Optimal recovery of holomorphic functions from inaccurate information about integration type operators

Arthur James Degraw

University at Albany, State University of New York, ad971334@albany.edu

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Physical Sciences and Mathematics Commons

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/540

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
OPTIMAL RECOVERY OF HOLOMORPHIC FUNCTIONS FROM INACCURATE INFORMATION ABOUT INTEGRATION TYPE OPERATORS

by


A Dissertation
Submitted to the University at Albany, State University of New York
in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

College of Arts & Sciences
Department of Mathematics and Statistics
2012
Contents

1 Introduction 1
  1.1 Historical development ........................................ 1
  1.1.1 Structure of the dissertation ................................ 8

2 Statements of Results 10
  2.1 Radial Integral Information Operator .......................... 10
     2.1.1 Inaccuracy in $L_2(T)$ Norm ............................... 11
     2.1.2 Inaccuracy in $l_2^N(T)$ Norm ............................. 13
     2.1.3 Varying Levels of Accuracy Termwise .................... 15
     2.1.4 Applications: Hardy–Sobolev and Bergman–Sobolev Spaces ... 18
  2.2 Secant Integral Information Operator .......................... 26
     2.2.1 Applications: Hardy–Sobolev and Bergman–Sobolev Spaces ... 30
  2.3 Directions For Further Study .................................... 31

3 Background Material 33
  3.1 General Setting .................................................. 37
  3.2 Construction of Optimal Method and Error ..................... 40
  3.3 Dual Problem Equivalence ...................................... 44

4 Proofs of Results 48
  4.1 Radial Integral Information Operator Problems ............... 48
4.1.1 Theorem 1: Inaccuracy in $L_2(\mathbb{T})$ Norm  

4.1.2 Theorem 2: Inaccuracy in $l^N_2(\mathbb{T})$ Norm  

4.1.3 Theorem 3: Varying Levels of Accuracy Termwise  

4.2 Secant Integral Information Operator Problem  

5 Examples  

5.1 Radial Information  

5.2 Secant Information
ABSTRACT

This paper addresses the optimal estimation of functions from Hilbert spaces of functions on the unit disc. The estimation, or recovery, is performed from inaccurate information given by a linear information operator. The information operators considered are of integration type, along radial and secant paths. The results are applied to the Hardy–Sobolev and Bergman–Sobolev classes.
Acknowledgements

The journey of this dissertation has not taken the most direct route nor been without the occasional bump. Many people have guided me in the right direction, redirected me when I got lost, and helped me to overcome bumps in the road. I would like to specifically thank some of those people.

I have had many professors during my graduate studies that have given their time and effort to ensure that my education was successful. Now teaching full time myself, I recognize the amount of dedication that is needed to make each class a fruitful learning experience. I thank each professor for their time both in and out of the classroom.

This dissertation could not be complete without tremendous support and feedback from two advisors: Dr. Konstantin Osipenko and Dr. Michael Stessin. This research began while Dr. Osipenko visited the university and continued long after both he and I left. Through Dr. Osipenko’s insight into my many questions, his ability to direct my studies from afar, as well as his extraordinary patience and persistant motivational prompts, this research was possible while we both resided on different continents. Through the revision and publication process both Dr. Stessin’s and Dr. Osipenko’s guidance has been invaluable. Sometimes life tells us where we need to go instead of the other way around. Without Dr. Osipenko and Dr. Stessin patiently working with me to complete my studies in absentia, none of this would have been possible.

On this path I have made many valuable friends along the way that warrant my gratitude. Many due thanks to Arthur Lubovsky for always saying I would finish, even when I had my own doubts. To Stacy Newman, for finding a way to keep the research within school parameters and her always open door to lend an ear in rough times. She never failed to find an answer for a confused graduate student. If not for Stacy’s friendship and desire to see me continue my studies I may have taken the all-too-easy route.

Outside the walls of the university, my family is my support. Without them, I never would have made it this far. I miss you all.
Lastly, I want to thank my wife, Desma. Sometimes life throws curve balls but you have taught me to always keep swinging for the fences. Your encouragement and refusal to accept limitations have pushed me further than I ever thought I could go. Life’s an adventure.

Arthur DeGraw, September 2012
Chapter 1

Introduction

1.1 Historical development

Optimal recovery theory developed from the ideas of quadrature formulas. In the 1950’s, with the advent of basic computing machines, numerical integration techniques became important. Quadrature formulas deal with the numerical approximation of integrals. A basic scenario is as follows: Let \( f \) be defined on a closed interval \([a, b] \subseteq \mathbb{R}\), estimate the value of

\[
\int_{a}^{b} f(x) \, dx
\]  

(1.1)

from the values of \( f \) at some distinct points in \([a, b]\). To compute or approximate 1.1, quadrature formulas have been developed. Typically, one first chooses, or is given, a set of nodes \( x_1, \ldots, x_n \in [a, b] \) at which \( f \) is known precisely. A quadrature formula will use the values \( f(x_j) \) for \( j = 1, \ldots, n \) to approximate the integral. As a simple example, if \( x_1 = a, \, x_n = b \) then the left endpoint summation of basic Integral Calculus is given by the approximation

\[
\int_{a}^{b} f(x) \, dx \approx \sum_{j=1}^{n-1} c_j f(x_j)
\]

with \( c_j = (x_{j+1} - x_j) \).
Various quadrature formulas have been developed for numerical integration. One form of the Gaussian quadrature formula is a method to approximate \( \int_{-1}^{1} f(x) \, dx \) for \( f : [-1, 1] \to \mathbb{R} \). The information that will be available is a sampling of the function \( f \) at some nodes: \( f(x_1), f(x_2), \ldots, f(x_n) \) where \(-1 \leq x_1 < x_2 < \ldots < x_n \leq 1\). Using the Lagrange interpolating polynomial through the points \( x_1, \ldots, x_n \):

\[
P(x) = \sum_{j=1}^{n} f(x_j) \prod_{k=1, k \neq j}^{n} \frac{x - x_j}{x_k - x_j}
\]

satisfies \( P(x_j) = f(x_j) \). Then

\[
\int_{-1}^{1} P(x) = \sum_{j=1}^{n} w_j f(x_j)
\]

where the problem is to optimally choose the set of nodes \( x_j \) which will yield the weights \( w_j \), assuming that equality holds for polynomials up to order \( 2n-1 \). The approximation problem amounts to solving a system of \( 2n \) equations with the \( 2n \) unknowns, \( x_1, \ldots, x_n, w_1, \ldots, w_n \). For example, let \( f(x) = x^3 - 2x^2 \) and \( n = 2 \). Then the goal is to determine \( x_j, w_j, j = 1, 2 \) such that the error of the approximation

\[
\int_{-1}^{1} f(x) \, dx \approx w_1 f(x_1) + w_2 f(x_2)
\]

is minimal. Assuming that equality holds for polynomials up to order \( 2n-1=3 \), the values of \( x_j, w_j \) are solutions to the system

\[
\begin{align*}
\int_{-1}^{1} 1 \, dx &= 2 = w_1 + w_2, & \int_{-1}^{1} x \, dx &= 0 = w_1 x_1 + w_2 x_2 \\
\int_{-1}^{1} x^2 \, dx &= \frac{2}{3} = w_1 x_1^2 + w_2 x_2^2, & \int_{-1}^{1} x^3 \, dx &= 0 = w_1 x_1^3 + w_2 x_2^3.
\end{align*}
\]
This system will have solutions \( x_1 = 1/\sqrt{3}, x_2 = -1/\sqrt{3}, w_1 = w_2 = 1 \). By evaluating the summation

\[
w_1 f(x_1) + w_2 f(x_2) = \frac{1}{3^{3/2}} - \frac{2}{3} + \frac{1}{-3^{3/2}} - \frac{2}{3} = -\frac{4}{3}
\]

and the integral

\[
\int_{-1}^{1} x^3 - 2x^2 \, dx = \frac{1}{4}x^4 - \frac{2}{3}x^3 \bigg|_{x=-1}^{1} = -\frac{4}{3}
\]

we can see the approximation is exact. For the class of polynomial functions of degree at most 2n-1, the Gaussian quadrature method will give the precise value of the integral.

The precision of any quadrature formula approximation will depend heavily on the class of functions being considered. Let us restrict our function domain to the unit interval \( I = [0, 1] \). If the integrand could be any function \( f \) in a class \( X \) of functions, a simple way to define the error in the approximation \( A(f) \) to \( \int_I f \) is

\[
\sup_{f \in X} |A(f) - \int_I f|.
\]

(1.2)

Some classes of functions are easier to approximate than others. As a particularly poor example consider the class of continuous functions \( C(I) \) with \( |f| \leq 1 \). Let it also be given, \( f(x_j) = 0 \) for \( j = 1, \ldots, n \). For any \( \epsilon > 0 \) there are functions \( f_1, f_2 \in C(I) \) with \( \int_I f_1 > 1 - \epsilon \) and \( \int_I f_2 < -1 + \epsilon \) and so the error given by 1.2 of any approximation is at least 1.

Integration is an example of a linear functional, that is a mapping from a vector space to the base field. From the ideas of quadrature formulas, in which the approximation to an integral is obtained, the generalization to approximation of general linear functionals begins.

Two groups of researchers, Ehrenpreis et. al. at International Business Machines, and the famous seminars of Luzin at the Moscow State University, began to consider questions
of approximation and the generalizations of quadrature notions. The group of Luzin lead
to several important results in the development of the field of Optimal Recovery, which
included those of A. Sard [28], S. M. Nikolskij [17], S.A. Smolyak [29] and N.S. Bakhvalov
[1].

In [28], Sard considered the approximation to integrals of functions by linear methods
where the value of the integrand was known precisely at a finite number of fixed points.
In his 1949 paper, Sard also discusses the extension of his results to any linear functional.
Nikolskij [17] with the setting of Sard allowed the points of evaluation to be optimally
chosen. In a sense, Nikolskij, was showing a construction of how to choose optimal
information.

For the results of Sard and Nikolskij, only linear methods of recovery were considered.
It became of particular interest to determine under what conditions a linear method will
be among the best methods to use for approximation. In his 1959 dissertation, Smolyak
[29] was able to show that there is a linear method amongst the optimal methods to
approximate a real linear functional from values of a finite collection of linear functionals
whose values are known precisely. The proof of his result is based on the idea of if
given a balanced convex set, one can construct a set of support hyperplanes. The first
hyperplane of support is constructed to intersect the boundary of a convex set by choosing
the supremum over all values of the linear functional to be estimated. The hyperplane is
a linear combination of the \( n \) linear functionals. The proof then demonstrates that since
the set is balanced and convex an antipodal hyperplane of support can be constructed
thus trapping the convex set between the hyperplanes. To conclude, Smolyak is able to
show that the distance from any point in the convex set to one of the hyperplanes is at
most the error of an arbitrary method of recovery. Since the hyperplanes are constructed
as linear functions of the values of \( n \) linear functionals the result is obtained.

In [2], Bakhvalov was able to further the ideas of Smolyak to linear operators on convex
classes of functions. He was able to show that on convex classes of functions, the error
when considering only linear methods of recovery was not greater than when the methods were unrestricted.

For the recovery settings described, the information about the function, such as the values at a system of nodes, or values of linear functionals, are known precisely. The introduction of a perturbation in the values of the information was introduced in [13] by A. G. Marchuk and K. Yu. Osipenko. In this paper the authors solved the problem of approximating a linear functional on classes of functions that are convex and centrally symmetric when the information is given by the imprecise values of $n$ linear functionals.

The extension of Smolyak’s work to the complex case is due to K. Yu. Osipenko, [18], in 1976. In this paper Osipenko was able to show the existence of a linear method amongst the optimal methods of recovery for approximating a linear complex functional.

Research into Optimal Recovery conducted at IBM by C.A. Micchelli and T.J. Rivlin lead to their survey paper [15] in 1977. In this paper the authors lay the framework and terminology into which many of the problems of optimal recovery can be stated. Their focus is on the optimal methods for recovery of linear problems where the information is given by a linear operator.
The use of duality in approximation problems had been implemented since the 1940s. For convex optimization problems there exists a dual problem. Solutions to the dual problem are closely related to solutions to the original optimization problem, however the dual problem is often easier to solve. In 1979, Melkman and Micchelli [14] addressed the problem of optimal recovery of linear operators on Hilbert spaces by linear methods from inaccurate information and a duality relation was developed. This result concerned the coincidence of the extremal elements for the dual and mixed dual problems. With this coincidence, a linear optimal method of recovery will exist.

In the 1950’s and 1960’s, as the field of Optimal Recovery began to develop so too did the efficiency of computing machines. Not only was accuracy in the approximation critical but so was the computational efficiency of an algorithm for recovery. Computational complexity is an important area of approximation theory (however this paper does not investigate these topics). In their monograph [30], J.F. Traub and H. Wozniakowski, develop the ideas of \( \epsilon \)-complexity. The idea is to determine the computational cost of obtaining an estimate that is within \( \epsilon > 0 \) of the actual value. For instance, considering Gaussian quadrature again, if it is desired that

\[
| \int_{-1}^{1} f(x) dx - \sum_{j=1}^{n} w_j f(x_j) | < \epsilon
\]

the computational cost of the recovery algorithm \( \sum_{j=1}^{n} w_j f(x_j) \) is in computing, \( x_j, f(x_j) \) and \( w_j \) for \( j = 1, \ldots, n \) and large enough \( n \).

Through the early 1990s work was completed by Osipenko and Stessin in the non-Hilbert space setting. Specifically, [23] addresses recovery of the evaluation of the derivatives of functions belonging to the Hardy spaces, \( H_p \), and Bergman spaces, \( A_p \), on the unit disc in \( \mathbb{C} \) for \( 1 \leq p < \infty \). The extremal elements found in the Bergman case, which are contractive zero-divisors, played a role in the factorization theory in Bergman spaces (see [7]). In the following year, [24] obtained results for the optimal recovery of a linear
functional on $H_p$ and $A_p$, where each space is generalized to functions in the unit ball of $\mathbb{C}^n$. It was demonstrated that the optimal method of recovery was related to weighted reproducing kernels. Obtaining results about the zeros of the kernels and representing the kernels as hypergeometric functions helped Weir to obtain his result on factorization in weighted Bergman spaces. In addition, in 1992 Osipenko and Stessin [25] obtained results for optimal recovery of an operator from inaccurate data in the spaces $H_p$ for arbitrary $1 < p < \infty$ including non-Hilbert cases.

Recently, in 2010, Osipenko and Stessin [22] introduced the idea of a spectral function and the spectrum of the dual problem. If the dual problem has a solution then the spectrum, dependent on the solution, will be non-empty. Furthermore, the error of optimal recovery of a linear operator can be found by solving an extremal problem involving the spectrum and it was found that the spectrum of the dual problem is related to the spectrum of Toeplitz operators.

Problems that can be stated in the framework of optimal recovery are many and varied. This is partially due to the ability to state recovery of general functionals and operators on linear spaces in this context. Problems involving best approximation of differentiation, integration, series expansions, solutions to differential and integral equations, interpolation, quadrature formulas and more can be stated in the broadest sense in the ideas of optimal recovery.

One particular area of application is in the calculation of $n$-widths. Consider the best approximation of a set $W$ by an $n$-dimensional subspace $X_n$, in a linear space $X$, with norm $\| \cdot \|_X$. The Kolmogorov $n$-width of $W$ is defined to be the value

$$d_n(W, X) = \inf_{X_n} \sup_{x \in W} \inf_{y \in X_n} \| x - y \|_X.$$  \hspace{1cm} (1.3)

For the linear $n$-width, let $L_n$ be the set of all bounded linear operators on $Y$ whose image
is a subspace of $X$ of at most dimension $n$. Then the linear $n$-width is given by

$$s_n(W, X, Y) = \inf_{L \in L^*_n} \sup_{f \in W} \| f - Lf \|_X. \quad (1.4)$$

The expression 1.4 is the best approximation of a function $f$ by a linear operator $L$. Micchelli and Fisher [8] were able to demonstrate that in some cases 1.3 and 1.4 are equal thus making the problem of $n$-widths into one of optimal recovery.

Another application is to the Fourier series expansion of a real valued function $f(x)$ given by $f(x) = \sum_{j=-\infty}^{\infty} c_ne^{inx}$. The coefficients, $c_n$ are obtained from

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)e^{-inx}dx.$$

The estimation of Fourier coefficients with a priori information given by another linear operator can be stated in an optimal recovery setting as the Fourier Transform is a linear operator. Reversing the roles of what is information and what is to be recovered, the Fourier coefficients may be used as the a priori information in an optimal recovery problem such as estimation of a function or its derivatives (see [10], [27].)

1.1.1 Structure of the dissertation

The second chapter of this dissertation presents the statements of the main results. The results are concerned with the optimal recovery of classes of analytic functions with various integration type information operators. The spaces of analytic functions are a natural setting for optimal recovery theory and are ideal to work with as each function may be represented as a series. The integration type operators that are considered are of line integrals along radial and secant paths. Connections with imaging processes and image reconstruction are discussed. Applications to the Hardy–Sobolev and Bergman–Sobolev classes of functions are demonstrated as a direct consequence of the main results and potential directions for further study are presented. Chapter three develops the mathe-
matical background of the results in chapter 2. A proof of the result contained in the dissertation of Smolyak is presented. A general setting for problems of optimal recovery of linear operators from inaccurate information is constructed as well as several specific theorems that will be used in the proofs of the results of chapter 2. Chapter four contains the proofs of results stated in chapter 2. The final chapter looks at the application of the results to the original question that was considered at the onset of research. Chapter five shows how the results are applied to a particular class of functions and provides a construction of an optimal method under the settings for the different integral operators, computes the optimal error, and graphically demonstrates the accuracy of the method.
Chapter 2

Statements of Results

The problems stated in this chapter involve information from integral type operators of two distinct types. Each integral operator is defined for functions on the unit disc $\mathbb{D} \subset \mathbb{C}$. The first type is a radial integral and the other a secant integral that is essentially the Radon transform. These results are then applied to the Hardy–Sobolev and Bergman–Sobolev spaces of functions over $\mathbb{D}$.

2.1 Radial Integral Information Operator

Consider the class of functions defined on the unit disc given by

$$X = X_\gamma = \left\{ f(z) = \sum_{j=0}^{\infty} a_j z^j : \sum_{j=0}^{\infty} \gamma_j |a_j|^2 < \infty \right\}$$

for $\gamma_j \geq 0$ satisfying

$$\lim_{j \to \infty} \gamma_j^{1/j} \geq 1$$

and

$$\lim_{j \to \infty} (\gamma_j(j + 1))^{-1} = 0.$$
Therefore, any \( f \in X \) is holomorphic in the unit disc by (2.2). We define the semi-norm in \( X \) as

\[
\|f\|_X = \left( \sum_{j=0}^{\infty} \gamma_j |a_j|^2 \right)^{1/2}
\]

and

\[
W = \{ f \in X : \|f\|_X \leq 1 \}
\]

be the unit ball in \( X \).

2.1.1 Inaccuracy in \( L_2(\mathbb{T}) \) Norm

Let \( K : X \to L_2(\mathbb{T}), \mathbb{T} = [-\pi, \pi] \), be a linear operator given by

\[
Kf(\phi) = \int_0^1 f(re^{i\phi}) dr.
\]

That is, \( Kf \) is the radial integral of \( f \). To see that \( Kf \in L_2(\mathbb{T}) \), by (2.3) we have for all but finitely many \( j \), \( \gamma_j \geq \frac{c}{(j+1)^2} \) for some \( c > 0 \). Thus if \( \|f\|_X < \infty \) then \( \|Kf\|_{L_2(\mathbb{T})} < \infty \).

We assume to know \( Kf(\phi) \) given with a level of accuracy. That is, for a given \( \delta > 0 \), we know a \( \tilde{y} \in L_2(\mathbb{T}) \) such that

\[
\|Kf - \tilde{y}\|_{L_2(\mathbb{T})} \leq \delta.
\]

The problem of optimal recovery is to find an optimal recovery method of the function \( f \) in the class \( W \) from the information \( \tilde{y} \) satisfying (2.5). The error of a given method is measured in the \( L_2(\mathbb{D}) \) norm defined by

\[
\|f\|_{L_2(\mathbb{D})} = \left( \frac{1}{2\pi} \int_0^1 \int_{-\pi}^{\pi} |f(re^{i\phi})|^2 r d\phi dr \right)^{1/2}.
\]

Any method \( m : L_2(\mathbb{T}) \to L_2(\mathbb{D}) \) is admitted as a recovery method. The error of any one such method, \( m \), is defined as a worst case scenario over all \( f \in W \) and \( \tilde{y} \) satisfying
\begin{equation}
\ e(W, K, \delta, m) = \sup_{f \in W} \sup_{y \in L_2(T)} \|m(y) - f\|_{L_2(D)}.
\end{equation}

The error of optimal recovery is the smallest such error of any admissible method:

\begin{equation}
E(W, K, \delta) = \inf_{m: L_2(T) \to L_2(D)} e(W, K, \delta, m).
\end{equation}

A method attaining the error of optimal recovery is called an optimal method of recovery.

Consider the points in \(\mathbb{R}^2\) given by \(\{(\gamma_j(j + 1)^2, j + 1)\}_{j \in \mathbb{N}}\) and define the convex hull of the origin and this collection of points as \(M\):

\begin{equation}
M = \co\{(0,0) \cup \{(\gamma_j(j + 1)^2, j + 1)\}_{j \in \mathbb{N}}\}. \quad (2.6)
\end{equation}

Let

\begin{equation}
\theta(x) = \max\{y : (x, y) \in M\}, \quad (2.7)
\end{equation}

thus \(\theta\) is a piecewise linear function. Let \((x_s, y_s), s = 0, 1, \ldots\) be the points of break of \(\theta\) with \(0 = x_0 < x_1 < \ldots\).

**Theorem 1.** Suppose that \(x_s < \delta^{-2} \leq x_{s+1}\) with \(y_s > 0\). Let

\begin{equation}
\hat{\lambda}_1 = \frac{y_{s+1} - y_s}{x_{s+1} - x_s}, \quad \hat{\lambda}_2 = \frac{y_s x_{s+1} - y_{s+1} x_s}{x_{s+1} - x_s}. \quad (2.8)
\end{equation}

Then the error of optimal recovery is

\begin{equation}
E(W, K, \delta) = \sqrt{\hat{\lambda}_1 + \hat{\lambda}_2 \delta^2}, \quad (2.9)
\end{equation}

and

\begin{equation}
\hat{m}(y) = \sum_{j=0}^{\infty} (1 + \hat{\lambda}_1 \hat{\lambda}_2^{-1} \gamma_j(j + 1)^2)^{-1}(j + 1)\tilde{y}_j z^j \quad (2.10)
\end{equation}
is an optimal method of recovery. If \( y_s = 0 \) then \( E(W, K, \delta) = \sqrt{\frac{m}{x_1}} \) and \( \hat{m}(\tilde{y}) = 0 \) is an optimal method.

**Proof.** (see 4.1.1)

It should be noted that for fixed \( \hat{\lambda}_1, \hat{\lambda}_2 \neq 0 \), that is for a fixed \( \delta > 0 \), the terms

\[
\epsilon_j = \left( 1 + \hat{\lambda}_1 \hat{\lambda}_2^{-1} \gamma_j (j + 1)^2 \right)^{-1} (j + 1)
\]

will have the property, \( 0 < \epsilon_j < 1 \) and \( \lim_{j \to \infty} \epsilon_j = 0 \) as \( \lim_{j \to \infty} (\gamma_j (j + 1))^{-1} = 0 \). So \( \hat{m} \) smooths approximate values of the coefficients of \( \tilde{y} \) by the filter \( \epsilon_j \).

### 2.1.2 Inaccuracy in \( l^N_2(\mathbb{T}) \) Norm

Our next problem of optimal recovery remains to recover \( f \in X = X_\gamma \) from inaccurate information pertaining to the radial integral of \( f \). However, the inaccurate information we are given are the values \( \tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_{N-1} \in \mathbb{C} \) such that

\[
\sum_{j=0}^{N-1} |K_j f - \tilde{y}_j|^2 \leq \delta^2
\]

where \( K_j f \) is the \( j \)th coefficient of the radial integral \( K f \),

\[
K_j f = \langle K f(\phi), e^{-ij\phi} \rangle_{L^2(\mathbb{T})}.
\]

Denote

\[
K^N = (K_0, K_1, \ldots, K_{N-1}).
\]

We again consider the space of functions \( X = X_\gamma \) given by (2.1) and \( M \) and \( \theta \) defined by (2.6) and (2.7) respectively but now add the condition

\[
\gamma_j > 0, \quad j \geq N. \quad (2.11)
\]
The problem of optimal recovery on the class $W$ given by (2.4) is to determine the optimal error

$$E(W, K^N, \delta) = \inf_{m : \mathbb{C}^N \to L_2(\mathbb{D})} \sup_{f \in W, (\tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_{N-1}) \in \mathbb{C}^N, \sum_{j=0}^{N-1} |K_j f - \tilde{y}_j|^2 \leq \delta^2} \| f - m(\tilde{y}) \|_{L_2(\mathbb{D})}$$

(2.12)

and an optimal method $\hat{m} : \mathbb{C}^N \to L_2(\mathbb{D})$ obtaining this error.

Define $l_0 \in \mathbb{N}$ as the largest index such that

$$(\gamma_{l_0}(l_0 + 1))^{-1} = \max_{l \geq N} \{ (\gamma_l(l + 1))^{-1} \},$$

(2.13)

which by (2.3) exists, and

$$s_0 = \min \left\{ s \geq 0 : (\gamma_{l_0}(l_0 + 1))^{-1} \geq \frac{y_{s+1} - y_s}{x_{s+1} - x_s} \right\}.$$  

(2.14)

**Theorem 2.** Suppose $x_s < \delta^{-2} \leq x_{s+1}$ with $s < s_0$. If $y_s > 0$ let $\hat{\lambda}_1, \hat{\lambda}_2$ be given by (2.8). Then the optimal error is given by (2.9) and

$$\hat{m}(\tilde{y}) = \sum_{j=0}^{N-1} (1 + \hat{\lambda}_1 \hat{\lambda}_2^{-1} \gamma_j(j + 1)^2)^{-1} (j + 1) \tilde{y}_j z_j$$

(2.15)

is an optimal method. If $y_s = 0$ then $E(W, K^N, \delta) = \sqrt{\frac{y_{s+1}}{x_{s+1}}}$ and $\hat{m}(\tilde{y}) = 0$ is an optimal method.

If $\delta^{-2} > x_{s_0}$ with $y_{s_0} > 0$ then the error of optimal recovery is (2.9) and (2.15) is an optimal method with $\hat{\lambda}_1 = (\gamma_{l_0}(l_0 + 1))^{-1}$ and $\hat{\lambda}_2 = y_{s_0} - x_{s_0} \hat{\lambda}_1$. For $\delta^{-2} > x_{s_0}$ with $y_{s_0} = 0$, $E(W, K^N, \delta) = \sqrt{\frac{y_1}{x_1}}$ and $\hat{m}(\tilde{y}) = 0$ is an optimal method.

**Proof.** (see 4.1.2)

One may be able to reduce the amount information needed without affecting the error of optimal recovery. Therefore, by reducing the number of terms in the optimal method
we reduce the computations needed. The following ideas are in [27]. We consider the subset \( J_s \subseteq C_X \), \( s < s_0 \) as the set of all points whose slope to the origin is greater than the slope of \( \theta(x) \) for \( x \in [x_s, x_{s+1}] \), that is the slope of the line segment between points \((x_s, y_s)\) and \((x_{s+1}, y_{s+1})\). Define the sets

\[
J_s = \left\{ 0 \leq j \leq N - 1 : (\gamma_j(j + 1))^{-1} \geq \frac{y_{s+1} - y_s}{x_{s+1} - x_s} \right\}, \quad s = 0, 1, \ldots, s_0 \tag{2.16}
\]

where if \( \gamma_j = 0 \) define \((\gamma_j(j + 1))^{-1} = \infty\). Now consider the same problem as stated in Theorem 2 using only information \( K_{J_s}f \). For \( y_s > 0 \), \( \frac{y_s}{x_s} > \frac{y_{s+1}}{x_{s+1}} \geq \frac{y_{s+1} - y_s}{x_{s+1} - x_s} \) and so \((x_s, y_s), (x_{s+1}, y_{s+1}) \in \{(\gamma_j(j + 1)^2, j + 1)\}_{j \in J_s}\). In this situation, \( 0 \leq s < s_0 \) with \( y_s > 0 \), it was shown that the error of optimal recovery only involves the two points \((x_s, y_s), (x_{s+1}, y_{s+1})\) then the reduction in information from \( K^N \) to \( K_{J_s} \) will not change the error. That is \( E(W, K^N, \delta) = E(W, K_{J_s}, \delta) \) and if \( |J_s| = \tilde{N} \), an optimal method is

\[
\hat{m}(y) = \sum_{k=0}^{\tilde{N}} \left(1 + \lambda_1 \lambda_2 \gamma_{j_k}(j_k + 1)^2\right)^{-1} (j_k + 1) y_{j_k} z^{j_k} \tag{2.17}
\]

where \( y = (y_0, \ldots, y_{\tilde{N}}) \).

### 2.1.3 Varying Levels of Accuracy Termwise

In Theorems 1 and 2 the inaccuracy of the information given is a total inaccuracy. That is, the inaccuracy \( \delta^2 \) is an upper bound on the sum total of the inaccuracies in each term, be it a finite or infinite sum. For Theorems 1 and 2 however, there is no way to tell how the inaccuracy is distributed. In particular, with regards to Theorem 2, the situations in which the given information \( \tilde{y} = (\tilde{y}_0, \ldots, \tilde{y}_{N-1}) \) satisfies

\[
|K_j f - \tilde{y}_j| \leq \delta/\sqrt{N}, \quad j = 0, \ldots, N - 1
\]
or for some particular \( m \in \mathbb{N} \) satisfying \( 0 \leq m \leq N - 1 \)

\[
|K_j f - \tilde{y}_j| \leq \begin{cases} 
\delta : j = m \\
0 : j \neq m 
\end{cases}
\]

are treated the same. For the next problem of optimal recovery we address this ambiguity.

The problem of optimal recovery is to determine an optimal method and the optimal error of recovering \( f \in X_\gamma \), from the information \( \tilde{y} = (\tilde{y}_0, \ldots, \tilde{y}_{N-1}) \in \mathbb{C}^N \) satisfying

\[
|K_j f - \tilde{y}_j| \leq \delta_j
\]

for some prescribed \( \delta_j \geq 0 \) and \( j = 0, \ldots, N - 1 \).

To define \( X_\gamma \) use conditions (2.2) and (2.11) as previously but impose an additional restriction. We add the condition

\[
\gamma_j(j + 1) \leq \gamma_{j+1}(j + 2), \quad j = 0, 1, \ldots.
\]

Define \( \tilde{\delta} = (\delta_0, \ldots, \delta_{N-1}) \) where \( \delta_j \geq 0 \) are the levels of accuracy. If \( \gamma_0 \delta_0^2 \leq 1 \) define

\[
p_0 = \max \left\{ p \geq 0 : \sum_{j=0}^{p} \delta_j^2 \gamma_j(j + 1)^2 \leq 1, \quad p \leq N - 1 \right\}.
\] (2.18)

**Theorem 3.** If \( \gamma_0 \delta_0^2 \leq 1 \) let

\[
\hat{\lambda} = \frac{1}{\gamma_0(p_0 + 1)}, \quad \hat{\lambda}_j = \begin{cases} 
\frac{j + 1}{\gamma_0 + 1} - \hat{\lambda} \gamma_j(j + 1)^2 & : j \leq p_0 \\
0 & : p_0 + 1 \leq j \leq N - 1
\end{cases}
\] (2.19)

then the error of optimal recovery is given by

\[
E(W, K^N, \tilde{\delta}) = \sqrt{\hat{\lambda} + \sum_{j=0}^{p_0} \delta_j^2 \hat{\lambda}_j}
\] (2.20)
and

\[ \hat{m}(\tilde{y}) = \sum_{j=0}^{p_0} \hat{\lambda}_j \tilde{y}_j z_j \]  \hspace{0.5cm} (2.21)

is an optimal method.

If \( \gamma_0 \delta_0^2 > 1 \) then \( E(W, K^N, \delta_0) = \gamma_0^{-1/2} \) and \( \hat{m}(\tilde{y}) = 0 \) is an optimal method.

Proof. (see 4.1.3)

The optimal method may not use all of the information provided as \( p_0 \) may be less than \( N - 1 \). Thus increasing \( N \) may not change \( p_0 \) and hence not change the error or the method. If \( p_0 < N - 1 \), then

\[ E(W, K^{p_0+1}, \delta) = E(W, K^N, \delta) \]

and we can reduce the amount of information needed for a given optimal error.

If \( p_0 = N - 1 \) we may be able to reduce the error of optimal recovery if we have more information available. Fix \( \delta = \delta_\infty = (\delta_0, \delta_1, \ldots) \). The greater number of terms we have of \( Kf \) then the better we may be able to approximate \( f \), that is the smaller the optimal error of recovery. Let

\[ N_\delta = \max\{p \geq 0 : \sum_{j=0}^{p} \delta_j^2 \gamma_j (j+1)^2 \leq 1\} \]  \hspace{0.5cm} (2.22)

and for \( N_\delta < \infty \)

\[ E(W, K^{N_\delta}, \delta) \leq E(W, K^N, \delta) \]

for any \( N \geq 0 \). If we know the first \( N_\delta \) terms with some errors, then further increasing the terms will not yield a decrease in the error of optimal recovery.
2.1.4 Applications: Hardy–Sobolev and Bergman–Sobolev Spaces

We now apply the general results to the Hardy–Sobolev and Bergman–Sobolev spaces of functions on the unit disc. Let $\mathcal{H}(\mathbb{D})$ denote the set of functions holomorphic on the unit disc. Define the Hardy space of functions $\mathcal{H}^2(\mathbb{D})$ as the set of all $f \in \mathcal{H}(\mathbb{D})$, $f(z) = \sum_{j=0}^{\infty} a_j z^j$ with $\|f\|_{\mathcal{H}^2(\mathbb{D})} < \infty$ where

$$\|f\|_{\mathcal{H}^2(\mathbb{D})}^2 = \sum_{j=0}^{\infty} |a_j|^2.$$

The Hardy-Sobolev space of functions, $\mathcal{H}^{2,r}(\mathbb{D})$, are those $f \in \mathcal{H}(\mathbb{D})$ such that $f^{(r)} \in \mathcal{H}^2(\mathbb{D})$ and $H^{2,r}(\mathbb{D})$ is the class consisting of those $f \in \mathcal{H}^{2,r}(\mathbb{D})$ with $\|f^{(r)}\|_{\mathcal{H}^2(\mathbb{D})} \leq 1$.

The Bergman space of functions $\mathcal{A}^2(\mathbb{D})$ is the space of all $f \in \mathcal{H}(\mathbb{D})$ such that

$$\|f\|_{\mathcal{A}^2(\mathbb{D})}^2 = \int_{\mathbb{D}} |f(z)|^2 dA < \infty.$$

That is, $\mathcal{A}^2(\mathbb{D})$ is the space of all holomorphic functions in $L^2(\mathbb{D})$. The Bergman-Sobolev space of functions, $\mathcal{A}^{2,r}(\mathbb{D})$, consists of $f \in \mathcal{H}(\mathbb{D})$ with $f^{(r)} \in \mathcal{A}^2(\mathbb{D})$ and $A^{2,r}(\mathbb{D})$ as the class of all $f \in \mathcal{A}^{2,r}(\mathbb{D})$ with $\|f^{(r)}\|_{\mathcal{A}^2(\mathbb{D})} \leq 1$.

So each space can be considered as the space $X_\gamma$ with

$$\gamma_j = \gamma_j(X) = \begin{cases} 0 & : j < r \\ \left(\frac{j!}{(j-r)!}\right)^2 & : r \leq j, \quad X = \mathcal{H}^{2,r}(\mathbb{D}) \\ \left(\frac{j!}{(j-r)!}\right)^2 \frac{1}{j-r+1} & : r \leq j, \quad X = \mathcal{A}^{2,r}(\mathbb{D}) \end{cases}.$$ 

For each space of functions we have the collection of points $C_X = \{\alpha_j, \beta_j\} = \{(\gamma_j(j+1)^2, j+1)\}_{j \in \mathbb{N}}$. If $X = \mathcal{H}^{2,r}(\mathbb{D})$ then for $j \geq r$

$$\gamma_j(j+1)^2 = \left(\frac{j!}{(j-r)!}\right)^2 (j+1)^2 = \left(\frac{(j+1)!}{(j-r)!}\right)^2.$$ 

18
Therefore for $r = 0, 1, \ldots$

$$\lim_{j \to \infty} (\gamma_j(j + 1))^{-1} = \lim_{j \to \infty} \left(\frac{(j - r)!}{j!}\right)^2 \frac{1}{j + 1} = 0.$$ 

In this case we consider the collection of points

$$C_{H^2,r} = \left\{(0, 0) \cup \left\{ \gamma_j(H^2)(j + 1)^2, j + 1 \right\} \right\}_{j \in \mathbb{N}}.$$ 

It is easy to see that if $M = \text{co}(C_{H^2,r})$ then the piecewise linear function $\theta(x) = \max\{y : (x, y) \in M\}$ will have points of break

$$\left\{(0, r), ((r + 1)!^2, r + 1), \ldots, \left(\frac{(j + 1)!}{(j - r)!}, j + 1\right), \ldots\right\}. \quad (2.23)$$

As an example, if $r = 0$, then the problem is to recover $f \in H^2$ from information about the radial integral, $Kf$. In this case the problem simplifies as $\gamma_j = 1$ for all $j \geq 0$ and the points of break of $\theta$ are given by

$$\{(0, 0), (1, 1), (2^2, 2), (3^2, 3), \ldots, ((s + 1)^2, s + 1), \ldots\}.$$ 

Similarly for $X = A^{2,r}(\mathbb{D})$ and $r = 1, 2, \ldots$ we have

$$\lim_{j \to \infty} (\gamma_j(j + 1))^{-1} = \lim_{j \to \infty} \left(\frac{(j - r)!}{j!}\right)^2 \frac{j - r + 1}{j + 1} = 0.$$ 

For the space $A^{2,r}$, the points to consider are

$$C_{A^{2,r}} = \left\{(0, 0) \cup \left\{ \gamma_j(A^{2,r})(j + 1)^2, j + 1 \right\} \right\}_{j \in \mathbb{N}}.$$ 

Again let $\theta(x) = \max\{y : (x, y) \in \text{co}(C_{A^{2,r}})\}$ and thus the points of break of $\theta$ will be
For the special case of \( A^{2,0} \), the function \( \theta \) has only a single point of break at the origin as

\[
C_{A^{2,0}} = \{(0, 0) \cup \{(j + 1, j + 1)\}_{j \in \mathbb{N}}\}
\]

so that \( \theta(x) = x \) for \( x \geq 0 \). Furthermore, \( A^{2,0} \) does not satisfy (2.3) as

\[
\lim_{j \to \infty} \left( \frac{1}{j + 1} \right)^{-1} = \lim_{j \to \infty} \frac{1}{j + 1} = 1.
\]

Thus, in the applications of the general results, this case will be treated separately.

For notational purposes, let \((x_s, y_s) = (x_s(X), y_s(X)), s = 0, 1, \ldots\) be the points of break of \( \theta \) for the space \( X \).

**Corollary 1.** Let \( X = H^{2,r} \) or \( A^{2,r} \). If \( x_s < \delta^{-2} \leq x_{s+1} \) with \( s > 0 \) or \( r > 0 \) then the error of optimal recovery is given by (2.9) and (2.10) is an optimal method. If \( s = 0 \) and \( r = 0 \) then \( E(W, K, \delta) = 1 \) and \( \hat{m}(y) = 0 \) is optimal.

**Proof.** For the spaces \( X = A^{2,r} \) or \( H^{2,r} \), \( y_s = 0 \) if and only if \( s = 0 \) and \( r = 0 \). Thus \( y_s > 0 \) if and only if \( s > 0 \) or \( r > 0 \). Thus apply Theorem 1 to obtain the result for all spaces except \( A^{2,0} \). The dual problem in the case \( X = A^{2,0} \) leads to a simple Lagrange function. The dual problem is specifically

\[
\sum_{j=0}^{\infty} \frac{|a_j|^2}{j + 1} \rightarrow \max, \quad \sum_{j=0}^{\infty} \frac{|a_j|^2}{j + 1} \leq 1, \quad \sum_{j=0}^{\infty} \frac{|a_j|^2}{(j + 1)^2} \leq \delta^2.
\]

Therefore the Lagrange function is simply given by

\[
\mathcal{L}(f, \lambda_1, \lambda_2) = \sum_{j=0}^{\infty} \frac{|a_j|^2}{(j + 1)^2} (\lambda_1(j + 1) + \lambda_2 - (j + 1)).
\]
Now if we let $\lambda_1 = 1$ and $\lambda_2 = 0$ then $L(f, \lambda_1, \lambda_2) = 0$ for any $f \in X$. So now proceed as in Theorem 1. As any $\hat{f} \in A^{2,0}$ will minimize $L$, choose $\hat{f}$ as in (4.4). The extremal problem (4.5) is solved similarly, and as $\lambda_2 = 0$ then $a_j = 0$ for $j = 0, 1, \ldots$.

It should be noted that the optimal method described is stable with respect to the inaccurate information data. That is, if $\delta^2 \in (x_s, x_{s+1}]$ then $\delta^2$, and hence the inaccuracy, can vary within this interval and the method (2.10) will still be an optimal method.

We now apply Theorem 2 to the Hardy–Sobolev spaces $H^{2,r}$ and Bergman–Sobolev spaces $A^{2,r}$ in which $s_0$ is explicitly defined by

$$s_0 = \min \left\{ s \geq 0 : \frac{1}{\gamma_N(N + 1)} \geq \frac{1}{\gamma_{s+1}(s + 2)^2 - \gamma_s(s + 1)^2} \right\}.$$

For the case $W = A^{2,0}$, $\gamma_j(j + 1) = 1$ for all $j \geq 0$. Thus $s_0 = 0$ does not depend on $N$. So $x_{s_0}(A^{2,0}) = x_0(A^{2,0}) = 0$ and hence for any $\delta$ we are in the case $\delta^2 > x_{s_0}$.

**Corollary 2.** Let $X = H^{2,r}$ or $A^{2,r}$. Suppose $x_s < \delta^2 \leq x_{s+1}$ with $s < s_0$. If $s > 0$ or $r > 0$ then let $\lambda_1, \lambda_2$ be given by (2.8) and the optimal error is given by (2.9) and (2.15) is an optimal method. If $s = 0$ and $r = 0$ then $E(W, K^N, \delta) = 1$ and $\hat{m}(\tilde{y}) = 0$ is an optimal method.

Otherwise suppose $\delta^2 > x_{s_0}$. If $s_0 > 0$ or $r > 0$ then the optimal error is given by (2.9) and (2.15) is an optimal method with $\lambda_1 = (\gamma_N(N + 1))^{-1}$ and $\lambda_2 = y_{s_0} - x_{s_0} \lambda_1$.

If $s_0 = 0$ and $r = 0$ then $E(W, K^N, \delta) = 1$ and $\hat{m}(\tilde{y}) = 0$ is an optimal method.

**Proof.** As previously stated, if $X = A^{2,0}$ the only break point of $\theta$ is $(0, 0)$ and furthermore as $\gamma_j(A^{2,0}) = (j + 1)^{-1}$ then $l_0$ given by (2.13) does not exist so we treat this special case. In this case, the dual extremal problem is

$$\|f\|_{L^2(D)}^2 \to \max, \quad \sum_{j=0}^{\infty} \frac{|a_j|^2}{j + 1} \leq 1, \quad \sum_{j=0}^{N-1} \frac{|a_j|^2}{(j + 1)^2} \leq \delta^2.$$
and the corresponding Lagrange function is simply

\[ \mathcal{L}(f, \lambda_1, \lambda_2) = \sum_{j=0}^{\infty} \frac{|a_j|^2}{(j+1)^2} (\lambda_1(j+1) + \chi_j^N \lambda_2 - (j+1)) \]

where \( \chi_j^N \) is the characteristic function of \( \{ j \in \mathbb{N} : j < N \} \). If \( \hat{\lambda}_1 = 1 \) and \( \hat{\lambda}_2 = 0 \) then \( \mathcal{L}(f, 1, 0) = 0 \) for any \( f \in A^{2,0} \). Now proceed as in the proof of Theorem 2 to obtain the result.

Example 1. \( X = H^{2,0} \)

For the class \( H^{2,0} \), \( \gamma_j = 1 \) for all \( j \geq 0 \). So for \( W = H^{2,0} \) we have \( (x_s, y_s) = (s^2, s) \) for all \( s \geq 0 \). Clearly

\[ \max_{j \geq N} \{(\gamma_j(j+1))^{-1}\} = \max_{j \geq N} \{(j+1)^{-1}\} = (N+1)^{-1} \]

and so

\[ s_0 = \min \left\{ s \geq 0 : \frac{1}{N+1} \geq \frac{y_{s+1} - y_s}{x_{s+1} - x_s} = \frac{1}{2s+1} \right\} \]

\[ = \min \{ s \geq 0 : N \leq 2s \} . \]

That is \( s_0 = \left\lceil \frac{N}{2} \right\rceil \) and furthermore,

\[ x_{s_0} = \left\lceil \frac{N}{2} \right\rceil^2 . \] (2.25)

We may be able to reduce the amount of information we need in order to obtain a given optimal error. That is, we may not need all of the information. For example, consider the case \( W = H^{2,0} \), \( N = 30 \), and let \( \delta > 0 \) satisfy \( \frac{1}{10} \leq \delta < \frac{1}{9} \) with \( s_0 > 9 \). In this case, as \( E(H^{2,0}, K^N, \delta) \) only depends on the values \( (x_s, y_s) = (9^2, 9) \) and \( (x_{s+1}, y_{s+1}) = (10^2, 10) \) then we can reduce the number of terms used until \( s_0 = s + 1 = 10 \) without changing
the values of the points, \((x_s, y_s), (x_{s+1}, y_{s+1})\). By \((2.25)\), the value of \(N\) can be reduced to \(N_\delta = 19\) without changing the optimal error. Furthermore, the optimal method can now be taken over only \(N_\delta \leq N\) terms.

The sets \(J_s\) given by \((2.16)\) were defined with the same reasoning

\[
J_s = \{0 \leq j \leq 29 : \frac{1}{j+1} \geq \frac{1}{2s+1}\}.
\]

If we let \(s = 9\) we obtain

\[
J_9 = \{0 \leq j \leq 29 : \frac{1}{j+1} \geq \frac{1}{19}\} = \{j \in \mathbb{N} : 0 \leq j \leq 18\}
\]

so that the method of optimal recovery \((2.17)\) only uses the first 19 terms even though 30 are given.

**Example 2.** \(X = A^{2,1}\).

In this example we have \(\gamma_j = j\) for all \(j \geq 0\). After noticing

\[
\max\{\gamma_j(j + 1)^{-1}\} = \max\{((j(j+1))^{-1}\} = (N(N+1))^{-1}
\]

and some arithmetic we can show that \(s_0\) is the smallest non-negative integer satisfying the inequality

\[
3s^2 + 7s + 4 - N(N + 1) \geq 0.
\]

As the left hand side is simply a quadratic in \(s\), we can show the inequality will be satisfied if

\[
s \geq -\frac{7}{6} + \frac{1}{6} (1 + 12N(N + 1))^{1/2}
\]

which will be non-negative if \(N \geq 2\). As one of our assumptions is \(0 \leq r \leq N - 1\) then
in this case we already know, as \( r = 1 \), \( N \geq 2 \). Thus
\[
s_0 = \left\lceil \frac{1}{6} \left( (1 + 12N(N+1))^{1/2} - 7 \right) \right\rceil.
\]

**Example 3.** \( X = \mathcal{H}^{2,1} \).

We now let \( r = 1 \) and consider the class \( \mathcal{H}^{2,1} \). We know \( \gamma_0 = 0 \) and \( \gamma_j = j^2 \) for \( j = 1, 2, \ldots \) so that \( (x_s, y_s) = ((s(s+1))^2, s+1) \) for all \( s \geq 0 \). To find \( s_0 \) we first note that
\[
\max_{j \geq N} \{ (\gamma_j(j+1))^{-1} \} = \max_{j \geq N} \{ (j^2(j+1))^{-1} \} = (N^2(N+1))^{-1}.
\]

From the definition of \( s_0 \) we have
\[
s_0 = \min \left\{ s \geq 0 : N^2(N+1) \leq x_{s+1} - x_s = s^2(s+1)^2 - s^2(s-1)^2 \right\}
\]
\[
= \min \left\{ s \geq 0 : N^2(N+1) \leq 4s^3 \right\}.
\]

Solving for \( s \) over the nonnegative integers we have
\[
s_0 = \left\lceil \left( \frac{N^2(N+1)}{4} \right)^{1/3} \right\rceil.
\]

Thus
\[
x_{s_0} = (s_0(s_0+1))^2 = \left( \left\lceil \left( \frac{N^2(N+1)}{4} \right)^{1/3} \right\rceil \right) \left( \left( \frac{N^2(N+1)}{4} \right)^{1/3} + 1 \right)^2.
\]

For any \( \delta' \) with \( \delta'^{-2} \geq \delta^{-2} > x_{s_0} \) the optimal method of recovery of \( x \in \mathcal{H}^{2,1} \) will not change. That is, knowing \( Kf \) with more precision than \( \delta_{x_{s_0}} = \frac{1}{\sqrt{x_{s_0}}} \) will not change the constructed method of optimal recovery but will reduce the optimal error.

We now apply Theorem 3 to the spaces \( X = \mathcal{H}^{2,r} \) or \( X = \mathcal{A}^{2,r} \) for \( 0 \leq r \leq N - 1 \). In this situation \( \gamma_j(j+1) \) will be a non-decreasing sequence for all \( j \geq 0 \). Also, for any \( r > 0 \) we have \( \gamma_0 = 0 \) and we are always in the case \( \gamma_0 \delta_0^2 \leq 1 \). For \( r = 0 \) then for both
the Hardy and Bergman spaces $\gamma_0 = 1$ and so the condition $\gamma_0 \delta_0^2 \leq 1$ will be satisfied if we know $\tilde{y}_0$ satisfying
$$|a_0 - \tilde{y}_0| = |K_0 f - \tilde{y}_0| \leq \delta_0 \leq 1.$$

**Corollary 3.** Let $X = H^{2,r}$ or $X = A^{2,r}$ with $1 \leq r \leq N - 1$ or $r = 0$ and $\delta_0 \leq 1$ and $p_0$ given by (2.18). Let $\tilde{\lambda}, \tilde{\lambda}_j, j = 0, \ldots, N - 1$ be given by (2.19). Then the error of optimal recovery is given by (2.20) and (2.21) is an optimal method. If $r = 0$ and $\delta_0 > 1$ then $E(W, K^N, \delta) = 1$ and $\hat{m}(\tilde{y}) = 0$ is an optimal method.

**Proof.** For Theorem 3 we simply used conditions (2.2) and (2.11), both of which are satisfied by $H^{2,r}$ and $A^{2,r}$ for all $0 \leq r \leq N - 1$.

As a direct consequence of Theorem 3, we consider the situation in which we have a uniform bound on the inaccuracy of each of the first $N$ terms of $K_j f$. That is we take $\delta_j = \delta$ for every $0 \leq j \leq N - 1$. If $\delta^2 \gamma_0 \leq 1$ we define $p_0$ similarly as
$$p_0 = \max \left\{ p \geq 0 : \delta^2 \sum_{j=0}^{p} \gamma_j (j+1)^2 \leq 1, \ p \leq N - 1 \right\}$$
and the apriori information is given by the values $\tilde{y}_0, \tilde{y}_1, \ldots, \tilde{y}_{N-1}$ such that
$$|K_j f - \tilde{y}_j| \leq \delta.$$

Again we will only need the values $\tilde{y}_0, \ldots, \tilde{y}_{p_0}$ for an optimal method.

As previously noted, as the optimal method and error of optimal recovery only use up to the $p_0$ term then any information beyond may be disregarded if $p_0 < N - 1$ as additional information will not decrease the error of optimal recovery.
Example 4. Considering the space $A^{2,0}$,

\[
p_0 = \max \left\{ p \geq 0 : \delta^2 \sum_{j=0}^{p} j + 1 \leq 1, \quad p \leq N - 1 \right\}
\]

\[
= \max \left\{ p \geq 0 : p^2 + p - 2\delta^{-2} \leq 0 \right\}
\]

\[
= \left\lfloor \frac{1}{2} (-1 + \sqrt{1 + 8\delta^{-2}}) \right\rfloor
\]

For the space $H^{2,0}$, $p_0$ is defined by

\[
p_0 = \max \left\{ p \geq 0 : \delta^2 \sum_{j=0}^{p} (j + 1)^2 \leq 1, \quad p \leq N - 1 \right\}
\]

\[
= \max \left\{ p \geq 0 : \delta^2 \frac{(p + 1)(p + 2)(2p + 3)}{6} \leq 1, \quad p \leq N - 1 \right\}
\] (2.26)

using the sum of squares relationship

\[
\sum_{k=1}^{n} k^2 = \frac{1}{6} n(n + 1)(2n + 1).
\]

2.2 Secant Integral Information Operator

In engineering, medicine and other fields it is often useful to “see” inside a particular object. In medicine, a non-invasive procedure to assist in diagnosing a patient is highly desirable. One of the problems with a standard X-ray is that it is simply a 2D graph of the average density of a 3D object along a straight line path, so objects become superimposed. One way around the loss of 3D information is to take many images from different angles and reconstruct a 3D image of the internal structure. If one of the three dimensions is fixed and many images are taken of this cross-section at different angles the internal nature of the object can be reconstructed. The idea is to then vary the fixed dimension and combine each reconstructed cross-section to obtain a 3D model.

In 1917, Johan Radon developed what would be later coined the Radon transform, or
Radon line transform to represent the integral of a function over a straight line path in $\mathbb{R}^2$, or hyperplane in $\mathbb{R}^n$. The amazing fact shown by Radon was that if the straight line path integrals were taken over all possible angles then from all of these resulting projections, a complete and exact reconstruction of the function could be obtained, through the inverse Radon transform.

In the X-ray tomography setting, the object being scanned has a particular density function $\mu : \mathbb{R}^3 \to \mathbb{R}$. The absorption or scattering of energy of the incident X-ray beam is dependent upon the density $\mu(x, y, z = c)$, where $c$ is constant, along the straight line path it takes through the object, this is referred to as the attenuation of the beam. For each fixed value of $z = c$ and each incident beam angle, $\zeta$, a projection $\mathbb{R}^3 \to \mathbb{R}^2$ is obtained. In practice, several problems/difficulties occur. First, not all values for $\zeta \in [-\pi, \pi]$ can be used so there is incomplete data. Secondly, there are inaccuracies in the measurement of the attenuation. This can be due to scattering of the x-ray beam, variations in incident X-ray energy, and simply movement of the patient [9]. In what follows, the problem considered is optimal recovery of an analytic function on the unit disc from inaccurate information about the Radon transform. This can be likened to reconstruction of the "density" function in one slice, $z = c$, from projections that contain some error.

Again we consider the space of functions $X = X_\gamma$ as previously defined. The integral operator to consider is the integral of $f \in X$ along the secant path on the unit disc given by parameters $h \in [0, 1]$ and $\phi \in [-\pi, \pi]$ as shown. Then

\begin{figure}[h]
\centering
\includegraphics[width=0.3\textwidth]{secant_integral_info.png}
\caption{Secant Integral Information}
\end{figure}
\[ Lf = \int_{p_1 p_2} f(z)|dz| \]

is the Radon transform of \( f \) along path \( p_1 p_2 \). Parametrizing the segment \( p_1 p_2 \) we use

\[ z(t) = (1 - t)p_1 + tp_2 = (1 - t)e^{i(\phi + \cos^{-1} h)} + te^{i(\phi - \cos^{-1} h)}, \]

and thus

\[ |dz| = 2\sqrt{1 - h^2}dt \]

and

\[ Lf(\phi, h) = \sum_{j=0}^{\infty} a_j c_j(h)e^{ij\phi} \quad (2.27) \]

where \( c_j(h) = \frac{2}{j+1} \sin((j + 1) \cos^{-1} h) \). In the \( L_2(\mathbb{D}) \) norm we obtain

\[ \|Lf\|^2_{L_2(\mathbb{D})} = \frac{1}{2\pi} \int_0^1 \int_{-\pi}^{\pi} |Lf(\phi, h)|^2 h d\phi dh = \sum_{j=0}^{\infty} |a_j|^2 \cdot \eta_j^{-1} \]

where

\[ \eta_j = \begin{cases} (j + 1)^2 & : j \equiv 0 \mod 2 \\ j(j + 2) & : j \equiv 1 \mod 2 \end{cases} \quad (2.28) \]

To see that \( Lf \in L_2(\mathbb{D}) \) we have as \( \left( \gamma_j(j + 1) \right)^{-1} \to 0 \) then there is some \( c > 0 \) such that for all but finitely many \( j \) we have \( \gamma_j \geq \frac{c}{j+1} \geq \frac{c}{\eta_j} \). Hence, for some \( M \in \mathbb{R} \)

\[ \|Lf\|^2_{L_2(\mathbb{D})} = \sum_{j=0}^{\infty} \frac{|a_j|^2}{\eta_j} \leq M + \frac{1}{c} \sum_{j=0}^{\infty} \gamma_j |a_j|^2 < \infty \]

as \( f \in X \) and so \( Lf \in L_2(\mathbb{D}) \).

Let \( J_n \) be the Bessel function of the first kind of order \( n \in \mathbb{Z} \),

\[ J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(n + m + 1)!} \left( \frac{x}{2} \right)^{2m+n} \]
\[ y(h, \phi) = \sum_{j \in \mathbb{Z}} \sum_{k=1}^{\infty} y_{jk} J_{\mu_k(h)}(\mu_k h) e^{ij\phi}. \] \tag{2.29}

Let \( d_j(h) = \sum_{k=1}^{\infty} y_{jk} J_{\mu_k(h)} \) so \( y(h, \phi) = \sum_{j \in \mathbb{Z}} d_j(h) e^{ij\phi} \) for \( j \in \mathbb{Z} \). For \( j \geq 0 \), denote
\[ \nu_j = \int_0^1 d_j(h) c_j(h) h \, dh. \tag{2.30} \]

Now let \( C = \{(0, 0) \cup \{\gamma_j \eta_j, \eta_j(j + 1)^{-1}\}_{j \geq 0}\} \) and \( M = \text{co} \, C \). First notice that
\[ \lim_{j \to \infty} \eta_j(j + 1)^{-1} = \infty \quad \text{and the slope from } (0, 0) \text{ to a point } (\gamma_j \eta_j, \eta_j(j + 1)^{-1}) \text{ will have the property} \]
\[ \lim_{j \to \infty} \frac{\eta_j(j + 1)^{-1}}{\gamma_j \eta_j} = \lim_{j \to \infty} \frac{1}{\gamma_j(j + 1)} = 0 \]
by assumption on the sequence \( \{\gamma_j\}_{j \geq 0} \). So by Lemma 1, if \( \theta(x) = \max\{y : (x, y) \in M\} \) has points of break \( (x_s, y_s), s = 0, 1, \ldots \) with \( 0 = x_0 < x_1 < \cdots \) then any linear piece of \( \theta \) will have non-negative slope and intercept. Since these points satisfy Lemma 1 then we can solve similar problems as in the case of recovery of \( f \in X \) from information given inaccurately about the radial integral of \( f \), and as such the proofs will have much of the same style.

Let it be given, \( y \in L_2(\mathbb{D}) \), such that
\[ ||L_f - y||_{L_2(\mathbb{D})} \leq \delta \tag{2.31} \]

\[
\nu_j = \sum_{k=1}^{\infty} \sum_{n=0}^{2k+j+1} \frac{(-1)^k y_{jk}}{(j+1)k!(k+j+1)!} (\mu_{j}^{(k+j)})^{2k+j} 2^{-(4k+2j+2)} \binom{2k+j+1}{n} \left( \frac{e^{(2k+2j+3-2n)\frac{\pi}{2} i}}{2k+2j+3-2n} - \frac{e^{(2k+2j+1-2n)\frac{\pi}{2} i}}{2k+2j+1-2n} - \frac{e^{(2k+1-2n)\frac{\pi}{2} i}}{2k+1-2n} + \frac{e^{(2k-1-2n)\frac{\pi}{2} i}}{2k-1-2n} \right)
\]
for some $\delta \geq 0$. The problem of optimal recovery is to determine an optimal recovery method of the function $f$ from the inaccurate information $y$ and to determine the error of optimal recovery.

Any mapping $m : L_2(\mathbb{D}) \to L_2(\mathbb{D})$ is admitted as a recovery method. The error of a method $m$ is

$$e(W, L, \delta, m) = \sup_{f \in W, \|y\|_{L_2(\mathbb{D})} \leq \delta} \|f - m(y)\|_{L_2(\mathbb{D})}$$

and so the error of optimal recovery is

$$E(W, L, \delta) = \inf_{m : L_2(\mathbb{D}) \to L_2(\mathbb{D})} e(W, L, \delta, m).$$

**Theorem 4.** Suppose $x_s < \delta^{-2} \leq x_{s+1}$ with $y_s > 0$. Let

$$\hat{\lambda}_1 = \frac{y_{s+1} - y_s}{x_{s+1} - x_s}, \quad \hat{\lambda}_2 = \frac{x_{s+1}y_s - x_s y_{s+1}}{x_{s+1} - x_s}. \quad (2.32)$$

Then

$$E(W, L, \delta) = \sqrt{\hat{\lambda}_1 + \delta^2 \hat{\lambda}_2} \quad (2.33)$$

and

$$\hat{m}(y) = \sum_{j=0}^{\infty} \frac{\hat{\lambda}_2 \nu_j}{\hat{\lambda}_1 \gamma_j + \hat{\lambda}_2 \eta_j} z^j. \quad (2.34)$$

If $y_s = 0$ then $\hat{m}(y) = 0$ is optimal and $E(W, L, \delta) = \sqrt{y_1/x_1}$.

**Proof.** (see 4.2)

### 2.2.1 Applications: Hardy–Sobolev and Bergman–Sobolev Spaces

For the Hardy–Sobolev and Bergman–Sobolev spaces of functions, Theorem 4 can be applied directly for $\mathcal{H}^{2,r}$ for $r \geq 0$ and for $\mathcal{A}^{2,t}$ for $t \geq 1$ but $\mathcal{A}^{2,0}$ must be treated separately. For $\mathcal{A}^{2,0}$, as stated for Corollary 1, condition (2.3) is not satisfied. Essentially,
the same argument as in Corollary 1 will be used.

**Corollary 4.** Let $X = \mathcal{H}^{2,r}$ or $\mathcal{A}^{2,r}$. If $x_s < \delta^{-2} \leq x_{s+1}$ with $s > 0$ or $r > 0$ then a method of optimal recovery is (2.34) and the error of optimal recovery is (2.33). If $s = r = 0$ then $E(W, K, \delta) = 1$ and $\hat{m}(y) = 0$ is optimal.

**Proof.** For the case $X = \mathcal{A}^{2,0}$ the dual problem to consider is

$$\sum_{j \geq 0} \frac{1}{j+1} |a_j|^2 \rightarrow \max, \quad \sum_{j \geq 0} \frac{1}{j+1} \leq 1, \quad \sum_{j \geq 0} \frac{1}{\eta_j} |a_j|^2 \leq \delta^2$$

which leads to the simple Lagrange function

$$\mathcal{L}(f, \lambda_1, \lambda_2) = \sum_{j \geq 0} \frac{|a_j|^2}{\eta_j} \left( \lambda_1 \frac{\eta_j}{j+1} + \lambda_2 - \frac{\eta_j}{j+1} \right).$$

Let $\hat{\lambda}_1 = 1$ and $\hat{\lambda}_2 = 0$ and for all $f \in X$, $\mathcal{L}(f, \hat{\lambda}_1, \hat{\lambda}_2) = 0$. Construct $\hat{f}(z)$ using the system (4.15) and proceed as in Theorem 4.

$$\square$$

### 2.3 Directions For Further Study

There are several problems that have arisen and could be addressed in future research.

1. Each problem addressed herein has been completed where each of the spaces $X, Y, Z$ is a Hilbert space. What can be said of the optimal recovery of these problems in a non-Hilbert space setting?

2. As Osipenko and Stessin did in [24], these ideas might be extended to the cases where the spaces are of multi-variable functions instead of single variable ones.

3. The condition on the values of $\gamma_j$, $j = 0, 1, \ldots$ given by 2.3 could possibly be weakened and similar results obtained.
I would like to close the gap between Theorem 4 and application. One way to accomplish this could be to address a more sampling type of problem.

4. Given some $\epsilon > 0$ and some inaccuracy in the data $\delta > 0$ is it possible to determine if the error of optimal recovery can be made smaller than $\epsilon$ with a finite set of projections. That is determine if the set of angles $\phi_1, \ldots, \phi_N$ with inaccuracy of information at most $\delta$ is enough to obtain $E(T, W, I) \leq \epsilon$.

5. If instead we are given a predetermined finite set of angles at which projections are taken, what can be said of the optimal error and method of recovery?

6. Many CT scanners are using a fan beam type of radiation emitter so that the incident X-rays are not parallel but instead are in a fan or cone shape through the object. I would like to see how the ideas of this paper can be also addressed with the new shape.
Chapter 3

Background Material

In his unpublished doctoral dissertation [29], the work of Smolyak addresses the problem of optimal recovery of a real linear functional from exact values of $n$ linear information functionals and under what conditions there would exist a linear optimal method. We place this result in the context of our general setting and use it as a starting point for the problems addressed herein. To consider Smolyak’s result we will have the need of developing some surrounding material.

**Definition 1.** If $X$ is a real or complex linear space, a set $B$ is a *balanced convex set* if and only if for every $x \in B$ and every $\alpha \in X$, $|\alpha| \leq 1$, we have $\alpha x \in B$.

Let $X$ be a real linear space and $T, I_1, \ldots, I_n$ real linear functionals. The problem is to recover $Tx$ for $x \in W$, where $W \subset X$ is a balanced convex set, from the information $Ix = (I_1x, \ldots, I_nx)$. The problem of optimal recovery is to determine the error of optimal recovery and an optimal recovery method of $Tx$ on the set $W$ from the precise information $Ix$. Any operator $m : \mathbb{R}^n \to \mathbb{R}$ is admitted as a recovery method.

In the current setting the information will be known precisely. For a given $f \in W$, let

$$Vf = \{g \in W : Ig = If\},$$
that is the pre-image of the set $If$ in $W$. Pushing this set forward through the operator $T$ let

$$Uf = \{Tg : g \in Vf\}.$$  

As more than one function can have the same information, then $m(If) = m(Ig)$ for each $f, g \in Vf$ and so a method of recovery needs to approximate all $g \in Uf$. As the value of $\|m(Ix) - Tx\|$ will vary depending on the chosen $x \in Vf$, we have the following definitions.

**Definition 2.** The *error of the method* $m$, is given by

$$e(T, W, I, m) = \sup_{x \in W} \|m(Ix) - Tx\|.$$  

**Definition 3.** The *optimal error* is defined as

$$E(T, W, I) = \inf_{m : I(W) \to Z} e(T, W, I, m)$$

and any method $\hat{m}$ obtaining this error is an *optimal method of recovery*, that is if and only if

$$E(T, W, I) = e(T, W, I, \hat{m}).$$

**Theorem 5.** *(Smolyak [29])* Let $X$ be a real vector space, $W \subseteq X$ a balanced convex
set, $T$ a linear functional on $W$ and $l_1, \ldots, l_n$ linear real information functionals, $I = (l_1, \ldots, l_n)$. Then among optimal methods of recovery there exists a linear method. Let

$$
\psi_k(y) = \sup_{x \in W} \frac{Tx}{l_k x = y\delta_{kj}} \quad (3.1)
$$

where $k = 1, \ldots, n$ and $\delta_{jk}$ is the Kronecker $\delta$ function. If $\psi_k$ is differentiable at 0 for each $1 \leq k \leq n$ then

$$
\hat{m}(Ix) = \sum_{j=1}^{n} \psi_j'(0)l_jx
$$

is the unique optimal method of recovery.

The theorem shows that to determine the optimal error of recovery we consider values of $Tx$ where $x \in W \cap \ker I$. The following proof is a modified form of the proofs found in [2] and [30].

**Proof.** First, if the error of optimal recovery is infinite then any method of recovery is optimal and hence there exists a linear method amongst optimal ones.

Let us assume the error of optimal recovery belongs to the interval $(0, \infty)$. Define

$$
Y = \{(Tx, l_1f, \ldots, l_nf) : x \in W\} \subset \mathbb{R}^{n+1}
$$

which is a closed balanced convex set.

Now, for any convex set $A$, and any point $p \in \partial A$ there exists a support hyperplane that contains $p$. A support hyperplane seperates the space into two pieces, one that contains all points of $A$ and one that contains none. Let us consider the boundary point $(D_0, 0, \ldots, 0) \in \partial Y$ where

$$
D_0 = \sup_{(y_0,0,\ldots,0) \in Y} y_0.
$$
Let the support hyperplane to $Y$ through the boundary point $(D_0, 0, \ldots, 0)$ be given by

$$C_0(y_0 - D_0) + \sum_{j=1}^{n} C_j l_j x = 0$$

for some $C_j \in \mathbb{R}$. Note that the hyperplane of support can be chosen such that $C_0 > 0$. With $C_0 > 0$, for any $(y_0, \ldots, y_n) \in Y$,

$$y_0 - \sum_{j=1}^{n} \frac{C_j}{C_0} l_j x \leq D_0$$

and as $Y$ is symmetric with respect to the origin then a similar argument shows

$$y_0 - \sum_{j=1}^{n} \frac{C_j}{C_0} l_j x \geq -D_0$$

and so $Y$ lies between the two support hyperplanes. So for any $x \in W$

$$|Lx - \sum_{j=1}^{n} D_j l_j x| \leq D_0, \quad D_j = -\frac{C_j}{C_0}. \quad (3.2)$$

Define $\tilde{m}(Ix) = \sum_{j=1}^{n} \frac{C_j}{C_0} l_j x$.

Let $x_0 \in W$ be an element such that $\sup_{(Tx, 0, \ldots, 0) \in Y} Tx_0 = D_0$. Then for any method $m$,

$$2D_0 = 2|Tx_0| = |Tx_0 + Tx_0| = |Tx_0 - T(-x_0)|$$

$$= |Tx_0 - m(0) - (T(-x_0) - m(0))|$$

$$\leq |Tx_0 - m(0)| + |T(-x_0) - m(0)|$$

$$\leq 2e(T, W, I, m). \quad (3.3)$$
Since $m$ was arbitrary, $D_0 \leq E(T,W,I)$ and appealing to 3.2 and 3.3,

$$E(T,W,I) \leq e(T,W,I,\tilde{m}) \leq D_0 \leq E(T,W,I).$$

Thus $e(T,W,I,\tilde{m}) = E(T,W,I)$ and so $\tilde{m}$ is an optimal method of recovery.

Now let $\phi(Ix) = \sum_{j=1}^{n} q jl_jx$ be an optimal error algorithm. Choose $x \in W$ satisfying $l_kx = y$ and $l_jx = 0$ for $j \neq k$ and small $y$. Then $|Tx - y\psi_k| \leq e(T,W,I,\phi) = \psi_k(0) < \infty$. By 3.1, $Tx$ can be made arbitrarily close to $\psi_k(y)$ and so we have $\psi_k(y) - yq_k \leq \psi_k(0)$. Rearranging and dividing by $y$ we obtain,

$$\frac{\psi_k(|y|) - \psi_k(0)}{|y|} \leq q_k \leq \frac{\psi_k(-|y|) - \psi_k(0)}{-|y|}. \quad (3.4)$$

Assuming $\psi_k(0)$ exists, then letting $|y| \to 0$ we obtain $q_k = \psi'_k(0)$. Therefore any linear optimal method of recovery is precisely the unique linear optimal method of recovery $\phi(Ix) = \sum_{j=1}^{n} \psi'_j(0)l_jx$.

\[\Box\]

This theorem of Smolyak was instrumental in the development of the current ideas of optimal recovery. In what follows, the ideas are generalized to complex linear spaces, recovery of linear operators, and to recovery involving inaccurate information. There are many important results that were developed as a result of 5. The results that are to be explicitly used in Chapter 4 are presented here.

### 3.1 General Setting

Let $W$ be a subset of a linear space $X$, let $Z$ be a normed linear space, and $T$ the linear operator $T : X \to Z$ that we are trying to recover on $W \subset X$ from given information. This information is provided by a linear operator $I : W \to Y$ where $Y$ is a normed linear space. For any $x \in W$ we know some $\tilde{y} \in Y$ that is near $Ix$. That is, we know $\tilde{y} \in Y$
such that
\[ \| Ix - \tilde{y} \|_Y \leq \delta \] (3.5)
for some \( \delta > 0 \). The value \( \tilde{y} \) is our inaccurate information. Now we try to approximate the value of \( Tx \) from \( \tilde{y} \) using an algorithm or method, \( m \). Define a method to be any mapping \( m : Y \rightarrow Z \), and regard \( m(\tilde{y}) \) as the approximation to \( Tx \) from the information \( \tilde{y} \in Y \).

\[ W \subseteq X \xrightarrow{T} Z \]
\[ \xleftarrow{I} \]
\[ \xleftarrow{m} \]
\[ \downarrow \]
\[ \downarrow \]
\[ Y \]

Figure 3.2

Our goal is to minimize the difference of \( Tx \) and \( m(\tilde{y}) \) in \( Z \), i.e. minimize
\[ \| Tx - m(\tilde{y}) \|_Z. \]

However, the size of \( \| Tx - m(\tilde{y}) \|_Z \) varies since \( \tilde{y} \) can be chosen to be any \( \tilde{y} \in Y \) satisfying (3.5). Furthermore \( Ix \) varies depending on the \( x \in W \) chosen. So the error of any single method is defined as
\[ e(T, W, Y, m) = \sup_{\substack{x \in W, \ y \in Y \ \|Ix - \tilde{y}\|_Y \leq \delta}} \| Tx - m(\tilde{y}) \|_Z. \]

Now the optimal error is that of the method with the smallest error. Thus the error of
optimal recovery is defined as

$$E(T, W, Y) = \inf_{m: Y \to Z} \sup_{x \in W, \bar{y} \in Y} \| Tx - m(\bar{y}) \|_Z = \inf_{m: Y \to Z} e(T, W, Y, m).$$

More specifically, let $Y_1, \ldots, Y_n$ be linear spaces with semi-inner norms $\| \cdot \|_{Y_k}$ and $I_k : X \to Y_k$ linear operators, $I = (I_1, \ldots, I_n)$. We want to recover $Tx$ for

$$x \in W_k = \{ x \in X : \| I_j x \|_{Y_j} \leq \delta_j, 1 \leq j \leq k, 0 \leq k \leq n \},$$

(where if $k = 0$ we let $W_0 = X$), if we know the values $\bar{y}_j$ satisfying $\| I_j x - \bar{y}_j \|_{Y_j} \leq \delta_j$ for $j = k+1, \ldots, n$. So the information operators are only the operators $I_j$ for $j = k+1, \ldots, n$ as we know estimated values of $I_j x$ to within $\delta_j$. Using the data provided inaccurately by the information operators $I_j$, $k < j \leq n$, we try to recover $Tx$ on the set $W_k$ defined using $I_j$ for $j \leq k$. Let $F = (I_{k+1}, \ldots, I_n)$ be the multi-valued operator given by

$$F(x) = \{ y = (y_{k+1}, \ldots, y_n) \in Y_{k+1} \times \cdots \times Y_n : \| I_j x - y_j \|_{Y_j} \leq \delta_j, j = k+1, \ldots, n \}.$$

Consequently, a method of recovery is any operator $m : Y_{k+1} \times \cdots \times Y_n \to Z$ and the error of this method is:

$$e(T, W_k, I, \delta, m) = \sup_{x \in W_k} \sup_{(\bar{y}_{k+1}, \ldots, \bar{y}_n) \in Y_{k+1} \times \cdots \times Y_n} \| Tx - m(\bar{y}) \|_Z$$

where $\delta = (\delta_1, \ldots, \delta_n)$. The intrinsic error or error of optimal recovery for the given problem is defined as the value:

$$E(T, W_k, I, \delta) = \inf_{m: Y_{k+1} \times \cdots \times Y_n \to Z} e(T, W_k, I, \delta, m)$$

and if $\hat{m}$ satisfies such an equality then $\hat{m}$ is an optimal method. In the following theorems
due to Magaril-II’yaev and Osipenko, this problem of optimal recovery is closely related to the duality extremal problem

\[ \|Tx\|_Z^2 \rightarrow \max, \quad \|I_jx\|_{Y_j}^2 \leq \delta_i^2, \quad j = 1, \ldots, n, \quad x \in X. \tag{3.6} \]

Application of duality to extremal problems with linear functionals may be found in [6].

### 3.2 Construction of Optimal Method and Error

The following results of G. G. Magaril-II’yaev and K. Yu. Osipenko are applied to several problems of optimal recovery and so proofs are presented. The included proofs can be found in [20].

**Theorem 6.** Assume that there exist \( \hat{\lambda}_j \geq 0, \quad j = 1, \ldots, n \) such that the value of the extremal problem

\[ \|Tx\|_Z^2 \rightarrow \max, \quad \sum_{j=1}^{n} \hat{\lambda}_j \|I_jx\|_{Y_j}^2 \leq \sum_{j=1}^{n} \hat{\lambda}_j \delta_{ij}^2, \quad x \in X \tag{3.7} \]

is the same as in (3.6). Also assume for each \( y = (y_1, \ldots, y_n) \in Y_1 \times \cdots \times Y_n \) there exists \( x_y = x(y_1, \ldots, y_n) \) which is a solution to

\[ \sum_{j=1}^{n} \hat{\lambda}_j \|I_jx - y_j\|_{Y_j}^2 \rightarrow \min \quad x \in X. \tag{3.8} \]

Then for all \( k, 0 \leq k < n, \)

\[ E(T, W_k, I, \delta) = \sup_{x \in X} \|Tx\|_Z \quad \text{subject to} \quad \|I_jx\|_{Y_j} \leq \delta, \quad j = 1, \ldots, n \]

and the method

\[ \hat{m}(y_{k+1}, \ldots, y_n) = Tx(0, \ldots, 0, y_{k+1}, \ldots, y_n) \tag{3.9} \]
is optimal.

Proof. (The following proof is essentially taken from [20].)

The set

$$F^{-1}(0) = \{ x \in W_k : \| I_j x \|_{Y_j} \leq \delta_j, j = k + 1, \ldots, n \} =$$

is centrally symmetric. That is for each $x \in F^{-1}(0)$ we have $-x \in F^{-1}(0)$. Therefore for any method $m$,

$$2\|Tx\|_Z = \|Tx + T(-x)\|_Z = \|Tx - m(0) - (T(-x) - m(0))\|_Z \leq \|Tx - m(0)\|_Z + \|-T(-x) - m(0)\|_Z \leq 2e(T, W, F, m).$$

As the last inequality is independent of $x \in F^{-1}(0)$ we have for any method $m$,

$$e(T, W, F, m) \geq \sup_{x \in F^{-1}(0)} \|Tx\|_Z.$$

Since $m$ was arbitrary, then inequality holds for optimal $m$ and hence

$$E(T, W, F) \geq \sup_{x \in F^{-1}(0)} \|Tx\|_Z \geq \sup_{x \in W_k} \|Tx\|_Z \| I_j x \|_{Y_j} \leq \delta_j, j = k + 1, \ldots, n \geq \sup_{x \in X} \|Tx\|_Z \| I_j x \|_{Y_j} \leq \delta_j, j = 1, \ldots, n.$$ (3.10)

This then yields a lower bound on the error of optimal recovery. For an upper bound, let $Y = Y_1 \times \cdots Y_n$ with semi-inner product defined by

$$((y_1^1, \ldots, y_n^1), (y_1^2, \ldots, y_n^2))_Y = \sum_{j=1}^n \lambda_j (y_j^1, y_j^2)_Y.$$
Then
\[ \|Ix - y\|^2_Y = \sum_{j=1}^n \lambda_j \|I_jx - y_j\|^2_{Y_j} \]
and so extremal problem (3.8) can be rewritten as
\[ \|Ix - y\|^2_Y \to \min, \quad x \in X. \quad (3.11) \]

Let \(x_y\) be a solution to this extremal problem. Assume \((Ix_y - y, Ix)_Y = \alpha \neq 0\) for some \(x \in X\). Then \(Ix_y - \frac{\alpha}{\|Ix\|^2_Y}Ix \in Y\) and
\[
\left\|Ix_y - \frac{\alpha}{\|Ix\|^2_Y}Ix - y\right\|^2_Y = \left(Ix_y - \frac{\alpha}{\|Ix\|^2_Y}Ix - y, Ix_y - \frac{\alpha}{\|Ix\|^2_Y}Ix - y\right)_Y \\
= \|Ix_y - y\|^2_Y - 2 \text{Re} \left( Ix_y - y, \frac{\alpha}{\|Ix\|^2_Y}Ix \right)_Y + \frac{\|\alpha\|^2}{\|Ix\|^4_Y} \|Ix\|^2_Y \\
= \|Ix_y - y\|^2_Y - 2 \text{Re} \left( \frac{\pi}{\|Ix\|^2_Y} (Ix_y - y, Ix)_Y \right) + \frac{\|\alpha\|^2}{\|Ix\|^4_Y} \\
= \|Ix_y - y\|^2_Y - 2 \frac{\|\alpha\|^2}{\|Ix\|^2_Y} + \frac{\|\alpha\|^2}{\|Ix\|^4_Y} \\
< \|Ix_y - y\|^2_Y.
\]
This contradicts \(x_y\) being a solution to extremal problem (3.11) and hence \(0 = \alpha\) and so for any \(x \in X\), \((Ix_y - y, Ix)_Y = 0\). Therefore
\[
\|Ix - y\|^2_Y = \|Ix - Ix_y + Ix_y - y\|^2_Y \\
= \|Ix - Ix_y\|^2_Y + 2 \text{Re}(Ix - Ix_y, Ix_y - y)_Y + \|Ix_y - y\|^2_Y \\
= \|Ix - Ix_y\|^2_Y + 2 \text{Re}(Ix_y - y, I(Ix - x_y))_Y + \|Ix_y - y\|^2_Y \\
= \|Ix - Ix_y\|^2_Y + \|Ix_y - y\|^2_Y.
\]
and hence for any $x \in X$

$$\|Ix - Ix_y\|_Y^2 \leq \|Ix - y\|_Y^2 = \sum_{j=1}^n \hat{\lambda}_j \|I_j x - y_j\|_Y^2.$$  \hfill (3.12)

Now let $x \in X$ be arbitrary and $y \in Y$ be given such that $y = (0, \ldots, 0, y_{k+1}, \ldots, y_n)$ with $\|I_j x - y_j\|_Y \leq \delta_j$ for $j = k + 1, \ldots, n$ and $x_y$ a solution to (3.11) for the given $y$. Define $z = x - x_y$ and then

$$\|Iz\|_Y^2 = \|Ix - Ix_y\|_Y^2 \leq \sum_{j=1}^n \hat{\lambda}_j \|I_j x - y_j\|_Y^2 \leq \sum_{j=1}^n \hat{\lambda}_j \delta_j.$$

For the method $\hat{m}(y) = Tx_y$ we can estimate the error by

$$\|Tx - \hat{m}(y)\|_Z = \|Tx - Tx_y\|_Z = \|T(x - x_y)\|_Z = \|Tz\|_Z \leq \sup_{x \in X} \|Tz\|_Z \leq \sum_{j=1}^n \hat{\lambda}_j \|I_j x\|_Y^2 \leq \sum_{j=1}^n \hat{\lambda}_j \delta_j = \sup_{x \in X} \|Tx\|_Z \|I_j x\|_Y \leq \delta_j, \quad j = 1, \ldots, n$$

The last equality is due to the assumption of coincidence of the values of extremal problems (3.6) and (3.8). Therefore

$$E(T, W_k, I, \delta) \leq e(T, W_k, I, \delta, \hat{m}) \leq \sup_{x \in X} \|Tx\|_Z \|I_j x\|_Y \leq \delta_j, \quad j = 1, \ldots, n$$

Together with the lower bound, (3.10), we obtain the error of optimal recovery

$$E(T, W_k, I, \delta) = \sup_{x \in X} \|Tx\|_Z \|I_j x\|_Y \leq \delta_j, \quad j = 1, \ldots, n$$

and hence the method $\hat{m}$ is an optimal method of recovery.
Theorem 6 gives a constructive approach to finding an optimal method $\hat{m}$ from the information. It follows from results obtained in [10]–[21] (see also [22] where this theorem was proven for one particular case.)

3.3 Dual Problem Equivalence

In order to apply Theorem 6 the values of extremal problems (3.7) and (3.6) must agree. The following result, also due to G. G. Magaril-I'lyaev and K. Yu Osipenko, provides conditions under which the values of problems (3.7) and (3.6) will agree.

Typically, when one encounters extremal problems, one approach is to construct the Lagrange function $L$. For an extremal problem of the form of (3.7), the corresponding Lagrange function is

$$L(x, \lambda_1, \ldots, \lambda_n) = -\|Tx\|_Z^2 + \sum_{j=1}^{n} \lambda_j \|I_jx\|_{Y_j}^2.$$ 

Furthermore, $\hat{x} \in X$ is called an extremal element if $\|I_j\hat{x}\|_{Y_j}^2 \leq \delta_j^2$ for $j = 1, \ldots, n$ and thus admissible in (3.7) and

$$\|Tx\|_Z^2 = \sup_{x \in X, \|I_jx\|_{Y_j}^2 \leq \delta_j^2} \|Tx\|_Z^2.$$ 

**Theorem 7.** If $\hat{\lambda}_j \geq 0$, $j = 1, \ldots, n$ and $\hat{x} \in X$ satisfies $\|I_j\hat{x}\|_{Y_j}^2 \leq \delta_j^2$ and

(a) $\min_{x \in X} L(x, \hat{\lambda}_1, \ldots, \hat{\lambda}_n) = L(\hat{x}, \hat{\lambda}_1, \ldots, \hat{\lambda}_n)$

(b) $\sum_{j=1}^{n} \hat{\lambda}_j (\|I_j\hat{x}\|_{Y_j}^2 - \delta_j^2) = 0$
then \( \hat{x} \) is an extremal element and

\[
\sup_{x \in X, \|I_j x\|_{Y_j}^2 \leq \delta_j^2} \|Tx\|_Z^2 = \sup_{x \in X, \sum_{j=1}^n \lambda_j \|I_j x\|_{Y_j}^2 \leq \sum_{j=1}^n \lambda_j \delta_j^2} \|Tx\|_Z^2 = \sum_{j=1}^n \lambda_j \delta_j^2.
\]

**Proof.** (Taken from [20].)

Let \( x, \hat{x} \in X \) be admissible in (3.6), that is \( \|I_j x\|_{Y_j}^2 \leq \delta_j^2 \) for \( j = 1, \ldots, n \), with \( \hat{x} \) satisfying conditions (a) and (b) of the theorem. Then by condition (a) we have

\[
-\|Tx\|_Z^2 \geq -\|Tx\|_Z^2 - \sum_{j=1}^n \lambda_j (\delta_j^2 - \|I_j x\|_{Y_j}^2) = (-\|Tx\|_Z^2 + \sum_{j=1}^n \lambda_j \|I_j x\|_{Y_j}^2) - \sum_{j=1}^n \lambda_j \delta_j^2.
\]

\[
\geq (-\|T\hat{x}\|_Z^2 + \sum_{j=1}^n \lambda_j \|I_j \hat{x}\|_Z^2) - \sum_{j=1}^n \lambda_j \delta_j^2.
\]

Furthermore, by (b) we have

\[
(-\|T\hat{x}\|_Z^2 + \sum_{j=1}^n \lambda_j \|I_j \hat{x}\|_Z^2) - \sum_{j=1}^n \lambda_j \delta_j^2 = -\|T\hat{x}\|_Z^2 + \sum_{j=1}^n \lambda_j (\|I_j x\|_{Y_j}^2 - \delta_j^2) = -\|T\hat{x}\|_Z^2.
\]

So for any \( x \in X \),

\[
\|Tx\|_Z \leq \|T\hat{x}\|_Z
\]

and \( \hat{x} \) is extremal in (3.6). To see that \( \hat{x} \) is extremal in (3.8) notice that if \( x, \hat{x} \) are admissible in (3.8) then the same string of relations holds. Thus the values of extremal problems (3.6) and (3.8) coincide. To show \( \mathcal{L}(\hat{x}, \lambda) = 0 \) assume not. If \( \mathcal{L}(\hat{x}, \lambda) = \epsilon > 0 \) then let \( x_0 = \alpha \hat{x} \) with \( \alpha < 1 \). Then

\[
\mathcal{L}(x_0, \lambda) = \mathcal{L}(\alpha \hat{x}, \lambda) = \alpha^2 \mathcal{L}(\hat{x}, \lambda) < \mathcal{L}(\hat{x}, \lambda),
\]

and if \( \mathcal{L}(\hat{x}, \lambda) = \epsilon < 0 \), choose \( \alpha > 1 \) so that

\[
\mathcal{L}(x_0, \lambda) = \mathcal{L}(\alpha \hat{x}, \lambda) = \alpha^2 \mathcal{L}(\hat{x}, \lambda) < \mathcal{L}(\hat{x}, \lambda).
\]
In either case, the minimality of condition (a) is contradicted and hence \( \|T\hat{x}\|_Z^2 = \sum_{j=1}^n \hat{\lambda}_j \|I_j\hat{x}\|_{Y_j}^2 \). As \( \hat{x} \) is extremal in (3.6) we obtain by condition (b)

\[
\sup_{x \in X, \|I_jx\|_{Y_j}^2 \leq \delta_j^2} \|Tx\|_Z^2 = \|T\hat{x}\|_Z^2 = \sum_{j=1}^n \hat{\lambda}_j \|I_j\hat{x}\|_Z^2 = \sum_{j=1}^n \hat{\lambda}_j \delta_j^2.
\]

\[
(-\|T\hat{x}\|_Z^2 + \sum_{j=1}^n \hat{\lambda}_j \|I_j\hat{x}\|_{Y_j}^2) - \sum_{j=1}^n \hat{\lambda}_j \delta_j^2 = -\|T\hat{x}\|_Z^2 + \sum_{j=1}^n \hat{\lambda}_j (\|I_jx\|_{Y_j}^2 - \delta_j^2) = -\|T\hat{x}\|_Z^2.
\]

So for any \( x \in X \),

\[
\|Tx\|_Z \leq \|T\hat{x}\|_Z
\]

and \( \hat{x} \) is extremal in (3.6). To see that \( \hat{x} \) is extremal in (3.8) notice that if \( x, \hat{x} \) are admissible in (3.8) then the same string of relations holds. Thus the values of extremal problems (3.6) and (3.8) coincide. To show \( L(\hat{x}, \hat{\lambda}) = 0 \) assume not. If \( L(\hat{x}, \hat{\lambda}) = \epsilon > 0 \) then let \( x_0 = \alpha \hat{x} \) with \( \alpha < 1 \). Then

\[
L(x_0, \hat{\lambda}) = L(\alpha \hat{x}, \hat{\lambda}) = \alpha^2 L(\hat{x}, \hat{\lambda}) < L(\hat{x}, \hat{\lambda}),
\]

and if \( L(\hat{x}, \hat{\lambda}) = \epsilon < 0 \), choose \( \alpha > 1 \) so that

\[
L(x_0, \hat{\lambda}) = L(\alpha \hat{x}, \hat{\lambda}) = \alpha^2 L(\hat{x}, \hat{\lambda}) < L(\hat{x}, \hat{\lambda}).
\]

In either case, the minimality of condition (a) is contradicted and hence \( \|T\hat{x}\|_Z^2 = \sum_{j=1}^n \hat{\lambda}_j \|I_j\hat{x}\|_{Y_j}^2 \). As \( \hat{x} \) is extremal in (3.6) we obtain by condition (b)

\[
\sup_{x \in X, \|I_jx\|_{Y_j}^2 \leq \delta_j^2} \|Ix\|_Z^2 = \|I\hat{x}\|_Z^2 = \sum_{j=1}^n \hat{\lambda}_j \|I_j\hat{x}\|_Z^2 = \sum_{j=1}^n \hat{\lambda}_j \delta_j^2.
\]

\[
\square
\]

If we wish to combine Theorems 6 and 7 to determine an optimal error and method
then we must show the posed problem is able to satisfy equating extremal problems (3.7) and (3.6). Through Theorem 7 we have such a means available. If it can be shown that the problem satisfies (a), (b) of Theorem 7, then Theorem 6 provides a way to determine an optimal method and Theorem 7 also specifies the error of such a method.

Each of the previous theorems gives conditions under which a linear optimal method of recovery is guaranteed in varying situations. For an example in which no linear method of optimal recovery exists, refer to [14].
Chapter 4

Proofs of Results

4.1 Radial Integral Information Operator Problems

4.1.1 Theorem 1: Inaccuracy in $L_2(\mathbb{T})$ Norm

To obtain the optimal error and a method of optimal recovery we will first need a short lemma concerning sequences of real numbers.

Let $\{\alpha_j\}_{j \in \mathbb{N}}, \{\beta_j\}_{j \in \mathbb{N}}$ be sequences of non-negative real numbers that satisfy

$$\lim_{j \to \infty} \beta_j = \infty, \quad \lim_{j \to \infty} \frac{\beta_j}{\alpha_j} = 0.$$ 

Define $C = \text{co}\{(0,0) \cup \{(\alpha_j, \beta_j)\}_{j \in \mathbb{N}}\}$ to be the convex hull. Define $\theta(x)$ for $x \in [0, \infty)$ by

$$\theta(x) = \max\{y : (x, y) \in C\}.$$ 

Lemma 1. The piecewise linear function $\theta$ with points of break $(x_s, y_s) \ s = 0, 1, \ldots$, with $x_0 < x_1 < \ldots$, for $x \in [x_s, x_{s+1}]$ given by $\theta(x) = \lambda_1 x + \lambda_2$ is such that $\lambda_1, \lambda_2 \geq 0$.

Proof. Assume that

$$\lambda_1 = \frac{y_{s+1} - y_s}{x_{s+1} - x_s} < 0.$$ 

It means that $y_{s+1} < y_s$. Since $\beta_j \to \infty$ and $\alpha_j \to \infty$ as $j \to \infty$ there is a $k \in \mathbb{N}$ such that $\alpha_k > x_s$ and $\beta_k > y_s$. Then the interval between $(x_s, y_s)$ and $(\alpha_k, \beta_k)$ belongs to $C$. Consequently, $\theta(x_{s+1}) > y_{s+1}$ and $(x_{s+1}, y_{s+1})$ is not a point of break of $\theta$.

Assume that $\lambda_2 < 0$. Since $(0,0) \in C$ the interval between $(0,0)$ and $(x_{s+1}, y_{s+1})$ belongs to $C$. Geometrically, the line $(0,0)$ to $(x_{s+1}, y_{s+1})$ will lie above the line $\theta(x) = \lambda_1 x + \lambda_2$. It means that $\theta(x_s) > y_s$ contradicting that $(x_s, y_s)$ is a point of break of $\theta$.

Note that as $\lim_{j \to \infty} \frac{\beta_j}{\alpha_j} = 0$ then for any fixed $(\alpha_b, \beta_b)$ the ratio of slopes between points $(\alpha_b, \beta_b)$ and $(\alpha_j, \beta_j)$ also tends to 0 as

$$\lim_{j \to \infty} \frac{\beta_j - \beta_b}{\alpha_j - \alpha_b} = \lim_{j \to \infty} \frac{\beta_j}{\alpha_j} - \frac{\beta_b}{\alpha_b} = 0.$$  

Consequently, there will be an infinite number of distinct points of break for $\theta$.

Proof. (Theorem 1)
Consider the dual extremal problem

$$\|f\|_{L_2(D)}^2 \to \max, \quad \|f\|_X^2 \leq 1, \quad \|Kf\|_{L_2(T)}^2 \leq \delta^2$$

(4.1)

which can be written as

$$\sum_{j=0}^{\infty} \frac{1}{j+1} |a_j|^2 \to \max, \quad \sum_{j=0}^{\infty} \gamma_j |a_j|^2 \leq 1, \quad \sum_{j=0}^{\infty} \frac{1}{(j+1)^2} |a_j|^2 \leq \delta^2$$

where \( f(z) = \sum_{j=0}^{\infty} a_j z^j \). Define the corresponding Lagrange function as

$$\mathcal{L}(f, \lambda_1, \lambda_2) = \sum_{j=0}^{\infty} \frac{|a_j|^2}{(j+1)^2} (\lambda_1 \gamma_j (j+1)^2 + \lambda_2 - (j+1)).$$

Let the line segment between successive points \((x_s, y_s)\) and \((x_{s+1}, y_{s+1})\) be given by \( \theta_s(x) = \lambda_1 x + \lambda_2 \). That is \( \theta_s = \theta|_{[x_s,x_{s+1}]} \). Thus \( \lambda_1, \lambda_2 \) are given by (2.8). Take any \((\gamma_j (j+1)^2, j+1)\), then by definition of the function \( \theta \) we have

$$j + 1 \leq \theta(\gamma_j (j+1)^2) \leq \theta_s(\gamma_j (j+1)^2).$$

Thus for all \( j = 0, 1, \ldots \)

$$j + 1 \leq \lambda_1 \gamma_j (j+1)^2 + \lambda_2$$

and hence \( \mathcal{L}(f, \lambda_1, \lambda_2) \geq 0 \) for any \( f \in X \).

We proceed to the construction of a function \( \hat{f} \) admissible in (4.1) that also satisfies

$$\lambda_1(\|\hat{f}\|_X^2 - 1) + \lambda_2(\|K\hat{f}\|_{L_2(T)}^2 - \delta^2) = 0.$$

Assume \( y_s > 0 \). As \( y_s = 0 \) if and only if \( s = 0 \) and \( y_0 = 0 \) then \( y_s > 0 \) if and only if \( s > 0 \) or \( y_0 > 0 \). Let \( k, k' \in \mathbb{N} \) be the indices that satisfy \((x_s, y_s) = (\gamma_k (k+1)^2, k+1)\) and \((x_{s+1}, y_{s+1}) = (\gamma_{k'} (k'+1)^2, k'+1)\). We let \( a_j = 0 \) for \( j \neq k, k' \), and choose \( a_k, a_{k'} \).
so that they satisfy the conditions

\[ \gamma_k |a_k|^2 + \gamma_{k'} |a_{k'}|^2 = 1 \quad \text{and} \quad \frac{1}{(k+1)^2} |a_k|^2 + \frac{1}{(k'+1)^2} |a_{k'}|^2 = \delta^2. \] (4.2)

From these conditions let

\[ a_k = (k + 1) \left( \frac{\delta^2 \gamma_{k'} (k' + 1) - 1}{\gamma_{k'} (k' + 1)^2 - \gamma_k (k + 1)^2} \right)^{1/2} = y_s \left( \frac{\delta^2 x_{s+1} - 1}{x_{s+1} - x_s} \right)^{1/2} \]

\[ a_{k'} = (k' + 1) \left( \frac{1 - \delta^2 \gamma_k (k + 1)^2}{\gamma_{k'} (k' + 1)^2 - \gamma_k (k + 1)^2} \right)^{1/2} = y_{s+1} \left( \frac{1 - \delta^2 x_s}{x_{s+1} - x_s} \right)^{1/2} \] (4.3)

and

\[ \hat{f}(z) = a_k z^k + a_{k'} z^{k'}. \] (4.4)

Now if \( x_s < \delta^{-2} \leq x_{s+1} \) with \( s > 0 \) or \( s = 0 \) and \( y_0 > 0 \) the function \( \hat{f} \) is admissible in (4.1) and \( \mathcal{L}(\hat{f}, \hat{\lambda}_1, \hat{\lambda}_2) = 0 \), that is \( \hat{f} \) minimizes \( \mathcal{L}(f, \hat{\lambda}_1, \hat{\lambda}_2) \) and condition (a) of Theorem 7 is satisfied. Furthermore, by construction, \( \hat{f} \) satisfies condition (b) of Theorem 7.

If \( y_s = 0 \), that is \( s = 0 \) and \( y_0 = 0 \), define \( a_{k'} \) as in (4.3). Then as \( x_0 = 0 \)

\[ a_{k'} = y_1 \left( \frac{1 - \delta^2 x_0}{x_1 - x_0} \right)^{1/2} = y_1 x_1^{-1/2} = \gamma_{k'}^{-1/2}. \]

So let \( \hat{f}(z) = a_{k'} z^{k'} \) and we have

\[ \| \hat{f} \|^2_X = 1, \quad \| K \hat{f} \|^2_{L^2(\mathbb{T})} = \frac{1}{\gamma_{k'} (k' + 1)^2} = \frac{1}{x_1} \leq \delta^2. \]

Thus the function \( \hat{f} \) is admissible in (4.1) and satisfies (a) and (b) of Theorem 7. It should be noted that in this case \( \hat{\lambda}_1, \hat{\lambda}_2 \) are simply \( \hat{\lambda}_1 = y_1 / x_1 \) and \( \hat{\lambda}_2 = 0 \).

Now we proceed to the extremal problem

\[ \hat{\lambda}_2 \| K f - \tilde{y} \|^2_{L^2(\mathbb{T})} + \hat{\lambda}_1 \| f \|^2_X \rightarrow \min, \quad f \in W, \quad \tilde{y} \in L^2(\mathbb{T}). \] (4.5)
This problem may be rewritten as
\[
\hat{\lambda}_2 \sum_{-\infty}^{\infty} |\tilde{y}_j|^2 + \hat{\lambda}_2 \sum_{j=0}^{\infty} \frac{1}{j+1} a_j - \tilde{y}_j|^2 + \hat{\lambda}_1 \sum_{j=0}^{\infty} \gamma_j |a_j|^2 \rightarrow \min
\]
which has solution
\[
a_j = \begin{cases} 
(1 + \hat{\lambda}_1 \hat{\lambda}_2^{-1} \gamma_j(j+1)^2)^{-1} (j+1)\tilde{y}_j : j \geq 0, \hat{\lambda}_2 \neq 0 \\
0 : \text{otherwise}
\end{cases}
\]

So for \( x_s < \delta^{-2} \leq x_{s+1}, \ y_s > 0 \) by Theorems 6 and 7, (2.10) is an optimal method and the error of optimal recovery is given by (2.9). If \( y_s = 0 \) then \( E(W, K, \delta) = \sqrt{\hat{\mu}/x_1} \) and \( \hat{\mu}(\tilde{y}) = 0 \) is an optimal method.

4.1.2 Theorem 2: Inaccuracy in \( l^N_2(\mathbb{T}) \) Norm

Proof. For the cases \( x_s < \delta^{-2} \leq x_{s+1} \) with \( s < s_0 \) we simply apply the same structure of proof as in Theorem 1. For the case \( \delta^{-2} > x_{s_0} \) there remains some work.

Our construction will depend on whether or not \( y_{s_0} = 0 \), that is whether or not \( s_0 = 0 \) with \( (x_{s_0}, y_{s_0}) = (0, 0) \).

First we notice \( y_{s_0} \leq N \). Assume not. Then if \( y_{s_0} > N > 0 \) we also know \( x_{s_0} > 0 \) since for all \( j \geq N \) we assumed \( \gamma_j > 0 \). Since \( x_0 = 0 \) we know \( s_0 > 0 \). Then by definition of \( s_0 \) we know for \( 0 \leq s < s_0 \),
\[
\frac{y_{s+1} - y_s}{x_{s+1} - x_s} > (\gamma_0(l_0 + 1))^{-1}
\]
and substituting \( s = s_0 - 1 \) we have
\[
\frac{y_{s_0}}{x_{s_0}} \geq \frac{y_{s_0} - y_{s_0-1}}{x_{s_0} - x_{s_0-1}} > (\gamma_0(l_0 + 1))^{-1}
\]
which contradicts the definition of \( l_0 \). Therefore \( y_{s_0} \leq N \) and if \( (x_{s_0}, y_{s_0}) = (\gamma_c(c + 1)^2, c + 1) \) then \( c < N \).

In either case, \( y_{s_0} = 0 \) or \( y_{s_0} > 0 \), the dual problem is of the form

\[
\|f\|_{L_2(\mathbb{D})}^2 \to \max, \quad \|f\|_X^2 \leq 1, \quad \sum_{j=0}^{N-1} |K_j f|^2 \leq \delta^2, \quad f \in X. \tag{4.6}
\]

The corresponding Lagrange function is then

\[
\mathcal{L}(f, \lambda_1, \lambda_2) = \sum_{j=0}^{\infty} \frac{|a_j|^2}{(j+1)^2} \left( \lambda_1 \gamma_j(j+1)^2 + \chi_j N \lambda_2 - (j + 1) \right) \tag{4.7}
\]

where \( \chi_j N \) is the characteristic function of \( \{ j \in \mathbb{N} : j < N \} \).

**Case i: \( y_{s_0} > 0 \)**

If \( y_{s_0} > 0 \) let \( c \) correspond to the index satisfying

\[
(x_{s_0}, y_{s_0}) = (\gamma_c(c + 1)^2, c + 1).
\]

To determine \( \hat{\lambda}_1, \hat{\lambda}_2 \) let \( y = \hat{\lambda}_1 x + \hat{\lambda}_2 \) be the line through the point \( (x_{s_0}, y_{s_0}) \) that is parallel to the line from the origin to \( (\gamma_{l_0}(l_0 + 1)^2, l_0 + 1) \). That is, let

\[
\hat{\lambda}_1 = (\gamma_{l_0}(l_0 + 1))^{-1}, \quad \hat{\lambda}_2 = y_{s_0} - x_{s_0} \hat{\lambda}_1. \tag{4.8}
\]

So for any point of break we have \( y_s \leq \hat{\lambda}_1 x_s + \hat{\lambda}_2 \) and for any index \( j \leq N - 1 \), we obtain

\[
j + 1 \leq \theta(\gamma_j(j + 1)^2) \leq \hat{\lambda}_1 \gamma_j(j + 1)^2 + \hat{\lambda}_2.
\]

If \( j \geq N \) then

\[
\hat{\lambda}_1 \gamma_j(j + 1) - 1 = (\gamma_{l_0}(l_0 + 1))^{-1} \gamma_j(j + 1) - 1 \geq (\gamma_j(j + 1))^{-1} \gamma_j(j + 1) - 1 = 0.
\]
Thus for the chosen \( \hat{\lambda}_1 \) and \( \hat{\lambda}_2 \) and any \( f \in X \) we have \( \mathcal{L}(f, \hat{\lambda}_1, \hat{\lambda}_2) \geq 0 \).

To construct \( \hat{f} \in X \) admissible in (4.6), let \( \hat{a}_j = 0 \) for \( j \neq c, l_0 \) and define \( \hat{a}_c, \hat{a}_{l_0} \) by the system

\[
\| \hat{f} \|_X^2 = 1, \quad \sum_{j=0}^{N-1} |K_j \hat{f}|^2 = \delta^2
\]

and since \( 0 \leq c < N \leq l_0 \) this becomes

\[
\gamma_c |\hat{a}_c|^2 + \gamma_{l_0} |\hat{a}_{l_0}|^2 = 1, \quad \frac{|\hat{a}_c|^2}{(c + 1)^2} = \delta^2.
\]

So let \( \hat{a}_c = \delta (c + 1) \) and \( \hat{a}_{l_0} = \left( \frac{1 - \delta^2 \gamma_c (c + 1)^2}{\gamma_{l_0}} \right)^{1/2} \). Then for \( \delta^2 \geq \gamma_c (c + 1)^2 = x_{s_0} \) the function \( \hat{f}(z) = \hat{a}_c z^c + \hat{a}_{l_0} z^{l_0} \) is admissible in (4.6) with

\[
\mathcal{L}(\hat{f}, \hat{\lambda}_1, \hat{\lambda}_2) = \frac{|\hat{a}_c|^2}{(c + 1)^2} \left( \hat{\lambda}_1 \gamma_c (c + 1)^2 + y_{s_0} - x_{s_0} \hat{\lambda}_1 - (c + 1) \right) + \frac{|\hat{a}_{l_0}|^2}{(l_0 + 1)^2} \left( \hat{\lambda}_1 \gamma_{l_0} (l_0 + 1)^2 - (l_0 + 1) \right) = 0.
\]

Therefore \( \mathcal{L}(\hat{f}, \hat{\lambda}_1, \hat{\lambda}_2) = 0 = \min_{f \in X} \mathcal{L}(f, \hat{\lambda}_1, \hat{\lambda}_2) \) and by construction we have \( \| \hat{f} \|_X^2 = 1 \) and \( \sum_{j=0}^{N-1} |K_j \hat{f}|^2 = \delta^2 \) so that

\[
\hat{\lambda}_1 (\| \hat{f} \|_X^2 - 1) + \hat{\lambda}_2 \left( \sum_{j=0}^{N-1} |K_j \hat{f}|^2 - \delta^2 \right) = 0
\]

and conditions (a) and (b) of Theorem 7 are satisfied.

**Case ii:** \( y_{s_0} = 0 \)

If \( y_{s_0} = 0 \) then \( (x_{s_0}, y_{s_0}) = (0, 0) \), and \( s_0 = 0 \), as this is the only point in the set \( \{(0, 0) \cup \{ \gamma_j (j + 1)^2, j + 1 \} \}_{j \in \mathbb{N}} \) with a \( y \)-coordinate of 0. Furthermore, as \( (0, 0) \) is a point of break of \( \theta \) we know \( \gamma_j > 0 \) for all \( j \in \mathbb{N} \). Since \( s_0 = 0 \) then by the definition of
so we know \( \frac{y_1}{x_1} \leq (\gamma_{l_0}(l_0 + 1))^{-1} \). As

\[
\frac{y_1}{x_1} = \max_{j \geq 0} \{(\gamma_j(j + 1))^{-1}\} \geq \max_{j \geq N} \{(\gamma_j(j + 1))^{-1}\} = (\gamma_{l_0}(l_0 + 1))^{-1}
\]

(4.9)

then we obtain equality, \( \frac{y_1}{x_1} = (\gamma_{l_0}(l_0 + 1))^{-1} \).

Define \( \tilde{\lambda}_1, \tilde{\lambda}_2 \) by (4.8) so \( \tilde{\lambda}_1 = \frac{y_1}{x_1} \) and \( \tilde{\lambda}_2 = 0 \). If we let \( \tilde{f} \in X \) be \( \tilde{f}(z) = (\gamma_{l_0})^{-1/2} z^{l_0} \) then

\[
\tilde{\lambda}_1(\|\tilde{f}\|_X^2 - 1) + \tilde{\lambda}_2(\sum_{j=0}^{N-1} |K_j \tilde{f}|^2 - \delta^2) = \tilde{\lambda}_1(1 - 1) + 0(0 - \delta^2) = 0.
\]

In addition \( \tilde{f} \) is admissible in extremal problem (4.6) as \( \|\tilde{f}\|_X \leq 1 \) and \( \sum_{j=0}^{N-1} |K_j \tilde{f}|^2 = 0 \leq \delta^2 \).

To justify \( \mathcal{L}(f, \tilde{\lambda}_1, \tilde{\lambda}_2) \geq 0 \) simply note that as \( (x_1, y_1) \) satisfies (4.9) and \( \gamma_j > 0 \) for all \( j \geq 0 \) then \( \frac{y_1}{x_1} \gamma_j(j + 1) - 1 \geq 0 \). So we have

\[
\mathcal{L}(f, \tilde{\lambda}_1, \tilde{\lambda}_2) = \sum_{j=0}^{\infty} \left| a_j \right|^2 \left( \frac{1}{j+1} - \tilde{y}_j \right)^2 \rightarrow 0
\]

since \( \mathcal{L}(\tilde{f}, \tilde{\lambda}_1, \tilde{\lambda}_2) = 0 \) then \( \tilde{f} \) minimizes \( \mathcal{L} \).

For both cases, we now consider extremal problem

\[
\tilde{\lambda}_1 \|f\|_X^2 + \tilde{\lambda}_2 \sum_{j=0}^{N-1} |K_j f - \tilde{y}_j|^2 \rightarrow \min, \quad f \in X.
\]

(4.10)

This problem can be written as

\[
\lambda_1 \sum_{j=0}^{\infty} \gamma_j |a_j|^2 + \tilde{\lambda}_2 \sum_{j=0}^{N-1} \left| \frac{a_j}{j+1} - \tilde{y}_j \right|^2 \rightarrow \min,
\]

which will have solution

\[
a_j = \begin{cases}
(1 + \lambda_1 \tilde{\lambda}_2^{-1} \gamma_j(j + 1)^2)^{-1}(j + 1)\tilde{y}_j & : 0 \leq j \leq N - 1, \quad \tilde{\lambda}_2 \neq 0 \\
0 & : j \geq N, \text{ or } \tilde{\lambda}_2 = 0
\end{cases}
\]

\( \square \)
4.1.3 Theorem 3: Varying Levels of Accuracy Termwise

**Proof.** The dual problem in this situation is thus

\[ \|f\|_{L_2(D)}^2 \rightarrow \max, \quad \|f\|_{X\tilde{}}^2 \leq 1, \quad \frac{|a_j|^2}{(j + 1)^2} \leq \delta_j^2, \quad j = 0, 1, \ldots, N - 1 \]  

(4.11)

with the corresponding Lagrange function

\[
L(f, \lambda) = -\sum_{j=0}^{\infty} \frac{|a_j|^2}{j + 1} + \lambda \sum_{j=0}^{\infty} \gamma_j |a_j|^2 + \sum_{j=0}^{N-1} \lambda_j \frac{|a_j|^2}{(j + 1)^2} \\
= \sum_{j=0}^{N-1} \frac{|a_j|^2}{(j + 1)^2} (\lambda \gamma_j (j + 1)^2 + \lambda_j - (j + 1)) \\
+ \sum_{j=N}^{\infty} \frac{|a_j|^2}{(j + 1)^2} (\lambda \gamma_j (j + 1)^2 - (j + 1))
\]

The method of proof will be to first determine \( \lambda = (\lambda, \lambda_1, \ldots, \lambda_{N-1}) \) with \( \lambda, \lambda_j \geq 0 \) and \( \hat{f} \in X \) admissible in (4.11) and satisfying (a) and (b) of Theorem 7.

If \( \gamma_0 \delta_0^2 \leq 1 \), define \( \lambda \) and \( \hat{f} \) as follows:

\[
\hat{\lambda} = \frac{1}{\gamma_{p_0+1}(p_0 + 2)}, \quad \hat{\lambda}_j = \begin{cases} 
\lambda + 1 - \gamma_j (j + 1)^2 & : j \leq p_0 \\
0 & : p_0 + 1 \leq j \leq N - 1
\end{cases}
\]

\[
\hat{\alpha}_j = \begin{cases} 
\delta_j (j + 1) & : j \leq p_0 \\
\left( \left( 1 - \sum_{j=0}^{p_0} \gamma_j \delta_j^2 (j + 1)^2 \right) \frac{1}{\gamma_{p_0+1}} \right)^{1/2} & : j = p_0 + 1 \\
0 & : j > p_0.
\end{cases}
\]

To verify \( \hat{\lambda}_j \geq 0 \) assume \( j \leq p_0 \) in which case \( \gamma_j (j + 1)^2 \leq \gamma_{p_0+1}(p_0 + 2)(j + 1) \) and hence

\[
\hat{\lambda}_j = \lambda + 1 - \frac{1}{\gamma_{p_0+1}(p_0 + 2)} \gamma_j (j + 1)^2 \geq j + 1 - \frac{1}{\gamma_{p_0+1}(p_0 + 2)} \gamma_{p_0+1}(p_0 + 2)(j + 1) = 0.
\]
To show, for the chosen $\hat{\lambda}$ and any $f \in X$, $L(f, \hat{\lambda}) \geq 0$, we consider the cases $j \leq p_0$ or $j > p_0$. For $j > p_0$ we know by assumption $\gamma_{p_0+1}(p_0 + 2) \leq \gamma_j(j + 1)$ and hence

$$\hat{\lambda}\gamma_j(j + 1) - 1 = \frac{1}{\gamma_{p_0+1}(p_0 + 2)} \gamma_j(j + 1) - 1 \geq 1 - 1 = 0$$

For $j \leq p_0$

$$\hat{\lambda}\gamma_j(j + 1)^2 + \hat{\lambda}_j - (j + 1) = \hat{\lambda}\gamma_j(j + 1)^2 + (j + 1 - \hat{\lambda}\gamma_j(j + 1)^2) - (j + 1) = 0.$$

Thus for any $f \in X$, $L(f, \hat{\lambda}) \geq 0$. For the constructed $\hat{f}$

$$L(\hat{f}, \hat{\lambda}) = \left( \sum_{j=0}^{p_0} \frac{|\hat{a}_j|^2}{(j + 1)^2} (\hat{\lambda}\gamma_j(j + 1)^2 + (j + 1 - \hat{\lambda}\gamma_j(j + 1)^2) - (j + 1)) \right)$$

$$+ \frac{|\hat{a}_{p_0+1}|^2}{(p_0 + 2)^2} \left( \frac{1}{\gamma_{p_0+1}(p_0 + 2)} \gamma_{p_0+1}(p_0 + 2)^2 - (p_0 + 2) \right)$$

$$= 0$$

and thus $\hat{f}$ minimizes the Lagrange function.

To show $\hat{f}$ is admissible in (4.11) we can clearly see that for $j \leq p_0$, $\frac{|\hat{a}_j|^2}{(j + 1)^2} \leq \delta_j^2$. It remains to show $\frac{|\hat{a}_{p_0+1}|^2}{(p_0 + 2)^2} \leq \delta_{p_0+1}^2$ for $p_0 < N - 1$. Assume not, then

$$\left( \left( 1 - \sum_{j=0}^{p_0} \delta_j^2 \gamma_j(j + 1)^2 \right) \frac{1}{\gamma_{p_0+1}(p_0 + 2)^2} \right) > \delta_{p_0+1}^2$$

which occurs if and only if

$$1 > \sum_{j=0}^{p_0+1} \delta_j^2 \gamma_j(j + 1)^2$$

which contradicts the definition of $p_0$ unless $p_0 = N - 1$. If $p_0 = N - 1$ then $p_0 + 1 = N$ and hence we no longer need the condition $\frac{|\hat{a}_{p_0+1}|^2}{(p_0 + 2)^2} \leq \delta_{p_0+1}^2$ in order for $\hat{f}$ to satisfy (4.11).
Furthermore

\[ \|\hat{f}\|_X^2 = \left( \sum_{j=0}^{p_0} \gamma_j |\hat{a}_j|^2 \right) + \gamma_{p_0+1} |\hat{a}_{p_0+1}|^2 = \left( \sum_{j=0}^{p_0} \gamma_j |\hat{a}_j|^2 \right) + \left( 1 - \left( \sum_{j=0}^{p_0} \gamma_j |\hat{a}_j|^2 \right) \right) = 1 \]

and so \( \hat{f} \) is admissable in (4.11).

By the construction of \( \hat{f} \) we also have the results \( \|\hat{f}\|_X = 1 \) and \( \|\hat{a}_j\| = \delta_j \) for \( j \leq p_0 \) while \( \hat{\lambda}_j = 0 \) for \( j > p_0 \). Thus \( \hat{f} \) satisfies (b) of Theorem 7 as

\[ \hat{\lambda}(\|\hat{f}\|_X^2 - 1) + \sum_{j=0}^{N-1} \hat{\lambda}_j \left( \frac{|\hat{a}_j|^2}{(j+1)^2} - \delta_j^2 \right) = 0. \]

We now proceed to the extremal problem

\[ \hat{\lambda}\|f\|_X^2 + \sum_{j=0}^{p_0} \hat{\lambda}_j |K_j f - \tilde{y}_j|^2 \to \min, \quad f \in X. \]

Notice the upper bound on the sum is \( p_0 \leq N - 1 \) as \( \hat{\lambda}_j = 0 \) for any \( j > p_0 \). This extremal problem will have solution

\[ a_j = \begin{cases} \frac{\hat{\lambda}_j(j+1)}{\hat{\lambda}_j(j+1)^2 + \hat{\lambda}_j} \tilde{y}_j = \hat{\lambda}_j \tilde{y}_j : j \leq p_0 \\ 0 : j > p_0 \end{cases}. \]

Therefore the error of optimal recovery is given by

\[ E(W, K^N, \delta) = \sqrt{\hat{\lambda} + \sum_{j=0}^{p_0} \delta_j^2 \hat{\lambda}_j} \]

and

\[ \tilde{m}(\tilde{y}) = \sum_{j=0}^{p_0} \hat{\lambda}_j \tilde{y}_j z^j \]

is an optimal method.

Now we proceed to the case \( \gamma_0 \delta_0^2 > 1 \). Choose \( \hat{\lambda} = \gamma_0^{-1} \) and \( \hat{\lambda}_j = 0 \) for \( j = \)
0, 1, \ldots, N - 1. Then as \( \gamma_0 \leq \gamma_j(j + 1) \) for all \( j \geq 0 \) we have

\[
\hat{\lambda}\gamma_j(j + 1)^2 - (j + 1) = (j + 1)(\gamma_0^{-1}\gamma_j(j + 1) - 1) \geq (j + 1)(1 - 1) = 0.
\]

Thus \( \mathcal{L}(f, \overline{\lambda}) \geq 0 \) for all \( f \in X_\gamma \). Let \( \hat{a}_0 = \gamma_0^{-1/2} \) and \( \hat{f}(z) = \hat{a}_0 \) and notice \( \|\hat{f}\|_X^2 = 1 \) and clearly \( |\hat{a}_0|^2 = \gamma_0^{-1} \leq \delta_0^2 \) so \( \hat{f} \) is admissible in (4.11). Furthermore

\[
\mathcal{L}(\hat{f}, \overline{\lambda}) = |\hat{a}_0|^2(\hat{\lambda}\gamma_0 - 1) = 0
\]

and so \( \mathcal{L}(\hat{f}, \overline{\lambda}) = \min_{f \in X}(f, \lambda) \). Also,

\[
\hat{\lambda}\left(\|\hat{f}\|_X^2 - 1\right) + \sum_{j=0}^{N-1} \hat{\lambda}_j \left(\frac{|a_j|^2}{(j+1)^2} - \delta_j^2\right) = \hat{\lambda}(1 - 1) + 0 = 0.
\]

Therefore \( E(W, K^N, \overline{\delta}) = \sqrt{\hat{\lambda}} = \gamma_0^{-1/2} \) and \( \hat{m}(\overline{y}) = 0 \) is an optimal method.

\[\square\]

### 4.2 Secant Integral Information Operator Problem

**Proof.** The dual problem in this case is

\[
\|f\|_{L_2(D)}^2 \to \max, \quad \|f\|_X^2 \leq 1, \quad \|Lf\|_{L_2(D)}^2 \leq \delta^2 \tag{4.12}
\]

that is

\[
\sum_{j=0}^{\infty} \frac{1}{j+1} |a_j|^2 \to \max, \quad \sum_{j=0}^{\infty} \gamma_j|a_j|^2 \leq 1, \quad \sum_{j=0}^{\infty} \frac{1}{\eta_j} |a_j|^2 \leq \delta^2.
\]

Then

\[
\mathcal{L}(f, \lambda_1, \lambda_2) = \sum_{j=0}^{\infty} \frac{|a_j|^2}{\eta_j} \left(\lambda_1\gamma_j\eta_j + \lambda_2 - \frac{\eta_j}{j+1}\right)
\]

is the corresponding Lagrange function. Now if \( \hat{\lambda}_1, \hat{\lambda}_2 \) are given by (2.32) then by Lemma 1, \( \hat{\lambda}_1, \hat{\lambda}_2 \geq 0 \). As \( x_s \to \infty \) then for any \( j \) there is some \( k \) with \( x_k \leq \gamma_j\eta_j < x_{k+1} \) in
which case
\[ \eta_j(j + 1)^{-1} \leq \theta(\gamma_j \eta_j) \leq \theta_k(\gamma_j \eta_j) = \hat{\lambda}_1 \gamma_j \eta_j + \hat{\lambda}_2 \]

where \( \theta_k = \theta|_{[x_s, x_{s+1}]} \) and \( \theta \) is given by (2.7). Hence \( 0 \leq \hat{\lambda}_1 \gamma_j \eta_j + \hat{\lambda}_2 - \eta_j(j + 1)^{-1} \).

Therefore \( L(f, \hat{\lambda}_1, \hat{\lambda}_2) \geq 0 \) for any \( f \in X \).

Our next step is to construct \( \hat{f} \in X \) admissible in (4.12) that satisfies the conditions

\[
L(\hat{f}, \hat{\lambda}_1, \hat{\lambda}_2) = \min_{f \in X} L(f, \hat{\lambda}_1, \hat{\lambda}_2) \tag{4.13}
\]

\[
\hat{\lambda}_1(\|\hat{f}\|_X^2 - 1) + \hat{\lambda}_2(\|L\hat{f}\|_{L^2(\mathbb{D})} - \delta^2) = 0 \tag{4.14}
\]

of Theorem 7. Note, \( y_s = 0 \) if and only if \( s = 0 \) and \( y_0 = 0 \). Assuming \( y_s > 0 \), let \( m, m' \) satisfy \((\gamma_m \eta_m, \eta_m(m + 1)^{-1}) = (x_s, y_s)\) and \((\gamma_{m'} \eta_{m'}, \eta_{m'}(m' + 1)^{-1}) = (x_{s+1}, y_{s+1})\). Let \( \hat{a}_j = 0 \) for \( j \neq m, m' \) and \( \hat{a}_m, \hat{a}_{m'} \) defined by the equations

\[
\gamma_m |\hat{a}_m|^2 + \gamma_{m'} |\hat{a}_{m'}|^2 = 1, \quad \eta_m^{-1} |\hat{a}_m|^2 + \eta_{m'}^{-1} |\hat{a}_{m'}|^2 = \delta^2. \tag{4.15}
\]

Let

\[
\hat{a}_m = \left( \frac{\eta_{m'}(\delta^2 \gamma_m \eta_m - 1)}{\gamma_{m'} \eta_{m'} - \gamma_m \eta_m} \right)^{1/2}
\]

\[
\hat{a}_{m'} = \left( \frac{\eta_m(1 - \delta^2 \gamma_{m'} \eta_{m'})}{\gamma_m \eta_m - \gamma_{m'} \eta_{m'}} \right)^{1/2}
\]

and for \( \gamma_m \eta_m < \delta^{-2} \leq \gamma_{m'} \eta_{m'} \), ie.

\[
x_s < \delta^{-2} \leq x_{s+1}, \quad \tag{4.16}
\]

the function \( \hat{f}(z) = \hat{a}_m z^m + \hat{a}_{m'} z^{m'} \) is admissible in (4.12) and by construction satisfies...
Furthermore, we can show \( \hat{f} \) also satisfies (4.13) as

\[
\mathcal{L}(\hat{f}, \hat{\lambda}_1, \hat{\lambda}_2) = \frac{|a_m|^2}{\eta_m} \left( \frac{y_{s+1} - y_s}{x_{s+1} - x_s} \gamma_m \eta_m + \frac{x_{s+1}y_s - x_s y_{s+1}}{x_{s+1} - x_s} - \eta_m(m + 1)^{-1} \right) \\
+ \frac{|a_{m'}|^2}{\eta_m'} \left( \frac{y_{s+1} - y_s}{x_{s+1} - x_s} \gamma_{m'} \eta_{m'} + \frac{x_{s+1}y_s - x_s y_{s+1}}{x_{s+1} - x_s} - \eta_{m'}(m' + 1)^{-1} \right) \\
= \frac{|a_m|^2}{\eta_m} \left( \frac{y_{s+1} - y_s}{x_{s+1} - x_s} x_s + \frac{x_{s+1}y_s - x_s y_{s+1}}{x_{s+1} - x_s} - y_s \right) \\
+ \frac{|a_{m'}|^2}{\eta_{m'}} \left( \frac{y_{s+1} - y_s}{x_{s+1} - x_s} x_{s+1} + \frac{x_{s+1}y_s - x_s y_{s+1}}{x_{s+1} - x_s} - y_{s+1} \right) \\
= 0.
\]

If \( y_s = 0 \), then \( s = 0 \) and \( y_0 = 0 \) so we know \( x_0 = 0 \). Let \( a_j = 0 \) for all \( j \neq m' \), \( a_{m'} = \gamma_{m'}^{-1/2} \) and \( \hat{f}(z) = a_{m'} z^{m'} \) to obtain \( \|\hat{f}\|_X^2 = 1 \) and

\[
\|Lf\|_{L_2(D)}^2 = \frac{|a_{m'}|^2}{\eta_{m'}} = (\eta_{m'} \gamma_{m'})^{-1} = (x_1)^{-1} \leq \delta^2
\]

by (4.16).

We proceed to the extremal problem

\[
\hat{\lambda}_1 \|f\|_X^2 + \hat{\lambda}_2 \|Lf - y\|_{L_2(D)}^2 \to \min \tag{4.17}
\]

with \( f \in X_\gamma \) and \( y \) satisfying (2.31). To solve (4.17), let \( y \in L_2(D) \) be given by (2.29).

Extremal problem (4.17) can be written

\[
\hat{\lambda}_1 \sum_{j=0}^{\infty} |a_j|^2 + \hat{\lambda}_2 \left\| \sum_{j \in \mathbb{Z}} (a_j c_j(h) - d_j(h)) e^{ij\phi} \right\|_{L_2(D)}^2 \to \min. \tag{4.18}
\]
where $\chi_j = \begin{cases} 0 & : j < 0 \\ 1 & : j \geq 0 \end{cases}$. Problem (4.18) can be rewritten

$$
\hat{\lambda}_1 \sum_{j=0}^{\infty} \gamma_j |a_j|^2 + \hat{\lambda}_2 \int_0^1 \sum_{j \in \mathbb{Z}} |a_j \chi_j c_j(h) - d_j(h)|^2 h \, dh \rightarrow \min,
$$

which will have solution

$$
a_j = \begin{cases} \frac{\hat{\lambda}_2 \int_0^1 d_j(h)c_j(h)h \, dh}{\lambda_1 \gamma_j + \hat{\lambda}_2 \int_0^1 c_j^2(h)h \, dh} & : j \geq 0 \\ 0 & : j < 0 \end{cases}.
$$

To simplify $a_j$, with $\eta_j$ and $\nu_j$ as previously defined,

$$
a_j = \begin{cases} \frac{\hat{\lambda}_2 \nu_j}{\lambda_1 \gamma_j + \hat{\lambda}_2 \eta_j} & : j \geq 0 \\ 0 & : j < 0 \end{cases}.
$$

Thus by Theorems 6 and 7, (2.34) is an optimal method and (2.33) the error of optimal recovery.

In the case $y_s = 0$; $\hat{\lambda}_1 = \frac{y_s}{\lambda} > 0$ and $\hat{\lambda}_2 = 0$. Then $\hat{m}(y) = 0$ is optimal and $E(W, L, \delta) = \sqrt{\frac{y_s}{\lambda}}$ is the error of optimal recovery.
Chapter 5

Examples

5.1 Radial Information

Example 5. Cat’s Eyes: For several scenarios we have developed an optimal method of recovery. The apriori information has dealt with the radial integral on the unit disk over different spaces of functions. In this section we examine how closely the optimal method approximates the original function. The function we consider is

\[ f(z) = z^2 - z + \frac{3}{16}. \]  

(5.1)

This function, referred to as the Cat’s Eyes function, has zeros at \( z = \frac{1}{4} \) and \( z = \frac{3}{4} \). As each space is assumed to be over the disc we will use the notation \( H^2_r = H^2_r(D) \), \( H^2_r = H^2_r(D) \), and for \( r = 0 \) simply \( H^2 \) and \( H^2 \).

For the first application, we consider the situation of Corollary 1. The first assumption we have for the function \( f \) is that \( f \in H^2 \), with \( \|f\|_{H^2} \leq 1 \), i.e. \( f \in H^2 \). For the space \( X_\gamma = H^2 \) we have \( \gamma_j = 1 \) for \( j = 0, 1, \ldots \). Our collection of points \( C_{H^2} \) is given by

\[ C_{H^2} = \{(0, 0) \cup \{(j + 1)^2, j + 1)\}_{j=0,1,...}. \]
If
\[ \theta(x) = \max\{y : (x, y) \in \text{co}(C_{\mathcal{H}^2})\} \]
then the points of break of \( \theta \) are precisely the points in the set \( C_{\mathcal{H}^2} \). That is, \( x_s = s^2 \) and \( y_s = s \). We now restate Corollary 1 in this context.

**Corollary 5.** Suppose \( \frac{1}{s+1} \leq \delta < \frac{1}{s} \) for some integer \( s \geq 1 \). Let \( \hat{\lambda}_1 = \frac{1}{2s+1} \) and \( \hat{\lambda}_2 = \frac{s(s+1)}{2s+1} \). Then
\[ E(H^2, K, \delta) = \sqrt{ \frac{1 + \delta^2 s(s + 1)}{2s + 1} } \]
and
\[ \hat{m}(y) = \sum_{j=0}^{\infty} \left( 1 + \frac{(j+1)^2}{s(s+1))} \right)^{-1} (j+1)\bar{y}_j z^j \]
is an optimal method. If \( \delta \geq 1 \) then \( E(H^2, K, \delta) = 1 \) and \( \hat{m}(y) = 0 \) is an optimal method.

In general, for an unknown function \( g \in H^2 \) given by \( g(z) = \sum_{j=0}^{\infty} g_j z^j \) we have apriori information \( \bar{y} \in L^2(\mathbb{T}) \) where \( \bar{y} = \sum_{j=0}^{\infty} \bar{y}_j e^{ij\phi} \) such that
\[ \|Kg - \bar{y}\|_{L^2(\mathbb{T})} \leq \delta^2. \]
More explicitly we have the values \( \bar{y}_0, \bar{y}_1, \ldots \) such that
\[ \sum_{j=0}^{\infty} \left| \frac{g_j}{j+1} - \bar{y}_j \right|^2 \leq \delta^2. \]
For simplicity the coefficients are real valued, although this is not a necessary condition. Assume we are given
\[ \bar{y} = \frac{1}{4} - \frac{11}{10} e^{i\phi} + \frac{14}{15} e^{2i\phi} + \frac{1}{20} e^{3i\phi} \]
with
\[ \|Kg - \bar{y}\|_{L^2(\mathbb{T})} \leq \frac{1}{25}. \] (5.2)
As \( \delta = \frac{1}{5} \) then \( s = 4 \) satisfies \( \frac{1}{s+1} \leq \delta < \frac{1}{s} \). Our error of optimal recovery for this class \( W = H^2 \) and \( \delta = \frac{1}{5} \) is

\[
E(H^2, K, \delta) = \sqrt{\frac{1}{2s+1} + \delta^2 \frac{s(s+1)}{2s+1}} = \frac{1}{\sqrt{5}}.
\]

A method of optimal recovery is then

\[
\hat{m}(\tilde{y}) = \frac{20}{21} \tilde{y}_0 + \frac{5}{3} \tilde{y}_1 z + \frac{60}{29} \tilde{y}_2 z^2 + \frac{25}{9} \tilde{y}_3 z^3
\]

\[
= \frac{5}{21} - \frac{11}{9} z + \frac{56}{29} z^2 + \frac{5}{36} z^3
\]

Notice that we could have \( g = f \) as any \( g \in H^2 \) with \( Kg \) belonging to the set

\[
B_\delta(\tilde{y}) = \{ y \in Y : \| \tilde{y} - y \|_{L_2(T)} \leq \delta \},
\]

that is the \( \delta \) ball about the information \( \tilde{y} \), satisfies (5.2).

We can compare the modulus of the function \( \hat{m}(\tilde{y}) \) to the original \( f \) at each \( z \in \mathbb{D} \). The figures below are representative of the modulus of each function. The color scheme is graded to the magnitude of the modulus of the described function.

![Figure 5.1](attachment:figure5.1.png)
Lastly we consider the absolute value of the difference in $f$ and the optimal method of recovery $\hat{m}(\tilde{y})$ at each point. We already know $E(H^2, K, \delta) = 1/\sqrt{5}$, which is the worst case scenario over all $g \in W$ with $Kg \in B_\delta(\tilde{y})$, for the value $\|\hat{m}(\tilde{y}) - g\|_{L^2(\mathbb{D})}$. Thus, using information $\tilde{y}$ and the Cat’s Eyes function $f$,

$$\|\hat{m}(\tilde{y}) - f\|_{L^2(\mathbb{D})} \leq 1/\sqrt{5}.$$ 

We have

$$\|\hat{m}(\tilde{y}) - f\|_{L^2(\mathbb{D})}^2 = \frac{1}{2\pi} \int_0^1 \int_{-\pi}^{\pi} |\hat{m}(\tilde{y})(re^{i\phi}) - f(re^{i\phi})|^2 r\,d\phi\,dr$$

$$= \frac{1}{2\pi} \int_0^1 \int_{-\pi}^{\pi} | - \frac{1}{84} - \frac{2}{9}r^2e^{2i\phi} + \frac{27}{29}r^2e^{3i\phi} + \frac{5}{36}r^3e^{3i\phi}| r\,d\phi\,dr$$

$$= \frac{840261}{5274752} < \frac{1}{5} = E(H^2, K, \delta)^2$$

as expected.

In Theorem 1 we stipulated a bound on the normed difference of $Kf$ and $\tilde{y}$ in the $L^2(\mathbb{T})$ norm. Theorems 2 and 3 dealt with more specific distributions in the inaccuracy of the information. We proceed to the setting of Theorem 3 in which we compare the
function $f \in H^2$ with optimal recovery method $\hat{m}(y)$ where $y = (y_0, \ldots, y_{N-1}) \in \mathbb{C}^N$ and each $y_j$ approximates $K_j f$ to within $\delta_j > 0$.

**Example 6.** Let $W = H^2$ and $x \in W$. In the general setting, assume to be given $\tilde{y}_j \in \mathbb{C}$ such that

$$|K_j x - \tilde{y}_j| \leq \delta_j$$

(5.3)

for $\delta_j = (\alpha^{1/2}(j+1))^{-1}$ with $\alpha > 0$ and $j = 0, 1, \ldots, N-1$ for some integer $N$. Refering back to the definition of $p_0$ we have

$$p_0 = \max\{p \geq 0 : \sum_{j=0}^{p} \delta_j^2 \gamma_j(j+1)^2 \leq 1, \ j \leq N-1\}$$

$$= \max\{p \geq 0 : p \leq \alpha, \ p \leq N-1\}$$

$$= \min\{\lfloor \alpha \rfloor, N-1\}.$$ 

If we instead want to know how many terms we would need approximations for in order to obtain the smallest error of optimal recovery we refer back to (2.22). In which case we would need $N_\delta$ terms such that $N_\delta - 1 = \lfloor \alpha \rfloor$. If we have more than $\lfloor \alpha \rfloor + 1$ terms there will be no further reduction in the optimal error.
Example 7. For our next example we will again vary $\delta_j$ in the setting of Theorem 3. Let $\delta_j = (\alpha^{1/2}(j + 1)^{3/2})^{-1}$ for $j = 0, \ldots, N - 1$ for some integer $N$. Again assume for some $x \in H^2$ we have $\tilde{y}_j$ satisfying (5.3). Then

$$p_0 = \max\{p \geq 0 : \sum_{j=0}^{p} (j + 1)^{-1} \leq \alpha, \quad p \leq N - 1\}.$$

If we have approximations for $K_j x$ for $j = 0, \ldots, N_\delta$ where $N_\delta$ is given by (2.22) this will yield the least optimal error. The harmonic numbers, that is the partial sums, $S_p = \sum_{j=1}^{p} \frac{1}{j}$, grow proportional to the natural logarithm function. As Euler has shown,

$$\lim_{n \to \infty} \left( \sum_{j=1}^{n} \frac{1}{j} - \ln n \right) = \gamma \approx 0.57722.$$

Then for larger $\alpha$ we will have $N_\delta \approx \lceil e^{\alpha - \gamma} \rceil - 1$. The exact values of $N_\delta$ and the approximations for $\alpha = 1, \ldots, 10$ are shown below

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_\delta$</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>29</td>
<td>81</td>
<td>225</td>
<td>614</td>
<td>1672</td>
<td>4548</td>
<td>12365</td>
</tr>
<tr>
<td>$\lceil e^{\alpha - \gamma} \rceil - 1$</td>
<td>0</td>
<td>3</td>
<td>10</td>
<td>29</td>
<td>82</td>
<td>225</td>
<td>614</td>
<td>1672</td>
<td>4548</td>
<td>12365</td>
</tr>
</tbody>
</table>

Figure 5.4

For simplicity, in this example choose $\alpha = 3$, that is $\delta_j = (3(j + 1)^3)^{-1/2}$ so that $N_\delta = 9$. So to minimize the error of optimal recovery we would need the approximate values $y_j \in \mathbb{C}$ to $K_j x$ for $j = 0, \ldots, 8$. Let us first consider the case in which we have only 5 terms approximated. That is we have $y_j$ satisfying (5.3) for $j = 0, 1, 2, 3, 4$ and let $y_j = j$ be given. In this case $p_0 = 4$ and hence $p_0 = N - 1$ as $N_\delta > N - 1$. By Theorem 3 with $\hat{\lambda} = \frac{1}{6}$ and $\hat{\lambda}_j = j + 1 - \frac{(j+1)^2}{6}$ the method

$$\hat{m}(y) = \sum_{j=0}^{p_0} \hat{\lambda}_j y_j z^j = \sum_{j=0}^{4} = \frac{4}{3} z + \frac{3}{2} z^2 + \frac{4}{3} z^3 + \frac{5}{6} z^4$$

68
is an optimal one and the error of optimal recovery is

\[ E(H^2, K^N, \delta) = \sqrt{\hat{\lambda} + \sum_{j=0}^{4} \delta_j^2 \hat{\lambda}_j} = \sqrt{\frac{5699}{10800}} \approx 0.72642. \]

In contrast if we assumed to have the first \( N_\delta = 9 \) terms \( y_j \in \mathbb{C} \) for \( j = 0, \ldots, 8 \) satisfying (5.3) keeping \( \delta_j = (3(j + 1)^3)^{-1/2} \) fixed then the error of optimal recovery would be

\[ E(H^2, K^{N_\delta}, \delta) = \sqrt{\hat{\lambda} + \sum_{j=0}^{8} \delta_j^2 \hat{\lambda}_j} = \sqrt{\frac{9886753}{19051200}} \approx 0.72039. \]

So the additional information is only able to reduce the error slightly. Approximate values of the optimal error are shown in the table for all \( N \).

<table>
<thead>
<tr>
<th>( N )</th>
<th>( E(H^2, K^N, \delta) )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E(H^2, K^N, \delta) )</td>
<td>0.81650</td>
<td>0.76376</td>
<td>0.74224</td>
<td>0.73188</td>
<td>0.72642</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( N )</th>
<th>( E(H^2, K^N, \delta) )</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9+</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E(H^2, K^N, \delta) )</td>
<td>0.72341</td>
<td>0.72176</td>
<td>0.72083</td>
<td>0.72039</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 5.5**

**Example 8.** As a final example of this type we again let \( W = H^2 \), but now let the error in each information term be constant. In the setting of Corollary 2, we let \( \delta_j^2 = 1/\alpha \) with \( 0 < 1/\alpha \leq 1 \). Then \( p_0 \) is given by (2.26). Notice that for \( t \in [0, \infty) \), \( g(t) = 2t^3 + 9t^2 + 13t + 6(1 - \alpha) \) is a non-decreasing real function. Furthermore, for \( \alpha \geq 1 \), we have \( g(0) = 6(1 - \alpha) \leq 0 \). Thus \( g \) has exactly one real root in the interval \([0, \infty)\), call it \( \eta \). Then

\[
p_0 = \min\{\lceil \eta \rceil, N - 1\}
\]

where clearly \( \eta \) depends on \( \alpha \).

For this problem we concern ourselves with obtaining the minimal optimal error for a given situation in which we can either fix the number of terms available and vary the
inaccuracy of each term, that is $\alpha$, or we fix $\alpha$ and vary the number of terms available. If

$$N_\alpha = \lfloor \eta \rfloor$$

then $N_\alpha$ is the least number of terms such that the optimal error is minimized for a given $\delta = 1/\alpha$. Knowing more terms than $N_\alpha$ will not yield a decrease in the optimal error.

In contrast fix the number of terms that can be given to $n$ and define

$$\alpha_n = \min\{\alpha \in [1, \infty) : N_\alpha = n\}.$$ 

That is, $1/\alpha_n$ is the largest inaccuracy for each term such that if we know more than $n$ terms the optimal error will not decrease further.

The tables of values, Table ?? and ??, give a sample of values for each situation described.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>100</th>
<th>1000</th>
<th>10000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_\alpha = \lfloor \eta \rfloor$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>12</td>
<td>29</td>
</tr>
<tr>
<td>$E(H^2, K^{N_\alpha}, \delta)$</td>
<td>1</td>
<td>0.6218</td>
<td>0.5888</td>
<td>0.5062</td>
<td>0.4490</td>
<td>0.3790</td>
<td>0.2673</td>
<td>0.1796</td>
</tr>
</tbody>
</table>

Figure 5.6: Values of $N_\alpha$ and optimal error for various $\alpha$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_n$</td>
<td>5</td>
<td>14</td>
<td>30</td>
<td>55</td>
<td>91</td>
<td>140</td>
<td>204</td>
<td>285</td>
</tr>
<tr>
<td>$E(H^2, K^n, \delta)$</td>
<td>0.6218</td>
<td>0.5126</td>
<td>0.4522</td>
<td>0.4106</td>
<td>0.3792</td>
<td>0.3545</td>
<td>0.3338</td>
<td>0.3165</td>
</tr>
</tbody>
</table>

Figure 5.7: Values of $\alpha_n$ and optimal error for various $n \in \mathbb{N}$.
5.2 Secant Information

When considering the radial integral operator the *Cat’s Eyes* function \( f(z) = z^2 - z + \frac{3}{16} \) was considered. For the secant integral operator, consider \( g(z) = \frac{1}{4} z^4 - \frac{1}{64} \in H^{2,1} = W \) which has \( g(z) = 0 \) for \( z = \pm \frac{1}{2}, \pm \frac{i}{2} \). Applying secant integral operator \( L \) given by (2.27) we obtain

\[
Lg(\phi, h) = \frac{1}{10} \sin(5 \cos^{-1} h) e^{4i\phi} - \frac{1}{32} \sin(\cos^{-1} h) \\
= \left( \frac{1}{10} - \frac{6}{5} h^2 + \frac{8}{5} h^4 \right) \sqrt{1 - h^2} e^{4i\phi} - \frac{1}{32} \sqrt{1 - h^2}.
\]

Recall, \( g \), and hence \( Lg \), is unknown information. All that is available is inaccurate information \( \tilde{y} \) that approximates \( Lg \) to within some \( \delta > 0 \) and that \( g \) belongs to a certain class of analytic functions \( W \). For the space \( X = H^{2,1} \), \( \gamma_j = j^2 \) for \( j \geq 0 \). The inaccurate information \( \tilde{y}(h, \phi) \) in the form of (2.29) with \( y_{jk} \) for \( 0 \leq j \leq 4, 1 \leq k \leq 10 \) given by the matrix entries

\[
\begin{pmatrix}
-.4087795834e -1 & .1666666667e -1 & .1000000000 & -.8000000000e -3 & -.1159473947 \\
-.8097891722e -1 & .2143347051e -5 & .8000000000e -3 & -.1000000000 & -.2907270483 \\
-.9132723701e -2 & -.1000000000 & -.3333333333e -1 & -.3703703704e -2 & 1.484878308 \\
.6597384541e -2 & -.3125000000e -2 & -.1000000000 & -.1600000000e -3 & -.5919077619e -1 \\
.9500590969e -1 & -.2143347051e -5 & .1000000000 & -.1666666667e -1 & .1196169416 \\
-.9592259076e -1 & .3200000000e -4 & -.1250000000e -1 & .1666666667e -1 & -.2725847374e -1 \\
.2155563626e -2 & .1000000000 & .6400000000e -5 & .3200000000e -4 & .9036049717e -1 \\
.6106047233e -2 & -.1000000000 & -.1000000000 & .6250000000e -2 & -.1457313932e -1 \\
-.2529740483e -2 & -.1562500000e -2 & .1000000000 & .7716049383e -4 & .8947003838e -1 \\
.3911144766e -2 & .2777777778e -2 & .9765625000e -4 & -.1600000000e -3 & -.9106937183e -2
\end{pmatrix}
\]

and \( y_{jk} = 0 \) otherwise. These values were calculated by first determining the coefficients, \( g_{jk} \), for the expansion of \( g \) in a Bessel-Fourier series in \( h \) and \( \phi \). Then an error was introduced, for a finite number of terms, by simply adding \( e(t) = \frac{1}{10}(-t)^{-4} \) where \( t \) is a random variable uniformly distributed on the integers 1 through 6 to the values of \( g_{jk} \).

Thus

\[
y_{jk} = \begin{cases} 
  g_{jk} + e(t) & \text{for } 0 \leq j \leq 4, \ 1 \leq k \leq 10 \\
  0 & \text{otherwise}
\end{cases}
\]

It is given that \( \|Lf - \tilde{y}\|_{L^2(\mathbb{D})} \leq 0.08 \) (in fact \( \|Lf - \tilde{y}\|_{L^2(\mathbb{D})} \approx 0.065 \)). Now as
$$\delta = 0.08 \text{ and } (x_s, y_s) = (\gamma_s \eta_{s+1}, \eta_{s+1}(s + 1)^{-1}) \text{ (for } \eta_s \text{ see eq. (2.28)) then } s = 4 \text{ satisfies } x_s < \delta^{-2} \leq x_{s+1} \text{ and furthermore}$$

$$\hat{\lambda}_1 = \frac{y_{s+1} - y_s}{x_{s+1} - x_s} = \frac{\eta_4 - \eta_3}{\gamma_4 \eta_3 - \gamma_3 \eta_3} = \frac{1}{212}$$

$$\hat{\lambda}_2 = \frac{x_{s+1}y_s - x_s y_{s+1}}{x_{s+1} - x_s} = \frac{\gamma_4 \eta_4 \eta_3 - \gamma_3 \eta_3 \eta_4}{\gamma_4 \eta_3 - \gamma_3 \eta_3} = \frac{925}{212}.$$  

Thus by Theorem 4 the method (2.34):

$$\hat{m}(\tilde{y}) = \sum_{j=0}^{\infty} \frac{925 \nu_j}{j^2 + 925 \eta_j^2} \tilde{z}^j,$$

where $\eta_j, \nu_j$ given by (2.28) and (2.30) respectively, is an optimal method. Using the values for $y_{jk}, \hat{m}$ can be written as the approximation

$$\hat{m}(\tilde{y}) = -0.008824964826 + 0.004137275250z + 0.04973290473z^2$$

$$+ 0.05119266624z^3 + 0.2022473264z^4.$$  

Now we can compare the plots of the modulus of an optimal method $|\hat{m}(\tilde{y})|$, having been constructed from the inaccurate information $\tilde{y} \approx Lg$, to the actual function $g$ on the unit disc $\mathbb{D}$.

The theorem states that $E(W, L, \delta) = \sqrt{\hat{\lambda}_1 + \delta^2 \hat{\lambda}_2}$ and hence in this case the error of optimal recovery is

$$\sup_m e(H^{2,1}, L, 0.08, m) = E(H^{2,1}, L, 0.08) = \sqrt{\frac{1}{212} + 0.08^2 \frac{925}{212}} \approx 0.1806696140.$$  

In fact

$$e(H^{2,1}, L, 0.08, \hat{m}) = \|\hat{m}(\tilde{y}) - g\|_{L_2(\mathbb{D})} \approx 0.07907784087.$$
Figure 5.8

(a) $|\tilde{m}|$ overhead view

(b) $|\tilde{m}|$ perspective view

Figure 5.9

(a) $|g|$ overhead view

(b) $|g|$ perspective view
Figure 5.10

(a) $|g(z) - \hat{m}(z)|$ overhead view

(b) $|g(z) - \hat{m}(z)|$ perspective view
Bibliography


