 0$$

$$\phi(1) \geq 0$$

$$\phi(2) \geq 0$$

and the remaining entries of ϕ have to be zero as a result of the alignment condition (Lemma 4.1).

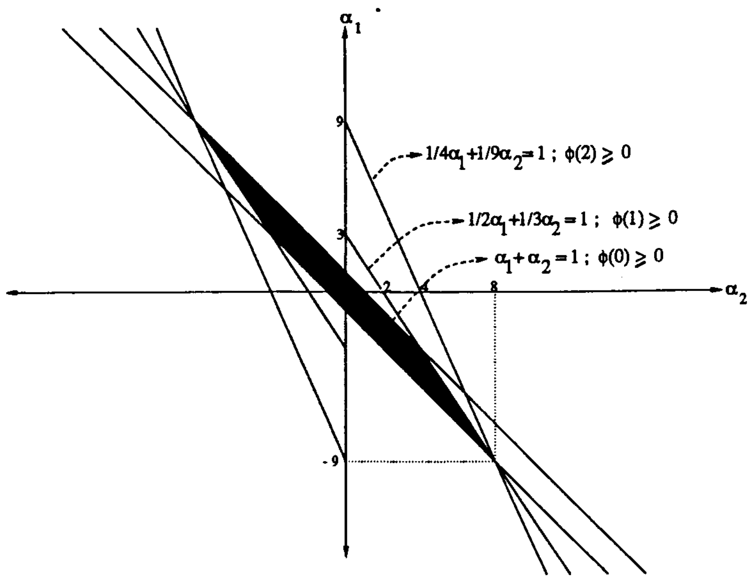

Figure 5.1: Constraint region denoted by the first two inequalities

At this point, it should be obvious why the third inequality is an active one, although it is irrelevant in the sense it does not effect the solution of the linear programming problem: since its border intersects the solution point, it enables $\phi(2)$ to be nonzero, adding extra flexibility to the closed-loop transfer function - hence controller - selection. Indeed, the equalities that have to be satisfied by the possibly nonzero entries depicted above are in the form

$$\begin{bmatrix} 1 & 1/2 & 1/4 \\ 1 & 1/3 & 1/9 \end{bmatrix} \begin{bmatrix} \phi(0) \\ \phi(1) \\ \phi(2) \end{bmatrix} = \begin{bmatrix} 1/4 \\ 1/9 \end{bmatrix}$$

whose solution, with the previous sign constraints, can be found to be as

$$\phi(0) = \rho \quad \phi(1) = -5\rho \quad \phi(2) = 1 + 6\rho \quad -1/6 \leq \rho \leq 0$$

Hence, there are multiple solutions to the l^1-optimization problem even though the associated linear programming problem has a unique solution. Now we can impose H_2-optimization as a secondary optimization problem and try to pick the l^1-optimal controller that yields the minimum H_2-norm for the closed-loop transfer function. This can be done by using the Wolfe's algorithm for quadratic programming which was mentioned before:

but for this simple case we do not need such elaborate schemes but a little bit of algebra and calculus to obtain

$$||\phi||_2^2 = 62\rho^2 + 12\rho + 1$$

$$\Rightarrow \frac{d}{d\rho}(||\phi||_2^2) = 124\rho + 12 = 0 \rightarrow \rho = -3/31 \in [-1/6, 0]$$

and thus the minimum H_2-norm is achieved by choosing the closed-loop transfer function entries as

$$\phi(0) = -3/31$$

$$\phi(1) = 15/31$$

$$\phi(2) = 13/31$$

from which the l^1-optimal controller can be obtained in a straightforward manner from (11), yielding an "optimal" H_2-norm of $\sqrt{13/31}$.

2. In the above example we have shown that multiple solutions for the l^1-optimal controller might exist even when the associated linear programming problem has a unique solution. To show that the converse might also be true, consider the same example with a different cost functional; namely let

$$H(a_1) = 1/2$$

$$H(a_2) = 1/3$$

then the $H = [1/2\ 1/3]'$ vector is parallel to the border of the second inequality constraint and as we would expect, the linear programming problem has multiple solutions in the form

$$\alpha_1 = \gamma \quad \alpha_2 = 3 - 3/2\gamma \quad 4 \leq \gamma \leq 8$$

Now, if α_i are chosen somewhere other than the two corners $(4, -3)$ and $(8, -9)$, the only possibly nonzero entry of ϕ is $\phi(1)$ which has to satisfy

$$\begin{bmatrix} 1/2 \\ 1/3 \end{bmatrix} \phi(1) = \begin{bmatrix} 1/2 \\ 1/3 \end{bmatrix} \quad \phi(1) \geq 0$$

whereas at the corner $(4, -3)$ we have both $\phi(0)$ and $\phi(1)$ possibly nonzero with

$$\begin{bmatrix} 1 & 1/2 \\ 1 & 1/3 \end{bmatrix} \begin{bmatrix} \phi(0) \\ \phi(1) \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/3 \end{bmatrix} \quad \phi(0) \geq 0 \quad \phi(1) \geq 0$$

and at the corner $(8,-9)$ we have even more flexibility since in addition $\phi(2)$ can also be nonzero with

$$\begin{bmatrix} 1 & 1/2 & 1/4 \\ 1 & 1/3 & 1/9 \end{bmatrix} \begin{bmatrix} \phi(0) \\ \phi(1) \\ \phi(2) \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/3 \end{bmatrix} \quad \phi(0) \leq 0 \quad \phi(1) \geq 0 \quad \phi(2) \geq 0$$

The extra flexibility in the choice of the entries of the closed-loop impulse response when the α_i's are selected at a corner rather than another point in the set throughout which the l^1-optimal norm is achieved should be evident from the above discussion. Moreover, the simple solution

$$\phi(k) = \begin{cases} 1 & \text{if } k = 1 \\ 0 & \text{otherwise} \end{cases}$$

is obviously a solution as we had predicted. It can easily be proved that this is also the only solution, which shows that it is possible to have a unique solution for the l^1-optimization problem even when its associated linear programming problem has multiple solutions. This example also gives us an inkling of how unlikely it is to find multiple solutions for the l^1-optimization problem: even though three constraint surfaces intersect at a single point - which is improbable by itself - and this point is within the set throughout which the l^1-optimal performance is attained, still we cannot find multiple solutions for the l^1-optimal controller.

To avoid complexity and to be able to visualize the geometric interpretations of the arguments, the examples above have been chosen in R^2. Nevertheless, most of the ideas developed throughout this study including the existence and almost uniqueness of l^1-optimal controllers and the relationship between the degeneracies of the primary and dual problems have been clearly illustrated in this simple setting. Our findings are summarized and options for further research are discussed in the following conclusion section.

CONCLUSION

In this study we have first obtained an alternative formulation of the l^1-optimal linear dynamic controller design for SISO discrete time systems and have showed this approach leads to the same linear programming problem as in [3]. The main point of this effort was to show that some basic system level arguments could be used quite effectively to formulate the problem with considerably more understanding of the underlying principles of the l^1-optimal controllers. Then we have proved that the l^1-optimal controller always exists and have proceeded to investigate the case where the linear programming problem associated with the primary problem has multiple solutions. Our observations led to the interesting conclusion that even in this case, any secondary optimization problem imposed on the l^1-optimal controllers - under the presumption that a multitude of them existed - would have its solution(s) at one of the corners of the parameter set throughout which the l^1-optimal performance is achieved. We have also shown by construction that a simple solution to the primary problem in this case exists. Then we have proved that the l^1-optimal controllers are almost always unique regardless of the nature (degenerate/non-degenerate) of the associated linear programming problem, and that the existence of multiple solutions for one of the problems neither necessitates nor prohibits the existence multiple solutions for the other one. Under the light of these findings, we have imposed the minimization of the H_2-norm of the closed-loop transfer function as a secondary performance goal in the case of multiple l^1-optimal controllers and discussed solution techniques. Finally, all these arguments have been demonstrated by appropriate examples to clarify the various aspects of the l^1-optimal linear dynamic feedback controllers.

The existence results and the arguments about a secondary optimization problem imposed on l^1-optimal controllers can be expanded to MIMO case as well under proper rank conditions, but the possibility of having redundant inputs in such a case disrupts uniqueness results. Even though it is highly unlikely (with probability zero) to have multiple solutions for the l^1-optimal controllers, the study does provide useful guidelines in the off-chance that such an event occurs.

The results obtained in this study can easily be generalized to cover the cases where F

(3.5) has zeros with higher multiplicities within the unit circle or zeros with single multiplicity on the unit circle. Although our approach seems to be making an extreme use of the discrete nature of the systems under consideration, similar arguments could also prove quite useful while dealing with continuous time systems. However, an analogous approach to continuous time systems would result in a constraint region specified by a continuum of inequalities rather than the countable inequalities in the discrete time case, leading to a nonlinear programming problem for exact solution. Nevertheless, this region would still be a convex balanced compact one and thus could be approximated as closely as desired by finitely many linear inequalities. Consequently, the resulting approximately dual linear programming problem could be solved using the procedures discussed in this study to obtain suboptimal controllers.

While referring to the uniqueness of l^1-optimal controllers, the term *almost always* has been used in analogous fashion to the term *almost everywhere* in real analysis. Indeed, the likelihood of existence of multiple solutions for the l^1-optimal controller is identical to that of an $n+1$ by $n+1$ square matrix being singular, which is an extremely rare case in practice since even the slightest change in one of the entries might disrupt the singularity. Indeed, such an event can be shown to have probability zero under an appropriately defined Lebesgue-type continuous probability measure on $R^{n \times n}$.

Although it is a nice theoretical result, the uniqueness of l^1-optimal controllers leaves much to be desired for designers. Unfortunately, stability and l^1-optimality are not likely to be the only requirements in real life: order and complexity restrictions, robustness, reliability are just a few of the issues that immediately come to mind as further expectations. Consequently, it is evident that rather than forcing l^1-optimality, further research in this area should be more focused on keeping the l^1-norm performance index below a prescribed value while seeking to realize other performance goals. Hence, techniques to construct suboptimal rather than optimal controllers and understanding the properties of such controllers become more crucial from a practical point of view. Nevertheless, it is my sincere hope that the results and the discussions in this thesis might prove useful and serve as a first stepping stone for those who might be willing to do further research in this area.

BIBLIOGRAPHY

1. J.C. Doyle, K. Glover, P.P. Khargonekar, B.A. Francis; *State-Space Solutions to Standard H_2 and H_∞ Control Problems*; IEEE Trans. on Auto. Control; Vol. 34; No. 8; Aug. 1987

2. M. Vidyasagar; *Optimal Rejection of Persistent Bounded Disturbances*; IEEE Trans. on Auto. Control; Vol. AC-31; No. 6; June 1986

3. M.A. Dahleh and J.B. Pearson; *l^1 Optimal Controllers for Discrete Time Systems*; Proc. Amer. Contr. Conf.; Seattle, WA; June 1986

4. M.A. Dahleh and J.B. Pearson; *l^1 Optimal Feedback Controllers for MIMO Discrete-Time Systems*; IEEE Trans. on Auto. Control; Vol. AC-32; No. 4; April 1987

5. M.A. Dahleh and J.B. Pearson; *L^1 Optimal Compensators for Continuous-Time Systems*; IEEE Trans. on Auto. Control; Vol. AC-32; No. 10; Oct. 1987

6. M.A. Dahleh and J.B. Pearson; *Optimal Rejection of Persistent Disturbances, Robust Stability and Mixed Sensitivity Minimization*; IEEE Trans. on Auto. Control; Vol. 33; Aug. 1988

7. D.G. Meyer; *Two Properties of l^1 Optimal Controllers*; IEEE Trans. on Auto. Control; Vol. 33; No. 9; Sept. 1988

8. G. Deodhare and M. Vidyasagar; *l^1-optimality of Feedback Control Systems: The SISO Discrete-time case*; IEEE Trans. on Auto. Control; Vol. 35; Sep. 1990

9. M. Dahleh and M.A. Dahleh; *Optimal Rejection of Persistent Bounded Disturbances: Continuity Properties and Adaptation*; IEEE Trans. on Auto. Control; Vol. 35; June 1990

10. M. Vidyasagar; *Further Results on the Optimal Rejection of Persistent Bounded Disturbances*; IEEE Trans. on Auto. Control; Vol. 36; No. 6; June 1991

11. J.S. Shamma and M. Dahleh; *Time-varying versus Time-invariant Compensation for Rejection of Persistent Bounded Disturbances and Robust Stabilization*; IEEE Trans. on Auto. Control; Vol. 36; July 1991

12. M.A. Dahleh and J.S. Shamma; *Rejection of Persistent Bounded Disturbances: Nonlinear Controllers*; Systems & Control Letters 18; 1992

13. J.S. Shamma; *Nonlinear State Feedback for l^1 Optimal Control*; Systems & Control Letters 21; 1993

14. I.J. Diaz-Bobillo and M.A. Dahleh; *Minimization of the Maximum Peak-to-Peak Gain: The General Multiblock Problem*; IEEE Trans. on Auto. Control; Vol. 38; Oct. 1993

15. F. Blanchini and M. Sznaier; *Rational L^1 Suboptimal Compensators for Continuous-Time Systems*; IEEE Trans. on Auto. Control; Vol. 39; No. 7; July 1994

16. J.C. Doyle, B.A. Francis. A.R. Tannenbaum; *Feedback Control Theory*; MacMillan Publishing Co.; 1992

17. M.A. Dahleh and I.J. Diaz-Bobillo; *Control of Uncertain Systems: A Linear Programming Approach*; Prentice Hall; 1995

18. D.G. Luenberger; *Optimization by Vector Space Methods*; J. Wiley & Sons; 1969

19. Joel Franklin; *Methods of Mathematical Economics*; Springer-Verlag; 1980