Appendix C

Pure Mathematics and Problems

We have taken the viewpoint of engineering and computation in several parts of the book. Much of it could be looked at simply as mathematics. One implication of this work is that it suggests a new class of problems in functional analysis and operator theory or in several complex variables if that is your viewpoint.

First we emphasize that a pure mathematician might read Part III, then follow up with part IV. One who despises computation might read the how to run the program part quickly, but the parts which describe numerical algorithms actually could have been cast in much more mainstream mathematical terms. To describe this, first focus on a statement which formulates what we have done to the old paradigm Nehari problem.

Claim.

For the Nehari problem of approximating a given function $k \in L^\infty$ by functions $f \in H^\infty$,

$$\inf_{f \in H^\infty} ||k - f||_\infty,$$

there are two different sets of optimality conditions for solutions f each having the property that their associated differential is invertible at almost all locations.

1. *Equations (16.5) and (16.6) in Theorem 21 due to Nehari.*
2. *The key optimality condition $T(f, \beta) = 0$ determined by equations (13.16) and (13.18) in Chapter 13), or equations (16.8) and (16.9) in Chapter 16.*

We call this a claim rather than the theorem because the differentials are operators on spaces of functions and we did not say exactly what the spaces are.

Idea of Proof. Here we shall denote the Nehari-Commutant Lifting conditions (16.5) and (16.6) in Theorem 21 as a system of equations

$$\mathcal{N}(\alpha, \lambda) = 0$$

where $\lambda = \tau^2 > 0$ and $\alpha \in H^2_N$. The differential

$$\mathcal{N}'_{\alpha,\lambda} \begin{pmatrix} \varepsilon \\ \delta \end{pmatrix} := (R_k - \lambda)\varepsilon - \delta\alpha$$

where $R_k : H^2_N \longrightarrow H^2_N$ is defined by $R_k \varepsilon := P_{H^2_N} \left[\overline{k}^t P_{H^{2\perp}_N} k \varepsilon \right]$. for each $\varepsilon \in L^2_N$. Now R_k is a selfadjoint operator on L^2_N and λ is an eigenvalue of R_k which for almost all k has multiplicity one. Thus range of $R_k - \lambda$ has codimension one, so it is not invertible. Fortunately, the map $\delta \longrightarrow \delta\alpha$ has one dimensional range in L^2_N and since α is the null vector of $R_k - \lambda$ it is perpendicular to range $(R_k - \lambda)$. Therefore, range $\mathcal{N}^1_{\alpha,\lambda}$ equals L^2_N.

The optimality conditions for the OPT problem after projecting are denoted $T(f, \beta) = 0$ in chapters 13 and 16. The objective is to prove that $T'_{f,\beta}$ is invertible for almost all f, β which are solutions to OPT. The formula for $T'_{f,\beta}$ is given in Propostion 4 of Chapter 16. There are various spaces on which $T'_{f,\beta}$ could act. The easiest setting to analyze is when $T'_{f,\beta}$ acts from H^2_{N+1} to H^2_{N+1} (see [**HMW**]). Then it roughly has the form selfadjoint Toeplitz plus Hankel. The Hankel operator is compact and *the Toeplitz because of the special structure of* M_1 *can be shown to be invertible provided* wA *is positive definite and* $a(e^{j\theta}) = \Xi^n \lambda(e^{j\theta}) F(e^{j\theta})$ *where* F *is analytic inside the disk,* $\lambda \geq 0$ *and* $n \leq 1$. A is positive definite because Γ is strictly plurisubharmonic; indeed for the Nehari example $A = 1$. The optimality conditions guarantee that the condition on a holds and thus for "most" choices of symbol k we have invertibility. See Theorem 6 [**HMW**].
■

Unfortunately $H^2_N + 1$ is the wrong space for analyzing Newton's method since a nonlinear map rarely maps $H^2_N + 1$ into itself. Thus we did the analysis of invertibility in the space of C^∞ functions with a topology imposed by a family of Sobolev norms. The proof and precise formulation can be found in [**HMW**]. ■

The technique applied here makes sense for many optimization problems in function theory. Find optimality conditions $\mathcal{C}$ for the problem. These are a set of equations involving primal and dual variables. In our problem and probably in many others. The differential of $\mathcal{C}$ is not invertible. Find some way to modify $\mathcal{C}$ and produce new optimality conditions $\widetilde{\mathcal{C}}$ with invertible [1]left invertible will do differential $\widetilde{\mathcal{C}}'$. Thus we have the pure mathematics problem: find optimality conditions with invertible differential. Possibly this is a natural type of mathematical result to seek for many function theory problems.

What we are saying is hardly shocking from the viewpoint of optimization theory. The surprising thing is that for the class of problems we study one does not immediately get the invertibility condition. Also it is a bit surprising that such basic principles of optimization theory were only recently applied to the very old Nevanlinna-Pick-Nehari problem.

For perspective and to make sure that the reader interprets this speculation widely (wildly) enough we mention some of the more interesting work on optimization in several complex variables. This concerns N-P-N type problems on the polydisk. There is striking work of Agler which derives beautiful matrix inequality conditions for checking if a particular N-P interpolation problem is solvable but for a norm stronger than the supnorm. Agler's view implemented successfully several problems is that instead of doing supnorm optimization one modifies the norm in some way so that the answer to the problem can be gotten from an eigenvalue problem or something similar. A very elegant result of this type, due to Cotlar and Sadosky, converts the Nehari problem in the polydisk for a BMO type norm to an eigenvalue problem.

The suggestion here is that we liberalize our notion of what it means to answer a question in function theory beyond insisting that an answer must have the form of an eigenvalue problem. It seems like a very reasonable pursuit for function theory problems from the pure mathematics point of view is to find optimality conditions whose differential is invertible. While this seems like a considerable loosening of the eigenvalue condition if we insist that our conditions use only explicit formulas, then it may not be that liberal a condition. Only time will tell if these ideas extend in many other directions.

Also we mention that there are extremely pretty connections between the theory of analytic disks developed in this book and ongoing work in Several Complex Variables. The article **[HMer1]** describes one line of connections while the paper **[HV]** quickly indicates some others. Readers in the area of Several Complex Variables might find these articles interesting.

---

[1] 1

Appendix D

Uncertainty

D.1 Introduction

So far we have considered the system $\mathcal{S}$ where the plant P is known. However in most physical systems typically there is a certain amount of "plant uncertainty", i.e., lack of knowledge about the plant.

To explain what uncertainty is, we use the example of an airplane. We can model the control of an airplane with a system of differential equations. These equations are the result of an idealization of the airplane that invariably does not include some elements, either because we are not aware that they affect the airplane, or because we choose not to include them for simplicity and are deemed to be relatively unimportant. The differential equations have a series of parameters that have physical significance and whose values we can tell only approximately.

We may further simplify the model by linearizing the system of differential equations. The model is surely not perfect, indeed at high frequencies it is highly inaccurate. Also there are factors affecting the system, among which we have changes in the properties of the components or even malfunction of some components.

A common approach to treating uncertainty when doing system design is to assume that there is available a known *reference* or *nominal* plant P_0, and that the true plant P of the system is, in the frequency domain description,

$$P(j\omega) = P_0(j\omega) + \Delta_0 \tag{D.1}$$

where Δ_0 is unknown. We may think of P as the result of perturbing P_0.

Uncertainty can have a devastating effect on a system. For example, a stable system may cease to be stable after a perturbation of the plant. This definitely is something that we do not want under any circumstances, so we shall require that reasonable amounts of uncertainty should never make a system that we have designed unstable. Thus we insist that all feasible plants make the system *internally stable*. This is called *robust stability*.

Uncertainty also affects the performance of the system. Ideally, we would like to design the system with the property that when the plant is perturbed, the resulting system satisfies the performance requirements originally set. This must hole for any feasible perturbation of the plant. Sometimes this is referred to as *robust performance*.

D.2 Types of uncertainty

The first type of uncertainty arises when a set $\mathcal{P}_\omega$ of possible plant values at each frequency ω is known, but no specific information on the nature of the uncertainty is available. This is called *unparametric uncertainty*. Knowledge of the uncertainty set $\mathcal{P}_\omega$ is the main input to the design procedure. Here is one scenario in which the set $\mathcal{P}$ is determined: we carry out several experiments to measure $P(j\omega)$. This produces a (discrete) set of plant values $P(j\omega)$. This discrete set is approximated by a set $\mathcal{P}_\omega$ of possible plants as shown in Fig. D.1.

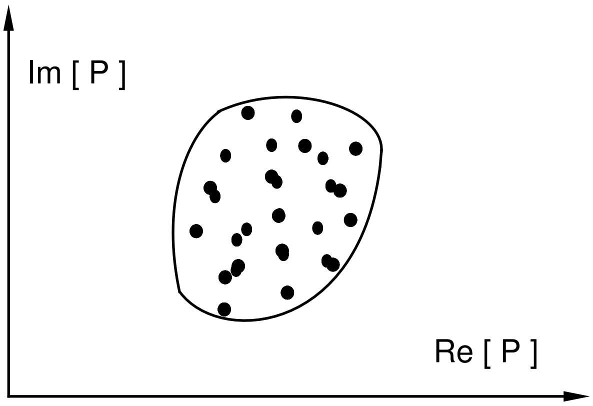

Figure D.1: The uncertainty set $\mathcal{P}_\omega$

The second type of plant uncertainty is called *parametric uncertainty*. It arises when it is assumed that the plant P may vary according to certain parameters ranging in some known sets. As an example, suppose that we are given the plant

$$P(s) = \frac{\beta}{s - \alpha}$$

where we also know that $0.5 < \alpha < 0.7$ and $1 < \beta < 2$. In this case we know something about P: it is a rational function with relative degree 1 and only one pole. We also know a range for the parameters α and β.

Finally, one may have a mixed type of uncertainty, where there are a parametric part and an unparametric part in ΔP. Either type of uncertainty can be converted to the unparametric type, this is not discussed further.

D.3 Dealing with uncertainty: worst case analysis

Now we treat unparametric uncertainty. We assume that at frequency ω, there is a *known* set $\mathcal{P}_\omega$ of possible plant values $p = P_0(j\omega) + \Delta$. In the discussion below we also assume that the compensator C is fixed, and the symbol C will be supressed from now on, so that *many expressions that actually depend on C will not show it explicitly.*

We also have a given performance (or cost) function. If we know the exact value of the plant P then the performance is denoted by

$$\tilde{G}(\omega, P(j\omega)) \tag{D.2}$$

The problem is that $P(j\omega)$ is unknown, so we must use a mathematical trick to produce a conservative estimate of the performance. We form a new performance function G by taking the worst case possible for plant value $P(j\omega)$:

$$G(\omega, P_0(j\omega)) = \sup_{p=P_0+\Delta\in\mathcal{P}_\omega} \tilde{G}(\omega, p) \tag{D.3}$$

Observe that whatever the true plant $P(j\omega)$ is, we have that

$$\tilde{G}(\omega, P(j\omega) \le G(\omega, P_0(j\omega)) \qquad (\forall\omega)$$

Thus when $G(\omega, P_0(j\omega))$ is small, so too is $\tilde{G}(\omega, P(j\omega))$.

The maximization in (D.3) is quite difficult to handle analytically, even with functions $\tilde{\Gamma}$ and sets $\mathcal{P}_\omega$ which are simple. Later we will calculate G explicitly for certain cases, but for now we concentrate on developing a practical alternative.

D.4 A method to treat plant uncertainty

If the sets $\mathcal{P}_\omega$ are known, sometimes an explicit formula can be obtained for the function G. However, this is almost always a very complicated calculation to do (when it is possible!).

Here we present a method for treating uncertainty that is conceptually simple (it also generalizes to the MIMO case), and is relatively easy to handle numerically. It is based on the construction of a performance function G^* that depends on P_0 and not on P, but which nevertheless includes uncertainty information. It is possibly the simplest compromise one might think of. Much more sophisticated approaches exist. Probably the most effective is J. Doyles' μ-synthesis, which is harder to describe. The exposition here is designed to encourage the reader to try his own compromises.

The idea is simple: to form the modified performance G^*, we use a Taylor expansion of $\tilde{G}(\omega, P_0(j\omega) + \Delta)$ about $\Delta = 0$.

$$\tilde{G}(\cdot, P_0 + \Delta) = \tilde{G}(\cdot, P_0) + 2\ Re\ (\frac{\partial \tilde{G}}{\partial p}(\cdot, P_0)\ \Delta) + \mathcal{O}(|\Delta|^2) \tag{D.4}$$

For simplicity, the variable ω is supressed in (D.4). The function G^* is formed by dropping from the r.h.s. of expression (D.4) all terms in Δ with order > 1, and by choosing Δ to have the direction of steepest ascent of $\tilde{G}$ at $p = P_0$. That is, Δ has the form

$$\Delta = \frac{\overline{\frac{\partial \tilde{G}}{\partial p}(\cdot, P_0)}}{|\frac{\partial \tilde{G}}{\partial p}(\cdot, P_0)|} * t \tag{D.5}$$

where t is chosen as the largest positive number such that $P_0 + \Delta$ belongs to $\mathcal{P}_\omega$. Thus

$$G^*(\cdot, P_0) = \tilde{G}(\cdot, P_0) + 2|\frac{\partial \tilde{G}}{\partial p}(\cdot, P_0)| * t, \tag{D.6}$$

where t depends on ω, P_0, $\tilde{G}$ and the set $\mathcal{P}_\omega$.

We shall see later that some simplification can be obtained by writing t as

$$t = \delta |P_0| \tag{D.7}$$

in which case we obtain

$$G^*(\cdot, P_0) = \tilde{G}(\cdot, P_0) + 2\delta \left| \frac{\partial \tilde{G}}{\partial p}(\cdot, P_0) \right|\ |P_0| \tag{D.8}$$

This is the case where the uncertainty sets $\mathcal{P}_\omega$ are disks about $p = P_0(j\omega)$, with (known) radius $\delta|P_0(j\omega)|$:

$$\mathcal{P}_\omega = \{\ P_0(j\omega) + \Delta\ :\ |\Delta| < \delta|P_0(j\omega)|\ \} \tag{D.9}$$

It is easy to see that in this case our t in the expression (D.6) is given by (D.7)

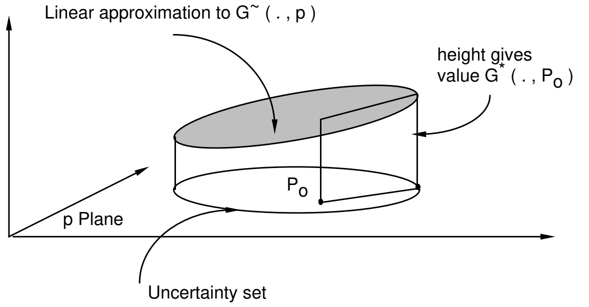

Figure D.2: The uncertainty set and the direction of largest increase in G.

D.5 An example with quasicircular Γ

To illustrate the method we will do the calculations for a performance function of the form

$$\tilde{G}(\omega, p) = W(j\omega)\ \left|\ k(j\omega) - \frac{p\ C(j\omega)}{1 + p\ C(j\omega)}\ \right|^2 \tag{D.10}$$

where C, W and k are given functions of frequency ω. A simple calculation yields

$$\left.\frac{\partial \tilde{G}}{\partial p}\right|_{p=P_0} = W\ \overline{\left(k - \frac{P_0\ C}{1 + P_0\ C}\right)}\ \frac{-C}{(1 + P_0\ C)^2} \tag{D.11}$$

Substituting (D.11) into (D.8), we obtain

$$G^*(\cdot, P_0) = W\ \left|\ k - \frac{P_0\ C}{1+P_0\ C}\ \right|^2 + 2\ \delta\ W\ \left|\ \left(k - \frac{P_0\ C}{1+P_0\ C}\right)\ \frac{P_0\ C}{(1+P_0\ C)^2}\ \right| \tag{D.12}$$

Note that the change of variables

$$T_0 = \frac{P_0\ C}{1 + P_0\ C}$$

in (D.12) produces an elegant formula in terms of the closed-loop transfer function T_0. Below we call Γ the resulting performance.

$$\Gamma(\cdot, T_0) = W \ \left(\mid k - T_0 \mid^2 + 2\, \delta \mid (k - T_0)\, T_0\, (1 - T_0) \mid \right) \tag{D.13}$$

D.6 Performance and uncertainty

We now derive an explicit formula for the worst case uncertainty performance G, when the uncertainty set is a disk about P_0 and the given performance function $\tilde{G}$ is quasicircular.

We will compare this formula to the modified performance G^* obtained in Subsection ??, and derive an estimate of the error for this type of performance function.

Theorem 26 *Fix a frequency ω and suppose that k, P_0, and C are known complex numbers. Let δ be such that $\delta|P_0| < 1$ and*

$$\mathcal{P}_\omega = \{P_0 + \Delta : |\Delta| < \delta|P_0|\} \tag{D.14}$$

Set

$$\tilde{G} = \left| k - \frac{(P_0 + \Delta)C}{1 + (P_0 + \delta)C)} \right| \tag{D.15}$$

and

$$G(\cdot, P_0) = \max_{P_0 + \Delta \in \mathcal{P}_\omega} \tilde{G}(\cdot, P_0 + \Delta) \tag{D.16}$$

Then

$$G(\cdot, P_0) = \frac{\left| k - T_0 - \delta^2 |T_0|^2 (k-1) \right| + \delta |T_0| |T_0 - 1|}{1 - \delta^2 |T_0|^2} \tag{D.17}$$

where

$$P_0 C = \frac{T_0}{1 - T_0} \tag{D.18}$$

Proof. Set $\Delta = \delta P_0 z$ where z ranges in the unit disk in the complex plane, i.e., $|z| \leq 1$. From this relation and from (D.15), (D.16) and (D.18) we get that

$$G(\cdot, P_0) = \max_{z \in D} |k-1| \left| \frac{\frac{(k-T_0)}{(k-1)\delta T_0} + z}{\frac{1}{\delta T_0} + z} \right| \tag{D.19}$$

The conclusion of the Theorem follows directly from relation (D.19) and the following lemma:

Lemma 6 *If a and b are constants such that $|b| > 1$, then*

$$\max_{|z|\leq 1}\left|\frac{a+z}{b+z}\right| = \frac{|1-a\overline{b}|+|a-b|}{|b|^2-1} \tag{D.20}$$

Proof of Lemma. The linear fractional transformation

$$z \rightarrow \frac{a+z}{b+z}$$

maps the unit disk one-to-one and onto a disk D_1 in the plane Note that the set D_1 is not a half plane since $|b| > 1$. If c_1 and r_1 denote the center and radius of D_1 respectively, then it is clear that the largest modulus possible for a point in D_1 is $|c_1| + r_1$. We now determine c_1 and r_1. Note that for all $z = e^{i\theta}$ we have

$$\left|\frac{a+z}{b+z} - c_1\right| = r_1 \tag{D.21}$$

Some algebra leads to

$$|a - c_1 b|\left|\frac{1+\frac{1-c_1}{a-c_1 b}e^{i\theta}}{b+e^{i\theta}}\right| = r_1, \qquad \forall\theta \tag{D.22}$$

A result from complex analysis says that rational transformations on the circle that have constant modulus must have pole and zeros that come in conjugate pairs. This means that

$$\overline{b} = \frac{1-c}{a-cb} \tag{D.23}$$

Solving for c_1 in (D.23) we obtain

$$c_1 = \frac{1-a\overline{b}}{1-|b|^2}$$

and

$$r_1 = \frac{|a-b|}{|b|^2-1}$$

Recall that the modified performance function G^* is a practical approximation to the actual worst case uncertainty performance G. The following corollary gives an estimate for the difference between the two functions.

Corollary 3 *Suppose that $\delta|T_0| < 1$. Then*

$$\begin{aligned} |G(\cdot, P_0) - G^*(\quad \cdot, P_0)| \leq \\ & 2\delta^2|T_0|^2|k-1|(2|k-T_0| + |1-T_0| + |k-1|) \\ & +2\delta^2(\tfrac{9}{4}|T_0|^2(|k-T_0| + 2|1-T_0|)^2 + +\tfrac{1}{4}|1-T_0|^2) \end{aligned} \tag{D.24}$$

Proof: (To be filled in).

See Figures D.3, D.4 and D.5 for a graphical comparison between the functions G and G^* from equations (D.17) and (D.13).

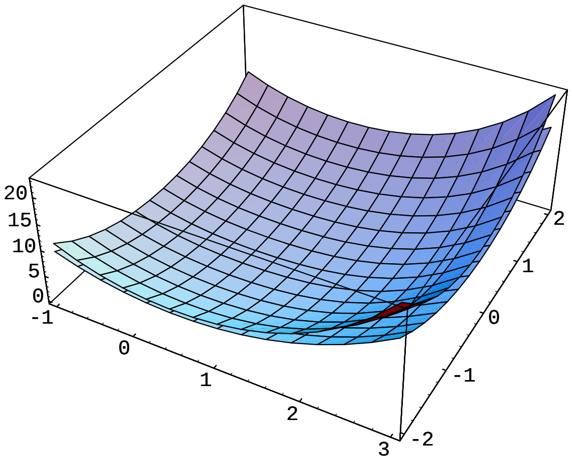

Figure D.3: Plot of the worst case uncertainty performance G from (D.17) and of the modified performance G^* from (D.13) as functions of T_0. Here $k = 0.5$, $\delta = 0.1$, and $P_0 = 1$.

D.7 Extensions

We now obtain a modified performance by using second order terms from a Taylor expansion of the performance. The basic formula we need is the order 2 expansion of $\tilde{G}$:

$$\begin{aligned} \tilde{G}(\cdot, P_0 + \Delta) = \tilde{G}(\cdot, P_0) \quad & +2\mathrm{Re}[a\Delta p] \\ & +A|\Delta|^2 + \mathrm{Re}[B(\Delta)^2] + \mathcal{O}(|\Delta|^3) \end{aligned} \tag{D.25}$$

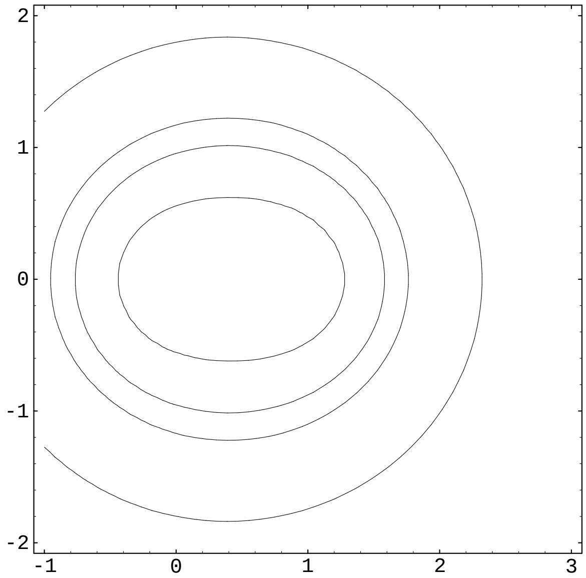

Figure D.4: Plot of the levels 0.01, 0.05, 0.1, 0.5 for the difference $G - G^*$. The functions G and G^* are given by equations (D.17) and (D.13). Here $k = 0.5$, $\delta = 0.1$, and $P_0 = 1$.

where

$$\begin{aligned} a &= \left.\frac{\partial \tilde{G}}{\partial p}\right|_{p=p_0} \\ A &= \left.\frac{\partial^2 \tilde{G}}{\partial \overline{p} \partial p}\right|_{p=p_0} \\ B &= \left.\frac{\partial^2 \tilde{G}}{\partial p^2}\right|_{p=p_0} \end{aligned}$$

From the expression (D.25) we can produce several variations of the method in Subsection ??. To illustrate we carry out the calculations for functions G with the form

$$\tilde{G} = \left| k - \frac{P_0 C}{1 + P_0 C} \right|^2$$

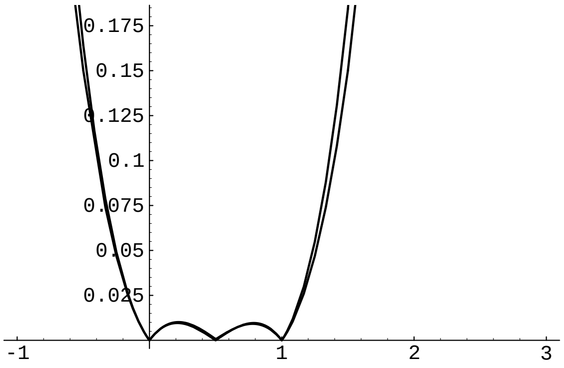

Figure D.5: Plot of G and G^* when T_0 ranges on the x-axis. The functions G and G^* are given by equations (D.17) and (D.13). Here $k = 0.5$, $\delta = 0.1$, and $P_0 = 1$.

For such functions we have

$$\begin{cases} a &= \left(\overline{k - \frac{P_0C}{1+P_0C}}\right) \cdot \frac{C}{1+P_0C} \cdot \frac{1}{1+P_0C} \cdot (-1) = \frac{(\overline{k-T_0})T_0(1-T_0)}{P_0} \\ A &= \frac{|C|^2}{|1+P_0C|^4} = \frac{|T_0|^2|1-T_0|^2}{|P_0|^2} \\ B &= \frac{2c^2}{(1+P_0C)^4} \cdot \left(\overline{k - \frac{P_0C}{1+P_0C}}\right) = \frac{2\ T_0^2(1-T_0)(\overline{k-T_0})}{P_0^2} \end{cases} \tag{D.26}$$

The expansion of $\tilde{G}$ in this case is

$$\begin{aligned} G^*(\cdot, P_0 + \Delta P) = \quad & |k - T_0|^2 + 2\mathrm{Re}\left[\frac{(\overline{k-T_0})T_0(1-T_0)}{p_0}\Delta p\right] + \frac{|T_0|^2|1-T_0|^2}{|P_0|^2}\ |\Delta|^2 \\ & +\mathrm{Re}\left[\frac{2\ T_0^2(1-T_0)(\overline{k-T_0})}{P_0^2}(\Delta)^2\right] + \mathcal{O}(|\Delta|^3) \end{aligned} \tag{D.27}$$

If Δ is chosen as in (D.5) and (D.7), then after some simplification we obtain the following modified performance function (now in terms of T_0 instead of P_0):

$$\begin{aligned} \Gamma^*(\cdot, T_0) = \quad & |k - T_0|^2 + 2|T_0||1 - T_0|\delta \\ & + \left(|1 - T_0|^2 + 2\mathrm{Re}[(1 - \overline{T_0})(k - T_0)]\right)|T_0|^2\delta^2 \end{aligned} \tag{D.28}$$

Note that expression (D.28) has a second order term (in ΔP) that *may be negative* for some T_0's.

A variation of (D.28) is obtained by dropping some of the second order terms in (D.28):

$$\Gamma^*(\cdot, T_0) = |k - T_0|^2 + 2|T_0||1 - T_0|\delta + |1 - T_0|^2|T_0|^2\delta^2 \qquad \text{(D.29)}$$

Appendix E

Getting OPTDesign and Anopt via anonymous ftp

To get the packages OPTDesign and Anopt through anonymous ftp, type:

```
ftp osiris.ucsd.edu
```

When the remote system requests the accountname, you reply

```
anonymous
```

When the system requests the password, type your email address:

```
myadress.edu
```

Then do this:

```
cd pub/helton/OPTDesign+Anopt
```

The OPTDesign.README file tells what is available. To view it, type

```
get OPTDesign.README -
```

To transfer everything in the directory type

```
binary

mget *
```

and hit return key repeatedly. To end your session type

```
bye
```

Appendix F

Anopt.ma Notebook

Appendix G

NewtonInterpolant.ma Notebook

Appendix H

NewtonFit.ma Notebook