Multidisk Problems in H^∞ Optimization: a Method for Analysing Numerical Algorithms *

Harry Dym
The Weizmann Institute
Rehovot 76100, Israel
dym@wisdom.weizmann.ac.il

J. William Helton
Univ. of Calif. San Diego
La Jolla, CA, 92093
jwhelton@ucsd.edu

Orlando Merino
Univ. of Rhode Island
Kingston, RI 02881
merino@math.uri.edu

September 11, 2001

Abstract

The theory of block Toeplitz operators and block Hankel operators is exploited to analyze numerical algorithms for solving optimization problems of the form

$$\gamma^* = \inf_{f \in A_N} \sup_{e^{i\theta}} \|\Gamma(e^{i\theta}, f(e^{i\theta}))\|_{m\times m} ,$$

where $\Gamma(e^{i\theta}, z)$ is a smooth positive semidefinite matrix valued function of $e^{i\theta}, z = (z_1, \ldots, z_n)$ and $\overline{z} = (\overline{z_1}, \ldots, \overline{z_N})$, and A_N is a prescribed set of N-tuples $f = (f_1, \ldots, f_N)$ of functions that are analytic in the open unit disk $\mathbf{D}$ of the complex plane $\mathbf{C}$. The algorithms under consideration are based on writing the equations that an optimum must satisfy (in terms of "primal" and "dual" variables f, γ, Ψ) as $T(\Psi, f, \gamma) = 0$ and then invoking a Newton algorithm (or something similar) to solve these equations. The convergence of Newton's method depends critically upon whether or not the differential T' is invertible. For the class of problems under consideration, this is a very difficult issue to resolve. However, it is relatively easy to determine when T' is a Fredholm operator with Fredholm index equal to zero. Fortunately, it turns out that this weaker condition seems to characterize effective numerical algorithms and is reasonably easy to check. Explicit tests for the differential T' to be a

*Partially supported by the NSF and the ONR

Fredholm operator of index zero. are presented and compared with numerical experiments on a few randomly chosen two and three disk problems. The experimental results lend credence to our contention that whether or not the differential T' is a Fredholm operator of index zero determines the numerical behavior for "almost all" multidisk problems.

Keywords H^∞ Optimization, H^∞ Control, Optimization over Analytic Functions, Semidefinite Programming, Primal-Dual Optimization.

Contents

1 Introduction

1.1 The optimization problem MOPT

A number of optimization problems can be formulated as follows:

> **(MOPT)** Given a nonnegative scalar valued function $\mathbf{\Gamma}$ on $\partial\mathbf{D} \times \mathbf{C}^N$ (that measures performance and will be termed a performance function), find $\gamma^* \geq 0$ and f^* in A_N which solve
>
> $$\gamma^* = \inf_{f \in A_N} \sup_{e^{i\theta}} \mathbf{\Gamma}(e^{i\theta}, f(e^{i\theta})) \,.$$

Here A_N is a prescribed set of N-tuples $(f_1, \ldots, f_N)$ of functions that are analytic in the open unit disk $\mathbf{D}$ of the complex plane $\mathbf{C}$.

In this paper we consider the case where the performance function is of the form

$$\mathbf{\Gamma}(e^{i\theta}, z) = \|\Gamma(e^{i\theta}, z)\|_{m\times m},$$

where Γ is a smooth positive semidefinite (and hence automatically selfadjoint) $m\times m$ matrix valued function and $\|M\|_{m\times m}$ is the largest singular value of the matrix M. This is a mathematically appealing type of performance function that is continuous but typically is not differentiable, since the matrix norm is not differentiable for most matrix valued functions. (Think of $|x|$.) Such performance functions $\mathbf{\Gamma}$ arise in engineering, cf. [HMer:98]. For example, the well known Nehari problem and the H^∞ multidisk problem can be incorporated into this framework, as can many other problems.

This paper develops the mathematics needed to understand and develop numerical algorithms. These algorithms are based on writing the equations that an optimum must satisfy (in terms of "primal" and "dual" variables f, γ, Ψ) as

$$T(\Psi, f, \gamma) = 0$$

and then invoking a Newton algorithm (or something similar) to solve these equations. Such algorithms are called *primal-dual algorithms* and play a major role in the optimization of matrix valued functions. Newton's method involves the differential T' and uses its inverse critically. Indeed the "second order convergence" of Newton's method is completely governed by whether or not this differential is invertible.

In this paper we obtain the differential T' of T and find that it is difficult to determine when it is invertible. However, it is relatively easy to determine when T' is a Fredholm operator with Fredholm index equal to zero. Fortunately, it turns out that this weaker condition seems to characterize effective numerical algorithms (and is reasonably easy to check for multidisk problems). This is perhaps not so surprising if one recalls that Fredholm operators of index zero are compact perturbations of invertible operators.

Recall that the H^∞ multidisk problem is one where we have a special class of Γ that are block diagonal with block diagonal entries equal to the matrix valued performance functions

$$\Gamma^p(e^{i\theta}, f(e^{i\theta})) = (K^p(e^{i\theta}) \; - \; f(e^{i\theta}))^T(K^p(e^{i\theta}) \; - \; f(e^{i\theta})), \tag{1}$$

where $K^p(e^{i\theta}), p = 1, \ldots, v$, and $f(e^{i\theta})$ are $m \times m$ matrix valued functions. In other coordinates **MOPT** with such Γ are often called Integral Quadratic Constraint (IQC) problems. Here we determine explicit tests for the differential T' to be a Fredholm operator of index zero for the multidisk problem and compare the implications of these tests with numerical experiments on a few randomly chosen two and three disk problems.

The experimental results lend credence to our contention that whether or not the differential T' is a Fredholm operator of index zero determines the numerical behavior for "almost all" problems in the class under consideration. Moreover, both our theory and our experiments lead us to assert that the v disk problem in $m \times m$ matrix function space is well behaved for a broad range of Newton type methods when $v = m$, in contrast with other values of v.

Earlier work, cf. [HMer:98] parts III and IV, gave a reasonably complete theory of **MOPT** problems for performance functions $\mathbf{\Gamma}$ which are smooth. Also, [HMW:98], and [HMer:98] part V gave optimality conditions for **MOPT** and some numerical algorithms based on it. However, the cited sources did not analyze these algorithms. That is the task of the present paper.

1.2 Assumptions

We shall assume throughout this paper that $\Gamma(e^{i\theta}, z)$ is a positive semidefinite matrix valued function that is twice continuously differentiable in $z = (z_1, \ldots, z_N)$ and $\overline{z} = (\overline{z_1}, \ldots, \overline{z_N})$, and that

$$\Gamma(e^{i\theta}, z), \quad \frac{\partial \Gamma}{\partial z_\ell}(e^{i\theta}, z), \quad \frac{\partial \Gamma}{\partial \overline{z_\ell}}(e^{i\theta}, z), \quad \frac{\partial^2 \Gamma}{\partial \overline{z_\ell} \partial z_j}(e^{i\theta}, z), \quad \frac{\partial^2 \Gamma}{\partial z_\ell^2}(e^{i\theta}, z), \quad \frac{\partial^2 \Gamma}{\partial \overline{z_\ell}^2}(e^{i\theta}, z)$$

are at least continuously differentiable in θ (for all points $e^{i\theta}$ on the unit circle $\mathbf{T} = \partial \mathbf{D}$).

Some sample problems are sketched below.

1.2.1 Everything is independent of θ

An important special case of the **MOPT** problem that has received considerable attention, cf. the survey [VB:96], is the case where Γ and f do not depend on θ. Then f is an N- tuple of complex numbers or a $2N$-tuple of real numbers and Γ is a matrix valued function of f alone. The optimization problem **MOPT** becomes

$$\gamma^* = \inf_{f \in \mathbf{C}^N} \; \|\Gamma(f)\|_{m \times m} \, .$$

1.2.2 The matrix Nehari problem

To illustrate **MOPT** we recall the classical Nehari problem: *Given K, a bounded $m \times m$ matrix valued function on the unit circle, find its distance to the Hardy space of bounded matrix valued analytic functions $H^\infty_{m\times m}$.* That is, with some poetic license[1], find

$$\operatorname{dist}(K, H^\infty_{m\times m}) := \inf_{f\in H^\infty_{m\times m}} \sup_\theta \| \, K(e^{i\theta}) \; - \; f(e^{i\theta}) \, \|_{m\times m}, \tag{2}$$

or, equivalently,

$$\operatorname{dist}(K, H^\infty_{m\times m})^2 = \inf_{f\in H^\infty_{m\times m}} \sup_\theta \| \, K(e^{i\theta}) \; - \; f(e^{i\theta}) \, \|^2_{m\times m}. \tag{3}$$

Since $\|B^T B\|_{m\times m} = \|B\|^2_{m\times m}$ for any $m \times m$ matrix B and its conjugate transpose B^T, we may rewrite (3) as

$$\operatorname{dist}(K, H^\infty_{m\times m})^2 = \inf_{f\in H^\infty_{m\times m}} \sup_\theta \| \, (K(e^{i\theta}) \; - \; f(e^{i\theta}))^T (K(e^{i\theta}) \; - \; f(e^{i\theta})) \, \|_{m\times m}. \tag{4}$$

To put the Nehari Problem in **MOPT** notation, take

$$\Gamma(e^{i\theta}, Z) = (K(e^{i\theta}) \; - \; Z)^T (K(e^{i\theta}) \; - \; Z) \;, \tag{5}$$

where $Z = (z_{ij})^m_{i,j=1}$ denotes a matrix with $N = m^2$ independent entries. It is clear that $\Gamma(e^{i\theta}, Z)$ is analytic in z_{ij} and $\overline{z_{ij}}$, $i, j = 1, \ldots, m$, and continuous in θ if K is continuous in θ, and that in this case **MOPT** gives an optimal value γ such that

$$\gamma = \operatorname{dist}(K, H^\infty_{m\times m})^2. \tag{6}$$

Also, in view of the convexity of the $L^\infty_{m\times m}$ norm, a local solution is a global solution too. Hence solutions to **MOPT** correspond to solutions to the Nehari problem: the minimizers $f \in H^\infty_{m\times m}$ are the same, while the optimal values are related by equation (6).

We now use the Nehari problem to illustrate one important aspect of **MOPT** problems, namely, that solutions are often not unique when $m > 1$. To illustrate this we consider the function

$$K(e^{i\theta}) = \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & 0 \end{pmatrix} \;. \tag{7}$$

One can easily check with the help of Theorem 2 of [HMW:98] (which also appears as Theorem 17.1.1 of [HMer:98]) that the function zero gives the optimal distance from K to $H^\infty_{m\times m}$, which turns out to be 1. If $f \in H^\infty$ satisfies $|f(e^{i\theta})| \le 1$ for all θ, then

$$\sup_\theta \left\| K(e^{i\theta}) - \begin{pmatrix} 0 & 0 \\ 0 & f(e^{i\theta}) \end{pmatrix} \right\|_{2\times 2} = 1, \tag{8}$$

so any such f gives rise to a solution to **MOPT**.

[1] We are using *supremum* instead of *essential supremum*.

1.2.3 A two disk problem

The two disk problem is a natural generalization of the Nehari problem (which we look at as a one disk problem) to "two disks". *Given a pair K^1 and K^2 of continuous $m \times m$ matrix valued functions on the unit circle (which we think of as the centers of matrix function disks) and two performance functions*

$$\Gamma^p(e^{i\theta}, Z) = (K^p(e^{i\theta}) \ - \ Z)^T (K^p(e^{i\theta}) \ - \ Z) \ , \quad p = 1, 2, \tag{9}$$

the two disk problem is to find the smallest number γ and a function f in the space of bounded analytic functions $H^\infty_{m\times m}$, so that

$$\Gamma^1(e^{i\theta}, f(e^{i\theta})) \leq \gamma I_m \quad \text{and} \quad \Gamma^2(e^{i\theta}, f(e^{i\theta})) \leq \gamma I_m. \tag{10}$$

In formula (9), Z denotes a matrix $Z = (z_{ij})_{i,j=1}^m$ and $N = m^2$. Note that (10) holds if and only if f simultaneously lies inside both of the matrix function disks. This is the **MOPT** problem for the performance function

$$\Gamma := \begin{pmatrix} \Gamma^1 & 0 \\ 0 & \Gamma^2 \end{pmatrix}.$$

1.2.4 The multidisk problem

The analysis in the preceding subsection is easily extended to v disks,

$$\Gamma^p(e^{i\theta}, f(e^{i\theta})) = (K^p(e^{i\theta}) \ - \ f(e^{i\theta}))^T (K^p(e^{i\theta}) \ - \ f(e^{i\theta})), \qquad p = 1, \ldots v. \tag{11}$$

Here K^p and f are $m \times m$ matrix valued functions and we seek the smallest γ satisfying

$$\Gamma^p(e^{i\theta}, f(e^{i\theta})) \leq \gamma I_m$$

for $p = 1, \ldots, v$ and all θ.

The *multidisk problem* is the **MOPT** problem for the performance function

$$\Gamma := \operatorname{diag}(\Gamma^1, \ldots, \Gamma^v) = \begin{pmatrix} \Gamma^1 & 0 & \ldots & 0 \\ 0 & \Gamma^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \Gamma^v \end{pmatrix}, \tag{12}$$

where Γ^p, $p = 1, \ldots, v$, is given by (11).

1.2.5 Notation and usage

- The acronym mvf denotes matrix valued function. If B is a matrix (or a mvf), then B^T stands for the conjugate transpose of B, B^τ stands for the plain transpose of B and, if B is square, then $\operatorname{tr} B = \text{trace } B$.

- $\Gamma(e^{i\theta}, z) = \Gamma(e^{i\theta}, z_1, \ldots, z_N)$ is a positive semidefinite (and hence automatically selfadjoint) smooth $m \times m$ mvf on $\mathbf{T} \times \mathbf{C}^N$. Thus m and N are positive integers that are determined by Γ.

- $$\frac{\partial \Gamma}{\partial z}(\cdot, f) = \left[\frac{\partial \Gamma}{\partial z_1}(\cdot, f), \ldots, \frac{\partial \Gamma}{\partial z_N}(\cdot, f)\right]$$
  is an $m \times Nm$ mvf with N blocks of size $m \times m$ in a row. Multiplication on the right by an $m \times m$ matrix (or mvf) B, and the application of the trace, should be understood as performed on each block entry. Thus, for example,
  $$\operatorname{tr}\left\{\frac{\partial \Gamma}{\partial z}(\cdot, f)B\right\} = \left[\operatorname{tr}\left\{\frac{\partial \Gamma}{\partial z_1}(\cdot, f)B\right\}, \ldots, \operatorname{tr}\left\{\frac{\partial \Gamma}{\partial z_N}(\cdot, f)B\right\}\right]$$
  is a $1 \times N$ mvf.

- $(\frac{\partial \Gamma}{\partial z})^T$ is a block column mvf with N blocks of size $m \times m$: $\left(\frac{\partial \Gamma}{\partial z_j}\right)^T$, $j = 1, \ldots, N$. Since Γ is selfadjoint,
  $$\left(\frac{\partial \Gamma}{\partial z_j}\right)^T = \frac{\partial \Gamma}{\partial \overline{z}_j}, \qquad j = 1, \ldots, N.$$

- The abbreviations
  $$a := \frac{\partial \Gamma}{\partial z}(\cdot, f) = [a_1, a_2, \ldots, a_N] \quad \text{and} \quad a\varphi = \sum_{j=1}^N a_j \varphi_j$$
  will be used when it is clear from the context what f is and when $\varphi \in H_N^\infty$.

- $\Gamma(e^{(i\theta)}, \cdot)$ is said to be plurisubharmonic (PLUSH) if the $Nm \times Nm$ matrix
  $$\Gamma_{\overline{z}z}(e^{i\theta}, \cdot) = \left[\frac{\partial^2 \Gamma}{\partial \overline{z}_i \partial z_j}(e^{i\theta}, \cdot)\right]_{i,j=1}^N$$
  is positive semidefinite on $\mathbf{C}^{\mathbf{Nm}}$. Γ is said to be strictly PLUSH if $\Gamma_{z\overline{z}}$ is strictly positive definite on $\mathbf{C}^{Nm}$. If $m = 1$, then it is often the case that Γ is strictly PLUSH, while if $m > 1$ it is more reasonable to expect that only the weaker condition PLUSH holds.

- $H^2_{r \times n}$ stands for the Hardy space of $r \times n$ mvf's with entries in the (scalar) Hardy space H^2 for the open unit disc with inner product
  $$\langle F, \ G \rangle = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{tr}\left\{G(e^{i\theta})^T F(e^{i\theta})\right\} d\theta.$$

- $(H^2_{k \times m})_+ = \{F \in H^2_{k \times m}$: the first $k \times k$ block of the constant term $f(0)$ is upper triangular (not strictly) $\}$ when $k \leq m$.

- $H^2_+ = (H^2_{k \times m})_+$

1.2.6 Optimality equations for MOPT

The optimality conditions that are the source of our computer algorithms are obtained by associating a mixture of the primal problem **MOPT** and a dual problem which we now state:

$$\textbf{(PDMOPT)} \quad \text{Find} \qquad \gamma := \min_f \max_\Psi \frac{1}{2\pi}\int_0^{2\pi} \operatorname{tr}\left\{\Gamma(e^{i\theta}, f(e^{i\theta}))\Psi(e^{i\theta})\right\} d\theta$$

$$\text{subject to:} \quad \Psi \geq 0, \ \Psi \in L^1_{m\times m}, \quad \int_0^{2\pi} \operatorname{tr}\Psi(e^{i\theta})d\theta = 2\pi, \quad f \in H^\infty_N \ .$$

In this paper we shall always assume that the indicated minimum and maximum exist.

Solutions f, Ψ to the *primal -dual problem* produce f which are solutions to **MOPT**, providing that f is continuous. Solving the first order optimality conditions for the primal-dual problem gives excellent candidates for local solutions to **MOPT**. Earlier results on optimality conditions assume that for each θ the columns of the $N \times m^2$ mvf

$$\mathcal{U}_f(e^{i\theta}) = \begin{pmatrix} \frac{\partial\Gamma_{1,1}}{\partial z_1}(\cdot, f) & \cdots & \frac{\partial\Gamma_{\ell,k}}{\partial z_1}(\cdot, f) & \cdots & \frac{\partial\Gamma_{m,m}}{\partial z_1}(\cdot, f) \\ \vdots & & & & \vdots \\ \frac{\partial\Gamma_{1,1}}{\partial z_N}(\cdot, f) & \cdots & \frac{\partial\Gamma_{\ell,k}}{\partial z_N}(\cdot, f) & \cdots & \frac{\partial\Gamma_{m,m}}{\partial z_N}(\cdot, f) \end{pmatrix} \tag{13}$$

are linearly independent; see e.g., Theorem 2 in [HMW:98] and Theorem 17.1.1 in [HMer:98]. But this is only possible if

$$N \ \geq \ \text{the number of entries in the mvf } \Gamma \ , \tag{14}$$

which is not the case for the multidisk problem when $v > 1$, since then $N = m^2$ and Γ is a $vm \times vm$ mvf. Fortunately, it turns out that we can get by with less stringent assumptions that are formulated in terms of the mvf

$$\mathcal{V}_f(e^{i\theta}, B) \ = \ \begin{bmatrix} \operatorname{tr}\left\{\frac{\partial\Gamma}{\partial z_1}(e^{i\theta}, f(e^{i\theta}))B\right\} \\ \vdots \\ \operatorname{tr}\left\{\frac{\partial\Gamma}{\partial z_N}(e^{i\theta}, f(e^{i\theta}))B\right\} \\ \operatorname{tr}\left\{\left[\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta}))\right] B\right\} \end{bmatrix} . \tag{15}$$

Here Γ is an $m \times m$ mvf and B is an $m \times m$ matrix. For each θ, $\mathcal{V}_f(e^{i\theta}, B)$ defines a linear map from $\mathbf{C}^{m\times m}$ into $\mathbf{C}^{N+1}$. The basic condition that we want is that the null space of this map be zero, at least when restricted to appropriate classes of matrices B (or $m \times m$ matrix valued measures $d\mu$). In order to compare this condition with the condition imposed on $\mathcal{U}_f$ that was discussed earlier, it is useful to note that the first N entries of $\mathcal{V}_f(e^{i\theta}, B)$ can be reexpressed as

$$\begin{bmatrix} \mathrm{tr}\{a_1 B\} \\ \vdots \\ \mathrm{tr}\{a_N B\} \end{bmatrix} = \mathcal{U}_f(e^{i\theta}) \ \mathrm{vec}(B) \ , \tag{16}$$

where $\mathrm{vec}(B)$ is the $m^2 \times 1$ vector that is obtained by stacking the successive columns of B on top of each other. Thus,

$$\text{null space}\{\mathcal{U}_f(e^{i\theta})\} = 0 \Longrightarrow \text{null space}\{\mathcal{V}_f(e^{i\theta}, B)\} = 0 \ , \tag{17}$$

but not vice versa.

If f is a continuous function on the circle and $\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta})) \geq 0$, then we shall say that the triple $(\gamma, f(e^{i\theta}), \Gamma(e^{i\theta}, f(e^{i\theta}))$ is **measure nondegenerate** if for every $m \times m$ matrix valued measure $d\mu$ on the circle $\mathbf{T}$

$$\mathcal{V}_f(e^{i\theta}, d\mu) = 0 \text{ and } d\mu \geq 0 \Longrightarrow d\mu = 0 \ . \tag{18}$$

Here too, the condition null space$\{\mathcal{U}_f(e^{i\theta})\} = 0$ automatically implies the validity of (18), but not vice versa.

Theorem 1.1 *Let $\Gamma(e^{i\theta}, z)$ be a positive semidefinite mvf that satisfies the smoothness conditions specified in Section 1.2, and assume that the* **PDMOPT** *problem has a solution $(\Psi, f, \gamma) \in L^1_{m\times m} \times H^\infty_N \times \mathbf{R}$ such that:*

(1) *$\Psi \in L^2_{m\times m}$ and is positive semidefinite for a.e. point $e^{i\theta} \in \mathbf{T}$.*

(2) *f is continuous and $\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta})) \geq 0$ a.e..*

(3) *The triple (γ, f, Γ) is measure multidisk nondegenerate (i.e., (18) holds).*

Then Ψ, f and γ must satisfy the following conditions:

$$\begin{array}{lrcl} \text{(a)} & \Psi(\gamma I_m - \Gamma(\cdot, f)) & = & 0 \ \ a.e.. \\ \text{(b)} & P_{H^2_N}[\mathrm{tr}[\frac{\partial \Gamma}{\partial z}^T(\cdot, f)\Psi]] & = & 0 \ . \\ \text{(c)} & \frac{1}{2\pi}\int \mathrm{tr}\{\Psi\} d\theta - 1 & = & 0 \ . \end{array} \tag{19}$$

The proof of this theorem is given in Section 6.1.

1.2.7 Complementarity and strict complementarity

Condition (a) in Theorem 1.1 holds if and only if the function Ψ satisfies

$$\mathrm{range}\Psi \subset \mathrm{null}\{\gamma I - \Gamma(\cdot, f)\}, \tag{20}$$

for a.e. θ, or equivalently, if and only if

$$\mathrm{range}\Psi \perp \mathrm{range}\{\gamma I - \Gamma(\cdot, f)\} \tag{21}$$

for a.e. θ. Because of this, condition (a) is often called a **complementarity condition**. A stronger condition on a solution (Ψ, f, γ) to the primal–dual problem (**PDMOPT**) is the **strict complementarity condition**. It states that, in addition to (21), the two range spaces are orthogonal complements. In other words their sum spans $\mathbf{C}^m$. This forces equality in (20).

Strict complementarity is an extremely important condition here as well as in many areas of optimization. It is the main ingredient in the assumption **(SCOM)** which is formulated below in Section 2.2 and is invoked in the main theorems of this paper.

Remark. Assumption (3) of Theorem 1.1 requiring measure non-degeneracy may not be necessary. We do not know an example which insures that it is needed.

1.2.8 Factoring the dual variable

Throughout the rest of the paper we assume that the dual variable
$\Psi = G^T G$ **has an outer spectral factor** $G \in (H^2_{k\times m})_+$ **with rank** $G = k$ **a.e..**
We remark that if $\Psi = G^T G$, where $G \in H^2_{k\times m}$, then there is no loss of generality in assuming that $G \in (H^2_{k\times m})_+$. Indeed, one may choose a (constant) unitary matrix U so that $G_1 = UG$ is in H^2_+; this uses the QR factorization of matrices. Note that $G_1^T G_1 = G^T G$.
We now rexpress the optimality condition (a) of Theorem 1.1 in terms of G.

Lemma 1.2 *If $G \in H^2_+$ and if $\gamma I - \Gamma(\cdot, f) \geq 0$, then the following statements are equivalent:*

(1) $(\gamma I - \Gamma(\cdot, f))G^T G = 0$ *a.e..*

(2) $P_{H^2_+}[G(\gamma I - \Gamma(\cdot, f))] = 0$.

(3) $G(\gamma I - \Gamma(\cdot, f)) = 0$ *a.e..*

The proof appears in Section 6.2 .

1.3 Numerical Algorithms

Many numerical algorithms can be based on solving the optimality equations (a)—(c) in Theorem 1.1 or a variation of them. We shall list two algorithms in this paper and analyze the second of them. We hope that the method we use for this analysis can be readily adapted to most natural algorithms, although this remains to be seen.

The main idea is that solving equations (a)—(c) in Theorem 1.1 for G, f, and γ, is equivalent to solving an operator equation of the form

$$T \begin{pmatrix} G \\ f \\ \gamma \end{pmatrix} = 0 \tag{22}$$

for the same unknowns. It is common to approach such problems using Newton's method or some variation thereof.

1.3.1 Newton's method

Newton's method is an iterative scheme for solving operator equations of the form

$$T[x] = 0. \tag{23}$$

In terms of the differential T' of T, the Newton step for updating a given x is:

$$\tilde{x} = x - (T'_x)^{-1} T[x] \ , \tag{24}$$

whenever the differential is invertible. We refer to (24) as the *Newton iteration*, and to the repeated application of (24) as the *Newton algorithm*.

A standard fact [K:52] about Newton iteration is

Theorem 1.3 *Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be normed linear spaces, and suppose the operator $T : \mathcal{B}_1 \to \mathcal{B}_2$ is two times continuously Frechet differentiable in a neighborhood V of $x^* \in \mathcal{B}_1$. Assume also that $T[x^*] = 0$, and that T'_{x^*} has a bounded inverse $(T'_{x^*})^{-1}$. Then there exist a neighborhood W of x^* in $\mathcal{B}_1$ and a constant $c > 0$ such that for every $x \in W$ the linear operator T'_x has a bounded inverse, and $\tilde{x} := x - (T'_x)^{-1} T[x]$ satisfies*

$$||\, \tilde{x} - x^* \,||_{\mathcal{B}_1} \leq c \,||\, x - x^* \,||^2_{\mathcal{B}_1}.$$

In the present setting, the differential

$$T'_{(G,f,\gamma)} : \begin{pmatrix} H^2_+ \\ H^2_N \\ \mathbf{R} \end{pmatrix} \to \begin{pmatrix} H^2_+ \\ H^2_N \\ \mathbf{R} \end{pmatrix}$$

is the key term in the linearization

$$0 = T(G, f, \gamma) + T'_{(G,f,\gamma)}[(\Delta, \varphi, \eta)] \tag{25}$$

of (22) that is used to solve for (Δ, φ, η) at each iteration. Variations are to modify the operator T in some way which has better numerical properties, but which still allows one to construct a primal–dual solution from the answer. If $T'_{(G,f,\gamma)}$ is invertible at the optimum (G, f, γ), then Newton's method has local second order convergence (which is very good).

Interior point methods often mix following "the central path" with Newton steps. The theory of interior point methods [Wr:98] [AHO:96] tells us that one can expect local second order convergence to a solution (G, f, γ) providing that $T'_{(G,f,\gamma)}$ is invertible at this solution. Indeed invertibility of $T'_{(G,f,\gamma)}$ is sufficient but not necessary for good local behavior (even without following the central path). Our concern in this article is the theoretical analysis of such invertibility, since it is often extremely informative.

1.3.2 Algorithms for solving MOPT

The Algorithm G^TG

The first algorithm we present uses the operator

$$T : \begin{pmatrix} G \\ f \\ \gamma \end{pmatrix} \to \begin{pmatrix} P_{H^2_+}[G(\gamma I - \Gamma(\cdot, f))] \\ \\ P_{H^2_N}[\mathrm{tr}[-\frac{\partial \Gamma}{\partial z}(\cdot, f)^T G^T G]] \\ \\ \frac{1}{2\pi}\int \mathrm{tr}\{G^T G\} d\theta - 1 \end{pmatrix}. \tag{26}$$

Algorithm $\boldsymbol{G^TG}$. Given $x^{(r)} = (G^{(r)}, f^{(r)}, \gamma^{(r)})$, to update to $x^{(r+1)}$ carry out steps I – III below:

I. **Subproblem.** Solve $T(x) + T'_x(\delta x) = 0$ for $\delta x = (\delta G, \delta f, \delta \gamma)$.

II. **Line search.** Perform a linear search to determine a step parameter $t \geq 0$ that minimizes $\|T(x + t\delta x)\|$.

III. **Update.** Set $x^{(r+1)} = (G^{(r+1)}, f^{(r+1)}, \gamma^{(r+1)})$, where $f^{(k+1)} = f^{(k)} + t\delta f$, $\gamma^{(k+1)} = \gamma^{(k)} + t\delta\gamma$, and $G^{(k+1)} = G^{(k)} + t\delta G$. If $\gamma I - \Gamma(\cdot, f) \geq 0$ is not satisfied, $\gamma^{(k+1)}$ is chosen as $\sup_\theta \Gamma(e^{i\theta}, f^{(k+1)}(e^{i\theta}))$.

The Algorithm $G + G^T$

The second algorithm we present uses the additive decomposition $\Psi = G + G^T$. In this setting G is always square, that is, $k = m$, and the operator

$$T : \begin{pmatrix} G \\ f \\ \gamma \end{pmatrix} = \begin{pmatrix} P_{H^2_+}[(G + G^T)(\gamma I - \Gamma(\cdot, f)) + (\gamma I - \Gamma(\cdot, f))(G + G^T)] \\ \\ P_{H^2_N}[\mathrm{tr}[-\frac{\partial \Gamma}{\partial z}(\cdot, f)^T (G + G^T)]] \\ \\ \frac{1}{2\pi}\int \mathrm{tr}\{G + G^T\} d\theta - 1 \end{pmatrix}. \tag{27}$$

Algorithm $G + G^T$. Given $x^{(r)} = (G^{(r)}, f^{(r)}, \gamma^{(r)})$, to update to $x^{(r+1)}$ carry out steps I – III below:

I. **Subproblem.** Solve $T(x) + T'_x(\delta x) = 0$ for $\delta x = (\delta G, \delta f, \delta\gamma)$.

II. **Line search.** Perform a linear search to determine a step parameter $t \geq 0$ that minimizes $\sup_\theta \Gamma(e^{i\theta}, f^{(k)}(e^{i\theta})) + t\delta f(e^{i\theta}))$.

III. **Update.** Set $x^{(r+1)} = (G^{(r+1)}, f^{(r+1)}, \gamma^{(r+1)})$, where $f^{(k+1)} = f^{(k)} + t\delta f$, $\gamma^{(k+1)} = \sup_\theta \Gamma(e^{i\theta}, f^{(k+1)}(e^{i\theta}))$, and, providing that $G + G^T \geq 0$ is satisfied, $G^{(k+1)} = G^{(k)} + t\delta G$. If $G + G^T \geq 0$ does not hold for this choice of t, then a smaller step is selected for G to ensure positivity.

This gives a function space analog of Haeberly and Overton's $XZ + ZX$ algorithm [HO:94], [AHO:96] given by those authors for finite dimensional matrix optimization.

In this paper we describe a theory for analyzing algorithms of the type given above. We apply the theory to Algorithm G^TG since it is the first one we looked at. A theoretical study of Algorithm $G + G^T$ is in progress. Numerical experiments (see Section 1.5) strongly suggest that Algorithm $G + G^T$ has a considerably broader range of effectiveness than Algorithm G^TG.

1.3.3 The main problem in analyzing algorithms

A wide range of mathematical questions arise in the analysis of numerical algorithms. However, our philosophy in this paper is different from the traditional one which is commonplace in mathematics, wherein one wants as many theorems as possible about the situation. By contrast, when dealing with numerical algorithms, if (as is usually the case) the goal is to have a theory of a class of algorithms which helps to develop new algorithms, then (although having many theorems is satisfying) what is really important is knowing the smallest set of theorems which correlate strongly with the performance of algorithms. The point is that when presented with a new candidate for an algorithm we want a few simple calculations that can be carried out quickly and serve to predict whether the algorithm will be successful or not. A major objective is to avoid many fruitless computer experiments. This puts a premium on identifying simple properties which are critical in practice and identifying other properties which can be ignored. As indicated above, a question that is central to analyzing the performance of optimization algorithms is:

When is $T'_{(G,f,\gamma)}$ invertible?

This is a very difficult question to answer in the settings under consideration. Instead, we shall focus on the following question:

When is $T'_{(G,f,\gamma)}$ a Fredholm operator with Fredholm index equal to zero?

An affirmative answer seems to be a good indicator of an effective algorithm. Fredholm index zero is a weaker notion than invertibility, but we found that in our setting it is easy to check, and moreover, we believe that it gives the information we need to predict and evaluate the performance of a substantial class of algorithms. (See the conjecture in Section 1.4.3.)

In Section 1.5 we compare the implications of our tests with numerical experiments on a few randomly chosen two and three disk problems. The experimental results lend credence to our contention that *whether or not T' has Fredholm index zero determines numerical behavior for "almost all" multidisk problems.*

Definition 1.4 *Let X, Y be Banach spaces, let $A : X \to Y$ be a bounded linear operator from X into Y and let $\mathcal{B}(X,Y)$ denote the space of all such operators. Then A is a Fredholm operator if:*
(i) $n(A) := \dim \text{null } A < \infty$.
(ii) range A *is closed in* Y.
(iii) $r(A) := \dim\ (Y \ \setminus \ \text{range } A) < \infty$.
If A is a Fredholm operator, the number

$$\iota(A) := n(A) - r(A)$$

is called the Fredholm index of A. Let $\mathcal{F}_0$ denote the set of Fredholm operators of index 0.

If A and B are Hilbert spaces, then condition (iii) may be written as $\dim\ (\text{range}A)^\perp < \infty$. It is well known that compact perturbations of Fredholm operators are Fredholm. It is clear from the definition that when $X = Y$, the bounded invertible operators are Fredholm operators with index equal to zero. Also, a self adjoint bounded linear operator on a Hilbert space that is Fredholm clearly has index zero.

Another issue that is completely suppressed is a detailed analysis of the various choices of function spaces on which T and consequently T' can act. The reason for ignoring this can be seen in the light of well known theorems to the effect that the non zero spectrum of a Toeplitz operators generated by a fixed smooth symbol is reasonably independent of the choice of the space on which the operator acts; see e.g., [GoF74]. Also empirically we find that the behavior in the L^2 norm that is analysed here corresponds to the behavior observed experimentally. Further justification for not doing analysis on complicated spaces stems from [HMW93], which studied the (**OPT**) problem rather then the more complicated (**MOPT**) problem. That paper analysed a special case of the algorithm studied here, but took the underlying space to be C^∞. The analysis showed that the key Toeplitz operator there is invertible on L^2 if and only if it is invertible on the space C^∞ equipped with a Frechet norm. Thus, our advice to anyone interested in the development of algorithms is to focus their effort on how T' behaves on L^2.

1.4 T and T' for the multidisk problem

The main theorem of this paper, Theorem 2.3, is formulated for general **PDMOPT** problems. Instead of stating the fully general result in the introduction, we consider the special case of the Algorithm G^TG, applied to the H^∞ multidisk problem; see Theorem 1.8 . The first step is to specialize Theorem 1.1 to the multidisk case. Fortunately, the dual variable Ψ in the **PDMOPT** optimality equations for the multidisk problem can be taken block diagonal: $\Psi = \text{diag}(\Psi^1, \ldots, \Psi^v)$. We shall assume that each of these diagonal blocks is factorable. The main result is formulated in the next subsection.

1.4.1 Optimality conditions for the multidisk problem

We first observe that the measure non-degenerate condition (18) when specialized to the multidisk problem becomes the multidisk measure non-degenerate condition that is defined for $\gamma, f, K^p, p = 1, \ldots, v$, as follows. If $d\mu^p$, $p = 1, \ldots, v$, is a set of positive semidefinite $m \times m$ matrix valued measures on the unit circle such that

$$\sum_{p=1}^{v} \{K^p(e^{i\theta}) - f(e^{i\theta})\} d\mu^p = 0$$

and

$$\{\gamma I_m - \Gamma^p(e^{i\theta}, f(e^{i\theta}))\} d\mu^p = 0 \; for \; p = 1, \ldots, v \; ,$$

then

$$d\mu^p = 0 \; for \; p = 1, \ldots, v \; .$$

Theorem 1.5 *Let the functions $K^p(e^{i\theta})$, $p = 1, \ldots, v$, that appear in (11) be continuously differentiable, and assume that the* **PDMOPT** *problem for the H^∞ multidisk problem has a solution $(\Psi, f, \gamma) \in L^1_{vm \times vm} \times H^\infty_{m^2} \times \mathbf{R}$ such that:*

(1)
$$\Psi = \text{diag}\left((G^1)^T G^1, \ldots, (G^v)^T G^v\right),$$
where $G^p \in (H^2_{k_p \times m})_+$ is outer with rank $G^p = k_p$ a.e. for $p = 1, \ldots, v$.

(2) $\Psi \in L^2_{vm \times vm}$.

(3) *f is continuous and $\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta})) \geq 0$.*

(4) *γ, f, K^p, $p = 1, \ldots, v$, is multidisk measure non-degenerate.*

Then G^p, f and γ must satisfy the following conditions:

$$\begin{array}{lrcl} \text{(a)} & P_{H^2_+}[G^p(\gamma I - \Gamma^p(\cdot, f))] & = & 0 \quad for \; p = 1, \cdots, v. \\ \text{(b)} & \beta := \sum_{p=1}^{v} {G^p}^T G^p (K^p - f)^T & \in & e^{i\theta} H^2_{m \times m} \\ \text{(c)} & \sum_{p=1}^{v} \frac{1}{2\pi} \int \text{tr}\{(G^p)^T G^p\} d\theta - 1 & = & 0. \end{array} \qquad (28)$$

This theorem is an easy consequence of Theorem 1.1 and Lemma 1.2. The proof is given in Section 5.2. Similar conclusions are available in [HMer:98] and [HMW:98] under the more restrictive condition that the columns of $\mathcal{U}_f(e^{i\theta})$ are linearly independent. See also [OZ:93] for a related (though more abstract) result for the two disk problem.

The operator T for the Algorithm G^TG in the multidisk setting is

$$T : \begin{pmatrix} G \\ f \\ \gamma \end{pmatrix} \rightarrow \begin{pmatrix} \text{diagonal}\left\{ P_{H^2_+}[G^1(\gamma I - \Gamma^1(\cdot, f))], \ldots, P_{H^2_+}[G^v(\gamma I - \Gamma^v(\cdot, f))] \right\} \\ \\ P_{H^2_{m\times m}}[-\sum_{p=1}^{v}(K^p - f){G^p}^T G^p] \\ \\ \frac{1}{2\pi}\int \text{tr}\{G^TG\}d\theta - 1 \end{pmatrix}. \tag{29}$$

1.4.2 The Fredholmness of T'

Our goal is to determine when T' for the H^∞ multidisk problem is Fredholm of index 0. As we shall see the most valuable indicator of this is the number $\mu_{pnull}(\theta)$ that is given by Definition 1.6 below. We shall impose the following extra condition in addition to those that are imposed in the formulation of Theorem 1.5.

> *The outer factors G^p of Ψ^p are continuous with constant rank k_p for $p = 1, \ldots, v$.*

Definition 1.6 *For each θ, let $\mu_{pnull}(\theta)$ denote the dimension of the following vector subspace of $\mathbf{C}^{m\times m}$:*

$$\{B \in \mathbf{C}^{m\times m} : \Psi^p(e^{i\theta})B^T = 0 \quad and \quad \Psi^p(e^{i\theta})(K^p(e^{i\theta}) - f(e^{i\theta}))^T B = 0 \; for \; p = 1, \ldots, v\}.$$

Definition 1.7 *For each θ, let $\mu_{dnull}(\theta)$ denote the dimension of the following space of matrix tuples:*

$$\{\{\, B^1, \ldots, \, B^v\} \in \mathbf{C}^{k_1\times k_1} \times \cdots \times \mathbf{C}^{k_v\times k_v} : \sum_{p=1}^{v}(K^p(e^{i\theta}) - f(e^{i\theta}))G^p(e^{i\theta})^T B^p G^p(e^{i\theta}) = 0\}.$$

Note that if the functions K^p, f and G^p are given numerically, then the numbers $\mu_{pnull}(\theta)$ and $\mu_{dnull}(\theta)$ *can be computed numerically at each θ.* The notation $\mu_{pnull}(\theta)$ and $\mu_{dnull}(\theta)$ stems from the connection with the null space of σ_{primal} and σ_{dual} in the multidisk case; see Sections 5.2.1 and 5.2.2, respectively.

1.4.3 Conclusions for multidisk problems

Theorem 1.8 *Let the functions $K^p(e^{i\theta})$, $p = 1, \ldots, v$, that appear in (11) be continuously differentiable, and assume that the* **PDMOPT** *problem for the multidisk problem has a solution $(\Psi, f, \gamma) \in L^1_{vm\times vm} \times H^\infty_{m\times m} \times \mathbf{R}$ such that:*

(1) $\Psi = \mathrm{diag}\,(\Psi^1, \ldots, \Psi^v)$ *is block diagonal and continuously differentiable.*

(2) $\Psi^p = (G^p)^T G^p$ *has a continuous outer spectral factor $G^p \in (H^2_{k_p\times m})_+$ with* rank $G^p = k_p$ *for $p = 1, \ldots, v$ and all θ.*

(3) *f is continuously differentiable and $\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta})) \geq 0$.*

(4) *γ, f, K^p, $p = 1, \ldots, v$, is multidisk measure non-degenerate.*

(5) *Strict complementarity holds at (Ψ, f, γ).*

(6) *$\mu_{pnull}(\theta) = 0$ for all θ.*

(7) *$\mu_{dnull}(\theta) = 0$ for all θ.*

Then the operator T' is Fredholm with index zero. If $\mu_{pnull}(\theta) \neq 0$ or $\mu_{dnull}(\theta) \neq 0$ for some θ, i.e., if (6) or (7) *fails, then T' is not Fredholm.*

Remark. Measure non-degeneracy looks similar to $\mu_{dnull}(\theta) = 0$, enough so that Assumption (4) of Theorem 1.8 may eventually be subsumed by Assumption (7). This has not been fully confirmed yet; see Section 2.4.3 for some preliminary discussion. At this point, however, we do know, from the substantial number of computer experiments we have run, that in practice Assumption (4) plays no role. On the other hand, whether or not $\mu_{pnull}(\theta) = 0$ has a big effect.

The proof of Theorem 1.8 is given in Section 5.3. The extra smoothness assumptions that are imposed on $\Psi(e^{i\theta})$ and $f(e^{i\theta})$ in the formulation of this theorem are added to insure the existence of continuous factors $H(e^{i\theta})$ in item 3 of the definition of **(SCOM)** in Section 2.2. Much of the theory can be developed without these restrictions, but the added generality seems to have little practical importance in the analysis of numerical algorithms, at least at this time. For a theorem of wider scope, which is formulated for the general **MOPT** problem and not just the H^∞ multidisk problem, see Theorem 2.3.

At this point we state a conjecture that is based on both numerical experiments (that are summarized in Section 1.5) and theoretical considerations. If the conjecture is true, then Theorem 1.8 becomes a very powerful tool for predicting the effectiveness of Newton's method and its variants.

Conjecture. *For generic multidisk optimization problems, if T' at solutions to* **PDMOPT** *is a Fredholm operator of index zero, then T' is invertible.*

One rationale for this conjecture is the general fact that the class $\mathcal{F}_0$ of Fredholm operators of index 0 is an open dense subset of $\mathcal{B}(X,X)$, see Section 1.11 in [BS:90]. Naturally, the set of invertible operators is an open dense subset of $\mathcal{B}(X,X)$, and so is an open subset of $\mathcal{F}_0$ which is dense in $\mathcal{F}_0$. The presumption behind the conjecture is that in the class of K^p's that produce $T'_{(G,f,\gamma)}$ in $\mathcal{F}_0$, almost all K^p's produce T' in the open dense set of invertible operators. Further, in this paper we see that T' has the form

$$T' = L + C$$

where L is a block Toeplitz matrix and C is a compact operator. The conditions in Theorem 1.8 (primarily $\mu_{pnull}(\theta) = 0$ and $\mu_{dnull}(\theta) = 0$ for all θ), imply that L is Fredholm of index 0. The conjecture thus asserts that for most selections of K^p, the resulting C has no special relationship to L and so yields an operator $L + C$ which is invertible.

Also in the paper we do a little more work than is immediately necessary in analysing the Toeplitz operator L. Possibly this will be useful in further investigation. For example, in Section 2.4.4 we give the dimension of the null space of L. The proof is long and is not included.

1.4.4 Rules of thumb

We observe in experiments that in multidisk situations, when the computed values of $\Psi^p(e^{i\theta})$ have rank that is independent of θ, then the *Algorithm G^TG produces T' with Fredholm index zero if $v = m$.* This section gives a theoretical justification for this observation, namely, we prove this to be true in situations where the Ψ^p , $p = 1, \ldots, v$ all have rank one (that is, $k_p = 1$) and the Ψ^p and f are in "general position," a notion we now define. In computer experiments we see that when all disk constraints are active (that is, $k_p > 0$) there is a dichotomy:

either

(1) $\Psi^p(e^{i\theta})$ has rank independent of θ, in which case the Ψ^p , $p = 1, \ldots, v$, all have rank one and general position holds.

or

(2) for some p, $\Psi^p(e^{i\theta})$ has rank equal to 0 on some intervals in $[0, 2\pi]$.

Now we introduce some definitions. A set of matrices $L^p \in \mathbf{C}^{m \times k_p}, p = 1, \ldots, v$, is said to have linearly independent ranges provided that any selection of nonzero vectors

$$x_p \in \text{ range } L^p, \quad p = 1, \ldots, v,$$

is linearly independent. Note that the set of matrix v-tuples $\{L^1, \ldots, L^v\} \in \mathbf{C}^{m \times k_1} \times \cdots \times \mathbf{C}^{m \times k_v}$ that satisfy the constraint

$$\sum_{p=1}^{v} \text{rank} L^p \leq m \tag{30}$$

and have linearly independent ranges is an open dense subset of the set $\mathbf{C}^{m\times k_1} \times \cdots \times \mathbf{C}^{m\times k_v}$, i.e., this is a generic property.

Definition 1.9 *The functions $\Psi = \mathrm{diag}(\Psi^1, \ldots, \Psi^v)$, f are in general position if for each point $e^{i\theta}$:*

1. *The matrices Ψ^p for $p = 1, \ldots, v$ have linearly independent ranges.*
2. *The matrices $(K^1 - f)\Psi^1, \ldots, (K^v - f)\Psi^v$ have linearly independent ranges.*
3. *The matrices $(K^1 - f), \ldots, (K^v - f)$ are all invertible.*

Definition 1.10 *For each θ, let*

$$\mu_D(\theta) := m - \mathrm{rank}\{\Psi^1(e^{i\theta}) + \cdots + \Psi^v(e^{i\theta})\}$$

and

$$\mu_R(\theta) := m - \mathrm{rank}\{[(K^1 - f)\Psi^1(K^1 - f)^T + \cdots + (K^v - f)\Psi^v(K^v - f)^T](e^{i\theta})\}.$$

Proposition 1.11 *Under the hypotheses of Theorem 1.8,*

$$\mu_{pnull}(\theta) = \mu_D(\theta)\ \mu_R(\theta).$$

If Ψ and f are in general position, then $\sum_{p=1}^{v} \mathrm{rank}\Psi^p \leq m$, and $\mu_{pnull}(\theta)$ which is independent of θ, equals

$$\mu_{pnull} = (m - \sum_{p=1}^{v} \mathrm{rank}\Psi^p)^2$$

and then $\mu_{dnull}(\theta) = 0$. If in addition $v = m$ and all constraints are active, then $\mu_{pnull} = 0$.

The proof of this proposition is given in Section 6.4

When $\sum_p^v k_p < m$, Proposition 1.11 says $\mu_{pnull} \neq 0$ which says that the differential T' will not be invertible (see Theorem 1.8) Since $\sum_{p=1}^v k_p < m$ is what typically occurs when $v < m$, Proposition 1.11 suggests that for almost all v-disk problems with $v < m$, the differential T' will not be invertible. Consequently, if $v < m$, algorithms based solely on Newton's method will behave badly. In particular, this is so for the Nehari problem when $m > 1$.

When $v > m$, the spectral factors G of Ψ do not exist, so the hypotheses of our theorems break down.

When $v = m$, we conclude from the lemma and our studies that Newton based methods are well behaved for almost all situations.

Some functional analysts might be both a bit disappointed and a bit surprised that the conditions for the invertibility of $T'(f, \Psi, \gamma)$ for any given problem can not be checked a priori. At best they hold for "generic" problems, which does not say anything about a particular problem. This, however, is a precise analog of the situation in linear programming [ZTD:92], where the main result says that if "strict complementarity" holds at the solution to a given LP problem, then numerous interior point methods are better than first order convergent to that solution. Strict complementarity does not always hold, it is merely generic, and one never knows until after an algorithm is run whether or not the optimum satisfies strict complementarity. Nonetheless, the theorem alluded to is extremely valuable in analyzing and assessing interior point methods in LP.

It is important to bear in mind that even if conditions on (G, f, γ) are found to insure that $T'_{(G,f,\gamma)}$ is invertible on an open dense set of (G, f, γ) in some suitable topology, there is no guarantee that for an open dense set of Γ the optimizing (G, f, γ) correspond to invertible differentials $T'_{G,f,\gamma}$. However, it does suggest that. Proving such results could easily take many people many years, and would likely have little effect on computational practice. Even when we allow many assumptions, as we shall see, quite a bit of mathematics must be developed. Thus, we assume things like smoothness freely.

1.5 Numerical experiments

We implemented Algorithms G^TG and $G + G^T$. We always took G^p to be an $m \times m$ matrix valued function rather than the more aggressive choice of $k \times m$. **We found Algorithm $G + G^T$ to be the most broadly effective** and would recommend it to the interested reader.

One caveat about the experiments is that they were run on a course grid (at most 128 points.) This is all that was practical since for convenience we represent functions on the disk inefficiently as values on grid points, or as power series. Possibly the coarse grid compromises some of our findings.

For the multidisk problem, there are of course various cases. In one case, for example, not all the performance measures Γ^p matter; in another case they all do matter. This leads us to call a performance measure Γ^p active at an optimum if Ψ^p is not identically 0. For example, Γ^2 matters if and only if at an optimum it is active for at least one θ. In what follows we only report on what happens when all the constraints turn out to be active.

Our numerical experiments for $v \leq m$, in generic situations, led to the following observations:

• For both algorithms we always found that $\text{rank}(\Psi^p) \leq 1$, for all θ and indeed for $v \leq m$, that $\text{rank}(\Psi^p) = 1$. While higher rank Ψ^p are usually possible these algorithms have a strong disposition not to produce them.

• For both algorithms we found in a few randomly selected problems that $\mu_{pnull}(\theta)$ is independent of θ in many cases (though not all), and that $\mu_{pnull} = 0$ predicts that T' is invertible, thus substantiating the conjecture of Section 1.4.3.

• Also if $\mu_{pnull}(\theta)$ is independent of θ, we found it to be given by the formula

$$(m - \sum_p \text{ rank } \Psi^p)^2 = (m - v)^2.$$

• In practice μ_{pnull} determines a finer structure of the null space of T'. When T' is not Fredholm one would expect that its null space is within finite dimensions of being invariant under multiplication by $e^{i\theta}$, and thus it makes sense to count its dimension as a module over H^∞. This number is easy to compute experimentally via the formula

$$\text{moddim} := \frac{\text{dim null space of } T'_{(G^*,f^*,\gamma^*)}}{\text{grid size}}.$$

We found in all experiments that if $\mu_{pnull}(\theta)$ is constant, then $\mu_{pnull} = \text{moddim}$.

• The reason Algorithm $G + G^T$ works better than expected appears to be that even when T' at optimum is not invertible, e.g., if $v < m$, at each iteration the current T' has good conditioning until the very end of the computer run (then it explodes). The Algorithm G^TG certainly does not have this property and its behavior is indeed dominated by the invertibility of T' at optimum, just as our theory predicts.

The table below gives our experimental findings for both algorithms.

m	v	$(\mathrm{m}-\mathrm{v})^2$	moddim	dim null T$'$	Experim. rank(G^p)
2	1	1	1	∞	1 for $\forall_p$
2	2	0	0	0	1 for $\forall_p$
3	1	4	4	∞	1 for $\forall_p$
3	2	1	1	∞	1 for $\forall_p$

In this paper, we shall develop the theory underlying the G^TG algorithm but do not analyze the $G + G^T$ algorithm. Nevertheless the experiments reported on here were run for both algorithms and suggest that μ_{pnull} has a strong connection with the Algorithm $G + G^T$ as well as for the Algorithm G^TG. The Algorithm $G + G^T$ is currently under study.

Comparisons with the matrix case and [AHO:96] are interesting. [AHO:96] have primal and dual non degeneracy conditions that are similar to the constraint $m = v$ in our case. They prefer their $XZ + ZX$ algorithm which is the analog of our Algorithm $G + G^T$. Experimentally we found that in our setting $m = v$ is stronger than needed to get good results (but not necessarily second order convergence).

2 General Theorems

This section begins by computing the differential of T and then analyzing its invertibility. We ultimately state a result about Newton type methods for the general **MOPT** problem. The subsequent sections show how this result specializes to the multidisk problem, and gives the proof of the results stated in the introduction.

2.1 Calculation of the differential of T

In this section we shall calculate the differential of the basic operator T that is defined by formula (26). We shall assume the smoothness conditions specified in Section 1.1, and we shall make use of (the first two terms of) the formula

$$\begin{aligned}
F(z_0+w) &= F(z_0) + 2\mathrm{Re}\left\{\frac{\partial F}{\partial z}(z_0)w\right\} \\
&\quad +\frac{1}{2}\left\{2w^T\frac{\partial^2 F}{\partial \overline{z}\partial z}(z_0)w + +2\mathrm{Re}\ \overline{w}^T\frac{\partial^2 F}{\partial z^2}w\right\} + o(\|w\|^2) \\
&= F(z_0) + \sum_{j=1}^{N}\frac{\partial F}{\partial z_j}(z_0)w_j + \sum_{j=1}^{N}\frac{\partial F}{\partial \overline{z}_j}(z_0)\overline{w}_j \\
&\quad +\frac{1}{2}\left\{2\sum_{i,j=1}^{N}\overline{w}_i\frac{\partial^2 F}{\partial \overline{z}_i\partial z_j}(z_0)w_j + \sum_{i,j=1}^{N}w_i\frac{\partial^2 F}{\partial z_i\partial z_j}(z_0)w_j\right. \\
&\qquad \left. + \sum_{i,j=1}^{N}\overline{w}_i\frac{\partial^2 F}{\partial \overline{z}_i\partial \overline{z}_j}(z_0)\overline{w_j}\right\} + o(\|w\|^2),
\end{aligned}$$

which is valid for smooth functions F from $\mathbf{C}^{\mathbf{N}} \to \mathbf{R}$. Here,

$$\frac{\partial^2 F}{\partial \overline{z}\partial z} \quad \text{and} \quad \frac{\partial^2 F}{\partial z^2}$$

are $N \times N$ matrices with ij entries equal to

$$\frac{\partial^2 F}{\partial \overline{z}_i\partial z_j} \quad \text{and} \quad \frac{\partial^2 F}{\partial z_i\partial z_j},$$

respectively.

(It is perhaps useful to recall that, in terms of the notation $z_j = x_j + iy_j$ and $\overline{z}_j = x_j - iy_j$,

$$\frac{\partial F}{\partial z_j} = \frac{1}{2}\left\{\frac{\partial F}{\partial x_j} - i\frac{\partial F}{\partial y_j}\right\} \quad \text{and} \quad \frac{\partial F}{\partial \overline{z}_j} = \frac{1}{2}\left\{\frac{\partial F}{\partial x_j} + i\frac{\partial F}{\partial y_j}\right\}$$

(for real or complex valued functions F) and that if F is analytic in the variables $z_1, \ldots, z_N$, then $\frac{\partial F}{\partial \bar{z}_j} = 0$.) Therefore,

$$T'_{(G,f,\gamma)} : \begin{pmatrix} H^2_+ \\ H^2_N \\ \mathbf{R} \end{pmatrix} \to \begin{pmatrix} H^2_+ \\ H^2_N \\ \mathbf{R} \end{pmatrix}$$

is given by

$$T'_{(G,f,\gamma)} \begin{bmatrix} \Delta \\ \varphi \\ \eta \end{bmatrix} = \begin{pmatrix} P_{H^2_+}[\Delta(\gamma I - \Gamma(\cdot, f))] - P_{H^2_+}[G\frac{\partial \Gamma}{\partial z}(\cdot, f)\varphi + G\frac{\partial \Gamma}{\partial \bar{z}}(\cdot, f)\overline{\varphi}] + \eta G \\ -P_{H^2_N} \mathrm{tr} \left[\left\{ G^T \Delta + \Delta^T G \right\} \frac{\partial \Gamma}{\partial z}(\cdot, f)^T + G^T G \left\{ \frac{\partial^2 \Gamma}{\partial \bar{z} \partial z}(\cdot, f)\varphi + \frac{\partial^2 \Gamma}{\partial \bar{z}^2}(\cdot, f)\overline{\varphi} \right\} \right] \\ 2\mathrm{Re}\frac{1}{2\pi} \int \mathrm{tr}\{G^T \Delta\} d\theta \end{pmatrix}. \tag{31}$$

Here, the second row of $T'_{(G,f,\gamma)}$ is the $N \times 1$ mvf with entries

$$-P_{H^2} \mathrm{tr} \left\{ \left[\Delta^T G + G^T \Delta \right] \frac{\partial \Gamma}{\partial \bar{z}_i} \right\} - \sum_{j=1}^{N} P_{H^2} \mathrm{tr} \left[G \frac{\partial^2 \Gamma}{\partial \bar{z}_i \partial z_j} G^T \varphi_j + G \frac{\partial^2 \Gamma}{\partial \bar{z}_i \partial \bar{z}_j} G^T \overline{\varphi}_j \right]$$

for $i = 1, \ldots, N$.

2.2 The assumptions PSCON and SCOM

From now on we shall be working with Γ and tuples (Ψ, f, γ) which satisfy certain assumptions. The first assumption includes the smoothness of Γ that is presumed throughout.

(**PSCON**) Γ is a positive semidefinite mvf that meets the smoothness assumptions that are specified in Section 1.2, and is plurisubharmonic.

The next assumption guarantees that (Ψ, f, γ) satisfies a function theoretic form of the strict complementarity condition that was briefly mentioned in Section 1.2.7.

(**SCOM**) 1. The triple (Ψ, f, γ) is continuous, with Ψ positive semidefinite, $f \in H^\infty_N$ and $\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta})) \geq 0$. Moreover, it meets the strict complementarity condition

$$\mathbf{C}^m = \mathrm{range}\Psi(e^{i\theta}) \oplus \mathrm{range}(\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta}))) \tag{32}$$

at every point $e^{i\theta}$.

2. There exists a $G \in H^\infty_{k\times m}$ that is outer and continuous with rank k at every point $e^{i\theta}$ such that $\Psi = G^T G$.
3. There exists an $H \in H^\infty_{\ell\times m}$, $\ell = m-k$, that is outer and continuous with rank ℓ at every point $e^{i\theta}$, such that

$$H(\gamma I_m - \Gamma(\cdot, f))H^T = I_{\ell\times\ell}.$$

The function G is called the *outer spectral factor* of Ψ; it is unique up to a constant $k\times k$ unitary left multiplier.

2.3 $T' = L + C$ with L selfadjoint and C compact

Proposition 2.1 *If assumptions* **(SCOM)** *and* **(PSCON)** *are in force, then the differential of* T *may be written in the form*

$$T'_{(G,f,\gamma)} = L + C,$$

where L *and* C *map* $H^2_+ \oplus H^2_N \oplus \mathbf{C}$ *into itself,* C *is compact and* L *is the selfadjoint operator given by the formula*

$$L\begin{bmatrix}\Delta\\ \varphi\\ \eta\end{bmatrix} = \begin{pmatrix} P_{H^2_+}[\Delta(\gamma I - \Gamma(\cdot,f))] & - & P_{H^2_+}[G\frac{\partial\Gamma}{\partial z}(\cdot,f)\varphi] & + & \eta G \\ -P_{H^2_N}[\mathrm{tr}\left\{(G^T\Delta)\frac{\partial\Gamma}{\partial z}(\cdot,f)^T\right\}] & - & P_{H^2_N}[\mathrm{tr}\{G^T G\frac{\partial^2\Gamma}{\partial\overline{z}\partial z}(\cdot,f)\varphi\}] & + & 0 \\ \frac{1}{2\pi}\int \mathrm{tr}\{G^T\Delta\}d\theta & + & 0 & + & 0 \end{pmatrix}$$

Proof. The issue reduces to analyzing L because of the fact that a Hankel operator

$$\mathcal{H}_\vartheta : g \in H^2 \to P_{H^{2\perp}}\vartheta g$$

with symbol $\vartheta \in H^\infty + \mathcal{C}$ is compact, as is its adjoint. Here $\mathcal{C}$ stands for the class of continuous functions on $\mathbf{T}$. In the present setting the operator

$$C := T' - L$$

can be expressed in the form

$$C\begin{bmatrix}\Delta\\ \varphi\\ \eta\end{bmatrix} = \begin{bmatrix} C_{12}(\varphi)\\ C_{21}(\Delta) + C_{22}(\varphi)\\ C_{31}(\Delta)\end{bmatrix},$$

where

$$C_{12}(\varphi) = -P_{H^2_+}\left[G\frac{\partial\Gamma}{\partial\overline{z}}\overline{\varphi}\right] = -\sum_{j=1}^N P_{H^2_+}\left[G\frac{\partial\Gamma}{\partial\overline{z}_j}\overline{\varphi}_j\right],$$

$$C_{21}(\Delta) = -P_{H_N^2} \operatorname{tr} \left\{ G \left(\frac{\partial \Gamma}{\partial z} \right)^T \Delta^T \right\}$$

is an $N \times 1$ mvf with entries

$$-P_{H^2} \operatorname{tr} \left\{ G \frac{\partial \Gamma}{\partial \overline{z}_i} \Delta^T \right\}, \qquad i = 1, \ldots, N,$$

$$C_{22}(\varphi) = -P_{H_N^2} \operatorname{tr} \left\{ G \frac{\partial^2 \Gamma}{\partial \overline{z}^2} G^T \overline{\varphi} \right\}$$

is an $N \times 1$ mvf with entries

$$-\sum_{j=1}^{N} P_{H^2} \left[\operatorname{tr} \left\{ G \frac{\partial^2 \Gamma}{\partial \overline{z}_i \partial \overline{z}_j} G^T \overline{\varphi_j} \right\} \right], \qquad i = 1, \ldots, N,$$

and

$$C_{31}(\Delta) = \frac{1}{2\pi} \int_0^{2\pi} \operatorname{tr} \left\{ G \Delta^T \right\} d\theta.$$

Thus, in view of the preceding remarks, C is compact. It is also readily checked that L is selfadjoint with the help of the formula

$$\langle \Delta, L_{12}(\varphi) \rangle_{H_+^2} = \langle L_{21}(\Delta), \varphi \rangle_{H_N^2} \ . \tag{33}$$

We now verify (33).

$$\begin{aligned}
\langle L_{21}(\Delta), \varphi \rangle_{H_N^2} &= -\int \textstyle\sum_{j=1}^N \operatorname{tr}[G^T \Delta (\frac{\partial \Gamma}{\partial z_j}(\cdot, f))^T] \overline{\varphi_j} \\
&= -\int \textstyle\sum_{j=1}^N \operatorname{tr}[\frac{\partial \Gamma}{\partial z_j}(\cdot, f)^T \overline{\varphi_j} G^T \Delta] \\
&= -\int \textstyle\sum_{j=1}^N \operatorname{tr}[\Delta (G \frac{\partial \Gamma}{\partial z_j}(\cdot, f) \varphi_j)^T] \\
&= \langle \Delta, L_{12}(\varphi) \rangle_{H_+^2} \ .
\end{aligned}$$

□

2.4 Fredholmness of T'

Proposition 2.1 implies that T' is Fredholm of index 0 if and only if L is Fredholm of index 0. We shall give conditions guaranteeing that L is Fredholm, but first we need a few definitions.

2.4.1 Key definitions

Definition 2.2 *Let* $a = \frac{\partial \Gamma}{\partial z}(\cdot, f) = (a_1, \ldots, a_N) : \mathbf{T} \to (\mathbf{C}^{\mathbf{m}\times\mathbf{m}})^N$ *(so that* $a\varphi = \sum_{j=1}^N a_j \varphi_j$*) and* $G : \mathbf{T} \to \mathbf{C}^{k\times m}$ *be continuous functions and let* A *denote the* $m \times m$ *matrix with* ij *entry*

$$A_{ij} = \mathrm{tr}\{G \frac{\partial^2 \Gamma}{\partial \overline{z}_i \partial z_j}(\cdot, f) G^T\}.$$

i. *Let* σ_{primal} *and* σ_{dual} *denote the multiplication operators*

$$\sigma_{primal} : \varphi \in H^2_N \quad \longrightarrow \quad \begin{pmatrix} Ga\varphi \\ \\ A\varphi \end{pmatrix} \in L^2_{k\times m} \times L^2_N$$

$$\sigma_{dual} : \Delta_1 \in (H^2_{k\times k})_+ \quad \longrightarrow \quad \mathrm{tr}\{Ga^T G^T \Delta_1\} \in L^2_N$$

ii. *The multiplication operator* σ_{primal} *is said to be* **regular** *if it is generated by a function which has a trivial null space at every point* $e^{i\theta} \in \mathbf{T}$.

iii. *The multiplication operator* σ_{dual} *is said to be* **regular** *if the span of the* $k \times k$ *matrices* $(Ga_j^T G^T)(e^{i\theta})$ *is equal to* $\mathbf{C}^{k\times k}$ *at every point* $e^{i\theta} \in \mathbf{T}$.

2.4.2 A key result

Theorem 2.3 *Assume that* Γ, G, f *and* γ *meet the following constraints:*

(1) *The assumptions* **(PSCON)** *and* **(SCOM)** *are satisfied.*

(2) *The operators* σ_{primal} *and* σ_{dual} *are regular.*

Then $T'_{(G,f,\gamma)}$ *is Fredholm of index 0.*

Remark. The reader may wonder why the measure nondegeneracy condition does not appear explicitly in the preceding theorem (as well as the next one). The reason is that, although it intervenes in the theorems of Chapter 1 to help insure that $T(G, F, \gamma) = 0$, at solutions of the **PDMOPT**, we then calculate the differential T' of T at an arbitrary "point" (G, F, γ) and study its properties. The properties of T' and L etc as operators do not involve the measure non degeneracy per se. Therefore it does not appear in the statements of the current theorems. It is "only" when you try to connect with the original problem that these assumptions come into play.

The last theorem is a consequence of Theorem 3.9. The proof is given in Section 3.4.6. Note that the most damning assumption we made is that Ψ has a nice spectral factor G. However, the numerical experiments that are discussed in [HMW:98] indicate that this is often true.

2.4.3 Measure non-degeneracy vs. σ_{dual} regular

If ν is a selfadjoint matrix valued measure, then

$$d\mu = G^T d\nu\ G \tag{34}$$

is also, and, in addition, $d\mu$ satisfies

$$(\gamma I_m - \Gamma) d\mu = 0. \tag{35}$$

Strict complementarity of $\gamma I_m - \Gamma$ and Ψ suggests that (35) implies (34), although we have not checked this.

A matrix valued measure $d\mu$ of the form (34) satisfies the measure non-degeneracy condition $\mathcal{V}_f(e^{i\theta}, d\mu) = 0$ if and only if

$$\mathrm{tr}(a_j G^T d\nu G) = 0 \quad \text{for} \quad j = 1, \ldots, N.$$

This implies that $d\nu = 0$ if and only if the span of the matrices $(G a_j^T G^T)(e^{i\theta}),\ \ j = 1 \ldots, N$ is equal to $\mathbf{C}^{k\times k}$, the set of all complex $k \times k$ matrices. This is exactly the σ_{dual} regularity condition. Thus we see that measure non-degeneracy and σ_{dual} regularity are closely related. Measure non degeneracy is, however, a less restrictive condition because it only need hold for matrix valued measures $d\mu$ which are nonnegative.

2.4.4 The null space of L

We have considerably stronger theorems that tell us the size of the null space of L. We begin with a key definition.

Definition 2.4 *Let* $a = (a_1, \ldots, a_N) : \mathbf{T} \to (\mathbf{C}^{m\times m})^N$ *(so that* $a\varphi = \sum_{j=1}^N a_j \varphi_j$*) and let* $G : \mathbf{T} \to \mathbf{C}^{k\times m}$ *be continuous functions.*

i. *Let* $\mathcal{L}_{primal}$ *and* $\mathcal{L}_{dual}$ *denote the operators given by*

$$\mathcal{L}_{primal}\varphi := \begin{pmatrix} P_{H^2_+} G a \varphi \\ \\ P_{H^2_N} A\varphi \end{pmatrix} \quad \text{and} \quad \mathcal{L}_{dual}\Delta_1 := \begin{pmatrix} P_{H^2_N} \mathrm{tr}\{G a^T G^T \Delta_1\} \\ \\ \frac{1}{2\pi}\int\ \mathrm{tr}\ \{G G^T \Delta_1\} d\theta \end{pmatrix},$$

where $\Delta_1 \in H^2_{k\times k}$ *and* $\varphi \in H^\infty_N$*. We call* σ_{primal} *(resp.* σ_{dual}*) the symbol of the operator* $\mathcal{L}_{primal}$ *(resp.* $\mathcal{L}_{dual}$*).*

ii. *The operator* $\mathcal{L}_{primal}$ *(resp.* $\mathcal{L}_{dual}$*) is said to be* **regular** *if* σ_{primal} *(resp.* σ_{dual}*) is regular.*

Additional analysis leads to a refinement of Theorem 2.3, which is based on the observation that the null space of L effectively splits into two parts. One comes from the primal problem and one from the dual problem. The next theorem, which is stated without proof in order to save space, serves as a sample.

Theorem 2.5 *Assume that Γ, G, f and γ meet the following constraints:*

(1) *The assumptions* **(PSCON)** *and* **(SCOM)** *are satisfied.*

(2) *The operators σ_{primal} and σ_{dual} are regular.*

Then the dimension of the space of all triples $(\Delta, \varphi, \eta) \in H^2_+ \times H^2_N \times \mathbf{R}$ in the null space of L is equal to the dimension of the space of all triples $(\Delta_1, \varphi, \eta) \in (H^2_{k\times k})_+ \times H^2_N \times \mathbf{R}$ such that

$$\mathcal{L}_{primal}\varphi = \begin{pmatrix} G\eta \\ 0 \end{pmatrix} \quad \textit{and} \quad \mathcal{L}_{dual}\Delta_1 = 0.$$

We remark that $\mathcal{L}_{primal}\varphi = 0$ may not force φ to be 0.

3 Proofs

3.1 The null space of σ_{dual}

Before launching into proofs, we shall show that $\mathcal{U}_f(e^{i\theta})^T\mathcal{U}_f(e^{i\theta})$ invertible implies that σ_{dual} is regular. Although we shall not need this result in the rest of this paper, it does serve to illustrate the connections between a number of the conditions that are often imposed in optimization problems of this sort.

Proposition 3.1 *Let $\Gamma(e^{i\theta}, z)$ satisfy the smoothness conditions specified in Section 1.2, let $f \in H^\infty_N$ be continuous and recall the notation*

$$a_\ell(\cdot) := \frac{\partial \Gamma}{\partial z_\ell}(\cdot, f(\cdot)), \quad \ell = 1, \ldots, N, \quad \textit{and} \quad a := (a_1, \ldots, a_N).$$

The condition $\mathcal{U}_f(e^{i\theta})^T\mathcal{U}_f(e^{i\theta})$ is invertible for all θ, implies $N \geq m^2$ and is equivalent to each of the following two conditions:

i. *The span of the set $\{a_1^T, a_2^T, , \ldots, a_N^T\}$ is equal to $\mathbf{C}^{m\times m}$ for all θ.*

ii. *For each θ, the map of $B \in \mathbf{C}^{m\times m} \mapsto aB = (a_1 B, \ldots, a_N B) \in (\mathbf{C}^{m\times m})^N$ has a trivial nullspace.*

Proof. This is done by careful bookkeeping applied to the definition of $\mathcal{U}_f$. Formula (16) is helpful. $\square$

3.2 Changing the variables of L

We shall analyze the invertibility of L with the help of a change of variables. If $Q \in H^\infty_{m\times m}$, then

$$\Delta \longrightarrow \widetilde{\Delta} := \Delta\, Q$$

is a map of $(H^2_{k\times m})_+$ into $H^2_{k\times m}$. Moreover, if $Q(0)$ is upper triangular, then this map sends $(H^2_{k\times m})_+$ into itself. If both Q and its pointwise inverse Q^{-1} belong to $H^\infty_{m\times m}$, then these mappings are onto. The adjoint of this mapping with respect to the matrix inner product

$$\langle A, B\rangle = \operatorname{tr}\left\{AB^T\right\}$$

is

$$\Delta \to \Delta\, Q^T. \tag{36}$$

We consider the map M_0 from $(H^2_{k\times m})_+ \oplus H^2_N \oplus \mathbf{R}$ into itself that is defined by the rule

$$M_0(\Delta,\ \varphi,\ \eta) := (\Delta\, Q, \varphi, \eta).$$

The adjoint of this map in the $1/2\pi \int_0^{2\pi} \text{tr}\{AB^T\}d\theta$ inner product is the Toeplitz like operator M_0^T defined by

$$M_0^T(\Delta,\ \varphi,\ \eta) := (P_{H^2_+}\{\Delta\ Q^T\}, \varphi, \eta).$$

Our goal is to choose Q so that invertibility of the map

$$\widetilde{L} := M_0^T L M_0 \tag{37}$$

is easy to analyze. We shall assume Q is in $(H^\infty_{m\times m})_+$, and hence that the constant term $Q(0)$ is upper triangular.

The abbreviations

$$a := \frac{\partial \Gamma}{\partial z}(\cdot, f) = [a_1, a_2, \ldots, a_N] \quad \textit{and} \quad a\varphi = \sum_{j=1}^{N} a_j \varphi_j$$

that were introduced earlier will be useful.

It is clear from formula (37) that $\widetilde{L}$ is selfadjoint. Moreover,

$$\widetilde{L}\begin{bmatrix} \Delta \\ \varphi \\ \eta \end{bmatrix} = \begin{pmatrix} P_{H^2_+}[\Delta\ Q(\gamma I - \Gamma(\cdot, f))Q^T] & - & P_{H^2_+}[Ga\varphi Q^T] & + & \eta P_{H^2_+}[GQ^T] \\ -P_{H^2_N}[\text{tr}\{G^T \Delta Q a^T\}] & - & P_{H^2_N}[A\varphi] & + & 0 \\ \frac{1}{2\pi}\int \text{tr}\{G^T \Delta\ Q\}d\theta & + & 0 & + & 0 \end{pmatrix}. \tag{38}$$

The exhibited formula for $\widetilde{L}$ follows easily from the formula for L with the help of

Lemma 3.2 *If Q is in $(H^\infty_{m\times m})_+$ and $F \in L^2_{k\times m}$, then*

$$P_{H^2_+}\left\{\left(P_{H^2_+}F\right)Q^T\right\} = P_{H^2_+}\left\{FQ^T\right\} \ .$$

Proof. If $F \in L^2_{k\times m}$, then

$$F = F_+ + F_-,$$

where $F_+ \in H^2_+$ and $F_- \in (H^2_+)^\perp$. Therefore,

$$FQ^T = F_+Q^T + F_-Q^T.$$

But now as

$$\langle C, F_-Q^T\rangle = \langle CQ, F_-\rangle = 0$$

for every $C \in H^2_+$, it follows that $F_-Q^T \in (H^2_+)^\perp$. Thus

$$P_{H^2_+}\left\{FQ^T\right\} = P_{H^2_+}\left\{F_+Q^T\right\} \ ,$$

as claimed. $\square$

The next conclusion is immediate from formula (37).

Proposition 3.3 *If Q and Q^{-1} belong to $H^\infty_{m\times m}$, then L is invertible (resp. Fredholm of index k) if and only if $\widetilde{L}$ is invertible (resp. Fredholm of index k).*

3.3 A nice form for $\widetilde{L}$

3.3.1 The spectral factors G, H and Q

Now we select $Q \in (H^\infty_{m\times m})_+$ to meet our ends. There are two closely related ways to define Q.

1. Define ρ by

$$\rho := \Psi^{[-1]} + (\gamma I - \Gamma) \ ,$$

where $\Psi^{[-1]}$ denotes the Moore-Penrose inverse of Ψ. Then, in view of the strict complementarity assumption (**SCOM**), the inequality $\rho(e^{i\theta}) \geq \delta I_{m\times m}$ holds a.e. for some $\delta > 0$. Take Q to be the outer function in $H^\infty_{m\times m}$ satisfying

$$\rho^{-1} = Q^T Q.$$

Then Q is invertible in $H^\infty_{m\times m}$, and

$$\rho = Q^{-1}Q^{-T} \quad \text{and} \quad Q\,\rho\,Q^T = I_{m\times m} \ .$$

2. Note that Q may be expressed in terms of the outer mvf's $G \in (H^\infty_{k\times m})_+$ and $H \in H^\infty_{\ell\times m}$ that appear in the formulas $\Psi = G^T G$ and $H(\gamma I - \Gamma)H^T = I_{\ell\times\ell}$ that were introduced in the assumption (**SCOM**) as

$$Q = \begin{pmatrix} G \\ \\ H \end{pmatrix} .$$

Here rank $G = k$ and rank $H = \ell$. To see this, recall that at the optimum choice of (G, f, γ),

$$G\{\gamma I - \Gamma(\cdot, f)\} = 0 \ .$$

Therefore, upon introducing the singular value decomposition

$$G = U[D \quad 0]V,$$

where U is $k\times k$ unitary, D is $k\times k$ positive diagonal and V is $m\times m$ unitary at each point θ, it follows that

$$\Psi^{[-1]} = V^T \begin{bmatrix} D^{-2} & 0 \\ 0 & 0 \end{bmatrix} V \ ,$$

$$G\Psi^{[-1]}G^T = I_{k\times k}$$

and

$$\text{range}\Psi^{[-1]}G^T = \text{range}V^T \begin{bmatrix} I_{k\times k} \\ 0 \end{bmatrix} = \text{range}G^T = \text{range}\Psi^{[-1]} = \text{range}\Psi \ ,$$

whereas,

$$\mathrm{range}\{\gamma I_m - \Gamma(\cdot, f)\} = \mathrm{range} V^T \begin{bmatrix} 0 \\ I_{\ell\times\ell} \end{bmatrix} .$$

Thus,

$$\gamma I_m - \Gamma(\cdot, f) = V^T \begin{bmatrix} 0 & 0 \\ 0 & E^2 \end{bmatrix} V$$

for some positive definite $\ell \times \ell$ matrix E and hence, upon setting

$$H = [0 \quad E^{-1}]V ,$$

we see that

$$H\Psi^{[-1]} = 0$$

and

$$H\rho H^T = H\{\gamma I_m - \Gamma(\cdot, f)\}H^T = I_{\ell\times\ell} .$$

Therefore,

$$Q = \begin{pmatrix} G \\ H \end{pmatrix}$$

has the property

$$Q\{\gamma I_m - \Gamma(\cdot, f)\}Q^T = \begin{bmatrix} 0 & 0 \\ 0 & I_{\ell\times\ell} \end{bmatrix}.$$

3.3.2 $\widetilde{L}$ in nice coordinates

The next step is to partition Δ as

$$\Delta = (\Delta_1 \ \Delta_2)$$

with $\Delta_1 \in (H^2_{k\times k})_+$ and $\Delta_2 \in H^2_{k\times\ell}$ (so that $\Delta_1(0)$ is upper triangular, while $\Delta_2(0)$ is arbitrary). It then follows that $\widetilde{L}$ acting on

$$\begin{pmatrix} \Delta \\ \varphi \\ \eta \end{pmatrix}$$

can be reexpressed in the form

$$\begin{pmatrix} P_{H^2_+}[(0\ \Delta_2)] & - & P_{H^2_+}[(GaG^T\ GaH^T)\varphi] & + & \eta P_{H^2_+}[(GG^T\ 0)] \\ -P_{H^2_N}[\mathrm{tr}(\Delta_1 Ga^T G^T + \Delta_2 Ha^T G^T)] & - & P_{H^2_N}[A\varphi] & + & 0 \\ \frac{1}{2\pi}\int \mathrm{tr}\{GG^T\Delta_1\}d\theta & + & 0 & + & 0 \end{pmatrix}. \tag{39}$$

Here we used the fact that the first optimality condition implies that

$$G\ H^T = 0\ ,$$

that is, the range of G^T and the range of H^T are orthogonal complements. Moreover, the following conventions are in force:

$$P_{H_+^2}\left[\begin{pmatrix} GaG^T & GaH^T \end{pmatrix}\varphi\right] = P_{H_+^2}\left[\sum_{j=1}^{N}\begin{pmatrix} Ga_jG^T & Ga_jH^T \end{pmatrix}\varphi_j\right]$$

and

$$P_{H_N^2}\left[\operatorname{tr}\left(\Delta_1 G a^T G^T + \Delta_2 H a^T G^T\right)\right] = P_{H_N^2}\begin{bmatrix} \operatorname{tr}(\Delta_1 G a_1^T G^T + \Delta_2 H a_1^T G^T) \\ \vdots \\ \operatorname{tr}(\Delta_1 G a_N^T G^T + \Delta_2 H a_N^T G^T) \end{bmatrix}.$$

3.4 $\widetilde{L}$ and Fredholmness

We shall show in this section that the operator $\widetilde{L}$ given by formula (39) in Section 3.3 is a Toeplitz operator. Also, we establish conditions under which $\widetilde{L}$ is Fredholm of index zero, and we shall discuss invertibility issues.

3.4.1 Assumptions and notation

Throughout this section we suppose that $\Gamma(e^{i\theta}, z)$ satisfies **PSCON** and that (Ψ, f, γ) satisfies **SCOM**. In addition, we assume that G is a continuous outer function in $(H^2_{k\times m})_+$ and that Q is a continuous invertible outer function in $(H^2_{m\times m})_+$, which is obtained by stacking the given matrix function G with the $(k-m)\times m$ matrix valued analytic function H that was considered in the preceding subsection. Thus, $a = (a_1, \ldots, a_N)$ is a continuous mvf whose entries a_j take $m\times m$ matrix values and A is a continuous positive semidefinite mvf from $\mathbf{T}$ into $\mathbf{C}^{N\times N}$.

3.4.2 $\widetilde{L}$ is Toeplitz

We now introduce a multiplication operator that plays a key role in our discussion of $\widetilde{L}$ and T'.

Definition 3.4 *For $\ell = m-k$, let P_2 denote the projection operator*

$$\begin{aligned} P_2 : (H^2_{k\times m})_+ = (H^2_{k\times k})_+ \times H^2_{k\times \ell} \quad &\longrightarrow \quad (H^2_{k\times m})_+ \\ \Delta = (\Delta_1, \Delta_2) \quad &\longmapsto \quad \Delta\Pi_2 = (0, \Delta_2), \end{aligned} \tag{40}$$

where Π_2 *designates the matrix*

$$\begin{pmatrix} 0_{k\times k} & 0_{k\times \ell} \\ 0_{\ell\times k} & I_{\ell\times\ell} \end{pmatrix},$$

and let M *denote the operator*

$$\begin{aligned} M : (H^2_{k\times m})_+ \times H^2_N \quad &\longrightarrow \quad L^2_{k\times m}\times L^2_N \\ \begin{bmatrix} \Delta \\ \varphi \end{bmatrix} \quad &\longmapsto \quad \begin{bmatrix} \Delta\Pi_2 - GaQ^T\varphi \\ \\ -\mathrm{tr}[\Delta Q a^T G^T] - A\varphi \end{bmatrix} \end{aligned} \tag{41}$$

Note that we are representing the values of M in block column format. The map M is defined pointwise, so for fixed θ we may think of M as a map from $\mathbf{C}^{k\times m}\times\mathbf{C}^N$ to itself. Also, a similar comment applies to P_2, and one may view the pointwise action of P_2 as multiplication (on the right) of $k\times m$ matrices by the block matrix Π_2.

We now confirm that the map M is related to $\widetilde{L}$ in an obvious way.

Definition 3.5 *Let* $\mathcal{P}$ *denote the orthogonal projection operator*

$$\mathcal{P} : L^2_{k\times m}\times L^2_N \longrightarrow (H^2_{k\times m})_+ \times H^2_N \tag{42}$$

and let

$$\mathcal{L} : (H^2_{k\times m})_+ \times H^2_N \to (H^2_{k\times m})_+ \times H^2_N$$

be the operator given by

$$\mathcal{L}(\cdot) = \mathcal{P}(M(\cdot)) \tag{43}$$

Formula (43) suggests that $\mathcal{L}$ is a Toeplitz operator. This is indeed the case.

Proposition 3.6 *The operator* $\mathcal{L}$ *is a Toeplitz operator* $\widetilde{\mathcal{T}}$ *with continuous symbol defined on* $(H^2_{k\times m})_+\times H^2_N$*. Moreover,*

$$\widetilde{L}\begin{bmatrix} \Delta \\ \varphi \\ \eta \end{bmatrix} = \begin{bmatrix} \mathcal{L}\begin{pmatrix} \Delta \\ \varphi \end{pmatrix} \\ 0 \end{bmatrix} + \begin{bmatrix} P_{H^2_+}[GQ^T]\eta \\ 0 \\ \frac{1}{2\pi}\int \mathrm{tr}\{G^T\Delta\, Q\}d\theta \end{bmatrix}. \tag{44}$$

Proof. To see this, represent $H^2_{k\times m}\times H^2_N$ as H^2_{km+N} and use a shift on some of the entries of $H^2_{k\times m}\times H^2_N$ to obtain $(H^2_{k\times m})_+\times H^2_N$. Clearly, the action of M on $(H^2_{k\times m})_+\times H^2_N$ corresponds to the action of a suitable multiplication operator with continuous matrix symbol $\widetilde{M}$ on elements in H^2_{km+N}. Hence the operator $\mathcal{P}(M(\cdot)) = \mathcal{L}$ is a Toeplitz operator with continuous symbol. □

3.4.3 $\mathcal{L}$ is selfadjoint

Proposition 3.7 *The operators M and $\mathcal{L}$ are self adjoint with respect to the indicated inner products.*

Proof. For each fixed θ, one can view M as a mapping from $\mathbf{C}^{k\times m} \times \mathbf{C}^N$ to itself. We choose the inner product on this space to be the Euclidean one,

$$\Big\langle (\triangle, \varphi), (\delta, \psi) \Big\rangle_{\mathbf{C}^{k\times m}\times\mathbf{C}^N} = \mathrm{tr}(\delta^T \triangle) + (\psi^T \varphi)$$

We have that for all $(\triangle, \varphi), (\delta, \psi) \in \mathbf{C}^{k\times m} \times \mathbf{C}^N$,

$$\begin{aligned}
&\Big\langle (\triangle, \varphi), M[(\delta, \psi)] \Big\rangle_{\mathbf{C}^{k\times m}\times\mathbf{C}^N} \\
&\quad = \mathrm{tr}\left\{(\delta\Pi_2 - GaQ^T\psi)^T \triangle\right\} + (-\mathrm{tr}\left[\delta Q a^T G^T\right] - A\psi)^T \varphi \\
&\quad = \mathrm{tr}\left\{(\Pi_2\delta^T - \psi^T Q a^T G^T)\triangle\right\} + (-\mathrm{tr}(GaQ^T\delta^T) - \psi^T A^T)\varphi \\
&\quad = \mathrm{tr}(\Pi_2\delta^T\triangle - GaQ^T\delta^T\varphi) + \mathrm{tr}(-\psi^T Q a^T G^T \triangle) - \psi^T A^T \varphi \\
&\quad = \mathrm{tr}(\delta^T\triangle\Pi_2 - \delta^T GaQ^T\varphi) + \psi^T(-\mathrm{tr}(Qa^T G^T \triangle) - A\varphi) \\
&\quad = \mathrm{tr}\left\{\delta^T(\triangle\Pi_2 - GaQ^T\varphi)\right\} + \psi^T(-\mathrm{tr}(\triangle Q a^T G^T) - A\varphi) \\
&\quad = \Big\langle M[\triangle, \varphi], (\delta, \psi) \Big\rangle_{\mathbf{C}^{k\times m}\times\mathbf{C}^N}
\end{aligned}$$

If now $(\triangle, \varphi), (\delta, \psi)$ represent functions in $(H^2_{k\times m})_+ \times H^2_N$, then

$$\begin{aligned}
&\Big\langle (\triangle, \varphi), \mathcal{L}(\delta, \psi) \Big\rangle_{L^2_{k\times m}\times L^2_N} \\
&\quad = \Big\langle (\triangle, \varphi), \mathcal{P}(M[\delta, \psi]) \Big\rangle_{L^2_{k\times m}\times L^2_N} \\
&\quad = \Big\langle (\triangle, \varphi), M[\delta, \psi] \Big\rangle_{L^2_{k\times m}\times L^2_N} \\
&\quad = \frac{1}{2\pi}\int_0^{2\pi} \Big\langle (\triangle, \varphi), M[\delta, \psi] \Big\rangle_{\mathbf{C}^{k\times m}\times\mathbf{C}^N} d\theta \\
&\quad = \frac{1}{2\pi}\int_0^{2\pi} \Big\langle M[\triangle, \varphi], (\delta, \varphi) \Big\rangle_{\mathbf{C}^{k\times m}\times\mathbf{C}^N} d\theta \\
&\quad = \Big\langle M[\triangle, \varphi], (\delta, \psi) \Big\rangle_{L^2_{k\times m}\times L^2_N} \\
&\quad = \Big\langle \mathcal{P}[M[\triangle, \varphi]], (\delta, \psi) \Big\rangle_{L^2_{k\times m}\times L^2_N} \\
&\quad = \Big\langle \mathcal{L}(\triangle, \varphi), (\delta, \psi) \Big\rangle_{L^2_{k\times m}\times L^2_N}.
\end{aligned}$$

The rest of the proof is straightforward. $\square$

3.4.4 The null space of M

In this subsection we derive a characterization of the null space of M that is useful for our study of the operator $\widetilde{L}$.

Proposition 3.8 *Let $(\triangle, \varphi) \in \mathbf{C}^{k\times m} \times \mathbf{C}^N$. Partition Δ as $(\Delta_1, \Delta_2) \in \mathbf{C}^{k\times k} \times \mathbf{C}^{k\times(m-k)}$. Then $M[\triangle, \varphi] = 0$ if and only if the following conditions hold:*

$$\textbf{i.} \qquad \Delta_2 = 0.$$

$$\textbf{ii.} \qquad \sigma_{primal}(\varphi) = 0.$$

$$\textbf{iii.} \qquad \sigma_{dual}(\Delta_1) = 0.$$

Proof. If $M[\triangle, \varphi] = 0$, then

$$[0 \quad \triangle_2] - G \sum_{\ell=1}^{N} a_\ell \varphi_\ell [G^T \;\; H^T] = 0 \tag{45}$$

and

$$-\mathrm{tr}\left\{\triangle_1 G a^T G^T\right\} - \mathrm{tr}\left\{\triangle_2 H a^T G^T\right\} - A\varphi = 0. \tag{46}$$

From (45) it follows that

$$0 = G \sum_{\ell=1}^{N} a_\ell \varphi_\ell G^T = G a\varphi G^T \tag{47}$$

and

$$\triangle_2 = G \sum_{\ell=1}^{N} a_\ell \varphi_\ell H^T = G a\varphi H^T. \tag{48}$$

Combine (48) and (46) to obtain

$$\mathrm{tr}\left\{\triangle_1 G a^T G^T\right\} + \mathrm{tr}\left\{G \sum_{\ell=1}^{N} \; a_\ell \varphi_\ell H^T H a^T G^T\right\} + A\varphi = 0. \tag{49}$$

Next, multiply by φ^T on the left to obtain

$$\mathrm{tr}\left\{\triangle_1 G (a\varphi)^T G^T\right\} + \mathrm{tr}\left\{G a\varphi H^T H (a\varphi)^T G^T\right\} + \varphi^T A\varphi = 0. \tag{50}$$

The first term on the left hand side of (50) is zero by (47) and each of the remaining two terms is nonnegative. Thus, we see that

$$G a\varphi H^T = 0 \tag{51}$$

and

$$A\varphi = 0. \tag{52}$$

From (47), (51), the invertibility of $Q^T = (G^T \quad H^T)$ and (52), we conclude that $\sigma_{primal}(\varphi) = 0$.

The left hand side of (51) is precisely $\triangle_2$ by formula (48), that is, we have

$$\triangle_2 = 0. \tag{53}$$

Now substitute (52) and (53) into (46) to obtain $\sigma_{dual}(\Delta_1) = 0$.

We now prove the converse. Suppose that equations (i) -(iii) hold. Then, since $\Delta\Pi_2 = \Delta_2 = 0$, we see that

$$M \begin{bmatrix} \Delta \\ \varphi \end{bmatrix} = \begin{bmatrix} \Delta\Pi_2 - GaQ^T\varphi \\ -\mathrm{tr}[\Delta Q a^T G^T] - A\varphi \end{bmatrix} = \begin{bmatrix} -Ga\varphi Q^T \\ -\mathrm{tr}[\Delta_1 G a^T G^T] - A\varphi \end{bmatrix} = 0.$$

$\square$

3.4.5 $\widetilde{L}$ as a Fredholm operator of index zero

Theorem 3.9 *Suppose that $\Gamma(e^{i\theta}, z)$ satisfies* **(PSCON)** *and let (Ψ, f, γ) be such that* **(SCOM)** *is satisfied. Let G be a continuous outer function of rank k in $(H^2_{k\times m})_+$ such that $G^T G = \Psi$ (the existence of such an outer function is guaranteed by* **(SCOM)***) and assume that the nullspace of $\mathcal{V}_f(e^{i\theta})$ restricted to matrices of the form*

$$\{G^T(e^{i\theta}) B G(e^{i\theta}) : \ B \in \mathbf{C}^{k\times k}\}$$

is equal to zero for every point $e^{i\theta}$. Then σ_{dual} is regular.

If, in addition, the operator σ_{primal} is regular, then $\widetilde{L}$ and T' are Fredholm operators of index zero.

Proof. Suppose first that $B \in \mathbf{C}^{k\times k}$ is orthogonal to the span of

$$\{(Ga_j^T G^T)(e^{i\theta}) : \quad j = 1, \ldots, N\}$$

for some choice of $e^{i\theta}$. Then

$$\begin{aligned} 0 &= \langle B, (Ga_j^T G^T)(e^{i\theta})\rangle \\ &= \mathrm{tr}\{(Ga_j G^T)(e^{i\theta}) B\} \\ &= \mathrm{tr}\{a_j(e^{i\theta}) G^T(e^{i\theta}) B G(e^{i\theta})\} \end{aligned}$$

for $j = 1, \ldots, N$. Moreover, by the **(SCOM)** assumption

$$\{\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta}))\} G(e^{i\theta})^T = 0\ .$$

Consequently, $G(e^{i\theta})^T BG(e^{i\theta})$ belongs to the null space of $\mathcal{V}_f(e^{i\theta}, \cdot)$. Therefore, by assumption, $G(e^{i\theta})^T BG(e^{i\theta}) = 0$, and hence, as $G(e^{i\theta})$ is right invertible by assumption, $B = O_{k\times k}$. This completes the proof that σ_{dual} is regular.

Now assume in addition that σ_{primal} is also regular. We claim that for each θ, the operator M acting on $\mathbf{C}^{m\times m} \times \mathbf{C}^N$ has a trivial null space. To see this, suppose $M[\Delta, \varphi] = 0$ for an element $(\Delta, \varphi) \in \mathbf{C}^{m\times m} \times \mathbf{C}^N$. Then, by (i) and (ii) of Proposition 3.8, we must have $\Delta_2 = 0$ and

$$\sigma_{primal}(\varphi) = \begin{pmatrix} Ga\varphi \\ \\ A\varphi \end{pmatrix} = 0. \tag{54}$$

Since σ_{primal} is regular, it follows that $\varphi = 0$.

Next, from item (iii) of Proposition 3.8 we see that

$$\langle \Delta_1, Ga_\ell G^T \rangle = 0 \quad \text{for} \quad \ell = 1, \ldots, N. \tag{55}$$

Therefore, since σ_{dual} is regular, $\Delta_1 = 0$. Thus, $\Delta = 0$ and M has a trivial kernel. Since M is selfadjoint by Proposition 3.7, we conclude that M is an invertible multiplication operator for each point θ. Hence, $\mathcal{L}$ is a self adjoint Toeplitz operator with continuous, pointwise invertible symbol. Theorem 2.94 on page 96 of [BS:90] guarantees that such operators are Fredholm. Thus, we conclude that $\mathcal{L}$ is a Fredholm operator. Moreover, since $\mathcal{L}$ is selfadjoint by Proposition 3.7, the Fredholm index of $\mathcal{L}$ is zero.

By Proposition 3.6, $\widetilde{L}$ (and therefore T') is Fredholm with index 0 if and only if $\mathcal{L}$ is is Fredholm with index 0, since the operator that corresponds to the second column on the right hand side of formula (44) is compact.

□

3.4.6 Proof of Theorem 2.3

Proof. The desired conclusions are immediate from the last three paragraphs of the proof of Theorem 3.9.

□

4 The H^∞ one disk (Nehari) problem

Our main inerest in this paper is the multidisk problem. Nevertheless, the Nehari problem is a good place to start, since it is a one disk problem and hence the calculations for this setting serve as a model for the multidisk problem.

4.1 Derivatives

Recall that for the Nehari problem

$$\Gamma(e^{i\theta}, Z) = (K(e^{i\theta}) \ - \ Z)^T(K(e^{i\theta}) \ - \ Z), \tag{56}$$

where Γ, K (and f) are $m \times m$ mvf's and $Z = (z_{ij})_{i,j=1}^m$ with $N = m^2$ independent entries. We shall need formulas for the first and second derivatives of Γ. In particular, it is readily seen that

$$a_{k\ell} := \frac{\partial}{\partial z_{k\ell}}\Gamma(e^{i\theta}, Z) = -(K(e^{i\theta}) \ - \ Z)^T E_{k\ell} \tag{57}$$

and

$$\frac{\partial^2}{\partial \overline{z}_{rs}\partial z_{k\ell}}\Gamma(e^{i\theta}, Z) = E_{rs}^T E_{k\ell} = E_{sr}E_{k\ell}, \tag{58}$$

where $E_{k\ell}$ is the $m \times m$ constant matrix with a one in the $k\ell$ position and zeros elsewhere.

Lemma 4.1 *The performance function $\Gamma(e^{i\theta}, f)$ for the Nehari problem is plurisubharmonic. It is strictly plurisubharmonic if and only if $m = 1$.*

Proof. By formula (58),

$$\sum_{r,s,k,\ell=1}^m u_{rs}^T \frac{\partial^2\Gamma}{\partial\overline{z}_{rs}\partial z_{k\ell}} u_{k\ell} = \sum_{r,s,k,\ell=1}^m u_{rs}^T E_{sr}E_{k\ell}u_{k\ell} = \| \sum_{k,\ell=1}^m E_{k\ell}u_{k\ell}\|^2$$

for every choice of the $N = m^2$ column vectors $u_{11}, \ldots, u_{mm}$ of size $m \times 1$. Therefore, Γ is plurisubharmonic. Moreover, if $m > 1$, then there exist nonzero vectors $u_{k\ell}$ such that $E_{k\ell}u_{k\ell} = 0$. Thus, Γ is strictly plurisubharmonic if and only if $m = 1$. $\square$

We remark that, since

$$E_{sr}E_{k\ell} = \begin{cases} 0 & \text{when} \quad r \neq k \\ E_{s\ell} & \text{when} \quad r = k, \end{cases} \tag{59}$$

the last sum is also equal to

$$\sum_{s,k,\ell=1}^m u_{ks}^T E_{sk}E_{k\ell}u_{k\ell} = \sum_{k=1}^m \| \sum_{\ell=1}^m E_{k\ell}u_{k\ell}\|^2$$

Moreover, much the same sort of analysis leads to the auxiliary conclusion that the $m^2 \times m^2$ matrix with entries

$$\mathrm{tr}\left\{\frac{\partial^2\Gamma}{\partial\overline{z}_{rs}\partial z_{k\ell}}\right\} = \mathrm{tr}\{E_{sr}E_{k\ell}\}$$

is positive definite, since

$$\sum_{s,r,k,\ell=1}^m \overline{\varphi_{rs}}\mathrm{tr}\{E_{sr}E_{k\ell}\}\varphi_{k\ell} = \sum_{k,\ell=1}^m |\varphi_{k\ell}|^2.$$

4.2 The optimality condition

In the Nehari case, the **MOPT** optimality conditions given in Theorem 1.1 can be stated as follows:

$$\begin{array}{lrcl} \text{(a)} & [\gamma I - (K-f)^T(K-f)]G^T & = & 0\ . \\ \text{(b)} & P_{H^2_{m\times m}}\{[K-f]G^TG\} & = & 0\ . \\ \text{(c)} & \frac{1}{2\pi}\int \mathrm{tr}\{G^TG\}d\theta & = & 1\ . \end{array} \tag{60}$$

Proof. (a) is immediate from Lemma 1.2 and (c) is obvious. Therefore, we turn to (b). For each k,ℓ we have

$$\begin{aligned} 0 &= P_{H^2}\mathrm{tr}[a_{k\ell}^T G^T G] \\ &= -P_{H^2}\mathrm{tr}[E_{k\ell}^T(K-f)G^TG] \\ &= -P_{H^2}[(K-f)G^TG]_{k\ell}. \end{aligned}$$

□

4.3 The null space of σ_{primal}

We turn now to the analysis of the null space of σ_{primal} in the Nehari case. For this analysis it is convenient to think of φ as an $m\times m$ matrix with entries $\varphi_{k\ell}, k,\ell = 1,\ldots,m$, rather than an $m^2\times 1$ column vector. Then

$$\varphi = \sum_{k,\ell=1}^{m} E_{k\ell}\varphi_{k\ell}. \tag{61}$$

Lemma 4.2 *In the setting of the Nehari problem, a matrix $B\in \mathbf{C}^{m\times m}$ is in the null space of σ_{primal} at the point $e^{i\theta}\in\mathbf{T}$ if and only if*

$$G(e^{i\theta})(K(e^{i\theta})-f(e^{i\theta}))^TB = 0 \quad \text{and} \quad G(e^{i\theta})B^T = 0. \tag{62}$$

Proof. If $\sigma_{primal}(B) = 0$, then the first statement in (62) follows directly from the first block row in the definition of σ_{primal} and formula (57). Next, the second block row of the formula for σ_{primal} implies that

$$\begin{aligned} 0 &= \langle AB, B\rangle \\ &= \sum_{r,s}\sum_{k,\ell}\mathrm{tr}\{G(e^{i\theta})\overline{B}_{rs}E_{sr}E_{k\ell}B_{k\ell}G(e^{i\theta})^T\} \\ &= \mathrm{tr}\{G(e^{i\theta})B^TBG(e^{i\theta})^T\}. \end{aligned}$$

But this clearly implies that B $G(e^{i\theta})^T = 0$ also, as claimed. $\square$

We remark that if $K(e^{i\theta}) - f(e^{i\theta})$ is invertible when f is a solution to **MOPT**, then an alternate way of writing the first condition in (62) at an optimum is

$$G(K-f)^{-1}B = 0\ .$$

This is because the optimality condition

$$G(e^{i\theta})(\gamma I_m - (K(e^{i\theta})\ -\ f(e^{i\theta}))^T(K(e^{i\theta})\ -\ f(e^{i\theta}))) = 0$$

can be rewritten as

$$\gamma G(e^{i\theta})\ =\ G(e^{i\theta})(K(e^{i\theta})\ -\ f(e^{i\theta}))^T(K(e^{i\theta})\ -\ f(e^{i\theta})).$$

But this implies that

$$G(e^{i\theta})(K(e^{i\theta})\ -\ f(e^{i\theta}))^T = \gamma G(e^{i\theta})(K(e^{i\theta})\ -\ f(e^{i\theta}))^{-1},$$

which gives the result.

4.4 The null space of σ_{dual}

In the setting of the Nehari problem, σ_{dual} maps the $k \times k$ matrix B into the $m \times m$ matrix with entries

$$\begin{aligned}
\mathrm{tr}\left\{G\frac{\partial\Gamma}{\partial z_{ij}}(\cdot,f)^T G^T B\right\} &= -\mathrm{tr}\left\{G\left[(K-f)^T E_{ij}\right]^T G^T B\right\} \\
&= -\mathrm{tr}\left\{GE_{ji}(K-f)G^T B\right\} \\
&= -\mathrm{tr}\left\{E_{ji}(K-f)G^T BG\right\} \\
&= -\left((K-f)G^T BG\right)_{ij},
\end{aligned}$$

for $i, j = 1, \ldots, m$. Thus, we have established the following result:

Lemma 4.3 *In the setting of the Nehari problem, a matrix $B \in \mathbf{C}^{k\times k}$ is in the null space of σ_{dual} at the point $e^{i\theta} \in \mathbf{T}$ if and only if*

$$(K(e^{i\theta}) - f(e^{i\theta}))G(e^{i\theta})^T BG(e^{i\theta}) = 0.$$

4.5 The $\mathcal{U}$ condition

The condition for $\mathcal{U}_f$ given by (13) to be invertible specializes as follows.

Lemma 4.4 *In the Nehari case, $\mathcal{U}_f^T(e^{i\theta})\mathcal{U}_f(e^{i\theta})$ is invertible if and only if $K(e^{i\theta}) - f(e^{i\theta})$ is invertible.*

Proof. Direct calculation gives

$$\mathcal{U}_f(e^{i\theta}) = -\Pi^T \mathrm{diag}\left\{\overline{K(e^{i\theta}) - f(e^{i\theta})}, \dots, \overline{K(e^{i\theta}) - f(e^{i\theta})}\right\}\Pi,$$

where the mvf inside the curly brackets on the right is block diagonal with m identical $m \times m$ blocks and Π is an $m^2 \times m^2$ permutation matrix. □

4.6 Conclusions for the Nehari case

The conclusions for the Nehari case are subsumed in the conclusions for the multidisk problem: just take $v = 1$. In particular, it should be noted that if $m > 1$, then the differential T' is not a Fredholm operator in this setting.

5 Multidisk MOPT

The Nehari problem prescribes a disk in matrix function space and seeks an analytic mvf which lies inside it. Now we specify v disks

$$\Gamma^p(e^{i\theta}, f(e^{i\theta})) = (K^p(e^{i\theta}) \; - \; f(e^{i\theta}))^T (K^p(e^{i\theta}) \; - \; f(e^{i\theta})), \qquad p = 1, \dots v. \tag{63}$$

Here Γ^p and f are $m \times m$ matrix valued functions and we seek an analytic function f and the smallest γ satisfying

$$\Gamma^p(e^{i\theta}, f(e^{i\theta})) \leq \gamma I_m \quad \text{for } p = 1, \dots, v \text{ and all } \theta.$$

5.1 Multiperformance MOPT

To any problem with multiple performance functions Γ^p, $p = 1, \dots, v$, we can associate a single block matrix (now of size $mv \times mv$)

$$\Gamma := \begin{pmatrix} \Gamma^1 & 0 & \dots & 0 \\ 0 & \Gamma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Gamma^v \end{pmatrix}$$

and consider the corresponding **MOPT** problem. Fortunately, the dual variable Ψ in the corresponding **PDMOPT** problem (which is formulated in subsection 1.2.6) can be taken block diagonal.

Theorem 5.1 *Let $\Gamma(e^{i\theta}, z)$ be a positive semidefinite mvf that meets the smoothness conditions specified in Section 1.2 and assume that the primal dual problem* **PDMOPT** *has a solution $(\Psi, f, \gamma) \in L^1_{vm \times vm} \times H^\infty_N \times \mathbf{R}$ such that:*

(1) $\Psi = \operatorname{diag}\,(\Psi^1, \dots, \Psi^v)$ *is block diagonal.*

(2) $\Psi^p = (G^p)^T G^p$ *has an outer spectral factor* $G^p \in (H^2_{k_p\times m})_+$ *with* $\operatorname{rank} G^p = k_p$ *a.e. for* $p = 1, \dots, v$.

(3) $\Psi^p \in L^2_{m\times m}$ *for* $p = 1, \dots, v$.

(4) f *is continuous and* $\gamma I - \Gamma(e^{i\theta}, f(e^{i\theta})) \geq 0$.

(5) *The conditions* $\sum_{p=1}^{v} \operatorname{tr}\left\{\frac{\partial \Gamma^p}{\partial z_j} d\mu^p\right\} = 0$, $\operatorname{tr} \sum_{p=1}^{v} \left\{\gamma I_m - \Gamma^p(e^{i\theta}, f((e^{i\theta}))\right\} d\mu^p = 0$ *and* $d\mu^p \geq 0$ *for* $p = 1, \dots v$ *imply that* $d\mu^p = 0$ *for* $p = 1, \dots v$.

Then G^p, f *and* γ *must satisfy the following conditions:*

(a) $G^1(\gamma I - \Gamma^1) = \cdots = G^v(\gamma I - \Gamma^v) = 0$.

(b) $P_{H^2_N} \operatorname{tr}\left\{(G^1)^T G^1 \frac{\partial \Gamma^1}{\partial z}^T + \cdots + (G^v)^T G^v\ \frac{\partial \Gamma^v}{\partial z}^T\right\} = 0$.

(c) $\sum\limits_{p=1}^{v} \int \operatorname{tr}(G^{p^T} G^p) d\theta = 2\pi$.

Proof. This theorem is an easy consequence of Theorem 1.1 and the special structure of the multidisk problem. □

It is also readily checked by straightforward calculation that the mvf $\mathcal{U}_f(e^{i\theta})$ for the full performance function Γ is simply related to the mvf's $\mathcal{U}^p_f(e^{i\theta})$ for the performance functions $\Gamma^p, p = 1, \dots, v$: The set of columns of $\mathcal{U}_f$ is equal to the union of the set of columns $\mathcal{U}^p_f,\ p = 1, \dots, v$, supplemented by zero columns. Thus,

$$\mathcal{U}_f \mathcal{U}_f^T = \sum_{p=1}^{v} \mathcal{U}^p_f (\mathcal{U}^p_f)^T. \tag{64}$$

However, this does not seem to be useful at the moment, since the columns of $\mathcal{U}_f(e^{i\theta})$ are linearly independent if and only if $\mathcal{U}_f(e^{i\theta})^T \mathcal{U}_f(e^{i\theta})$ is invertible.

5.2 The H^∞ multidisk problem

We now take the block diagonal entries in $\Gamma(e^{i\theta}, f)$ to be of the form (56). It is then convenient to picture both f and φ as $m \times m$ matrices with entries $f_{k\ell}$ and $\varphi_{k\ell},\ k, \ell = 1, \dots, m$, instead of a column vectors of height $m^2 \times 1$, just as in the Nehari case. Thus, upon writing

$$\Gamma = \Gamma(\cdot, Z) = \operatorname{diag}\Big\{(K^1 - Z)^T (K^1 - Z), \dots, (K^v - Z)^T (K^v - Z)\Big\}$$

with

$$Z = [z_{k\ell}],\ k, \ell = 1, \dots, m\ ,$$

it is readily checked that

$$\frac{\partial \Gamma}{\partial z_{k\ell}} = -\mathrm{diag}\Big\{(K^1 - Z)^T E_{k\ell}, \dots, (K^v - Z)^T E_{k\ell}\Big\} \tag{65}$$

and

$$\frac{\partial^2 \Gamma}{\partial \overline{z}_{rs} \partial z_{k\ell}} = \mathrm{diag}\Big\{E_{sr} E_{k\ell}, \dots, E_{sr} E_{k\ell}\Big\} = \mathcal{E}_{sr} \mathcal{E}_{k\ell}, \tag{66}$$

where

$$\mathcal{E}_{k\ell} = \mathrm{diag}\Big\{E_{k\ell}, \dots, E_{k\ell}\Big\}\ , \tag{67}$$

has v identical block diagonal entries. These formulas lead easily to

Lemma 5.2 *The performance function Γ for the H^∞ multidisk problem is plurisubharmonic. In fact*

$$\sum_{r,s,k,\ell=1}^{m} u_{rs}^T \frac{\partial^2 \Gamma}{\partial \overline{z}_{rs} \partial z_{k\ell}} u_{k\ell} = \| \sum_{k,\ell=1}^{m} \mathcal{E}_{k\ell} u_{k\ell} \|^2$$

for every choice of vectors $u_{k\ell} \in \mathbf{C}^{mv}$, $\quad k, \ell = 1, \dots, m.$

Lemma 5.3 *Let $d\mu = \mathrm{diag}\{d\mu^1, \dots, d\mu^v\}$ with nonnegative block matrix measures $d\mu^p$ of size $m \times m$. Then, in the H^∞ multidisk case,*

$$\mathcal{V}_f(e^{i\theta}, d\mu) = 0 \quad \textit{if and only if}$$

$$\sum_{p=1}^{v} \{K^p(e^{i\theta}) - f(e^{i\theta})\} d\mu^p = 0$$

and

$$\{\gamma I_m - \Gamma^p(e^{i\theta}, f(e^{i\theta}))\} d\mu^p = 0$$

for $p = 1, \dots, v$.

Proof In view of formula (15), the first N alias m^2 entries in the constraint $\mathcal{V}_f(e^{i\theta}, d\mu) = 0$ yields the sequence of formulas

$$\begin{aligned}
0 &= \operatorname{tr}\left\{\frac{\partial \Gamma}{\partial z_{rs}} d\mu\right\} \\
&= \sum_{p=1}^{v} \operatorname{tr}\left\{\frac{\partial \Gamma^p}{\partial z_{rs}} d\mu^p\right\} \\
&= -\sum_{p=1}^{v} \operatorname{tr}\left\{(K^p - f)^T E_{rs} d\mu^p\right\} \\
&= -\sum_{p=1}^{v} \operatorname{tr}\left\{E_{rs} d\mu^p (K^p - f)^T\right\} \\
&= -\sum_{p=1}^{v} \left\{d\mu^p (K^p - f)^T\right\}_{sr}
\end{aligned}$$

for $s, r = 1, \ldots, m$. But this leads easily to the first condition in the forward implication. The second condition is selfevident, as is the converse. $\square$

Proof of Theorem 1.5. Our first objective is to verify that Theorem 1.1 is applicable. The main effort is to translate condition (4). But that is done in the preceding lemma.

Next, invoking Theorem 1.1 and Lemma 1.2, we see that conditions (a) and (c) are obvious. To prove (b), note that the second optimality condition in Theorem 1.1 or Theorem 5.1 implies that

$$-\operatorname{tr}(G a_{rs} G^T) = \operatorname{tr}\left\{\sum_{p=1}^{v} G^{p^T} G^p (K^p - f)^T E_{rs}\right\} \in e^{i\theta} H^2 \tag{68}$$

for all $r = 1, \ldots, m$ and $s = 1, \ldots, m$. Thus $\operatorname{tr}[\beta E_{rs}] \in e^{i\theta} H^2$ which is equivalent to $\beta \in e^{i\theta} H^2_{m\times m}$. $\square$

5.2.1 The null space of σ_{primal}

Now we turn to an analysis of the null space of T' for the H^∞ multidisk problem. Recall from Theorem 2.3 that the key to understanding the null space of T' is an understanding of the operators σ_{primal} and σ_{dual}.

Lemma 5.4 *(Primal) In the H^∞ multidisk case, assume that K^p, G^p for $p = 1, \ldots, v$, and f are continuous on $\mathbf{T}$. Then a matrix $B \in \mathbf{C}^{m\times m}$ is in the null space of σ_{primal} at the point $e^{i\theta}$ if and only if*

$$G^p(e^{i\theta})(K^p(e^{i\theta}) - f(e^{i\theta}))B = 0 \quad \textit{and} \quad G^p(e^{i\theta})B^T = 0 \quad \textit{for} \quad p = 1, \ldots, v.$$

Thus,
$$\mu_{pnull}(\theta) = \dim\left\{B \in \mathrm{C}^{m\times m} : \sigma_{primal}(B) = 0 \quad at \quad e^{i\theta}\right\}.$$

Proof. For the multidisk problem,

$$\sigma_{primal}(B) = \begin{pmatrix} -\mathrm{diag}\left\{G^1(K^1-f)^T B, \ldots, G^v(K^v-f)^T B\right\} \\ \Sigma_{p=1}^{\mathrm{v}} G^{p^T} G^p B^T \end{pmatrix}.$$

Thus, $\sigma_{primal}(B) = 0$ if and only if $G^p(K^p - f)^T B = 0$ for $p = 1, \ldots, v$ and $\sum_{p=1}^{\mathrm{v}} G^{p^T} G^p B^T = 0$. The second equation is equivalent to $G^p B^T = 0$ for $p = 1, \ldots, v$. This completes the proof of the first assertion. The second follows immediately from the definition of $\mu_{pnull}(\theta)$. □

5.2.2 The null space of σ_{dual}

Lemma 5.5 *In the H^∞ multidisk case, let $B^p \in \mathbf{C}^{k_p\times k_p}$ for $p = 1, \ldots, v$. Then the null space of σ_{dual} is equal to zero providing that $B^p = O_{k_p\times k_p}$, $p = 1, \ldots, v$, is the only solution of the equation*

$$\sum_{p=1}^{v}\{K^p(e^{i\theta}) - f(e^{i\theta})\}G^p(e^{i\theta})^T B^p G^p(e^{i\theta}) = 0, \tag{69}$$

that is, $\mu_{dnull}(\theta) = 0$ for all θ.

Proof. Recall that σ_{dual} is regular provided the null space of σ_{dual} is 0 at each point $e^{i\theta} \in \mathbf{T}$. This is an elaboration of the calculation for the Nehari case. It is an easy consequence of the definition of σ_{dual} and formula (65). □

5.2.3 The $\mathcal{U}$ condition for the multidisk problem

In view of formula (64) and the calculations for the Nehari problem, it is readily seen that

$$\mathcal{U}_f\mathcal{U}_f^T = \sum_{p=1}^{v}\Pi^T\mathrm{diag}\Big\{\overline{(K^p-f)(K^p-f)^T}, \ldots, \overline{(K^p-f)(K^p-f)^T}\Big\}\Pi,$$

where the block diagonal matrix inside the curly brackets on the right has m identical $m\times m$ blocks and Π is an $m^2\times m^2$ permutation matrix. Thus, $\mathcal{U}_f(e^{i\theta})\mathcal{U}_f(e^{i\theta})^T$ is invertible if and only if $K^p(e^{i\theta}) - f(e^{i\theta})$ is invertible for $p = 1, \ldots, v$. However, if $v > 1$, then $\mathcal{U}_f(e^{i\theta})^T\mathcal{U}_f(e^{i\theta})$ is never invertible.

5.3 Proof of Theorem 1.8

Proof. The first four assumptions of this theorem guarantee that the assumptions imposed in Theorem 1.5 are met. Therefore, all the conclusions of the latter are in force. In view of Lemma 5.2 and the presumed smoothness we have **(PSCON)**. Then, with the aid of assumption (5), we obtain the first two items of **(SCOM)**. The third follows from the construction in Section 3.3.1. In particular, since $\Psi(e^{i\theta})$ is continuous with constant rank, the Moore-Penrose inverse $\Psi^{[-1]}(e^{i\theta})$ is also a continuous function of θ on the unit circle (cf [Ste:77]). Therefore, ρ is continuous and strictly positive definite on the circle. This guarantees the existence of the factorization $\rho^{-1} = Q^T Q$ with Q outer. However, in order to obtain a continuous factor Q, we need a little more:

By a general theorem that has been obtained in the work of J. Plemelj, N. I. Mushelisvili and N. P. Vekua, a sufficient condition for the continuity of Q is the Holder continuity of ρ on the circle. Therefore, since $\gamma I - \Gamma(e^{i\theta}, f(e^{i\theta}))$ is Holder continuous by assumption, it remains only to check that the Moore-Penrose inverse $\Psi^{[-1]}(e^{i\theta})$ of $\Psi(e^{i\theta})$ is Holder continuous. To this end, let $\mathcal{C}$ be a simple closed contour in the open right half plane that encircles the nonzero spectrum $\sigma'(\Psi(e^{i\theta}))$ of the positive semidefinite matrix $\Psi(e^{i\theta})$ for all $\theta \in [0, 2\pi]$ and is such that

$$|\lambda - \mu| \geq \delta > 0 \tag{70}$$

for every choice of $\lambda \in \mathcal{C}$ and $\mu \in \sigma'(e^{i\theta})$. The presumed continuity and constant rank of $\Psi(e^{i\theta})$ guarantees the existence of such a contour. Then it is readily checked that

$$\Psi^{[-1]}(e^{i\theta}) = \frac{1}{2\pi i} \int_{\mathcal{C}} \lambda^{-1} (\lambda I - \Psi(e^{i\theta}))^{-1} d\lambda$$

and hence that

$$\frac{d}{d\theta} \Psi^{[-1]}(e^{i\theta}) = \frac{1}{2\pi i} \int_{\mathcal{C}} \lambda^{-1} (\lambda I - \Psi(e^{i\theta}))^{-1} \Psi'(e^{i\theta}) (\lambda I - \Psi(e^{i\theta}))^{-1} d\lambda.$$

Therefore, since

$$\|(\lambda I - \Psi(e^{i\theta})^{-1}\| \leq \delta^{-1} \tag{71}$$

for $\lambda \in \mathcal{C}$, thanks to (70), the last formula leads easily to the bound

$$\left\| \frac{d}{d\theta} \Psi^{[-1]}(e^{i\theta}) \right\| \leq \frac{1}{2\pi} \delta^{-2} \|\Psi'(e^{i\theta})\| \int_{\mathcal{C}} \frac{|d\lambda|}{|\lambda|},$$

which in turn guarantees the Holder continuity of $\Psi^{[-1]}$. The fact that $H(e^{i\theta})$ has constant rank follows easily from the construction in Section 3.3.1.

Now we return to the main thread of the proof. Hypotheses (6) and (7) state that σ_{primal} and σ_{dual} are regular. Thus, by Theorem 2.3, T' is Fredholm of index zero. On the otherhand, if either (6) or (7) fails, then, by Proposition 3.8, the symbol M is not invertible for all θ and hence the operator T' is not Fredholm. $\square$

6 Supplementary Proofs

We begin this section with a proof of the optimality theorem on which this paper is based. The proof is a refinement of the proof of Theorem 2 of [HMW:98] (which also appears as Theorem 17.1.1 of [HMer:98]). Then we turn to the $G^T G$ factorization of Ψ and establish Lemma 1.2, which serves to rephrase condition (a) of Theorem 1.1 in terms of G. Subsequently, we prove a few specializations of Theorem 1.1 to the multi performance case and then, finally, justify Proposition 1.11.

6.1 Proof of Theorem 1.1

Equation (20.34) on page 217 of [HMer:98] implies that if γ, f is a local solution of the **MOPT** problem with $\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta})) \geq 0$, then

$$\mathrm{tr}\left\{\frac{\partial\Gamma}{\partial z_\ell}(\cdot, f)(d\lambda)^\tau\right\} = e^{i\theta}\varphi_\ell\frac{d\theta}{2\pi}\ , \qquad \ell = 1, \ldots, N\ , \tag{72}$$

where φ_ℓ belongs to the Hardy space H^1 and $d\lambda$ is an $m \times m$ nonnegative matrix valued measure on the circle such that

$$\int_0^{2\pi} \mathrm{tr}\{d\lambda\} = 1\ . \tag{73}$$

Consider the Radon-Lebesgue-Nikodym decomposition

$$(d\lambda)^\tau = \Psi\frac{d\theta}{2\pi} + d\mu\ , \tag{74}$$

where Ψ is an $m \times m$ mvf that is summable on the circle and $d\mu$ is an $m \times m$ matrix valued measure whose entries are singular with respect to Lebesgue measure. Note that since $d\lambda$ is nonnegative, so are $(d\lambda)^\tau$, Ψ and $d\mu$. Substituting (74) into (72) and matching absolutely continuous and singular measures, yields the relations

$$\mathrm{tr}\left\{\frac{\partial\Gamma}{\partial z_\ell}(\cdot, f)\Psi\right\} = e^{i\theta}\varphi_\ell \tag{75}$$

and

$$\mathrm{tr}\left\{\frac{\partial\Gamma}{\partial z_\ell}(\cdot, f)d\mu\right\} = 0\ . \tag{76}$$

Next, by (20.37) of [HMer:98] (corrected–by transposing one of the factors–and) adapted to the present setting, we have

$$\int_0^{2\pi} \mathrm{tr}\{[\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta}))](d\lambda)^\tau\} = 0 \tag{77}$$

and hence, upon invoking the decomposition (74) and taking advantage of the fact that $\gamma I_m - \Gamma(\cdot, f)$, Ψ and $d\mu$ are all positive semidefinite, we obtain

$$\int_0^{2\pi} \mathrm{tr}\{[\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta}))]\Psi(e^{i\theta})\}d\theta = 0 \tag{78}$$

and

$$\int_0^{2\pi} \mathrm{tr}\{[\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta}))]d\mu = 0. \tag{79}$$

But (76) and (79) imply that

$$\mathcal{V}_f(e^{i\theta}, d\mu) = 0$$

and hence, by assumption (4), $d\mu = 0$. On the other hand, formula (78) implies that

$$\mathrm{tr}\{[\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta}))]\Psi(e^{i\theta})\} = 0$$

a.e. on the circle. Therefore, since both of the factors in the product are positive semidefinite, the full matrix

$$[\gamma I_m - \Gamma(e^{i\theta}, f(e^{i\theta}))]\Psi(e^{i\theta}) = 0$$

a.e. on the circle. This is conclusion (a) of (19). Condition (c) of (19) drops out easily from formulas (73) and (74) and the vanishing of $d\mu$.

Finally, (75) implies that

$$\mathrm{tr}\left\{e^{-i\theta}\frac{\partial \Gamma}{\partial z_j}(\cdot, f)\Psi\right\} \in H^1 \cap L^2, \qquad j = 1, \ldots, N.$$

Therefore, by the Smirnov maximum principle,

$$\mathrm{tr}\left\{e^{-i\theta}\frac{\partial \Gamma}{\partial z_j}(\cdot, f)\Psi\right\} \in H^2, \qquad j = 1, \ldots, N.$$

But this is easily seen to be equivalent to condition (b) of (19). $\square$

6.2 Proof of Lemma 1.2

If $(\gamma I_m - \Gamma(\cdot, f))G^T G = 0$, then $(\gamma I_m - \Gamma(\cdot, f))G^T G(\gamma I_m - \Gamma(\cdot, f)) = 0$, which implies that $G(\gamma I_m - \Gamma(\cdot, f)) = 0$. Therefore, (1) $\Rightarrow$ (3) $\Rightarrow$ $P_{H^2_+}[G(\gamma - \Gamma(\cdot, f)] = 0$. Conversely, if this last condition holds, then $G(\gamma I_m - \Gamma(\cdot, f)) = F$, with $F \in H^{2\perp}_+$. Thus,

$$G^T G(\gamma I_m - \Gamma(\cdot, f)) = G^T F$$

and

$$\frac{1}{2\pi}\int \mathrm{tr}\{G^T G(\gamma I - \Gamma(\cdot, f))\}d\theta = \frac{1}{2\pi}\int \mathrm{tr}\{G^T F\}d\theta = \langle F, G\rangle = 0,$$

since $F \in (H^2_+)^\perp$ and $G \in H^2_+$. But the trace of the product of nonnegative matrices is nonnegative. Since the trace of the product integrates to zero, it must be zero for almost all θ. Now the product of nonnegative matrices has trace zero if and only if the product itself is zero. Therefore

$$G^T G(\gamma I_m - \Gamma(\cdot, f)) = 0.$$

This proves that (2) $\Rightarrow$ (1) and serves to complete the proof. $\square$

Now we turn to multi-performance **MOPT**. In the setting of the multi performance **MOPT** that was introduced in Section 5.1, it is readily checked that the conditions in (19) can be reexpressed in terms of the positive semidefinite diagonal blocks $\Psi^p \in L^1_{m\times m}$ of Ψ and the performance functions Γ^p, $p = 1, \ldots, v$, as follows:

There exist a set of positive semidefinite mvf's $\Psi^p \in L^1_{m\times m}$ such that

$$\begin{array}{rcl} \Psi^1(\gamma I - \Gamma^1) \quad = \cdots = \Psi^v(\gamma I - \Gamma^v) & = & 0 \\ \mathrm{tr}\left\{\Psi^1 \frac{\partial \Gamma^1}{\partial z_j}\right\} \quad + \cdots + \mathrm{tr}\left\{\Psi^v \ \frac{\partial \Gamma^v}{\partial z_j}\right\} & \in & e^{i\theta} H^1 \\ \sum_{p=1}^{v} \int \mathrm{tr}(\Psi^p) d\theta \qquad & = & 2\pi \end{array} \tag{80}$$

6.3 Proof of Theorem 5.1

The proof is an immediate consequence of Lemma 1.2 and Theorem 1.1, specialized to the H^∞ multidisk case. Formula (80) translates the conclusions of that theorem to the present setting. $\square$

6.4 Proof of Proposition 1.11

In order to prove Proposition 1.11 we need a lemma.

Lemma 6.1 *Let L^p and R^{p^T} be $m \times k_p$ matrices of* rank *k_p, for $p = 1, \ldots, v$ and suppose that at least one of the sets $\{L^1, \ldots, L^v\}, \{R^{1T}, \ldots, R^{vT}\}$ has linearly independent ranges. Then:*

1. *The only matrices $B^p \in \mathbf{C}^{k_p \times k_p}, p = 1, \ldots, v$, which satisfy the condition*

$$\sum_{p=1}^{v} L^p B^p R^p = 0 \tag{81}$$

are the matrices $B^p = 0$ for $p = 1, \ldots, v$.

2. If

$$\sum_{p=1}^{v} k_p = m,$$

then the only matrix $C \in \mathbf{C}^{m\times m}$ that meets the conditions

$$R^p C^T = 0 \quad and \quad L^{p^T} C = 0 \quad for\ p = 1, \ldots, v,$$

is the matrix $C = 0$.

We remark that the formulation of part (2) of this lemma is a little deceptive. At first glance it appears that the conditions $R^p C^T = 0$ and $L^{p^T} C = 0$ are reenforcing each other. In fact they are invoked independently, according to which of the matrices R or L (that are defined below in the proof) is invertible.

Proof of Lemma 6.1

Proof of (1). Let

$$L = \left[L^1 \cdots L^v\right], \quad R^T = \left[R^{1T} \cdots R^{p^T}\right]$$

and

$$B = \mathrm{diag}\{B^1, \ldots, B^v\}.$$

Then the condition (81) is the same as to say that

$$L\ B\ R = 0.$$

Now, if the set L^p, $p = 1, \ldots, v$, has linearly independent ranges, then L is left invertible. Therefore

$$B\ R = 0,$$

or, what is the same,

$$B^p R^p = 0 \quad \text{for} \ \ p = 1, \ldots, v.$$

But this in turn implies that $B^p = 0$, since the presumed maximal rank condition implies that the R^p are all right invertible. This completes the proof of (1), when the $L^p, p = 1, \ldots, v$ have linearly independent ranges.

If the $R^{p^T}, p = 1, \ldots, v$, have linearly independent ranges, then R is right invertible and hence the condition $LBR = 0$ leads to $LB = 0$, i.e., $L^p B^p = 0$ for $p = 1, \ldots, v$, which again implies that $B^p = 0$ for all p.

Proof of (2). If the $L^p, p = 1, \ldots, v$, have linearly independent ranges, then the assumption $\sum\limits_{p=1}^{v} \mathrm{rank} L^p = m$ implies that L is an invertible matrix and the condition

$$L^{p^T} C = 0, p = 1, \ldots, v,$$

implies that

$$L^T C = 0,$$

which clearly forces $C = 0$. On the other hand, if the R^{p^T} have linearly independent ranges, then the asumption $\sum_{p=1}^{v} \text{rank} R^{p^T} = m$ implies that R^T is an invertible matrix and the condition

$$R^p C^T = 0,\ p = 1, \ldots, v,$$

implies that

$$RC^T = 0.$$

Therefore, $C = 0$ in this case also. $\square$

Proof of Proposition 1.11. Since $\Psi^p(e^{i\theta}) \geq 0$, the constraints on B in Definition 1.6 are also valid if Ψ^p is replaced by $(\Psi^p)^{\frac{1}{2}}$. They imply that at each point $e^{i\theta} \in \mathbf{T}, B$ maps the orthogonal complement of

$$\text{range}\{\Psi^1(e^{i\theta})\} + \cdots + \text{range}\{\Psi^v(e^{i\theta})\} = \text{range}\{\Psi^1(e^{i\theta}) + \cdots + \Psi^v(e^{i\theta})\}$$

into the orthogonal complement of

$$\text{range}\{(K^1 - f)\Psi^1(K^1 - f)^T + \cdots + (K^v - f)\Psi^v(K^v - f)^T\}(e^{i\theta})\ .$$

These spaces have dimension $\mu_D(\theta)$ and $\mu_R(\theta)$, respectively, which serves to prove the first formula in the proposition.

The independence of $\mu_{pnull}(\theta)$ from θ is true because rank $\Psi^p(e^{i\theta})$ is independent of θ for almost all θ, since Ψ^p has the analytic factor G^p.

Next, general position implies that

$$\mu_D(\theta) = m - \sum_{p=1}^{v} \text{rank} \Psi^p(e^{i\theta})$$

and

$$\begin{aligned} \mu_R(\theta) &= m - \sum_{p=1}^{v} \text{rank}(K^p(e^{i\theta}) - f(e^{i\theta}))\Psi^p(e^{i\theta})(K^p(e^{i\theta}) - f(e^{i\theta}))^T \\ &= m - \sum_{p=1}^{v} \text{rank} \Psi^p(e^{i\theta}). \end{aligned}$$

Thus general position implies the second formula.

That $\mu_{pnull}(\theta) = 0 = \mu_{dnull}(\theta)$ follows immediately from part (2) and part (1) of Lemma 6.1, respectively. $\square$

7 References

[AHO:96] F. Alizadeh, J.A. Haeberly and M.L. Overton, *"Primal-Dual Interior-Point Methods for Semidefinite Programming: Convergence Rates, Stability and Numerical Results"*, *NYU Computer Science Department Technical Report 721*, May 1996.

[BS:90] A. Bottcher and B. Silbermann, Analysis of Toeplitz Operators. Springer Verlag, Berlin, 1990.

[BuG:68] M.S. Budjanu and I.C. Gohberg, General theorems on the factorization of matrix-valued functions. I. Fundamental Theorem. Mat. Issled. **3** (1968), no. 2, 87-103; English Transl. Amer. Math. Soc. Translations, (2) **102** (1973), 1-14.

[GoF:74] I. C. Gohberg and I. A. Feldman, Convolution Equations and Projection Methods for their Solution, Transl. Math. Monographs, vol. 41, American Mathematical Society, Providence, 1974.

[DHM:99] H. Dym, J. W. Helton, and O. Merino, Algorithms for solving multidisk problems in H^∞ optimization, Conference on Decision and Control 1999

[HO:94] J.-P.A. Haeberly and M. L. Overton, Optimizing Eigenvalues of Symmetric Definite Pencils. Proceedings of the American Control Conference, Baltimore, July 1994.

[Hel:64] H. Helson, Lectures on Invariant Subspaces. Academic Press. 1964

[Hel:87] J. W. Helton, Operator Theory, Analytic Functions, Matrices, and Electrical Engineering. CBMS Regional Conference Series in Mathematics. American Mathematical Society. Providence, 1987.

[HMer:93] J. W. Helton and O. Merino, Conditions for Optimality over $\mathcal{H}^\infty$. SIAM J. Cont. Opt., Vol. 31, No. 6, November 1993.

[HMer:98] J. W. Helton and O. Merino, "Classical Control using H^∞ Methods: Theory, Optimization and Design". SIAM, Philadelphia, 1998.

[HMW:93] J. W. Helton, O. Merino and T. Walker, Algorithms for Optimizing over Analytic Functions, Indiana U. Math. J., Vol. 42, No. 3 (1993), 839 - 874.

[HMW:98] J. W. Helton, O. Merino and T. Walker, H^∞ Optimization with Plant Uncertainty and Semidefinite Programming, pp1-41 I Jour Nonlin Robust Cont. vol 8, 763-802 (1998).

[K:52] L. V. Kantorovich, Functional Analysis and Applied Mathematics, Uspekhi Mat. Nauk, **3**(1948), pp. 89-185. Translated by C. Benster as Nat. Bureau of Standards Report IS09, Washington, D. C. 1952.

[K:82] S. Krantz, Function Theory of Several Complex Variables. Wiley. 1982

[OZ:93] J. G. Owen and G. Zames, Duality theory for MIMO robust disturbance rejection IEEE TAC, May 1993 (38) No. 5, pp 743 – 752.

[SI:95] Robert E. Skelton and T. Iwasaki, *Increased Roles of Linear Algebra in Control Education*, IEEE Control Systems, 8/95, pp. 76–90.

[Ste:77] G. W. Stewart, On the perturbation of pseudo-inverses, projections and and linear least square problems, SIAM Review, **19**, (1977) 634–662.

[VB:96] L. Vandenberghe and S. Boyd, Semidefinite Programming, SIAM Review, Vol 38 (1996), no.1, 49-95.

[Wr:98] S.Wright, Primal Dual Optimization. SIAM. 1998

[ZTD:92] Y. Zhang, R. A. Tapia, and J. E. Dennis *On the super linear and quadratic primal dual interior point linear programming algorithms SIAM Journal On Optimazation*, **2**(1992), pp. 304-324.