$\mathcal{H}^\infty$ Optimization With Uncertainty in the Plant *

J. William Helton
Dept of Mathematics
U. Calif. San Diego
La Jolla, CA 92093

Orlando Merino
Dept of Mathematics
U. of Rhode Island
Kingston, RI 02881

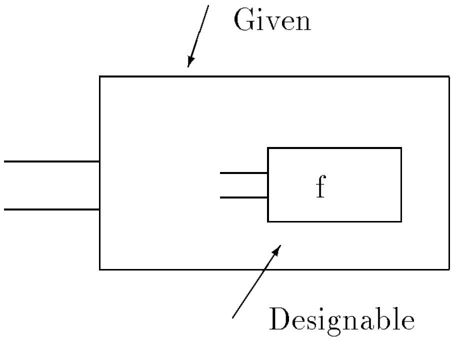
Trent E. Walker
Dept of Mathematics
U. Calif. Berkeley
Berkeley, CA

Abstract

The main result presented here gives the first order (neccesary) conditions for solutions to H^∞ optimization problems where there is an uncertainty parameter. These conditions can be stated as a set of equations which when solved produce excellent candidates for solutions, thus our result should lead to concrete algorithms in the spirit of the highly successful ones found by the authors for the certain plant case. How one crafts such algorithms for this uncertain case is indicated, but not analysed or tested.

1 Introduction

Suppose our objective is to design a linear, time invariant, system in the frequency domain description, with stability a major consideration. Typically part of the system is given and we are forced to use (in control it is called the plant), and another part f we get to choose (or designable part) (see Fig. 1). The stipulation that the designable part of the circuit be stable amounts to requiring that f has no poles in the right half plane (R.H.P.)



Suppose also that to carry out the design (i.e., choose f) we use a given performance function Γ. In many cases the performance takes the form of a positive function of frequency ω and of the designable part f at that frequency.

Thus we write $\Gamma(\omega, f(i\omega))$ for the performance at frequency ω and for designable function f. The performance is a cost function: big is bad. The "worst case" is the frequency ω at which $\sup_{\omega \in R} \Gamma(\omega, f(i\omega))$ occurs. We minimize this over the set A of R.H.P. stable functions f. Hence we have the problem

OPT Given a performance $\Gamma(\omega, z)$, find

$$\inf_{f \in A} \sup_{\omega \in R} \Gamma(\omega, f(i\omega))$$

and find a minimizer f^* if it exists.

A minimizer f^* is also called solution. A local solution is a solution over some open subset of A.

In physical systems there is always uncertainty. A way to treat the uncertainty in the context of performance functions is to write the performance as a function of an uncertainty parameter α. We denote a performance function that is also a function of an uncertainty parameter α by

$$\tilde{\Gamma}(\ \omega\ ,\ \alpha\ ,\ f(\ i\omega\ )\ )$$

The range $\mathcal{P}_\omega$ of the uncertainty parameter α may change with the frequency ω and is often known in advance. We will assume that $\mathcal{P}_\omega \subset \mathbf{R}^k$

If we fix the designable function f and the frequency ω, the worst case in terms of uncertainty is $\Gamma(\omega, f((i\omega))) := \sup_{\alpha \in \mathcal{P}_\omega} \tilde{\Gamma}(\omega, \alpha, f(i\omega))$. Thus we have:

UncOPT Given a performance $\tilde{\Gamma}(\omega, \alpha, z)$, find

$$\inf_{f \in A} \sup_{\omega \in R} \sup_{\alpha \in \mathcal{P}_\omega} \tilde{\Gamma}(\omega, \alpha, f(i\omega))$$

and find a minimizer f if it exists.

*Supported in part by the AFOSR and the NSF

One connection between OPT and UncOPT [HM] is that UncOPT can always be viewed as an OPT problem, by defining Γ as

$$\Gamma(\omega, z) = \sup_{\alpha \in \mathcal{P}_\omega} \tilde{\Gamma}(\omega, \alpha, f(i\omega)) \tag{1}$$

The problem OPT has been studied for more than 10 years and is becoming well understood, see [HM], [HMW], [OZ] and much earlier references cited there. Much less is known about UncOPT. [HM] gives optimality conditions for the problem under strong hypotheses.

2 Background

Compromise methods. A common approach to treating plant uncertainty is to model the uncertainty α so that the frequency response function of the system is a linear fractional transformation of α and of the designable function f.

This gives a class of performance functions which equal the norm of linear fractional transformations $\mathcal{L}$ of the frequency response of the system.

$$\tilde{\Gamma}(\omega, \alpha, f(i\omega)) = ||\mathcal{L}_\omega(f(i\omega), \alpha)||$$

That a f meets a performance specification simply means that

$$\tilde{\Gamma}(\omega, \alpha, f(i\omega)) \leq 1 \qquad (\forall\alpha)(\forall\omega)$$

Thus one wants to calculate

$$\inf_f \sup_\omega \sup_\alpha \tilde{\Gamma}(\omega, \alpha, f(i\omega)) \tag{2}$$

The famous μ of John Doyle gives a performance function Γ_μ and a function $W_\omega(z)$ which is a l.f.t. in z for each ω, so that[1] $\Gamma_\mu(\omega, f(i\omega)) = \mu_{\mathcal{Q}}(W_\omega(f(i\omega)))$. Here $\mathcal{Q}$ = is a suitable algebra of block diagonal matrices. It is known that the expression (2) is equivalent to

$$\inf_f \sup_\omega \Gamma_\mu(\omega, f(i\omega)) \tag{3}$$

The calculation of (3) is difficult and Doyle proposes a very effective compromise which replaces the true performance by

$$\inf_f \sup_\omega \inf_{D\, comm.\, \mathcal{Q}} \tilde{\Gamma}_\nu(\omega, D, f(i\omega)) \tag{4}$$

where $\tilde{\Gamma}_\nu(\omega, D, f) = ||DW_\omega(f(i\omega))D^{-1}||$, and the D's are diagonal matrices that commute with $\mathcal{Q}$.

[1]The *spectrum of a matrix W with respect to a set* $\mathcal{Q}$ of matrices is $\sigma_{\mathcal{Q}} = \{\xi \in \mathcal{Q} : W - \xi \text{ is not invertible}\}$ and the spectral radius $\mu_{\mathcal{Q}}(W)$ of W w.r.t. $\mathcal{Q}$ is $\mu_{\mathcal{Q}}(W) = \sup \inf_x \frac{||\xi x||}{||x||}$ $\mu_{\mathcal{Q}}(W)$ is also referred to as the *structured singular value.*

Calculation of solutions. We see that (2) is an UncOPT problem, (3) is an OPT problem, and that (4) is an UncOPT problem (or at least equivalent to one). The UncOPT problems have performance of the form matrix norm of a smooth function. Newton type methods do not apply directly and currently we are working on extending them to these cases.

Computational efforts to do synthesis with (4), called μ-synthesis, are currently based on iterating on each coordinate: Fix f and ω and minimize $\tilde{\Gamma}_\nu$ over D's to get $D_0(\omega)$. Then solve an OPT problem for the performance $\tilde{\Gamma}_\nu(\omega, D_0(\omega), f)$ to obtain an update to f, and so on. The mathematics in this talk is aimed at developing a theory for updating D and f simultaneously: *optimality conditions are used explicitly in order to determine the update.* Usually such schemes have a advantages over coordinate descent.

Change of coordinates For mathematical convenience we will use the unit circle instead of the imaginary axis in the complex plane, and the unit disk instead of the R.H.P. It is well known that this change of coordinates is accomplished with a linear fractional transformation.

Suppose that a one parameter family $\mathcal{P} = \{\mathcal{P}_\theta : 0 \leq \theta < 2\pi\}$ of subsets of R^k is given. We refer to $\mathcal{P}_\theta$ as an *uncertainty set.* Here we consider only the case when $\mathcal{P}_\theta$ is given in terms of a smooth function $g(e^{i\theta}, \alpha)$:

$$\mathcal{P}_\theta = \{\ \alpha \in R^k : g(e^{i\theta}, \alpha) \leq 0\ \}$$

and

$$\partial\mathcal{P}_\theta = \{\ \alpha \in R^k : g(e^{i\theta}, \alpha) = 0\ \}$$

Furthermore, assume that $g(e^{i\theta}, \alpha) = g(e^{-i\theta}, \alpha)$ for all θ and all $\alpha \in \mathcal{P}_\theta$, so in particular, $\mathcal{P}_\theta = \mathcal{P}_{-\theta}$. The function g is referred to as *defining function* for $\mathcal{P}_\theta$.

A *performance function* is a function $\tilde{\Gamma} : \partial D \times R^k \times C^N \to R^+$. The performance $\tilde{\Gamma}$ is *real symmetric* if $\tilde{\Gamma}(e^{i\theta}, \alpha, z) = \tilde{\Gamma}(e^{-i\theta}, \alpha, \overline{z})$. All performance functions treated here are assumed to be real symmetric. We reserve the use of the symbol Γ to denote a performance function that is not a function of α.

By A_N we denote the space of all C^N-valued functions f that are continuous on the closed unit disk D^- and analytic on the open disk D, such that $\overline{f(\zeta)} = f(\overline{\zeta})$. *The problems OPT and UncOPT are now assumed to be given in terms of $e^{i\theta}$ instead of $i\omega$, $f(e^{i\theta})$ instead of $f(i\omega)$ etc.*

Notation. Elements $z \in \mathbf{C}^N$ are represented as column vectors. By z^t we denote the transpose of z, so that $\bar{z}^t z = ||z||^2_{C^N}$. By $\mathcal{L}^\infty_N$ we represent the set of all $\mathbf{C}^N$ valued measurable functions f on ∂D with finite essential supremum norm $||f||_\infty = \sup_\theta ||f(e^{i\theta})||_{C^N}$, and $\mathcal{H}^\infty_N$ consists of those elements of $\mathcal{L}^\infty_N$ which extend to analytic functions on $\mathbf{D}$. The spaces $\mathcal{L}^p_N$ and

$\mathcal{H}_N^p$ ($p \geq 1$) are defined similarly in terms of the norm $\|f\|_p = \left(\int_0^{2\pi} \|f(e^{i\theta})\|_{C^N}^p \frac{d\theta}{2\pi}\right)^{1/p}$. Functions $f \in \mathcal{L}_N^1$ have an associated Fourier series expansion

$$f(e^{i\theta}) \sim \sum_{\ell=-\infty}^{\infty} \hat{f}_\ell e^{i\ell\theta}$$

where the constants $\hat{f}_\ell \in C^N$ are called Fourier coefficients. In engineering applications only $\hat{f}_\ell \in R^N$ occur, hence we assume that this relation holds for all functions considered in this talk.

Also, when $N = 1$ (the scalar case), by $\frac{\partial\Gamma}{\partial z}(e^{i\theta}, z)$ we denote the complex partial derivative of Γ with respect to z ($e^{i\theta}$ is fixed). When $N > 1$, $\frac{\partial\Gamma}{\partial z}(e^{i\theta}, z)$ is an N-vector with entries $\frac{\partial\Gamma}{\partial z_\ell}(e^{i\theta}, z)$.

3 Results

Before presenting the main result, we give a theorem that gives necessary conditions for solutions of problems OPT where the performance is of the form a maximum of several smooth functions. Thus there is no uncertainty parameter. This extension to multiple performance functions is a variant on results found in Helton and Vityaev [HV].

Theorem 1 *Let $\Gamma_\ell(e^{i\theta}, z)$, $\ell = 1, \ldots, k$ be smooth positive valued performance functions, and let $f^* \in A_N$. Set*

$$\Gamma(e^{i\theta}, z) = \max_{1 \leq \ell \leq k} \Gamma_\ell(e^{i\theta}, z) \tag{5}$$

and

$$\gamma^* = \sup_\theta \Gamma(e^{i\theta}, f^*(e^{i\theta})) \tag{6}$$

and suppose that for each $\theta \in Q(f^*) = \{\theta : \Gamma^*(e^{i\theta}, f^*(e^{i\theta})) = \gamma^*\}$, condition (7) below holds:

$$(\exists z \in C^N)\ (\forall \ell)\ \mathrm{Re}(\frac{\partial \Gamma_\ell}{\partial z}(e^{i\theta}, f^*(e^{i\theta}))^t z) < 0 \tag{7}$$

Let χ be the function $\chi(e^{i\theta}) = e^{i\theta}$. If f^* is a local solution to OPT, then there exist nonnegative valued measurable functions $\mu_1, \ldots, \mu_k$, with $\sum_{\ell=1}^k \mu_\ell > 0$ and normalized by $\int \sum_{\ell=1}^k \mu_\ell = 1$, and there exists $F \in H_N^1$ such that

$$\begin{array}{rcll} \mu_\ell(\gamma^* - \Gamma_\ell(\cdot, f^*)) &=& 0, \quad 1 \leq \ell \leq k & \text{(i)} \\ \mu_1 \frac{\Gamma_1}{\partial z}(\cdot, f) + \ldots + \mu_k \frac{\Gamma_k}{\partial z}(\cdot, f) &=& \chi F & \text{(ii)} \end{array} \tag{8}$$

When taken together with the relations $\sum_{\ell=1}^k \mu_\ell > 0$ and $\gamma^* - \Gamma_\ell(\cdot, f^*) \geq 0$, condition (8.i) is the well known flatness necessary condition for solutions [HH], [HMar], [HMer]. We remark that conditions (8.i) may be stated in the following equivalent form:

$$\int \sum_{\ell=1}^k \mu_\ell\, \Gamma_\ell\,(\,\cdot\,,\, f\,)\ =\ \gamma^*$$

Condition (8.ii) (from [HV]) generalizes the known condition on the complex gradient of smooth performance functions [HMer]. A geometric version of it is that the dual cone to the level sets of $\Gamma(e^{i\theta}, z)$ in z at $z = f^*(e^{i\theta})$ contains $\overline{F(e^{i\theta})}$, where F is an analytic function. The hypothesis (7) is the natural one if one generalizes the known results in [HM]. This is a non-degeneracy type conditions which guarantee that the sublevel sets $\mathcal{S}_\theta = \{z \in C^N : \Gamma(e^{i\theta}, z) \leq \gamma^*\}$ have a well defined dual cone at $z = f^*(e^{i\theta})$ with nonempty interior. Theorem 1 can be proved with dual extremal techniques (also see [HV]), but we do not have space here so we refer the reader to the paper [HMW3] (in preparation).

Our second result gives necessary conditions for solutions to UncOPT.

Theorem 2 *Let $\tilde{\Gamma}(e^{i\theta}, \alpha, z)$ be smooth positive functions, and let $\mathcal{P}_\theta = \{\alpha \in R^k : g(e^{i\theta}, \alpha) \leq 0\}$. Let $f^* \in A_N$ and set*

$$\gamma^* = \sup_\theta \sup_{\alpha \in \mathcal{P}_\theta} \tilde{\Gamma}(e^{i\theta}, \alpha, f^*(e^{i\theta})) \tag{9}$$

Suppose that for each θ in the peak set $Q(f^*) = \{\theta : \sup_{\alpha \in \mathcal{P}_\theta} \tilde{\Gamma}(e^{i\theta}, \alpha, f^*(e^{i\theta})) = \gamma^*\}$, the equation

$$\gamma^* = \sup_{\alpha \in \mathcal{P}_\theta} \tilde{\Gamma}(e^{i\theta}, \alpha, z) \tag{10}$$

has at most k solutions $\alpha = \beta_1$, $\alpha = \beta_2, \ldots$. For every $e^{i\theta} \in Q(f^*)$, set

$$\Gamma_\ell(e^{i\theta}, z) = \tilde{\Gamma}(e^{i\theta}, \beta_\ell, f^*(e^{i\theta})) \tag{11}$$

and suppose that for all $e^{i\theta}$ in the peak set $Q(f^*)$, relation (7) holds. Let χ be the function $\chi(e^{i\theta}) = e^{i\theta}$.

If f^* is a local solution to Unc-OPT, then there exist real valued measurable functions $\mu_\ell \geq 0$, $\ell = 1, \ldots, k$, such that $\sum_{\ell=1}^k \mu_\ell > 0$ and $\int \sum_{\ell=1}^k \mu_\ell = 1$, there exist R^k-valued functions $\alpha_1, \ldots, \alpha_k$, there exist real valued measurable functions $\eta_\ell \geq 0$, $\ell = 1, \ldots, k$, and there exists $F \in H_N^1$ such that

$$\begin{array}{rcll} \mu_\ell\ (\ \gamma^* - \tilde{\Gamma}(\cdot, \alpha_\ell, f^*)\ ) &=& 0, \quad 1 \leq \ell \leq k & \text{(i)} \\ \sum_{\ell=1}^k\ \mu_\ell\ \frac{\partial \tilde{\Gamma}}{\partial z}(\cdot, \alpha_\ell, f^*) &=& \chi F & \text{(ii)} \\ \frac{\partial \tilde{\Gamma}}{\partial \alpha}(\cdot, \alpha_\ell, f^*) - \eta_\ell \frac{\partial g}{\partial \alpha}(e^{i\theta}, \alpha_\ell) &=& 0, \quad 1 \leq \ell \leq k & \text{(iii)} \end{array} \tag{12}$$

4 Algorithms

Optimality conditions for H^∞ problems give insight and often lead to algorithms for finding solutions, see [HMW2]. A theme common thruout numerical optimization is to look at the first order optimality conditions for your problem as a collection of nonlinear of equations $\mathcal{O}$ which one can then attack with Newton's method. For this to work the linearized $\mathcal{O}$ equations must be an invertible system.

Unfortunately, the equations (12) for UncOPT (as we found with OPT [HWM2]), have highly degenerate linearizations; there are too many unknowns. Thus it is important to do some transformations which cut down the number of unknowns in (12) before applying Newton's method. Now we give an approach to doing this.

Let $\epsilon_l > 0$ be a small number. We mention first a modification of the flatness condition (12.i) that helps with numerics and which allows continuous solutions:

$$\mu_\ell \; (\; \gamma^* - \tilde{\Gamma}(\cdot, \alpha_\ell, f^*) \;) = \epsilon_l, \quad 1 \leq \ell \leq k \tag{13}$$

By $\mathbf{P}[\cdot]$ we denote the orthogonal projection of L^2 onto the space $(H_0^2)^\perp$. That is, $\mathbf{P}[\; f \;] \; \sim \sum_{n=-\infty}^{0} \hat{f}_n e^{in\theta}$. By taking projections in (12.ii) we eliminate F. This, with (12.iii), and (13) give:

$$\begin{aligned} \mu_\ell \; (\; \gamma^* - \tilde{\Gamma}(\cdot, \alpha_\ell, f^*) \;) \; - \epsilon_l \;\; &= \;\; 0, \quad 1 \leq \ell \leq k \\ \textstyle\sum_{\ell=1}^{k} \; \mathbf{P}[\; \mu_\ell \, \frac{\partial \tilde{\Gamma}}{\partial z}(\cdot, \alpha_\ell, f^*) \;] \;\; &= \;\; 0 \\ \tfrac{\partial \tilde{\Gamma}}{\partial \alpha}(\cdot, \alpha_\ell, f^*) - \eta_\ell \tfrac{\partial g}{\partial \alpha}(e^{i\theta}, \alpha_\ell) \;\; &= \;\; 0, \quad 1 \leq \ell \leq k \end{aligned} \tag{14}$$

Let us denote by $T(f, \mu, \alpha, \gamma, \eta)$ the left-hand-side of (14). The algorithm we describe is based on applying Newton's Method to solve $T(f, \mu, \alpha, \gamma, \eta) = 0$ for $f \in A_N$, α_ℓ, $\eta_\ell \geq 0$, $\mu_\ell \geq \epsilon$, and $\gamma \in R^+$, such that $\sum_\ell \mu_\ell \geq 0$ and $\gamma - \Gamma(\cdot, \alpha_\ell, f) \geq \epsilon$. This is in the spirit of the highly successful primal dual type algorithms discussed in [HMW1], [HMW2]. Also one sees parallels with modern algorithms (like [HO]) for solving matrix optimization problems ,so possibly a unified approach will emerge.

Note that the derivation of (14) is formal since at optimum the function F that appears in the optimality conditions may not be in L^2, so $\mathbf{P}$ may not apply to it. Any practical control solution must be smooth (of modest order as well). The leaves much room for regularizing bad $F's$.

Algorithm Let $x^0 = (f^0, \mu^0, \alpha^0, \gamma^0, \eta^0)$ be given, where $\mu_\ell^0 \geq \epsilon$ and

$$\gamma^0 = \epsilon + \sup_{\theta} \; \sup_{\alpha : g(e^{i\theta}, \alpha) \leq 0} \; \tilde{\Gamma}(e^{i\theta}, \alpha(e^{i\theta}), f^0(e^{i\theta}))$$

To produce an update x^1, write $\Delta x = (\Delta f^0, \Delta \mu^0, \Delta \alpha^0, \Delta \gamma^0, \Delta \eta^0)$, and:

1. Solve
$$T'_{x^0}[\Delta x] = -T[x^0]$$
2. Update $f^1 = f^0 + \Delta f$, $\mu^1 = Max(\mu^0 + \Delta\mu, \epsilon)$, $\eta^1 = Max(\eta^0 + \Delta\eta, 0)$, $\alpha^1 = \alpha^0 + \Delta\alpha$.

Also, a natural thing to try in combination of (1) and (2) is a linesearch to improve on global behavior. Implementation of the algorithm proposed here and it's analysis is work in progress. Based on the previous experience [HMW1], [HMW2], we expect that an algorithm similar to this one will exhibit good local convergence behavior until singularities of the optimum become pronounced.

Note that using the ideas discussed here (many occuring in [HV]) it is a simple matter to generate an algorithm for solving the problem OPT (no uncertainty) where the performance function is of the form the maximum of k given performance functions, as is the case in Theorem 1. We have yet to study such algorithms.

5 References

[HH] J. W. Helton and R. Howe. A bang-bang principle for the frequency domain. J. Approx. Th. **47** (1986), 101-121.

[HMar] J. W. Helton and D. Marshall. Optimization in the frequency domain and analytic selections. U. of Indiana Math. J. 1991.

[HM] J. W. Helton and O. Merino. Conditions for Optimality over $\mathcal{H}^\infty$. SIAM J. Cont. Opt., 1993.

[HMW1] J. W. Helton, O. Merino and T. E. Walker. Conditions for optimality over H^∞: Numerical Algorithms. Proceedings IEEE's CDC 1992, Tucson, Arizona.

[HMW2] J. W. Helton, O. Merino and T. E. Walker. Algorithms for Optimizing over Analytic Functions. U. of Indiana Math. J. (42) No. 3, pp 839–874.

[HMW2] J. W. Helton, O. Merino and T. E. Walker. Paper in preparation.

[HO] J. P. Haeverly and M. Overton. Optimizing eigenvalues of symmetric definite pencils. Proceedings of the 1994 ACC. Baltimore, June 30-July2.

[HV] J. W. Helton and A. Vityaev. In preparation.

[OZ] J. G. Owen and G. Zames, Duality Theory for MIMO Robust Disturbance Rejection. IEEE TAC, May 1993 (38) No. 5, pp 743 – 752.