



Solución al Problema Inverso de Dispersión de Ondas en el Contexto Bayesiano

Tesis desarrollada para obtener el grado de:

**Maestro en Ciencias con Especialidad en Computación y
Matemáticas Industriales**

presentada por:

Abel Palafox González

Lic. en Computación

agosto 2012

Tutor: Dr. Marcos Capistrán Ocampo

“I do not fear computers. I fear the lack of them”. Isaac Asimov

Dedicada a mi padre, mi heroe.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por todo el apoyo brindado durante el posgrado. Así mismo agradezco a los investigadores, personal administrativo y de servicios del Centro de Investigación en Matemáticas, A. C. (CIMAT) por contribuir sustancialmente a mi desarrollo profesional. Un agradecimiento especial a Mary mi esposa por ser mi fortaleza durante todos estos años. Finalmente y no menos importante, gracias a mi familia y amigos por su apoyo incondicional.

Resumen

La teoría de dispersión de ondas concierne el efecto que tiene un medio inhomogeneo sobre una partícula incidente u onda. En particular, si el campo total es visto como la suma de un campo incidente y un campo dispersado, entonces el problema directo de dispersión de ondas consiste en determinar el campo dispersado a partir del campo incidente y de la ecuación diferencial que gobierna el movimiento de la onda. Por otro lado, el problema inverso de dispersión de ondas es determinar la inhomogeneidad a partir del comportamiento asintótico del campo dispersado. El problema inverso no puede ser resuelto por medio de una expresión analítica. En éste trabajo se propone abordar el problema inverso en un contexto Bayesiano. Un esquema Bayesiano provee una distribución de probabilidad que nos permite estudiar la estimación de la forma del obstáculo dispersor, dada una discretización para la misma. Presentamos una manera eficiente de resolver el problema directo y una versión paralelizada del mismo, así como una implementación para aproximar la distribución de probabilidad de la solución del problema inverso, basada en un método Markov Chain Monte Carlo. Mostramos al final los resultados obtenidos para el enfoque Bayesiano del problema inverso y las distribuciones de probabilidad involucradas en dicho esquema.

Índice general

Dedicatoria	IV
Agradecimientos	V
Resumen	VI
Lista de Figuras	IX
Lista de Tablas	XI
Lista de Algoritmos	XII
1. Introducción	1
2. Planteamiento del problema	3
2.1. Modelo Físico	4
2.2. Solución Fundamental	5
2.3. Potenciales Acústicos	7
2.4. El Caso \mathbb{R}^2	8
2.5. Método de Nystrom	10
2.6. Problema Inverso	13
2.6.1. Estadística Bayesiana	15

2.6.2. Enfoque Bayesiano	16
3. Tratamiento del problema	18
3.1. Distribución a priori	18
3.2. Verosimilitud	21
3.3. Distribución Posterior	22
3.4. Implementación Numérica	23
3.4.1. Análisis de la Matriz A	25
3.4.2. Generalized Conjugate Residuals, GCR	26
3.4.3. Esquema de Paralelización	30
4. Resultados	37
4.1. Caso de Prueba Fish	40
4.2. Caso de Prueba Kite	43
5. Conclusiones	59
5.1. Problema Directo	59
5.2. Problema Inverso	60
5.3. Trabajo a Futuro	62
Bibliografía	64

Lista de Figuras

2.1. Fenómeno de Dispersión de Ondas.	5
2.2. Curva de Prueba para el Problema Directo: Kite	13
3.1. Análisis de la constante C	20
3.2. Eigenvalores de las Matrices A^s (.)y A^c (+) con diferentes números de onda.	27
3.3. Número de condición para las matrices de eigenvectores para el Potencial de Capa Simple (a) y el Potencial Combinado (b). . .	28
3.4. Tiempos de Ejecución obtenidos con diferentes implementaciones	36
4.1. Figura Fish	40
4.2. Resultados para $n = 1$ con 1 onda incidente ((a) y (b)) y 3 ondas incidentes ((c) y (d)).	41
4.3. Resultados para $n = 4$ con 1 onda incidente ((a) y (b)) y 3 ondas incidentes ((c) y (d)).	42
4.4. Distribuciones marginales para los coeficientes para $n = 4$ con 1 onda incidente ((a), (b) y (c)) y 3 ondas incidentes ((d), (e) y (f)).	43
4.5. Resultados para $n = 5$ con 1 onda incidente ((a) y (b)) y 3 ondas incidentes ((c) y (d)).	44
4.6. Curva de Prueba para el Problema Inverso: Kite	45

4.7. Valores de los vectores de parámetros θ_n con diferentes valores de n para 3 ondas incidentes	47
4.8. Valores de los vectores de parámetros θ_n con diferentes valores de n para 5 ondas incidentes	47
4.9. Valores de los vectores de parámetros θ_n con diferentes valores de n para 7 ondas incidentes	48
4.10. Valores de los vectores de parámetros θ_n con diferentes valores de n para 16 ondas incidentes	48
4.11. Parámetros estimados incrementando el número de coeficientes para distintos números de ondas incidentes.	49
4.12. Error relativo en las estimaciones con diferente números de ondas incidentes	50
4.13. Regiones de confianza de la estimación con 3 ondas incidentes u^i .	51
4.14. Curvas estimadas con 3 ondas incidentes u^i	52
4.15. Regiones de confianza de la estimación con 5 ondas incidentes u^i .	53
4.16. Curvas estimadas con 5 ondas incidentes u^i	54
4.17. Regiones de confianza de la estimación con 7 ondas incidentes u^i .	55
4.18. Curvas estimadas con 7 ondas incidentes u^i	56
4.19. Regiones de confianza de la estimación con 16 ondas incidentes u^i .	57
4.20. Curvas estimadas con 16 ondas incidentes u^i	58

Lista de Tablas

2.1. Convergencia del Método de Nystrom para el Potencial de Capa Simple	14
2.2. Convergencia del Método de Nystrom para el Potencial Combinado	14
3.1. Números de Condición para la Matriz del Potencial de Capa Simple A^s	25
3.2. Números de Condición para la Matriz del Potencial Combinado A^c	25

Lista de Algoritmos

1.	Problema Directo	24
2.	Esquema General Metodos de Descenso para Matrices no Simétricas	29
3.	Algoritmo “Cascada”	31
4.	Construcción de la Matriz A^c	32
5.	Inicialización de GCR	33
6.	Actualización	34
7.	Cálculo de p_{k+1}	34
8.	Esquema de Paralelización del Método GCR(k) para resolver el Problema Directo con Potencial Combinado	35
9.	Esquema General para la Solución al Problema Inverso	39

Capítulo 1

Introducción

Consideremos el siguiente fenómeno: un aparato emisor produce un cambio de presión en un medio homogéneo e isotrópico, este cambio de presión se propaga en forma de onda acústica; suponemos ahora que se tiene un objeto dentro del medio, la onda acústica incide sobre la superficie del objeto interactuando con él para posteriormente ser reflejada o dispersada. La onda reflejada cambia su forma dependiendo de la superficie del objeto y de cómo es su interacción con el mismo. La teoría de dispersión de ondas se refiere, en términos generales, al estudio de las ondas reflejadas en éste fenómeno. En el contexto de la teoría de dispersión de ondas, es posible predecir el comportamiento de las ondas reflejadas o dispersadas. Es objeto de nuestro estudio, estudiar tales ondas para recuperar la información correspondiente a la superficie del objeto que causa la dispersión.

El fenómeno arriba descrito, se presenta frecuentemente en la naturaleza. Un ejemplo claro es la forma como ciertos animales (murciélagos, delfines) obtienen información de su entorno mediante la emisión y recepción de ondas acústicas de diferentes frecuencias que son dispersadas por objetos presentes en el medio.

De igual forma, aplicaciones como radar o el ultrasonido en medicina se basan en el estudio de ondas dispersadas.

El modelo físico asociado a este fenómeno, ha sido ampliamente estudiado [3], [4], [7], [8], [9], [10], [11]. Es un tema maduro en lo que se refiere al aspecto teórico. Es necesario proporcionar métodos computacionales adecuados para tener una estrategia de solución robusta y eficiente.

El problema de dispersión de ondas presenta varias dificultades. No existe un procedimiento analítico para recuperar la información del obstáculo dispersor. En nuestro modelo, recuperar tal información implica resolver ecuaciones integrales en las que el dominio de integración es precisamente la superficie del obstáculo que desconocemos. Adicionalmente, las ecuaciones integrales presentan singularidades logarítmicas en sus kernels de integración que además son no lineales. Una vez tratadas tales singularidades, se requiere invertir un sistema de la forma $Ax = b$ donde la matriz A es densa, no simétrica y con valores complejos.

Hablaremos en el capítulo 2 del modelo físico asociado al fenómeno de nuestro interés, así como de las ecuaciones integrales que resuelven el problema directo, además definiremos el problema inverso en cuestión y su formulación como un problema Bayesiano. Describiremos en el capítulo 3 las distribuciones de probabilidad asociadas, bajo nuestros supuestos, para proporcionar la información necesaria al esquema Bayesiano en la solución del problema inverso. También mostramos en el capítulo 3, aspectos importantes de la implementación numérica. Presentaremos los resultados de nuestra solución al problema inverso dentro del esquema Bayesiano en el capítulo 4. Concluiremos nuestras aportaciones en el capítulo 5 y hablaremos del trabajo a futuro para obtener mejores resultados utilizando el mismo enfoque de solución al problema inverso.

Capítulo 2

Planteamiento del problema

Para estudiar problemas inversos se requiere un sólido conocimiento de la teoría del problema directo correspondiente. En este capítulo se hablará del modelo físico asociado al fenómeno descrito en el capítulo 1. Hablaremos entonces de las ecuaciones que gobiernan el comportamiento de las ondas en el modelo físico asociado al problema de dispersión de ondas acústicas. Describiremos las soluciones a la ecuación de Helmholtz, así como el comportamiento asintótico de las ondas dispersadas y la interacción de las mismas con el objeto dispersor. Mostraremos las ecuaciones integrales a resolver y hablaremos de la unicidad de las soluciones del problema directo. Abordaremos el problema inverso como un problema de inferencia Bayesiana y hablaremos de los aspectos de la estadística Bayesiana que hacen atractivo éste enfoque para resolver el problema de nuestro interés.

2.1. Modelo Físico

Consideremos una onda acústica de amplitud pequeña que se propaga en un medio homogéneo isotrópico en \mathbb{R}^3 visto como un fluido no viscoso. Esta onda produce variaciones en la presión del medio como:

$$p = -\frac{\partial U}{\partial t},$$

donde U es un potencial de velocidad que satisface la ecuación de onda:

$$\frac{1}{c^2} \frac{\partial^2 U}{\partial t^2} = \Delta U.$$

El potencial U es una onda armónica en el tiempo, de la forma:

$$U(x, t) = \operatorname{Re}\{u(x)e^{-i\omega t}\},$$

con frecuencia $\omega > 0$. La parte espacial que depende de los valores complejos u satisface la *ecuación de onda reducida* o *ecuación de Helmholtz*:

$$\Delta u + k^2 u = 0, \tag{2.1}$$

donde el número de onda k está dado por la constante positiva $k = \frac{\omega}{c}$.

Supongamos ahora que el medio tiene un objeto D en su interior, como se muestra en la Figura 2.1. En este caso, una onda incidente u^i es dispersada en una onda u^s dependiendo de la naturaleza del objeto D . La onda total $u = u^i + u^s$, debe satisfacer la ecuación de onda en el exterior $\mathbb{R}^3 \setminus \bar{D}$ de D y una condición de frontera en ∂D . En nuestro caso, estamos suponiendo que el obstáculo D es de frontera impenetrable y no absorbente y que por tanto, la onda total se anula en la frontera ∂D . Esto es $u^i = -u^s$ o dicho en otras palabras, se tiene condición de frontera tipo Dirichlet $u = 0$ en ∂D .

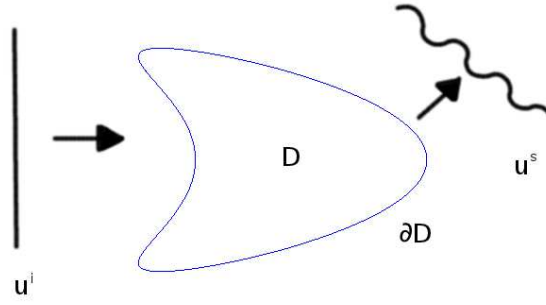


Figura 2.1: Fenómeno de Dispersión de Ondas.

Las ondas dispersadas u^s deben satisfacer una condición de radiación o decaimiento:

$$\lim_{r \rightarrow \infty} r \left(\frac{\partial u^s}{\partial r} - iku^s \right) = 0 \quad , \quad r = |x|. \quad (2.2)$$

Esta condición, llamada *condición de radiación de Sommerfeld* garantiza unicidad para los problemas de dispersión.

2.2. Solución Fundamental

Hablaremos ahora de las propiedades de la ecuación de Helmholtz (2.1) con número de onda positivo k . Muchas de estas pueden ser deducidas de la solución fundamental:

$$\Phi(x, y) = \frac{1}{4\pi} \frac{e^{ik|x-y|}}{|x-y|}, \quad x \neq y.$$

Para un y fijo, la solución fundamental satisface la ecuación de Helmholtz en $\mathbb{R}^3 \setminus \{y\}$.

Por la fórmula de Green tenemos que:

$$u(x) = \int_{\partial D} \left\{ \frac{\partial u}{\partial \nu}(y) \Phi(x, y) - u(y) \frac{\partial \Phi(x, y)}{\partial \nu(y)} \right\} ds(y)$$

$$- \int_D \{ \Delta u(y) + k^2 u(y) \} \Phi(x, y) dy, \quad x \in D,$$

en particular tenemos que si u satisface la ecuación de Helmholtz entonces:

$$u(x) = \int_{\partial D} \left\{ \frac{\partial u}{\partial \nu}(y) \Phi(x, y) - u(y) \frac{\partial \Phi(x, y)}{\partial \nu(y)} \right\} ds(y), \quad x \in D. \quad (2.3)$$

Una solución u de la ecuación de Helmholtz que cumpla la condición de radiación de Sommerfeld (2.2) es llamada solución radiante de la ecuación de Helmholtz. Las soluciones radiantas de la ecuación de Helmholtz tienen el comportamiento de ondas esféricas salientes:

$$u(x) = \frac{e^{ik|x|}}{|x|} \left\{ u_\infty(\hat{x}) + O\left(\frac{1}{|x|}\right) \right\}, \quad |x| \rightarrow \infty, \quad (2.4)$$

uniformemente en todas direcciones $\hat{x} = \frac{x}{|x|}$ donde la función u_∞ definida en la esfera unitaria es conocida como el mapeo de campo lejano de u .

El campo lejano u_∞ , en el contexto del modelo físico asociado al fenómeno físico, modela el decaimiento de la amplitud de las ondas dispersadas. Existe una correspondencia uno a uno entre una solución radiante u y su campo lejano u_∞ bajo un mapeo $F : u \mapsto u_\infty$ definido por las ecuaciones (2.3) y (2.4) de la siguiente forma:

$$u_\infty(\hat{x}) = \frac{1}{4\pi} \int_{\partial D} \left\{ u(y) \frac{\partial e^{-ik\hat{x}\cdot y}}{\partial \nu(y)} - \frac{\partial u}{\partial \nu}(y) e^{-ik\hat{x}\cdot y} \right\} ds(y).$$

donde $\hat{x} = \frac{x}{|x|}$. El mapeo F es mal planteado en el sentido de Hadamar debido a la no linealidad. No obstante, la inyectividad de F nos permite resolver el problema directo garantizando unicidad en el traslado de una solución radiante u a su campo lejano.

2.3. Potenciales Acústicos

Dada una función integrable φ , las integrales:

$$u(x) = \int_{\partial D} \varphi(y) \Phi(x, y) ds(y), \quad x \in \mathbb{R}^3 \setminus \partial D, \quad (2.5)$$

y

$$v(x) = \int_{\partial D} \varphi(y) \frac{\partial \Phi(x, y)}{\partial \nu(y)} ds(y), \quad x \in \mathbb{R}^3 \setminus \partial D, \quad (2.6)$$

son llamados respectivamente potencial acústico de Capa Simple y potencial acústico de Capa Doble con densidad φ . Ambos potenciales son soluciones a la ecuación de Helmholtz en \mathbb{R}^3 y en $\mathbb{R}^3 \setminus \bar{D}$ y satisfacen la condición de radiación de Sommerfeld.

Como consecuencia de la fórmula de Green, cualquier solución de la ecuación de Helmholtz puede ser expresada como combinación de potenciales de capa simple y capa doble:

$$u(x) = \int_{\partial D} \left\{ \frac{\partial \Phi(x, y)}{\partial \nu(y)} - i\eta \Phi(x, y) \right\} \varphi(y) ds(y), \quad x \in \mathbb{R}^3 \setminus \partial D, \quad (2.7)$$

con una densidad $\varphi \in C(\partial D)$ y un parámetro real de acoplamiento $\eta \neq 0$

Ahora bien, por otro lado tenemos las siguientes relaciones:

$$\begin{aligned} u(x) &= \int_{\partial D} \varphi(y) \Phi(x, y) ds(y), \quad x \in \partial D, \\ \frac{\partial u_{\pm}}{\partial \nu}(x) &= \int_{\partial D} \varphi(y) \frac{\partial \Phi(x, y)}{\partial \nu(y)} ds(y) \mp \frac{1}{2} \varphi(x), \quad x \in \partial D, \\ v_{\pm}(x) &= \int_{\partial D} \varphi(y) \frac{\partial \Phi(x, y)}{\partial \nu(y)} ds(y) \mp \frac{1}{2} \varphi(x), \quad x \in \partial D. \end{aligned}$$

Entonces el potencial combinado (2.7) puede ser escrito como la ecuación integral:

$$\varphi + K\varphi - i\eta S\varphi = 2f, \quad (2.8)$$

con:

$$(S\varphi)(x) = \int_{\partial D} \varphi(y)\Phi(x, y)ds(y), \quad x \in \partial D,$$

y

$$(K\varphi)(x) = \int_{\partial D} \varphi(y)\frac{\partial\Phi(x, y)}{\partial\nu(y)}ds(y), \quad x \in \partial D,$$

donde S y K son dos veces los potenciales acústicos de capa simple (2.5) y capa doble (2.6) respectivamente.

2.4. El Caso \mathbb{R}^2

En \mathbb{R}^2 , la solución fundamental a la ecuación de Helmholtz está dada por

$$\Phi(x, y) = \frac{i}{4}H_0^{(1)}(k|x - y|), \quad x \neq y,$$

con $H_0^{(1)}$ función de Hankel de tipo 1 y orden 0. La condición de radiación de Sommerfeld es reemplazada por:

$$\lim_{r \rightarrow \infty} \sqrt{r} \left(\frac{\partial u}{\partial r} - iku \right) = 0, \quad r = |x|,$$

uniformemente en todas direcciones $x/|x|$. La definición de mapeo de campo lejano es reemplazada entonces por:

$$u_\infty(\hat{x}) = \frac{e^{-i\frac{\pi}{4}}}{\sqrt{8\pi k}} \int_{\partial D} \left\{ u(y)\frac{\partial e^{-ik\hat{x}\cdot y}}{\partial\nu(y)} - \frac{\partial u}{\partial\nu}(y)e^{-ik\hat{x}\cdot y} \right\} ds(y),$$

para $|\hat{x}| = x/|x|$.

Supondremos ahora que la curva frontera ∂D tiene una representación paramétrica y analítica, de periodo 2π :

$$x(t) = (x_1(t), x_2(t)), \quad 0 \leq t \leq 2\pi,$$

en el sentido de las manecillas del reloj. Además satisface $[x'_1(t)]^2 + [x'_2(t)]^2 > 0$ para todo t . Entonces es posible transformar la ecuación integral (2.8) en la forma paramétrica:

$$\psi(t) - \int_0^{2\pi} \{L(t, \tau) + i\eta M(t, \tau)\} \psi(\tau) d\tau = g(t), \quad 0 \leq t \leq 2\pi,$$

donde $\psi(t) = \varphi(x(t))$, $g(t) = 2f(x(t))$ y los kernels estan dados por:

$$L(t, \tau) = \frac{ik}{2} x'_2(\tau) [x_1(\tau) - x_1(t)] - x'_1(\tau) [x_2(\tau) - x_2(t)] \frac{H_1^{(1)}(kr(t, \tau))}{r(t, \tau)},$$

$$M(t, \tau) = \frac{i}{2} H_0^{(1)}(kr(t, \tau)) \{[x'_1(\tau)]^2 + [x'_2(\tau)]^2\}^{1/2},$$

para $t \neq \tau$. Aquí definimos $r(t, \tau) = \{[x_1(t) - x_1(\tau)]^2 + [x_2(t) - x_2(\tau)]^2\}^{1/2}$.

Ahora bien, notemos que los kernels tienen singularidades logarítmicas en $t = \tau$. Entonces, para su tratamiento numérico, se dividen los kernels en:

$$L(t, \tau) = L_1(t, \tau) \ln \left(4 \sin^2 \frac{t - \tau}{2} \right) + L_2(t, \tau),$$

$$M(t, \tau) = M_1(t, \tau) \ln \left(4 \sin^2 \frac{t - \tau}{2} \right) + M_2(t, \tau),$$

donde:

$$L_1(t, \tau) = \frac{k}{2\pi} \{x'_2(\tau) [x_1(t) - x_1(\tau)] - x'_1(\tau) [x_2(t) - x_2(\tau)]\} \frac{J_1(kr(t, \tau))}{r(t, \tau)},$$

$$L_2(t, \tau) = L(t, \tau) - L_1(t, \tau) \ln \left(4 \sin^2 \frac{t - \tau}{2} \right),$$

$$M_1(t, \tau) = -\frac{1}{2\pi} J_0(kr(t, \tau)) \{[x'_1(\tau)]^2 + [x'_2(\tau)]^2\}^{1/2},$$

$$M_2(t, \tau) = M(t, \tau) - M_1(t, \tau) \ln \left(4 \sin^2 \frac{t - \tau}{2} \right).$$

Aquí, J_n corresponde a la función de Bessel de orden n .

Los kernels L_1, L_2, M_1 , y M_2 son analíticos. Los terminos diagonales se escriben entonces como:

$$L_2(t, t) = L(t, t) = \frac{1}{2\pi} \frac{x_1'(t)x_2''(t) - x_2'(t)x_1''(t)}{[x_1'(t)]^2 + [x_2'(t)]^2},$$

y

$$M_2(t, t) = \left\{ \frac{i}{2} - \frac{C}{\pi} - \frac{1}{2\pi} \ln \left(\frac{k^2}{4} \{ [x_1'(t)]^2 + [x_2'(t)]^2 \} \right) \right\} \{ [x_1'(t)]^2 + [x_2'(t)]^2 \}^{1/2},$$

para $0 \leq t \leq 2\pi$ y C la constante de Euler.

Por lo tanto, debemos resolver numéricamente una ecuación integral de la forma:

$$\psi(t) - \int_0^{2\pi} K(t, \tau)\psi(\tau)d\tau = g(t), \quad 0 \leq t \leq 2\pi, \quad (2.9)$$

donde el kernel puede ser escrito en la forma:

$$K(t, \tau) = K_1(t, \tau) \ln \left(4 \sin^2 \frac{t - \tau}{2} \right) + K_2(t, \tau),$$

con funciones analíticas K_1, K_2 y g . Es importante señalar que la singularidad logarítmica debe ser dividida de tal forma que se preserve la periodicidad de los kernels K_1 y K_2 .

2.5. Método de Nystrom

Supongamos que tenemos $f : \mathbb{R} \rightarrow \mathbb{R}$ y queremos calcular:

$$\int_0^{2\pi} \ln \left(4 \sin^2 \frac{t - \tau}{2} \right) f(\tau)d\tau, \text{ y } \int_0^{2\pi} f(\tau)d\tau.$$

En el caso de ecuaciones integrales de funciones analíticas periódicas, usar aproximaciones globales a través de polinomios trigonométricos, es superior a utilizar aproximaciones locales con splines polinomiales de bajo orden.

El método de Nystrom consiste en aproximación de las integrales por fórmulas de cuadratura. En nuestro caso, para los integrandos periódicos con periodo 2π , elegimos un conjunto de puntos equidistantes $t_j = \pi j/n, j = 0, \dots, 2n - 1$, y usamos la regla de cuadratura:

$$\int_0^{2\pi} \ln \left(4 \sin^2 \frac{t - \tau}{2} \right) f(\tau) d\tau \approx \sum_{j=0}^{2n-1} R_j^{(n)}(t) f(t_j), \quad 0 \leq t \leq 2\pi,$$

con los pesos de cuadratura dados por:

$$R_j^{(n)} = -\frac{2\pi}{n} \sum_{m=1}^{n-1} \frac{1}{m} \cos m(t - t_j) - \frac{\pi}{n^2} \cos n(t - t_j), \quad j = 0, \dots, 2n - 1,$$

y la regla trapezoidal:

$$\int_0^{2\pi} f(\tau) d\tau \approx \frac{\pi}{n} \sum_{j=0}^{2n-1} f(t_j).$$

Puesto que f es analítica, de acuerdo a las estimaciones libres de error en la derivada para el término de residuo en interpolación trigonométrica para funciones periódicas analíticas, los errores para las reglas de cuadratura mencionadas anteriormente, decrecen al menos exponencialmente conforme el numero de puntos $2n$ se incrementa. Más precisamente, el error es de orden $O(\exp(-n\sigma))$ donde σ denota la mitad del ancho de una franja paralela en el plano complejo dentro del cual la función analítica real f puede ser homológicamente extendida.

Usando el método de Nystrom, la ecuación integral (2.9) es reemplazada por la ecuación aproximante:

$$\psi^{(n)}(t) - \sum_{j=0}^{2n-1} \left\{ R_j^{(n)}(t) K_1(t, t_j) + \frac{\pi}{n} K_2(t, t_j) \right\} \psi^{(n)}(t_j) = g(t),$$

para $0 \leq t \leq 2\pi$, con:

$$R_j^{(n)} = R_j^{(n)}(0) = -\frac{2\pi}{n} \sum_{m=1}^{n-1} \frac{1}{m} \cos \frac{mj\pi}{n} - \frac{(-1)^j \pi}{n^2}, \quad j = 0, \dots, 2n - 1. \quad (2.10)$$

Puesto que la ecuación integral tiene solución única, y los kernels K_1 y K_2 y el lado derecho g son continuos, tenemos que:

1. El sistema lineal tiene solución única para n suficientemente grande.
2. Como $n \rightarrow \infty$ la solución aproximada $\psi^{(n)}$ converge uniformemente a la solución de la ecuación integral ψ .
3. El orden de convergencia de los errores de la cuadratura, se traslada al error de $\psi^{(n)} - \psi$.

Esto en particular significa que el error decrece exponencialmente, es decir, que existen constantes positivas C y σ tales que:

$$|\psi^{(n)}(t) - \psi(t)| \leq Ce^{-n\sigma}, \quad 0 \leq t \leq 2\pi$$

para todo n .

Una vez obtenida la densidad φ , podemos calcular el campo lejano de acuerdo a la siguiente relación:

$$u_\infty(\hat{x}) = \frac{e^{-i\frac{\pi}{4}}}{\sqrt{8\pi k}} \int_{\partial D} \{k\nu(y) \cdot \hat{x} + \eta\} e^{-ik\hat{x} \cdot y} \varphi(y) ds(y), \quad |\hat{x}| = 1, \quad (2.11)$$

la cual puede ser evaluada nuevamente con la regla trapezoidal.

Verificamos la precisión del método de Nystrom resolviendo el problema directo para la curva *kite* mostrada en la Figura 2.2. Los resultados se muestran en la Tabla 2.1 para el potencial de capa simple (2.5), y 2.2 para el potencial combinado (2.7). Mostramos únicamente el valor del campo lejano en las direcciones $d = (1, 0)$ y $-d = (-1, 0)$ para valores de número de onda de $k \in \{1, 5\}$. Podemos apreciar la convergencia superalgebraica, pues al incrementar al doble el número de puntos en cada caso, se tienen al menos dos dígitos correctos más

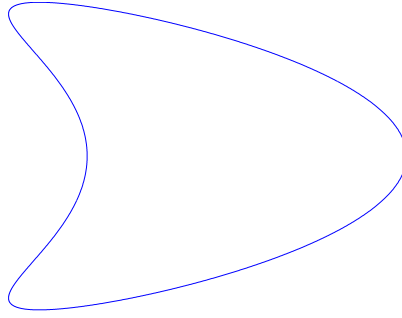


Figura 2.2: Curva de Prueba para el Problema Directo: Kite

para el valor del campo lejano. Cabe señalar que estos resultados coinciden con los valores mostrados en [4] y verifica la precisión de nuestra implementación.

Observamos en las Tablas 2.1 y 2.2, que los valores para el campo lejano coinciden para ambos potenciales conforme incrementamos el valor de n . Esto nos permite utilizar cualquiera de los dos potenciales para generar datos sintéticos en nuestros casos de prueba. Para la solución del problema inverso sintético, es indispensable que el método de solución del mismo sea ajeno al método con que se generaron los datos de prueba.

2.6. Problema Inverso

Recordemos el problema directo de dispersión de ondas, hacemos referencia a la Figura 2.1 para fines ilustrativos: dada la frontera del obstáculo dispersor ∂D , una o más ondas planas incidentes u^i y una condición de frontera, encontrar la onda dispersada u^s y en particular, su decaimiento a una distancia lejana al objeto dispersor. Existe una gran variedad de problemas inversos posibles en la dispersión de ondas acústicas. Sin embargo, nosotros nos enfocaremos en resolver

k	n	$\text{Re } u_\infty(d)$	$\text{Im } u_\infty(d)$	$\text{Re } u_\infty(-d)$	$\text{Im } u_\infty(-d)$
$k = 1,0$	8	-1,61838169568	1,41690941077	1,41690941077	1,41690941077
	16	-1,62747685461	1,3972218378	1,3972218378	1,3972218378
	32	-1,62745744541	1,3969446487	1,3969446487	1,3969446487
	64	-1,6274575037	1,39694488231	1,39694488231	1,39694488231
	128	-1,62745750369	1,39694488231	1,39694488231	1,39694488231
	256	-1,62745750369	1,39694488231	1,39694488231	1,39694488231
$k = 5,0$	8	-1,37709039719	-0,38452531273	-0,38452531273	-0,38452531273
	16	-2,36517464786	-0,253616783423	-0,253616783423	-0,253616783423
	32	-2,4655349367	-0,199212087324	-0,199212087324	-0,199212087324
	64	-2,47554379663	-0,199457879732	-0,199457879732	-0,199457879732
	128	-2,47554380143	-0,199457879742	-0,199457879742	-0,199457879742
	256	-2,47554380143	-0,199457879742	-0,199457879742	-0,199457879742

Cuadro 2.1: Convergencia del Método de Nystrom para el Potencial de Capa Simple

k	n	$\text{Re } u_\infty(d)$	$\text{Im } u_\infty(d)$	$\text{Re } u_\infty(-d)$	$\text{Im } u_\infty(-d)$
$k = 1,0$	8	-1,58134143988	1,45375214825	1,45375214825	1,45375214825
	16	-1,62642413032	1,39015283444	1,39015283444	1,39015283444
	32	-1,62745909725	1,39696610392	1,39696610392	1,39696610392
	64	-1,62745750365	1,39694488235	1,39694488235	1,39694488235
	128	-1,62745750369	1,39694488231	1,39694488231	1,39694488231
	256	-1,62745750369	1,39694488231	1,39694488231	1,39694488231
$k = 5,0$	8	-1,0489560285	-0,045366098255	-0,045366098255	-0,045366098255
	16	-1,94985346954	-0,393504341441	-0,393504341441	-0,393504341441
	32	-2,46742701064	-0,198925741834	-0,198925741834	-0,198925741834
	64	-2,47554379315	-0,199457879648	-0,199457879648	-0,199457879648
	128	-2,47554380143	-0,199457879742	-0,199457879742	-0,199457879742
	256	-2,47554380143	-0,199457879742	-0,199457879742	-0,199457879742

Cuadro 2.2: Convergencia del Método de Nystrom para el Potencial Combinado

el siguiente: dado un campo lejano para una o varias ondas planas incidentes u^i y sabiendo que el obstáculo dispersor es impenetrable y no absorbente, con condición de frontera Dirichlet, determinar la forma del objeto dispersor ∂D .

El problema inverso es mal planteado porque, en particular, el mapeo $Fu = u_\infty$ es mal planteado en el sentido de Hadamar pues es fuertemente no lineal. Esto hace que el problema inverso en cuestión sea difícil de resolver. Adicionalmente, el problema inverso no tiene una solución analítica. No obstante, se puede demostrar la unicidad de la solución del problema inverso partiendo de la inyectividad del operador F [4].

Los métodos clásicos para resolver problemas inversos, no proporcionan evidencia de la solución en la forma de una distribución de probabilidad y de igual forma, tampoco se agrega información acerca de la distribución de los parámetros que generan a la curva ∂D .

Una distribución de probabilidad para la curva ∂D dado un campo lejano u_∞ , nos dice qué tan probable es que la curva ∂D haya generado dicho campo lejano, a partir de una o varias ondas incidentes u^i conocidas. Resulta obvio pensar que la curva más probable sea la curva que soluciona el problema inverso.

Por tal motivo, plantearemos el problema inverso como un problema Bayesiano. Mostraremos conceptos básicos de estadística Bayesiana así como nuestra formulación del problema inverso en dicho contexto.

2.6.1. Estadística Bayesiana

La estadística Bayesiana es un enfoque de la Estadística que se basa en el teorema de Bayes. La estadística Bayesiana permite expresar la incertidumbre en términos de una medida de probabilidad. Más aun, se piensa que la probabilidad

es una medida de lo que se sabe acerca de un evento. Además toda probabilidad depende de una serie de supuestos que describen al problema de interés.

Teorema 2.6.1 (Teorema de Bayes) *Sean A, B dos eventos cualesquiera tales que $P(B) > 0$. Entonces*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$P(A)$ es llamada probabilidad a priori y representa el conocimiento que se tiene inicialmente sobre el evento A . $P(B|A)$ es el modelo observacional, indica cómo sería la probabilidad del evento B si se conociera A . $P(A|B)$ es llamada probabilidad posterior o posteriori y muestra cuál es la probabilidad de A después de haber observado B . La probabilidad $P(B)$ es un término de normalización que hace que $P(A|B)$ sea en efecto una medida de probabilidad.

El esquema Bayesiano permite incorporar conocimiento previo al problema que queremos resolver, a través de la probabilidad a priori. Este planteamiento vuelve al enfoque Bayesiano atractivo para los fines de ésta tesis.

2.6.2. Enfoque Bayesiano

Supongamos que tenemos mediciones del campo lejano $\hat{u}_\infty = (u_{\infty_0}, \dots, u_{\infty_n})$ con distribución $\pi_{\hat{u}_\infty|\theta}$ y se desea hacer inferencia sobre el vector de parámetros θ que es tal que $\theta \mapsto \partial D_\theta$, o bien, que el vector θ genera a la curva ∂D_θ . La solución Bayesiana a este problema es la siguiente: primero se propone una distribución a priori $\pi_0(\theta)$ para el vector de parámetros θ ; después se establece un modelo observacional $\pi(\hat{u}_\infty|\theta)$ que representa cómo es la probabilidad de los datos suponiendo que conocemos θ ; finalmente se utiliza la regla de Bayes para resumir el conocimiento previo y la evidencia que se tiene sobre θ en una

distribución de probabilidad dada por:

$$\pi(\theta|\hat{u}_\infty) = \frac{\pi(\hat{u}_\infty|\theta)\pi_0(\theta)}{\int \pi(\hat{u}_\infty|\theta)\pi_0(\theta)d\theta},$$

o equivalentemente:

$$\pi(\theta|\hat{u}_\infty) \approx \pi(\hat{u}_\infty|\theta)\pi_0(\theta),$$

pues el término $\int \pi(\hat{u}_\infty|\theta)\pi_0(\theta)d\theta$ es una constante de normalización. La distribución $\pi(\theta|\hat{u}_\infty)$ recibe el nombre de distribución posterior y es el objeto de nuestro estudio.

Recordemos nuevamente el problema inverso de nuestro interés: dadas mediciones del campo lejano u_∞ con una o más ondas incidentes u^i conocidas, determinar la curva frontera ∂D del obstáculo dispersor. En el contexto Bayesiano, lo anterior se traduce a identificar la distribución posterior $\pi(\theta|\hat{u}_\infty)$ a partir de $\hat{u}_\infty = (u_{\infty_0}, \dots, u_{\infty_n})$ una vez introducida la información a priori en forma de una distribución π_0 y de un modelo observacional. Dicho en otras palabras, consideramos que las observaciones de campo lejano son realizaciones de una variable aleatoria \hat{u}_∞ , definiremos una distribución de probabilidad para identificar cómo son las curvas que generan tales realizaciones de \hat{u}_∞ , así como determinar la forma cómo se obtienen tales mediciones.

Hablaremos entonces en el capítulo 3 de cada uno de los términos, con excepción de la constante de normalización, involucrados en la formulación Bayesiana del problema inverso.

Capítulo 3

Tratamiento del problema

3.1. Distribución a priori

Tener una distribución para los parámetros θ proporciona ventaja sobre otros planteamientos porque provee una gran cantidad de información sobre θ , además permite hacer uso de resultados de teoría de probabilidad. Necesitamos definir entonces, adecuadamente, el mapeo $\theta \mapsto \partial D_\theta$ de tal forma que podamos tener control en la distribución $\pi_0(\theta)$.

Para fines de esta tesis restringiremos nuestros estudios al caso \mathbb{R}^2 . Empezamos suponiendo que los objetos dispersores tienen frontera suave, con periodicidad 2π y que tiene simetría horizontal. Estos objetos pueden ser descritos de la siguiente manera:

$$\partial D_\theta(\alpha) = (r(\alpha) \cos(\alpha), r(\alpha) \sin(\alpha)), \quad \text{con } 0 \leq \alpha < 2\pi$$

donde r se escribe como la serie de Fourier:

$$r(\alpha) = \frac{a_0}{2} + \sum_{j=1}^n a_j \cos(j\alpha) + b_j \sin(j\alpha),$$

y entonces el vector de parámetros θ es:

$$\theta = (a_0, a_1, \dots, a_n, b_1, \dots, b_n).$$

Escribiremos θ_n para indicar el número de coeficientes $2n + 1$ del vector de parámetros θ .

Al ser los objetos dispersores $\partial D \in C^2(\mathbb{R})$, entonces el radio r puede ser expresado con pocos coeficientes de la serie de Fourier, o bien, con n pequeño. Entonces, la distribución π_0 debe hacer que los coeficientes de θ_n decaigan rápidamente conforme crece el valor de n . Además, al suponer que ∂D tiene simetría horizontal, los coeficientes b_i son cero en nuestros experimentos.

Como consecuencia de la identidad de Parseval, tenemos que:

$$\|\partial D_{\theta_n} - \partial D\|^2 = \sum_{m>n} |(a_m, b_m)|^2, \quad (3.1)$$

es el error cuadrático medio de la estimación de la serie ∂D_{θ_n} a la curva ∂D [12]. Más aun, al ser $\partial D \in C^2(\mathbb{R})$ entonces $|(a_m, b_m)| = O(m^{-2})$. Esto es,

$$|(a_m, b_m)| < C m^{-2},$$

o bien,

$$|(a_m, b_m)| < C \frac{1}{m^2},$$

para alguna constante $C > 0$.

Resulta entonces natural suponer que los coeficientes tienen una distribución normal:

$$a_m, b_m \sim N\left(0, \frac{C}{m^2}\right). \quad (3.2)$$

Notemos que este planteamiento reduce cada uno de los términos del error cuadrático medio (3.1), promoviendo así que este error disminuya y por consiguiente ∂D_{θ_n} aproxime a la curva ∂D .

El valor de la constante C no esta especificado. Como hemos mencionado, la distribución a priori contiene la información que conocemos del problema. Haremos entonces un análisis del tipo de curvas que se pueden obtener con diferentes valores de la constante C . Para esto, generamos vectores aleatorios $\theta_n = (a_0, \dots, a_n, b_1, \dots, b_n)$ donde los coeficientes a_m, b_m son muestras aleatorias de la distribución (3.2). Presentamos en la Figura 3.1 las curvas obtenidas para valores de $C \in \{0,5, 1, 2, 2,5\}$. Observamos que para los valores $C = 0,5$ y $C = 1$, las curvas resultantes presentan pocas oscilaciones y representan objetos esencialmente elipticos y con pocos detalles. Por otro lado, el valor $C = 2,5$ presenta curvas con demasiadas oscilaciones. Considerando el caso de prueba para el problema directo, es decir, la curva mostrada en la Figura 2.2, elegiremos el valor de $C = 2$ por porporcionar curvas similares a la curva de nuestro caso de prueba.

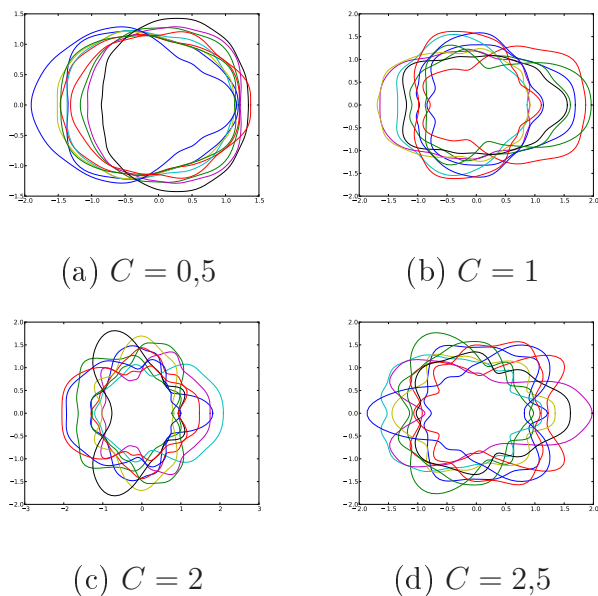


Figura 3.1: Análisis de la constante C

Consecuentemente, la distribución a priori para cada parámetro estará dada por:

$$a_m \sim e^{-m^2 \frac{\|a_m\|^2}{2}},$$

de forma análoga se tiene la distribución para b_m .

Observemos que lo anterior contempla coeficientes a_i con $i > 0$. El coeficiente a_0 tiene un comportamiento distinto pues éste término tiene que ver con la escala del objeto a estimar. Supondremos que éste término tiene una distribución uniforme y por tanto en nuestro planteamiento de la distribución a priori no incluimos el coeficiente a_0 .

Entonces, la distribución para todos los coeficientes a_i con $i > 0$, estaría dada como un producto de exponenciales:

$$\pi_0(\theta_n) \sim e^{-\frac{\|a_1\|^2}{2}} e^{-\frac{\|b_1\|^2}{2}} \dots e^{-n^2 \frac{\|a_n\|^2}{2}} e^{-n^2 \frac{\|b_n\|^2}{2}},$$

el cual, equivale a $\sum_{i=1}^n \frac{i^2}{2} (\|a_i\|^2 + \|b_i\|^2)$ tomando el negativo del logaritmo de ésta distribución.

3.2. Verosimilitud

El modelo observacional está dado por la verosimilitud $\pi(\hat{u}_\infty|\theta)$. En nuestro planteamiento, suponemos que las mediciones del campo lejano \hat{u}_∞ tienen un ruido observacional gaussiano η con media cero y varianza σ^2 . Esto es,

$$\hat{u}_\infty = F(\partial D) + \eta, \quad \text{con: } \eta \sim N(0, \sigma^2).$$

Entonces el modelo observacional estará dada por la verosimilitud

$$\eta = \hat{u}_\infty - F(\partial D)$$

y por consiguiente, la máxima logverosimilitud se tiene cuando

$$\|\hat{u}_\infty - F(\partial D)\|^2$$

es mínimo. Equivalentemente, en nuestro planteamiento, cuando:

$$\|\hat{u}_\infty - F(\partial D_{\theta_n})\|^2$$

es mínimo.

Resulta conveniente para los propósitos de ésta tesis, tener una forma eficiente de resolver el problema directo. Una evaluación no precisa de $F(\partial D_{\theta_n})$ generaría errores que pueden ser mas grandes en magnitud, que el error que induce el nivel de ruido y por consiguiente, una incorrecta interpretación de los resultados.

3.3. Distribución Posterior

La inferencia Bayesiana presenta dos principales dificultades: es difícil seleccionar la distribución a priori adecuada para cada problema y por otra parte, generalmente es complicado realizar los cálculos para la distribución posterior. Afortunadamente, estamos en posición para utilizar métodos MCMC (Markov Chain Monte Carlo). Los métodos MCMC permiten tener una muestra de una distribución de probabilidad π sin tener que simular directamente de dicha distribución. Para ello, estos métodos construyen una cadena de Markov Ergódica cuya distribución estacionaria es π .

Utilizando entonces un método MCMC, podemos tener una muestra de la distribución posterior $\pi(\theta|\hat{u}_\infty)$ y poder dar evidencia de los alcances de la solución al problema inverso. Entonces, el principal esfuerzo de ésta tesis consiste en

definir un conjunto de parámetros θ_n que generan a las curvas ∂D_{θ_n} y estudiar el comportamiento de la distribución posterior, bajo el muestreo dado por el método MCMC, conforme variamos el número n de éstos parámetros.

El algoritmo *t-walk* es un método MCMC de propósito general para distribuciones continuas arbitrarias[2]. Las distribuciones a estimar por el t-walk pueden ser de escala y estructura de correlación arbitrarias. Además, el algoritmo t-walk no requiere parámetros de ajuste o adaptativos y converge debido a que está construido como un algoritmo Metrópolis-Hastings. El t-walk proporciona buenos resultados para algunos ejemplos con numeros de parámetros relativamente grandes (≈ 200)[2].

La finalidad del t-walk, es precisamente permitirnos enfocarnos más en el análisis de los datos que en el algoritmo MCMC. Por tanto, utilizaremos este algoritmo como herramienta para estudiar la distribución posterior en la solución del problema inverso.

3.4. Implementación Numérica

Como hemos mencionado con anterioridad, requerimos una implementación que resuelva el problema directo de forma precisa. Mostraremos entonces el procedimiento utilizado que se basa en el enfoque de ecuaciones integrales y el método de Nyström mencionado en el capítulo 2. Presentamos también los algoritmos utilizados. Es importante señalar que estamos considerando que el problema directo puede ser resuelto utilizando el Potencial de Capa Simple o el Potencial Combinado. En la solución del problema inverso, realizaremos múltiples evaluaciones del problema directo con distintas curvas, por tanto, para dar resultados que no dependan del método numérico con que generamos

los datos sintéticos, utilizaremos el Potencial de Capa Simple para generar el campo lejano asociado al caso de prueba y el Potencial Combinado para las evaluaciones del problema directo.

Algoritmo 1 Problema Directo

Require: $\partial D_{\theta_n}, u^i, k$

Ensure: u_∞

Calcula los pesos de Nyström: R

if Potencial Combinado **then**

Calcula los potenciales: L, M

Calcula la matriz $A = I - L - ikM$

else

Calcula el potencial de Capa Simple: M

Calcula la matriz $A = M$

end if

Resuelve $A\varphi = -u^i$, para φ

Calcula $u^\infty = Z\varphi$ $\{Z$ es el mapeo de campo lejano (2.11) $\}$

Obsevemos ahora que, en el Algoritmo 1, se necesita resolver el sistema:

$$A\varphi = -u^i \quad \text{para } \varphi, \quad (3.3)$$

donde $A = A^s$ ó $A = A^c$ definiendo A^s la matriz generada con el potencial de capa simple (2.5) y A^c la matriz generada con el potencial combinado (2.7). Ambas matrices son densas, no simétricas y con coeficientes complejos, además son de tamaño $N \times N$ con N el número de puntos t_j , $j = 0, \dots, N-1$, con que definimos la curva $\partial D_{\theta_n}(t_j)$. Es necesario entonces, hacer un análisis de ambas matrices para definir un método numérico adecuado para resolver el sistema (3.3).

3.4.1. Análisis de la Matriz A

Primeramente, calculamos el número de condición de ambas matrices A^s y A^c . Presentamos en la Tabla 3.1 el número de condición para la matriz A^s . Análogamente, la Tabla 3.2 muestra los números de condición correspondientes a A^c . En ambos casos fijamos el número de onda en $k = 1$, aunque cabe mencionar que el comportamiento mostrado en las Tablas 3.1 y 3.2 es similar al obtenido variando el número de onda.

Número de puntos N	Número de condición
8	8,59231804803
16	20,1673779247
32	49,574140683
64	113,561337804
128	249,056139536
256	528,588651494
512	1097,93428681

Cuadro 3.1: Números de Condición para la Matriz del Potencial de Capa Simple A^s .

Número de puntos N	Número de condición
8	2,68204675516
16	3,50171633788
32	3,59780748843
64	3,598659748
128	3,59865866317
256	3,59865866317
512	3,59865866317

Cuadro 3.2: Números de Condición para la Matriz del Potencial Combinado A^c .

Observamos en las Tablas 3.1 y 3.2 indican que las matrices no están mal condicionadas en ambos casos. Sin embargo, A^c tiene un mejor número de condición y por tal motivo elegimos el potencial combinado (2.7) para hacer las evaluaciones del problema directo y el potencial de capa simple (2.5) para generar los datos sintéticos. Los métodos directos, como la factorización PLU,

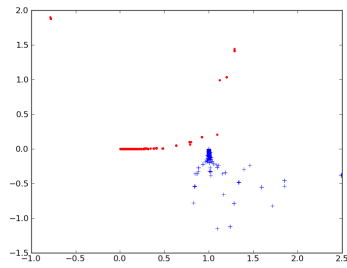
pueden resolver el sistema (3.3) sin embargo, estos métodos suelen ser lentos cuando se incrementa la dimensión de la matriz a factorizar.

Elegiremos entonces un método iterativo. Para esto, vale la pena hacer un análisis de los eigenvalores de las matrices de ambos potenciales para elegir un método que nos proporcione resultados adecuados. Mostramos en la Figura 3.2 los eigenvalores para ambos potenciales con números de onda $k = 1, 2, 5, 10$. En dicha Figura, los símbolos '.' corresponden a los valores de los eigenvalores de la matriz A^s , dibujados en el plano complejo. Por otra parte, los símbolos '+' indican los eigenvalores de la matriz A^c . Observamos que en todos los casos, los eigenvalores se acumulan en el punto $(0, 0)$ para la matriz A^s y en el punto $(1, 0)$ para A^c conforme se incrementan las dimensiones de ambas matrices. Esto en particular nos dice que están acotados y que podemos utilizar el método Generalized Conjugate Residuals (GCR) para resolver eficientemente el sistema (3.3).

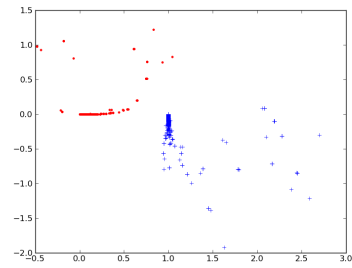
3.4.2. Generalized Conjugate Residuals, GCR

Los métodos iterativos basados en espacios de Krylov, como Gradiente Conjugado, GMRES o GCR en nuestro caso, ofrecen buenos resultados con un número pequeño de iteraciones. Recordemos que las matrices A^s y A^c no son simétricas, por lo que no es posible utilizar un método clásico como Gradiente Conjugado. Se puede pensar entonces, en utilizar un método para matrices generales como lo es GCR.

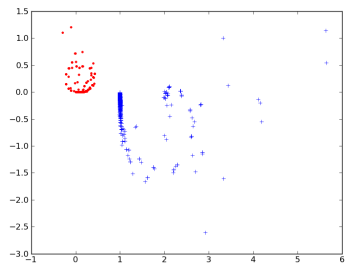
Se puede demostrar que si A no está demasiado lejos de ser normal, en el sentido de que el número de condición $\kappa(V)$ con $A = V\sigma(A)V^{-1}$ no es muy grande, y si se pueden encontrar polinomios de grado n propiamente norma-



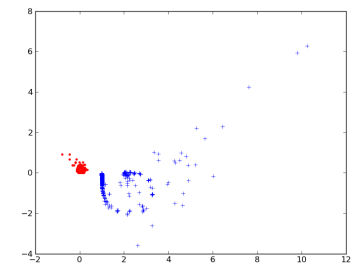
(a) $k = 1$



(b) $k = 2$

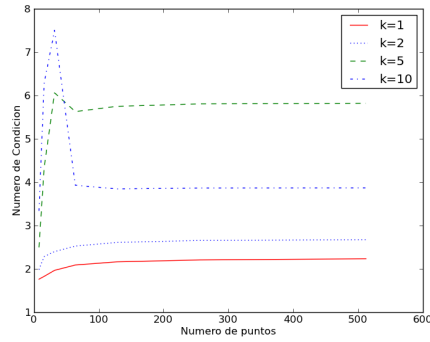


(c) $k = 5$

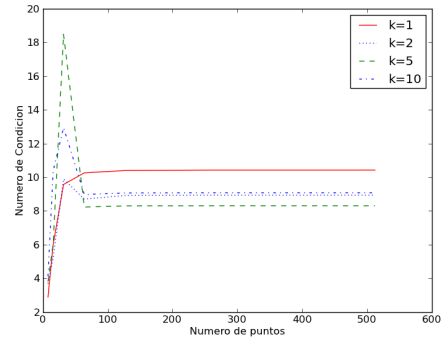


(d) $k = 10$

Figura 3.2: Eigenvalores de las Matrices A^s (.) y A^c (+) con diferentes números de onda.



(a) $\kappa(V)$ para A^s



(b) $\kappa(V)$ para A^c

Figura 3.3: Número de condición para las matrices de eigenvectores para el Potencial de Capa Simple (a) y el Potencial Combinado (b).

lizados, tales que su tamaño en el espectro $\sigma(A)$ decrezca rápidamente con n , entonces el método GCR converge rápidamente [6] [14].

Observemos entonces los números de condición de las matrices de eigenvectores para cada uno de los potenciales. En la Figura 3.3.(a) mostramos el número de condición de la matriz de eigenvectores asociada a A^s y en 3.3.(b) la correspondiente a A^c . Como podemos ver, los números de condición son aceptables y se vuelven casi constantes conforme se incrementa la dimensión de las matrices para los números de onda $k \in \{1, 2, 5, 10\}$.

Presentamos pues en el Algoritmo 2 el esquema general de los métodos de descenso para matrices no simétricas. La diferencia entre los distintos métodos (GCR, GMRES, MR, etc.) se encuentra en la elección de la dirección p_{i+1} .

Para el método GCR, a partir del Algoritmo 2, basta elegir la dirección p_{i+1} como:

$$p_{i+1} = r_i + \sum_{j=0}^i b_j^{(i)} p_j, \quad (3.4)$$

Algoritmo 2 Esquema General Metodos de Descenso para Matrices no Simétricas

Require: A, φ_0, u^i

Ensure: φ

Calcular $r_0 = -u^i - A\varphi_0$

Hacer $p_0 = r_0$

for $i \leq n$ **do**

$$a_i = \frac{(r_i, Ap_i)}{(Ap_i, Ap_i)}$$

$$x_{i+1} = x_i + a_i p_i$$

$$r_{i+1} = r_i - a_i Ap_i$$

Calcula p_{i+1}

end for

donde

$$b_j^{(i)} = -\frac{(Ar_{i+1}, Ap_j)}{(Ap_j, Ap_j)}, \quad j \leq i.$$

La implementación del método GCR es bastante simple pero, como podemos observar, el cálculo de (3.4) implica un alto costo computacional tanto en almacenamiento como tiempo de cómputo. El costo para la i -ésima iteración es de $(3(i+1) + 4)n + 1$ productos matriz vector y de $(2(i+2) + 2)n$ localidades de memoria para almacenamiento. Es posible sin embargo, utilizar una alternativa que consiste en reiniciar el método GCR cada $k+1$ iteraciones. El vector φ_{k+1} es ahora el nuevo punto inicial. El costo se reduce entonces considerablemente pues en general menos de k vectores son necesarios para calcular (3.4). Esta variante es conocida como GCR(k).

El costo computacional para el método GCR(k) es de $(3(i+1) + 4)n + 1$ operaciones matriz vector y de $(2(i+2) + 2)n$ localidades de memoria en

la i -ésima iteración. Se tiene además un resultado de convergencia para este método [6].

Tenemos entonces que el método $\text{GCR}(k)$ converge rápidamente para ambas matrices A^s y A^c . Aun utilizando $\text{GCR}(k)$ el costo computacional resulta elevado. Sin embargo, es posible hacer una implementación en paralelo que permite reducir el tiempo de computo requerido y proporcionar tiempos accesibles para la resolución del sistema (3.3).

3.4.3. Esquema de Paralelización

Un muestreo estadísticamente significativo requiere un gran número de evaluaciones del problema directo. Esto se traduce en la práctica, en alto costo computacional. Por tanto, consideramos necesario paralelizar cada evaluación de dicho problema. Observamos que cada entrada de la matriz A del sistema (3.3) es resultado de la evaluación del potencial correspondiente en el punto (t_i, t_j) , $i, j = 0, \dots, N - 1$ y que la evaluación de cada punto es independiente de los demás. Entonces podemos construir un arreglo de hilos en los que se evalúe cada punto en paralelo. Después, el sistema (3.3) puede ser resuelto para φ utilizando una versión paralelizada del método $\text{GCR}(k)$.

Presentaremos primero, 4 etapas que componen nuestro esquema de paralelización: construcción de la matriz A^c , inicialización de $\text{GCR}(k)$, actualización y calculo de p . Pero antes, mostramos en el Algoritmo 3 un procedimiento que nos permite paralelizar los productos punto.

Presentamos entonces en el Algoritmo 8 nuestro esquema general basado de paralelización para resolver el problema directo usando la matriz del Potencial Combinado A^c y el método $\text{GCR}(k)$ utilizando los Algoritmos 4, 5, 6 y 7.

Algoritmo 3 Algoritmo “Cascada”

Require: $u, v \in \mathbb{C}^n$, $n/2$ hilos en 1 bloque

Ensure: $u \cdot v$

```
for hilo  $i$  do
   $w_i = \bar{u}_i v_i$ 
   $w_{i+n/2} = \bar{u}_{i+n/2} v_{i+n/2}$ 
  if  $n \geq 512$  then
    if  $i \leq 256$  then
       $w_i = w_i + w_{i+256}$ 
    end if
  end if
  if  $n \geq 256$  then
    if  $i \leq 128$  then
       $w_i = w_i + w_{i+128}$ 
    end if
  end if
  if  $n \geq 128$  then
    if  $i \leq 64$  then
       $w_i = w_i + w_{i+64}$ 
    end if
  end if
  if  $i \leq 32$  then
    if  $n \geq 64$  then
       $w_i = w_i + w_{i+32}$ 
    end if
    if  $n \geq 32$  then
       $w_i = w_i + w_{i+16}$ 
    end if
    if  $n \geq 16$  then
       $w_i = w_i + w_{i+8}$ 
    end if
    if  $n \geq 8$  then
       $w_i = w_i + w_{i+4}$ 
    end if
    if  $n \geq 4$  then
       $w_i = w_i + w_{i+2}$ 
    end if
    if  $n \geq 2$  then
       $w_i = w_i + w_{i+1}$ 
    end if
  end if
end for

return  $w_0$ 
```

Algoritmo 4 Construcción de la Matriz A^c

Require: $\partial D, k, n$ hilos y n bloques**Ensure:** A^c

```
for hilo  $l$  en bloque  $j$  do
  Calcula  $L = R_{|l-j|}^{(n/2)} L_1(t_l, t_j) + \frac{2\pi}{n} L_2(t_l, t_j)$ 
  Calcula  $M = R_{|l-j|}^{(n/2)} M_1(t_l, t_j) + \frac{2\pi}{n} M_2(t_l, t_j)$ 
  if  $i = j$  then
     $A_{c_{ij}} = 1 - (L - ikM)$ 
  else
     $A_{c_{ij}} = -(L - ikM)$ 
  end if
end for
return  $A^c$ 
```

El criterio de convergencia del Algoritmo 8 es $\|r_k\|/\|u^i\| \leq \varepsilon$. En la Figura 3.4 se muestran los tiempos obtenidos con cuatro implementaciones realizadas para resolver el problema directo con la matriz A^c . Las implementaciones hechas en los lenguajes Matlab y Python, resuelven el sistema (3.3) con métodos proporcionados por cada lenguaje (GMRES en Matlab y SVD con Python). Por otro lado, la implementación en C utiliza nuestra implementación del método GCR tal y como se describe en el Algoritmo 2, mientras que la implementación hecha en C con CUDA utiliza la versión paralelizada de $GCR(k)$ descrita en el Algoritmo 8 con $k = n/2$. Podemos observar que nuestra implementación hecha en C es bastante costosa en cuanto a tiempo de ejecución se refiere; le sigue la implementación en Matlab, posteriormente la que fue hecha en Python y por último la versión paralelizada. Este resultado nos indica que la diferencia que existe

Algoritmo 5 Inicialización de GCR

Require: $A^c, u^i, \varphi_0, n/2$ hilos en 1 bloque

Ensure: r_0, p_0

for hilo l **do**

 Calcula $r_l = (A_l^c, x_0)$ { A_l^c el l -ésimo renglon de A^c }

 Calcula $r_{l+n/2} = (A_{l+n/2}^c, x_0)$

$p_l = r_l$

$p_{l+n/2} = r_{l+n/2}$

 Calcula (A_l^c, r)

 Calcula $(A_{l+n/2}^c, r)$

 Calcular $\|r\|$ { con Algoritmo 3}

 Calcular $\|-u^i\|$ { con Algoritmo 3}

end for

Algoritmo 6 Actualización

Require: $A, p_k, r_k, \varphi_{k+1}, n/2$ hilos en 1 bloque

Ensure: $r_{k+1}, A^c p_{k+1}, \|A^c p_{k+1}\|^2, \|r_{k+1}\|^2$

for hilo l **do**

Calcular $A^c p_{(k+1)l} = (A_l^c, p_k)$

Calcular $A^c p_{(k+1)(l+n/2)} = (A_{l+2}^c, p_k)$

Calcular $(A^c p_{k+1}, r)$ { con Algoritmo 3}

Calcular $(A^c p_{k+1}, A^c p_{k+1})$ { con Algoritmo 3}

$$a_i = -\frac{(A^c p_{k+1}, r)}{(A^c p_{k+1}, A^c p_{k+1})}$$

$$\varphi_{(k+1)l} = \varphi_{kl} + a_i p_{kl}$$

$$r_{(k+1)l} = r_{kl} - a_i A p_{kl}$$

Calcular $A^c r_{(k+1)l} = (A_l^c, r_k)$

Calcular $A^c r_{(k+1)(l+n/2)} = (A_{l+2}^c, r_k)$

Calcular (r_{k+1}, r_{k+1}) { con Algoritmo 3}

Hacer $p_{(k+1)l} = r_{(k+1)l}$

end for

Algoritmo 7 Cálculo de p_{k+1}

Require: $A, p_k, u^i, n/2$ hilos en k bloques

Ensure: p_{k+1}

for hilo l en bloque j **do**

Calcula $b_j = \frac{(A p_j, A r_{(k+1)})}{A p_j}$ { con Algoritmo 3}

$$p_{(k+1)l} = p_{(k+1)l} + b_j p_{jl}$$

$$p_{(k+1)(l+n/2)} = p_{(k+1)(l+n/2)} + b_j p_{j(l+n/2)}$$

end for

Algoritmo 8 Esquema de Paralelización del Método GCR(k) para resolver el

Problema Directo con Potencial Combinado

Require: $\partial D_{\theta_n}, u^i, k$

Ensure: u_∞

Se definen n hilos y n bloques

Se calcula A_c con el Algoritmo 4

Se definen n hilos 1 bloque

for hilo l **do**

$$u_l^i = -2 * e^{ik\hat{x}\cdot x_i} \text{ \{onda incidente\}}$$

end for

while $counter < n$ & no convergencia **do**

Se definen $n/2$ hilos en 1 bloque {Inicialización}

Se inicializa GCR(k) con el Algoritmo 5

while $k < reinicio$ & no convergencia **do**

Actualizar φ con el Algoritmo 6

Actualizar p con el Algoritmo 7

$$k = k + 1$$

end while

$$counter = counter + 1$$

end while

Se definen n hilos 1 bloque

for hilo l **do**

Calcular $u_{\infty l}$ con la regla de trapecio.

end for

return $\sum_{l=0}^n u_{\infty l}$

entre la implementación en paralelo y las demás, permite ahorrar tiempo de cómputo al realizar el muestreo estadístico necesario para resolver el problema inverso. Consecuentemente, podemos proporcionar información de la solución al problema inverso en tiempos aceptables y con una mejor aproximación a la distribución posterior.

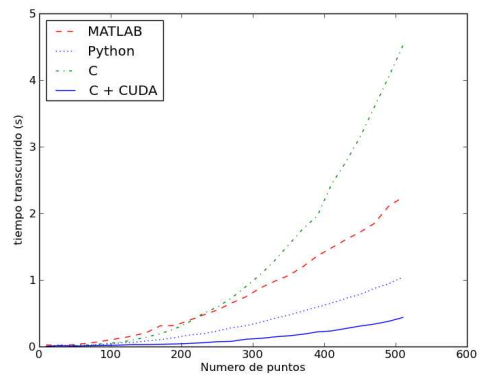


Figura 3.4: Tiempos de Ejecución obtenidos con diferentes implementaciones

Capítulo 4

Resultados

Presentaremos en este capítulo los resultados obtenidos en nuestra implementación para resolver el problema inverso. Primero, construimos un ejemplo totalmente sintético que nos permite aprender a controlar e interpretar los resultados que proporcionan las distribuciones a priori supuestas, así como la herramienta t-walk.

Es importante señalar que, no obstante que la implementación del problema directo hecha con C + CUDA proporciona resultados en menor tiempo, utilizamos nuestras implementaciones en Python. Este lenguaje proporciona herramientas útiles para fines ilustrativos y para presentar de forma adecuada los resultados obtenidos.

Presentamos en el Algoritmo 9 el procedimiento a seguir para resolver el problema inverso utilizando t-walk. Señalamos nuevamente el hecho de que se generen los datos para el campo lejano a estimar utilizando el potencial de capa simple (2.5) mientras que las evaluaciones de las muestras θ_n utilizan el potencial combinado (2.7). Esto permite que nuestro resultado no dependa de los métodos numéricos empleados. Cabe mencionar que sólo utilizamos las últimas

M evaluaciones resultantes del t-walk debido a que la distribución estacionaria converge a la distribución posterior y por tanto, a partir de que el t-walk se estabiliza es cuando podemos tener una buena aproximación. Los resultados que proporciona el t-walk son histogramas que nos muestran las distribuciones marginales de cada uno de los coeficientes estimados. Podemos pensar en elegir la media de cada una de estas distribuciones como nuestra mejor estimación, pues cuando la estimación se estabiliza, las distribuciones marginales se asemejan a Gaussianas tales que la media se encuentra muy cerca de los valores más probables para cada coeficiente.

Los valores de la distribución posterior son el resultado de la evaluación de un funcional de “energía” definido por la logverosimilitud y el logaritmo de la distribución a priori como:

$$E = \|u_\infty - \tilde{u}_\infty\|^2 + \sum_{i=1}^n \frac{i^2}{2} (\|a_i\|^2 + \|b_i\|^2).$$

Nuestro objetivo principal en esta tesis, es estudiar la dependencia de la distribución posterior en la discretización de la representación del radio en series de Fourier. Por tanto, mostraremos resultados de diversos experimento manteniendo constantes el número de onda, número de puntos, número de angulos de observación, numero de iteraciones de t-walk, numero M , y variaremos el número de coeficientes $2n + 1$ del vector θ_n y el número de ondas incidentes para ilustrar la mejora en los resultados.

Algoritmo 9 Esquema General para la Solución al Problema Inverso

Require: $\partial D, \hat{x}, u^i$

Ensure: $\tilde{\theta}_n$

- 1: Calcula u_∞ asociado a ∂D para todas las ondas incidentes u^i en los puntos de observación \hat{x} utilizando el potencial de Capa Simple y le agregamos ruido aditivo Gaussiano $\eta \sim N(0, 0,01)$.
 - 2: Definimos el número de coeficientes $2n + 1$ del vector de parámetros θ_n
 - 3: Definimos la logverosimilitud como $\|u_\infty - \tilde{u}_\infty\|^2$ donde \tilde{u}_∞ es el campo lejano asociado a una curva $\partial \tilde{D}$ generada con una muestra aleatoria en el espacio de parámetros θ_n utilizando el Potencial Combinado
 - 4: Definimos la distribución a priori π_0 para los parámetros θ_n
 - 5: Definimos el conjunto del que se va a muestrear θ_n
 - 6: Se evalúa con t-walk para un número fijo suficientemente grande de iteraciones
 - 7: Tomamos los últimos M resultados de las evaluaciones del t-walk con menor valor de energía
 - 8: Calculamos la media para cada coeficiente de los M vectores θ_n
 - 9: **return** $\tilde{\theta}_n$ el vector de medias de los vectores θ_n
-

4.1. Caso de Prueba Fish

Para nuestro primer experimento, utilizaremos la curva mostrada en la Figura 4.1. Esta curva fue generada con los parametros siguientes:

$$\theta_4 = (2,5, 0,5, 0, 0,25, -0,25, 0, 0, 0, 0),$$

recordando que

$$\partial D_\theta(\alpha) = (r(\alpha) \cos(\alpha), r(\alpha) \sin(\alpha)), \quad \text{con } 0 \leq \alpha < 2\pi$$

donde r se escribe como la serie de Fourier:

$$r(\alpha) = \frac{a_0}{2} + \sum_{j=1}^n a_j \cos(j\alpha) + b_j \sin(j\alpha),$$

y entonces el vector de parámetros θ_n es:

$$\theta_n = (a_0, a_1, \dots, a_n, b_1, \dots, b_n).$$

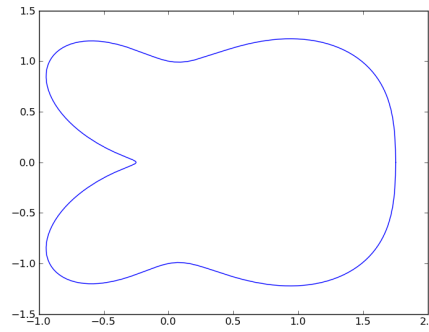
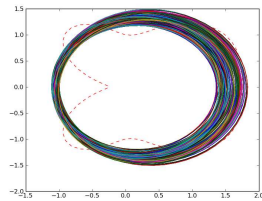


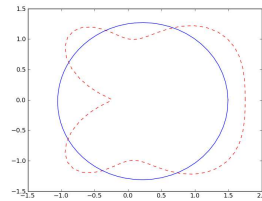
Figura 4.1: Figura Fish

Presentamos en la Figura 4.2 los resultados obtenidos para $\theta_1 = (a_0, a_1, b_1)$, es decir, tratando de recuperar la curva original utilizando sólo tres coeficientes

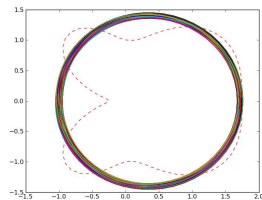
de Fourier. En la Figura 4.2.(a) observamos que la región de confianza de la estimación es más delgada por la parte derecha que por la izquierda. Esto se debe a que la onda incidente viaja con dirección $(1,0)$ y por tanto hay un fenómeno de oclusión con el objeto que no permite que las mediciones de los puntos de observación sean correctas. Para corroborar este hecho incrementamos a 3 el número de ondas incidentes con direcciones en los ángulos $0, \frac{2}{3}\pi$ y $\frac{4}{3}\pi$. Observamos que en éste caso 4.2.(c) la región de confianza es más delgada que en el caso anterior. Las curvas estimadas en ambos casos se muestran en las Figuras 4.2.(b) y 4.2.(d), en ambos casos tenemos elipses, pues es el único tipo de curvas que podemos escribir con nuestra representación con $n = 1$.



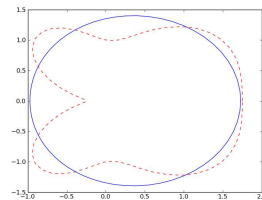
(a) Región de Confianza



(b) Curva estimada



(c) Región de Confianza



(d) Curva estimada

Figura 4.2: Resultados para $n = 1$ con 1 onda incidente ((a) y (b)) y 3 ondas incidentes ((c) y (d)).

Probamos ahora con $n = 4$, entonces tenemos 9 coeficientes a estimar que coinciden con el número de coeficientes con que generamos la curva 4.1. Los

resultados se muestran en la Figura 4.3. En este caso notamos, que las regiones de confianza obtenidas no están muy bien definidas y aunque proporcionen un resultado ligeramente aceptable, no es el tipo de resultados en los que podemos garantizar como correctos. Efectivamente podemos observar las distribuciones marginales en la Figura 4.4 y observamos que no es posible definir que valores son los más probables para nuestro resultado. Cabe señalar sin embargo que el resultado mostrado en la Figura 4.3.(d) es aceptable.

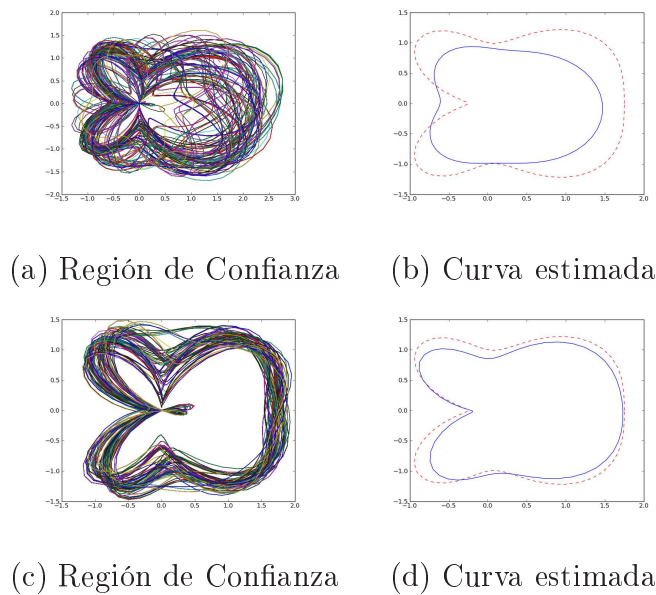


Figura 4.3: Resultados para $n = 4$ con 1 onda incidente ((a) y (b)) y 3 ondas incidentes ((c) y (d)).

Un aspecto importante de la formulación Bayesiana consiste en la elección de la distribución a priori para los parámetros θ_n . Los resultados mostrados en las Figuras 4.3 y 4.4 no tienen una distribución a priori definida, son resultados obtenidos únicamente utilizando información de la verosimilitud. Esto nos indica, que aun en un caso sintético relativamente sencillo y con un bajo nivel

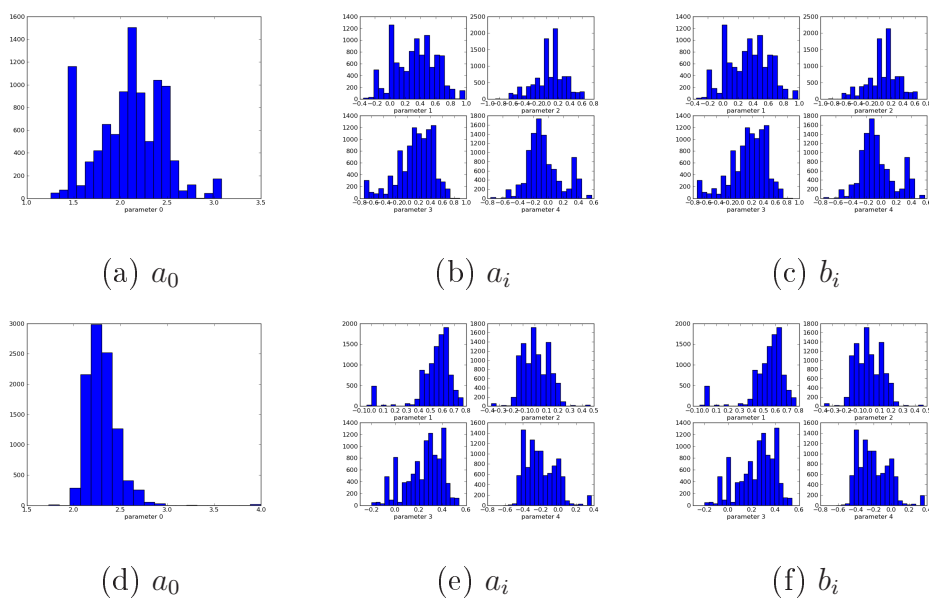


Figura 4.4: Distribuciones marginales para los coeficientes para $n = 4$ con 1 onda incidente ((a), (b) y (c)) y 3 ondas incidentes ((d), (e) y (f)).

de ruido, es fundamental una buena elección de la distribución a priori para tener aproximaciones más adecuadas. Incrementando el número de coeficientes a $n = 5$ se vuelve más notoria esta problemática, pues como observamos en la Figura 4.5 los resultados son totalmente incorrectos. Este tipo de resultados incorrectos pueden obtenerse incluso con número adecuado de coeficientes pero con una mala elección de la distribución a priori.

4.2. Caso de Prueba Kite

Realizamos experimentos para la curva mostrada en la Figura 4.6. Recordemos que ésta curva fue utilizada como caso de prueba para nuestra solución del problema directo. Para éste caso, utilizamos las condiciones mencionadas en el

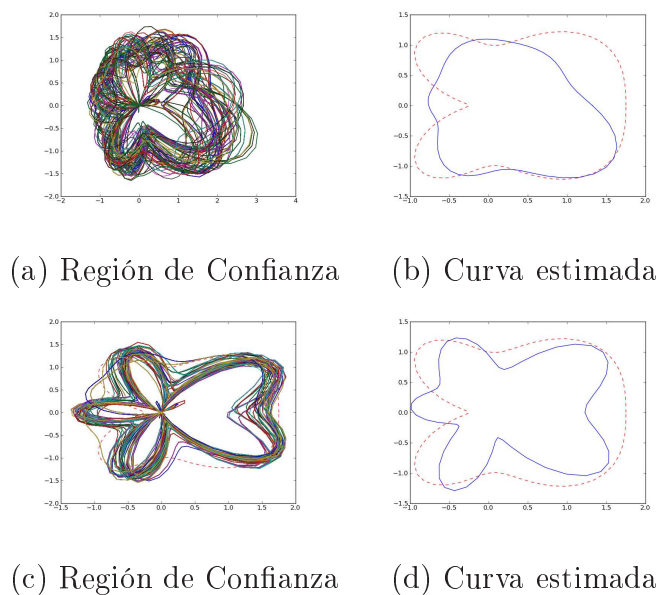


Figura 4.5: Resultados para $n = 5$ con 1 onda incidente ((a) y (b)) y 3 ondas incidentes ((c) y (d)).

capítulo 3, a saber, la verosimilitud como modelo observacional y la distribución a priori basada en la identidad de Parseval. Para este ejemplo utilizamos 64 puntos para definir la curva y 64 direcciones de observación.

Un hecho importante que se deduce de los resultados mostrados en éste capítulo, es que nuestro planteamiento no depende fuertemente del número de coeficientes del vector θ_n . Esto es, conforme se aumenta el número n , nuestro esquema promueve que los coeficientes se acerquen rápidamente a cero.

Haremos primeramente un análisis de los coeficientes obtenidos en nuestras estimaciones. Mostramos en las Figuras 4.7, 4.8, 4.9 y 4.10 los vectores θ_n resultantes de nuestras estimaciones para 3,5,7 y 16 ondas incidentes. Como podemos observar, la condición que impusimos a los coeficientes a través de la distribución a priori obliga a que los coeficientes decaigan con rapidez conforme

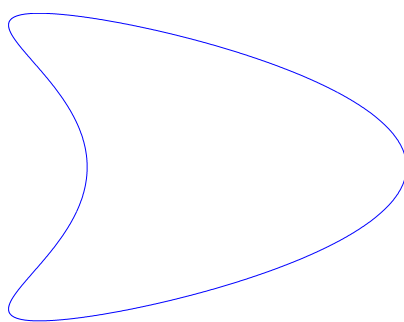


Figura 4.6: Curva de Prueba para el Problema Inverso: Kite

aumenta el valor de n . Un aspecto importante en éste punto es que, a juzgar por las Figuras 4.7, 4.8, 4.9 y 4.10, los vectores son prácticamente los mismos fijando el número de ondas incidentes y variando el el número de coeficientes utilizados. Más aún, en la Figura 4.11 observamos que las estimaciones para los 10 primeros coeficientes no muestran grandes diferencias al incrementar el número n de elementos del vector θ_n y tampoco utilizando distintos números de ondas incidentes. Esto nos indica que la distribución a priori elegida es razonable para este problema.

Debido a que los coeficientes estimados presentan sólo pequeñas variaciones, las curvas resultantes estimadas son ligeramente distintas entre si. Presentamos en la Figura 4.12.(a), 4.12.(b), 4.12.(c) y 4.12.(d) el error relativo obtenido entre la estimación y la curva 4.6 a estimar con 3, 5, 7 y 16 ondas incidentes. En la Figura 4.12.(e) presentamos las Figuras 4.12.(a), 4.12.(b), 4.12.(c) y 4.12.(d) juntas para fines comparativos. No es de sorprender que los resultados con 16 ondas incidentes tengan un menor relativo en la estimación, pues con 16 ondas incidentes se tiene más información del campo dispersado y consecuentemente,

mas detalle en la estimación.

Mostraremos entonces las estimaciones obtenidas. Presentamos en las Figuras 4.13, 4.15, 4.17 y 4.19 las regiones de confianza estimadas utilizando 3, 5, 7 y 16 ondas incidentes u^i respectivamente. De igual forma, en 4.14, 4.16, 4.18 y 4.20 tenemos las curvas estimadas para 3, 5, 7 y 16 ondas incidentes.

Como podemos observar en las Figuras 4.14, 4.16, 4.18 y 4.20, los resultados mejoran en la medida en que incrementamos el número de coeficientes $2n + 1$ pero también es de vital importancia tener suficientes ondas incidentes u^i . Este hecho en particular nos indica los alcances de la solución, dependiendo del número de ondas incidentes, en un experimento físico. Es decir, dependiendo de la calidad deseada para la estimación para el objeto dispersor, elegiríamos el número de ondas incidentes a utilizar. En términos generales, los resultados son aceptables considerando la complejidad del problema. Cabe señalar que los resultados mejoran incrementando tanto el número de coeficientes como el de ondas incidentes y que, debido a la convergencia dada por el método de Nystrom, los resultados obtenidos son esencialmente los mismos incrementando el número de puntos en la curva, así como las direcciones de observación \hat{x} .

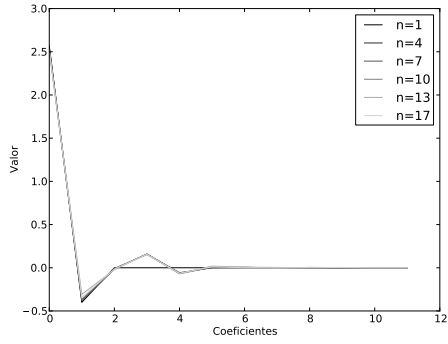


Figura 4.7: Valores de los vectores de parámetros θ_n con diferentes valores de n para 3 ondas incidentes

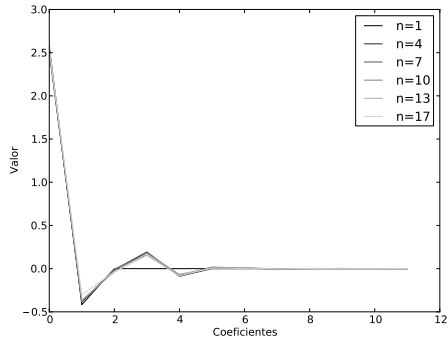


Figura 4.8: Valores de los vectores de parámetros θ_n con diferentes valores de n para 5 ondas incidentes

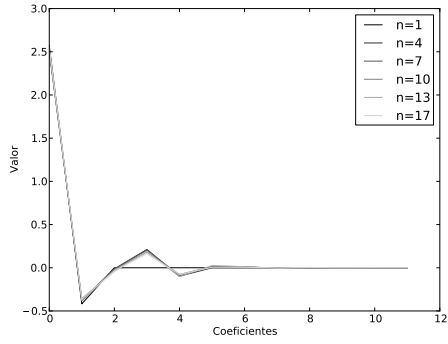


Figura 4.9: Valores de los vectores de parámetros θ_n con diferentes valores de n para 7 ondas incidentes

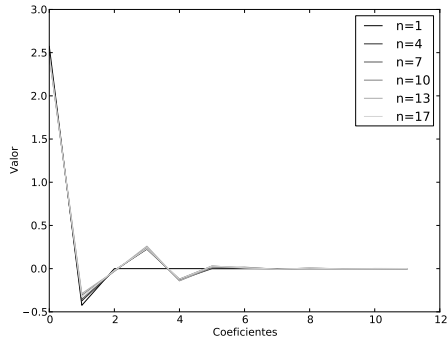
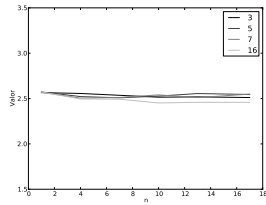
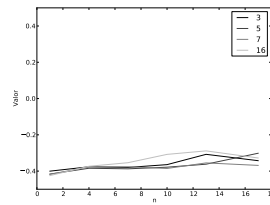


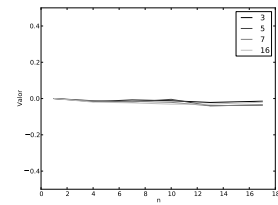
Figura 4.10: Valores de los vectores de parámetros θ_n con diferentes valores de n para 16 ondas incidentes



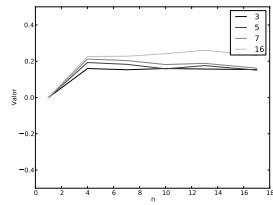
(a) a_0



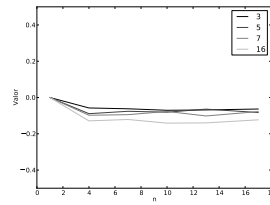
(b) a_1



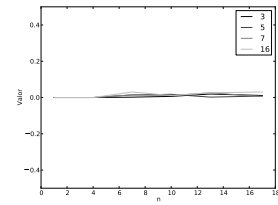
(c) a_2



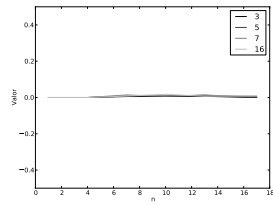
(d) a_3



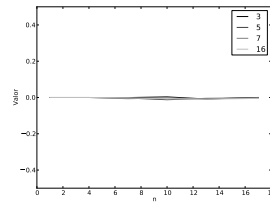
(e) a_4



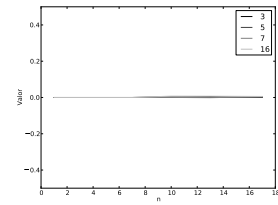
(f) a_5



(g) a_6

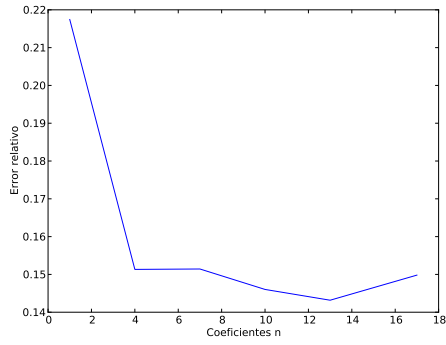


(h) a_7

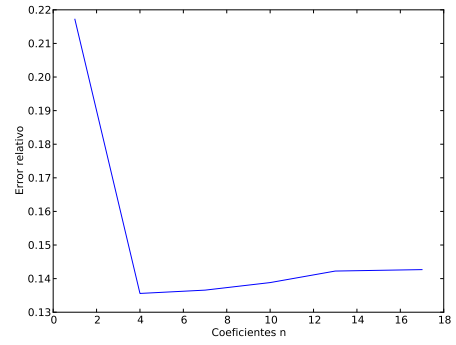


(i) a_8

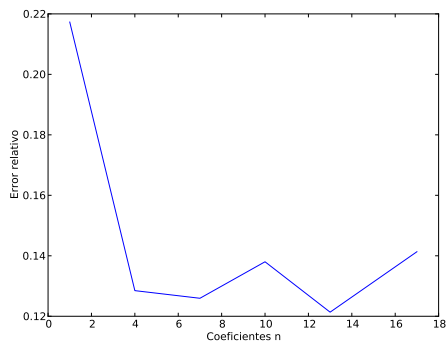
Figura 4.11: Parámetros estimados incrementando el número de coeficientes para distintos números de ondas incidentes.



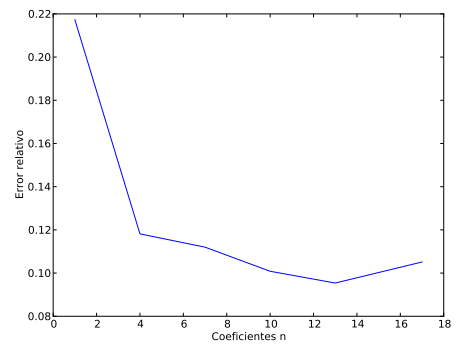
(a) 3 ondas incidentes



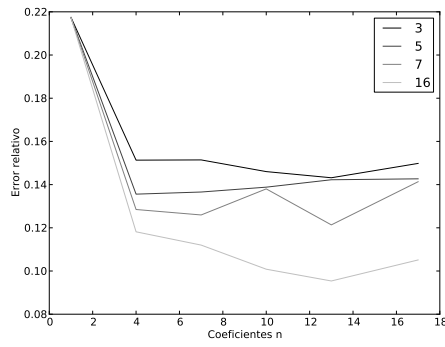
(b) 5 ondas incidentes



(c) 7 ondas incidentes

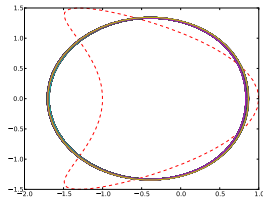


(d) 16 ondas incidentes

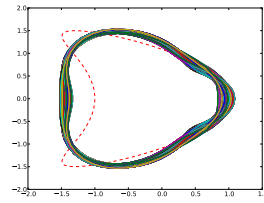


(e) 3, 5, 7 y 16 ondas incidentes

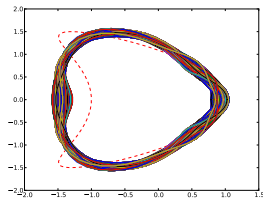
Figura 4.12: Error relativo en las estimaciones con diferente números de ondas incidentes



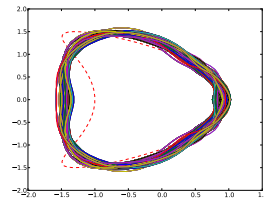
(a) $n = 1$



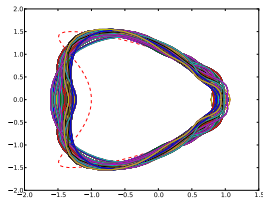
(b) $n = 4$



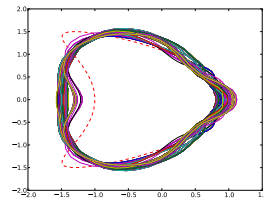
(c) $n = 7$



(d) $n = 10$

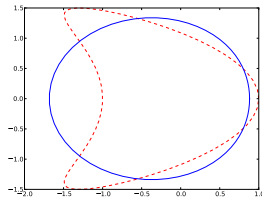


(e) $n = 13$

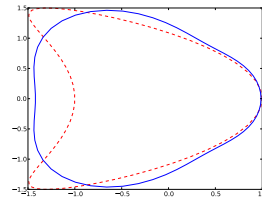


(f) $n = 17$

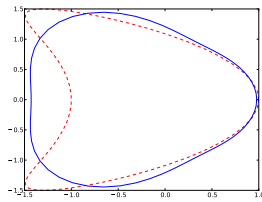
Figura 4.13: Regiones de confianza de la estimación con 3 ondas incidentes u^i .



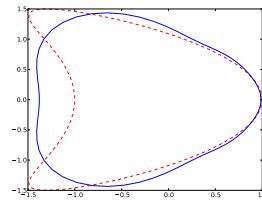
(a) $n = 1$



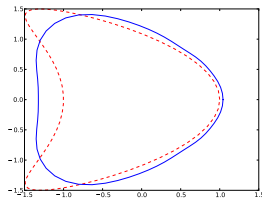
(b) $n = 4$



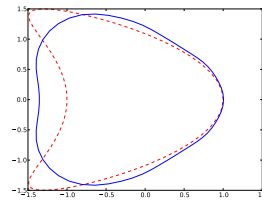
(c) $n = 7$



(d) $n = 10$

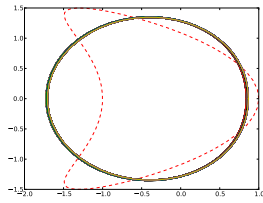


(e) $n = 13$

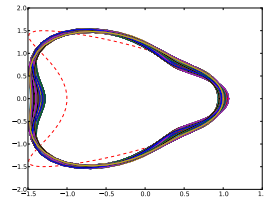


(f) $n = 17$

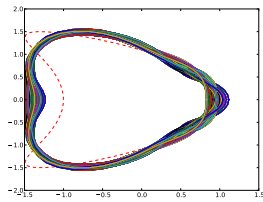
Figura 4.14: Curvas estimadas con 3 ondas incidentes u^i .



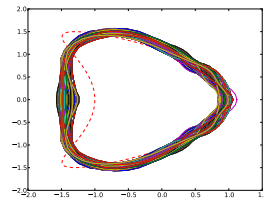
(a) $n = 1$



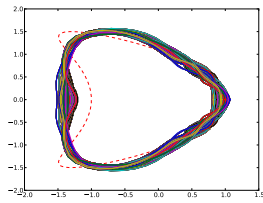
(b) $n = 4$



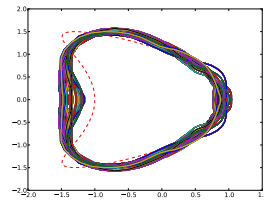
(c) $n = 7$



(d) $n = 10$

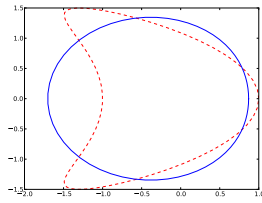


(e) $n = 13$

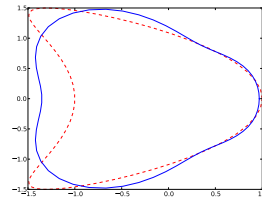


(f) $n = 17$

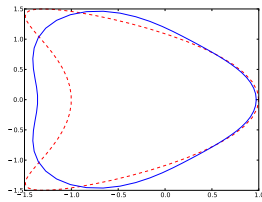
Figura 4.15: Regiones de confianza de la estimación con 5 ondas incidentes u^i .



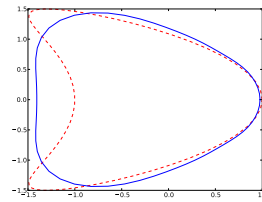
(a) $n = 1$



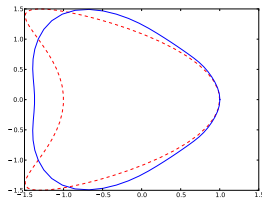
(b) $n = 4$



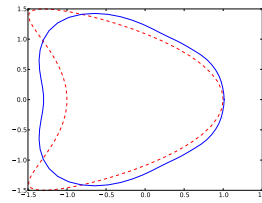
(c) $n = 7$



(d) $n = 10$

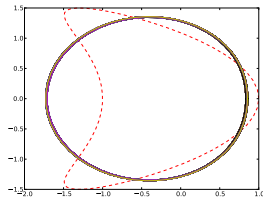


(e) $n = 13$

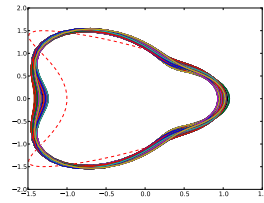


(f) $n = 17$

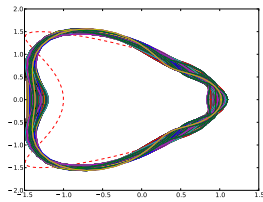
Figura 4.16: Curvas estimadas con 5 ondas incidentes u^i .



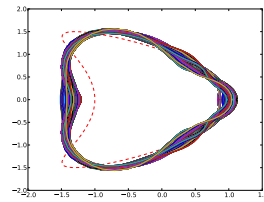
(a) $n = 1$



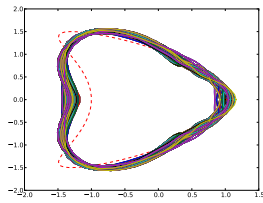
(b) $n = 4$



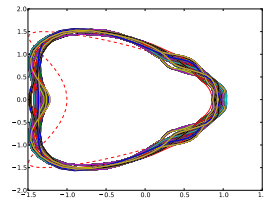
(c) $n = 7$



(d) $n = 10$

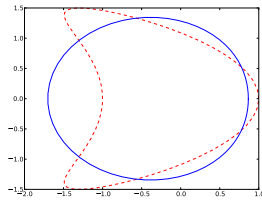


(e) $n = 13$

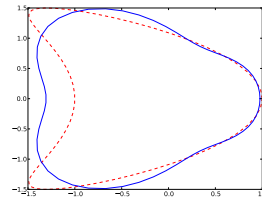


(f) $n = 17$

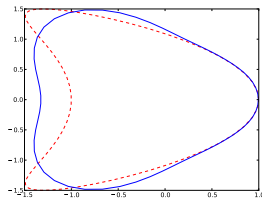
Figura 4.17: Regiones de confianza de la estimación con 7 ondas incidentes u^i .



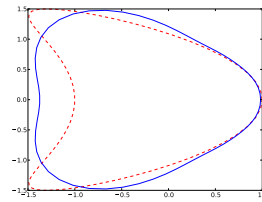
(a) $n = 1$



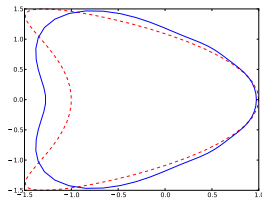
(b) $n = 4$



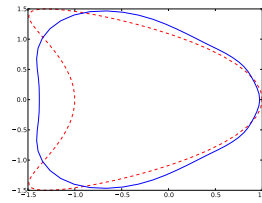
(c) $n = 7$



(d) $n = 10$

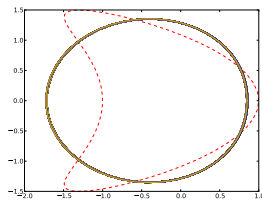


(e) $n = 13$

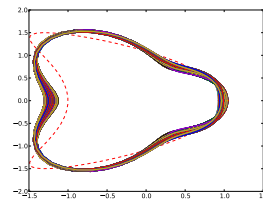


(f) $n = 17$

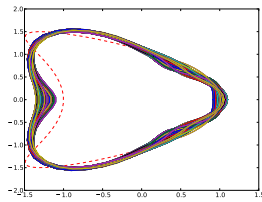
Figura 4.18: Curvas estimadas con 7 ondas incidentes u^i .



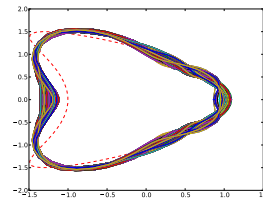
(a) $n = 1$



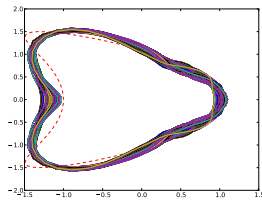
(b) $n = 4$



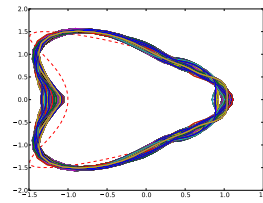
(c) $n = 7$



(d) $n = 10$

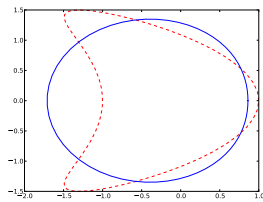


(e) $n = 13$

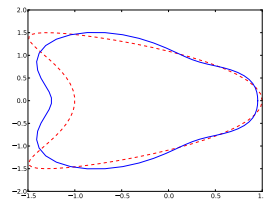


(f) $n = 17$

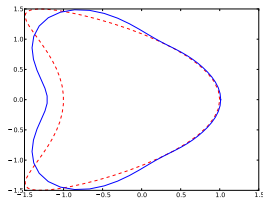
Figura 4.19: Regiones de confianza de la estimación con 16 ondas incidentes u^i .



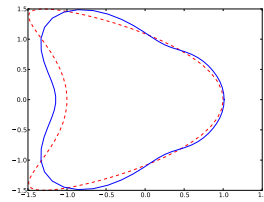
(a) $n = 1$



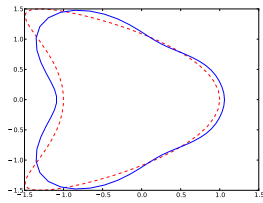
(b) $n = 4$



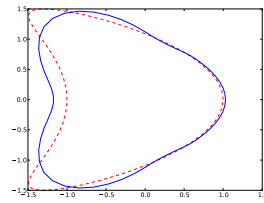
(c) $n = 7$



(d) $n = 10$



(e) $n = 13$



(f) $n = 17$

Figura 4.20: Curvas estimadas con 16 ondas incidentes u^i .

Capítulo 5

Conclusiones

La teoría de dispersión de ondas ha sido ampliamente estudiada. Actualmente las aportaciones en este campo van dirigidas a implementar estrategias computacionales que permitan resolver, tanto el problema directo como el problema inverso, de forma eficiente. En este sentido, el desarrollo de herramientas de cómputo es una pieza fundamental para dar solución a problemas como el de dispersión de ondas. Hablaremos a continuación de aspectos que consideramos destacables en la solución, tanto del problema directo, como del problema inverso.

5.1. Problema Directo

El método de Nystrom proporciona ventajas significativas para la solución de ecuaciones integrales. La primera de ellas es que reduce el problema de encontrar el potencial en un todo el dominio que contiene al obstáculo dispersor, a resolver una ecuación integral sobre la frontera del objeto. Aquí se presenta una reducción importante del costo computacional. Posteriormente, debido a la pe-

riodicidad de la frontera del obstaculo dispersor, el método de Nystrom aproxima la solución a la ecuación integral con convergencia superalgebraica. Tenemos entonces otra reducción importante del costo computacional al no requerir una gran cantidad de puntos sobre la frontera para dar resultados bastante precisos. Consecuentemente las matrices de los potenciales de Capa Simple y Combinado sean matrices con dimensiones no tan grandes.

Por otro lado, la construcción de las matrices para ambos potenciales es paralelizable. Aunado a ello, se pueden implementar métodos computacionales que a su vez sean paralelizables para resolver el sistema no simétrico, complejo y denso en tiempos muy reducidos. Resolver el problema inverso, típicamente requiere de una gran cantidad de evaluaciones del problema directo, por tanto, la paralelización del problema directo es un aspecto importante para tener resultados el problema inverso en tiempos viables.

Existen diversas aplicaciones en las que el lenguaje de programación o la estrategia de implementación no son relevantes. En nuestro caso, requerimos hacer grandes cantidades de evaluaciones, así que se debe elegir adecuadamente la implementación a utilizar tratando de que ésta sea eficiente tanto en tiempo de computo como en administración de recursos de almacenamiento. Por supuesto, también se tiene que elegir la implementación considerando que proporcione el resultado más preciso posible.

5.2. Problema Inverso

Los problemas inversos típicamente son mal planteados. Además, las soluciones al problema directo no siempre son suficientemente precisas. Esto los vuelve aún más complicados de resolver. Ahora bien, en general se tiene un

modelo de ruido en las observaciones o datos del problema. Los métodos clásicos como Tykhonov introducen un término de regularización que compense la información perdida a causa del ruido y que permita incorporar información conocida al problema. Es decir, que la estrategia de regularización refleje lo que sabemos del problema y que con esto se reduzca adecuadamente el espacio de búsqueda para la solución. Desafortunadamente, la elección incorrecta de una estrategia de regularización puede provocar que los resultados que obtengamos sean totalmente distintos a la solución exacta.

El planteamiento Bayesiano, nos proporciona un entorno en el que podemos tener un poco más de control sobre la información a priori al escribirla como una distribución de probabilidad. Además, la incorrecta elección de la distribución a priori puede ser identificada al observar la aproximación a la distribución posterior, así como las distribuciones marginales de los parámetros. Por otro lado, resolver el problema para casos sencillos como el mostrado en esta tesis, nos ayuda a controlar las distribuciones involucradas y posteriormente pensar en abordar el problema en una aplicación práctica manteniendo las mismas distribuciones. El planteamiento Bayesiano, permite además tener medidas sobre la información que se tiene o resulta de la implementación y en ese sentido utilizar resultados de teoría de probabilidad para saber qué tanto podemos confiar en nuestras estimaciones.

En los problemas de optimización, en general, resulta complicado elegir el parámetro de regularización. Un parámetro de regularización grande en comparación con el nivel de ruido, provoca que la estrategia de regularización tenga mayor peso en la minimización y por tanto, que la solución sea incorrecta. Contrariamente, un parámetro pequeño no es suficiente para compensar el error inducido por el ruido. En nuestro esquema de solución, un planteamiento ade-

cuado de la distribución a priori controla la influencia de la distribución misma evitando el problema de la regularización.

5.3. Trabajo a Futuro

El problema inverso que presentamos aún tiene muchas áreas de oportunidad. La ejecución del algoritmo t-walk es bastante costosa. Una mejora inmediata para el presente trabajo es utilizar la versión paralelizada del problema directo para el algoritmo t-walk. Alternativamente, podemos estudiar el algoritmo t-walk para identificar secciones del mismo que puedan ser paralelizadas y reducir así el costo de todas las evaluaciones o en su defecto, incrementar el número de estas para garantizar la convergencia de la distribución estacionaria.

En esta tesis, se realizaron experimentos para resolver el problema inverso utilizando diferentes distribuciones a priori para los parámetros. Los mejores resultados fueron obtenidos con la a priori basada en la identidad de parseval. Sin embargo, es complicado tener una noción adecuada de distancia en el espacio generado por los coeficientes de Fourier. Es necesario estudiar entonces diferentes planteamientos que proporcionen distribuciones a priori adecuadas para los parámetros. Recordemos que nuestros experimentos utilizaron curvas cerradas simples y suaves, entonces, en la medida que la complejidad de la curva aumente, nuestro planteamiento no podrá ofrecer buenos resultados.

Los objetos dispersores que podemos encontrar en problema de dispersión de ondas, pueden ser curvas no suaves o con esquinas. En este caso, la cantidad de coeficientes de Fourier para representar la curva se incrementa considerablemente y no necesariamente define la curva con exactitud. Es importante pues, pensar en una forma distinta de escribir la parametrización de la curva de tal

forma que sigamos teniendo pocos coeficientes para definirla y que estos sean suficientes para describir la curva completa. Una alternativa es escribir el radio como una combinación lineal de bases wavelet. Las bases wavelet son escalamientos y traslaciones de una wavelet madre, en este sentido, podemos usar wavelets madre que sí permitan representar esquinas (como Haar por ejemplo). Además, los coeficientes de la expresión en la base wavelet se encuentran en un espacio de Bessov por lo que la noción de distancia y por consiguiente la distribución a priori de los mismos, estaría en terminos de la norma de Bessov.

Finalmente, queremos agregar que el esquema Bayesiano para problemas inversos, plantea alternativas en busca de la solución a estos. Los resultados que se obtienen con la formulación Bayesiana pueden llegar a ser bastante precisos con la información adecuada para cada una de las distribuciones.

Bibliografía

- [1] Nvidia cuda reference manual.
- [2] J. A. Christen and C. Fox. A general purpose sampling algorithm for continuous distributions (the t-walk). *Journal of Bayesian Analysis*, 5(2):263–282, 2010.
- [3] D. Colton and R. Kress. *Integral Equation Methods in Scattering Theory*. John Wiley & Sons, 1983.
- [4] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*, volume 93 of *Applied Mathematical Sciences*. Springer, second edition, 1998.
- [5] R. G. Duque. *Python para todos*. <http://mundogeek.net/tutorial-python/>, 2011.
- [6] S. C. Eisenstat, H. C. Elman, and M. H. Schultz. Variational iterative methods for nonsymmetric systems of linear equations. *SIAM Numerical Analysis*, 20(2):345–357, 1983.
- [7] V. Isakov. *Inverse Problems for Partial Differential Equations*, volume 127 of *Applied Mathematical Sciences*. Springer, 2006.

- [8] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*, volume 190 of *Applied Mathematical Sciences*. Springer, 2005.
- [9] R. Kress. Integral equation methods in inverse obstacle scattering. *AN-ZIAM*, 42:65–78, 1995.
- [10] R. Kress. *Linear Integral Equations*, volume 82 of *Applied Mathematical Sciences*. Springer, 2 edition, 1999.
- [11] R. G. Newton. Inverse scattering. ii. three dimensions. *Journal of Mathematical Physics*, 21(7), 1980.
- [12] M. A. Pinsky. *Introducción al análisis de Fourier y las ondas*. Thompson, 2003.
- [13] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, second edition edition, 1992.
- [14] L. N. Trefethen and D. Bau. *Numerical Linear Algebra*. SIAM, 1997.
- [15] C. Vogel. *Computational methods for inverse problems*, volume 23. Society for Industrial Mathematics, 2002.