there exists a constant $\delta_2 > 0$ such that

$$\text{Re } a(e^{j\theta})h(e^{j\theta}) < -\delta_2 < 0 \qquad \forall e^{j\theta} \in \tau^c \tag{12.6}$$

We use the Taylor expansion of Γ to see that for $\epsilon > 0$ small enough[1] the function $f = f^* + \epsilon h$ satisfies

$$\sup_\theta \Gamma(e^{j\theta}, f(e^{j\theta})) < \sup_\theta \Gamma(e^{j\theta}, f^*(e^{j\theta})) - \epsilon\delta_2 \qquad \forall\theta \in \tau^c \tag{12.8}$$

Of course ϵh might be very large on τ, and if ϵ is not small enough, then $\Gamma(e^{j\theta}, f(e^{j\theta}))$ may be larger than γ^* for $\theta \in \tau$. However, we can take ϵ so small that

$$\sup_{\theta\in\tau} |\Gamma(e^{j\theta}, f^*(e^{j\theta}) + \epsilon h(e^{j\theta}))| < \gamma^* - \delta_1/2 \tag{12.9}$$

Then $\Gamma(e^{j\theta}, f(e^{j\theta}) + \epsilon h(e^{j\theta}))$ is uniformly less than γ^* for *all* θ, contradicting the assumption that f^* is optimal.

12.3 Solutions satisfy the winding number condition

We now may assume $\Gamma(e^{j\theta}, f^*(e^{j\theta}))$ is constant $= \gamma^*$. If $m :=$ wind (a ; 0) ≤ 0, then the function $b(e^{j\theta}) = (e^{j\theta})^m a(e^{j\theta})$ is never 0 and has winding number 0 about 0. Then by Theorem 18 there is a function h in A whose phase approximates the phase of $-\overline{b}$, so that there exists a constant $\delta_2 > 0$ such that

$$2\text{Re}a(e^{j\theta})h(e^{j\theta}) < -\delta_2 < 0 \;\; \forall\theta \tag{12.10}$$

Compare (12.10) with (12.6): the former is valid for all θ and the latter is valid for $\theta \in \tau^c$. We can proceed with steps similar to those that follow (12.6) and with a small enough ϵ we obtain

$$\sup_\theta \Gamma(e^{j\theta}, f^*(e^{j\theta}) + \epsilon h(e^{j\theta})) < \gamma^* - \epsilon\delta_2/2, \;\forall\theta$$

This produces the contradiction

$$\sup_\theta \Gamma(e^{j\theta}, f(e^{j\theta})) < \sup_\theta \Gamma(e^{j\theta}, f^*(e^{j\theta}))$$

[1] A precise argument is, choose $\epsilon > 0$ such that $\epsilon c_0 ||h||^2_\infty < -\delta_2$, that is, $c_0||\epsilon h||^2_\infty < -\epsilon\delta_2$. Combining this with inequality (12.5) we have that

$$\begin{aligned} \Gamma(e^{j\theta}, f^*(e^{j\theta}) + \epsilon h(e^{j\theta})) &\leq \gamma^* + 2\text{Re}a(e^{j\theta})\epsilon h(e^{j\theta}) + c_0|\epsilon h(e^{j\theta})|^2 \\ &< \gamma^* - 2\delta_2\epsilon + \epsilon\delta_2 = \gamma^* - \epsilon\delta_2, \;\forall\theta \in \tau^c \end{aligned} \tag{12.7}$$

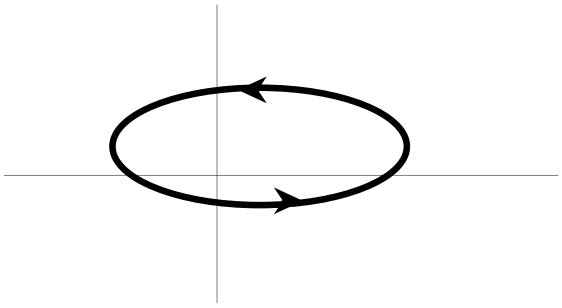

Figure 12.2: The function ah_n winds about 0.

12.4 Flatness and winding number conditions are sufficient

Suppose now that conditions (i) and (ii) hold, but that f^* is not an optimizer. To obtain a contradiction we assume that $\{h_n\}_{n\in N}$ is a sequence in A such that $||h_n||_\infty \to 0$ and

$$\sup_\theta \Gamma(e^{j\theta}, f^*(e^{j\theta}) + h_n(e^{j\theta})) < \sup_\theta \Gamma(e^{j\theta}, f^*(e^{j\theta})) =: \gamma^* \tag{12.11}$$

Before launching into heavy details we give the idea of the proof. Taylor approximation and the flatness condition (i) of Theorem 13 imply that

$$\Gamma(e^{j\theta}, f^*(e^{j\theta}) + h_n(e^{j\theta})) \approx \gamma^* + 2\mathrm{Re}(a(e^{j\theta})h_n(e^{j\theta})). \tag{12.12}$$

The key point is that since h_n is analytic on the disk, then

$$\text{wind}\,(ah_n; 0) = \text{wind}\,(a; 0) + \text{wind}\,(h_n; 0) > 0$$

provided h_n has no zeros on the circle. (If it does, then one can replace it by a function with zeros on the circle that is a small perturbation of h_n). This says that the function ah_n must wind about 0 (see Fig. 12.2), so in particular $\mathrm{Re}(ah_n)$ must be *positive* for some $e^{j\theta_1}$. Hence

$$\Gamma(e^{j\theta_1}, f^*(e^{j\theta_1}) + h_n(e^{j\theta_1})) \approx \gamma^* + 2\mathrm{Re}(a(e^{j\theta_1})h_n(e^{j\theta_1})) > \gamma^*$$

But this contradicts the assumption (12.11).

Now we give a formal proof which includes the second order term in the Taylor expansion. Since $||h_n||_\infty \to 0$ we may assume that $||h_n||_\infty < \delta_0$

without loss of generality, and we have from (12.5) that

$$\begin{aligned} -c_0|h_n(e^{j\theta})|^2 \quad &\leq \quad \Gamma(e^{j\theta}, f^*(e^{j\theta}) + h_n(e^{j\theta})) - \gamma^* - 2\mathrm{Re}(a(e^{j\theta})h_n(e^{j\theta})) \\ &\leq c_0 \quad |h_n(e^{j\theta})|^2, \quad \forall\ e^{j\theta} \end{aligned} \tag{12.13}$$

Hence

$$\begin{aligned} -c_0|h(e^{j\theta})|^2 + \gamma^* \quad &+ \quad 2\mathrm{Re}(a(e^{j\theta})h_n(e^{j\theta})) \\ &\leq \quad \Gamma(e^{j\theta}, f(e^{j\theta}) + h_n(e^{j\theta})) \leq \gamma^* \quad \forall e^{j\theta} \end{aligned} \tag{12.14}$$

i.e.,

$$-c_0|h(e^{j\theta})|^2 + 2\mathrm{Re}a(e^{j\theta})h_n(e^{j\theta}) < 0 \quad \forall e^{j\theta}. \tag{12.15}$$

Suppose for a moment that for each n there exists θ such that $a(e^{j\theta})h_n(e^{j\theta}) = t_n > 0$. Substituting in (12.15) we get

$$-\frac{c_0}{|a(e^{j\theta})|}t_n^2 + 2t_n < 0 \tag{12.16}$$

i.e.,

$$-\frac{c_0}{|a(e^{j\theta})|}t_n + 2 < 0 \tag{12.17}$$

But since $||h_n||\infty \rightarrow 0$, we have that $t_n \rightarrow 0$, so the left hand side of 12.17 cannot hold for all n.

To justify the existence of t_n as above, we note that the complex function $H_n(z) = zh_n(z)$ has a zero at $z = 0$. By complex function theory, the function H maps onto some disk centered at 0 so for some θ we have $H_n(e^{j\theta}) > 0$. Since the function $a(e^{j\theta})/e^{j\theta}$ is continuous and never 0, it follows that for some θ,

$$\frac{a(e^{j\theta})}{e^{j\theta}}H_n(e^{j\theta}) = a(e^{j\theta})h_n(e^{j\theta}) = t_n > 0 \tag{12.18}$$

This completes the proof of the Theorem. ∎

Chapter 13

Computer Solutions to OPT

One of the big payoffs of having results such as Theorem 13 available is that one can develop optimality tests. These tests that can be used as a measure of how close a candidate to solution to OPT is to being optimal. This topic is presented in Section 13.1. Section 13.3 describes how Theorem (13) leads to an effective computer algorithm. This requires some basic definitions which are given in Section 13.2.

13.1 Computer diagnostics

An extremely valuable tool in computation, especially with iterative methods, is a collection of *diagnostics* which indicate if the iteration is close to a solution or not close to a solution. The diagnostics which over the years we found effective for OPT are based directly on the general mathematical Theorem 13. This has the advantage that the diagnostics are not algorithm specific, so we recommend them to anyone producing H^∞ software. Indeed we make heavy use of them in *Anopt*, our H^∞ optimization software program.

Suppose that $f \in A$ is a candidate for a solution to OPT. To test Flatness Condition of Theorem 13, we can calculate

$$OT_1(f) := \frac{\sup_\theta \ \Gamma(e^{j\theta}, f(e^{j\theta})) - \inf_\theta \ \Gamma(e^{j\theta}, f(e^{j\theta}))}{\sup_\theta \ \Gamma(e^{j\theta}, f(e^{j\theta}))}$$

To implement [1] condition II, compute

$$OT_2(f) := \text{wind} \;(\; a \;;\; 0\;) - 1$$

by evaluating the phase of f on a grid and taking differences of these values on adjacent grid points. At optimum $OT_1(f) = 0$ and $OT_2(f) \geq 0$. Optimization programs for solving OPT should print $OT_1(f)$ and $OT_2(f)$ at each iteration so that the user can monitor how the approximate solutions f are progressing. Of course, if f is near the solution we must have $OT_1(f) \approx 0$ and $OT_2(f) \geq 0$.

To illustrate a practical set of diagnostics we solve the problem OPT when Γ is given by

$$\Gamma(e^{j\theta}, z) = |0.8 + (1/e^{j\theta} + z)^2|^2$$

A computer run produces the numbers shown in Table 1.

Table 1

It	Current Value SupGamma	Optimality Tests OT1	OT2	True error
0	3.24 E+00	9.9E-01	0 E+00	??????
1	1.3456107405895E+00	4.8E-01	0 E+00	??????
2	1.0242722917577E+00	4.9E-02	0 E+00	??????
3	1.0100847099506E+00	1.8E-02	0 E+00	??????
4	1.0005821314497E+00	2. E-03	0 E+00	??????

In our run above, the diagnostic OT1 went from 0.99 down to 0.002 in 4 iterations, while OT2 was zero throughout the iteration (this is not unusual for "scalar" cases). The column "true error" was added to illustrate that, typically, the true error $||f_k - f^*||_\infty$ is of the same order of magnitude than the numbers under OT1 and OT2.

The numbers in Table 1 were obtained with the program Anopt. The actual input used in this example is:

```
<<Anopt`;
```

[1] Since the winding number of a function can only take integer values, small changes in a function typically do not change the winding number of it with respect to 0. Condition II is automatically satisfied for all functions f that are close (in the norm of the supremum) to a solution f^*!

```
g = Abs[0.8 + (1/e + z[1] )^2 ]^2;

    Anopt[g, 0.01];
```

The reader who wants more details may see the document Anopt.ma in the Appendix.

13.2 Spaces of functions

This section serves as background for Section 13.3 and for the appendices.

Functions $f(z)$ in A have a power series expansion about 0 valid for any z in the unit disk:

$$f(z) = \sum_{\ell=0}^{\ell=\infty} c_\ell z^\ell \tag{13.1}$$

where the coefficients a_ℓ are real numbers. On the other hand, if we now consider $f(e^{j\theta})$ as a function on the circle, it is a continuous function and thus is has a Fourier series expansion,

$$f(e^{j\theta}) = \sum_{\ell=-\infty}^{\ell=\infty} \hat{f}_\ell (e^{j\theta})^\ell \tag{13.2}$$

where the Fourier coefficients $\hat{f}_\ell$ are real numbers. The relationship between (13.1) and (13.2) becomes clear when we formally substitute $e^{j\theta}$for z in (13.1):

$$f(e^{j\theta}) = \sum_{\ell=0}^{\ell=\infty} c_\ell (e^{j\theta})^\ell \tag{13.3}$$

Now, by the theory of Fourier series we know that the representation (13.2) is *unique*. If (13.3) is valid, then we must have that $c_n = \hat{f}_n$ for $n = 0, 1, 2, \ldots$, and that $c_n = 0$ for $n = -1, -2, \ldots$. To prove that (13.3) holds for functions $f \in A$, one can consider first polynomials in z for which this relation is obvious. Then the general result follows from the fact that functions in A are limits of polynomials.

Now suppose that $f(e^{j\theta})$ is any function on the circle that is continuous. If the Fourier expansion of $f(e^{j\theta})$ has the form

$$f(e^{j\theta}) = \sum_{\ell=-\infty}^{\ell=\infty} \hat{f}_\ell (e^{j\theta})^\ell \tag{13.4}$$

we see that we can generate a function $f(z)$ defined on the disk as a power series in z, by formally replacing $e^{j\theta}$ by z in (13.4). It is possible to show that the radius of convergence is 1 (see **[Hof]**). Thus, roughly speaking, functions defined by relation (13.1) and functions defined by relation (13.2) can be identified.

If we drop from the definition of A the requirement that $f(z)$ be continuous on the closed unit disk, then we obtain a normed vector space called the Hardy space H^∞. It is obvious that some functions $f(z)$ in H^∞ have boundary values $f(e^{j\theta})$ (those in A for example). But it is not clear, a priori, that one can talk about boundary values of an arbitrary function $f(z)$ in H^∞. There is however a way to define $f(e^{j\theta})$ mathematically with the help of *Measure Theory* (see **[Hof]**).

If $f(e^{j\theta})$ is a *measurable* function $f : \partial D \to C$, the norms below are well defined and give a nonnegative value or infinity:

$$||f||_p = \{\int_0^{2\pi} |f(e^{j\theta})|^p \frac{d\theta}{2\pi}\}^{1/p} \tag{13.5}$$

and

$$||f||_\infty = esssup_\theta |f(e^{j\theta})| \tag{13.6}$$

The norms above are used to define the normed vector spaces

$$L^p = \{f : f \text{ is a measurable function on } \partial D \text{ and } ||f||_p < \infty\}, \tag{13.7}$$

which are Banach spaces.

The spaces that occur most often correspond to $p = 1, 2, \infty$. In particular, the following relation holds:

$$||f||_1 \leq ||f||_2 \leq ||f||_\infty, \tag{13.8}$$

therefore $L^\infty \subset L^2 \subset L^1$.

The space L^2 is particularly important because it is also a *Hilbert space* with inner product

$$< f, g >= \int_0^{2\pi} f(e^{j\theta})\overline{g(e^{j\theta})}\frac{d\theta}{2\pi} \tag{13.9}$$

All functions in L^1 do have an associated Fourier series expansion[2] (13.2). This fact can be used to define the Hardy spaces H^p,

$$H^p = \{f \in L^p : f(e^{j\theta}) = \sum_{\ell=0}^{\ell=\infty} \hat{f}_\ell e^{j\ell\theta}\}$$

[2] But this expansion may fail to give $f(e^{j\theta})$ for some values of $e^{j\theta}$.

Again, functions in H^p can be thought of as functions defined and analytic on the open disk by means of a Taylor series.

Since H^2 is a closed subspace of L^2 it has an orthogonal complement ${H^2}^\perp$ in L^2. It is easy to check that

$$ {H^2}^\perp = \{ \, f \in L^2 \; : \; f(e^{j\theta}) = \sum_{\ell=-1}^{\ell=-\infty} \hat{f}_\ell \, e^{j\ell\theta} \, \}. $$

If $f = \sum_{\ell=-\infty}^{\ell=\infty} \hat{f}_\ell e^{j\ell\theta}$ is in L^2, we define the operators P^- and P_0^- on L^2 by

$$ \begin{aligned} P^-[\, f \,] \; &= \; P^-[\, \textstyle\sum_{\ell=-\infty}^{\ell=\infty} \hat{f}_\ell \, e^{j\ell\theta} \,] = \; \textstyle\sum_{\ell=-\infty}^{\ell=-1} \hat{f}_\ell \, e^{j\ell\theta} \\ P_0^-[\, f \,] = \; &\; P_0^-[\, \textstyle\sum_{\ell=-\infty}^{\ell=\infty} \hat{f}_\ell \, e^{j\ell\theta}] = \; \textstyle\sum_{\ell=-\infty}^{\ell=0} \hat{f}_\ell \, e^{j\ell\theta} \end{aligned} \tag{13.10} $$

P^- is the orthogonal projection of L^2 onto ${H^2}^\perp$.

13.3 A Numerical Algorithm for solving OPT

Suppose that a function $T : R \to R$ is given. In calculus, students learn that the points $x \in R$ at which T attains a minimum satisfy $T'(x) = 0$ (i.e., x is a critical point), and that one can determine excellent candidates for x by finding these critical points. We will do exactly this in order to find minimizers to OPT, but in a setting where the element x belongs to function space, and T is a function on this space. Consequently, our strategy is to write down some equations $T'(f) = 0$ that minimizers f^* satisfy, and then solve the equations with an iterative method.

The most common method for solving nonlinear equations $T'(x) = 0$ is Newton's method. It is based on the following simple idea. Say x_k is your current guess at the solution x^*. To update you want to find h so that $T'(x_k + h) = 0$, but since you cannot solve for h, you settle for solving for h in the linearized equation

$$ T(x_k) + T'_{x_k}(h) = 0 \tag{13.11} $$

This equation is linear in h and so it is readily solvable for h provided that the linear map T'_{x_k} is invertible. So one solves (13.11) for h, and then set[3] $x_{k+1} = x_k + h$.

[3] Often $x_{k+1} = x_k + t * h$ is used, where $t > 0$ is a parameter chosen according to a prespecified criterion. The process of finding t is called linesearch.

One virtue of Newton's method is that it applies to a wide variety of situations, even when the unknowns belong to function spaces. In particular, it can be used to solve OPT by iteration, and this is the basis for the method presented in this section. We treat the scalar case here, but most of what we say also applies to the case of optimization over vector valued functions f (Chapter 15).

The first thing we do now is to rewrite the winding number condition (ii) of Theorem 13 in a more algebraic form. This makes it more suited to a Newton method solution.

Lemma 2 *In addition to the hypotheses to Theorem 13, assume that the function $a(e^{j\theta})$ is smooth. Then the conditions (i) and (ii) of Theorem 13 are equivalent to:*
There exist $\gamma \in R$, $f^* \in A$, $F \in A$, and λ a smooth positive valued function on the circle such that

$$\begin{aligned} \Gamma(\cdot, f) &= \gamma \\ \lambda^{-1} \tfrac{\partial \Gamma}{\partial z}(\cdot, f) &= e^{j\theta} F \end{aligned} \tag{13.12}$$

Proof. Suppose that f^* is such that conditions (i) and (ii) of Theorem (13) are satisfied. From (i) we immediately get the existence of the constant γ in (13.12). If wind $(\ a\ ;\ 0\ ) = n$, we have that the smooth function $b(e^{j\theta}) = a(e^{j\theta})/(e^{j\theta})^n$ has winding number 0 about 0. By Theorem 18 there exists an analytic function F_1 such that phase$(b) =$ phase(F_1). Another way to describe this relation is

$$\frac{b(e^{j\theta})}{|b(e^{j\theta})|} = \frac{F(e^{j\theta})}{|F_1(e^{j\theta})|}, \qquad \forall\theta \tag{13.13}$$

Thus by the definition of b we obtain

$$\frac{a(e^{j\theta})}{|a(e^{j\theta})|(e^{j\theta})^n} = \frac{F(e^{j\theta})}{|F_1(e^{j\theta})|}, \qquad \forall\theta \tag{13.14}$$

i.e.,

$$a(e^{j\theta}) = \frac{|a(e^{j\theta})|}{|F_1(e^{j\theta})|}(e^{j\theta})^n F_1(e^{j\theta}), \qquad \forall\theta \tag{13.15}$$

Now set $F = (e^{j\theta})^{(n-1)} F_1(e^{j\theta})$ and $\lambda = \frac{|a(e^{j\theta})|}{|F(e^{j\theta})|}$. Note that since by hypothesis $n \geq 1$, we have that $F \in A$. This proves half of the Lemma. The other direction is obvious. ∎

We now try to solve (13.12). The unknowns in equation (13.12) are the functions λ, F and f and the scalar γ. The difficulty with applying Newton's method to equation (13.12) is that, in a sense, it has more unknown functions (four) than constraints (two). This means that local solutions are not isolated, rather there is a whole continuum of them. Newton's method is well known for its bad behavior in such circumstances, because in the linearization (13.11) the differential T'_{x_k} will not be invertible. Consequently, rather than working with (13.12) directly, we modify it first to obtain a problem with only two unknowns. This is the key step in our method to finding solutions to (13.12).

We proceed with formal calculations for a moment. Begin by eliminating indeterminacy in (13.12) by assuming $\lambda^{-1} \in L^2_1$ is such that $\int_0^{2\pi} \lambda(e^{j\theta})^{-1} \frac{d\theta}{2\pi} = 1$. Note that $P^-[\ \Gamma(\cdot, f)\ ] = 0$ since $\Gamma(\cdot, f)$ is a constant function, and that $P_0^-[\ e^{j\theta} F\ ] = 0$ (provided $F \in H^2$). Thus we obtain the system

$$\begin{aligned} P^-[\ \Gamma(\cdot, f)\ ] &= 0 \\ P_0^-[\ \lambda^{-1}\ \tfrac{\partial \Gamma}{\partial z}(\cdot, f)\ ] &= 0 \end{aligned} \tag{13.16}$$

in two unknowns, λ and f. Because of the way λ is normalized, one can write

$$\lambda^{-1} = 1 + \Re(e^{j\theta}\beta) \tag{13.17}$$

where β is a scalar valued function which is analytic on the disk. Now, substituting (13.17) into (13.16) gives an operator equation of the form

$$T(f, \beta) = \mathbf{0} \tag{13.18}$$

At this point the basic idea is clear: to find solutions to the optimality conditions (i) and (ii) of Theorem 13, it is enough to solve equation (13.18). Thus we state the following:

Critical Point Problem.

Find analytic functions f^*, β^* with $1 + 2\Re(e^{j\theta}\beta^*)$ positive on the circle such that (f^*, β^*) is a solution to equation (13.18).

We have found that Newton's method applied to the Critical Point Problem is extremely effective:

Newton's Method Given (f, β). Solve the equation

$$T(f, \beta) + T'_{(f,\beta)}[\epsilon, \eta] = \mathbf{0}$$

for ϵ, η and produce an update $(f + \epsilon, \beta + \eta)$.

A standard fact about Newton iteration is that when the linearizations $T'_{(f,\beta)}$ are uniformly invertible, the iterates have an excellent convergence property called second order convergence.

This is analysed in **[HMW]** where the linearizations are shown to be invertible and consequently our algorithm is second order convergent. It requires quite a bit of functional analysis to find the correct setting and estimates.

While it is well beyond the scope of our presentation here, we give the formula for T' in Chapter 16.

Chapter 14

References

The optimization problem of approximating in the norm of the supremum a given continuous function on the circle by functions in the disk algebra is a primitive version of the problem OPT considered in this part of the book. As such it was treated first by Nehari in 1949, who proved that the distance is the largest singular value of certain Hankel operator. Many workers made contributions, one specially worth mentioning is the team Adamjan-Arov-Krein, who developed an extensive theory in a series of papers [**AAK1-AAK3**].

The mathematical theory described in this part of the book started with Helton and Howe [**HH**]. Theorem 13 was stated first and proved in [**H6**]. Theoretical work and other results pertaining Theorem 13 can be found in [**HMer2**], [**M**], [**Hu**], [**HMar**].

Numerical work describing steepest descent type of algorithms for solving OPT is in [**HMer3**]. The algorithm described here is in [**HMW**].

[**A**] Ahlfors. Complex Analysis.

[**AAK1**] V. M. Adamjam, D. Z. Arov and M. G. Krein. Infinite Hankel matrices and generalized problems of caratheodory-Fejer and F. Riesz, Functional Anal. Appl. 2 (1968), pp. 1-18.

[**AAK2**] V. M. Adamjam, D. Z. Arov and M. G. Krein. Analytic properties of Schmidt paris for Hankel operator and the generalized Scur-Takagi problem. Math. USSR-Sb. 15 (1972), pp. 15-78.

[**AAK3**] V. M. Adamjam, D. Z. Arov and M. G. Krein. Infinite block Hankel matrices and related extension problems, Amer. Math. Soc. Translations 111 (1978), 133-156.

[H1] J. W. Helton. "Operator theory and broadband matching," Proceedings of Eleventh Annual Allerton Conference on Circuits and Systems Theory, 1976.

[H2] J. W. Helton. "A mathematical view of broadband matching," IEEE International Conference on Circuits and Systems Theory, New York, 1978.

[H3] J. W. Helton. "Broadbanding gain equalization directly from data," IEEE Trans. Circuits and Systems, CAS-28 (1981), No. 12, 1125–1137.

[H4] J. W. Helton. "An H-infinity approach to control," IEEE Conference on Decision Control, San Antonio, Texas, December 1983

[H5] J. W. Helton. "Worst case analysis in the frequency domain: The H-infinity approach to control," IEEE Trans. Auto. Cont., AC-30 (December 1985), No. 12, 1154–1170.

[H6] J. W. Helton. "Optimization over spaces of analytic functions and the Corona problem," J. Operator Theory, 15 (1986), No. 2, 359–375.

[H7] J. W. Helton: Optimal frequency domain design vs. an area of several complex variables, pleniary address. MTNS, 1989.

[HH] J. W. Helton and R. Howe. "A bang-bang theorem for optimization over spaces of analytic functions," J. Approx. Theory, 47 (1986), No. 2, 101–121.

[HMar] J. W. Helton and D. Marshall: "Frequency domain design and analytic selections," Indiana Univ. Math. Journal, 39 (Spring 1990), No. 1, 157–184.

[HMer1] J. W. Helton and O. Merino: "Optimal Analytic Disks," Collection: Several complex variables and complex geometry, Part 2, (Santa Cruz, CA, 1991) Proc. Sympos. Pure Math., 5 251–262.

[HMer2] J. W. Helton and O. Merino. Conditions for Optimality over $\mathcal{H}^\infty$. SIAM J. Cont. Opt., Vol. 31, No. 6, November 1993.

[HMer3] A Novel Approach to Accelerating Newton's Method for sup-norm optimization arising in H^∞-control. J. Optim. Theory Apps. 1993

[HMW] J. W. Helton, O. Merino and T. E. Walker. Algorithms for Optimizing over Analytic Functions. Indiana U. Math. J., Vol 42, No.3 (1993) .

[HV] J. W. Helton, and A. Vityaev. Preprint

[Hof] K. Hoffmann. Banach spaces of analytic functions. Dover

[Hui] S. Hui. Qualitative Properties of Solutions to H^∞-Optimization. Problems. J. of Functional Analysis **75** (1987) 323,348.

[M] O. Merino. Optimization over spaces of analytic functions. Thesis. University of California, San Diego. 1988.

[R] W. Rudin. Real and complex analysis.

Part IV

H^∞ Theory: Vector Case

Chapter 15

Many Analytic Functions: Optimization for MIMO system design

The design methods presented in this book extend directly to multiple input, multiple output (MIMO) systems. Indeed one can reformulate design problems as optimization problems with interpolation constraints on matrix-valued analytic functions. Parameterizations of the type

$$T(s) = T_1(s) + T_0(s)F(s), \tag{15.1}$$

discussed in Chapter 8 for the scalar case, also arise but now the functions involved are matrix valued analytic functions. Just as we had in the scalar case, design problems correspond to OPT problems, but this time over N-tuples of analytic functions $(f_1, f_2, \ldots, f_n)$ instead of a single analytic function. These are the most challenging problems which arise in MIMO design, and fortunately much is known about them.

In this chapter, while we do not lay out MIMO control theory, we do give one the tools to solve the H^∞ optimization problem that arises from this subject. This tool is the computer program Anopt. It is written to optimize over analytic functions on the unit disk rather than the R.H.P. Anopt applies most directly to the design of discrete time systems. We warn the reader that the constrained optimization required to codify internal stability in the vector case requires additional programming, for example to produce automatically a formula analogous to (15.1).

We include in the appendix the Mathematica notebook AnoptTutorial.ma, which is an introductory manual for using the program Anopt.

15.1 The OPT problem

We denote with A_N the (real) vector space of all N-tuples $f = (f_1, \ldots, f_N)$ where f_ℓ is in A. The space A_N is a complete normed space (or Banach space), equipped with the norm

$$||f||_\infty = \max_\theta ||F(e^{j\theta})||_{C^N},$$

where $||z||_{C^N}$ represents the euclidean norm on C^N. A performance function is a positive valued function Γ on $\partial D \times C^N$. We will use the notations $\Gamma(e^{j\theta}, z)$ and $\Gamma(e^{j\theta}, z_1, \ldots, z_N)$ to denote the same object. The optimization problem OPT over the set A_N is

OPT Given a performance function $\Gamma(e^{j\theta}, z)$, find $\gamma^* \geq 0$ and f^* in A_N (if it exists) such that

$$\gamma^* = \inf_{f \in A_N} \sup_\theta \Gamma(e^{j\theta}, f(e^{j\theta})) = \sup_\theta \Gamma(e^{j\theta}, f^*(e^{j\theta})).$$

The function f^* in the statement OPT is called a *solution* or *minimizer*.

The theoretical study of OPT with most practical implications is the characterization of solutions. We already saw in the study of the scalar case how conditions for optimality lead quickly to practical tests for optimality, and to algorithms. The situation in the vector case is quite similar, and this will become apparent in the following sections.

15.2 Solutions only a keystroke away

The theory and algorithms of OPT are mature enough to permit the use of computers to obtain approximate solutions as well as diagnostics with few keystrokes. This might look like an exaggeration[1], so to convince the reader we present a computer solved example.

Consider the performance function

$$\begin{aligned}\Gamma(e^{j\theta}, \quad & z_1, z_2) = \\ & |\tfrac{1}{e^{j\theta}+0.5} - z_1|^2 + 2\mathrm{Im}(z_1\ z_2) \\ & +(5 + \mathrm{Re}(\tfrac{1}{2+e^{j\theta}}))|\tfrac{1}{e} + z_2|^2\end{aligned} \tag{15.2}$$

The following is the input the user types in a Mathematica session in order to solve OPT.

[1] indeed it is an exageration in cases which cause numerical difficulties.